This is to certify that the

dissertation entitled

presented by

Douglas Scott Burdette

has been accepted towards fulfillment
of the requirements for

___Ph. D.___ degree in Biochemistry

_____
Major professor

Date__March 19, 1996__

PLACE IN RETURN BOX to remove this checkout from your record.
TO AVOID FINES return on or before date due.

DATE DUE	DATE DUE	DATE DUE
NOV 20 ? ｱ ｴ		

MSU is An Affirmative Action/Equal Opportunity Institution

c:\circ\datedue.pm3-p.1

SEC

THERMOANAEROBACTER ETHANOLICUS 39E SECONDARY-ALCOHOL DEHYDROGENASE: MOLECULAR BASIS FOR STABILITY AND CATALYSIS

By

Douglas S. Burdette

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Biochemistry

1996

dehydr
DHSa
precipi
subunit
propan-2
H5N. D
daz and
diethyl
Hydroph
catalytic Z
protein b
emission
residues i
are impor
prediction
Rossmann
distinct be
thermoph
8°C. An

ABSTRACT

THERMOANAEROBACTER ETHANOLICUS 39E SECONDARY-ALCOHOL DEHYDROGENASE: MOLECULAR BASIS FOR STABILITY AND CATALYSIS

By

Douglas S. Burdette

The *adhB* gene encoding *Thermoanaerobacter ethanolicus* 39E secondary-alcohol dehydrogenase (2° ADH) was cloned, sequenced, and overexpressed in *Escherichia coli* DH5α (~15% total prot.). The protein was purified to homogeneity by heating and precipitation. The 1056 bp gene encoded a homotetrameric recombinant enzyme (37.7 kDa subunits) that displayed a 170-fold greater catalytic efficiency toward NADP(H)-dependent propan-2-ol than toward ethanol oxidation. The 2° ADH site directed mutants C37S, H59N, D150N, D150E, and D150C displayed ≤ 3% of wild type catalytic activity. These data and the wild type enzyme inactivation by dithionitrobenzoate (DTNB) and diethylpyrocarbonate chemical modification, supported sequence predictions based on Hydrophobic Cluster Analysis (HCA) that Cys37, His59, and Asp150 residues are catalytic Zn ligands. X-ray absorption spectrometry data supported the presence of a protein bound Zn with a $ZnS_1(N-O)_{3-4}$ coordination sphere. Induction coupled plasma emission spectrometry data of wild type and mutant enzyme Zn binding, implicated these residues in 2° ADH Zn liganding. Thus, a catalytic Zn atom and its immediate environment are important in 2° ADH activity. Analysis of the G198D mutant supported the HCA based prediction that this 2° ADH binds NADP(H) in a common nucleotide binding motif, a Rossmann fold. The substrate and cofactor binding sites in the 2° ADH are structurally distinct because catalytic Zn ligand mutations did not change cofactor affinity. The thermophilic 2° ADH) was optimally active at ~90°C and displayed a half-life of 1.2 days at 80°C. Analysis of temperature dependent unfolding of 2° ADH in guanidine hydrochloride

(G.H

enzy

incre

activi

denat

KNO:

hydro

plots f

90°C s

energy

Conse

these m

(GuHCl) indicated the enzyme was very rigid (50% unfolded at ~115°C). Thus, the enzyme is nearly 100% folded at 90°C where it is maximally active. 2° ADH Activity increased 2-fold at 37°C to 75°C in low concentrations of GuHCl (120-190 mM); and, activity remained in 1.6M GuHCl demonstrating its high resistance to chemical denaturants. Catalytic rate enhancement was also seen with weakly chaotropic (i.e., KNO_3) and neutral (i.e., KCl) inorganic salts, indicating that specific ionic and not hydrophobic interactions are required for the catalytic rate enhancement. Linear Arrhenius plots for the oxidation of propan-2-ol by the native and recombinant 2° ADHs from 30°C to 90°C suggested that the thermal activity relationships are related to binding high kinetic energy substrates and not to temperature dependent changes in enzyme unfolding. Consequently it is suggested that the 2° ADH has evolved high rigidity for stability and these molecular determinants are distinct from those which control catalysis.

To Kate and to our families

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

1° ADH	primary-alcohol dehydrogenase
2° ADH	secondary-alcohol dehydrogenase
ADH	alcohol dehydrogenase
DEPC	diethylpyrocarbonate
DTNB	dithionitrobenzoate
DTT	dithiothreitol
Ea	reaction activation energy
EDTA	(ethylenedinitrilo)tetraacetic acid trisodium salt
EPR	Electron paramagnetic resonance
EXAFS	X-ray absorption spectrometry
FTIR	Fourier transformed infra-red spectroscopy
GAPDH	glyceraldehyde-3-phosphate dehydrogenase
GuHCl	Guanidine hydrochloride
HCA	Hydrophobic Cluster Analysis
ICP	induction coupled plasma emission spectrometry
MALDI	matrix associated laser desorption ionization mass spectrometry
ORF	open reading frame
PEG	polyethylene glycol
PCR	Polymerase chain reaction
RBS	ribosome binding site
T_m	Enzyme melting temperature in the absence of denaturant
$T_{m[Gu]}$	Enzyme melting temperature in the presence of guanidine

Chapter I

Literature Review

ALCOHOL DEHYDROGENASES

Alcohol dehydrogenase structure and function

Alcohol dehydrogenases (ADHs) are integral to both prokaryotic and eukaryotic metabolism. Categorized as 1° or 2° based on their displaying higher activity toward 1° or 2° alcohols, these enzymes typically act on a broad range of substrates. The 1° ADH isolated from mammalian liver is proposed to oxidize alcohols and possibly steroids during their detoxification [1]. ADHs are also involved in catabolic alcohol consumption by acetic acid batceria [2]. More commonly however, catabolic ADHs mediate the terminal electron transfer in solventogenic fermentations by yeasts [3-5] and bacteria [1]. Alcohol formation by these organisms provides the oxidized nicotinamide cofactor necessary for glycolysis, eliminating the need for an exogenous terminal electron acceptor. Catabolic 1° ADHs have been identified in mesophilic (optimal growth between 25°C and 50°C) [1,6], thermophilic (optimal growth between 60°C and 80°C) [7-9], and hyperthermophilic (optimal growth above 80°C) [10,11] microorganisms. However, catabolic 2° ADHs have only been reported in mesophilic [12,13] and thermophilic [8,14-16] bacteria. The mesophilic bacterium, *Clostridium beijerinckii*, expresses a 2° ADH during propan-2-ol production whereas, the thermophilic *Thermoanaerobacter sp.* and *Thermoanaerobacterium sp.* use a 2° ADH to form ethanol - a 1° alcohol - from acetylCoA [7,8]. These thermophilic bacteria ferment sugars to ethanol and acetic acid (neither butanol, acetone, nor propan-2-ol have been detected) and express a functional 1° ADH concurrently with the 2° ADH. The *Thermoanaerobacter ethanolicus* 39E 1° ADH is proposed to function in electron transfer between NAD(H) and NADP(H) and in ethanol consumption while the ethanol producing 2° ADH can function to reduce both acetylCoA and ethanal [8] using the generalized enzymatic reaction:

$$R-\overset{\overset{\displaystyle O}{\|}}{C}-CH_3 \ + \ NADPH+H^+ \ \longrightarrow \ R-\overset{\overset{\displaystyle OH}{\vdots}}{\underset{\overset{\vdots}{H}}{C}}-CH_3 \ + \ NADP^+$$
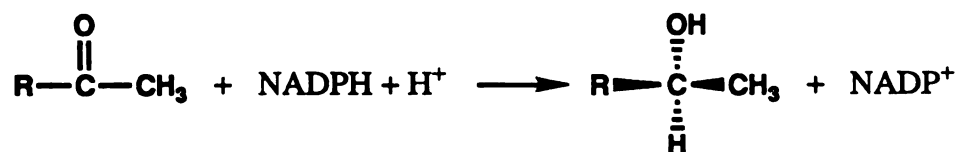
ADHs are predominantly nicotinamide cofactor dependent Zn metalloproteins although Fe [17] and ferredoxin linked enzymes have been reported [18]. Theorell and Chance proposed an ordered 1° ADH kinetic mechanism where cofactor binding is followed by substrate binding, hydride transfer, product release, then cofactor release [1]; however a semi-random mechanism for product and cofactor release has also been reported [19]. Isomerization of enzyme bound to reduced cofactor is believed to be the slow step in liver ADH catalysis [20,21]. Cofactor binding is proposed to induce a slow enzyme conformational change after a rapid initial binding step, completing active site pocket formation. Even the NAD^+ molecular geometry required for horse liver ADH catalytic activity has been determined [22]. Zn's catalytic role as a Lewis acid in the electrophilic ADH mechanism is well established, with the metal mediating direct hydride transfer between substrate and cofactor [23]. Kinetic and spectrophotometric measurements provide direct evidence that the alcohol or carbonyl substrate binds to the catalytic Zn atom [23-25]. This may explain both the broad ADH substrate specificities and why decades of research into 1° ADH structure-function relationships have failed to identify residues responsible for substrate binding and discrimination.

ADHs are either dimeric or tetrameric, composed of identical or nearly identical 35 kDa to 45 kDa subunits. Both 1° and 2° ADHs have been reported to contain catalytic Zn atoms and Cys residues critical to catalytic activity [1,8,15]. The exact protein structure has been determined for horse liver 1° ADH by x-ray crystallography [26] but no 3-dimensional 2° ADH structure has been reported. X-ray structural data confirmed that 1° ADHs bind a structural Zn atom with 4 Cys residues and a catalytic Zn with a His plus 2 Cys residues. The broad specificity of liver 1° ADH is attributed to a wide and deep hydrophobic binding pocket seen in the x-ray structure [1]. Nicotinamide cofactor was

shown to associate with a Rossmann fold in the 1° ADH structures as reported for other

NAD(P)(H) binding proteins [27]. The identification of Rossmann fold consensus

sequences {Gly-Xaa-Gly-Xaa-Xaa-Gly-(Xaa)$_{18-20}$[negatively charged amino acid for

NAD(H) dependent or neutral amino acid for NAD(P)(H)]} [28] provide sequence based

predictions of cofactor specificity linked to this particular structural fold.

Stereoselectivity is both a hallmark of enzymatic catalysis and an area of great

biotechnological interest. ADHs are highly enentioselective catalysts despite their

characteristically broad substrate specificities. In pyridine dinucleotide containing

enzymes, stereospecificity results from the cofactor attack angle according to Prelog [29].

Carbonyl reduction can result from hydride addition to the Re face, following Prelog's

Rule, or to the Si face, following Anti-Prelog's Rule. So based on the orientation of the

carbonyl relative to the cofactor and which cofactor hydrogen is transfered, there are 4

possible mechanisms for enantiospecific ADH reaction (E$_1$-E$_4$) [30]. The E$_1$ mechanism is

defined as the pro-R cofactor hydrogen (H$_R$) on the substrate Si face, E$_2$ involves cofactor

pro-S hydrogen (H$_S$) attack on the substrate Re face, H$_R$ transfer to the substrate Si face

results from an E$_3$ mechanism, and reduction on the carbonyl Re face by the cofactor H$_S$

describes mechanism E$_4$. The horse liver and yeast 1° ADHs operate by the E$_3$ pathway,

producing (S)-alcohols. E$_1$ reactions catalyzed by *L. kefir* and *Psuedomonas sp.* (SBD6)

ADHs form (R)-alcohols [30,31].

Overall structural similarity among ADHs is presumed based on their similar gross

morphologies and functional characteristics despite their dissimilar peptide sequences [1].

The proposed structural similarity between ADHs and the difficulty in determining residues

involved in substrate specificity makes comparative analysis between enzymes with

overlapping but different specificities important. Cloning the genes encoding these

enzymes would provide adequate supplies of protein for structural studies and the means to

perform site directed mutagenesis experiments. Mesophilic (GenBank Acc. No. D90004,

L02104, M91440, X17065, X59263), thermophilic (GenBank Acc. No. D90421), and

c

se

se

etra

w

B.

con

val

1ª a

Alc

synth

legis

and r

toxic

oxido

neuro

(e.g.,

polya

fatty a

lipases

Enzy

hyperthermophilic (GenBank Acc. No. S51211) 1° ADHs have been cloned and characterized. The 2° ADH from the obligately aerobic mesophile *Alcaligenes eutrophus* has been cloned and characterized [32] and the mesophilic *C. beijerinckii* 2° ADH gene sequence has been deposited in GenBank (Acc. No. M84723) but no characterization has been published. Alignment of peptide sequences derived from the translated gene sequences indicated the conservation of putative catalytic and critical core residues among the *A. eutrophus* and liver enzymes [32]. However, an alignment between the *A. eutrophus* 2° ADH, liver 1° ADH, and *T. brockii* 2° ADH (The *T. brockii* peptide sequence was determined by Edman degradation of purified protein fragments) led Peretz and Burstein to conclude that insufficient similarity existed between the 3 sequences to make comparative structural predictions [33]. These conflicting conclusions indicate that the validity of 2° ADH structure-function hypotheses based on structural comparisons between 1° and 2° ADHs awaits assessment by genetic manipulation.

Alcohol dehydrogenase industrial applications

While industrial organic syntheses have typically been dominated by chemical-synthetic processes, the rising public environmental awareness (reflected by new legislation) plus the recent developments in enzyme biotechnology suggest that enzymatic and mixed chemo-enzymatic processes will progressively be substituted for polluting or toxic chemical processes. Four enzyme groups - carbohydrases, lipases, proteases, and oxidoreductases - have a high potential for synthesizing peptides (e.g., pharmaceuticals, neuropeptides, and specific peptides for research purposes) [34]; flavors and fragrances (e.g., benzaldehyde, naringin) [35]; non-peptide polymers (e.g., polyesters, polyphenols, polyacrylates) [36]; chiral compounds [37]; and surfactants (e.g., monoglycerides, sugar fatty acid esters, alkyl glucosides) [38a]. Typically known for their hydrolytic activity, lipases and proteases can, in specific environments, be used as synthetic enzymes. Enzymatic syntheses present numerous advantages over chemical syntheses: they are

usually highly specific (i.e., enantio-, regio-, and stereo-specific), environmentally friendly, and their products are usually easily biodegradable.

Chiral alcohols are common in biological effectors (eg. the neurotransmitter norepinephrine) and chiral alcohols can be used as stereospecific reaction centers in bioactive compound manufacturing. With the top ten optically active drugs representing sales of $10 billion (U.S. dollars) annually and with the recent FDA policy focused on pharmaceutical enantiomeric purity [37], strategies to produce chiral compounds as either drugs or as synthetic intermediates have growing industrial potential. ADHs, enantioselectively active on a wide range of substrates, have been the focus of substantial research into enzymatic synthesis of industrially relevant chiral compounds. The horse liver 1° ADH is optimized for the interconversion of aldehydes and 1° alcohols but it will stereospecifically reduce ketones [1]. Mesophilic liver 1° ADH has already been used in the analytical scale production of chiral cyclic alcohols, polyalcohols, alcohol containing aldehydes (eg. L-glyceraldehyde) or ketones (cis-cyclohexanone-2-ol), lactones, and organosilicates [see 38,39]. Similarly diverse and numerous potentially valuable chiral bioconversions have been performed in vitro and in vivo with yeast ADH [see 38,40,41]. Resting cells of the hyperthermophile S. sulfataricus were shown to stereospecifically reduce ketones and to express a 1° ADH with activity toward ketones and 2° alcohols [42]. Ketones and 2° alcohols however, were poor substrates for this enzyme. The mesophilic, (R)-specific ADH from L. kefir has high activity toward alcohols near cyclic or aromatic structures [31,43]. A range of chiral alcohols was also created using a mesophilic (R)-specific 1° ADH from Pseudomonas sp. (ATCC 49794) [30].

The potential for enzymatic conversions in specialty chemical synthesis and racemic resolution is widely recognized [38,44-51]. However, the expensive cofactor requirements (NAD, NADP, FAD) of oxido-reductases such as ADHs may limit their applications [37,52]. Several systems [51-54], have been demonstrated for cofactor recycling, but current biotransformations are typically performed in whole cells, where the cellular

machinery retains and regenerates the cofactor [37,40,41,52]. The specific cases of chiral diol synthesis using a host of microbial cells have been compiled and reviewed [55]. Many important specialty chemicals and their precursors are sparingly soluble in water, requiring a solvent stable biocatalyst. Horse liver ADH was shown to be catalytically active in numerous organic solvents with specific activities similar to that in aqueous solution [56,57]. The activity if horse liver 1° ADH was also examined in water-oil microemulsions to enhance conversion rates for substrates with low water solubility [58]. The hyperthermophile S. solfataricus 1° ADH [10], and its resting cells have been used in chiral biotransformations [42]. This enzyme has high stability in the presence of solvents [59], but its low catalytic efficiency toward 2° alcohols and ketones might limit its applications.

The importance of biocatalyst longevity and stabilty to their large scale utilization makes necessary the development of strategies to stabilize mesophilic enzymes or the isolation of highly stable analogs to them. Intrinsically stable and active at high temperatures, thermophilic enzymes offer major biotechnological advantages over mesophilic enzymes: (i) once expressed in mesophiles, thermozymes may be purified by heat treatment [60-62]; (ii) their thermostability is associated with a higher resistance to chemical denaturants (such as a solvent or guanidine-HCl); (iii) performing enzymatic reactions at high temperatures can allow higher reaction rates, higher substrate concentrations, and lower viscosity; and (iv) there is a higher product yield during certain reactions due to chemical equilibrium shifts with high temperature. Isolating and characterizing thermophilic enzymes therefore, provides natural examples of structurally stable enzymes with high active temperature ranges that are potentially robust industrial biocatalysts. Comparisons of these thermophilic enzymes to their mesophilic analogs will also provide insights into what molecular interactions confer high folded protein stability. The extensive ADH structure and function research literature have been well reviewed [1].

T

H
in
th
en
[6
ina
[7
op
ab
and
(eg
and
enz

incl
60°C
consi
eukar

The emerging ideas regarding protein thermostability, thermophilicity, and folding molecular mechanisms will be reviewed here.

THERMOPHILIC ENZYMES

Thermophilic enzyme sources and diversity

Originally, thermophilicity was a property associated exclusively with spore forming bacteria. Thermophilic enzymes were believed to be unstable, with high protein turnover rates explaining why thermophiles did not grow faster than mesophiles [66]. However, *Thermus aquaticus*, a non-sporulating thermophile, was shown to express inherently thermostable enzymes [63], overturning these hypotheses. Subsequently, most thermophiles and hyperthermophiles [64-69] have been shown to possess inherently stable enzymes that function at temperatures above the organism's optimal growth temperature [65,70]. Enzyme thermostability is the protein's capacity to resist irreversible thermal inactivation, and is commonly reported as the enzyme's half-life at a given temperature [71]. Enzyme thermophilicity is defined as the temperature at which the enzyme is optimally active [71]. An enzyme is thermophilic if it is optimally active at temperatures above 60°C [71]. Although thermophilic enzymes typically originate from thermophiles and hyperthermophiles, some mesophiles produce enzymes active and stable above 60°C (eg. pancreatic ribonuclease A). Mesophilic enzymes are optimally active between 20°C and 60°C [71]. Mesophilic enzymes, typically from mesophiles, include most eukaryotic enzymes and those from mesophilic bacteria and archaea.

All known thermophiles (optimal growth from 60°C to 80°C) are microbial including bacteria, archea, and some blue-green algae which grow at temperatures up to 60°C. The predominance of prokaryotic life forms at thermophilic temperatures is consistent with the appearance of prokaryotes while the Earth was much warmer, with eukaryotic life forms evolving much later. Thermophiles have been isolated from hot

environments inclu

springs, soils, shal

(ii) microbially se'

and (iii) industrial

sludge systems, o

diverse as their m

heterotrophic, che

among others.

Tables 1

thermophilic and

thermostability p

reflects the lack

specific method

cases, it include

potential biotec

thermophiles w

enzymes have

thermostable w

growth temper

stringently for

alcohol dehyd

[GAPDHs], e

function of the

thermophilic (

substrate or c

Despi

usually have

environments including: (i) natural volcanic environments (continental solfataras, hot springs, soils, shallow marine and deep-sea hot sediments, submarine hydrothermal vents); (ii) microbially self-heated environments (e.g., manure, coal refuse piles, compost piles); and (iii) industrial environments (e.g., food industry effluents, hot-water lines, sewage sludge systems, oil drilling injection water systems). Thermophiles, as physiologically diverse as their mesophilic counterparts, include species that are aerobic and anaerobic, heterotrophic, chemoorganotrophic, chemolithotrophic, autotrophic, and phototrophic among others.

Tables 1 and 2 list enzymes which have been characterized and/or cloned from thermophilic and hyperthermophilic organisms, respectively. Thermophilicity and thermostability properties are included where available. The heterogeneity of this data reflects the lack of consensus on the way to guage these properties (we have proposed specific methods for standardizing these data [71]). This list is not exhaustive, in most cases, it includes only examples of each enzyme type and it focuses on enzymes with potential biotechnological applications. An extensive list of enzymes purified from thermophiles was published by Coolbear et al. [70] and detailed descriptions of individual enzymes have also been compiled [64,70,72]. Thermophilic enzymes are inherently thermostable with optimal activities at temperatures near the original organism's optimal growth temperature. The proximity to organism optimal growth temperature holds more stringently for enzymes within a single structural family (i.e., α–amylases, proteases, alcohol dehydrogenases [ADHs], glyceraldehyde-3-phosphate dehydrogenases [GAPDHs], etc.) than across a range of proteins, suggesting that stability is partly a function of the protein structural fold. Note that the optimal activity temperatures for some thermophilic enzymes, oxidoreductases in particular, have not been determined because of substrate or coenzyme (e.g., NAD, NADP) instability.

Despite being synthesized in a mesophilic host, recombinant thermophilic enzymes usually have kinetic and thermal stability characteristics identical to those of the native

Table 1. Thermozymes from thermophilic organisms

Enzyme	Organism (optimal growth temperature)	Enzyme thermophilicity	Enzyme thermostability	General comments	Ref. No.
OXIDOREDUCTASES					
Formylmethanofuran dehydrogenase	*Methanobacterium wolfei* (60°C)	65°C/pH7.4	stable at 65°C (+1M KCl) 20 min/75°C		73,74
GAPDH	*Bacillus stearothermophilus (55°C)*				

Table 1. Thermozymes from thermophilic organisms

Enzyme	Organism (optimal growth temperature)	Enzyme thermophilicity	Enzyme thermostability	General comments	Ref. No.
OXIDOREDUCTASES					
Formylmethanofuran dehydrogenase	*Methanobacterium wolfei* (60°C)	65°C/pH7.4	stable at 65°C (+1M KCl)		73,74
GAPDH	*Bacillus stearothermophilus* (70°C)		20 min/75°C		see 70
MTHF dehydrogenase	*Clostridium thermoaceticum* (60°C)	>64°C	stable at 50°C		75
Secondary alcohol dehydrogenase	*Thermoanaerobacter ethanolicus* (69°C)			C, S, expr.	
Sulfite reductase	*Thermodesulfobacterium commune* (70°C)	65-70°C			76
Glutamate dehydrogenase	*S. shibatae*			C, S	77
TRANSFERASES					
Adenylate kinase	*S. acidocaldarius* (75°C,pH2-3)	90°C/pH5.3-6.0	48h/75-80°C	C, S, expr.	78,79
DNA polymerase I	*Thermus aquaticus* (70°C)		40 min/95°C	C, S, expr.	73,80
3-Phosphoglycerate kinase	*B. stearothermophilus*			C, S, expr.	81
3-Phosphoglycerate kinase	*T. thermophilus* (65-72°C)			C, S	82
HYDROLASES					
β-N-Acetylhexosaminidase	*B. stearothermophilus*	75°C/pH6.5	10 min/73°C	E	83
Alkaline phosphatase	*Thermus sp.* Rt41A		5 min/85°C		84
α-Amylase	*Bacillus caldovelox* (70°C)		5.5-28 min/90°C (+ 5 mM CaCl₂)	E	85
α-Amylase	*B. stearothermophilus*	70°C	2 min/90°C/pH5.0 or 12.5 min at pH8.0	E, C, S, expr.	86, 87
α-Amylase (AmyA)	*Dictyoglomus thermophilum* (78°C)	90°C/pH5.5	70% active after 1h/90°C	E, C, S, expr.	88
α-Amylase (AmyB)	*D. thermophilum*	80°C/pH5.5	nd	E, C, S, expr.	89
α-Amylase (AmyC)	*D. thermophilum*	70°C/pH5.5	nd	E, C, S, expr.	89
β-Amylase	*Thermoanaerobacterium thermosulfurigenes* 4B (60°C)	75°C/pH5.5	>80% active after 1h/75°C	E, C, S, expr.	90

Amylopullulanase	Clostridium thermohydrosulfuricum E101 (65°C)	85-90°C/pH5.6	several h/85°C	E	91
Amylopullulanase	T. ethanolicus (69°C)	90°C/pH5.5	40 min/90°C	E, C, S, expr.	92
Amylopullulanase	Thermoanaerobacterium sp. Tok6-B1 (65°C)	80°C	several h/80°C	E	93

Enzyme	Organism	Temperature/pH	Stability	Characterization	Ref.
Amylopullulanase	Clostridium thermohydrosulfuricum E101 (65°C)	85-90°C/pH5.6	several h/85°C	E	91
Amylopullulanase	T. ethanolicus (69°C)	90°C/pH5.5	40 min/90°C	E, C, S, expr.	92
Amylopullulanase	Thermoanaerobacterium sp. Tok6-B1 (65°C)	80°C	several h/80°C	E	93
Amylopullulanase	Thermoanaerobacterium saccharolyticum (58°C)	75°C	45 min/75°C	E, C, S, expr.	94
Amylopullulanase	T. thermosulfurigenes EM1 (60°C)	70-75°C	nd	E, C, S, expr.	95
D-Asparaginase	T. aquaticus	>80°C/pH9.5	25 min/85°C		96
ATPase	S. acidocaldarius	75-85°C/pH6.2			see 97
Cyclodextrinase	T. ethanolicus	65°C/pH6.0	75 min/70°C	MB, C, S, expr.	98
Cyclodextrin glycosyltransferase	T. thermosulfurigenes EM1	hydrolysis (90-95°C) cyclization (80-85°C)	75% active after 5h/90°C (+ starch)	C, S, expr.	99
Endo β-1,4-glucanase	Clostridium thermocellum	62°C/pH5.2	nd		see 100
Endo β-1,4-glucanase	C. thermocellum	65°C/pH6.0	inactive after 30 min/80°C		see 100
Endo β-1,4-glucanase	C. thermocellum	60°C/pH6.4	1h/85°C		see 100
Endo β-1,4-glucanase	aCaldicellulosiruptor saccharolyticus (70°C)				
Endo-1,4-β-xylanase	C. saccharolyticus	70°C/pH5.5-7.7	20 min/75°C/pH6.0 (+BSA)	E, C, S, expr.	101, 102
Endo-1,4-β-xylanase	T. saccharolyticum	70°C/pH6.0	35 min/80°C or 40h/75°C or 8 min/90°C	E, C, S, expr.	103, 60
β-Galactosidase	Thermus 41A				see 70
α-Glucosidase	Bacillus sp.	75°C/pH5.5	10 min/75°C	C, S, expr.	104
α-Glucosidase	T. ethanolicus	75°C	35 min/75°C		105
β-Glucosidase	C. saccharolyticus		70 min/90°C		see 70
β-Glucosidase	C. thermocellum (60°C)	nd/pH6.0-6.5	60% active after 7h/60°C		see 100
α-Glucuronidase	Thermoanaerobacterium sp. (60°C)	60°C/pH5.4	1h/62°C		106
D-Hydantoinase	B. stearothermophilus	65°C/pH8.0	30 min/80°C	C, S	107
β-Mannanase	C. saccharolyticus				108

				C. S. expr.	
Neopullulanase	*B. stearothermophilus*	60-65°C			109
Proteinase	*Thermus sp. Rt4A2*	pH9.0	90% active after 1h/60°C	E	110
Proteinase	*Thermus sp. Rt41A*	90°C/pH8.0	90 min/80°C (+ 5 mM CaCl$_2$)	E	111
Proteinase	*Bacillus sp. Ak.1*	pH7.5	20 min/80°C (+ 5 mM CaCl$_2$) 13h/80°C or		112

Enzyme	Organism	Conditions	Stability	Methods	Ref.
Neopullulanase	*B. stearothermophilus*	60–65°C	90% active after 1h/60°C	C, S, expr.	109
Proteinase	*Thermus sp.* Rt4A2	pH9.0	90 min/90°C (+5 mM CaCl$_2$)	E	110
Proteinase	*Thermus sp.* Rt41A	90°C/pH8.0	20 min/90°C (+5 mM CaCl$_2$)	E	111
Proteinase	*Bacillus sp.* Ak.1	pH7.5	13h/80°C or 19 min/90°C (+5mM CaCl$_2$)		112
Pullulanase	*B. stearothermophilus*	65°C/pH6.0	stable for 1h/65°C	E	113
cPullulanase	*Bacillus sp.*	75°C	96h/70°C	E	114
Pullulanase	*Thermus sp.* AMD-33	70°C/ pH5.5–5.7		E, C, S	115, 116
cPullulanase	*Thermus aquaticus* YT-1	70–85°C?	>80% active after 10 h at 85°C		117
Pyrophosphatase	*Thermoplasma acidophilum* (60°C)	85°C/pH6.7		C, S	118,119
β-1,4-Xylanase	*Thermoanaerobacterium sp.*	80°C/pH6.2	1h/70°C		120
β-Xylosidase	*T. saccharolyticum*	70°C/pH5.5	55 min/75°C	C, S, expr.	121
β-Xylosidase	*C. saccharolyticus*	70°C/pH6.0–6.5	45 min/80°C	C, expr.	122
LYASES, ISOMERASES, AND LIGASES					
Fructose-1,6-diphosphate aldolase	*C. thermosaccharolyticum*	70°C	stable 4h/57°C		123
Citrate synthase	*T. acidophilum*	55°C	stable after 10 min/78°C	C, S, expr., CS	see 124
Glutamine synthetase	*B. stearothermophilus*		stable 5h/70°C (+Mn^{2+}/Mg^{2+}, glutamine, and NH$_4$Cl)		see 70
Phosphoenolpyruvate	*Methanobacterium thermoautotrophicum* (70°C)	65°C/pH7.4			125
Triose phosphate isomerase Xylose isomerase	*C. thermosaccharolyticum* *B. stearothermophilus*	62°C/pH8.6	10 min/64°C		126

Xylose isomerase	*T. saccharolyticum*	80°C/pH7.5	nd	C. S. expr.	127,128
	T. thermosulfurigenes	80°C/pH7.5	42 h/70°C (+ MgCl$_2$ and CoCl$_2$)	C. S. expr.	62,127
Xylose isomerase	*T. aquaticus* HB8	85°C/pH7.0	4d/70°C		129
Xylose isomerase	*T. thermophilus*	nd	nd	C. S. expr.	130,131

a Formerly *Caldocellum saccharolyticum* [225].

c The activity of this pullulanase on starch has not been characterized.

Xylose isomerase	T. saccharolyticum	80°C/pH7.5	nd	C, S, expr.	127,128
Xylose isomerase	T. thermosulfurigenes	80°C/pH7.5	42 h/70°C (+ MgCl2 and CoCl2)	C, S, expr.	62,127
Xylose isomerase	T. aquaticus HB8	85°C/pH7.0	4d/70°C		129
Xylose isomerase	T. thermophilus	nd	nd	C, S, expr.	130,131

[a] Formerly *Caldocellum saccharolyticum* [225].

[c] The activity of this pullulanase on starch has not been characterized.

C: cloned; S: sequenced; CS: crystal structure available; E: extracellular; MB: membrane associated; MTHF: methyltetrahydrofolate ;
GAPDH: Glyceraldehyde-3-phosphate dehydrogenase; exp.: recombinantly expressed.

Table 2. *Thermozymes from hyperthermophilic organisms*

Enzyme	Organism (optimal growth temperature)	Enzyme thermophilicity	Enzyme thermostability	General comments	Ref. No.
OXIDOREDUCTASES					
Alcohol dehydro...					

Table 2. Thermozymes from hyperthermophilic organisms

Enzyme	Organism (optimal growth temperature)	Enzyme thermophilicity	Enzyme thermostability	General comments	Ref. No.
OXIDOREDUCTASES					
Alcohol dehydrogenase (NAD-specific)	*Sulfolobus solfataricus* (70–85°C)	>95°C/pH7.5-8.5	5h/70°C		10
Alcohol dehydrogenase (NADP-specific)	*Thermococcus litoralis* (88°C)	80°C/pH8.8	2h/85°C		11
Aldehyde ferredoxin oxido-reductase	*Pyrococcus furiosus* (100°C)	>90°C/pH9.0-10.0	6h/80°C	CS	132,133
Ferredoxin (protein electron carrier)	*P. furiosus*	>95°C	stable for 12h (95°C)		134
Formaldehyde oxidoreductase	*T. litoralis* ES4 (100°C)	95°C	2h/80°C		see 64
Glutamate dehydrogenase (NADP-specific)	*P. furiosus*		3.5h/105°C		135
Glutamate dehydrogenase (NAD/NADP)		75-95°C/pH9.0	2-12 h/100°C		136,137,138
Glutamate dehydrogenase (NAD/NADP)	*S. solfataricus*	70°C/pH10.0	15h/80°C		139
Glutamate dehydrogenase (NADP-specific)	*T. litoralis*	>95°C/pH8.0	2h/98°C		11
GAPDH	*Thermotoga maritima* (80°C)	>75°C/pH6.0-8.0	>2h/100°C/pH6.0	C, S, expr.	140,141,142
GAPDH	*Pyrococcus woesi* (100°C)	nd	44 min/100°C	C, S, expr.	143,144
GAPDH (NAD)	*Thermoproteus tenax* (88°C)	nd	≥20 min/100°C		145
GAPDH (NADP)	*T. tenax*	nd	35 min/100°C		145
GAPDH	*Methanothermus fervidus* (82°C)	nd	20 min/83°C		144
Hydrogenase (F420-reactive)	*Methanococcus jannaschii* (85°C)	80-90°C/pH7.0-10.0	16-25 min/75°C		146
Hydrogenase (F420-non reactive)	*M. jannaschii*	80°C/pH9.0	107 min/75°C		146
Hydrogenase (H2 producing)	*P. furiosus*	>95°C	2h/100°C		147
Hydrogenase	*Pyrodictium brockii* (105°C)	>90°C	15 min/76°C		148

Hydrogenase (H₂ producing)					
L-Lactate dehydrogenase	T. maritima	>95°C/ pH8.6-9.5	50 min/80°C		149
Malate dehydrogenase	T. maritima	>95°C/pH7.0	30min/85°C	C, S, expr.	150,151
CH₂=H₄MPT dehydrogenase	M. fervidus			C, S	152
(F₄₂₀-dependent)	Methanopyrus kandleri (98°C)	75°C/pH4.5-6.5/2.0 M salt	100% active after 1h/90°C		153
CH₂=H₄MPT dehydrogenase	M. kandleri	>90°C/1.3M salt	cell extract 100%		154
(H₂-forming)					

Enzyme	Organism	Conditions	Stability		Ref.
Hydrogenase (H₂ producing)	T. maritima	>95°C/pH8.6-9.5	50 min/90°C		149
L-Lactate dehydrogenase	T. maritima	>95°C/pH7.0	30min/85°C	C, S, expr.	150,151 152
Malate dehydrogenase	M. fervidus			C, S	152
CH₂=H₄MPT dehydrogenase (F₄₂₀-dependent)	Methanopyrus kandleri (98°C)	75°C/pH4.5-6.5/2.0 M salt	100% active after 1h/90°C		153
CH₂=H₄MPT dehydrogenase (H₂-forming)	M. kandleri	>90°C/1.3M salt	cell extract 100% active after 1h/90°C pure enzyme rapidly inactivated (90°C)		154
CH₂=H₄MPT reductase (F₄₂₀-dependent)	M. kandleri	90°C/pH6.5-7.0 in 2.2-2.5 M SO₄²⁻, PO₄²⁻	rapidly inactivated at 90°C (no salt) 100% active after 1h/90°C (100 mM K₂HPO₄)		155
F420 dependent NADP reductase	Archaeoglobus fulgidus (83°C)	80°C/pH8.0	16-17 min/90°C (no salt) stable at 90°C (1M K2HPO4)		156
Pyruvate: ferredoxin oxidoreductase	A. fulgidus	>90°C/pH7.5	60 min/90°C (+2M KCl)		157
Pyruvate:ferredoxin oxidoreductase	P. furiosus	>90°C/pH8.0	23 min/80°C or 18 min/90°C		158,159
Pyruvate:ferredoxin oxidoreductase	T. maritima	>90°C/pH6.3	15h/80°C or 11h/90°C		159
Rubredoxin	P. furiosus	>95°C	stable for 24h (95°C)		160

TRANSFERASES

Enzyme	Organism	Conditions	Stability		Ref.
ATP sulfurylase	A. fulgidus	90°C, pH8.0			see 64
4-a-Glucanotransferase	T. maritima	70°C	3h/80°C		161,see 64
DNA polymerase II (vent polymerase)	T. litoralis	75°C	7h/95°C	C, S, expr.	64
DNA polymerase II	P. furiosus	>75°C	20h/95°C	C, S, expr.	162,see 64

Enzyme	Organism	Conditions	Stability	Ref.
DNA Polymerase III	*Desulfurococcus amylolyticus* (87°C)			163
CHO-H$_4$MPT formyltransferase	*M. kandleri* (98°C)	90°C/pH6.5/ 2 M SO$_4^{2-}$, PO$_4^{2-}$	~100% stable (90°C) + 1.5 M K$_2$HPO$_4$ (unstable (-) salt)	164
2-Phosphoglycerate kinase	*M. fervidus*		C, S, expr.	61

Enzyme	Organism	Conditions	Stability/Activity	Notes	Ref.
DNA Polymerase III	*Desulfurococcus amylolyticus* (87°C)				163
CHO-H4MPT formyltransferase	*M. kandleri* (98°C)	90°C/pH6.5/ 2 M SO4²-, PO4²-	~100% stable (90°C) + 1.5 M K2HPO4 (unstable (-) salt)		164
2-Phosphoglycerate kinase	*M. fervidus*			C, S, expr.	61
HYDROLASES					
α-Amylase (extracellular)	*P. furiosus*	100°C/pH5.5	3.2h/110°C		165
α-Amylase (intracellular)	*P. furiosus*	100°C/ pH6.5-7.5	85% active after 3h at 100°C	C, S, expr.	166,167
α-Amylase	*P. woesi*	100°C/pH5.5	4h/110°C		168
α-Amylase	*Thermococcus profundus* (80°C)	80°C/pH5.5-6.0	4h/90°C (5mM CaCl2)		169
β-Amylase	*T. maritima*	95°C/pH5.0	30 min/90°C		170
α-Amylase/glucoamylase	*T. maritima*	90°C/pH6.0	30 min/90°C		170
Amylopullulanase	*P. furiosus*	125°C/pH5.5 (5mM CaCl2)	12 min/120°C (5mM CaCl2 and PG7)		171
Amylopullulanase	ES4	110-125°C	20h/98°C (5mM CaCl2)		172
Amylopullulanase	*T. litoralis*	117°C/pH5.5 (5mM CaCl2)	5 min/120°C (5mM CaCl2 and PG7)		171
ATPase complex	*Pyrodictium occultum* (105°C)	100°C	30 min/110°C		173,174
Carboxypeptidase	*S. solfataricus*	85°C/pH5.5-9	13-14 min/90°C (holoenzyme) 14 min/80°C (apoenzyme)		175,176
Endo-1,4-β-glucanase	*T. maritima*	95°C/pH6.0-7.5	2h/95°C		177
Endo-1,4-β-xylanase	*Thermotoga sp.* (80°C)	105°C/ pH5.0-5.5	90 min/95°C		178
Exo-1,4-β-cellobiohydrolase	*Thermotoga sp.*	105°C/PH7.0	70 min/108°C		179
β-Galactosidase	*S. solfataricus*		24h/75°C or 3h/85°C		see 64
Exo-1,4-β-glucanase	*T. maritima*	95°C/pH6.0-7.5	30 min/95°C		177

Enzyme	Organism	Conditions	Stability	Ref.
β-Glucosidase	*Thermococcus celer* (87°C)		20 min/105°C	see 64
β-Glucosidase	*Thermotoga sp.*	nd/PH7.0	2.5h/98°C	180
α-Glucosidase	*P. furiosus*	110°C/ pH5.0-6.0	(40µg/ml BSA) 46–48h/98°C	181
Membrane-bond ATPase	*Sulfolobus*			182
CH2=H4MPT cyclohydrolase	*A. fulgidus*	8.5°C/pH18.5	100% stable after	164

17

Enzyme	Organism	Temp/pH	Stability	Notes	Ref
β-Glucosidase	Thermococcus celer (87°C)	nd/PH7.0	20 min/105°C		see 64
β-Glucosidase	Thermotoga sp.		2.5h/98°C (40μg/ml BSA)		180
α-Glucosidase	P. furiosus	110°C/ pH5.0-6.0	46-48h/98°C		181
Membrane-bond ATPase	Sulfolobus				182
CH2=H4MPT cyclohydrolase	A. fulgidus	85°C/pH8.5	100% stable after 50 min (90°C) with 1 M K2HPO4 (unstable (-) salt)		164
CH2=H4MPT cyclohydrolase	M. kandleri	95°C/pH8.0/ 1-2 M salt	100% active after 60 min/90°C		183
Protease	Desulfurococcus mucosus (85°C)	100°C	1.5h/95°C		see 64
Serine protease (pyrolysin)	P. furiosus	115°C	4h/100°C or 20 min/105°C		184
Sucrose α-hydrolase	P. furiosus	105°C	48h/95°C		see 64
Thiol protease	Pyrococcus sp. (95°C)	110°C/pH7.0	60 min/100°C		185
β-Xylosidase/ arabinofuranosidase	Thermotoga sp.	nd/pH7.0	4h/98°C (40μg/ml BSA)		186,180

LYASES, ISOMERASES, AND LIGASES

Enzyme	Organism	Temp/pH	Stability	Notes	Ref
DNA topoisomerase V	M. kandleri	active to 100°C			187
Glutamine synthetase	P. furiosus	85-90°C	2h/100°C		188
Glutamine synthetase	P. woesei	nd	nd	C, S	189
Xylose isomerase	T. neapolitana	97°C/pH7.1	24 min/95°C	C, S, expr.	190
Xylose isomerase	T. maritima	105-110°C	10 min/120°C		191

C: cloned; S: sequenced; expr.: recombinantly expressed; CH2=H4MPT: N5,N10-methylenetetramethanopterin; CHO-H4MPT: N5-formyl-methanopterin; GAPDH: Glyceraldehyde-3-phosphate dehydrogenase.

protein, allowing

thermophilic enzy

were extensively

fluorescence emis

(GuHCl)-dependa

constants, the san

enzyme [141]. *P*

the same Michae

P. woesei enzyme

temperature depe

thermostability w

dependence on pe

enzyme adaptatio

As more se

proteins, most evic

thermophilic orgar

enzymes contain t

[135,152,160,190]

the same 3-dimens

structural features

temperatures and t

to define amino aci

trying to explain th

stabilization by sul

destabilizing resid

by introducing pro

increased number

protein, allowing the use of standard molecular biology techniques and hosts for

thermophilic enzyme studies. *T. maritima* and *Pyrococcus woesei* recombinant GAPDHs

were extensively studied. The *T. maritima* recombinant GAPDH had the same

fluorescence emission and circular dichroism (CD) spectra, the same Guanidine-HCl

(GuHCl)-dependent denaturation and heat-inactivation behaviors, the same Michaelis

constants, the same allosteric inhibitor effect, and same specific activity as the native

enzyme [141]. *P. woesei* recombinant GAPDH had the same heat-inactivation behavior,

the same Michaelis constants and V_{max}, and the salt effect on stabilization as the native

P. woesei enzyme [143]. Thus, proper thermophilic protein folding was not strongly

temperature dependent, and all the information necessary for protein thermophilicity and

thermostability was encoded in the primary structure of the enzyme. This thermal property

dependence on peptide sequence suggests a direct genetic mutational mechanism for

enzyme adaptation to the thermal environment of its original source.

As more sequence information is gathered on thermophilic and thermostable

proteins, most evidence indicates that the catalytic machinery from thermophiles, like

thermophilic organism genetics, is homologous to that from mesophiles. Thermophilic

enzymes contain the same amino acids, the same catalytic consensus regions

[135,152,160,190], and, in many cases, enzyme structural analysis reveals that they share

the same 3-dimensional backbone as their mesophilic counterparts [151,160]. No distinct

structural features unique to thermophilic enzymes can account for their activity at elevated

temperatures and their resistance to heat denaturation. However, attempts have been made

to define amino acid "traffic rules" for high temperature molecular adaptation. Hypotheses

trying to explain the increased stability of thermostable proteins include: (i) α–helix

stabilization by substituting helix-stabilizing residues (such as alanines) for helix-

destabilizing residues (most often glycines) [192]; (ii) short-loop and β–turn rigidification

by introducing prolines [193]; (iii) increased average protein hydrophobicity [85,143]; (iv)

increased number of salt-bridges [194]; and (v) a decrease in cysteine and deamidable

residues. Howev

and sequenced, ev

rules.

No trend i

from the thorough

spanning an 82°C

mesophilic, therm

families are comp

that features accou

another. Enzymes

temperatures above

Table 3, with few

their cysteine or de

thermophilic count

isomerases. The A

their temperature fo

Asn+Gln content ar

authors acknowledg

hyperthermophiles

belief that thermoph

composition, enzym

differences. These

their immediate envi

flexibility. These pro

show that no simple,

Therefore, elucidating

thermophilicity will r

residues. However, as more genes from thermophiles and hyperthermophiles are cloned and sequenced, examples accumulate of thermophilic enzyme sequences contradicting these rules.

No trend in thermophilic GAPDH total amino acid compositions could be identified from the thorough comparison of GAPDH sequences from twenty-six different organisms spanning an 82°C growth temperature range [195]. In Table 3 the amino acid contents of mesophilic, thermophilic, and hyperthermophilic enzymes from nine different enzyme families are compared. No prominent general trend is observed and it is highly probable that features accounting for protein thermostabilization differ from one enzyme class to another. Enzymes from hyperthermophiles, in particular extracellular enzymes stable at temperatures above 100°C, show a high resistance to amino acid covalent modification. In Table 3, with few exceptions, enzymes from hyperthermophiles show a slight decrease in their cysteine or deamidable residues contents compared to their mesophilic and thermophilic counterparts. This trend is particularly noticeable among type II xylose isomerases. The Asn+Gln contents of four type II xylose isomerases was compared to their temperature for maximal activity [190], showing an inverse relationship between Asn+Gln content and temperature for maximal activity. Although it is interesting, the authors acknowledge that more information is needed on proteins from thermophiles and hyperthermophiles to confirm the statistical significance of this trend. Contrary to the early belief that thermophilic enzymes obtained their thermostability from a different amino acid composition, enzyme thermostability appears to be enhanced by numerous subtle sequence differences. These differences protect the highly labile residues and bonds (by modifying their immediate environment), they limit destabilizing interactions, and they restrain protein flexibility. These protein alignments and amino acid composition comparisons clearly show that no simple, general protein thermostabilization traffic rule can be defined. Therefore, elucidating the molecular mechanisms of protein thermostability and thermophilicity will require careful study of protein molecular structural interactions and

Table 3. Amino acid composition comparison of enzymes from mesophilic, thermophilic, and hyperthermophilic prokaryotes

Protein Class	Gly	Pro	Ala	Val	Leu	Ile	Phe	Trp	Tyr	His	Cys	Met	Ser	Thr	Lys	Arg	Asp	Glu	Asn	Gln	Asn + Gln	H-phob	Polar	Charged
Hydrolases																								
α-amylases																								
meso (13)	8.5	4.5	8.8	5.7	8.0	4.4	3.8	2.9	4.1	2.5	1.4	2.0	7.6	6.8	4.1	5.5	5.8	3.7	5.2	4.5	9.7	51	20	29
therm (4)	8.9	4.8	8.1	6.4	6.9	3.8	4.4	4.0	5.9	2.3	0.75	1.9	5.6	8.3	5.1	4.1	7.4	3.6	4.0	3.3	7.3	53	19	28
hyper (1)	6.8	4.3	4.0	9.2	9.1	6.3	5.7	2.6	6.3		0.31	1.8	5.1	1.8	7.4	5.4	5.4	10	3.7	2.3	6.0	54	11	34
neutral proteases																								
meso (11)	7.7	3.7	8.4	6.6	7.4	4.2	3.2	0.90	5.2	2.3	0.57	1.2	8.6	6.7	7.3	4.4	5.9	4.7	6.8	4.3	11	47	19	33
therm (2)	12	3.3	9.4	7.7	7.0	4.3	3.0	1.3	6.9	2.0	0.37	1.9	6.2	6.3	3.7	4.8	6.0	4.2	5.4	4.3	9.7	55	17	28
Oxidoreductases																								
1° ADHs																								
meso (5)	8.2	7.4	11	6.4	7.4	4.9	2.7	1.4	2.6	2.1	1.6	2.8	7.2	5.3	5.0	8.4	5.0	4.0	3.8	2.5	6.3	52	19	29
therm (1)	10	4.7	9.9	12	8.5	6.4	1.9	0.88	3.5	2.3	2.0	1.7	3.4	3.5	7.5	3.8	4.7	7.0	2.9	3.2	6.1	58	13	29
hyper (1)	11	3.5	8.4	7.5	7.9	8.9	3.2	0.63	3.0	2.1	2.3	3.0	4.8	5.1	5.5	4.9	5.7	4.9	5.3	2.7	8.0	54	17	29
2° ADHs																								
meso (2)	11	3.5	8.4	7.5	7.9	8.9	3.2	0.63	3.0	2.1	2.3	3.0	4.8	5.1	5.5	4.9	5.7	4.9	5.3	2.7	8.0	54	17	29
therm (1)	12	6.2	9.9	10	6.5	7.4	4.0	0.14	0.17	2.8	0.11	4.3	0.26	3.7	6.8	3.1	5.7	6.0	2.8	0.11	2.9	60	14	26
Glu DHs																								
meso (4)	11	3.6	9.5	8.8	7.0	4.9	4.0	1.6	3.1	1.5	1.2	3.6	4.9	5.0	6.1	4.5	4.6	7.8	4.3	3.5	7.8	53	16	31
hyper (3)	8.1	4.6	9.4	8.4	5.5	7.5	2.3	2.4	4.8	0.87	0.32	3.4	3.4	5.7	8.1	4.7	6.0	8.3	3.6	2.5	6.1	53	14	33
LDHs																								
meso (10)	6.8	4.6	10	6.8	8.8	7.3	3.8	1.1	3.1	2.4	0.94	2.4	6.1	5.4	6.3	4.9	5.1	6.0	4.5	3.6	8.1	52	17	30
therm (5)	9.4	3.6	12	10	8.9	6.6	3.7	0.95	3.3	2.2	0.38	1.3	4.1	4.1	3.5	6.8	5.6	6.8	3.2	2.6	5.8	59	12	28
hyper (1)	9.6	3.5	8.6	9.5	10	7.5	4.0	n.d.	2.8	2.0	n.d.	2.4	4.1	4.9	6.5	4.3	n.d.	n.d.	n.d.	n.d.	10	56	13	32
MDHs																								
meso (1)	11	4.5	12	10	10	6.4	0.29	0.0	0.13	0.96	0.96	0.13	5.8	5.4	7.0	0.26	3.8	6.4	3.5	4.2	7.7	58	14	28
therm (1)	8.2	4.9	15	8.6	9.2	5.2	3.0	0.12	0.21	0.12	0.30	3.4	0.28	4.3	5.5	5.8	5.2	7.3	3.4	3.7	7.1	57	12	31
hyper (1)	9.7	4.4	7.7	6.5	8.0	11	0.26	0.29	0.24	0.18	0.59	0.26	4.7	5.0	9.7	0.24	5.6	8.8	5.3	0.88	2.2	52	15	33
Isomerases																								
D-xylose isomerases																								
meso (8)	8.7	5.8	11	5.5	8.8	3.4	4.1	2.0	2.5	3.2	1.5	2.1	4.4	5.1	4.4	9.6	5.6	5.3	3.1	3.9	7.0	52	16	32
therm (3)	6.5	3.0	9.0	4.8	7.4	5.1	8.0	1.1	4.8	2.6	0.53	2.7	3.8	5.3	8.5	3.9	7.7	7.5	5.1	2.4	6.5	50	15	35
hyper (1)	7.6	3.6	7.9	5.0	9.0	5.6	7.9	0.11	4.3	2.5	0.68	0.18	2.7	5.6	8.6	4.7	7.2	9.0	3.4	0.18	3.6	52	13	35
Ligases																								
Gln synthetases																								
meso (12)	7.0	5.3	8.1	6.2	8.3	5.6	5.0	1.1	3.8	2.7	1.2	3.0	5.7	5.3	6.9	4.3	6.0	7.2	4.4	2.7	7.1	50	18	32
therm (1)	7.2	6.8	6.8	6.6	8.7	5.5	4.7	0.13	5.9	2.1	0.21	2.3	5.5	5.9	6.6	5.1	6.8	7.0	2.7	2.1	4.8	54	16	30
hyper (2)	7.3	5.9	7.1	6.3	8.2	7.2	5.5	1.4	5.1	2.4	0.34	2.6	4.4	4.0	7.3	4.7	5.8	10	3.5	0.91	4.4	54	14	32

n.d. - not determined

meso, enzymes from organisms with optimal growth temperatures below 60°C; therm, enzymes from organisms with optimal growth temperatures from 60-80°C; hyper, enzymes from organisms with optimal growth temperatures above 80°C.

Numbers in parentheses indicate the number of proteins represented in the category.

GenBank database accession numbers or specific references for the proteins used in the comparison: Primary alcohol dehydrogenase (1° ADH): meso - D90004, L02104, M91440, X17065, X59263; therm - D90421; hyper - S51211. Secondary alcohol dehydrogenase (2° ADH): meso - J03362, M84723; therm - [196]. Lactate dehydrogenase (LDH): meso - D13405, L29327, M22305, M72545, M82881, M95919, X01067, X55118, X55119, Z22737; therm - D00585, M19394, M19396, M28336, X04519; hyper - ref #320. Glutamate dehydrogenase (Glu DH): meso - K02499, M24021, M65250, M76403; hyper - L12408, L19995, M97860. Malate dehydrogenase (MDH): meso - M95049; therm - J02598; hyper - X51714. Glutamine synthetase (Gln synth): meso - D00513, D10020, L05609, L08256, M14536, M16626, M18966, M22811, M26107, M57275, X04880, X53509; therm - X53263; hyper - X60160, X60161. Neutral protease: meso - D10773, K01985, K02497, M36694, M36723, M62845, M64809, M83910, X61286, X73315, X75070; therm - M11446, M63575. α-amylase: meso - J01542, K00563, L19299, M18244, M24516, M25263, U04956, X07796, X12725, X12726, X52755, X52756, X55799; therm - M11450, M34957, M57457, X02769; hyper - L22346. D-xylose isomerase (XI): meso - L12967, M15050, M36269, M73789, M84564, X00772, X02795, X61059; therm - J05650, L09699, M91248; hyper - L38994.

energetics. Therm

organisms have be

Thermostability

Protein the

Inactivation is due

protein regions, or

believed to be a ra

Enzyme-specific, t

particular environm

remaining (determi

range) in a sample a

Tomazic and Kliban

native enzyr

(active)

Representations of th

usually linear, fitting

t=incubation time).

ln(res

Equation (1) contains

step is intramolecular

order rate law reflects

inactivation at various

energetics. Thermostable enzymes isolated from thermophilic and hyperthermophilic organisms have been termed thermozymes [71].

Thermostability

Protein thermostability is the protein's resistance to irreversible thermal inactivation. Inactivation is due either to protein chemical modification which destroys specific critical protein regions, or to protein denaturation (i.e., irreversible unfolding). Denaturation is believed to be a rapid cooperative process, once a threshold molecular energy is reached. Enzyme-specific, this threshold is related to the characteristic protein stabilizing energy in a particular environment. Thermostability is typically measured as the fraction of activity remaining (determined by standard enzyme assay in the enzyme's activity temperature range) in a sample after heating it at a constant temperature for a specific period of time. Tomazic and Klibanov proposed a general model for enzyme thermoinactivation [197]:

native enzyme	$\Leftrightarrow$	non-native	$\rightarrow$	scrambled structures	[scheme 1]
(active)		(inactive)		(often precipitate)	
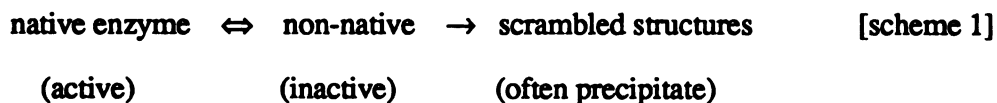
Representations of the natural log of enzyme residual activity plotted versus time are usually linear, fitting a pseudo-first order equation {Equation (1), with k=rate constant and t=incubation time}.

$$\ln(\text{residual activity}) = -kt \qquad\qquad [1]$$

Equation (1) contains no protein concentration term, suggesting that the rate-determining step is intramolecular with respect to protein. The only way to ensure that a pseudo-first order rate law reflects a truly intramolecular process however, is to measure thermal inactivation at various initial protein concentrations, verifying that k remains constant.

Tomazic an~

thermal inac~

Prote~

protein mole~

thermophilic

recombinant

and in differe~

native enzym

thermostabili

translational ~

Irreversible ir

denaturation,

protein chemi~

information re

(typically 30-(

forces [203].

approximately

hydrogen bon~

modification ra

.

thermostability

significant at te~

asparagine, and~

enzymes to tem~

temperatures ab~

strong enough to

by peptide bond

Tomazic and Klibanov's model is consistent with an intramolecular rate-determining step in thermal inactivation if the native to non-native protein step rate determining.

Protein stability is due to numerous ionic and nonionic interactions within the protein molecule and between the protein and the environment [198]. While some thermophilic enzymes have been shown to be stabilized by glycosylation [199,200], most recombinant thermozymes - expressed in mesophilic hosts, in the absence of glycosylation, and in different cellular environments - remain as thermophilic and thermostable as the native enzymes [141,143]. This observation demonstrates that thermophilicity and thermostability are encoded in the peptide sequence; they are neither a consequence of post-translational modifications, nor of non-covalent interactions with cellular components. Irreversible inactivation of mesophilic proteins typically results from irreversible denaturation, due to the disruption of numerous non-covalent interactions rather than to protein chemical decomposition [201]. Thus, the original protein conformational information remains intact in many denatured proteins. The protein stabilizing energy (typically 30-65 kJ mol^{-1} [202]) is 10-fold smaller than the magnitude of the opposing forces [203]. Since each additional hydrogen bond or salt bridge can contribute approximately 2-20 kJ mol^{-1} to the stabilizing energy, and since the possibilities to add hydrogen bonds or salt bridges in a protein are vast, it is likely that the protein covalent modification rates, rather than unfolding, determine the theoretical upper limit of protein thermostability. Ahern and Klibanov [204] reported that peptide depolymerization becomes significant at temperatures above 100°C, but oxidative decomposition of cysteine, asparagine, and glutamine residues has been shown to limit the thermostability of some enzymes to temperatures below 100°C. The identification of enzymes active at temperatures above 120°C suggests that the non-covalent protein structural interactions are strong enough to allow stability to approach the theoretical maximum temperatures defined by peptide bond destruction.

Thermophilic

While th

from 65 to 70°C

Thermophilic an

with variations i

conditions of hig

directly involved

denaturation are

must be more res

activities similar

atom positions kn

active temperature

Differential scanni

reasonably constar

absence of signific

enzymes are typica

functional architect

temperature ranges

expected to cause n

mesophilic and ther

typical Arrhenius-li

thermophilic enzym

been proposed that t

and thermophilicity [

are verified by carefu

Arrhenius plots obtai

respond similarly to te

Thermophilicity

While the system kinetic energy increase (~RΔT) is the same from 25 to 30°C or from 65 to 70°C, the total molecular kinetic energy is greater at high temperatures. Thermophilic and mesophilic enzymes similarly resist the structural changes associated with variations in their molecular energy, but thermophilic enzymes must do so under conditions of higher total kinetic energy. Therefore, while protein denaturation is not directly involved in enzyme thermophilicity, structural interactions which oppose denaturation are very important for thermophilicity. In other words, thermophilic enzymes must be more resistant to denaturation, while having dynamic structural states allowing activities similar to those of their mesophilic counterparts. Because of sub-Å variations in atom positions known to be tolerated in a functional active site, it is believed that, within its active temperature range, an enzyme maintains its average structure within strict limits. Differential scanning calorimetry measurements indicate that protein heat capacities remain reasonably constant within their catalytically active temperature ranges, indicating an absence of significant structural changes. Arrhenius plots for thermophilic and mesophilic enzymes are typically linear, suggesting that mesophilic and thermophilic enzyme functional architectures are similarly, tightly controlled throughout their respective temperature ranges (significant structural changes that alter the functional architecture are expected to cause non-Arrhenius behavior). Biphasic Arrhenius plots reported for some mesophilic and thermophilic enzymes [82,145,205] represent an important exception to the typical Arrhenius-like behavior, however, discontinuities are not a specific trait of thermophilic enzymes. In the cases of yeast and *Thermoproteus tenax* GAPDHs, it has been proposed that the discontinuity extent indicates the degree of enzyme thermostability and thermophilicity [145]. It is critical however, that such interpretations of Arrhenius data are verified by careful biophysical and kinetic experimentation. More likely, the similar Arrhenius plots obtained for mesophilic and thermophilic enzymes suggest that all proteins respond similarly to temperature.

In their res

enzymes are more

personal communi

mesophilic counter

temperatures is obs

temperature rigidit

activity at these tem

structural change o

enzyme V_{max} (k_{cat})

throughout their ac

relationship predicts

temperature, a set n

predominantly a fur

predicts that a more

activity curve than a

Therefore, the Arrhe

thermophilic analog

energy is controlling

change). Recent wor

T. thermosulfurigene

[206]. Thus, the perc

temperature for an en

These data clearly sh

Also, the excellent ag

Arrhenius theory sug

no catalytically signifi

In their respective temperature ranges, psychrophilic, mesophilic, and thermophilic enzymes are more rigid at low temperatures than at higher temperatures (G. Petsko, personal communication). Thermophilic enzymes are significantly more rigid than their mesophilic counterparts at room temperature [198], however, increased rigidity at low temperatures is observed in all enzyme categories. It has been proposed that the mesophilic temperature rigidity of thermophilic enzymes [198,203] was responsible for their low activity at these temperatures [140]. This implies that a catalytically significant enzyme structural change occurs between low and high temperatures. However, the increasing enzyme V_{max} (k_{cat}) values with temperature seen for thermophilic and mesophilic enzymes throughout their active ranges are typically consistent with the Arrhenius relationship. This relationship predicts that the percent of maximal enzyme activity measured at some temperature, a set number of degrees below the optimal activity temperature, is predominantly a function of the reaction activation energy (Ea) (Fig. 1). Thus, the theory predicts that a more thermophilic enzyme would have only a slightly broader temperature-activity curve than an enzyme with a lower optimal temperature and the same Ea. Therefore, the Arrhenius equation can explain poor low temperature activity for thermophilic analogs to mesophilic enzymes while assuming that only the system thermal energy is controlling the reaction rate (there is no catalytically significant enzyme structural change). Recent work in our laboratory demonstrated that *E. coli*, *B. stearothermophilus*, *T. thermosulfurigenes*, and *T. neapolitana* xylose isomerase Arrhenius plots were linear [206]. Thus, the percent of maximal activity at any temperature below the optimal temperature for an enzyme depended principally on the activation energy of the reaction. These data clearly showed low thermophilic enzyme activity at mesophilic temperatures. Also, the excellent agreement between the observed temperature-activity dependence and Arrhenius theory suggested that throughout the active temperature range for each enzyme no catalytically significant structural changes occurred.

Figure 1. Arrhenius dependence of reaction rate on the temperature and activation energy for mesophilic and hyperthermophilic enzymes. T_{max} is defined as the temperature for maximal enzyme activity and Ea is the reaction activation energy.

Relative activity

Figure legend:
- ■ Ea = 80 kJ/mol, Tmax = 90°C
- □ Ea = 80 kJ/mol, Tmax = 37°C
- ● Ea = 20 kJ/mol, Tmax = 90°C
- ○ Ea = 20 kJ/mol, Tmax = 37°C

Y-axis: Relative activity

X-axis: Change in temperature from maximal (°C)
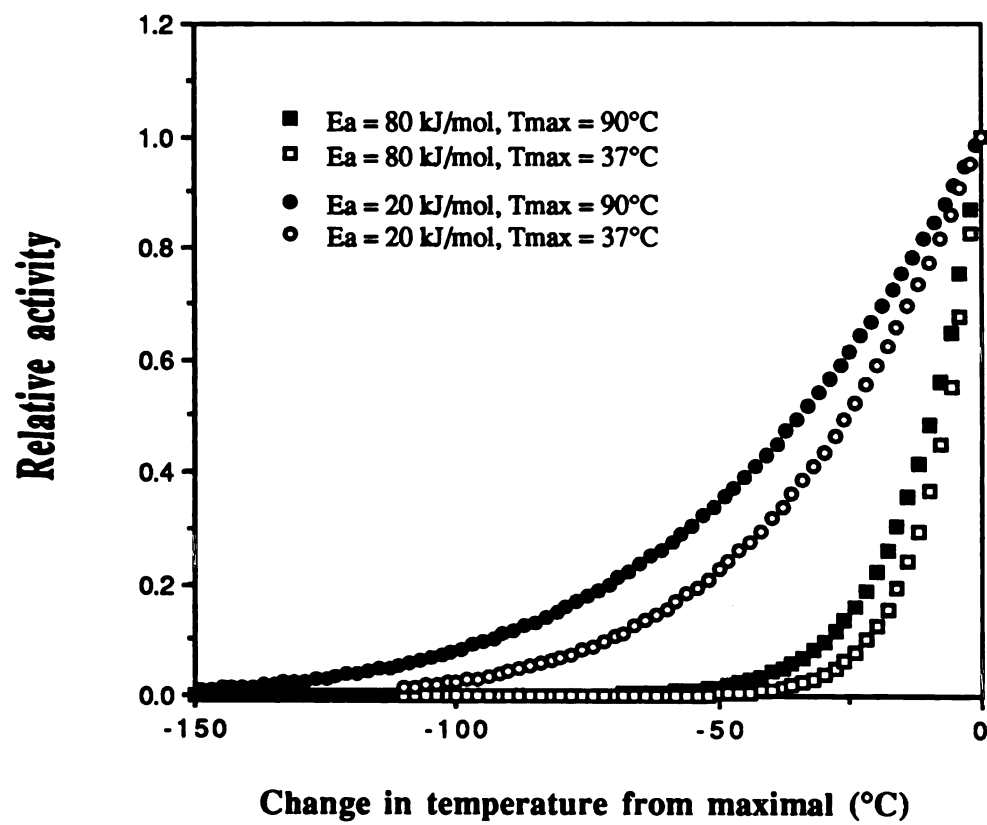
The Arrhe

been attributed to a

rigid, inefficient ca

Although a discon:

discontinuities can

other phenomena r.

thermophilic enzyr:

in enzyme structure

temperature activity.

As kinetic c

despite their activit:

V_{max} and K_m value

optimal temperature

have high catalytic e

assumption of simil.

Instead, thermophili

temperatures are typ

Substrate affinities f

temperatures are also

yield similar optimal

mesophilic enzymes (

be attributed to incre

possessing kinetic ene

dependence of K_m is

(equation 1)

The Arrhenius plot discontinuities reported for some thermophilic enzymes have been attributed to a catalytically significant enzyme structural change from an excessively rigid, inefficient catalyst to a less rigid sturcture optimized for enzymatic activity [140]. Although a discontinuous plot is not absolutely consistent with theory, Arrhenius discontinuities can result from temperature dependent changes in the reaction slow step or other phenomena not linked to enzyme structure [207,208]. However, whether thermophilic enzyme Arrhenius plot discontinuities indicate a temperature dependent change in enzyme structure and whether such a change can account for the low mesophilic temperature activity of thermophilic enzymes, remains to be determined.

As kinetic data accumulate on thermophilic enzymes, it becomes evident that, despite their activity at high temperatures, thermophilic enzymes catalyze reactions with V_{max} and K_m values similar to those of their mesophilic counterparts at their respective optimal temperatures. Thermophilic enzyme-catalyzed reactions were initially expected to have high catalytic efficiencies based on extrapolations of mesophilic enzyme rates and the assumption of similar substrate affinities for the mesophilic and thermophilic enzymes. Instead, thermophilic and mesophilic enzyme V_{max} values at their respective optimal temperatures are typically similar, as is typified by a series of xylose isomerases (Table 4). Substrate affinities for thermophilic and mesophilic enzymes at their respective optimal temperatures are also usually similar (Table 4,5) [215]. These similar K_m and V_{max} data yield similar optimal temperature catalytic efficiencies among analogous thermophilic and mesophilic enzymes (Table 4). While higher reaction rates with increasing temperature can be attributed to increased rates of collision and a larger fraction of the substrate population possessing kinetic energies above the reaction activation energy, the temperature dependence of K_m is less clear [208]. Even for the simple Michaelis-Menten case (equation 1)

Table 4. Kinetic constants for type II xylose isomerases purified from mesophiles and thermophiles[a]

Organism	Temperature (°C)	Km for xylose (mM)	Vmax (µmol/min/mg)	Kcat (1/min)	Kcat/Km (1/mM·min)	Ref. No.
E. coli	22-25	10	6	300	30	179
B. stearothermophilus	60	100	44.5	4900	49.0	209
T. thermosulfurigenes	65	20	15.7	3140	157.0	127
T. saccharolyticum	65	16	17.6	3520	220.0	127
T. maritima	90	74	68.4	8550	73.5	

Table 4. Kinetic constants for type II xylose isomerases purified from mesophiles and thermophiles[a]

Organism	Temperature (°C)	Km for xylose (mM)	Vmax (µmol/min/mg)	Kcat (1/min)	Kcat/Km (1/mM•min)	Ref. No.
E. coli	22-25	10	6	300	30	129
B. stearothermophilus	65	100	44.5	4900	49.0	209
T. thermosulfurigenes	65	20	15.7	3140	157.0	127
T. saccharolyticum	65	16	17.6	3520	220.0	127
T. maritima	90	74	68.4	8550	72.5	191
T. neapolitana	90	16	52.2	2654	166.8	190

[a] Activities for conversion of xylose to xylulose are reported

Table 5. Kinetic constants for glutamate dehydrogenases purified from mesophiles and thermophiles

Organism	Temperature tested (°C)	Km (mM)					Specific activity (µmol/min/mg)	Ref. No.
		L-glutamate	2-Oxoglutarate	NH4+	NADPH	NADP+		
E. coli	25	1.3	0.64	1.1	0.04	0.042	217	210
S. typhimurium	25	50	4.0	0.29	0.019	0.013	340	211
N. europaea	23-25	6.7	4.3	16	0.049	b0.0079	1030	212
T. novellus	30	35.5	7.4	7.5	0.077	0.08	nd	213
[a]S. solfataricus	60	2.5	1.4	4.2	0.01	0.05	35	137
[a]Thermococcus NA1	80	9.12	1.7	15.5	0.066	0.038	nd	214
	60	0.6	0.33	6-27	0.012	0.02	nd	139
[a]P. furiosus	85	1.6	4.2	9.5	nd	0.18	130	138

[a] These glutamate dehydrogenases utilize both NADPH and NADH as cofactors, but utilize preferentially NADPH.
b The reverse reaction (NADP+ reduction) is inhibited by NADPH and NH4+.

E

the K_m value, con

for the temperatur

Kr

Symbols A_1, A_{-1},

activation energies

respectively. While

predicted without a

with temperature. I

temperatures for bo

dehydrogenase (Ta

analogs to mesophi

extrapolation of an

relationship yields

substrate affinities

similar substrate af

enzymes at low ten

energy or employ a

effect described by

catalytic rate enhan

substrate binding af

structural limitation

may account for the

$$E + S \underset{k_{-1}}{\overset{k_1}{\longleftrightarrow}} ES \overset{k_2}{\longrightarrow} E + P \qquad\qquad (1)$$

the K_m value, composed of 3 fundamental rate constants k_1, k_{-1}, and k_2, yields equation 2 for the temperature dependence of K_m.

$$Km = (A_{-1}A_2/A_1{}^2)e[(2Ea_1 - Ea_{-1} - Ea_2)/RT] \qquad\qquad (2)$$

Symbols A_1, A_{-1}, A_2, Ea_1, Ea_{-1}, and Ea_2 represent the preexponential factors (A) and activation energies (Ea) for the reactions corresponding to rate constants k_1, k_{-1}, and k_2, respectively. While the specific effect of temperature on enzyme K_m values cannot be predicted without a great deal of mechanistic information, these values are expected to vary with temperature. Reported K_m values were indeed significantly higher at higher temperatures for both *T. maritima* xylose isomerase [190] and *P. furiosis* glutamate dehydrogenase (Table 5). If thermophilic enzymes are exact structural and functional analogs to mesophilic enzymes excepting their greater resistance to denaturation, the extrapolation of an enzyme's K_m values to different temperatures using the Arrhenius relationship yields results inconsistent with the assumption of generally similar enzyme substrate affinities at mesophilic and thermophilic temperatures. The observations of similar substrate affinities for thermophilic enzymes at high temperature and mesophilic enzymes at low temperatures therefore, argue that thermophilic enzymes must expend more energy or employ a different mechanistic strategy to so efficiently bind substrate. The circe effect described by Jencks contends that excess energy from substrate binding is used for catalytic rate enhancement [216]. Increased partitioning of this total energy to maintain substrate binding affinity versus catalytic rate enhancement in thermophilic enzymes or their structural limitations reducing the total efficiency of substrate energy utilization therefore, may account for the activity and affinity differences between Arrhenius theory prediction

and observation.

increasing thermo

stability and reduc

Protein folding

Thermoph

enzymes fold prop

by similar mechan

enzyme stability ar

enzyme characteris

and involves the ra

Folding is believed

structure is stabilize

between parts of the

staggering number c

rapidly into their act

series of well-define

partially folded inter

structural conformer

structures seen in x-r

folding energetics su

steep-walled and dee

folding pathway.

Freire et al. [2

unfavorable energetic

folded/unfolded states

stabilizing forces. The

and observation. Reports of enzyme mutants that have consistently lower k_{cat} values with increasing thermostability supports a direct structure-function link between folded protein stability and reduced efficiency of substrate energy utilization [217,218].

Protein folding

Thermophilic and mesophilic proteins are typically similar and thermophilic enzymes fold properly at mesophilic temperatures arguing that both classes of proteins fold by similar mechanisms. The link between protein folded structure, enzyme activity, and enzyme stability argues that understanding protein folding may yield insights into these enzyme characteristics. Protein folding begins as the peptide is synthesized at the ribosome and involves the rapid condensation of particular regions, or nuclei, into native-like states. Folding is believed to be driven primarily by hydrophobic forces, and the native protein structure is stabilized by a variety of hydrophobic, covalent, and coulombic interactions between parts of the protein and between the protein and the solvent [219]. Despite the staggering number of theoretical structures they may occupy, the fact that proteins fold rapidly into their active conformation indicates that they reach their final structure through a series of well-defined intermediate states. Rapid, cooperative protein folding suggests that partially folded intermediates are not stable and that the native state includes very few structural conformers of similar energy. This conclusion is supported by the conserved structures seen in x-ray crystallography and NMR. Extending this conclusion to protein folding energetics suggests that the free-energy well containing the protein native state is steep-walled and deep, comprising most of the free energy range that describes the actual folding pathway.

Freire *et al.* [220] proposed that at least two factors are specifically related to the unfavorable energetics of exposing complementary surfaces to the solvent in the partially folded/unfolded states: (i) the driving energy in protein folding and (ii) the protein stabilizing forces. The authors define complementary surfaces as surfaces which are not

solvent-exposed i

unfolded states. T

partly folded/unfo

states. Their mod

aqueous solution o

solvent, and that th

protein characteris

individual protein

regardless of their

authors also noted

percent of the secon

and have a highly d

existence of these s

interactions cannot,

stability. Haynie an

maximizing the stab

predicts that, in the p

the free-energy varia

absence of denaturan

parameters of the unf

protein folding is *de*

thermodynamic prope

Furthermore, contrary

process relies on a sm

Examining pro

unfolded states furthe

been established that n

solvent-exposed in the native state, but which become solvent-exposed in the partly unfolded states. These result from regions of the folded core remaining condensed in partly folded/unfolded proteins that expose surfaces uniquely present in these intermediate states. Their model predicts that protein folding intermediates are highly unstable in aqueous solution due to the exposure of hydrophobic complementary surfaces to the solvent, and that these states are poorly populated. This CORE model relies on universal protein characteristics, not on structural motifs (α–helices and β–sheets) specific to individual protein folds, and is consistent with the observation that most proteins, regardless of their secondary structural composition, fold in a two-state process. These authors also noted reports that protein condensed phases which retain a "...significant percent of the secondary structure content of the native state, exhibit considerable flexibility and have a highly disrupted tertiary structure" are stable under certain conditions. The existence of these stable condensed cores indicates that secondary structural element interactions cannot, by themselves, control the protein folding pathway and folded protein stability. Haynie and Freire proposed a thermodynamic model to determine the conditions maximizing the stability of folded protein intermediate states [221,222]. Their model predicts that, in the presence of a denaturant, intermediate-state stability is independent of the free-energy variation between the intermediate and the native state, and that, in the absence of denaturant, the intermediate state ΔH is derived from the thermodynamic parameters of the unfolded state alone. This model implies that, thermodynamically, protein folding is *de facto* a one-way process. The successive steps depend only on their thermodynamic properties relative to the previous state and not to the subsequent one. Furthermore, contrary to the current belief, they predict that the two-state protein folding process relies on a small entropy contribution to the intermediate state stability.

Examining protein folding as a function of the total energies of the folded and unfolded states further complicates this scenario [223]. At 25°C and high dilution it has been established that non-polar solute transfer into water was opposed mainly by entropy

and not enthalpy

driven. However

transfer to water

temperature-inde

highest between

transfer is propos

loosely approxim

small, non-polar n

temperature. Abo

significant. There

entropy and enthal

approximately equ

experimental resul

of partitioning entr

expressed in meso

entropy and enthal

contain local, cond

so the free energies

corresponding regi

Protein unfoldi

Protein unf

unfolding a direct r

by thermal energy

Extensive irreversi

the removal of the

denaturation (when

and not enthalpy. Under these conditions, protein folding can be considered to be entropy driven. However, because of the large, positive heat-capacity change of non-polar-solute transfer to water [223,224], the ΔH and ΔS for this process cannot be considered temperature-independent. The ΔG of non-polar-solute transfer into water is predicted to be highest between 130°C and 160°C, where $T\Delta S \approx 0$ [223,224]. At these temperatures the transfer is proposed to be completely enthalpy driven. Considering that the cell cytoplasm loosely approximates a dilute water solution, and that a protein loosely approximates a small, non-polar molecule, protein folding can be considered entropy driven only at room temperature. Above this temperature, both entropy and enthalpy contributions become significant. Therefore, at temperatures where many proteins are folded (37–100°C), entropy and enthalpy contributions to the folding free-energy are expected to be approximately equal [224,225]. These predicted energy contributions are consistent with experimental results [225], and may help explain why, despite the temperature dependence of partitioning entropy and enthalpy, thermophilic proteins are correctly folded when expressed in mesophilic hosts. Protein folding is therefore, a robust process resistant to entropy and enthalpy variations. Unfortunately, folding intermediates do not necessarily contain local, condensed conformations identical to their counterparts in the native enzyme, so the free energies of structural intermediates cannot predict the partial free energy of the corresponding region in the native protein [226].

Protein unfolding

Protein unfolding is fundamental to protein stability with the Gibbs energy of unfolding a direct measure of the folded protein stabilizing energy. Protein destabilization by thermal energy and chemical denaturing agents has been well documented [227]. Extensive irreversible denaturation (i.e., loss of active architecture that is not recovered by the removal of the denaturing force) is far more common than extensive reversible denaturation (where active structure is regained upon removal of the denaturing force).

Since the system :

reversible unfold:

irreversible prote:

image of protein f

folding intermedia

in protein thermal

folding. For syster

pathway must be t

unfolding must exa

process at equilibri

processes does not

completely unfolde

for both folding and

unfolding experime

folded enzyme stabi

mechanisms may be

asymmetry.

Protein dena

Enzyme activity usu

temperature of max:

enzyme's active str

at temperatures abov

activity upon coolin

incomplete unfoldin

observation that den

that denatured prote:

is the loss of native :

Since the system is usually chemically or thermally altered to initiate protein unfolding, reversible unfolding is not necessarily a mirror image of protein folding. The existence of irreversible protein denaturation also indicates that irreversible unfolding is not a mirror image of protein folding. Thus, unfolding intermediates do not necessarily represent folding intermediates. For example, the molecular aggregation and precipitation common in protein thermal denaturation did not occur for these protein molecules during proper folding. For systems at equillibrium, the flux through each specific fundamental reaction pathway must be the same in both directions. Thus, the mechanism of reversible protein unfolding must exactly mirror the folding mechanism under the same conditions. Not a process at equilibrium however, this microreversibility principle describing equilibrium processes does not apply to irreversible protein denaturation. The ΔG between the completely unfolded and native states in the same environment, however, must be identical for both folding and unfolding. Thus, studies of protein folding energetics using data from unfolding experiments can yield direct insights into the molecular determinants controlling folded enzyme stability. But, extrapolations from these data to the protein folding mechanisms may be misleading due to the potential for folding versus unfolding asymmetry.

Protein denaturation is directly related to thermophilicity and thermostability. Enzyme activity usually increases with temperature until it falls precipitously above the temperature of maximal activity. This rapid loss of activity is consistent with the loss of an enzyme's active structure by denaturation. Many enzymes, however, have long half-lives at temperatures above their highest active temperature. For these enzymes, the recovered activity upon cooling into their active temperature range has been attributed to reversible, incomplete unfolding [228,229]. Denaturation is usually rapid and cooperative but the observation that denatured proteins in solution retain some condensed structure indicates that denatured protein is not necessarily complete unfolding. Mechanistically, denaturation is the loss of native 3-dimensional structure sustained by solvent molecules invading the

unfolding protein

protein solvent in

Molecular mec

Intrinsic Facto

The reten

mesophilic hosts

peptide. These in

entropy of unfold

others. The char

functionally simi

comparisons to d

denaturation. No

enzymes and mes

believed to be du

enzymes. An exp

conclusion that th

acids and decreas

Ser, Lys, and As

stabilization by n

Not all amino aci

interactions and r

structural stability

over 10°C [232],

from multiple am

acid substitutions

unfolding protein. This hydration reaction (in aqueous solutions) must originate at the protein solvent interface, its surface.

Molecular mechanisms of Thermostability
Intrinsic Factors

The retention of thermophilic properties by enzymes recombinantly expressed in mesophilic hosts argues that these protein characteristics result in part from the nature of the peptide. These intrinsic factors include sequence specific amino acid replacement, altered entropy of unfolding, hydrophobic core packing, and loop region architecture, among others. The characterization of thermophilic and thermostable enzymes, structurally and functionally similar to well characterized mesophilic enzymes, allows the use of comparisons to determine the factors involved in stabilizing enzyme architecture against denaturation. No universal mechanism explains the differences between thermophilic enzymes and mesophilic enzymes. Thermostability and thermophilicity properties are believed to be due to subtle changes throughout the amino acid sequences of thermophilic enzymes. An extensive comparative amino acid analysis by Argos *et al.* led to the conclusion that thermal stability was related to (i) increased internal hydrophobic amino acids and decreased external hydrophobic amino acids, (ii) the replacement of Gly, Ser, Ser, Lys, and Asp by Ala, Ala, Thr, Arg, and Glu, respectively, and (iii) to helix stabilization by more exclusive use of amino acids commonly found in helices [230]. Not all amino acid substitutions alter the function or stability of a protein. Specific interactions and residues, rather than all amino acids, contribute significantly to protein structural stability [231]. While single substitutions can increase the stability of an enzyme over 10°C [232], the thermostability intrinsic to thermophilic enzymes most often results from multiple amino acid substitutions [233-235]. The effects of single and multiple amino acid substitutions on enzyme thermostability are shown in Table 6. The examples were

Table 6. Representative site directed mutagenesis studies of enzyme thermostability mechanisms

protein (Reference)	mutation	effect	Conclusions
α₁-antitrypsin [236]	F51C	3to42 min stab. at 57°C	Low volume highly flexible hydrophobic sidechains allow better regional packing and prevent aggregation
	L	3to40 min stab. at 57°C	
	I	3to34 min stab. at 57°C	
	V	3to33 min stab. at 57°C	
	A	3to6.9 min stab. at 57°C	
Methanothermus fervidus	*Site directed mutants of M. fervidus*		

Table 6. Repersesentative site directed mutagenesis studies of enzyme thermostability mechanisms

protein (Reference)	mutation	effect	Conclusions
α1-antitrypsin [236]	F51C L I V A	3to42 min stab. at 57°C 3to40 min stab. at 57°C 3to34 min stab. at 57°C 3to33 min stab. at 57°C 3to6.9 min stab. at 57°C	Low volume highly flexible hydrophobic sidechains allow better regional packing and prevent aggregation
Methanothermus fervidus (Mf) and Methanobacterium bryantii (Mb) glyceraldehyde-3-phosphate dehydrogenases [237]	Site directed mutants of M. fervidus Y323S Y323W Chimeric mutants between Mf and Mb parental Mf Mf(1-294)/Mb(295-336) Mf(1-242)/Mb(243-336) Mf(1-176)/Mb(177-336) Mf(1-176)/Mb(177-294)/Mf(295-336) parental Mb	^{a}ΔT50 (°C) = -4.5 1.3 0.0 -9.0 -10.0 -10.5 -4.6 -12.3	Increased hydrophobicity of residues involved in interdomain contact increased thermostability
Mesophilic (M-wt) and thermophilic (T-wt) kanamycin nucleotidyl transferases [238]	M-wt T-wt M-D80Y T-Q102K	^{b}T_k (°C) solution / immobilized 51.0 / 58.2 55.3 / 62.9 59.4 / 66.5 57.2 / 60.9	1) Immobilization stabilized the enzymes 2) Stabilizing mutations in solution do not necessarily stabilize the immobilized enzyme
Mesophilic (M-wt) and thermophilic (T-wt) kanamycin nucleotidyl transferases [239]	L252 K130/L252 K130 Y80/L252 Y80 Y80/K130/L252 Y80/K130	^{c}ΔTopt (°C) / ^{d}ΔT$_k$ (°C) n.d. / -7.2 / -1.7 5.0 / 3.4 / 3.8 8.0 / 7.4 / 7.3 10.0 / 10.8	1) Multiple mutations can have a greater effect on thermostability that single mutations 2) Increasing enzyme thermophilicity by mutation reduced specific activity

Chicken egg-white lysozyme (240)	Position 40-55-91	ΔTm (°C)	ΣΔΔGu (kcal mol⁻¹)	1) Thermostability is directly related to side chain volume and hydrophobicity for conservative substitutions of buried core residues.
	TLT	3.3	1.25	
	TIT	2.6	0.99	
	SIT	1.6	0.61	
	TIS (wt)	---	---	
	TVT	0.0	0.0	
	TIV	0.0	0.0	
	TIA	0.2	0.08	

Chicken egg-white lysozyme [240]	Position 40-55-91	ΔT$_m$ (°C)	ΔΔG$_u$ (kcal mol^{-1})	1) Thermostability is directly related to side chain volume and hydrophobicity for conservative substitutions of buried core residues.
	TLT	3.3	1.25	
	TIT	2.6	0.99	
	SIT	1.6	0.61	
	TIS (wt)	---	---	
	TVT	0.0	0.0	
	TIV	-0.2	-0.08	
	TIA	-0.4	-0.15	
	SIS	-0.7	-0.27	
	SVT	-0.8	-0.30	
	TLS	-1.2	-0.45	
	SIV	-1.4	-0.53	
	SIA	-2.3	-0.87	
	TVS	-2.4	-0.91	
	TVA	-3.3	-1.25	
	TVV	-3.3	-1.25	
	SVS	-3.5	-1.33	
	SVA	-4.4	-1.67	
	IIS	-5.8	-2.20	
	TMS	-6.0	-2.27	
	TID	-6.1	-2.31	
	TFS	-6.5	-2.46	
	TAT	-7.3	-2.77	
	TIY	-8.1	-3.07	
	TAS	-11.6	-4.40	
	TTS	-13.1	-4.96	
Chicken egg-white lysozyme [241]	H15L	ΔTm (°C) = 2.0		1) The effects of thermostabilizing mutations can be independent or cooperative.
	R114H	1.9		
	A31V	3.0		
	H15L/A31V/R114H	6.0		
	I55L	-1.2		
	S91T	2.6		
	I55L/S91T	3.2		
	D101S	2.1		
	I55L/S91T/D101S	4.8		
	H15L/A31V/I55L/S91T/D101S/R114H	10.5		

Human lysozyme I24?L Human lysozyme (2.33.24.1)	C77AA/95A	ΔTm (°C) = -14.4	Stability is reduced by the loss of a disulfide bond
		ΔΔG ... ΔTm ...	Stability is enhanced by addition of a ...

		ΔTm (°C) / ΔΔG		
Human lysozyme [242]	C77A/C95A	ΔTm (°C) = -14.4		Stability is reduced by the loss of a disulfide bond
Human lysozyme [233,243]		fΔTd (°C)	gΔΔGu (kcal mol⁻¹)	Stability is enhanced by addition of a less flexible residue (Pro) and reduced by addition of a more flexible residue (Gly)
	P71G/P103G	-4.5	-1.1	
	P71G	-4.7	-1.2	
	P103G	-0.1	0.0	
	A47P	0.3	1.1	
	D91P	-1.1	-0.8	
	V110P	1.2	2.4	
T4 lysozyme [232]	E128A	ΔTm (°C) = 0.6 ± 0.25		1) α-helix stabilization by Ala
	V131A	1.0 ± 0.25		
	L133A	-17.0 ± 2.0		2) Site mutation effects are independent and additive
	D127A/E128A	0.9 ± 0.25		
	E128A/V131A	1.5 ± 0.25		
	V131A/N132A	2.3 ± 0.25		
	E128A/V131A/N132A	3.4 ± 0.22		
	D127A/E128A/V131A/N132A	4.0 ± 0.25		
	E128A/V131A/N132A/L133A	-10.3 ± 0.5		
	D127A/E128A/V131A/N132A/L133A	-9.4 ± 0.5		
T4 lysozyme [244]		ΔTm (°C) = pH 2.0	pH6.5	1) Better H-bond and charge interactions stabilize helix capping
	T59N	-2.1	-2.8	
	S	-2.6	-0.4	
	D	-3.1	-3.1	2) Better local core packing increases stability
	G	-7.7	-4.1	
	V	-10.0	-4.0	
	A	-10.1	-4.0	
T4 lysozyme [236]	L46A	ΔTm (°C) = -8.6		Cavity creating mutants destabilize proteins
	L99A	-15.7		
	L118A	-12.2		
	L121A	-9.3		
	L133A	-8.9		
	F153A	-12.3		
	L99A/F153A	-41.8		

| Tolymerzyme (245) | T-PON A-OTF T-151S | ΔTm (°C) = 1.35 0.13 0.93 | 1) Local packing and release of strain increase stability |
| | | | 2) Addition of a peptide linked to buried |

Protein	Mutation	ΔTm (°C)	Notes
T4 lysozyme [245]	T26S A93T T151S	ΔTm (°C) = 1.35 0.13 0.93	1) Local packing and release of strain increase stability 2) Addition of a peptide ligand to bound water increases stability
T4 lysozyme [246]	C54T/C97A (disulf. = 0) I3C/C54T (disulf. = 1) I9C/C54T/C97A/L164C (disulf. = 1) T21C/C54T/C97A/T142C (disulf. = 1) I3C/I9C/C54T/L164C (disulf. = 2) I9C/T21C/C54T/C97A/T142C/L164C (disulf. = 2) I3C/I9C/T21C/C54T/T142C/L164C (disulf. = 3)	ΔTm (°C) -DTT / +DTT n.d. / 0.0 4.8 / -1.9 6.4 / -6.5 11.0 / -2.7 15.7 / -10.3 17.0 / -5.9 23.4 / -8.5	Addition of disulfide bonds enhanced moderate temperature protein stability and more bonds made the protein more stable
T4 lysozyme [247]	L99A L99V L99I L99M L99F F153A F153V F153I F153L F153M L99A/F153A	ΔTm (°C) = -11.4 -5.2 -3.7 -1.5 -0.7 -9.3 -4.5 -0.5 0.8 -1.6 -22.8	Stability is enhanced by: 1) Increased hydrophobicity of the amino acid 2) Increased core packing efficiency
T4 lysozyme [248]	G77A A82P	ΔTm (°C) pH 2.0 / pH 6.5 -1.4 / 0.9 0.8 / 2.1	Stability is enhanced by addition of a less flexible residue (Pro) and reduced by addition of a more flexible residue (Gly)
T4 lysozyme [249]	Q105A Q105E Q105G	ΔTm (°C) pH 2.1 / pH 5.8 -3.5 / -1.6 -2.1 / -3.0 -7.2 / -3.9	1) Altered core packing alters thermostability 2) The (-) of Glu at high pH destabilized the mutant

T4 lysozyme
[2.34]

	ΔTm (°C) =
M102L	-2.54
V111F	-4.77
V111I	-2.32
L99F/M102L	-2.34

1) Effect of multiple mutations is less than the added effects of individual mutations (cooperativity)

		ΔTm (°C)	Comments
T4 lysozyme [234]	M102L	ΔTm (°C) = -2.54	1) Effect of multiple mutations is less than the added effects of individual mutations (cooperative)
	V111F	-4.77	
	V111I	-2.32	
	L99F/M102L	-2.34	
	L99F/V111I	-2.70	
	L99F/F153L	0.09	2) Compensatory local core packing accounts for the cooperativity
	M102L/V111I	-5.49	
	V111I/F153L	-3.52	
	L99F/M102L/V111I	-4.02	
	L99F/M102L/F153L	-0.68	
	L99F/V111I/F153L	-1.73	
	L99F/M102L/V111I/F153L	-1.82	
T4 lysozyme [250]	Q86D/A92D	ΔTopt = 17°C	Added Ca^{2+} binding site increases maximal active temperature
T4 lysozyme [251]	A41S	ΔTm (°C) = -1.77	Burying a polar group in the protein core is more destabilizing than a solvent exposed substitution so hydrophobicity is important to maintaining folded protein stability
	A42S	-7.49	
	A49S	-1.53	
	A73S	-1.27	
	A82S	-0.99	
	A93S	-0.52	
	A98S	-7.47	
	A130S	-2.89	
	A134S	-0.44	
	V75T	-3.70	
	V87T	-4.55	
	V149T	-10.08	
T4 lysozyme [252]		ΔTm (°C)	Proteins can accomodate structural perturbations by local core repacking making them robust to random mutation
		pH 3.0 pH 5.4	
	S117F	4.8 2.8	
T4 lysozyme [253]	L99A	ΔTm (°C) = -15.3	1) Internal cavity creating, destabilizing mutants can be stabilized by binding of a small hydrophobic molecule so core packing is important to protein stability
	L99A + benzene (7.5mM)	-9.3	
	F153A	-12.0	
	F153A + benzene	-12.1	
	L99A/F153A	-9.8	
	L99A/F153A + benzene	-6.8	

| B. stearothermophilus neutral protease (NP-ste) [254] | A60P T6.1P S65P | T_{AT50} (°C) = 5.6 -10.0 4.3 | Addition of Pro to restrict local flexibility can dramatically alter enzyme |

		ΔT50 (°C)	iΔt1/2 (min)	
B. stearothermophilus neutral protease (NP-site) [254]	A69P	bΔT50 (°C) = 5.6		Addition of Pro to restrict local flexibility can dramatically alter enzyme stability but depending on added conformational strain this can be negative or positive
	T63P	-10.0		
	S65P	4.2		
	Y66P	-16.0		
	A69P	5.6		
B. stearothermophilus neutral protease (NP-site) [255]	A166S	ΔT50 (°C) = 1.2		Release of a buried water molecule and enhanced H-bonding increased folded protein stability
B. stearothermophilus neutral protease (NP-site) [256]	T63R	ΔT50 (°C) = 7.1		Increased hydrophobic contacts at the protein surface stabilize the protein
	T63K	6.7		
	T63F	6.2		
	T63I	4.1		
	T63Y	3.6		
	T63M	1.5		
	T63V	-0.9		
	T63G	-1.4		
	T63S	-4.0		
	T63A	-7.5		
	T63D	-7.5		
	T63P	-10.0		
B. cereus oligo-1,6-glucosidase [235]	K121P	1.4	j19	1) Added prolines on the surface of a protein increase thermostability
	K121P/E175P	1.8	j95	
	K121P/E175P/E290P	2.6	j>95,k2.3	
	K121P/E175P/E290P/E208P	3.4	k7.3	2) 2nd sites in β-turns and N-caps of α-helices provide the largest effect
	K121P/E175P/E290P/E208P/E270P	3.1	k4.3	
	K121P/E175P/E290P/E208P/E270P/E378P	3.1	k5.2	
	K121P/E175P/E290P/E208P/E270P/E378P/T261P	3.6	k12.5	3) The effects of Pro residues are additive and not cooperative
	K121P/E175P/E290P/E208P/E270P/E378P/T261P/E216P	3.7	k11.5	
	K121P/E175P/E290P/E208P/E270P/E378P/T261P/E216P /N109P	5.1	k38.5	

| F GM RNase III | ORO-WR1 →ORO C(R1 WR2 | ΔTm (°C) = 1.2 | 1) GM stabilizes C-terminus of an α- |
| (257) | ORO-WR1 →ORO-AR1-WR2 | -1.5 | helix (turn→helix) |

Source	Mutations	ΔTm (°C)		Comments
E. coli RNase H1 [257]	Q80-W81→Q80-G81-W82 Q80-W81→Q80-A81-W82 Q80-W81→Q80-G81-W82/G77A G77A	ΔTm (°C) = 1.2 -1.5 2.7 -2.9		1) G81 stabilizes C-terminus of an α-helix (paperclip) 2) A77 in an α-helix adds bond strain and alters local packing to destabilize 3) G81 alters local packing around A77 and their dual stabilization is cooperative
E. coli RNase H1 [258]		**ΔTm (°C)** pH3.0	pH5.5	
	R1 = ML → MNPSPR	-2.0	-1.7	1) Mutagenic effects may be independent, or show (+) or (-) cooperativity.
	R2 = YRGR → FHAH	-2.3	-3.2	
	R3 = YTR → EAC			
	R4 = H62P	3.4	4.1	2) Pro substitution even based on sequence comparison is not always stabilizing
	R5 = VRQGITO → LKKAFEG	2.7	4.5	
	R6 = KTADK → REG	4.0	5.6	
	R7 = LGQHQIKWEW → MAPRVRFHF	-5.7	2.4	3) Pro insertion in loop regions was stabilizing in both cases
	R8 = A125T	0.4	0.0	
	R9 = TGYQVEV → CPPRAPTLFHEEA	-18.9	-13.8	
	G80b = WQ → QGW	1.2	0.8	4) Thermostabilization of mesophilic proteins by region replacement based on sequence comparisons between functionally similar thermophilic and mesophilic proteins may not yield the expected effect
	p113 = Q113P	-0.6	-2.1	
	R7E119 = LGQHQIKWEW → MAPRVRFEF	-1.1	1.9	
	R1/R2	-4.5	-4.9	
	R1/R4	0.4	1.9	
	R1/R2/R4	-2.1	-0.1	
	R4/R7	-2.0	6.9	
	R5/R7	-1.9	8.2	
	R5/R6	6.0	9.2	
	R4/R6	6.7	8.8	
	R4/R5/R7	2.0	12.5	
	R5/R6/R7	1.9	12.4	
	R4/R5/R6/R7	6.2	16.7	
E. coli RNase H1 [257]	V74L V74I V74A	**ΔTm (°C)** 2.1 to 3.7 2.1 to 3.7 -7.6 (pH 3.0) and -12.7 (pH 5.5)		1) Cavity filling mutants slightly stabilize the folded protein 2) A cavity creating mutant significantly destabilized the protein

| E. coli RNase III (160,259) | pH Atm (°C) pH 8.5 pH 3.0 | replacement of β-turn residues to... |

Organism [ref]	Mutation	ΔTm (°C)		Description
E. coli RNase H1 [160,259]		pH 3.0	pH 5.5	Replacement of β-turn residues in left handed conformations with Gly or Asn stabilizes proteins
	K95G	5.7	6.8	
	K95N	2.9	3.2	
E. coli RNase H1 [257]		pH 3.0	pH 5.5	Reduced internal cavities stabilize folded proteins
	V74L	3.7	3.3	
	V74I	2.4	2.1	
	V74A	-7.6	-12.7	
Aspergillus oryzae RNase T1 [260]		β-band	Tyr-band	1) Altered local core packing destabilizes mutants
	Y45W	-1.6	-1.9	
	W59Y	-4.2	-4.5	2) Loss of H-bond to bound water in the W59Y mutant further destabilizes it
	Y45W/W59Y	-4.5	-4.7	
B. amyloliquefaciens subtilisin BPN' [261]		Δt1/2 (min) at 61°C		1) Disulfide bond stabilization of folded proteins does not correlate to their resistance to reduction
		-DTT	+DTT	
	V26C/A232C	4	8	
	A29C/M119C	-42	-34	
	D36C/P210C	3	-81	2) Disulfide bonds may not stabilize irreversibly unfolded proteins because of kinetic, not thermodynamic factors
	V148C/N243C	n.d.	-101	
	D41C/G80C	-113	-115	
B. amyloliquefaciens subtilisin BPN' [262]		Δt1/2 (min) at 58-59°C		1) disulfide engineering did not significantly stabilize the protein relative to the wild type
		-DTT	+DTT	
	T22C	-98	n.d.	
	S24C	0	n.d.	
	S87C	-16	n.d.	2) the disulfide disrupted an intrapeptide H-bond causing as much destabilization as stabilization
	T22C/S87C	-51	-60	
	S24C/S87C	-3	-23	
subtilisin E [263]	G61C/S98C	ΔTm (°C) = 4.5 Δt1/2 at 55°C = 50 min		Addition of a disulfide bond enhanced moderate temperature protein stability

43

		Δt1/2 (h)		
Actinoplanes missouriensis xylose isomerase [264]		sol (85°C)	immob (70°C)	1) Effect of mutations on soluble vs. immobilized enzymes may differ
	K253R	0.8	176	
	K294R	-1.4	-14	2) Glycation of Lys by glucose (substrate) is major pathway of protein destabilization
	K253Q	n.s.	-39	
	K294Q	n.s.	n.d.	
	A73S	-0.1	-15	
	G70S/G74T	0.6	-19	
	G70S/A73S/G74T	1.1	17	
		Δt1/2 (h)		
A. missouriensis xylose isomerase [265]		+gluc.(60°C)	-gluc.(84°C)	1) Arg for Lys substitutions prevent destabilizing glycation
	K253R	160	1.0	
	K309R	20	1.8	
	K319R	10	1.8	2) Arg residues add stabilizing H-bonds
	K323R	0	0.0	
	K309R/K319R/K323R	50	0.4	
Thermoanaerobacterium thermosulfurigenes xylose isomerase [266]	W49R	Δt1/2 (min) = 20		Reduced water accessible surface area increases protein thermostability
	W139F	32		
	W139M	23		
	W139A	15		
	F60H	-10		

n.d. = not determined

[a]ΔT50 (°C) = the change in temperature where 50% of the activity remained after 10 min incubation

[b]T$_k$ is defined as the temperature at which the inactivation rate constant = 0.5

[c]ΔTopt = the change in the temperature for maximal activity

[d]T$_k$ is defined as the temperature at which the inactivation rate constant = 0.1

[e]ΔTm = the change in melting temperature

[f]ΔT$_d$ = the change in denaturation temperature

[g]ΔΔG$_u$ = change in free energy of unfolding

[h]ΔT50 = the change in temperature where 50% of the activity remained after 30 min incubation at pH 5.2

[i]Δt$_{1/2}$ = the time at which half of the activity remained at the temperature indicated

[j]determined at 45°C

[k]determined at 48°C

chos

thor

exch

Thes

mech

simil

meso

contr

therr

throu

local

have

mutar

const

mutat

an α-

they c

of the

cooper

has co

engine

engine

energy

substit

specific

chosen from site-directed mutagenesis experiments in which the mutational effects were thoroughly studied by biophysical structural analyses (e.g., crystallography, hydrogen exchange measurements, Fourier transformed infra-red spectroscopy [FTIR], calorimetry). These examples are also representative of the known mutational effects. Typically, several mechanisms are used to stabilize a single protein. Finally, because of the structural similarities between mesophilic and thermophilic proteins, reviewing selected literature on mesophilic protein thermostabilization may aid in understand the the general mechanisms controlling thermophilic-protein folded stability.

Multiple substitutions and protein stabilization - A great deal of protein thermostability research has focused on enhancing mesophilic enzyme thermostability through mutagenesis. Because mutataional effects on protein stability are proposed to be local structural phenomena, they can be independent (multiple amino acid substitutions have the same effect as the sum of the individual substitutions) or cooperative (the multiple mutant stability is either greater than or less than the sum of the single mutations) when constructive or destructive interferences occurs between the spheres of individual mutational effects [234]. When Matthews and collaborators introduced multiple alanines in an α-helix of T4 lysozyme, the mutation effects were additive (Table 6). However, when they constructed multiple mutations affecting a single internal cavity (where the side-chains of the target residues interacted with each other), the effects of the multiple mutations were cooperative and less predictable (Table 6). Further work on chicken egg-white lysozyme has corroborated this range of mutational effects [240,241]. Therefore, from an engineering standpoint, a continuum of thermostabilities should be attainable by protein engineering, bounded only by the thermodynamic limits of the protein's total molecular energy. However, the difficulty in predicting the stabilizing effects of multiple substitutions prevents the use of amino acid thermostabilizing "traffic rules" to identify specific amino acids responsible for one enzymes thermostability relative to another.

Si

ha

na

co

rej

res

the

ten

doc

app

dis

pro

con

tha

eith

unfo

stud

poor

redu

ther

were

rathe

for th

betwe

indica

Math

Substitutions and modification of the thermodynamics of unfolding -
Since entropy has traditionally been considered the main factor driving protein stability, it has been proposed that reducing the entropy gained by protein unfolding stabilizes the native structure [248]. Reducing the entropy gain upon protein unfolding by adding a constraint not removed upon denaturation may be achieved by disulfide bond formation, replacing Gly residues with any other amino acid, and replacing any amino acid with Pro residues. While disulfide bond engineering has been shown to affect mesophilic enzyme thermostability [242,246,261-263], the chemical instability of Cys residues at thermophilic temperatures makes this strategy ineffective for high temperature applications [267], and it does not appear as an evolutionary strategy for thermophilic enzymes. Moreover, enthalpy appears to be a significant factor (often dominating entropy) in protein stabilization by disulfide bonds even when these bonds remain intact in both the folded and unfolded protein forms [242]. Lacking a β-carbon, glycine residues in solution have more conformational flexibility and entropy than any other amino acid. Matthews *et al.* predicted that replacement of glycine by any other amino acid containing a β-carbon would stabilize either thermophilic or mesophilic protein folded structure by ~4 kJ mole^{-1} relative to the unfolded state [248]. Unfortunately, results of their subsequent exhaustive mutagenic study of T4 lysozyme were not completely consistent with the theory. While mutations of poorly mobile amino acids (as determined by low average crystallographic β values) with reduced solvent accessibility were far more likely to have a significant effect on protein thermostability, stability mutants primarily affecting the free energy of the unfolded state were rare [268]. Therefore, they concluded that folded protein local packing efficiency, rather than an entropy difference between the folded and unfolded states, was responsible for the observed thermostabilization. Furthermore, the lack of significant differences between the thermophilic and mesophilic Gly residue percent compositions (Table 3) indicates that this strategy is also not employed to stabilize thermophilic proteins. Matthews and colleagues, using similar arguments, predicted that the extra geometric
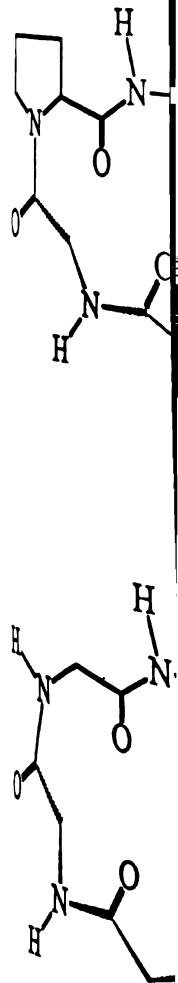
constraint of pro
stabilize the fold
configurations pr
been demonstrate
surrounding then
of a protein struc
proteins [233,24.
demonstrating tha
by nature. These
which can enhanc
appear to contain
counterparts. Also
entropic componer
[233,270]. The ex
smaller energy gair
seen.

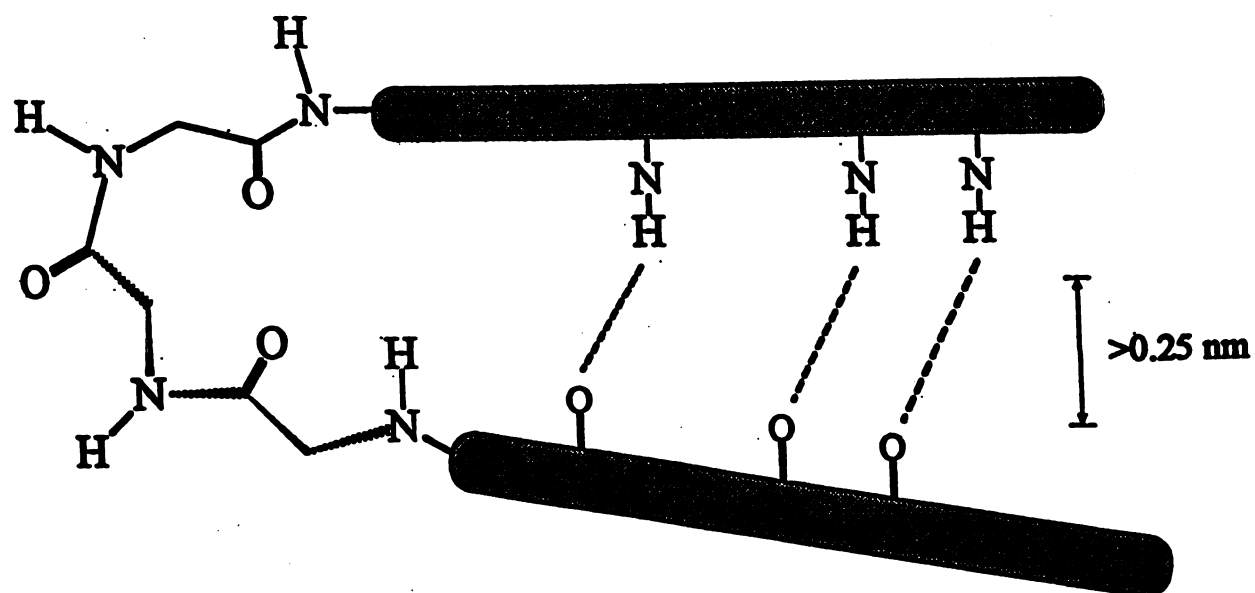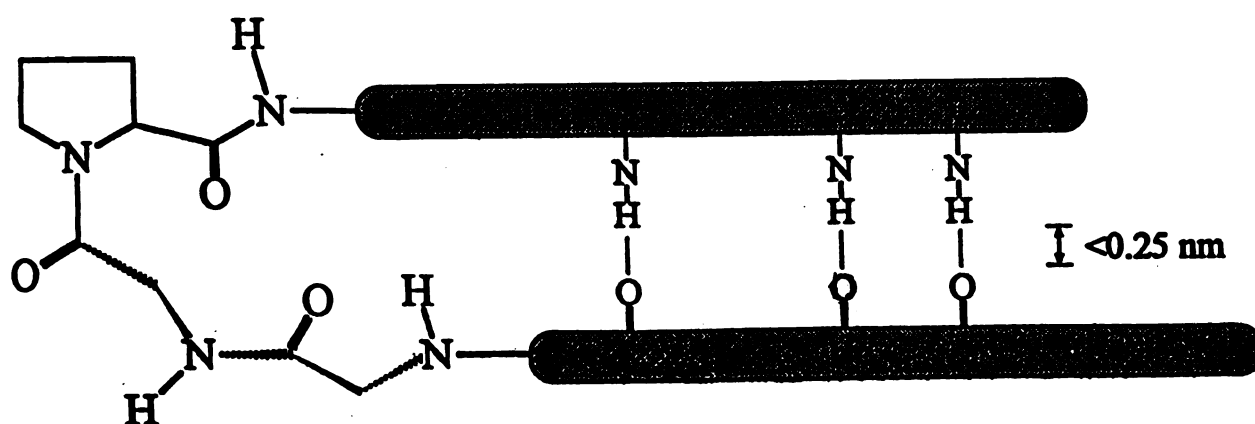**Prolines** ar
a zipper, prolines a
of numerous coulor
[71]. This proline-z
thermostability (Fig
insignificant, loops a
properly positioned
that loop regions can
elements, and that th
elements within an e
described here for the

constraint of proline residues present in both the folded and unfolded protein would stabilize the folded protein by ~4 kJ mole^{-1} [248]. The specific minimum-energy configurations prolines can occupy have been carefully calculated [269]. It has further been demonstrated that prolines play an important role in directing the local conformations surrounding them [269], and that proline insertions are only well tolerated in specific parts of a protein structure [269]. Proline substitutions have been shown to stabilize mesophilic proteins [233,243,248,258] and to further stabilize thermophilic proteins [254], demonstrating that the practical limits of enzyme stability are higher than those engineered by nature. These results suggest that careful proline substitution is a general technique which can enhance enzyme thermostability, and, in some cases, thermophilic proteins do appear to contain higher Pro residue percent compositions than their mesophilic counterparts. Also, as seen in disulfide stabilization of folded proteins, both enthalpic and entropic components were important to proline substitution thermostabilization effects [233,270]. The extent of proline steric constraints relative to other amino acids predicts a smaller energy gain from entropy due to proline introduction into a protein than is often seen.

Prolines and loop regions - It has been proposed that analogous to the stop on a zipper, prolines are used in constrained loop regions to prevent the sequential dissociation of numerous coulombic stabilizing interactions between the two adjacent core elements [71]. This proline-zipper model predicts a crucial role for loop regions in protein thermostability (Fig. 2), contradicting the traditional thinking that, typically structurally insignificant, loops act as flexible links maintaining a continuous peptide between two properly positioned core elements. This traditional view is supported by the observations that loop regions can withstand the accumulation of more neutral substitutions than can core elements, and that there is usually a greater sequence variability in loop regions than in core elements within an enzyme family [271]. However, the proline-zipper hypothesis described here for thermostabilization by loops follows from the mechanistic understanding

Figure 2. Role of proline residues in protein structural stabilization.

< 0.25 nm

> 0.25 nm

that the rapid coo

interface and tha

emerging data su

protein thermost

constrained loop

al. [259] constru

mesophilic/therm

that targeted a β-

dichroism measu

modeling, Kimur

created by the lys

core elements. T

hydrogen bond. T

introduced by pro

longer, constraine

or multiple Pro re

These prolines we

data predict that lo

structurally constr

number, will contr

regions appears to

represent a natural

protein engineering

Salt bridge

high, even compared

suggests that intram

in a highly polar sol

that the rapid cooperative protein unfolding process originates at the protein-solvent interface and that loops are usually found at the surface of protiens. Furthermore, emerging data suggests that manipulating turn structures may be a common method for protein thermostabilization in nature [259,272]. Studies focusing on β-turns suggest that constrained loop regions play a major role in protein resistance to denaturation. Kimura *et al.* [259] constructed *E. coli* RNaseH site-directed mutants, using the results of a previous mesophilic/thermophilic enzyme sequence comparison [258]. Mutations K95G and K95N that targeted a β-turn clearly stabilized the protein, as shown by activity assays and circular dichroism measurements. Based on the enzyme x-ray structure and on molecular modeling, Kimura *et al.* proposed that the K95G mutation eliminated some structural strain created by the lysine residue, and allowed a better interaction between the two neighboring core elements. The K95N mutation allowed the formation of an asparagine intra-residue hydrogen bond. This bond created a structural constraint analogous to the constraint introduced by prolines [259]. Additional prolines were observed in short loops and in longer, constrained loops (those whos flexibility is limited by factors such as salt bridges or multiple Pro residues) in two thermostable and one hyperthermostable xylose isomerase. These prolines were absent in the mesophilic enzymes (see Table 3). These unpublished data predict that loop stabilizing mutations are more effective when introduced into structurally constrained loops thus, the Pro residue positions in a peptide, as well as the number, will control their effect on folded protein stability. Proline introduction in turn regions appears to indicate a general importance for turn regions in protein stability and to represent a naturally occurring thermostability control mechanism that can be applied to protein engineering.

Salt bridges - Activities of ions immobilized on the same molecule are extremely high, even compared to those in bulk solution at molar concentrations [273]. This fact suggests that intramolecular salt bridges may be very stable, even at the surface of a protein in a highly polar solvent environment. Comparing the thermophilic *B. stearothermophilus*

and mesophilic y

extra salt bridges

thermostability [

phosphoglycerate

factors in these e

Actinoplanes mis

effects on *S. solf*

contributed signif

number of surface

in protein thermos

additional charged

thermostabilizing

in *B. stearotherm*

probability of the

irreversible denatu

that none of the ad

Hydrogen

folded protein struc

Subsequently, the r

thoroughly studied

addressed the impor

are formed between

significantly stronge

crystallographically

position shifts in NM

are strongly depende

H-bonding partners i

and mesophilic yeast phosphoglycerate kinase crystal structures led to the conclusion that extra salt bridges present in loops in the thermophilic protein contribute to its thermostability [274]. Crystal structure analysis of ribonuclease H [275] and 3-phosphoglycerate kinase [276] allowed the identification of salt-bridges as important factors in these enzymes's thermostabilities. An intra-peptide salt bridge also stabilizes the *Actinoplanes missouriensis* xylose isomerase [265]. A careful study of chaotropic salt effects on *S. solfataricus* carboxypeptidase similarly concluded that surface salt bridges contributed significantly to the enzymes resistance to thermal denaturation [175]. A number of surface residue to arginine substitutions also implicated electrostatic interactions in protein thermostability [265]. It is important to note that the presence of suitable additional charged residues in a peptide 1° structure is not sufficient to identify this thermostabilizing mechanism. Tomazic and Klibanov proposed that additional salt bridges in *B. stearothermophilus* α-amylase reduced reversible unfolding, thus reducing the probability of the partially unfolded enzyme forming scrambled structures and preventing irreversible denaturation [138]. However, subsequent crystallographic exidence indicated that none of the added Lys residues were involved in salt bridges [277].

Hydrogen bonds - The importance of hydrogen bonds (H-bonds) in maintaining folded protein structure was first recognized by Mirski and Pauling in 1936 [201]. Subsequently, the role of these coulombic interactions in protein stability has been thoroughly studied and reviewed [278]. A recent publication by Cleland and Kreevoy addressed the importance of low-barrier H-bonds in catalysis [279]. Low-barrier H-bonds are formed between residues carrying functional groups with similar pKa's, and are significantly stronger than ordinary double-well H-bonds. They can be detected crystallographically by either their characteristic short bond distances (< 2.5 Å) or by field position shifts in NMR [280]. It has long been known that the pKa's of ionizable groups are strongly dependent on environmental factors; in proteins, the environment surrounding H-bonding partners is often quite different from their environment in solution. These

observations, co[...]

thermophilic pro[...]

their mesophilic [...]

thermophilicity n[...]

thermophilic pro[...]

differences in stru[...]

stabilization crea[...]

the geometric cor[...]

Substantiating or [...]

structures.

Hydroph[...]

believed to provid[...]

B. caldovelox α-a[...]

forming residues t[...]

hydrophobic intera[...]

thermostability. M[...]

which contributed [...]

the engineered resid[...]

appeared clearly rel[...]

energy of unfolding[...]

energies for the subs[...]

in overall protein sta[...]

stabilizing factor at t[...]

hydrophobic content[...]

[247,251,253,255,2?[...]

in the core of T4 lyso[...]

residues at the surfac[...]

observations, combined with recent x-ray crystallography showing that the core of thermophilic proteins contains fewer and smaller cavities (i.e., higher atomic density) than their mesophilic counterparts [124], suggest that both enzyme thermostability and thermophilicity may be enhanced by the formation of stronger H-bonds. Thus, thermophilic proteins can substantially increase their stabilization energy with only small differences in structure or amino acid content. The excess enthalpic component of protein stabilization created by proline substitution in constrained loops may also be explained by the geometric constraint strengthening the H-bonds between adjacent core elements. Substantiating or refuting this hypothesis will require careful analysis of exact protein structures.

Hydrophobic interactions and core packing - Hydrophobic interactions are believed to provide the energy needed to fold proteins in aqueous solutions. Thermophilic *B. caldovelox* α-amylase [85] and ES4 GDH [135] contain more hydrophobic-interaction-forming residues than their mesophilic counterparts. It has been proposed, therefore, that hydrophobic interactions play a significant role in enzyme thermophilicity and thermostability. Matsumura *et al.* [281] created multiple substitutions of Ile3, a residue which contributed to the major hydrophobic core in T4 lysozyme. The hydrophobicity of the engineered residue (measured as the free energy of transfer from water to ethanol) appeared clearly related to protein stability (measured as the difference between the free energy of unfolding of the mutant and the wild-type protein). Water/ethanol transfer free energies for the substituted residues could account for approximately 80% of the difference in overall protein stability, indicating that hydrophobic interactions were the main stabilizing factor at this position. Other amino acid substitutions increasing the hydrophobic content of the protein core have increased protein thermostability [247,251,253,255,282]. Furthermore, the stability cost of burying a polar hydroxyl group in the core of T4 lysozyme was significant [251]. It is generally believed that hydrophobic residues at the surface of proteins are unfavorable for protein stability. However, the

protruding hydro

the protein surfac

stability. The *B.*

introducing bulky

[256].

Extensive

significant cavitie

core packing is su

substitutions [28

core environment

based mutagenesis

thermostability by

displacing buried H

Sequence analysis

was attributable to

enzyme's higher pe

have also shown a

efficiency in therm

thermostability [28

overall increase in l

enzymes since, in r

content exist betwee

indicates that increa

occuring proteins.

in some cases it can

strain [285,289,290]

still important in mai

protruding hydrophobic surfaces of some surface residues create a hydrophobic pocket in the protein surface, where the presence of a hydrophobic residue will favor protein stability. The *B. stearothermophilus* neutral protease was significantly stabilized either by introducing bulky hydrophobic residues or by Arg or Lys residues in such a surface pocket [256].

Extensive x-ray crystallographic studies have shown (i) that folded proteins contain significant cavities (some filled by water molecules, some not) [283], (ii) that protein local core packing is surprisingly able to compensate for deformations due to amino acid substitutions [284,285], and (iii) that aromatic-aromatic interactions in the hydrophobic core environment may stabilize the folded structure [234,252]. Other crystallography-based mutagenesis experiments have demonstrated the potential to increase protein thermostability by filling cavities in the folded structure [236,240,247,257,286], and by displacing buried H_2O molecules with either amino acid sidechains [255] or benzene [253]. Sequence analysis of *T. maritima* GAPDH was used to argue that its high thermostability was attributable to local packing interactions specifically involving the *T. maritima* enzyme's higher percentage of large hydrophobic groups [142]. Crystallographic studies have also shown a link between increased core hydrophobicity and better packing efficiency in thermostable proteins, supporting the importance of protein hydrophobicity in thermostability [287,288]. Some thermophilic enzymes are significantly stabilized by an overall increase in hydrophobicity but this trait is not universal among thermostable enzymes since, in many enzyme families, no significant differences in hydrophobic residue content exist between thermostable and thermolabile proteins (Table 3). This observation indicates that increased hydrophobicity is not the only stabilizing tool used in naturally occurring proteins. While better core packing is often linked to increased hydrophobicity, in some cases it can affect stability by more subtle means such as the reduction of bond strain [285,289,290]. Not necessarily stable in solution, secondary structural elements are still important in maintaining active enzyme architecture [224]. Based on the ability of high

densities of shor

increasing peptid

stabilization in c

Denser internal p

stabilizing force

Covalent

denaturation by te

or the oxidation [

protein thermosta

through catalytic

indirectly through

aggregation. For

hyperthermophilic

indicated that dear

after significant pr

the major mechani

egg-white lysozyn

violaceoniger and

rates of amino acid

(usually above 90°

significant cause of

Bacillus amylolique

unimolecular confor

licheniformis enzyn

An inverse relations

isomerases and their

95°C) [190]. The ge

densities of short peptides to form α-helices in crystals, it has been proposed that increasing peptide concentration increases the stability of helices [224], and that helix stabilization in core domains was implicated in protein thermostability [232,257,291]. Denser internal packing, therefore, may enhance protein thermostability either as a general stabilizing force or as a factor altering the stability of secondary structures.

Covalent destruction and irreversible denaturation - Enzyme irreversible denaturation by temperature-induced covalent modifications (i.e., Asn or Gln deamidation, or the oxidation [134,204,267] of Cys to cysteic acid [284,293,294]) directly affects protein thermostability. These covalent alterations possibly affect thermophilicity directly through catalytic residue decomposition (e.g., Cys residue in ADHs or GAPDHs) or indirectly through their role in irreversible denaturation due to protein degradation or aggregation. For example, peptide cleavage at the C-terminal end of arginine residues in a hyperthermophilic GAPDH limited its thermostability [151]. Klibanov and collaborators indicated that deamidation was probably involved in the formation of scrambled structures after significant protein unfolding occurred [295]. They also showed that deamidation was the major mechanism controlling thermoinactivation in *Bacillus* α-amylases [197] and hen egg-white lysozyme [138]. Sicard *et al.* reported a similar conclusion for *Streptomyces violaceoniger* and *S. olivochromogenes* glucose isomerase thermoinactivation [292]. The rates of amino acid covalent degradation become significant only at elevated temperatures (usually above 90°C), indicating that only under these conditions will such processes be a significant cause of enzyme inactivation [296]. Tomazic and Klibanov showed that *Bacillus amyloliquefaciens* and *B. stearothermophilus* α-amylases are inactivated by unimolecular conformational scrambling, whereas the more thermostable *Bacillus licheniformis* enzyme was inactivated by Asn/Gln deamidation [297].

An inverse relationship was seen between the percent of Asn plus Gln residues in D-xylose isomerases and their respective temperature for maximal activity (ranging from 55°C to 95°C) [190]. The general trend of fewer Cys, Asn, and Gln residues in the more

thermophilic pro

induced decomp

temperatures. In

isomerases [206]

50% unfolding) f

The mesophilic E

stearothermophi!:

and 85°C, respect

and 85°C, respect

85°C). The T. ne

higher than its ter

precipitation (95°(

appears to differ f

percentage of init

suspect that T. ned

another covalent d

unfolded [206]. T

require either the r

protein.

Extrinsic mecha

Extrinsic fa

protein stability (g!

pressure effects, etc

thermophilicity. It i

nature and concentr:

environmental facto

thermophilic proteins provides indirect support for the hypothesis that the thermally induced decomposition of these residues is undesirable for protein function at very high temperatures. In a comparative study of mesophilic and thermophilic type II D-xylose isomerases [206], we extrapolated the enzyme melting temperatures (i.e., temperature for 50% unfolding) from melting experiments performed in the presence of GuHCl (Table 7). The mesophilic *E.coli*, thermophilic *T. thermosulfurigenes*, and thermophilic *B. stearothermophilus* xylose isomerases were 50% unfolded at temperatures (53°C, 80°C, and 85°C, respectively) close to their respective maximal activity temperatures (55°C, 80°C and 85°C, respectively) and their precipitation initiation temperatures (52°C, 80°C, and 85°C). The *T. neapolitana* enzyme's melting temperature (120°C), however, was 25°C higher than its temperature for maximal activity (95°C) and for initiation of protein precipitation (95°C). Thus, the mechanism of *T. neapolitana* xylose isomerase inactivation appears to differ from that of the other three xylose isomerases, and does not involve a high percentage of initial unfolding. Further experimentation is required for verification, but we suspect that *T. neapolitana* xylose isomerase inactivation occurs through deamidation (or another covalent destruction mechanism) at temperatures where the enzyme is barely unfolded [206]. These data argue that the engineering of protein hyperthermostability will require either the removal or the modification of any thermolabile residues in the source protein.

Extrinsic mechanisms

Extrinsic factors, those due to the influence of non-protein effectors on folded protein stability (glycosylation, immobilization, stabilization by salts, metals, thermamines, pressure effects, etc.), have also been implicated in enzyme thermostability and thermophilicity. It is well known that *in vitro* enzyme activity is strongly affected by the nature and concentration of effector molecules present in solution as well as by environmental factors. Allosteric effectors, substrates, and ions (such as Ca^{2+} and PO_4^{3-})

Table 7. Comparison of xylose isomerase thermal parameters with the values obtained in

Table 7. Comparison of xylose isomerase thermal parameters with the optimal growth temperature of the respective microorganism

Microorganism	Organism	Xylose Isomerase			
	T_{opt}^a (°C)	T_{max} (°C)	half-life (min at 85°C)	T_m (°C)	T_{ppt} (°C)
mesophile					
E. coli	37	55	<1.0	52	53
thermophile					
T. thermosulfurigenes	60	80	45	80	80
B. stearothermophilus	65	85	ND	93	85
hyperthermophile					
T. neapolitana	80	95	330	120	90

ND: not determined

a Organism optimal growth temperatures (T_{opt}) are from the DSM catalog

b Specific activity was determined at T_{max} for each enzyme

are specifically bound by proteins. It is not surprising, therefore, that numerous enzyme thermostabilities are altered by such effectors [298]. Substrate molecules have long been known to stabilize enzymes, presumably by interaction in specific binding sites [250,266], and these effects are not reviewed here. Chemical cross-linking or immobilization, glycosylation, inorganic salts, and high pressure also stabilize enzymes. This discussion of these topics is confined to the current mechanistic theories and provides representative examples supporting or opposing these theories.

Glycosylation - An increasing number of bacterial extracellular enzymes have been shown to be glycosylated (for example, *C. thermocellum* cellulosome components and *T. saccharolyticum* endoxylanase [298a-300]). Most of these glycosylated enzymes retain their catalytic and stability properties when expressed in mesophilic hosts (not known to extensively glycosylate recombinant proteins). Also, a major source of thermal destabilization for the *A. missouriensis* glucose isomerase has been shown to be the glycosylation of a Lys residue by glucose - a glycoside as well as the glucose isomerase substrate [264]. While some glycosylated proteins have been shown to be more stable than their non-glycosylated forms [199,301], glycosylation is not a thermostabilization method commonly found in nature.

salts - Inorganic salts stabilize proteins by either a specific salt effect, where a metal ion interacts with the protein in a conformational manner or a general salt effect, which mainly affects water activity. Ca^{2+} binding by α-Amylases is an example of the former mechanism. The α-amylase catalytic site is located in a cleft between two domains (an $[\alpha/\beta]_8$ barrel and a large loop). Coordinated by ligands belonging to these two domains, Ca^{2+} is essential for the α-amylase's catalytic activity and thermostability [302]. Xylose isomerases bind two metal ions (either Co^{2+}, Mg^{2+}, or Mn^{2+}). One cation is directly involved in catalysis (the catalytic metal) and the second stabilizes the folded protein (the structural metal) [303,304]. The two metal-binding sites can have different specificities, and replacing one cation with another often significantly alters enzyme

activity, substra

lysozyme was

protein, relative

specific salt eff

size, coordinati

A study

K_3PO_4, Na_3PO

reduce the enzy

decreasing effec

precipitation is

Once precipitat

convenient way

thermostability

While the five e

varied from enz

in the presence

stabilized by 0.1

more efficiently

activating effect

satisfied by the i

(>1 M potassium

reductase activity

from being the o

salt) [155]. Thes

enzymes from dif

requirements, (ii)

activity, substrate specificity, and thermostability [303,305]. Ca^{2+} stabilization of

lysozyme was proposed to be due to a decrease in the entropy of the Ca^{2+}-bound unfolded

protein, relative to the entropy of the unfolded protein in the absence of Ca^{2+} [298]. This

specific salt effect was attributed to the characteristics of the specific binding sites (i.e.,

size, coordination number, and the nature of the liganding residues) [65].

A study of GAPDH thermostabilization by salt indicated that the relative effects of

K_3PO_4, Na_3PO_4, K_2SO_3, Na_2SO_3, KCl, and NaCl were consistent with their abilities to

reduce the enzyme solubility in aqueous solvent. Their action was attributed to their

decreasing effect on water activity, a general salt effect [76]. Ammonium sulfate protein

precipitation is a common application of the general salt effect for enzyme precipitation.

Once precipitated, enzymes are typically also stabilized. This procedure is a standard,

convenient way to store enzymes. Thauer and colleagues studied the effect of salts on the

thermostability and activity of five M. kandleri methanogenic enzymes [153-155,164,183].

While the five enzymes are activated and stabilized by salts, the extent of the salt effect

varied from enzyme to enzyme. CHO–H_4MPT formyltransferase was optimally stabilized

in the presence of 1.5 M K_2HPO_4 while the F_{420}-dependent CH_2=H_4MPT reductase was

stabilized by 0.1 M K_2HPO_4. K^+ and NH_4^+ cations typically improve enzyme stabilities

more efficiently than other cations. Of all the anions, SO_4^{2-} and HPO_4^{2-} had the strongest

activating effect [164]. Enzyme stabilizing salt requirements however, are not always

satisfied by the intracellular salt concentration. M. kandleri's intracellular salt concentration

(>1 M potassium plus 1.2 M 2,3-diphosphoglycerate) [164] seems to favor CH_2=H_4MPT

reductase activity (optimal at 2.0–2.5 M salt), but this intracellular salt concentration is far

from being the optimal enzyme stabilizing concentration (optimal between 0.1 and 1.5 M

salt) [155]. These studies illustrate the variety of situations encountered: (i) similar

enzymes from different organisms do not have the same stabilization and activation

requirements, (ii) enzymes from the same organism are not equally affected by an

environmental

necessarily cor

The hypothesis

between charge

data [306]. Ins

protein destabil

the interface be

This hypothesis

salt concentratio

M) required to s

polar uncharged

mesophilic cou

energy increase

with the less po

than the intera

theory predicts

their partially

Other

differences be

Methanosarc

correlated to

organisms. F

known to stab

thermamines (

proteins [307]

proteins. How

environmental (i.e., cellular) factor, and (iii) thermophilicity and thermostability are not necessarily controlled or favored by the same factors.

The hypothesis that the general salt effect was due to direct electrostatic interactions between charged amino acids and salt counterions was not corroborated by experimental data [306]. Instead, a smeared charge repulsion mechanism was proposed: Salt-induced protein destabilization was due to nonspecific charge repulsion and to ionization changes at the interface between complementary peptide surfaces in the partially unfolded protein. This hypothesis agrees with the well-documented protein stabilizing effect of intermediate salt concentrations, in particular with the relatively high salt concentrations (1.0 M to 2.0 M) required to stabilize some thermophilic enzymes [306]. It also agrees with the reduced polar uncharged residue content in some thermophilic enzymes, when compared to their mesophilic counterparts [85,230]. At high temperatures, the system's higher kinetic energy increases the potential for exposing buried surfaces to the solvent. Salt interactions with the less polar residues in the partially unfolded thermophilic enzyme are less favorable than the interactions with the more polar mesophilic enzyme residues. Furthermore, this theory predicts that greater solvent ionic character will stabilize folded proteins relative to their partially unfolded forms.

Other chemical effectors - Breitung *et al.* (1992) [164] found that the differences between *M. kandleri*, *M. thermoautotrophicum*, *Archaeoglobus fulgidus*, and *Methanosarcina barkeri* CHO-H$_4$MPT formyltransferase activation by salts were directly correlated to the intracellular 2,3-diphosphoglycerate concentration in the different organisms. Polyalcohols such as glycerol, sorbitol, mannitol, sucrose, and starch and are known to stabilize proteins [306]. Hyperthermophilic bacterial cytoplasms contain thermamines (polyamines) that have been shown to thermostabilize nucleic acids and proteins [307]. Organic polymers such as polyethylene glycol (PEG) also stabilize folded proteins. However, they can also induce precipitation and even inactivation.

Pressur

reported [308,30

crystalline solid

Unfolding a pro

stable than the f

compact , folde

pressure has on

thermophiles ha

from 200 to 40(

from organisms

reactions are pe

at high pressure

Molecular m

The mo

literature on thi

that if there are

they are either i

Arrhenius curve

these specific ca

occur. These di

thermophilic en

Arrhenius plots

discontinuous [9

not involve alter

[207,208]. Gene

the Arrhenius eq

Pressure effects - The ability of high pressure to stabilize proteins has been reported [308,309]. The observation that folded proteins have densities similar to crystalline solids is the theoretical basis for pressure as a general stabilizing force. Unfolding a protein would increase its volume, making a partially unfolded protein less stable than the folded protein at increased pressures. Thus, at higher pressures, the compact , folded protein form is more stable. Hei and Clark [308] have studied the effect pressure has on enzymes from barophilic organisms. They showed that some barophilic thermophiles have enzymes which are also barophilic (i.e., optimal activity at pressures from 200 to 400 atm) [308]. Hei and Clark also characterized barophilic enzymes isolated from organisms that were not, themselves, barophilic [308]. Since numerous chemical reactions are performed at high temperatures and pressures, demonstrating enzyme stability at high pressures is potentially important for enzyme biotechnological applications [309].

Molecular mechanisms of protein thermophilicity

The molecular basis for enzyme thermophilicity is not well understood, and the literature on this subject is limited. Linear Arrhenius plots for numerous enzymes indicate that if there are structural changes in the catalyst throughout the active temperature range, they are either not catalytically significant or they are offsetting. The discontinuous Arrhenius curves observed for catalysis by some enzymes [82,145,190] indicate that, in these specific cases, temperature dependent, functionally significant structural changes may occur. These discontinuous Arrhenius curves were believed to be characteristic of thermophilic enzymes, but, as mentioned previously, not all thermophilic enzyme Arrhenius plots are discontinuous and some mesophilic enzyme Arrhenius plots are also discontinuous [92]. Explanations for Arrhenius discontinuities have been proposed that do not involve altering the catalyst's structure (e.g., change in the slow step of the reaction) [207,208]. Generally, enzyme activity in its active temperature range is well described by the Arrhenius equation, and can be derived from the temperature for maximal activity,

magnitude of

equation, how

Thus, an enzy

the theoretical

protein unfold

independently

Most s

thermophilicit

thermostabilit

binding site in

increased its t

the absence of

and mutant en

activity of the

activity contin

its own tempe

enzyme therm

thermophilicit

insertion of 2

stability-only

Insertion of a

fold, the therm

support the hy

stability. Othe

reduced stabili

(70°C), sugges

can be structur

magnitude of maximal activity, and the activation energy for the reaction. The Arrhenius equation, however, cannot predict the optimal temperature for an enzymatic reaction. Thus, an enzyme's active temperature range appears to be determined at the lower end by the theoretical limitation on activity determined by the Arrhenius relationship. Partial protein unfolding or covalent chemical modifications appear to determine the upper limit, independently from the Arrhenius equation.

Most studies on enzyme thermostabilization do not comment on the enzyme thermophilicity. Only a few reports [250,310] describe mutations that altered both enzyme thermostability and thermophilicity. Kuroki et al. [250] demonstrated that adding a Ca^{2+} binding site in human lysozyme by site directed mutagenesis both stabilized the protein and increased its thermophilicity from 70°C (native enzyme) to 80°C (mutant holoenzyme). In the absence of Ca^{2+}, the mutant enzyme thermophilicity was reduced to 65°C. The native and mutant enzyme catalytic rates were superimposable up to the temperature for maximal activity of the wild-type enzyme (~70°C). Above that temperature the mutant enzyme activity continued to increase, with the same apparent Arrhenius behavior, until it reached its own temperature for maximal activity. The mutant behavior suggests that the wild-type enzyme thermophilicity was limited by its thermostability, and that the higher thermophilicity of the mutant was due to an increase in its thermostability. In one study, insertion of 2 to 4 amino acids into *Caldicellulosiruptor saccharolyticus* xylanase generated stability-only mutants and stability-plus-optimal temperature-altering mutants [310]. Insertion of a single Pro-Arg sequence reduced the enzyme's specific activity more than 2-fold, the thermostability 4-fold, and the optimal temperature by 20°C. These mutations support the hypothesis that, in this case, thermophilicity was also limited by folded protein stability. Other *C. saccharolyticus* xylanase mutants, displayed reduced activity and reduced stability (similar to the Pro-Arg mutant) while retaining native-like thermophilicity (70°C), suggested that the molecular bases for protein thermostability and thermophilicity can be structurally distinct. The Tomazic and Klibanov thermostability model (scheme (1),

and [197]) can potentially explain the observed difference between the two classes of *C. saccharolyticum* xylanase mutants. The mutations which only reduce enzyme thermostability would increase the rate of scrambled structure formation from the enzyme non-native state, while the mutations that reduce both thermophilicity and thermostability would result from the shift of native to nonnative enzyme at lower temperatures. This analysis predicts that a class of thermophilicity-reducing mutants that do not alter thermostability could be generated if the nonnative state (resulting from reversible unfolding) were stabilized with respect to the scrambled structure (the enzymes would only reversibly inactivate).

While the relationship between enzyme kinetics and protein thermophilicity/ thermostability remains unclear, it has been proposed that the poor activity of thermozymes at low temperatures is the result of excessive rigidity. This rigidity is believed to be necessary to maintain active enzyme architecture at high temperatures [198,203,215]. Data from crystallography, deuterium exchange, proteolytic susceptibility, and other experimental approaches have demonstrated that folded thermozymes are indeed more rigid at low temperatures than their mesophilic counterparts [198,203,215]. Tchernajenko *et al.* [206], however, reported linear Arrhenius plots for the four xylose isomerases they studied. Since these plots were perfectly fit by the Arrhenius equation, increased low-temperature protein rigidity was not required to explain the poor low-temperature activity of thermophilic enzymes [206]. The temperature dependence of thermophilic enzyme activity might, instead, only be determined by temperature-dependent substrate kinetic energy variations. Therefore, determining the importance of protein flexibility to thermophilic enzyme activity will require comparisons of corresponding thermophilic and mesophilic enzyme flexibilities and activities throughout their active temperature ranges.

Genetic engineering of thermozymes

Modification of enzyme catalytic properties

The accumulating protein structural data justifies the use of enzyme engineering as a rational research approach to make enzyme catalytic properties fit industrial processes. Knowing the L-lactate dehydrogenase structure and catalytic mechanism, Holbrook and colleagues attempted to systematically modify the B. stearothermophilus L-lactate dehydrogenas into a malate dehydrogenase by a point mutation proposed to better accommodate oxaloacetate in the substrate pocket [311], then to design a non-specific α-hydroxy acid dehydrogenase from the lactate dehydrogenase [312]. This non-specific α-hydroxy acid dehydrogenase mutant enzyme remained thermostable, and could be used for various chemical syntheses. A similar approach was used by Meng et al. [313] to switch the substrate preference of T. thermosulfurigenes xylose isomerase from xylose to glucose. Arnold and colleagues used PCR-mediated random mutagenesis to enhance subtilisin E activity in polar organic solvents [313a,313b]. Using several screening steps, they eventually selected a multiple mutant that was 256 times more efficient than the wild-type enzyme in 60% dimethylformamide. A similar PCR based technique, was successfully employed to enhance the Z. mobilis 1° ADH thermostability [217]. These studies provide proof of principle for the general application of directed enzyme design strategies using mutagenesis.

Thermostability and thermophilicity engineering

Research has shown that while thermophilicity and thermostability can be altered independently, they are often structurally related. Numerous molecular interactions might be responsible for these properties, however, only a few are of genetic engineering value: (i) protein sequences contain redundant information for proper folding, and protein structures can often accommodate amino acid substitutions without significantly altering the catalytic efficiency; (ii) a small number of amino acid substitutions throughout the protein

can significan

protein are m

ones) is critic

stabilization h

surface loop s

analysis has i

thermophilic

altering protei

helices by cap

same engineer

successfully st

methods (eg. H

dimensional str

identified (e.g.,

in constrained l

Preventing pept.

applications (>9

hyperthermophil

shown to be a m

appropriate strate

the project goal,

limiting enzyme

properties as clos

engineering. The

initial enzyme's hi

class that has therr

thermophilicity and

can significantly alter its thermophilicity and thermostability; and (iii) certain regions in a protein are more labile than others–stabilizing these regions (rather than the more stable ones) is critical to improve protein thermal properties. Four general strategies for thermal stabilization have been identified: more efficient protein core packing, α-helix stabilization, surface loop stabilization, and prevention of chemical degradation. Recent crystallographic analysis has indicated that more-rigid core structure (lower β-values) is a hallmark of thermophilic proteins [124,198]. However, systematic engineering of core rigidity without altering protein function is beyond the capabilities of current technology. Stabilizing α-helices by capping or by introducing alanines has had spotty success, and suffers from the same engineering problems as do the core packing strategies. Loop engineering has successfully stabilized proteins [233,243,248,254,258,259,272]. Sequence analysis methods (eg. HCA) now reliably predict surface loop regions, even in the absence of 3-dimensional structural information. Specific stabilizing amino acid substitutions have been identified (e.g., substitution of surface lysines with arginines and introduction of prolines in constrained loops, and salt bridges), making this approach a good engineering tool. Preventing peptide chemical degradation is also important for very high temperature applications (>90°C). Trends toward reduced Cys, Asn, and Gln contents in hyperthermophilic enzymes have been identified in nature [303], and deamidation has been shown to be a main factor responsible for irreversible enzyme inactivation [197]. The appropriate strategy for engineering protein thermostability or thermophilicity depends on the project goal, on the available structural information, and on the thermal mechanisms limiting enzyme stability or activity (Fig. 3). Selecting an initial enzyme with thermal properties as close to the target as possible clearly increases the chances of successful engineering. The probability of engineering success is further enhanced by having the initial enzyme's high-resolution 3-dimensional structure as well as a protein in the same class that has thermal characteristics close to the desired protein. Comparing thermophilicity and thermostability properties indicates whether the limiting step in protein

Figure 3. Flow chart of potential steps toward engineering enzyme thermophilicity and thermostability.

Design goal

Thermophilicity   Thermostability

Reversible protein inactivtion   Irreversible protein inactivtion
(Thermostability > Thermophilicity)   (Thermostability = Thermophilicity)

No precipitation   Precipitation

Thermophilicity

Precipitate   Precipitate
active   inactive

Prevent structural change   *Thermophilicity*

Complete   Partial

Thermophilicity

General:
1) Random mutagenesis with selection
2) sequence comparison directed mutagenesis
3) substitute Pro residues into loops
4) replace Lys with Arg in loops
5) chemically crosslink
6) immobilize
7) add stabilizing chemicals to the system
    a) protein specific
        i) salts (NaCl, KCl, etc.)
        ii) substrates/effectors
    b) reduce solvent entropy
        i) glycerol
        ii) Hofmeister series ions

Prevent precipitation:
1) Immobilize
2) Chemically modify
    a) succinylate Lys residues
    b) intramolecular crosslink
3) add polyalcohols to system

Only with 3-D structure information:
1) add divalent cation binding site
2) reduce bond strain by mutagenesis
3) fill internal cavities by mutagenesis
4) stabilize helices
    i) His/Gly to cap helix C-terminus
    ii) Glu/Asp to cap helix N-terminus
    iii) Ala to stabilize helix

Only for reversible inactivation:
1) add Gly to loops

Target temperatures above 80°C only:
1) reduce the number of Asn, Gln, and Cys residues

Target temperatures below 80°C with 3-D structure information only:
1) add disulfide bonds

inactivation is the initial partial unfolding or irreversible structural changes (such as precipitation or chemical degradation). If, in the temperature range of interest, the protein thermostability half-life is longer than the activity half-life, the reversibly partially unfolded protein is not immediately irreversibly inactivated. If, however, these two half-lives are similar, irreversible inactivation follows the initial unfolding almost instantly. The protein concentration in the target application should be estimated based on the protein function (receptor, enzyme, antibody, etc.), to provide a concentration range for these characterizations. Determination of the thermostability half-life at different initial protein concentrations within this range will indicate if inactivation is intramolecular (concentration independent) under these specific conditions.

Having increased enzyme thermostability allows applications of repeated thermal cycling (e.g., PCR), as well as permitting pre-application high-temperature processing steps (e.g., processing enzymes in food or feed additives). To increase enzyme thermostability, engineering efforts should focus on preventing irreversible protein inactivation. Strategies to prevent protein aggregation (e.g., chemical modification of lysine residues with succinic acid [197], addition of polyalcohols to the solution [306], or immobilization on an inert matrix) might significantly thermostabilize the protein. This approach would need to be undertaken if (i) enzyme precipitation coincides with inactivation and the precipitated protein does not retain complete activity, or (ii) if aggregation is undesirable for the target application. If these measures are insufficient or inappropriate for the design goal, then prevention of inactivating conformational changes through peptide modifications such as intramolecular cross-linking or through system additives (eg. salt or glycerol) that stabilize the folded protein structure should be examined. Salts have been shown to stabilize proteins through specific binding, charge shielding, and modification of the system entropy. The addition of organic chemicals such as glycerol and polyethyleneglycol has also been shown to stabilize some proteins, but they often reduce activity and can destabilize some proteins. The specificity of intramolecular

cross-linking is difficult to control and often reduces protein activity but it covalently tethers protein regions analogously to disulfide linkages.

Either random or site directed mutagenesis of the initial protein may be used to thermostabilize proteins. Random mutagenesis has been successfully used to increase enzyme thermostability and resistance to solvent-induced denaturation. Hageman and colleagues [314] developed a system in which a gene encoding a mesophilic enzyme is introduced into a thermophilic host, and variant enzymes with increased thermostability are selected during growth at increasing temperatures. Heat-stable kanamycin nucleotidyltransferase mutants were obtained using a *B. stearothermophilus* strain that expressed a mesophilic kanamycin nucleotidyltransferase to select kanamycin resistant variants at increasing temperatures [239]. Although attractive at first glance, this approach is limited by the absence of genetic tools (i.e., shuttle vectors, transformation methods) available for most thermophiles and hyperthermophiles, as well as by the limited enzymatic activities directly selectable in *B. stearothermophilus* cultures. The current lack of thermophilic cloning hosts also makes classic genetic complementation by thermophilicity mutants a significant challenge. Practical use of random mutagenesis requires a powerful mutant selection procedure or specific structural information (to limit the size of the target thus, reducing the total number of possible mutants to be tested).

Site-directed mutagenesis usually requires protein structural information. Strategies that include substitutions with residues present in structurally similar proteins with thermal properties similar to the desired mutant require sequence information from both proteins. If an enzyme targeted for hyperthermophilic applications (>80°C) is irreversibly inactivated, reduction of the number of noncatalytic Gln, Asn, and Cys residues can be used to stabilize the protein structure, using the single peptide sequence. Loop stabilization by introducing Pro or Arg residues also requires only knowing the protein sequence. Exact structural information from crystallography, NMR, or homology modeling allows the engineering of multi-residue motifs such as metal binding sites or disulfide bridges (for moderate

temperature stabilization) to stabilize the enzyme structure, and allows specifically tailored design schemes such as cavity filling mutations, helix stabilization by alanine insertions or capping (eg. His at the c-terminus), and alleviation of specific bond strains. Deuterium exchange, measured by NMR, can precisely identify the more labile protein regions to stabilize [315]. Muheim *et al.* [316] created a dihistidine metal-chelating site on a surface β-sheet of cytochrome *c*, and cross-linked it with the metal complex Ru^{II}(2,2'-bipyridine). This cross-linking increased the melting temperature of the mutant enzyme by more than 23°C.

Increasing enzyme thermophilicity requires preventing thermally induced unfolding. Since enzyme thermophilicity is often limited by thermostability, engineering thermophilicity usually requires enzyme thermostabilization. Also note that, for an enzyme that precipitates upon heating, if precipitation is not the first inactivation step, preventing aggregation is unlikely to enhance thermophilicity.

THESIS OBJECTIVES AND SIGNIFICANCE

Because archaeal and bacterial thermophiles were the first organisms to have evolved the study of these organisms and their enzymes is expected to provide insights into the origins of biocatalysis on Earth [317]. With their remarkable thermostability and activity, thermophilic enzymes provide a model for studying protein thermostability and biocatalysis at high temperatures. Structurally similar to their mesophilic counterparts, thermophilic enzymes are powerful tools for developing enzyme structural or functional models through comparative analysis. Their high-temperature activity and high stability make thermophilic enzymes excellent candidates for industrial enzymatic applications. Therefore, the fundamental scientific examination of thermophilic life may also yield novel catalysts of tremendous biotechnological value.

Numerous highly similar thermophilic 2° ADHs have been isolated from thermoanaerobic bacteria [7,8,15,16] but despite more that a decade of research since the first enzymes were purified in 1981 [14,15], no report of the cloning of a gene encoding a thermophilic 2° ADH has been published. There is also no record of either mesophilic or thermophilic 2° ADH characterization by mutagenesis. The only thermophilic 2° ADH peptide sequence avaliable was determined by Edman degradation [33]. Alignment of the liver 1° ADH and the *T. brockii* 2° ADH peptides by Peretz and Burstein [33] suggested that the structural Zn binding loop was absent in the 2° ADH, but they concluded that insufficient similarity existed between the *A. eutrophus*, horse liver and *Thermoanaerobacterium brockii* enzymes for further comparative analysis. Earlier, Lamed and Zeikus reported that *T. brockii* 2° ADH activity was both cysteine and Zn dependent [15] and Bryant and Ljungdahl measured Zn specifically bound to the *T. ethanolicus* (ATCC 31550) enzyme [14]. All thermophilic 2° ADHs identified are tetrameric with reported subunit molecular masses near 40 kDa. This coincidence of enzyme structural properties and the similarities seen in kinetic comparisons of these NADP(H) linked 2° ADHs led Nagata *et al.* to conclude that all thermophilic 2° ADHs were extremely structurally similar [16]. The greater than 75% sequence identity between the mesophilic *C. beijerinckii* and thermophilic *T. brockii* peptides further argues that all catabolic 2° ADHs may share significant 3-dimensional structural identity. These thermophilic and mesophilic 2° ADH peptide sequences therefore, make an excellent system for determination of protein thermostability structure-function relationships through comparative analysis. However, reciprocal mutagenesis experiments to verify the sequence based predictions require expressed clones of the genes encoding both proteins.

The *T. brockii* 2° ADH has been demonstrated to reduce aldehydes and ketones on aliphatic molecules containing up to 10 carbons [318]. This enzyme requires ketones on the second or third carbon, displaying higher activity toward unbranched noncyclic molecules. The enantiospecificity was lower and (R)- specific in the reduction of very

sho

cha

yea

AD

Ena

and

AD

con

bind

AD

enzy

was

abov

by th

chira

pote

have

oxyg

2'AD

1981

40%

tempe

cataly

nonaq

toward

range o

short ketones but greater than 90% and (S)- specific for reduction of ketones with aliphatic chains greater than 5 carbons in length. The *T. brockii* 2° ADH, like the horse liver and yeast 1° ADHs, produces (S)-alcohols using the E3 pathway [30]. For the *T. brockii* 2° ADH, lower reaction temperatures also yielded higher enantiomeric excess [319]. Enantiomeric excess at equillibrium is a function of the Gibbs energy ratio between the (R)- and (S)- specific reactions and this relationship indicated that the variation of *T. brockii* 2° ADH reaction stereospecificity with temperature resulted mainly from entropy considerations. An ordered kinetic mechanism with cofactor binding followed by substrate binding, hydride transfer, product release, then cofactor dissociation, similar to the liver 1° ADH mechanism, was identified for *T. brockii* 2° ADH catalysis [320]. The *T. brockii* enzyme is both thermostable and thermophilic [15]. Its optimal temperature for activity was reported to be above 70°C with significant inactivation rates occurring at temperatures above 90°C. A discontinuity in the Arrhenius plot for propan-2-ol and butan-2-ol oxidation by this enzyme was reported in the same study.

The *T. brockii* 2°ADH has already been used in the analytical scale production of a chiral constituent of civet (a perfume additive) [321] and of an insect pheromone [322]. Its potential value for other syntheses has also been examined [323]. Thermophilic 2° ADHs have high temperature optima (80 to 95°C), high thermostability, and reduced sensitivity to oxygen (they are only reversibly inactivated, unlike 1°ADHs) [8]. The extent of *T. brockii* 2°ADH stability in the presence of organic solvent is still controversial. In the original 1981 study [323] the enzyme remained 80% active after 15 min at 52°C in the presence of 40% propanol-2, whereas it was reviewed by Cowan [59] to rapidly inactivate at temperatures above 45°C and in the presence of 10% organic solvent. *T. brockii* enzyme catalysis was determined *in vitro* in supercritical CO_2 indicating that it is active in nonaqueous conditions under high pressure [324]. The *T. brockii* enzyme was also active toward both highly water soluble ketones and substrates with low water-solubility in a range of water-oil emulsions containing either cationic or neutral surfactants [325]. A

ra

Al

of

Th

the

the

chi

the

cha

ana

ove

prov

char

com

beije

site (

cata

bioch

.

ethan

range of potential chiral transformations has already been demonstrated for ADHs. 2° ADHs, optimized to interconvert ketones and 2° alcohols, are better suited to the production of chiral alcohols that 1° ADHs, optimized to interconvert aldehydes and 1° alcohols. Therefore, continued research into the molecular mechanisms responsible for 2° ADH thermostability, thermophilicity, and substrate specificity will advance our understanding of these general enzyme properties specifically using an enzyme with potential as an industrial chiral catalyst.

The research objectives presented in this thesis advances knowledge on thermophilic 2° ADH 2° ADH structure-function relationships in 5 ways. First, cloning and characterizing the gene encoding the *T. ethanolicus* 39E 2° ADH will make mutagenic analysis of thermophilic 2° ADH structure-function relationships possible. Second, stable overexpression of the properly active recombinant thermophilic 2° ADH in *E. coli* will provide sufficient quantities of easily purified protein for structural analysis. Third, characterization of enzyme thermostability and thermophilicity will allow sequence comparisons between this thermophilic enzyme and the similar mesophilic enzyme from *C. beijerinckii* for identification of 2° ADH thermal property molecular determinants. Fourth, site directed mutagenesis and biophysical analysis will test the predictions of 2° ADH catalytic architecture based on comparison to those documented for 1° ADHs. Finally, the biochemical mechanism of thermoinactivation and the forces stabilizing active the *T. ethanolicus* 2° ADH structure will be elucidated.

ACKNOWLEDGMENTS

72

REFERENCES

1. Brändén, C.-I., Jörnvall, H., Eklund, H. and B. Fururgren (1975) in The enzymes, 3rd ed., vol. XI part A" (P. D. Boyer, ed.) pp. 103-190. Academic Press, New York.

2. Gottschalk G. (1986) Bacterial metabolism, 2nd edition. Springer-Verlag, New York.

3. Schimpfessel, L. (1967) Arch. Int. Physiol. Biochim. **75**, 368-369

4. Schimpfessel, L. (1968) Biochim. Biophys. Acta **151**, 386-?.

5. Lutstorf, U. and Megnet, R. (1968) Arch. Biochem. Biophys. **126**, 933-944

6. Sund, H. and Theorell, H. (1963) in The enzymes, 2nd ed., vol. 7, (Boyer, P D., Lardy, H. and Myrbäck, K., ed.), pp. 25-83 Academic Press, New York.

7. Bryant, F. O., Weigel, J. and Ljungdahl, L. G. (1988) Appl. Environ. Microbiol. **54**, 460-465

8. Burdette, D. S. and Zeikus, J. G. (1994) Biochem. J. **302**, 163-170

9. Kolb, E. and Harris, J. I. (1971) Biochem J. 124, 76P-77P

10. Rella, R., Raia, C. A., Pensa, M., Pisani, F. M., Gambacorta, A., De Rosa, M., and Rossi, M. (1987) Eur. J. Biochem. **107**, 475-479

11. Ma, K., Robb, F. T., and Adams, M. W. W. (1994) Appl. Environ. Microbiol. **60**, 562-568

12. Ismaiel, A. A., Zhu, C.-X., Colby, G. D. and Chen, J.-S. (1993) J. Bacteriol. **175**, 5097-5105

13. Steinbüchel, A. and Schlegel, H. G. (1984) Eur. J. Biochem. **141**, 555-564

14. Bryant, F. O. and Ljungdahl, L. G. (1981) Biochem. Biophys. Res. Com. **100**, 793-799

15. Lamed R. J. and Zeikus, J. G. (1981) Biochem. J. **195**, 183-190

16. Nagata, Y., Maeda, K. and Scopes, R. K. (1992) Bioseparation 2, 353-362

17. Neale, A. D., Scopes, R. K., Kelly, J. M., and Wettenhall, R. E. H. (1986) Eur. J. Biochem. **154**, 119-124

18. Bleicher, K. and Winter, J. (1991) Eur. J. Biochem. **200**, 43-51

19. Ainsley, G. R. Jr. and Cleland, W. W. (1972) J. Biol. Chem. **247**, 946-951

20. Brooks, R. L. and J. D. Shore (1971) Biochemistry **10**, 3855-3858

73

21.   Bush, K., Shiner, V. J., and H. R. Mahler (1973) Biochemistry **12**, 4802-4805

22.   Beijer, N. A., Buck, H. M., Sluyterman, L. A. AE., and Meijer, E. M. (1990) Biochimica Biophysica Acta **1039**, 227-233

23.   Dunn, M. F. and Hutchison, J. S. (1973) Biochemistry **12**:4882-4892

24.   Jacobs, J. W., McFarland, Wainer, I., Jeanmaier, D., Ham, C., Hamm, K., Wnuk, M., and Lam, M. (1974) Biochemistry **13**, 60-64

25.   McFarland, Chu, Y.-H., and J. W. Jacobs (1974) Biochemistry **13**, 65-69

26.   Eklund, H., Nordström, B., Zeppezauer, E., Söderlund, G., Ohlsson, I., Boiwe, T. and Brändén, C.-I. (1974) FEBS Lett. **44**, 200-204

27.   Rossmann, M. G. Moras, D., and K. W. Olsen (1974) Nature **250**, 194-199

28.   Wierenga, R. K. and Hol, G. J. (1983) Nature **302**, 842-844

29.   Prelog, V. (1964) Pure Appl. Chem **9**, 119-130

31.   Bradshaw, C. W., Hummel, W. and Wong, C.-H. (1992) J. Organic Chem. **57**, 1533-1536

30.   Bradshaw, C. W., Shen, J.-G. and Wong, C.-H. (1992) J. Organic Chem. **57**, 1526-1532

32.   Jendrossek, D., Steinbuechel, A. and Schlegel, H. G. (1988) J. Bacteriol. **170**, 5248-5256

33.   Peretz, M. and Y. Burstein (1989) Biochemistry **28**, 6549-6555

34.   Kullmann, W. (1987) CRC Press, Inc., Boca Raton, FL

35.   Cheetham, P. S. J. (1993) Trends Biotechnol. **11**, 478-488

36.   Dordick, J. S. (1992) Trends Biotechnol. **10**, 287-293

37.   Faber, K. and Franssen, M.C.R. (1993) Trends Biotechnol. **11**, 461-470

38.   Jones, J. B. (1986) Tetrahedron **42**, 3351-3403

38a.   Sarney, D. B. and Vulfson, E. N. (1995) Trends Biotechnol. **13**, 164-172

39.   Zong, M.-H., Fukui, T., Kawamoto, T. and Tanaka, A. (1991) Appl. Microbiol. Biotechnol. **36**, 40-43

40.   Fuganti, C. and Grasselli, P. (1985) Tetrahedron Let. **26**, 101-104

41.   Nakamura, K., Inuoe, K., Ushio, K., Oka, S. and Ohno, A. (1987) Chem. let. **4**, 679-682

42.   Trincone, A., Lama, L., Lanzotti, V., Nicolaus, B., De Rosa, M., Rossi, M. and Gambacorta, A. (1990) Biotech. Bioeng. **35**, 559-564

43. Hummel W. (1990) Appl. Microbiol. Biotechnol. **34**, 15-19

44. Whitesides, G. M. and Wong, C.-H. (1983) Aldrichimica Acta **16**, 27-34

45. Rosazza, J. P. N. (1995) Amer. Soc. Microbiol. News **61**, 241-245

46. Meyer, H.-P. (1991) Biol. Ferm. Eng. **8**, 602-606

47. Wong, C.-H. (1989) Science **244**, 1145-1152

48. Irwin, J. B., Lok, K. P., Huang K. W. C. and Jones, J. B. (1978) J. Chem. Soc. Perkin I **12**, 1636-1641

49. Bradshaw, C. W., Hummel, W. and Wong, C.-H. (1992) J. Org. Chem. **57**, 1532-1536

50. Bradshaw, C. W., Fu, H., Shen, G.-J. and Wong, C.-H. (1992) J. Org. Chem. **57**, 1526-1532

51. Hummel, W. and Kukla, M. R. (1989) Eur. J. Biochem. **184**, 1-13

52. May, S. W. and Padgette, S. R. (1983) Bio/Technology **1**, 677-686

53. Persson, M., Månsson, M.-O., Bülow, L. and Mosbach, K. (1991) Biotechnology **9**, 280-284

54. Klibanov, M. (1983) Proc. Biochem. **18**, 13-23

55. Bel-Rhlid, R., Fauve, A., Renard, M. F. and Veschambre, H. (1992) Biocatalysis **6**, 319-337

56. Zaks, A. and Klibanov, A. M. (1988) J. Biol. Chem. **263**, 8017-8021

57. Klibanov, A. M. (1989) Trends Biol. Sci. **14**, 141-144

58. Samama, J. P., Lee, K. M. and Biellmann, J. F. (1987) Eur. J. Biochem. **163**, 609-617

59. Cowan, D.A. (1992) Trends Biotechnol. **10**, 315-323

60. Lee, Y.-E., Lowe, S. E. and Zeikus, J. G. (1993) Appl. Environ. Microbiol. **59**, 3134-3137

61. Lehmacher, A. and Hensel, R. (1994) Mol. Gen. Genet. **242**, 163-168

62. Lee, C., Bhatnagar, L., Saha, B. C., Lee, Y.-E., Takagi, M., Imanaka, T., Bagdasarian, M. and Zeikus, J. G. (1990) Appl. Environ. Microbiol. **56**, 2638-2643

63. Zeikus, J. G. and Brock, T. D. (1971) Biochim. Biophys. Acta **228**, 736-745

64. Adams, M. W. W. (1993) Ann. Rev. Microbiol. **47**, 627-658

65. Coolbear, T., Whittaker, J.M. and Daniel, R.M. (1992) Biochem. J. **287**, 367-374

66. Daniel, R. M. (1992) Origins Life Evol. Biosphere **22**, 33-42

67. Lowe, S. E., Jain, M. K., and Zeikus, J. G. (1993) Microbiol. Rev. **57**, 451-509

68. Stetter, K. O., Fiala, G., Huber, G., Huber, R., and Segerer, A. (1990) FEMS Microbiol. Rev. **75**, 117-124

69. Wiegel, J. and Ljungdahl, L. G. (1984) CRC Crit. Rev. Biotechnol. **3**, 39-108

70. Coolbear, T., Daniel, R. M. and Morgan, H. W. (1992) Adv. Biochem. Eng.-Biotechnol. **45**, 57-98

71. Vieille, C., Burdette, D. S. and Zeikus, J. G. (1996) in Biotech. Ann. Rev., vol. 2, (M. R. El Geweley, ed.) pp. 1-83. Elsevier, Amsterdam, Neth.

72. Berquist, P. L. and Morgan, H. W. (1992) in Molecular biology and Biotechnology of extremophiles, (Herbert, R. A. and Sharp, R. J., eds.), pp. 44-75. Chapman and Hall, NY

73. Winter, J., Lerp, C., Zabel, H.-P., Wildenauer, F. X., König, W. H., and Schindler, F. (1984) System. Appl. Microbiol. **5**, 457-466

74. Schmitz, R.A., Richter, M., Linder, D. and Thauer, R.K. (1992) Eur. J. Biochem. **207**, 559-565

75. O'Brien, W. E., Brewer, J. M. and Ljungdahl, L. G. (1973) J. Biol. Chem. **248**, 403-408

76. Hatchikian, E. C., and Zeikus, J. G. (1983) J. Bacteriol. **153**, 1211-1220

77. Benachenhou-Lahfa, N., Labedan, B. and Forterre, P. (1994) Gene **140**, 17-24

78. Lacher, K. and Schäfer, G. (1993) Arch. Biochem. Biophys. **302**, 391-397

79. Kath, T., Schmid, R., and Schäfer, G. (1993) Arch. Biochem. Biophys. **307**, 405-510

80. Lawyer, F. C., Stoffel, S., Saiki, R. K., Myambo, K., Drummond, R. and Gelfand, D. H. (1989) J. Biol. Chem. **264**, 6427-6437

81. Davies, G. J., Littlechild, J. A., Watson, H. C., and Hall, L. (1991) Gene **109**, 39-45

82. Fabry, S., and Hensel, R. (1987) Eur. J. Biochem. **165**, 147-155

83. Sakai, K., Narihara, M., Kasama, Y., Wakayama, M., and Moriguchi, M. (1994) Appl. Environ. Microbiol. **60**, 2911-2915

84. Hartog, A. T. and Daniel, R. M. (1992) Int. J. Biochem. **24**, 1657-1660

85. Bealin-Kelly, F., Kelly, C. T. and Fogarty, W. M. (1991) Appl. Microbiol. Biotechnol. **36**, 332-336

86. Brosnan, M. P., Kelly, C. T., and Fogarty, W. M. (1992) Eur. J. Biochem. **203**, 225-231

87. Gray, G. L., Mainzer, S. E., Rey, M. W., Lamsa, M. H., Kindle, K. L., Carmona, C. and Requadt, C. (1986) J. Bacteriol. **166**, 635-643

88. Fukusumi, S., Kamizono, A., Horinouchi, S., and Beppu, T. (1988) Eur. J. Biochem. **174**, 15-21

89. Horinouchi, S., Fukusumi, S., Ohshima, T. and Beppu, T. (1988) Eur. J. Biochem. **176**, 243-253

90. Hyun, H. H. and Zeikus, J. G. (1985) Appl. Environ. Microbiol. **49**, 1162-1167

91. Melasniemi, H. (1987) Biochem. J. **246**, 193-197

92. Mathupala, S. P., and Zeikus, J. G. (1993) Appl. Microbiol. Biotechnol. **39**, 487-493

93. Plant, A. R., Clemens, R. M., Morgan, H. W. and Daniel, R. M. (1987) Biochem. J. **246**, 537-541

94. Saha, B. C., Lamed, R., Lee, C.-Y., Mathupala, S. P. and Zeikus, J. G. (1990) Appl. Environ. Microbiol. **56**, 881-886

95. Madi, E., Antranikian, G., Ohmiya, K., and Gottschalk, G. (1987) Appl. Environ. Microbiol. **53**, 1661-1667

96. Guy, G. R. and Daniel, R. M. (1982) Biochem. J. **203**, 787-790

97. Schäfer, G. and Meyering-Vos, M. (1992) Ann. N.Y. Acad. Sci. **671**, 293-309

98. Saha, B. C. and Zeikus, J. G. (1990) Appl. Environ. Microbiol. **56**, 2941-2943

99. Wind, R. D., Liebl, W., Buitelaar, R. M., Penninga, D., Spreinat, A., Dijkhuizen, L., and Bahl, H. (1995) Appl. Environ. Microbiol. **61**, 1257-1265

100. Margaritis, A., and Merchant, R. F. J. (1986) CRC Crit. Rev. Biotechnol. **4**, 327-366

101. Schofield, L. R. and Daniel, R. M. (1993) Int. J. Biochem. **25**, 609-617

102. Lüthi, E., Jasmat, N. B. and Bergquist, P. L. (1990) Appl. Environ. Microbiol. **56**, 2677-2683

103. Lee, Y.-E., Lowe, S.E., Henrissat, B., and Zeikus, J.G. 1993. J. Bacteriol. **175**:5890-5898.

104. Nakao, M., Nakayama, T., Harada, M., Kakudo, A., Ikemoto, H., Kobayashi, S. and Shibano, Y. (1994) Appl. Microbiol. Biotechnol. **41**, 337-343

105. Saha, B. C., and Zeikus, J. G. (1991) Appl. Microbiol. Biotechnol. **35**, 568-571

106. Shao, W., Obi, S. K. C., Puls, J. and Wiegel, J. (1995) Appl. Environ. Microbiol. **61**, 1077-1081

107. Lee, S.-G., Lee, D.-C., Hong, S.-P., Sung, M.-H., and Kim, H.-S. (1995) Appl. Microbiol. Biotechnol. **43**, 270-276

108. Gibbs, M. D., Saul, D. J., Lüthi, E. and Bergquist, P. L. (1992) Appl. Environ. Microbiol. **58**, 3864-3867

109. Kuriki, T., Okada, S. and Imanaka, T. (1988) J. Bacteriol. **170**, 1554-1559

110. Freeman, S.-A., Peek, K., Prescott, M., and Daniel, R. (1993) Biochem. J. **295**, 463-469

111. Peek, K., Daniel, R. M., Monk, C., Parker, L., and Coolbear, T. (1992) Eur. J. Biochem. **207**, 1035-1044

112. Peek, K., Veitch, D. P., Prescott, M., Daniel, R. M., MacIver, B., and Bergquist, P. L. (1993) Appl. Environ. Microbiol. **59**, 1168-1175

113. Kuriki, T., Park, J.-H., Okada, S. and Imanaka, T. (1988) Appl. Environ. Microbiol. **54**, 2881-2883

114. Shen, G.-J., Srivastava, K. C., Saha, B. C. and Zeikus, J. G. (1990) Appl. Microbiol. Biotechnol. **33**, 340-344

115. Aubert, J.-P., Béguin, P. and Millet, J. (1993)Genetics and Molecular Biology of Anaerobic Bacteria 1993, (Sebald, M., ed.). Springer-Verlag, NY

116. Nashihara, N., Nakamura, N., Nagayama, H. and Horikoshi, K. (1988) FEMS Microbiol. Lett. **49**, 385-388

117. Plant, A. R., Morgan, H. W. and Daniel, R. M. (1986) Enz. Microbiol. Technol. **8**, 668-678

118. Richter, O.-M. H. and Schäfer, G. (1992) Eur. J. Biochem. **208**, 343-349

119. Richter, O.-M. H. and Schäfer, G. (1992.) Eur. J. Biochem. **209**, 351-355

120. Shao, W., DeBlois, S., and Wiegel, J. (1995) Appl. Environ. Microbiol. **61**, 937-940

121. Lee, Y.-E., and Zeikus, J. G. (1993) J. Gen. Microbiol. **139**, 1235-1243

122. Lüthi, E., and Bergquist, P. L. (1990) FEMS Microbiol. Lett. **67**, 291-294

123. Barnes, E. M., Akagi, J. M. and Himes, R. H. (1971) Biochim. Biophys. Acta **227**, 199-220

124. Russell, R. J. M., Hough, D. W., Danson, M. J. and Taylor, G. L. (1994) Structure **2**, 1157-1167

78

125. Kenealy, W. R. and Zeikus, J. G. (1982) FEMS Microbiol. Lett. **14**, 7-10

126. Shing, Y. W., Akagi, J. M. and Himes, R. H. (1975) J. Bacteriol. **122**, 177-184

127. Lee, C. and Zeikus, J. G. (1991) Biochem. J. **273**, 565-571

128. Lee, Y.-E., Ramesh, M. V. and Zeikus, J. G. (1993) J. Gen. Microbiol. **139**, 1227-1234

129. Lehmacher, A. and Bisswanger, H. (1990) J. Gen. Microbiol. **136**, 679-686

130. Dekker, K., Sugiura, A., Yamagata, H., Sakaguchi, K. and Udaka, S. (1992) Appl. Microbiol. Biotechnol. **36**, 727-732

131. Dekker, K., Yamagata, H., Sakaguchi, K. and Udaka, S. (1991) J. Bacteriol. **173**, 3078-3083

132. Chan, M. K., Mukund, S., Kletzin, A., Adams, M. W. W. and Rees, D. C. (1995) Science **267**, 1463-1469

133. Mukund, S. and Adams, M. W. W. (1991) J. Biol. Chem. **266**, 14208-14216

134. Ahern, T. and Klibanov, A. M. (1985) Science **228**, 1280-1284

135. DiRuggiero, J., Robb, F. T., Jagus, R., Klump, H. H., Borges, K. M., Kessel, M., Mai, X. and Adams, M. W. W. (1993) J. Biol. Chem. **268**, 17767-17774

136. Britton, K. L., Baker, P. J., Borges, K. M. M., Engel, P. C., Pasquo, A., Rice, D. W., Robb, F. T., Scandurra, R., Stillman, T. J. and Yip, K. S. P. (1995) Eur. J. Biochem. **229**, 688-695

137. Consalvi, V., Chiaraluce, R., Politi, L., Vaccaro, R., De Rosa, M., and Scandurra, R. (1991) Eur. J. Biochem. **202**, 1189-1196

138. Robb, F. T., Park, J.-B. and Adams, M. W. W. (1992) Biochim. Biophys. Acta **1120**, 267-272

139. Consalvi, V., Chiaraluce, R., Politi, L., Gambacorta, A., De Rosa, M., and Scandurra, R. (1991) Eur. J. Biochem. **196**, 459-467

140. Wrba, A., Schweiger, A., Schultes, V. and Jaenicke, R. (1990) Biochemistry **29**, 7584-7592

141. Tomschy, A., Glockshuber, R. and Jaenicke, R. (1993) Eur. J. Biochem. **214**, 43-50

142. Schultes, V., Deutzmann, R. and Jaenicke, R. (1990) Eur. J. Biochem. **192**, 25-31

143. Zwickl, P., Fabry, S., Bogedain, C., Haas, A. and Hensel, R. (1990) J. Bacteriol. **172**, 4329-4338

144. Hensel, R., Jakob, I., Scheer, H. and Lottspeich, F. (1992) Biochem. Soc. Symp. **58**, 127-133

145. Hensel, R., Laumann, S., Lang, J., Heumann, H. and Lottspeich, F. (1987) Eur. J. Biochem. **170**, 325-333

146. Shah, N. N. and Clark, D. S. (1990) Appl. Environ. Microbiol. **56**, 858-863

147. Bryant, F. O. and Adams, M. W. W. (1989) J. Biol. Chem. **264**, 5070-5079

148. Pihl, T. D., and Maier, R. J. (1991) J. Bacteriol. **173**, 1839-1844

149. Juszczak, A., Aono, S. and Adams, M. W. W. (1991) J. Biol. Chem. **266**, 13834-13841

150. Wrba, A., Jaenicke, R., Huber, R. and Stetter, K. O. (1990) Eur. J. Biochem. **188**, 195-201

151. Ostendorp, R., Liebl, W., Schurig, H. and Jaenicke, R. (1993) Eur. J. Biochem. **216**, 709-715

152. Honka, E., Fabry, S., Niermann, T., Palm, P. and Hensel, R. (1990) Eur. J. Biochem. **188**, 623-632

153. Breitung, J., Börner, G., Scholz, S., Linder, D., Stetter, K. O. and Thauer, R. K. (1992) Eur. J. Biochem. **210**, 971-981

154. Ma, K., Zirngibl, C., Linder, D., Stetter, K. O. and Thauer, R. K. (1991) Arch. Microbiol. **156**, 43-48

155. Ma, K., Linder, D., Stetter, K. O. and Thauer, R. K. (1991) Arch. Microbiol. **155**, 593-600

156. Kunow, J., Schwörer, B., Stetter, K. O. and Thauer, R. K. (1993) Arch. Microbiol. **160**, 199-205

157. Kunow, J., Linder, D. and Thauer, R. K. (1995) Arch. Microbiol. **163**, 21-28

158. Blamey, J. M. and Adams, M. W. W. (1993) Biochim. Biophys. Acta **1161**, 19-27

159. Blamey, J. M. and Adams, M. W. W. (1994) Biochemistry **33**, 1000-1007

160. Blake, P. R., Park, J.-B., Bryant, F. O., Aono, S., Magnuson, J. K., Eccleston, E., Howard, J. B., Summers, M. F. and Adams, M. W. W. (1991) Biochemistry **30**, 10885-10895

161. Perler, F. B., Comb, D. G., Jack, W. E., Moran, L. S., Qiang, B., Kucera, R. B., Benner, J., Slatko, B. E., Nwankwo, D. O., Hempstead, S. K., Carlow, C. K. S. and Jannasch, H. (1992) Proc. Natl. Acad. Sci. USA **89**, 5577-5581

162. Uemori, T., Ishino, Y., Toh, H., Asada, K. and Kato, I. (1993) Nucleic Acids Res. **21**, 259-265

163. Hooper, A. B., Hansen, J. and Bell, R. (1967) J. Biol Chem. **242**, 288-296

164. Klein, A. R., Breitung, J., Linder, D., Stetter, K. O. and Thauer, R. K. (1993) Arch. Microbiol. **159**, 213-219

165. Koch, R., Zablowski, P., Spreinat, A. and Antranikian, G. (1990) FEMS Microbiol. Lett. **71**, 21-26

166. Laderman, K. A., Davis, B. R., Krutzsch, H. C., Lewis, M. S., Griko, Y. V., Privalov, P. L., and Anfinsen, C. B. (1993) J. Biol. Chem. **268**, 24394-24401

167. Laderman, K.A., Asada, K., Uemori, T., Mukai, H., Taguchi, Y., Kato, I. and Anfinsen, C.B. (1993) J. Biol. Chem. **268**, 24402-24407

168. Koch, R., Spreinat, A., Lemke, K. and Antranikian, G. (1991) Arch. Microbiol. **155**, 572-578

169. Chung, Y. C., Kobayashi, T., Kanai, H., Akiba, T. and Kudo, T. (1995) Appl. Environ. Microbiol. **61**, 1502-1506

170. Schumann, J., Wrba, A., Jaenicke, R., and Stetter, K. O. (1991) FEBS Lett. **282**, 122-126

171. Brown, S.H., and Kelly, R.M. (1993) Appl. Environ. Microbiol. **59**, 2614-2621

172. Schuliger, J. W., Brown, S. H., Baross, J. A. and Kelly, R. M. (1993) Mol. Marine Biol. Biotechnol. **2**, 76-87

173. Phipps, B. M., Hoffman, A., Stetter, K. O. and Baumeister, W. (1991) EMBO J. **10**, 1711-1722

174. Heyne, R. I. R., de Vrij, W., Crielaard, W. and Konings, W. N. (1991) J. Bacteriol. **173**, 791-800

175. Villa, A., Zecca, L., Fusi, P., Columbo, S., Tedeschi, G. and Tortora, P. (1993) Biochem. J. **295**, 827-831

176. Colombo, S., D'Auria, S., Fusi, P., Zecca, L., Raia, C. A. and Tortora, P. (1992) Eur. J. Biochem. **206**, 349-357

177. Bronnenmeier, K., Kern, A., Liebl, W., and Staudenbauer, W. L. (1995) Appl. Environ. Microbiol. **61**, 1399-1407

178. Simpson, H. D., Haufler, U. R. and Daniel, R. M. (1991) Biochem. J. **277**, 413-417

179. Ruttersmith, L. D., and Daniel, R. M. (1991) Biochem. J. **277**, 887-890

180. Ruttersmith, L. D., and Daniel, R. M. (1993) Biochim. Biophys. Acta **1156**, 167-172

181. Costantino, H. R., Brown, S. H., and Kelly, R. M. (1990) J. Bacteriol. **172**, 3654-3660

182. Bowen, D., Littlechild, J. A., Fothergill, J. E., Watson, H. C. and Hall, L. (1988) Biochem. J. **254**, 509-517

183

18

18

18

18

18

18

190

191

192

193

194

195

196

197

198

199

200

201

202

183. Breitung, J., Schmitz, R. A., Stetter, K. O. and Thauer, R. K. (1991) Arch. Microbiol. **156**, 517-524

184. Eggen, R., Geerling, A., Watts, J. and deVos, W. M. (1990) FEMS Microbiol. Lett. **71**,17-20

185. Morikawa, M., Izawa, Y., Rashid, N., Hoaki, T. and Imanaka, T. (1994) Appl. Environ. Microbiol. **60**, 4559-4566

186. Ruttersmith, L. D., Daniel, R. M. and Simpson, H. D. (1992) Ann. N.Y. Acad. Sci. **672**, 137-141

187. Slesarev, A. I., Lake, J. A., Stetter, K. O., Gellert, M. and Kozyavkin, S. A. (1994) J. Biol. Chem. **269**, 3295-3303

188. Robb, F. T., Masuchi, Y., Park, J.-B. and Adams, M. W. W. (1992) in Biocatalysis at extreme temperatures, (Adams, M. W. W. and Kelly, R.M., eds.), pp. 74-85. American Chemical Society, Washington, DC

189. Tiboni, O., Cammarano, P. and Sanangelantoni, A. M. (1993) J. Bacteriol. **175**, 2961-2969

190. Vieille, C., Hess, J. M., Kelly, R. M. and Zeikus, J.G. (1995) Appl. Environ. Microbiol. **61**, 1867-1875

191. Brown, S. H., Sjøholm, C. and Kelly, R. M. (1993) Biotechnol. Bioeng. **41**, 878-886

192. Menéndez-Arias, L. and Argos, P. (1989) J. Mol. Biol. **206**, 397-406

193. Watanabe, K., Chishiro, K., Kitamura, K. and Suzuki, Y. (1991) J. Biol. Chem. **266**, 24287-24294

194. Lovell, C.R., Przybyla, A. and Ljungdahl, L. G. (1990) Biochemistry **29**, 5687-5694

195. Böhm, G. and Jaenicke, R. (1994) Int. J. Peptide Protein Res. **43**, 97-106

196. Burdette, D. S., Vieille, C. and Zeikus, J. G. (1996) Biochem. J. **313** (in press)

197. Tomazic, S. J. and Klibanov, A. M. (1988) J. Biol. Chem. **263**, 3086-3091

198. Fontana, A. (1990) in Life under extreme conditions: Biochemical adaptation, (di Prisco, G., ed.), pp. 89-113. Springer-Verlag, Heidelberg

199. Meldgaard, M. and Svendsen, I. (1994) Microbiol. **140**, 159-166

200. Olsen, O. and Thomsen, K. K. (1991) J. Gen. Microbiol. **137**, 579-585

201. Mirsky, A. E. and Pauling, L. (1936) Proc. Natl. Acad. Sci. USA **22**, 439-447

202. Pfeil, W. (1986) in Thermodynamic data for biochemistry and biotechnology, (Hinz, H.-J., ed.), pp. 349-376. Springer-Verlag, Heidelberg

203. Jaenicke, R. (1991) Eur. J. Biochem. **202**, 715-728

204. Ahern, T. J. and Klibanov, A. M. (1988) Meth. Biochem. Anal. **33**, 91-127

205. Steigerwald, V. J., Beckler, G. S., and Reeve, J. N. (1990) J. Bacteriol. **172**, 4715-4718

206. Tchernajenko, V., Vieille, C., Burdette, D. S. and Zeikus, J. G. (1996) manuscript in preparation

207. Londesborough, J. (1980) Eur. J. Biochem. **105**, 211-215

208. Segel, I. H. (1975) Enzyme Kinetics. John Wiley & Sons, NY

209. Smith, C. A., Rangarajan, M. and Hartley, B. S. (1991) Biochem. J. **277**, 255-261

210. Sakamoto, N., Kotre, A. M. and Savageau, M. A. (1975) J. Bacteriol. **124**, 775-783

211. Coulton, J. W. and Kapoor, M. (1973) Can. J. Microbiol. **19**, 439-450

212. Hooper, A. B., Hansen, J. and Bell, R. (1967) J. Biol. Chem. **242**, 288-296

213. LéJohn, H. B., Suzuki, I. and Wright, J. A. (1968) J. Biol. Chem. **243**, 118-128

214. Hudson, R. C., Ruttersmith, L. D. and Daniel, R. M. (1993) Biochim. Biophys. Acta **1202**, 244-250

215. Glaser, P., Presecan, E., Delepierre, M., Surewicz, W. K., Mantsch, H. H., Bârzu, O. and Gilles, A.-M. (1992) Biochemistry **31**, 3038-3043

216. Jencks, W. P. (1975) in Advances in Enzymology and related areas of molecular biology vol. 43, (Meister, A., ed.), pp. 219-410. John Wiley & Sons, NY

217. Rellos, P. and Scopes, R. K. (1994) Prot. Express. Purif. **5**, 270-277

218. Nagaro, T., Makiro, Y., Yamamoto, K., Urabe, I. and Okada , H. (1989) FEBS lett. **253**, 113-116

219. Kauzmann, W. (1959) Adv. Protein Chem. **14**, 1-47

220. Freire, E., Haynie, D. T. and Xie, D. (1993) Proteins: Struct. Funct. Genet. **17**, 111-123

221. Haynie, D. T. and Freire, E. (1993) Proteins: Struct. Funct. Genet. **16**, 115-140

222. Haynie, D. T. and Freire, E. (1994) Biopolymers **34**, 261-271

223. Privalov, P. L. and Gill, S. J. (1988) Adv. Protein Chem. **39**, 191-234

224. Dill, K.A. (1990) Biochemistry **29**, 7133-7155

83

225. Dill, K. A., Alonso, D. O. V. and Hutchinson, K. (1989) Biochemistry **28**, 5439-5449

226. Creighton, T. E. (1990) Biochem. J. **270**, 1-16

227. Creighton, T.E. (1993) Proteins: structures and molecular properties 2nd edition. W. H. Freeman and Company, NY

228. Tsuboi, M., Ohta, S. and Nakanishi, M. (1978) in Biochemistry of thermophily (Friedman, M., ed.), pp. 251-266. Academic Press, NY

229. Wütrich, K., Roder, H. and Wagner, G. (1980) in Protein folding, (Jaenicke, R., ed.), pp. 549-564. Elsevier/North Holland Biomedical Press, Amsterdam

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

231. Matthews, B.W. (1993) Ann. Rev. Biochem. **62**, 139-160

232. Zhang, X.-J., Baase, W. A. and Matthews, B. W. (1992) Protein Sci. **1**, 761-776

233. Herning, T., Yutani, K., Inaka, K., Kuroki, R., Matsushima, M. and Kikuchi, M. (1992) Biochemistry **31**, 7077-7085

234. Hurley, J. H., Baase, W. A. and Matthews, B. W. (1992) J. Mol. Biol. **224**, 1143-1159

235. Watanabe, K., Masuda, T., Ohashi, H., Mihara, H. and Suzuki, Y. (1994) Eur. J. Biochem. **226**, 277-283

236. Eriksson, A. E., Baase, W. A., Zhang, X.-J., Heinz, D. W., Blaber, M., Baldwin, E. P. and Matthews, B. W. (1992) Science **255**, 178-183

237. Biro, J., Fabry, S., Dietmaier, W., Bogedain, C. and Hensel, R. (1990) FEBS Lett. **275**, 130-134

238. Koizumi, J.-I., Zhang, M., Imanaka, T. and Aiba, S. (1990) Appl. Environ. Microbiol. **56**, 3612-3614

239. Matsumura, M., Yasumura, S. and Aiba, S. (1986) Nature **323**, 356-3358

240. Shih, P. and J. F. Kirsch (1995) Prot. Sci. **4**, 2050-2062

241. Shih, P. and J. F. Kirsch (1995) Prot. Sci. **4**, 2063-2072

242. Kuroki, R., Inaka, K., Taniyama, Y., Kidokoro, S.-I., Matsushima, M., Kikuchi, M. and Yutani, K. (1992) Biochemistry **31**, 8323-8328

243. Herning, T., Yutani, K., Taniyama, Y. and Kikuchi, M. (1991) Biochemistry **30**, 9882-9891

244. Bell, J. A., Becktel, W. J., Sauer, U., Baase, W. A. and Matthews, B. W. (1992) Biochemistry **31**, 3590-3596

245. Pjura, P. and Matthews, B. W. (1993) Protein Sci. 2, 2226-2232

246. Matsumura, M., Signor, G. and Matthews, B. W. (1989) Nature 342, 291-293

247. Eriksson, A. E., Baase, W. A., and Matthews, B. W. (1993) J. Mol. Biol. 229, 747-769

248. Matthews, B. W., Nicholson, H. and Becktel, W. (1987) Proc. Natl. Acad. Sci. USA 84, 6663-6667

249. Pjura, P., McIntosh, L. P., Wozniak, J. A. and Matthews, B. W. (1993) Proteins: Struct. Funct. Genet. 15, 401-412

250. Kuroki, R., Taniyama, Y., Seko, C., Nakamura, H., Kikuchi, M. and Ikehara, M. (1989) Proc. Natl. Acad. Sci. USA 86, 6903-6907

251. Blaber, M., Lindstrom, J. D., Gassner, N., Xu, J., Heinz, D. W. and Matthews, B. W. (1993) Biochemistry 32, 11363-11373

252. Anderson, D. E., Hurley, J. H., Nicholson, H., Baase, W. A. and Matthews, B. W. (1993) Protein Sci. 2, 1285-1290

253. Eriksson, A. E., Baase, W. A., Wozniak, J. A., and Matthews, B. W. (1992) Nature 355, 371-373

254. Hardy, F., Vriend, G., Veltman, O. R., van der Vinne, B., Venema, G. and Eijsink, V. G. H. (1993) FEBS Lett. 317, 89-92

255. Vriend, G., Berendsen, H. J. C., van der Zee, J. R., van den Burg, B., Venema, G. and Eijsink, V. G. H. (1991) Protein Eng. 4, 941-945

256. Van den Burg, B., Dijkstra, B. W., Vriend, G., van der Vinne, B., Venema, G. and Eijsink, V. G. H. (1994) Eur. J. Biochem. 220, 981-985

257. Ishikawa, K., Nakamura, H., Morikawa, K. and Kanaya, S. (1993) Biochemistry 32, 6171-6178

258. Kimura, S., Nakamura, H., Hashimoto, T., Oobatake, M. and Kanaya, S. (1992) J. Biol. Chem. 267, 21535-21542

259. Kimura, S., Kanaya, S. and Nakamura, H. (1992) J. Biol. Chem. 267, 22014-22017

260. Fabian, H., Schultz, C., Backmann, J., Hahn, U., Saenger, W., Mantsch, H. H. and Naumann, D. (1994) Biochemistry 33, 10725-10730

261. Mitchinson, C. and Wells, J. A. (1989) Biochemistry 28, 4807-4815

262. Wells, J. A. and Powers, D. B. (1986) J. Biol. Chem. 261, 6564-6570

263. Takagi, H., Takahashi, T., Momose, H., Inouye, M., Maeda, Y., Matsuzawa, H. and Ohta, T. (1990) J. Biol. Chem. 265, 6874-6878

264. Quax, W. J., Mrabet, N. T., Luiten, R. G. M., Schuurhuizen, P. W., Stanssens, P. and Lasters, I. (1991) Bio/Technology 9, 738-742

265. Mrabet, N. T., Van den Broeck, A., Van den brande, I., Stanssens, P., Laroche, Y., Lambeir, A.-M., Matthijssens, G., Jenkins, J., Chiadmi, M., van Tilbeurgh, H., Rey, F., Janin, J., Quax, W. J., Lasters, I., De Maeyer, M. and Wodak, S. J. (1992) Biochemistry 31, 2239-2253

266. Meng, M., Bagdasarian, M. and Zeikus, J. G. (1993) Bio/Technology 11, 1157-1161

267. Volkin, D. B. and Klibanov, A. M. (1987) J. Biol. Chem. 262, 2945-2950

268. Alber, T., Dao-pin, S., Nye, J. A., Muchmore, D. C. and Matthews, B. W. (1987) Biochemistry 26, 3754-3758

269. MacArthur, M. W. and Thornton, J. M. (1991) J. Mol. Biol. 218, 3397-3412

270. Dixon, M. M., Nicholson, H., Shewchuk, L., Baase, W. A. and Matthews, B. W. (1992) J. Mol. Biol. 227, 917-933

271. Perutz, M. F., Kendrew, J. C. and Watson, H. C. (1965) J. Mol. Biol. 13, 669-678

272. El Hawrani, A. S., Moreton, K. M., Sessions, R. B., Clarke, A. R. and Holbrook, J. J. (1994) Trends Biotechnol. 12, 207-211

273. Creighton, T.E. (1993) Proteins: structures and molecular properties 2nd ed., W.H. Freeman and Company, NY

274. Hensel, R., Fabry, S., Biro, J., Bogedain, C., Jakob, I. and Siebers, B. (1994) Biocatalysis 11, 151-164

275. Ishikawa, K., Okumura, M., Katayanagi, K., Kimura, S., Kanaya, S., Nakamura, H. and Morikawa, K. (1993) J. Mol. Biol. 230, 529-542

276. Davies, G. J., Gamblin, S. J., Littlechild, J. A. and Watson, H. C. (1993) Proteins: Struct. Funct. Genet. 15, 283-289

277. Machius, M., Weingand, G. and Huber, R. (1995) J. Mol. Biol. 246, 545-559

278. Baker, E. N., and Hubbard, R. E. (1984) Prog. Biophys. Mol. Biol. 44, 97-179

279. Cleland, W. W. and Kreevoy, M. M. (1994) Science 264, 1887-1890

280. Hibbert, F. and Emsley, J. (1990) in Advances in Physical Organic Chemistry, (Bethell, D., ed.), pp. 255-379. Academic Press Ltd., London

281. Matsumura, M., Becktel, W. J. and Matthews, B. W. (1988) Nature 334, 406-410

282. Yutani, K., Ogasahara, K., Tsujita, T. and Sugino, Y. (1987) Proc. Natl. Acad. Sci. USA 84, 4441-4444

283. Rashin, A. A., Iofin, M. and Honig, B. (1986) Biochemistry **25**, 3619-3625

284. McRee, D. E., Redford, S. M., Getzoff, E. D., Lepock, J. R., Hallewell, R. A. and Tainer, J. A. (1990) J. Biol. Chem. **265**, 14234-14241

285. Wilson, K. P., Malcolm, B. A. and Matthews, B. W. (1992) J. Biol. Chem. **267**, 10842-10849

286. Yamada, H., Kanaya, E., Inaka, K., Ueno, Y., Ikehara, M. and Kikuchi, M. (1994) Biol. Pharm. Bull. **17**, 192-196

287. Kanaya, S. and Itaya, M. (1992.) J. Biol. Chem. **267**, 10184-10192

288. Sandberg, W.S. and Terwilliger, T. C. (1989) Science **245**, 54-57

289. Baldwin, E. P., Hajiseyedjavadi, O., Baase, W. A., and Matthews, B. W. (1993) Science **262**, 1715-1718

290. Kwon, K.-S., Kim, J., Shin, H. S. and Yu, M.-H. (1994) J. Biol. Chem. **269**, 9627-9631

291. Heinz, D. W., Baase, W. A. and Matthews, B. W. (1992) Proc. Natl. Acad. Sci. USA **89**, 3751-3755

292. Sicard, P. J., Leleu, J.-B., Duflot, P., Drocourt, D., Martin, F., Tiraby, G., Petsko, G. and Glasfeld, A. (1990) Ann. N.Y. Acad. Sci. **613**, 371-375

293. Hallewell, R. A., Imlay, K. C., Lee, P., Fong, N. M., Gallegos, C., Getzoff, E. D., Tainer, J. A., Cabelli, D. E., Tekamp-Olson, P., Mullenbach, G. T. and Cousens, L. S. (1991) Biochem. Biophys. Res. Comm. **181**, 474-480

294. Lepock, J. R., Frey, H. E. and Hallewell, R. A. (1990) J. Biol. Chem. **265**, 21612-21618

295. Zale, S. E. and Klibanov, A. M. (1986) Biochemistry **25**, 5432-5444

296. Volkin, D. B. and Middaugh, C. R. (1992) in Stability of protein pharmaceuticals, Part A: Chemical and physical pathways of protein degradation, (Ahern, T.J. and Manning, M.C., eds.), pp. 215-247. Plenum Press, NY

297. Tomazic, S. J. and Klibanov, A. M. (1988) J. Biol. Chem. **263**, 3092-3096

298. Kuroki, R., Kawakita, S., Nakamura, H. and Yutani, K. (1992) Proc. Natl. Acad. Sci. USA **89**, 6803-6807

298a. Gerwig, G. J., Kamerling, J. P., Vliegenthart, J. F. G., Morag, E., Lamed, R. and Bayer, E. A. (1991) Eur. J. Biochem. **196**, 115-122

299. Lee, Y.-E., Lowe, S. E. and Zeikus, J. G. (1993) Appl. Environ. Microbiol. **59**, 763-771

300. Ng, T. K. and Zeikus, J. G. (1981) Biochem. J. **199**, 341-35

301. Chen, B.-l and King, J. (1991) in Protein Refolding, (Georgiou, G. and De Bernardez-Clark, E., eds.), pp. 119-132. American Chemical Society, Washington, DC

302. Boel, E., Brady, L., Brzozowksi, A. M., Derewenda, Z., Dodson, G. G., Jensen, V. J., Petersen, S. B., Swift, H., Thim, L. and Woldike, H. F. (1990) Biochemistry **29**, 6244-6249

303. Marg, G. A. and Clark, D. S. (1990) Enzyme Microbiol. Technol. **12**, 367-373

304. Whitlow, M., Howard, A. J., Finzel, B. C., Poulos, T. L., Winborne, E. and Gillialand, G. L. (1991) Proteins Struct. Funct. Genet. **9**, 153-173

305. Kasumi, T., Hayashai, K. and Tsumura, N. (1982) Agric. Biol. Chem. **46**, 21-30

306. Stigter, D., Alonso, D. O. V., and Dill, K. A. (1991) Proc. Natl. Acad. Sci. USA **88**, 4176-4180

308. Hei, D. J. and Clark, D. S. (1994) Appl. Environ. Microbiol. **60**, 932-939

309. Miller, J. F., Nelson, C. M., Ludlow, J. M., Shah, N. N. and Clark, D. S. (1989) Biotechnol. Bioeng. **34**, 1015-1021

310. Lüthi, E., Reif, K., Jasmat, N. B. and Bergquist, P. L. (1992) Appl. Microbiol. Biotechnol. **36**, 503-506

311. Wilks, H. M., Hart, K. W., Feeney, R., Dunn, C. R., Muirhead, H., Chia, W. N., Barstow, D. A., Atkinson, T., Clarke, A. R. and Holbrook, J. J. (1988) Science **242**, 1541-1544

312. Wilks, H. M., Halsall, D. J., Atkinson, T., Chia, W. N., Clarke, A. R. and Holbrook, J. J. (1990) Biochemistry **29**, 8587-8591

313. Meng, M., Lee, C., Bagdasarian, M. and Zeikus, J. G. (1991) Proc. Natl. Acad. Sci. USA **88**, 4015-4019

313a. Arnold, F.H. (1993) FASEB J. **7**, 744-749

313b. Chen, K. and Arnold, F. H. (1991) Bio/Technology **9**, 1073-1077

306. De Cordt, S., Hendrickx, M., Maesmans, G. and Tobback, P. (1994) Biotech. Bioeng. **43**, 107-114

307. Oshima, T. (1986) in Thermophiles: General, molecular, and applied microbiology (Brock, T. D., ed.) pp. 137-158. John Wiley & Sons, NY

314. Liao, H., McKenzie, T. and Hageman, R. (1986) Proc. Natl. Acad. Sci. USA **83**, 576-580

315. Bai, Y., Milne, J. S., Mayne, L. and Englander, S. W. (1994) Proteins: Struct. Funct. Genet. **20**, 4-14

316. Muheim, A., Todd, R. J., Casimiro, D. R., Gray, H. B. and Arnold, F.H. (1993) J. Am. Chem. Soc. **115**, 5312-5313

317. Woese, C. R. (1994) Microbiol. Rev. **58**, 1-9

318. Keinan, E., Hafeli, E. K., Seth, K. K. and Lamed, R. L. (1986) J. Am. Chem. Soc. **108**, 162-169

319. Pham, V. T., Phillips, R. S. and Ljungdahl, L. G. (1989) J. Amer. Chem Soc. **111**, 1935-1936

320. Periera, D. A., Gerson, F. P. and Oestreicher, E. G. (1994) J. Biotechnol. **34**, 43-50

321. Keinan, E., Seth, K. K. and Lamed, R. (1986) J. Am. Chem. Soc. **108**, 3474-3480

322. Keinan, E., Seth, K. K., Lamed, R., Ghirlando, R. and Singh, S. P. (1990) Biocatalysis **4**, 1-15

323. Lamed, R. J., Keinan, E. and Zeikus, J. G. (1981) Enz. Microbiol. Technol. **3**, 144-148

324. Aaltonen, O. and Rantakylä, M. (1991) Chemtech **1**, 240-248

325. Lee, K. M. and Biellmann, J.-F. (1987) Nuov. J. Chim. **11**, 1-4

Chapter II

Cloning and Expression of the Gene Encoding the *Thermoanaerobacter ethanolicus* 39E Secondary-Alcohol Dehydrogenase and Enzyme Biochemical Characterization

90

ABSTRACT

The *adhB* gene encoding *Thermoanaerobacter ethanolicus* 39E secondary-alcohol

dehydrogenase (2° ADH) was cloned, sequenced, and expressed in *Escherichia coli*. The

1056 bp gene encodes a homotetrameric recombinant enzyme consisting of 37.7 kDa

subunits. The purified recombinant enzyme is optimally active above 90°C with a half-life

of approximately 1.7 h at 90°C. An NADP(H)-dependent enzyme, the recombinant 2°

ADH has 170-fold greater catalytic efficiency in propan-2-ol versus ethanol oxidation. The

enzyme was inactivated by chemical modification using dithionitrobenzoate (DTNB) and

diethylpyrocarbonate, indicating that Cys and His residues are involved in catalysis. Zinc

was the only metal enhancing 2° ADH reactivation after DTNB modification, implicating

the involvement of a bound zinc in catalysis. Arrhenius plots for the oxidation of propan-2-

ol by the native and recombinant 2° ADHs were linear from 25°C to 90°C when the

enzymes were incubated at 55°C prior to assay. Discontinuities in the Arrhenius plots for

propan-2-ol and ethanol oxidations were observed, however, when the enzymes were

preincubated at 0°C or 25°C. The observed Arrhenius discontinuity, therefore, resulted

from a temperature dependent, catalytically significant 2° ADH structural change.

Hydrophobic Cluster Analysis comparisons of both mesophilic versus thermophilic 2°

ADH and 1° versus 2° ADH amino acid sequences were performed. These comparisons

predicted that specific proline residues may contribute to 2° ADH thermostability and

thermophilicity, and that the catalytic Zn ligands are different in 1° ADHs (two Cys and a

His) and 2° ADHs (Cys, His, and Asp).

INTRODUCTION

Alcohol dehydrogenases (ADHs) (EC 1.1.1.1 [NADH] or EC 1.1.1.2 [NADPH]) are well studied as a structurally conserved class of enzyme [1]. The X-ray structure of the horse liver primary-alcohol dehydrogenase (1° ADH) is known, and the properties of this enzyme have been extensively detailed [1,2]. ADHs are typically dimeric or tetrameric pyridine dinucleotide-dependent metalloenzymes with a zinc atom involved in catalysis. ADHs have been classified as 1° or 2° ADHs based on their relative activities toward 1° and 2° alcohols. It generally has been assumed that 1° and 2° ADHs are structurally similar, and that their substrate differences are due to relatively small changes in their active site architecture. Tetrameric 2° ADHs have been reported from a number of microorganisms [3-7]. The *Thermoanaerobacter ethanolicus* 39E 2° ADH is a bifunctional ADH/acetylCoA reductive thioesterase [7]. It has been proposed to function physiologically by oxidizing nicotinamide cofactor during ethanol formation, indirectly preventing glycolytic inhibition at the glyceraldehyde dehydrogenase step [8].

2° ADHs are attractive subjects for research into chiral chemical production because of their broad specificities and their highly enantiospecific conversion of prochiral ketones to alcohols [9-12]. Two issues facing the commercial scale application of 2° ADHs in chiral syntheses are the difficulty of regenerating and retaining expensive nicotinamide cofactor and the lack of inexpensive, highly stable enzymes. Cofactor regeneration and retention have been overcome by numerous strategies [13,14]. While thermophilic enzymes are generally much more stable than their mesophilic counterparts, the organisms that produce thermophilic 2° ADHs grow slowly and to low cell densities, making an alternative expression system for these enzymes crucial for both their commercial application and detailed protein structure-function studies.

The discovery of thermophilic bacteria has provided the opportunity to isolate thermostable enzymes directly. The thermophilic anaerobes *T. ethanolicus* and

Thermoanaerobacter brockii [15] express extremely enantiospecific 2° ADHs that are stable
above 70°C. These two 2° ADHs have been proposed to be extremely structurally similar
based on similar molecular weight and kinetic characteristics [16]. While the *Clostridium*
beijerinckii gene (*adhB*) encoding a mesophilic NADP(H)-dependent 2° ADH has been
cloned and sequenced (Genbank, Acc. no. M84723), and while the amino acid sequence of
the *T. brockii* 2° ADH has been determined by Edman degradation [17], no cloned
thermophilic 2° ADH is available for detailed biochemical studies.

The work reported here describes the cloning, sequencing, and expression of the
gene encoding the thermophilic *T. ethanolicus* 2° ADH. The kinetic and thermal properties
of the recombinant enzyme are determined. The biochemical basis for enzyme
discontinuous Arrhenius plots is investigated. Finally, the protein sequence information is
used to examine the structural similarities between thermophilic and mesophilic 1° ADHs
and 2° ADHs. These comparisons are used to predict the 2° ADH catalytic Zn liganding
residues and the potential involvement of specific prolines in 2° ADH thermal stability.

MATERIALS AND METHODS

Chemicals and reagents

All chemicals were of at least reagent/molecular biology grade. Gases were provided by
AGA specialty gases (Cleveland, OH), and made anaerobic by passage through heated
copper filings. Oligonucleotide syntheses and amino acid sequence analyses were
performed by the Macromolecular Structure Facility (Department of Biochemistry,
Michigan State University). The kanamycin resistance GenBlock (*EcoRI*) DNA cartridge
used in expression vector construction was purchased from Pharmacia (Uppsala, Sweden)
[18].

Media and strains

T. ethanolicus 39E (ATCC #33223) was grown in TYE medium as previously described [19]. All batch cultures were grown anaerobically under an N_2 headspace. *Escherichia coli* DH5α containing the recombinant *adhB* gene was grown in rich complex medium (20 g l^{-1} tryptone, 10 g l^{-1} yeast extract, 5 g l^{-1} NaCl) at 37°C in the presence of 25 μg ml^{-1} kanamycin and 100 μg ml^{-1} ampicillin.

DNA manipulations and library construction

Plasmid DNA purification, restriction analysis, PCR, and colony and DNA hybridizations were performed using conventional techniques [20,21]. *T. ethanolicus* chromosomal DNA was purified as previously described [22]. Partially digested 2–5 kb *Sau*3A I fragments were isolated by size fractionation from a 10–40% sucrose gradient [20] and ligated into pUC18 *Bam*H I/BAP (Pharmacia, Uppsala, Sweden). The ligation mixture was transformed into *E. coli* DH5α by electroporation [21]. Degenerate primers (1) (5'-ATGAA(R)GG(N)TT(H)GC(N) ATG(Y)T) and (2) (5'–G(W)(N)GTCATCAT(R)TC(N)G(K)(D)ATCAT) were used to synthesize the homologous probe for colony hybridization [23]. The DNA fragment containing the *adhB* gene was sequenced using the method of Sanger *et al.* [24].

Enzyme purification

The native 2° ADH was purified by established techniques [7]. The recombinant enzyme was purified from *E. coli* DH5a aerobically. The pelleted cells from batch cultures were resuspended (0.5 g wet wt. ml^{-1}) in 50 mM Tris:HCl [pH 8.0] (buffer A) containing 5 mM DTT, and 10 mM $ZnCl_2$, and lysed by passage through a French pressure cell. The clarified lysate was incubated at 65°C for 25 min and then centrifuged for 30 min at 15000g. The supernatant was applied to a DEAE-sephacryl column (2.5 cm x 15 cm) that was equilibrated with buffer A and eluted using a 250 ml NaCl gradient (0-300 mM).

Active fractions were diluted 4-fold in buffer A and applied to a Q-sepharose column (2.5 cm x 10 cm) equilibrated with buffer A. Purified enzyme was eluted using a 250 ml NaCl gradient (0-300 mM).

Molecular mass determination

Recombinant holoenzyme molecular mass was determined by comparison with protein standards (Sigma; St. Louis, MO) using gel filtration chromatography (0.5 ml min^{-1}) with a Pharmacia S300 column (110 cm x 1.2 cm) equilibrated with buffer A containing 200 mM NaCl. Subunit molecular mass values were determined by matrix associated laser desorption ionization mass spectrometry (MALDI) at the Michigan State University Mass Spectrometry Facility.

Kinetics and thermal stability

The standard 2° ADH activity assay was defined as NADP$^+$ reduction coupled to propan-2-ol oxidation at 60°C as previously described [7]. The enzyme was incubated at 55°C for 15 min prior to activity determination unless otherwise indicated. Assays to determine Km_{app} and $Vmax_{app}$ were conducted at 60°C with substrate concentrations between 20xKm_{app} and 0.2xKm_{app}. Kinetic parameters were calculated from nonlinear best fits of the data to the Michaelis-Menten equation using Kinzyme software [26]. Protein concentrations were measured using the bicinchoninic acid procedure (Pierce; Rockford, IL) [27].

2° ADH thermostability was measured as the residual activity after timed incubation at the desired temperatures. Thermal inactivation was stopped by incubation for 30 min at 25°C, and samples were prepared for activity assays by preincubation at 55°C for 15 min. Incubations were performed in 100 µl PCR tubes (cat. #72.733.050, Sarstedt; Newton, NC) using 200 µg ml^{-1} protein in 100 µl buffer A. Activity was determined using the unfractionated samples. The temperature effect on enzyme activity was studied using the substrates propan-2-ol or ethanol at 10xKm_{app} concentrations.

To study the effect of enzyme preincubation temperature on enzyme reaction rates, the 2° ADH was incubated at 0°C, 25°C, or 55°C for 15 min prior to activity determinations in the temperature range 30–90°C. Tris:HCl buffer pH values were adjusted at 25°C to be pH 8.0 at the temperature they were used (thermal correction factor = -0.031 ΔpH °C^{-1}). The statistical significance of the differences between the Arrhenius plot slopes above and below the discontinuity temperatures was determined by covariance analysis [25].

Chemical modification

Cysteine residues were reversibly modified using dithionitrobenzoate (DTNB) at 25°C and 60°C [28]. DTNB-inactivated 2° ADH was reactivated using DTT in the presence of 0.01 mM to 1.0 mM metal salts or 0.5 mM to 3.0 mM EDTA. Histidine residues were chemically modified with diethylpyrocarbonate (DEPC) by incubation with 20 mM or 40 mM DEPC in 50 mM phosphate buffer (pH 6.0) at 25°C for 1.0 h and the reaction was quenched by addition of 0.5 x volume of 0.5 M imidazole (pH 6.5) [29].

Protein sequence comparisons

The peptide sequences of the *Bacillus stearothermophilus* (acc. #D90421), *Sulfolobus solfataricus* (acc. #S51211), and *Zymomonas mobilis* (acc. #M32100) 1° ADHs and of the *Alcaligenes eutrophus* (acc. #J03362) and *Clostridium beijerinckii* (acc. #M84723) 2° ADHs were obtained from GenBank. The horse liver 1° ADH [1] and the *T. brockii* 2° ADH [17] peptide sequences were obtained from the literature. Access to GenBank, standard sequence alignments, and percentages of amino acid similarity/identity were performed using the Program Manual for the Wisconsin Package, Version 8, Sept. 1994 (Genetics Computer Group ; Madison, WI). Protein sequence alignments were performed using Hydrophobic cluster analysis (HCA) [30]. HCA plots of individual protein sequences were generated using HCA-Plot V2 computer software (Doriane; Le Chesnay, France).

Nucleotide sequence accession number

The GenBank accession number for the sequence published in this article is U49975.

RESULTS

Cloning and sequencing of the *T. ethanolicus adhB* gene

The *T. ethanolicus* 2° ADH was cloned from a *T. ethanolicus* chromosomal DNA library by homologous hybridization. The N-terminal sequence of the native *T. ethanolicus* 2° ADH (MKGFAML) was identical to those of *C. beijerinckii* and *T. brockii* 2° ADHs. This sequence was used to generate primer (1). Alignment of the *C. beijerinckii* and *T. brockii* 2° ADH peptide sequences indicated another conserved region (residues 147–153) that would reverse translate into a low-degeneracy oligonucleotide (primer [2]). The PCR product obtained using primers (1) and (2), and using *T. ethanolicus* chromosomal DNA as the template was 470 bp, as expected from the position of primers (1) and (2) in the *C. Beijerinckii* and *T. brockii* 2° ADHs. This PCR product was used as a homologous probe to screen the *T. ethanolicus* genomic library. The positive clone showing the highest 2° ADH activity at 60°C was selected for further studies. Plasmid pADHB25-C contained a 1.6 kb *Sau*3A I insert and was shown by subsequent sequencing and peptide analyses to carry the complete *adhB* gene. The 1.6 kb insert was subcloned into the pBluescriptIIKS(+) *Xba* I site to construct the expression plasmid pADHB25. The physical map of pADHB25 is shown in Figure 1. Plasmid pADHB25 was stabilized by insertion of a kanamycin resistance cartridge into the vector *Eco*R I site, allowing dual selection on kanamycin and ampicillin. This final construct (pADHB25-kan) was used for all subsequent work.

The nucleotide sequence of the pADHB25 insert is shown in Figure 2. A unique open reading frame (ORF) was identified that encoded a polypeptide which was highly homologous to *C. beijerinckii* 2° ADH, and which started with the N-terminal sequence of the native *T. ethanolicus* 2° ADH. Two consecutive ATG codons were identified as potential translation initiation codons. The N-terminus of the native 2° ADH starts with a single Met, suggesting that the first ATG codon is not translated or is post-translationally removed. A potential ribosome binding site (RBS) (positions 223-228) is located 10 bp upstream of the start codon. A potential promoter was identified approximately 70–100 bp upstream of the RBS. The "-35" and "-10" regions are highly similar to the *E. coli* consensus promoter sequences [31], and are separated by 16 bp. The "-10" region is duplicated with the second copy overlapping the first. Because of its improper distance from the "-35" region, this second copy may only provide an A+T rich sequence to aid in strand separation. No transcriptional stop site was identified downstream of *adhB*. In this region, instead, a truncated ORF, preceded by an RBS, was identified. It translated into a 45 amino acid peptide fragment 46% identical and 68% similar to the product of a similar truncated ORF located downstream of *C. beijerinckii adhB* gene. Little similarity was found with other sequences in GenBank, so the function of this ORF remains unknown.

Sequence comparison of 1° and 2° ADHs

Standard alignments of 1° and 2° ADH amino acid sequences indicated a high level of similarity among the three 2° ADHs from obligate anaerobes (Table 1). The *T. ethanolicus* 2° ADH differed from the *T. brockii* enzyme by only three residues, and was 75% identical to the enzyme from the mesophile *C. beijerinckii*. The similarity was lower for comparisons with the 2° ADH from the obligate aerobe *A. eutrophus*. The 1° ADHs showed less sequence conservation (25 to 54% identity and 49 to 71% similarity) than the 2° ADHs, and showed only 20 to 27% identity (and 48 to 51% similarity) to the 2° ADHs. Based on these standard alignments, the conservation of important core domain residues,

Figure 1. Restriction map of the *T. ethanolicus* 39E 2° ADH clone (pADHB25). The flanking restriction sites are from the plasmid polylinker.
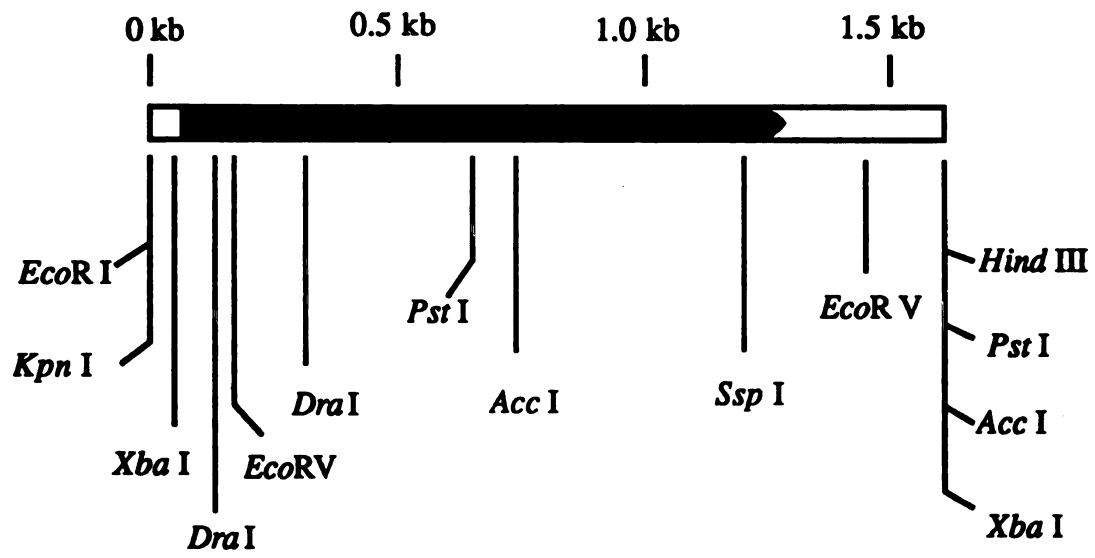
Figure 2. Nucleotide sequence and deduced amino acid sequence of *T.* *ethanolicus* 39E *adhB* and of the downstream open reading frame. The putative promoter -35 and -10 regions and RBSs are underlined. The *adhB* stop codon is indicated by three asterisks.

```
                                                                              90
TGAACAATAGACAACCCCTTTCTGTGATCTTGTTTTTTGCAAATGCTATTTTATCACAAGAGATTTCTCTAGTTCTTTTTTACTTAAAAA

                                                                             180
AACCCTACGAAATTTTAAACTATGTCGAATAAATTATTGATAATTTTTAACTATGTGCTATTATATTATTGCAAAAAATTTAACAATCAT
                               -35                      -10

                                                                             270
CGCGTAAGCTAGTTTTCACATTAATGACTTACCCAGTATTTTAGGAGGTGTTTAATGATGAAAGGTTTTGCAATGCTCAGTATCGGTAAA
                                            RBS           M  K  G  F  A  M  L  S  I  G  K

                                                                             360
GTTGGCTGGATTGAGAAGGAAAAGCCTGCTCCTGGCCCATTTGATGCTATTGTAAGACCTCTAGCTGTGGCCCCTTGCACTTCGGACATT
 V  G  W  I  E  K  E  K  P  A  P  G  P  F  D  A  I  V  R  P  L  A  V  A  P  C  T  S  D  I

                                                                             450
CATACCGTTTTTGAAGGAGCCATTGGCGAAAGAATAACATGATACTCGGTCACCGAAGCTGTAGGTGAAGTAGTTGAAGTAGGTAGTGAG
 H  T  V  F  E  G  A  I  G  E  R  H  N  M  I  L  G  H  E  A  V  G  E  V  V  E  V  G  S  E

                                                                             540
GTAAAAGATTTTAAACCTGGTGATCGCGTTGTTGTGCCAGCTATTACCCCTGATTGGTGGACCTCTGAAGTACAAAGAGGATATCACCAG
 V  K  D  F  K  P  G  D  R  V  V  V  P  A  I  T  P  D  W  W  T  S  E  V  Q  R  G  Y  H  Q

                                                                             630
CACTCCGGTGGAATGCTGGCAGGCTGGAAATTTTCGAATGTAAAAGATGGTGTTTTTGGTGAATTTTTTCATGTGAATGATGCTGATATG
 H  S  G  G  M  L  A  G  W  K  F  S  N  V  K  D  G  V  F  G  E  F  F  H  V  N  D  A  D  M

                                                                             720
AATTTAGCACATCTGCCTAAAGAAATTCCATTGGAAGCTGCAGTTATGATTCCCGATATGATGACCACTGGTTTTCACGGAGCTGAACTG
 N  L  A  H  L  P  K  E  I  P  L  E  A  A  V  M  I  P  D  M  M  T  T  G  F  H  G  A  E  L

                                                                             810
GCAGATATAGAATTAGGTGCGACGGTAGCAGTTTTGGGTATTGGCCCAGTAGGTCTTATGGCAGTCGCTGGTGCCAAATTGCGTGGAGCC
 A  D  I  E  L  G  A  T  V  A  V  L  G  I  G  P  V  G  L  M  A  V  A  G  A  K  L  R  G  A

                                                                             900
GGAAGAATTATTGCCGTAGGCAGTAGACCAGTTTGTGTAGATGCTGCAAAATACTATGGAGCTACTGATATTGTAAACTATAAAGATGGT
 G  R  I  I  A  V  G  S  R  P  V  C  V  D  A  A  K  Y  Y  G  A  T  D  I  V  N  Y  K  D  G

                                                                             990
CCTATCGAAAGTCAGATTATGAATCTAACTGAAGGCAAAGGTGTCGATGCTGCCATCATCGCTGGAGGAAATGATGACATTATGGCTACA
 P  I  E  S  Q  I  M  N  L  T  E  G  K  G  V  D  A  A  I  I  A  G  G  N  A  D  I  M  A  T

                                                                            1080
GCAGTTAAGATTGTTAAACCTGGTGGCACCATCGCTAATGTAAATTATTTTGGCGAAGGAGAGGTTTTGCCTGTTCCTCGTCTTGAATGG
 A  V  K  I  V  K  P  G  G  T  I  A  N  V  N  Y  F  G  E  G  E  V  L  P  V  P  R  L  E  W

                                                                            1170
GGTTGCGGCATGGCTCATAAAACTATAAAAGGCGGGCTATGTTCCGGTGGACGTCTAAGAATGGAAAGACTGATTGACCTTGTTTTTTAT
 G  C  G  M  A  H  K  T  I  K  G  G  L  C  P  G  G  R  L  R  M  E  R  L  I  D  L  V  F  Y

                                                                            1260
AAGCCTGTCGATCCTTCTAAGCTCGTCACTCACGTTTTCCAGGGATTTGACAATATTGAAAAAGCCTTTATGTTGATGAAAGACAAACCA
 K  P  V  D  P  S  K  L  V  T  H  V  F  Q  G  F  D  N  I  E  K  A  F  M  L  M  K  D  K  P

                                                                            1350
AAAGACCTTATCAAACCTGTTGTAATATTAGCATAAAAATGGGGACTTAGTCCATTTTTATGCTAATAAGGCTAAATACACTGGTTTTTT
 K  D  L  I  K  P  V  V  I  L  A  ***

                                                                            1440
TATATGACACATCGGCCAGTAAACTCTTGGTAAAAAAAATAACAAAAAATAGTTATTTTCTTAACATTTTTACGCCATTAACACTTGATAA

                                                                            1530
CATCATCGAAGAAGTAAATAAACAACTATTAAATAAAAGAAGAAGGAGGATTATCATGTTCAAAATTTTAGAAAAAAGAGAATTGGCACC
                                        RBS            M  F  K  I  L  E  K  R  E  L  A  P
                                                                            1620
TTCCATCAAGTTGTTTGTAATAGAGGCACCACTAGTAGCCAAAAAAGCAAGGCCAGGCCAATTCGTTATGCTAAGGATAAAAGAAGGAGG
  S  I  K  L  F  V  I  E  A  P  L  V  A  K  K  A  R  P  G  N  F  V  M  L  R  I  K  E  G  G

AGAAAGAATT  1630
 E  R  I
```

Table 1. Comparison of primary structural similarity between horse liver and bacterial alcohol dehydrogenases

Organism	% Identity (% similarity) of sequence with sequences from the following organisms:							
	T. eth	T. bro	C. bei	A. eut	B. stea	S. sul	Z. mob	Liver
2° Adhs								
*T. eth	100							
*T. bro	99.1 (99.4)	100						
C. bei	74.4 (86.6)	75.0 (86.9)	100					
A. eut	35.6 (61.5)	36.2 (62.1)	36.2 (59.5)	100				
1° Adhs								
*B. stea	26.7 (50.6)	27.2 (50.8)	24.1 (50.6)	31.9 (53.8)	100			
*S. sul	27.0 (51.6)	26.4 (51.0)	24.0 (50.6)	25.8 (49.0)	35.2 (55.8)	100		
Z. mob	24.2 (48.2)	25.1 (48.8)	20.5 (48.4)	25.6 (51.2)	54.0 (71.0)	33.9 (54.6)	100	
Liver	28.8 (53.8)	29.1 (54.4)	27.4 (52.6)	23.1 (49.7)	32.8 (55.4)	28.9 (50.9)	31.1 (50.4)	100

Percent identities and similarities were calculated by the GCG Gap program with penalties for gap insertion (3.0) and extension (0.1) T. eth: T. ethanolicus 39E; T. bro:T.brockii; C. bei: C.beijerinckii; A. eur: A. eutrophus; B. stea: B. stearothermophilus; S. sul: S. solfataricus; Z. mob: Z. mobilis; and liver: horse liver.
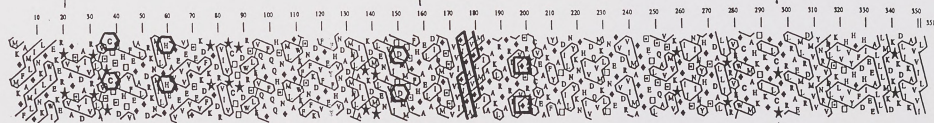*: Thermophilic Adh

Figure 3. HCA comparison of thermophilic and mesophilic 1° and 2° ADHs

centered on the putative nicotinamide binding motifs. The proposed catalytic Zn

liganding (O), structural Zn liganding (□), and Rossmann fold binding motif (□) residues

are indicated. Symbols:★: proline;◆: glycine;■: serine; □: threonine.
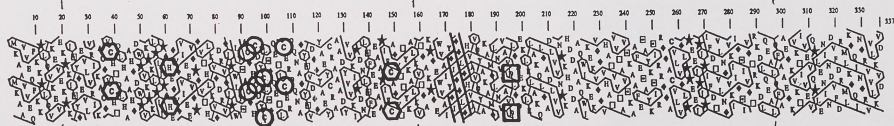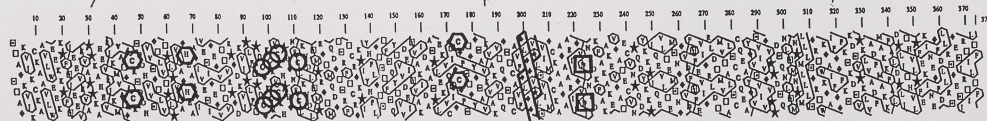
T. ethanolicus 39E 2° Adh
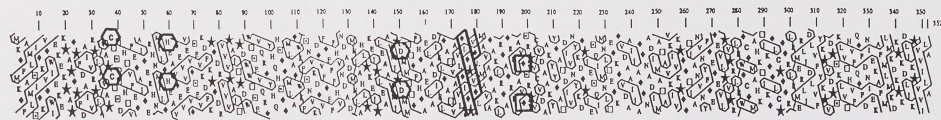
C. beijerinckii 2° Adh

Catalytic Zn binding   **Nicotinamide binding**

B. stearothermophilus 1° Adh

horse liver 1° Adh

amino acids lining the active site pocket, and structurally conserved glycines identified for the horse liver enzyme [1] was higher in all the peptides.

HCA comparisons allow the alignment of potentially similar 2° or 3° structural regions between enzymes with dissimilar amino acid sequences. The Rossmann fold consensus sequences {Gly-Xaa-Gly-Xaa-Xaa-Gly-(Xaa)$_{18-20}$[negatively charged amino acid for NAD(H) dependent or neutral amino acid for NAD(P)(H)]} [32] identified in each of the ADH peptides were used to direct HCA alignments of 1° and 2° ADHs. Figure 3 shows a representative HCA multiple comparison between a thermophilic (*T. ethanolicus*) and a mesophilic (*C. beijerinckii*) 2° ADH with a thermophilic (*B. stearothermophilus*) and a mesophilic (horse liver) 1° ADH based on the identified Rossmann fold consensus sequences. The consensus motifs of the NADP(H) linked *T. ethanolicus* and *C. beijerinckii* 2° ADHs were identified as Gly^{174} -Xaa-Gly^{176}-Xaa-Xaa-Gly^{179}, Gly^{198}, and those of the NAD(H) dependent *B. stearothermophilus* and horse liver 1° ADH were identified as Gly^{172} - Xaa - Gly^{174}-Xaa-Xaa-Gly^{177}, Asp^{195} and Gly^{199}-Xaa-Gly^{201}-Xaa-Xaa-Gly^{204}, Asp^{223}, respectively. This alignment also predicts 71 to 94% overall similarity between 1° and 2° ADHs for the critical hydrophobic core residues (71 to 94%), active site pocket residues (71 to 100%), and structurally critical glycines (100% in all cases) identified in the horse liver enzyme.

The HCA comparison allowed the identification of corresponding cluster regions and catalytically important residues in all four dehydrogenases based on the corresponding residues identified for the horse liver enzyme [1,2]. The horse liver 1° ADH contains two Zn atoms per subunit, one structural and one catalytic [1]. The thermophilic *B. stearothermophilus* 1° ADH contained a region involving cysteine residues 97, 100, 103, and 111 that is analogous to the structural Zn binding loop in the horse liver ADH (involving Cys^{92}, Cys^{95}, Cys^{98}, and Cys^{106}). However, no analogous structural Zn binding loop regions were identified in the *T. ethanolicus* and *C. beijerinckii* 2° ADHs (Fig. 3). The catalytic Zn ligands in the horse liver enzyme have been established (Cys^{46}, His^{68},

and Cys^{174}). The N-terminal cysteine and histidine ligands to the catalytic Zn atom appear to be conserved in both the 1° and 2° ADH sequences (Fig. 3). However, the second cysteine ligand in the horse liver 1° ADH (Cys^{174}), conserved in the *B. stearothermophilus* 1° ADH(Cys^{148}), is substituted with an aspartate residue (Asp^{150}) in both the *T. ethanolicus* and *C. beijerinckii* proteins. The *A. eutrophus* and *T. brockii* 2° ADHs also conserved corresponding aspartate residues and appeared to lack 1° ADH-like structural Zn binding loops (data not shown).

Sequence comparison of 2° ADH enzymes from a mesophile and thermophiles

The structural constraints introduced by proline residues have been proposed as a mechanism involved in protein thermostabilization [33]. Among the twelve nonconservative sequence substitutions between the mesophilic *C. beijerinckii* 2° ADH and the thermophilic *T. ethanolicus* 2° ADH, nine correspond to the introduction of prolines (22, 24, 149, 177, 222, 275, 313, 316, and 347) in the *T. ethanolicus* protein. All but Pro^{313} are also present in the thermophilic *T. brockii* 2° ADH. The mesophilic 2° ADH subunit contained a total of 13 prolines (3.7%) while the *T. ethanolicus* and *T. brockii* subunits contained 22 prolines (6.2%) and 21 prolines (6.0%), respectively. Prolines 20, 22, and 24 are in a 9 residue stretch of hydrophilic amino acids near the putative catalytic Zn ligand Cys^{37}. Pro^{149} interrupts a hydrophobic cluster and is next to another putative Zn ligand, Asp^{150}. The nicotinamide cofactor binding motif includes Pro^{177} which also interrupts a hydrophobic cluster region. Prolines 222, 275, 313, 316, and 347 are located in short (two to four residue) hydrophilic stretches that form putative turn regions.

Purification and characterization of the recombinant 2° ADH

The recombinant *T. ethanolicus* 2° ADH was highly expressed in *E. coli* in the absence of induction and no significant increase was seen upon induction with 5.0 mM isopropylthio-

b-D-galactoside. The enzyme expression level was similar in *E. coli* and in the native organism (1% to 5% of total protein). The recombinant enzyme was purified 36-fold to homogeneity (as determined by the presence of a single band on SDS-PAGE). The subunit molecular masses were calculated from the average of 3 (native enzyme) and 5 (recombinant enzyme) determinations using MALDI, yielding masses for the native and recombinant enzyme subunits of 37707 Da and 37854 Da, respectively. These values are within the generally accepted error of the technique (~1%), and are in agreement with the theoretical molecular mass for the native enzyme based on the gene sequence (37 644 Da). N-terminal amino acid analysis of the recombinant enzyme indicated that in $72 \pm 5\%$ of the recombinant protein the two N-terminal ATG codons had been translated, while $28 \pm 2\%$ of the protein contained a single N-terminal Met residue, like the native enzyme. The increased mass of 147 Da determined for the recombinant enzyme is consistent with the mass of one Met residue (149.2 Da), although this difference is small compared to the measurement error associated with MALDI. The recombinant holoprotein molecular mass was determined to be 160 kDa by gel filtration chromatography demonstrating that the recombinant *T. ethanolicus* 2° ADH, as the native enzyme, is a homotetramer.

The 2° ADH was completely inactivated at both 25°C and 60°C by cysteine-specific DTNB modification. Inactivation was partially reversed at 60°C by the addition of DTT allowing the recovery of 34% initial activity (initial activity = 54 ± 3.0 units mg^{-1}). The addition of $CdSO_4$, $FeCl_2$, $MnCl_2$, $CaCl_2$, $MgCl_2$, NaCl (0.01 mM to 1.0 mM), or EDTA (0.5 mM to 3.0 mM) did not affect enzyme reactivation by DTT, and the addition of $CoCl_2$ or $NiCl_2$ reduced the recovered activity to only 10% or 15% of the initial activity, respectively. Zinc was the only metal to enhance the reactivation (up to 48% activity recovery in the presence of 100 µM $ZnCl_2$). DEPC modification of histidine residues also completely inactivated the enzyme.

Characterization of thermal and kinetic properties of the recombinant *T. ethanolicus* 2° ADH

The activity of the recombinant 2° ADH toward ADH substrates was characterized. The $Vmax_{app}$ for propan-2-ol (68 units mg^{-1}) was 3.6-fold higher than for ethanol (19 units mg^{-1}), and the Km_{app} toward the 1° alcohol (53 mM) was almost 50-fold higher than for the 2° alcohol (1.1 mM). The catalytic efficiency of the recombinant enzyme was determined to be approximately 170-fold greater toward the 2° alcohol (0.062 ml min^{-1} mg^{-1}) than toward the 1° alcohol (0.00036 ml min^{-1} mg^{-1}). The Km_{app} for $NADP^+$ was 0.011 mM, and no NAD(H) dependent activity was detected. The temperature dependence of native and recombinant enzyme activities were determined to be similar. Thus, only the data obtained with the recombinant enzyme are reported here. *T. ethanolicus* 2° ADH activity was detected below 25°C and increased to beyond 90°C (Fig. 4A). The 2° ADH half-life at 90°C was 1.7 h. The Arrhenius plot for the oxidation of propan-2-ol was linear from 30°C to 90°C when the enzyme was incubated at 55°C prior to assay. Under the same conditions, however, a distinct discontinuity was seen in the Arrhenius plot for ethanol oxidation (Fig. 4B). Discontinuities were also observed at ~55°C and ~46°C for propan-2-ol and ethanol oxidations, respectively, when the enzyme was preincubated at 0°C or 25°C (Table 2). The slopes of the best fit regression lines above and below the discontinuity were significantly different beyond the 95% confidence level except for propan-2-ol oxidation by enzyme preincubated at 55°C, where the regression line slopes were similar at the 95% confidence level. The activation energies for propan-2-ol and ethanol oxidations were similar at assay temperatures above the discontinuities but were 15 to 20 kJ mol^{-1} higher for ethanol oxidation than for propan-2-ol oxidation at temperatures below the discontinuity temperatures (Table 2). Furthermore, the differences between activation energies above and below the discontinuity temperatures (D) decreased with increasing preincubation temperatures. and the differences for ethanol oxidation were at least 3-fold higher than for propan-2-ol oxidation.

Figure 4. **Arrhenius plots for the recombinant** *T. ethanolicus* **39E 2° ADH**

between 25°C and 90°C. (A) Temperature-activity data for propan-2-ol oxidation with

the enzyme preincubated at 55°C. Linear regression best fits to the data was determined to

be y = 13.647 - 3.3694x (R^2 = 0.993). (B) Temperature-activity data for ethanol oxidation

with the enzyme preincubated at 55°C. Points above and below the discontinuity are

indicated by ■ and●, respectively. Linear regression best fits to the data were determined

to be y = 9.7929 - 2.5036x (R^2 = 0.946) for points above and y = 15.730 - 4.3899x (R^2 =
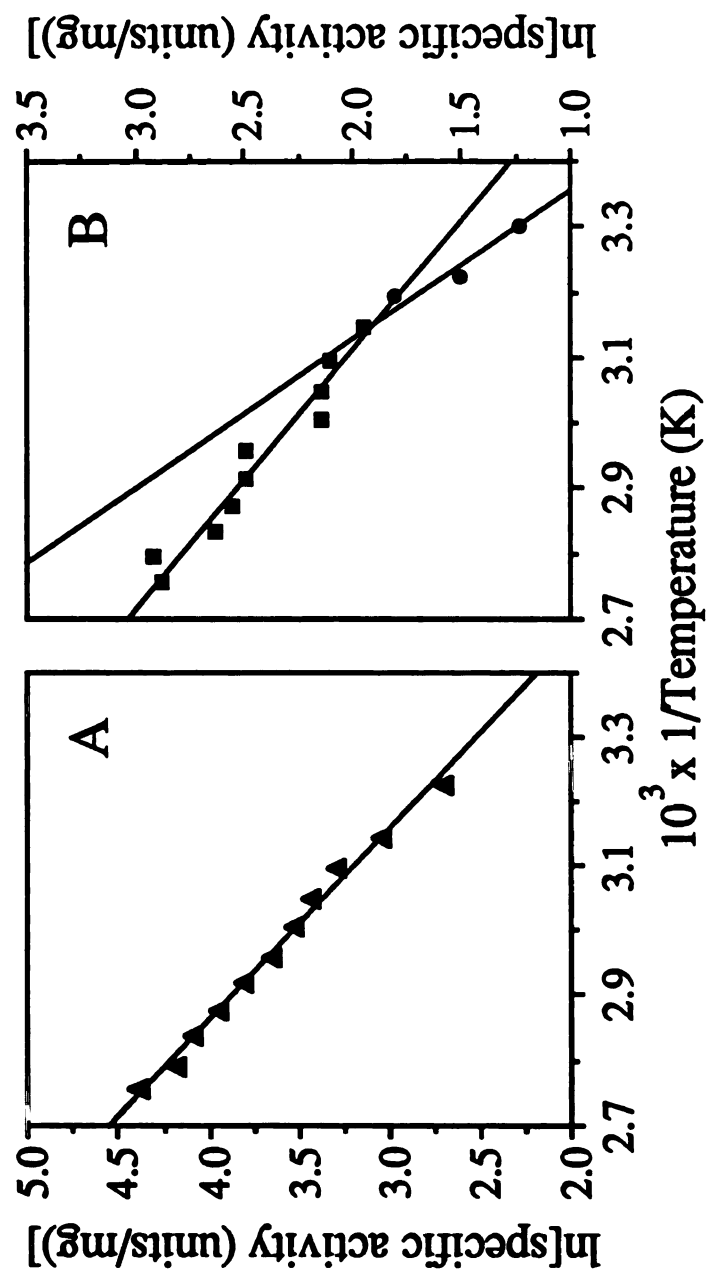
0.969) for points below the discontinuity.

Table 2. Effect of enzyme preincubation temperature on the activation energy for 2-propanol and ethanol oxidation by *T. ethanolicus* 39E 2° Adh

Substrate	Preincubation temperature (°C)	Discontinuity temperature (°C)	Activation energy (kJ mol^{-1})		
			above discontinuity	below discontinuity	[b]D
propan-2-ol	0	55	17	26	9
	25	54	20	26	6
	55	[c]ND	20	21	1
ethanol	0	48	18	45	27
	25	46	23	43	20
	55	45	21	36	16

[a]The existence of a biphasic Arrhenius plot was determined at the 95% confidence level by covariance analysis of the best fit regression lines to the data.

[b]D = (activation energy below the discontinuity) - (activation energy above the discontinuity).

[c]ND: Not detectable (the rate discontinuity was statistically insignificant at the 95% confidence level).

DISCUSSION

This first report of the cloning and expression of a thermophilic 2° ADH also provides evidence on the molecular bases for enzyme activity, thermophilicity, and discontinuous Arrhenius plots. Protein sequence comparisons predict that 1° and 2° ADHs have some structure-function similarities related to catalytic Zn and nicotinamide cofactor binding. However, differences in the hydrophobic clusters present in the overall 1° and 2° ADH sequences predict significant differences in the overall structures of these enzymes. The recombinant *T. ethanolicus* protein, with kinetic properties similar to those of the native enzyme [6], is a thermophilic and thermostable NADP(H) dependent enzyme that exhibits significantly greater catalytic efficiency toward 2° alcohols than 1° alcohols due to both lower Km_{app} and higher $Vmax_{app}$ values. Chemical modification experiments suggest that enzyme activity requires at least Cys and His residues and a tightly bound Zn atom. Finally, the magnitude of the bend in the Arrhenius plots varied inversely with enzyme preincubation temperature, indicating the existence of a catalytically significant, temperature dependent structural change in the enzyme.

The *adhB* gene encoding the thermophilic *T. ethanolicus* 2° ADH was cloned by hybridization and expressed in *E. coli*. Comparisons of the native and recombinant enzyme N-terminal sequences and molecular masses showed that the recombinant enzyme differed from the native protein only by the addition of an N-terminal methionine. While the first ATG is the main translation start codon in *E. coli*, it is unknown if it is the sole initiation codon (and the first methionine is processed later), or if the second ATG is also used as an initiation codon. The presence of a truncated ORF (preceded by an RBS) 180 nucleotides downstream of *adhB* and the absence of any potential transcriptional stop signal suggested that *adhB* may be the first gene of an operon. Although confirmation of this hypothesis will require additional experimentation, the similarity between the truncated ORF reported

here and the one downstream of *C. beijerinckii adhB* lends further support to this hypothesis.

HCA alignments of the 1° and 2° ADHs directed by their pyridine dinucleotide binding motifs indicated some structural similarity between these classes of enzymes. The negatively charged Asp[223, 195, 215] residues are consistent with the NAD(H) dependence of the horse liver, *B. stearothermophilus* , and *A. eutrophus* enzymes, respectively, while the presence of an uncharged residue (Gly[198]) at the analogous position in the *T. ethanolicus*, *T. brockii* and *C. beijerinckii* 2° ADHs is consistent with the NADP(H) dependence of these enzymes [3,6,7]. Furthermore, residues identified in the enzyme hydrophobic core, residues lining the active site cavity, and structurally important glycines previously identified in the horse liver 1° ADH [1], were better conserved between proteins in this alignment than the overall peptide sequences. Similarity of putative structurally critical regions has been reported for other ADH comparisons [34]. The correlation between enzyme structure-function properties and the predictions of this alignment support its use in further ADH structure-function studies. However, the apparent lack of a structural Zn binding loop in the 2° ADHs, the reports of only tetrameric 2° ADHs but both tetrameric and dimeric 1° ADHs, and the significantly greater similarity between the 2° ADHs than between the 1° and 2° ADHs predicted by HCA comparison also suggests that, while functionally similar, 1° and 2° ADHs may be less structurally similar than was it previously believed.

The thermophilic *T. ethanolicus* 39E 2° ADH incubated at 90°C retained detectable activity for more than one hour while the mesophilic *C. beijerinckii* 2° ADH was completely inactivated within 10 min at 70°C [6]. Still, these two enzymes shared more than 85% sequence similarity. Nine of the nonconservative substitutions between the subunits of these proteins correspond to prolines in the *T. ethanolicus* enzyme. The similarly thermophilic *T. brockii* 2° ADH contains 8 of these additional prolines. This difference in proline content of the thermophilic versus mesophilic 2° ADHs is consistent

with the hypothesis of Matthews *et al.* regarding the role of prolines in protein stabilization [33] and with the observations by numerous investigators of protein stabilization due to proline insertion [35,36], constrained loop regions [37], and proline substitution into loops [38-40]. The additional prolines in the *T. ethanolicus* 2° ADH were either in short putative loop regions or in longer putative loop regions containing multiple prolines, suggesting that the specific placement as well as the number of prolines may be critical to their stabilizing effect. The difference in proline content between the *T. ethanolicus* and *C. beijerinckii* enzymes and their overall high sequence similarity makes these 2° ADHs an excellent system for testing the effect of proline insertion on protein structure stabilization.

Comparative sequence analysis predicted the involvement of specific Cys, His, and Asp residues in 2° ADH catalysis. Chemical modification of the recombinant and native *T. ethanolicus* 2° ADHs with DTNB established the importance of cysteine residues in catalysis. The observation that Zn was the only metal enhancing the enzyme reactivation after DTNB modification suggested that the enzyme is Zn dependent. This finding is in agreement with the previous report that, once inactivated by the sulfhydryl modifying reagent p-chloromercuribenzoate, the *T. brockii* 2° ADH recovered activity only in the presence of $ZnCl_2$ [3]. Here, the recovery of 34% initial activity upon reversal of DTNB modification in the absence of added metal and the inability of EDTA to reduce the rate of reactivation, even at 60°C, suggested that the catalytic metal remained tightly bound to the protein. Histidine specific modification of the *T. ethanolicus* enzyme by DEPC was also accompanied by complete enzyme inactivation, implicating a histidine residue in catalysis. The apparent lack of a structural Zn binding region in the 2° ADH subunits argues that the DTNB linked inactivation and Zn dependent reactivation are not due to the loss and recovery of a cysteine-liganded structural Zn. Therefore, the catalytically important cysteine and histidine residues may act as ligands to the catalytic Zn as described for other ADHs [1]. Determination of the importance of Asp[150] to 2° ADH activity awaits mutagenic, crystallographic, or physical biochemical analysis.

The temperature dependence of catalytic activities for the native and recombinant enzymes were similar. The Arrhenius plot for propan-2-ol oxidation by $T.$ $ethanolicus$ 2° ADH preincubated at 55°C was linear, unlike that previously reported for the $T.$ $brockii$ 2° ADH [3]. However, statistically significant discontinuities in the Arrhenius plots for both ethanol and propan-2-ol oxidation by the $T.$ $ethanolicus$ enzyme were seen when the enzyme was preincubated at lower temperatures. A change in the rate determining step of the overall reaction and other explanations that do not invoke alterations in catalyst structure predict Arrhenius plot discontinuities [41,42]. A shift in the reaction slow step not related to an alteration in enzyme structure would be independent of the initial temperature of the enzyme, and is inconsistent with the inverse relationship observed here between enzyme preincubation temperature and the difference in activation energies above and below the discontinuity (D). At assay temperatures above the discontinuity, the reaction activation energies were similar for enzymes preincubated at 0°C, 25°C and 55°C, unlike those at lower assay temperatures. This observation is consistent with the enzyme attaining its optimally active conformation rapidly enough at higher assay temperatures not to affect the measured rate. The discontinuity temperatures, being below the lowest reported temperatures for $T.$ $ethanolicus$ growth, are physiologically irrelevant, but they underscore the importance of treating thermophilic enzymes differently from mesophilic enzymes when conducting kinetic analyses. Furthermore, the differences between the low temperature activation energies for ethanol and propan-2-ol argue that substrate/product-protein interactions are important to the rate determining step in catalysis.

The linear Arrhenius plot for propan-2-ol oxidation by $T.$ $ethanolicus$ 2° ADH preincubated at 55°C indicates that the low mesophilic temperature activity of this enzyme may be explained by the substrate energy alone. Arrhenius theory predicts increasing Km_{app} and $Vmax_{app}$ with increasing temperature [41], and this effect has been confirmed for the $Thermotoga$ $neopolitana$ D-xylose isomerase [43]. The thermophilic 2° ADH has lower $Vmax_{app}$ values than would be predicted from extension of the mesophilic enzyme

activity at 25°C. In fact, at 60°C the thermophilic 2° ADH maintains Km_{app} and $Vmax_{app}$ values toward substrates similar to those reported for the mesophilic enzyme at 25°C. These similar kinetic values suggest that the thermophilic enzyme channels more total reaction energy into substrate binding and less into turnover [44]. Therefore, an alternative explanation for the comparatively poor *T. ethanolicus* 2° ADH low temperature activity which argues that substrate affinity at high temperatures is maintained by sacrificing high turnover number must also be considered.

ACKNOWLEDGMENTS

REFERENCES

1.   Bränden, C.-I., Jörnvall, H., Eklund, H. and Furugren, B. (1975) in The enzymes, vol. XI part A (Boyer, P. D., ed.), pp. 103-190, Academic Press, NY

2.   Eklund, H., Nordström, B., Zeppezauer, E., Söderlund, G., Ohlsson, I., Boiwe, T. and Bränden, C.-I. (1974) FEBS Lett. **44**, 200-204

3.   Lamed, R. J. and Zeikus, J. G. (1981) Biochem. J. **195**, 183-190

4.   Steinbüchel, A. and Schlegel, H. G. (1984) Eur. J. Biochem. **141**, 555-564

5.   Bryant, F. O., Wiegel, J. and Ljungdahl, L. (1988) Appl. Env. Microbiol. **54**, 460-465

6.   Ismaiel, A. A., Zhu, C.-X., Colby,G. D. and Chen, J.-S. (1993) J. Bacteriol. **175**, 5097-5105

7.   Burdette, D. S. and Zeikus, J. G. (1994) Biochem. J. **302**, 163-170

8.   Lovitt, R. W., Shen, G.-S. and Zeikus, J. G. (1988) J. Bacteriol. **170**, 2809-2815

9.   Keinan, E., Hafeli, E. K., Seth, K. K. and Lamed, R. L. (1986) J. Am. Chem. Soc. **108**, 162-169

10.   Keinan, E., Hafeli, E. K., Seth, K. K. and Lamed, R. L. (1986) Ann. N. Y. Acad. Sci. **501**, 130-149

11.   Hummel, W. (1990) Appl. Microbiol. Biotech. **34**, 15-19

12.   Keinan, E., Seth, K. K., Lamed, R. L., Ghirlando, R. and Singh, S. P. (1990) Biocatalysis **4**, 1-15

13.   Hummel, W. and Kula, M.-R. (1989) Eur. J. Biochem. **184**, 1-13

14.   Persson, M., Månsson, M.-O., Bülow, L. and Mosbach, K. (1991) Biotechnology **9**, 280-284

15.   Lee, Y.-E., Jain, M. K., Lee, C., Lowe, S. E., and Zeikus, J. G. (1993) Int. J. Syst. Bacteriol. **43**, 41-51

16.   Nagata, N., Maeda, K. and Scopes, R. K. (1992) Bioseparation **2**, 353-362

17.   Peretz, M. and Burstein, Y. (1989) Biochemistry **28**, 6549-6555

18.   Oka, A., Sugisaki, H. and Takanami., M. (1981) J. Mol. Biol. **147**, 217-226

19.   Zeikus, J. G., Hegge, P. W. and Anderson, M. A. (1979) Arch. Microbiol. **122**, 41-48

20. Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecular cloning: A laboratory manual 2nd edition. (Nolan, C., ed.), Cold spring Harbor Press, NY

21. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. and Struhl, K. (1993) Current Protocols in Molecular Biology (Janssen, K., ed.), Current Protocols, NY

22. Doi, R. H. (1983) in Recombinant DNA Techniques: An Introduction (Rodriguez, R. L, and Tait, R. C., eds.), pp. 162-163, Benjamin-Cummings, Menlo Park, CA

23. Lee., C. C., Wu, X., Gibbs, R. A., Cook, R. G., Muzny, D. M. and Caskey, C. T. (1988) Science 239, 1288 - 1291

24. Sanger, F., Nicklen, S. and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467

25. Remington, R. D. and Schork, M. A. (1985) Statistics with applications to the biological and health sciences, pp. 296-309, Prentice-Hall, NJ

26. Brooks, S. (1992) Biotechniques 13, 906-911

27. Smith, P. K., Krohn, R. I., Hermanson, G. T., Mallia, A. K., Gartner, F. H., Provenzano, M. D., Fugimoto, E. K., Goele, N. M., Olson, B. J. and Klenk, D. C. (1985) Anal. Biochem. 150, 76-85

28. Habeeb, A. F. S. A. (1972) in Methods in enzymology vol. 25: Enzyme structure, part B (Hirs, C. H. W. and Timasheff, S. N., ed.), pp. 457-459, Academic press, NY

29. Miles, E. W. (1977) in Methods in enzymology vol. 47 (Hirs, C. H. W. and Timasheff, S. N., ed.), pp. 431-442, Academic press, NY

30. Lemesle-Verloot, L., Henrissat, B., Gaboriaud, C., Bissery, V., Morgat, A. and Mornon, J. P. (1990) Biochimie 72, 555-574

31. Hawley, D. K. and McClure, W. R. (1983) Nucl. Acids Res. 11, 2237-2255

32. Wierenga, R. K. and Hol, G. J. (1983) Nature 302, 842-844

33. Matthews, B. W., Nicholson, H. and Becktel, W. (1987) Proc. Natl. Acad. Sci. USA 84, 6663-6667

34. Jendrossek, D., Steinbuechel, A. and Schlegel, H. G. (1988) J. Bacteriol. 170, 5248-5256

35. Herning, T., Katsuhide, Y., Taniyama, Y. and Kikuchi, M. (1991) Biochemistry 30, 9882-9891

36. Herning, T., Katsuhide, Y., Inaka, K., Matsushima, M. and Kikuchi, M. (1992) Biochemistry 31, 7077-7085

37. Eijsink, V. G. H., Vriend, G., van de Burg, B., van de Zee, J. R., Veltman, O. R., Stulp, B. K. and Venema, G. (1993) Protein Engineer. 5, 157-163

38. Watanabe, K., Masuda, T., Ohashi, H., Mihara, H. and Suzuki, Y. (1994) Eur. J. Biochem. **226**, 277-283

39. Watanabe, K., Chishiro, K., Kitamura, K. and Suzuki, Y. (1991) J. Biol. Chem. **266**, 24287-24294

40. Hardy, F., Vriend, G., Veltman, O. R., van de Vinne, B., Venema, G. and Eijsink, V. G. H. (1993) FEBS Lett. **317**, 89-92

41. Segel, I. H. (1975) in Enzyme kinetics: behavior and analysis of rapid equilibrium and steady-state enzyme systems, pp. 929-934, John Wiley & Sons, NY

42. Londesborough, J. (1980) Eur. J. Biochem. **105**, 211-215

43. Vieille, C., Hess, J. M., Kelly, R. M., and Zeikus, J. G. (1995) Appl. and Environ. Microbiol. **61**, 1867-1875

44. Jencks, W. P. (1975) in Advances in enzymology and related areas of molecular biology, vol. 43, (Meister, A., ed.), pp. 219-411, John Wiley & sons, NY

Chapter III

Effect of thermal and chemical denaturants on *Thermoanaerobacter ethanolicus* 39E secondary-alcohol dehydrogenase stability

Prepared for submission to Biotechnology and Applied Biochemistry

121

ABSTRACT

Thermoanaerobacter ethanolicus secondary-alcohol dehydrogenase (2° ADH) was optimally active near 90°C with calculated thermostability half-lives of 1.2 days, 1.7 h, 19 min, 9.0 min, and 1.3 min at 80°C, 90°C, 92°C, 95°C, and 99°C, respectively. Enzyme activity loss upon heating (90-100°C) was accompanied by precipitation, but the soluble enzyme remaining after partial sample inactivation retained complete activity, mechanistically linking enzyme inactivation and precipitation. Enzyme thermoinactivation was modeled by a pseudo-first order rate equation suggesting that the rate determining step was unimolecular with respect to protein, thus, also implying that thermoinactivation preceded aggregation. Structural unfolding of the 2° ADH occurred with a T_m (50% unfolding) at ~115°C. This T_m was much higher than the temperature for maximal activity, indicating that the 2° ADH was rigid and tightly folded throughout its active temperature range. Thermodynamic calculations indicated that the active folded structure of the 2° ADH, like other proteins, is stabilized by a relatively small Gibbs energy ($\Delta G^{\ddagger}_{stab.}$ = 110 kJ mol^{-1}). 2° ADH was also active in > 1.0 M GuHCl, demonstrating its high resistance to chemical denaturation. Enzyme activities were increased 2-fold in low guanidine hydrochloride (GuHCl) concentrations (150-230 mM). The additive denaturing effects of GuHCl and temperature on *T. ethanolicus* 2° ADH $T_{m[Gu]}$ predict that at higher temperatures lower GuHCl concentrations should be required to reduce enzyme rigidity, thus lower GuHCl concentrations should result in optimal enzyme flexibility and activity. The increasing GuHCl concentration that provided maximal 2° ADH catalytic rate enhancement at higher temperatures was also seen with weakly chaotropic (KNO$_3$) or neutral (KCl) inorganic salts, suggesting that this 2° ADH catalytic rate enhancement results from specific ionic and not hydrophobic interactions.

INTRODUCTION

As industrial catalysts, enzymes must have robust activities that are stable under process chemical and thermal conditions. The saccharidases used in starch processing and the proteases used in detergents typify current industrial enzyme standards. The potential biotechnological application of enzymes as chiral chemical catalysts has long been recognized [1-10]. The chiral synthesis of high value bioactive compounds such as pharmaceuticals is increasingly important due to the recent FDA ruling mandating either enantiomeric purity or exhaustive toxicity testing to demonstrate the lack of any detrimental effects due to enantiomeric impurities [11]. Traditional chemical synthetic methods can often only ensure chiral purity using chiral feedstock chemicals because racemate resolution is often too difficult and expensive. Thus, the economic biocatalytic synthesis of either chiral feedstock chemicals or of the final chiral products can significantly advance current technologies.

Alcohol dehydrogenases (ADHs) are almost exclusively nicotinamide cofactor dependent, enantioselective enzymes with broad specificity [12]. While NAD(P)(H) is costly, the demonstrated effectiveness of numerous NAD(P)(H) retention and redox recycling strategies has removed this economic barrier to the scale-up of even moderate value adding processes [9,10,13]. Analytical scale chiral alcohol formation by 1° ADHs has been demonstrated for a diverse set of substrates [14,5,6]. These 1° ADHs are far less active on 2° alcohols or ketones than on 1° alcohols or aldehydes. The emerging class of 2° ADHs have greater catalytic efficiency for 2° alcohol oxidation and ketone reduction than for 1° alcohol oxidation and aldehyde reduction and they display broad substrate specificities [15-17]. The potential for 2° ADH catalyzed chiral chemical syntheses has been demonstrated using the *Thermoanaerobacterium brockii* and *Thermoanaerobacter ethanolicus* enzymes [18-22].

The kinetic similarity of the 2° ADHs isolated from a number of similar thermoanaerobes led Nagata *et al.* to conclude that the enzymes were also structurally similar [15]. A comparison of the *T. brockii* 2° ADH peptide sequence, determined by Edman degradation, to that of liver 1° ADH predicted some potential structure function similarities [21], however, without a 2° ADH clone these predictions could not be tested. Recently, the *adhB* gene encoding the *T. ethanolicus* 2° ADH was cloned, characterized, and overexpressed, demonstrating that the *T. ethanolicus* and *T. brockii* 2° ADHs differed by only 3 amino acids [23]. Therefore, research using the *T. ethanolicus* 2° ADH may provide knowledge representative of other 2° ADHs.

Thermozymes [24], with their greater resistance to both thermal and chemical denaturation than mesophilic enzymes, may alleviate some of the limiting process conditions for applications of enzymes and enhance enzyme lifespans. Irreversible mesophilic and thermophilic protein thermoinactivation has been shown to result from the loss of folded protein structure and from covalent peptide or amino acid destruction [see 24], suggesting that the higher stability of thermozymes is due to interactions mechanistically similar but stronger than those stabilizing mesophilic enzymes [see 24]. Thermophilic enzymes recombinantly expressed in mesophilic hosts retain their thermal characteristics, indicating that the peptide sequences encode these properties. Comparative analyses of thermophilic and mesophilic enzymes have typically concluded, from the lack of exotic amino acids or 1° structural differences, that the added stability of one enzyme compared to another results from numerous subtle differences in the folded protein structure [24,25]. Enzyme thermostabilizing mechanism theories including increased core hydrophobicity and atomic density, additional salt bridges, 2° structural element stabilization, and constrained surface loop architecture have been reported [see 24].

Detailed studies on thermal and chemical denaturation of a thermophilic 2° ADH have not been reported. The data presented here measure *T. ethanolicus* 2° ADH thermophilicity, thermostability, and stability to denaturation by GuHCl. The relationship

between *T. ethanolicus* 2° ADH rigidity and thermostability is determined through kinetic and biophysical characterization of enzyme activity, precipitation, and unfolding. Also, the effect of the chemical denaturant guanidine hydrochloride (GuHCl) on the activity of this robust biocatalyst is examined.

MATERIAL AND METHODS

Chemicals and reagents

All chemicals were of at least reagent/molecular biology grade. Gases were provided by AGA specialty gases (Cleveland, OH) and made anaerobic by passage though heated copper filings. Spectrophotometrically pure GuHCl was obtained as an 8.0 M aqueous solution (Cat. #24115: Pierce; Rockford, IL). Tris buffer pH values were measured at 25°C and formulated for pH 8.0 at the temperature they were used (thermal correction factor = -0.031 ΔpH °C^{-1}).

Media and strains

Escherichia coli (DH5α) containing the 2° ADH recombinant plasmid [23] was grown in rich, complex medium (20 g l^{-1} tryptone, 10 g l^{-1} yeast extract, 5 g l^{-1} NaCl) at 37°C with 25 μg ml^{-1} kanamycin and 100 μg ml^{-1} ampicillin.

Enzyme purification

The recombinant 2° ADH was purified from *E. coli* (DH5α) aerobically. The pelleted cells from batch cultures were resuspended (0.5 g wet wt. ml^{-1}) in buffer A {50 mM Tris:HCl (pH 8.0), 5 mM DTT, and 10 μM ZnCl$_2$} and lysed by passage though a French pressure cell. The clarified lysate was incubated at 65°C for 25 min then centrifuged for 30 min at 15000g. The clarified supernatant was applied to a DEAE-sephacryl column (2.5 cm x 15

cm) equilibrated with buffer B {50 mM Tris:HCl (pH 8.0)}. Active fractions were diluted

4-fold in buffer B and applied to a Q-sepharose column (2.5 cm x 10 cm) equilibrated with

buffer B. Purified recombinant enzyme was eluted using a 250 ml NaCl gradient (0-300

mM).

Enzyme activity

The standard assay for 2° ADH activity was defined as the reduction of $NADP^+$ using

propan-2-ol at 60°C as previously described [17]. Kinetic parameters were calculated from

nonlinear best fits of the data to the Michaelis-Menten equation using Kinzyme software

[26] on an IBM PC. Protein concentrations were measured in all cases by the

bicinchoninic acid (BCA) procedure (Pierce; Rockford, IL). The effect of GuHCl on 2°

ADH activity was determined at 37°C, 50°C, 60°C, 70°C, and 75°C in the absence or

presence of 100 mM, 200 mM, 400 mM, 800 mM, or 1.6 M GuHCl.

Ionic activities

Guanidinium ionic activities were determined using the Debye-Hückel limiting law [see

27]. The mean ionic activity ($a_\pm$) is defined as

$$a_\pm = \gamma_\pm m_\pm$$

where $m_\pm$ is the mean electrolyte molality and $\gamma_\pm$ is the mean activity coefficient. The mean

activity coefficient is determined as a function of temperature (T), ionic charge (z),

dielectric constant of water ($\varepsilon = 78.54$), and ionic strength (I) using the equation:

$$\log \gamma_\pm = \{-1.824 \times 10^6 / [(\varepsilon T)^{(3/2)}]\}(|z_+ z_-|)[(I)^{(1/2)}]$$

where

$$I = (1/2)\sum(m_i z_i^2).$$

The GuHCl concentration dependent enzyme activity experiments were performed in 45 mM to 50 mM Tris (pH 8.0) at 37°C, 50°C, 60°C, 70°C, and 75°C. Tris buffer was assumed to be 50% dissociation (pKa = 8.1). Because of GuHCl's pKa (> 10) it was assumed to be completely ionized. All species ionic charges were ±1.

Thermostability

Purified recombinant 2° ADH thermostability was evaluated by timed incubation at the desired temperatures, followed by incubation for 30 min at 25°C. Samples were then equilibrated at 55°C for at least 15 min prior to activity determination. Activity was determined using the unfractionated samples. Incubations were performed in 100 µl PCR tubes (cat. #72.733.050, Sarstedt; Newton, NC) using 0.2 mg ml^{-1} protein in 100 µl of buffer B. Protein precipitation due to heating for 2 min at 99°C was quantified in similar experiments by fractionation of the soluble and aggregated protein. First, the 70 µl of sample remaining after thermostability activity determination was centrifuged for 5 min at 12000g. The supernatant was then removed from the pelleted protein aggregate, the pellet was resuspended in 100 µl of buffer B, centrifuged (12000g, 5 min), the wash buffer removed, and the repelleted protein resuspended in 70 µl of buffer B. Enzyme activities and protein concentrations were determined for the initial sample, supernatant, pellet wash, and the resuspended pellet. Specific acivities and total protein values from the pellet wash and resuspended pellet fractions were combined and reported as the precipitate specific activity and total protein, respectively.

Heat induced enzyme precipitation

Heat-induced enzyme precipitation was monitored from 25°C to 100°C by light scattering ($\lambda = 580$ nm), using protein solutions (0.2 mg ml^{-1}) in buffer B. Absorbance measurements were conducted in 0.3 ml quartz cuvettes (pathlength = 1.0 cm), using a Gilford Response spectrophotometer (Corning; Oberlin, OH) equipped with a Peltier cuvette heating system. The apparent OD_{580} increase is due to light scattering from protein precipitation.

Heat-induced enzyme unfolding

2° ADH unfolding experiments were conducted in buffer B in the absence of GuHCl and in the presence of 2.0–6.0 M GuHCl. Unfolding was monitored from 25°C to 100°C by absorbance at 280 nm [28-30] in 0.3 ml quartz cuvettes (pathlength = 1.0 cm), using a Gilford Response spectrophotometer equipped with a Peltier cuvette heating system. Melting temperature (T_m) is defined as the temperature of the absorbance transition midpoint during an increasing thermal gradient (1.0°C min^{-1}). The 95% confidence interval error boundaries for the predicted 2° ADH T_m were determined by standard statistical treatment of the data [31].

Fluorescence spectrometry

Fluorimetric unfolding measurements were conducted with protein solutions (0.2 units $O.D._{280}$) in buffer C [10 mM Tris:HCl (pH 8.0)] at 25°C. Fluorescence spectra ($\lambda_{excitation} = 295$ nm, $\lambda_{emission} = 315\text{-}450$ nm) were recorded in 2.25 ml quartz cuvettes (pathlengths: excitation = 0.5 cm, emission = 1.0 cm), using an SLM spectrofluorimeter (Champain-Urbana, Il) equipped with a thermal, cell holder. Decreasing flourescence intensity indicates exposure of the aromatic amino acid residues to the polar solvent environment.

RESULTS

Thermophilicity and thermostability

T. ethanolicus 2° ADH activity increased with temperature to at least 90°C with no activity

observed at 100°C (Fig. 1). Accurately quantitating enzyme activity at temperatures above

90°C was difficult due to NADP(H) instability and equipment constraints, however, no

product formation was detected during reactions at 100°C. Recombinant 2° ADH

thermostability was determined at 65°C, 80°C, 90°C, 92°C, 95°C, and 99°C (Fig. 2). The

enzyme was stable at 65°C, and, it had calculated half-lives of 1.2 days, 1.7 h, 19 min, 9.0

min, and 1.3 min at 80°C, 90°C, 92°C, 95°C, and 99°C, respectively. The rate of

thermoinactivation was described by equation 1 for all temperature measured (k=rate

constant and t=time (min)).

$$\ln(\text{residual activity}) = \ln(\text{initial activity}) - kt \qquad (1)$$

This pseudo first order relationship indicated that the slow step in *T. ethanolicus* 2° ADH

thermoinactivation did not depend on enzyme concentration. Arrhenus analysis of the

inactivation rate constants (6.70×10^{-6} sec^{-1}, 1.34×10^{-4} sec^{-1}, 6.66×10^{-4} sec^{-1}, 1.60×10^{-3}

sec^{-1}, and 1.87×10^{-2} sec^{-1} at 80°C, 90°C, 92°C, 95°C, and 99°C, respectively) yielded a

transition state Gibbs energy barrier of 110 kJ mol^{-1} resulting from the difference between

$\Delta H^{\ddagger}$ = 440 kJ mol^{-1} and $\Delta S^{\ddagger}$ = 330 kJ mol^{-1} (temperature = 363 K).

Thermostability and unfolding

Enzyme inactivation was accompanied by precipitation and the effect of protein aggregation

on enzyme activity was investigated. Spectrophotometric determination of heat-induced

protein precipitation using a 1.0°C per minute thermal gradient indicated that the enzyme

Figure 1

Recombinant *T. ethanolicus* 39E 2° ADH thermophilicity. 2° ADH propan-2-ol oxidation

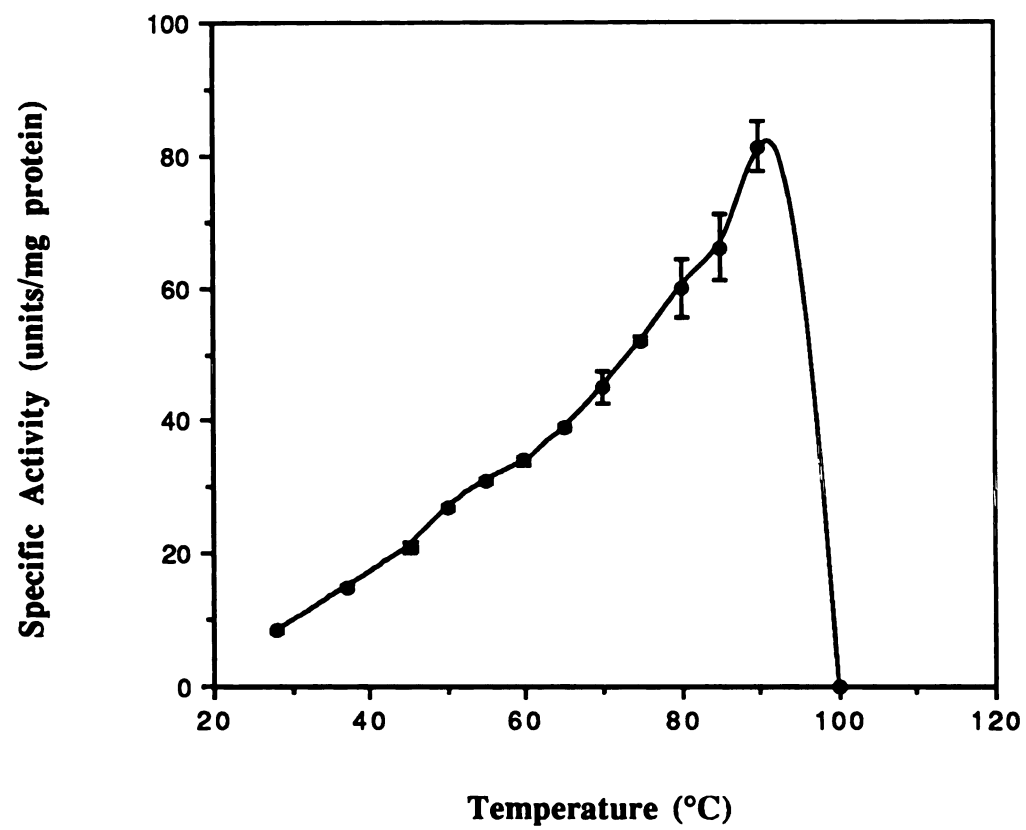activity was measured at temperatures spanning 30°C to 100°C.
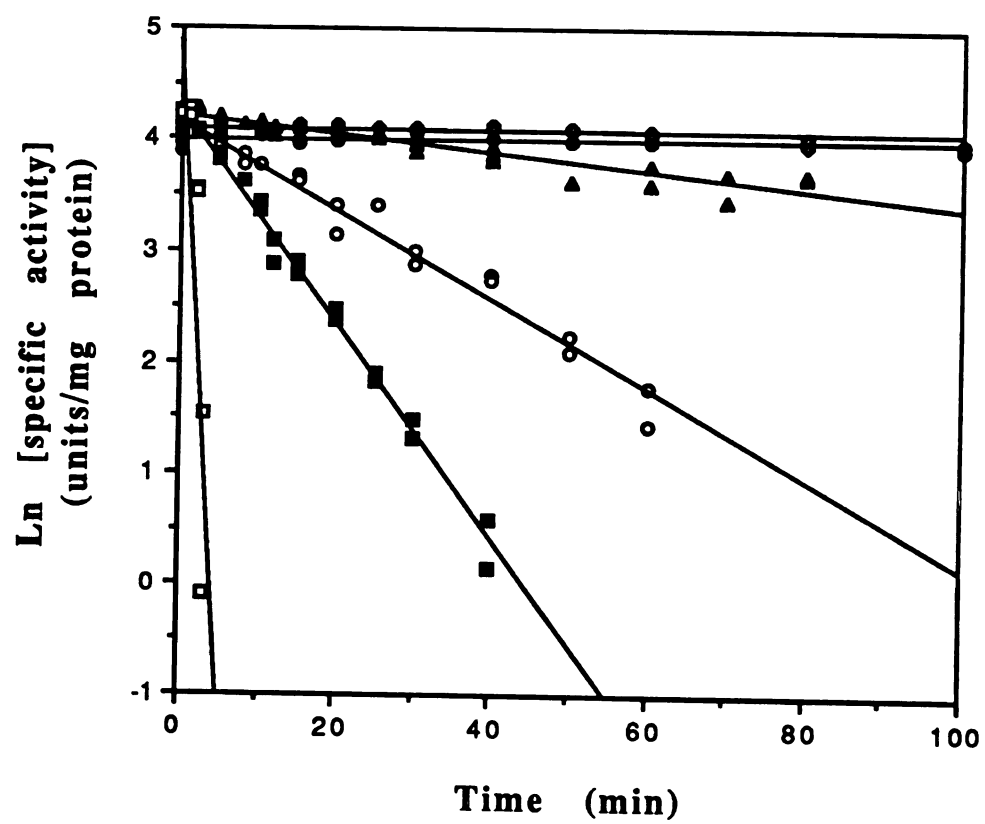
Figure 2

Recombinant *T. ethanolicus* 2° ADH thermostability. 2° ADH residual activity was determined after timed (1 min to 100 min) incubations at 65°C (●), 80°C (◇), 90°C (▲), 92°C (O), 95°C (■), and 99°C (□).

began to precipitate at temperatures above 85°C (Fig. 3). Precipitation was a rapid,

cooperative process that was essentially complete at 100°C with the transition midpoint at

~95°C. The potential correlation between enzyme precipitation and loss of catalytic activity

was then investigated by direct measurement (Table 1). After heat treatment for 2 min at

99°C (approximately the enzyme half-life at this temperature), the specific activity of the

unfractionated heat treated sample was half that of the unheated enzyme. Of the 11 ± 1.1

µg of protein originally added, 10 ± 1.0 µg were recovered after heat treatment.

Subsequent fractionation of the remaining sample (70 µl of the original 100 µl) by

centrifugation indicated that the supernatant contained half of the initial enzyme. This

soluble enzyme had similar specific activity to the unheated sample. The resuspended

pellet, which did not redissolve, also contained half of the protein but displayed no

detectible enzyme activity. Once precipitate had formed, enzyme continued to inactivate

and aggregate slowly upon standing at room temperature (data not shown).

The role of protein unfolding in 2° ADH thermoinactivation was investigated by

measuring the effect of GuHCl on temperature dependent enzyme unfolding.

Spectrophotometrically determined transition midpoints for protein unfolding (using a

1.0°C per minute thermal gradient) with 2.0 M, 2.25 M, 2.5 M, 2.75 M, and 3.0 M added

GuHCl yielded melting temperature ($T_{m[Gu]}$) values of 68°C, 59°C, 58°C, 53°C, and 48°C,

respectively. T_m is defined as the midpoint of the unfolding transition and is interpreted as

the point where the protein is 50% unfolded. The inset in figure 4 of a representative 2°

ADH melting curve used to obtain the $T_{m[Gu]}$ values indicates that unfolding was a rapid,

cooperative process. At GuHCl concentrations below 2.0 M enzyme precipitated prior to

completing the unfolding transition so no midpoint temperature could be identified. Also,

at concentrations above 3.0 M the absorbance change registered upon 2° ADH denaturation

became to small for reliable $T_{m[Gu]}$ determination. Enzyme $T_{m[Gu]}$ values decreased

linearly with increasing GuHCl concentrations (Fig. 4). The GuHCl dependence of 2°

Figure 3

Temperature dependence of 2° ADH precipitation. 2° ADH precipitation during an

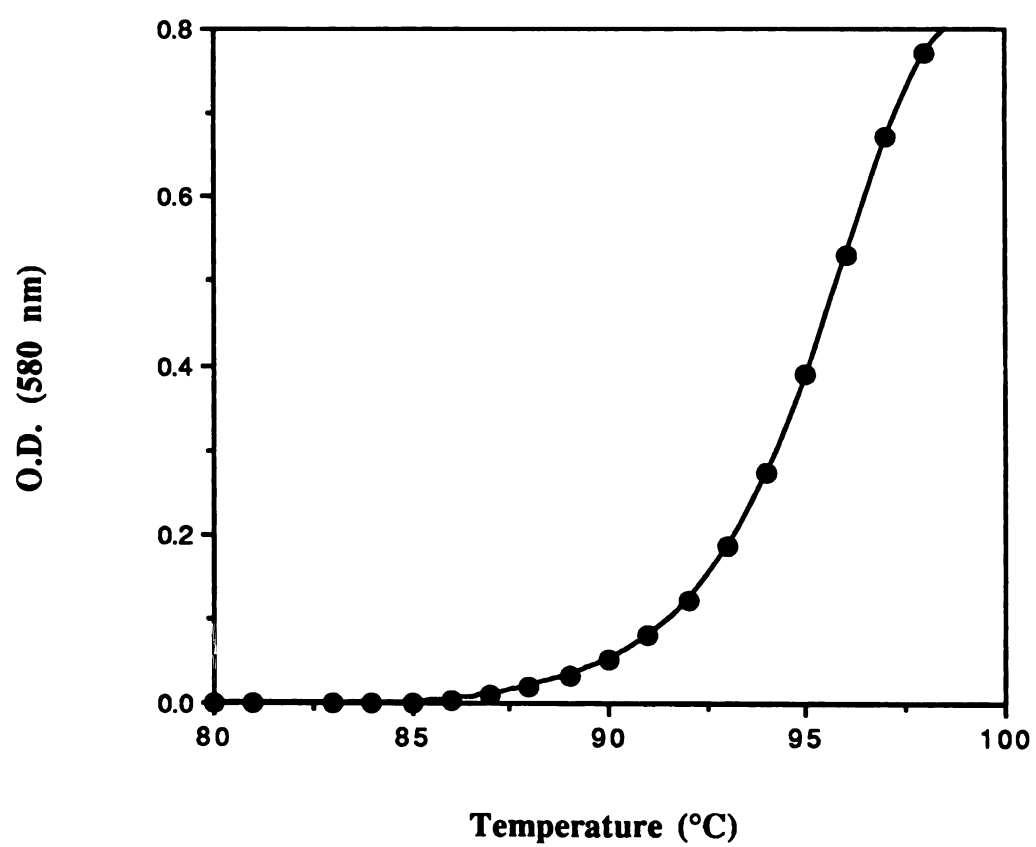incresaing thermal gradient (1.0°C min^{-1}) was determined by sample absorbance at 580 nm.

Table 1 Role of protein precipitation in *T. ethanolicus* 2° ADH thermal inactivation[a]

	Total protein (µg)	Total activity (units)	Specific activity (units mg⁻¹)
Initial	11 ± 0.050	0.38 ± 0.036	35 ± 2.3
Fractions[b]			
Supernatant	5.6 ± 0.56[c]	0.18 ± 0.010	33 ± 0.56
Pellet	5.6 ± 0.46[c]	< 0.00032	< 0.06
Complete[d]	10 ± 1.0	0.18 ± 0.018	18 ± 0.065

[a]100µl protein samples in buffer B were incubated at 99°C for 2 min.

[b]The precipitated material (pellet) was separated from the soluble material (supernatant) by centrifugation after heat treatment. The pelleted precipitate was washed with fresh buffer and repelleted by centrifugation. The values reported for the pellet fractions are the sum of those in the wash and the final pellet.

[c]Total measured protein values were multiplied by 1.4, normalizing them to the initial 100 µl reaction volumes (30 µl were removed for analysis of the complete sample prior to fractionation).

[d]Proir to precipitate and supernatant separation the homogenius sample enzyme activity and protein concentrations were determined.

Figure 4

Effect of GuHCl on *T. ethanolicus* 2° ADH melting temperature. The $T_{m[Gu]}$ values from absorbance were measured at 280 nm during an increasing thermal gradient ($1.0°C\ min^{-1}$) with 2.0 M, 2.25 M, 2.5 M, 2.75 M, and 3.0 M added GuHCl. The inset shows percent protein unfolding versus temperature for 2° ADH in 2.5 M GuHCl.
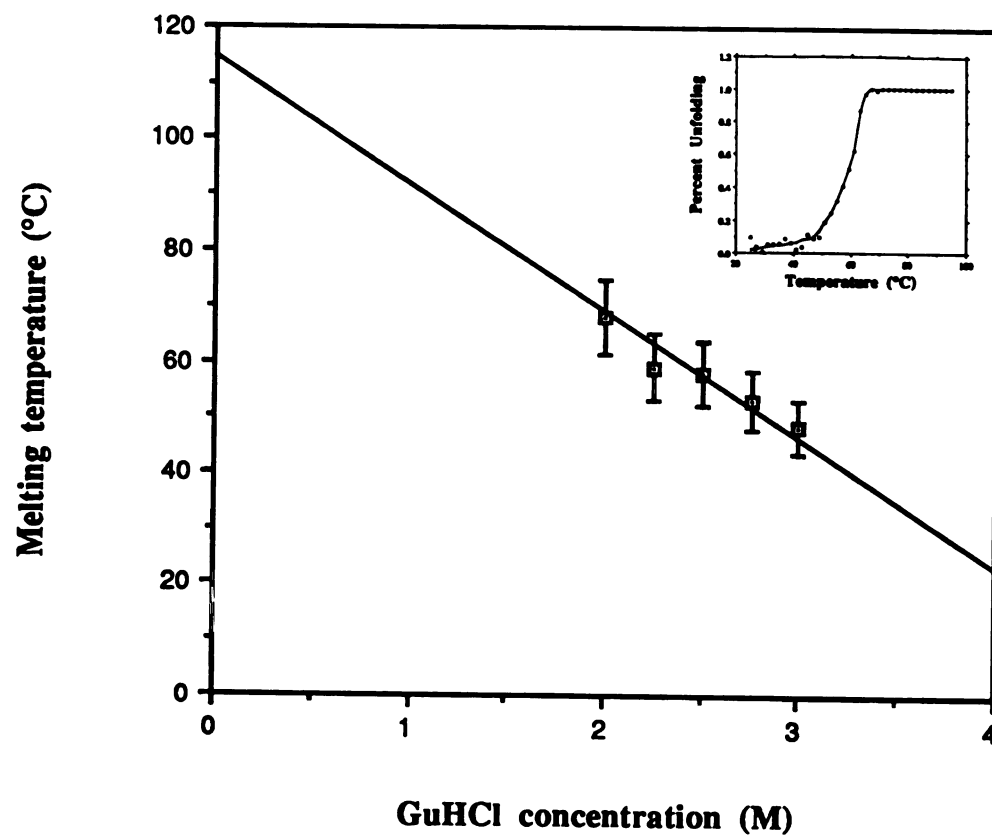
ADH denaturation determined by fluorescence intensity at 25°C indicated a melting

transition midpoint at 3.7 M (Fig. 5). Extrapolating to zero GuHCl from both the

absorbance and fluorimetric $T_{m[Gu]}$ data predicts a 2° ADH T_m of 115 ± 5.8°C (at the 95%

confidence level).

Chemical stability and 2° ADH activity

The generally poor low temperature activity of thermozymes has been attributed to their

high rigidity at these temperatures [32]. Therefore, reducing enzyme rigidity by adding a

chemical denaturant would be expected to enhance thermozyme low temperature activity.

The chemical denaturant GuHCl and thermal energy were also used to study the effect of

enzyme rigidity on catalysis. Figure 6 shows the effect of GuHCl concentrations up to 1.6

M on enzyme activity at temperatures between 37°C and 75°C. Maximal enzyme specific

activity was seen in the presence of 150 mM to 230 mM GuHCl at all temperatures tested

and it was approximately doubled that in the absence of GuHCl. The GuHCl concentration

corresponding to maximal enzyme activity increased with increasing temperature (147 mM,

184 mM, 206 mM, 237 mM, and 231 mM at 37°C, 50°C, 60°C, 70°C, and 75°C,

respectively). The mean ionic activity coefficients ($\gamma_\pm$) calculated from the Debye-Hückel

limiting law for these maximal activity conditions averaged to 0.808 ± 0.0051. The

calculated GuHCl activities also increased with increasing temperature (120 mM, 149 mM,

165 mM, 183 mM, and 187 mM at 37°C, 50°C, 60°C, 70°C, and 75°C, respectively).

Similar but less dramatic low concentration 2° ADH rate enhancements were seen for added

KCl and KNO_3 but not for K_2SO_4 (data not shown).

Figure 5

Effect of GuHCl on *T. ethanolicus* 2° ADH fluorescence. The flourescence measurements ($\lambda_{excitation}$=295 nm, $\lambda_{emission}$=340 nm) at 25°C were conducted using protein in GuHCl concentrations from 0.0 to 6.0 M.
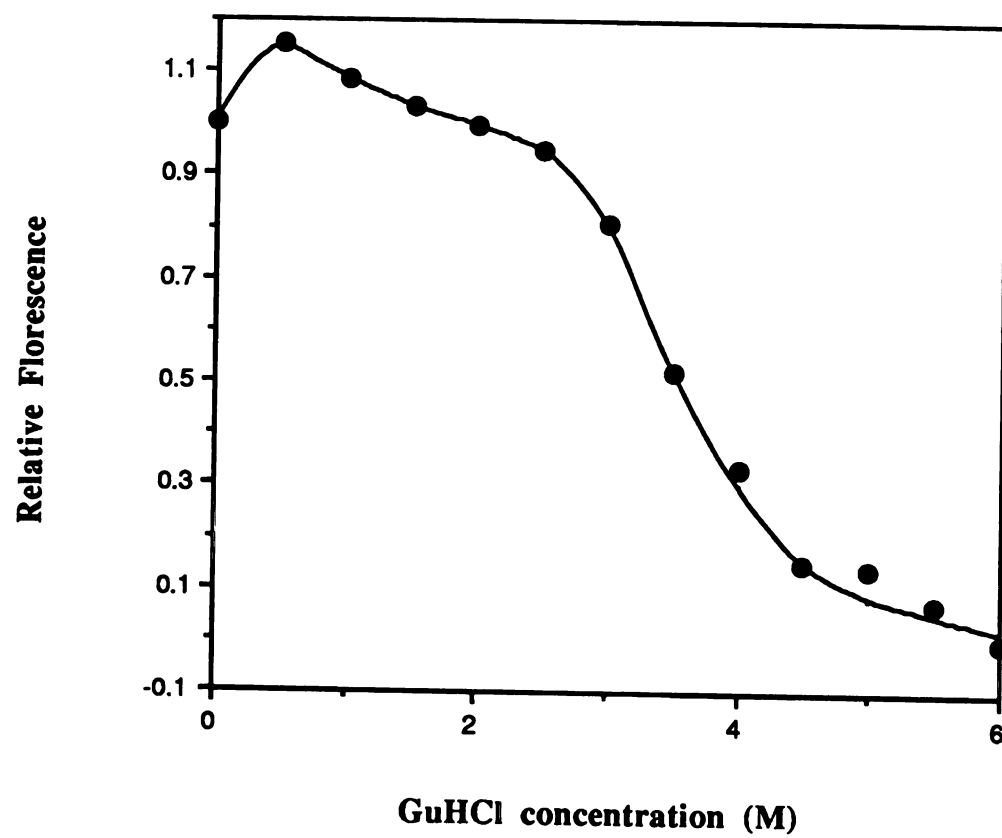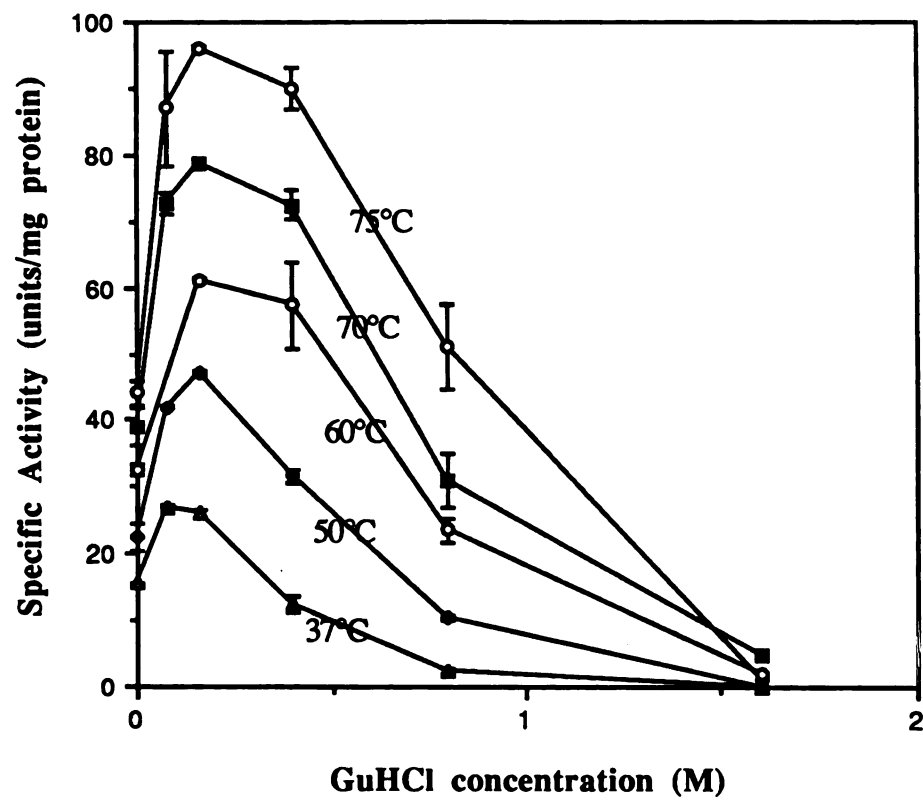
Figure 6

GuHCl and temperature dependence of 2° ADH activity. 2° ADH catalyzed propan-2-ol

oxidation rates were determined at 37°C (△), 50°C (◆), 60°C (●), 70°C (□) and 75°C (○) in

GuHCl concentrations from 0.0 to 1.6 M.

DISCUSSION

These data provide quantitative evidence that the *T. ethanolicus* 2° ADH is a highly thermophilic and thermostable enzyme with catalytic activity that is also robust to high levels of chemical denaturants. The enzyme was rapidly thermoinactivated only at temperatures above 90°C. Enzyme thermoinactivation was modeled by a pseudo-first order rate equation suggesting that the rate determining step was unimolecular with respect to protein. The large thermoinactivation transition state enthalpy and entropy changes are consistent with a significant protein structural change during this rate limiting step. Enzyme activity loss upon heating was accompanied by precipitation, but the soluble enzyme remaining after partial sample inactivation retained complete activity, mechanistically linking enzyme inactivation and precipitation. Finally, the increased GuHCl concentration necessary for enzyme rate enhancement at higher temperatures was also seen with certain inorganic salts arguing that this 2° ADH catalytic property results from specific ionic and not general hydrophobic interactions.

The highly thermophilic *T. ethanolicus* 2° ADH has a projected half-life of 2 months at 70°C which compares to the complete inactivation of the structurally similar mesophilic *Clostridium beijerinckii* enzyme [23,33] is also more stable than the *T. brockii* 2° ADH [33a] that differs by only 3 amino acids. Tomazic and Klibanov proposed that enzyme molecules reversibly enter structural states that can lead to essentially irreversible protein aggregation and inactivation [34]. With respect to protein, this process involves an initial intramolecular structural change followed by an intermolecular aggregation step. Thermoinactivation of the *T. ethanolicus* enzyme follows a pseudo-first order rate law, suggesting that the rate is independent of the protein concentration as reported for α-amylase [34] and D-xylose isomerase [35]. Thus, the process appears to be unimolecular with respect to 2° ADH protein. The soluble enzyme recovered after incomplete enzyme thermoinactivation retained specific activity comparable to the unheated enzyme, implying

complete inactivation of part of the enzyme population and not partial inactivation of the entire enzyme population. These data suggest that catalytically compromised 2° ADH intermediate structures did not persist, consistent with a rapid second step in irreversible thermoinactivation (precipitation). The absence of detectible enzyme activity in the precipitated protein demonstrated that loss of catalytic activity either accompanied or preceded protein precipitation. Therefore, the observed unimolecular *T. ethanolicus* 2° ADH thermoinactivation rate limiting step preceding precipitation of structurally scrambled protein molecules fits the two step Tomazic and Klibanov model [34] for protein thermoinactivation with a slow native to non-native enzyme structural transition followed by rapid aggregation of thermoinactivated enzyme. Identifying the exact molecular mechanism of irreversible 2° ADH thermoinactivation will require precise measurement of inactivated protein molecular mass or detection of covalent destruction products like ammonia to determine whether this rate limiting step involves covalent protein modification (eg., deamidation) or partial loss of native enzyme structure.

The inactivation slow step transition state thermodynamics were evaluated to determine the stabilizing energy for the 2° ADH. The transition state enthalpy, almost 20-fold that reported for enzymatic catalysis [23], is 100-fold greater than the expected molecular kinetic energy ($\sim$RT; where R = 8.341 J mol^{-1} K^{-1}, and T is the system temperature in Kelvin). Calculating the $\Delta S^{\ddagger}_{stability}$ from absolute rate theory requires the assumption that the inactivated, partly unfolded protein is an intermediate following the slow step transition state which is itself in rapid equilibrium with the native protein structure. Assuming this mechanism, the 2° ADH is stabilized by a relatively small energy difference (110 kJ mol^{-1}) between large enthalpic (440 kJ mol^{-1}) and entropic (330 kJ mol^{-1} at 393 K) contributions as described for other proteins [36]. The $\Delta H^{\ddagger}$ therefore, appears to be an effective barrier to 2° ADH thermoinactivation opposing a large destabilizing transition state entropy change.

Unfolding of *T. ethanolicus* 2° ADH was measured to examine the nature of this slow, inactivating 2° ADH structural transition. The $T_{m[Gu]}$ data from absorbance measurements predicted that a GuHCl concentration of 3.75 M would yield an enzyme $T_{m[Gu]}$ value of 25°C. This predicted value was corroborated by Trp fluorescence measurements. 2° ADH melting measured by both absorbance and fluorescence indicated an essentially linear $T_{m[Gu]}$ dependence on GuHCl concentration. Extrapolating this linear dependence of enzyme $T_{m[Gu]}$ on GuHCl concentration to zero GuHCl predicts a T_m value of 115 ± 5.8°C. This temperature is 20°C higher than the midpoint temperature for 2° ADH precipitation (~95°C) arguing that the precipitating protein is not significantly unfolded. Consequently, the high thermal stability of *T. ethanolicus* 2° ADH is associated with its high resistance (i.e., rigidity) to chemically and thermally induced unfolding.

Biophysical measurements have indicated that enzymes are more rigid at low temperatures than at high temperatures within their active temperature ranges [25,36,37]. However, the hypothesis that poor low temperature enzyme activity results from excessive structural rigidity [32] has been challenged based on the general Arrhenius theory adherence of enzyme catalytic rate temperature dependence [24,38] and specifically based on the temperature activity behavior of the *T. ethanolicus* 2° ADH [25]. Unlike other enzymes, the 2° ADH was totally folded at 90°C where it was optimally active. Here we attempted to alter 2° ADH rigidity by the addition of GuHCl. The additive denaturing effects of GuHCl and temperature on *T. ethanolicus* 2° ADH $T_{m[Gu]}$ predict that at higher temperatures lower GuHCl concentrations should be required to reduce enzyme rigidity, thus lower GuHCl concentrations should result in optimal enzyme flexibility and activity. While added GuHCl did significantly enhance enzyme activity, the effect was equally pronounced at low and at high temperatures. The slightly increased denaturant concentration corresponding to maximal 2° ADH activity at higher temperatures is inconsistant with the lower denaturant concentration predicted to yield optimal enzyme rigidity at higher temperatures. Similar 2° ADH activity enhancement was seen with low

concentrations of inorganic salts which are not strong protein denaturants, supporting the conclusion that ionic interactions and not optimal rigidity were responsible for the observed increase in the catalytic rate.

Industrial biocatalysts must have activities robust to a broad range reaction conditions. Both chemical and thermal stability are therefore important properties of a potential industrial chiral catalyst. The *T. ethanolicus* 2° ADH is active from 30°C to over 90°C with active half-lives greater than 1 hr to temperatures exceeding 90°C. GuHCl concentrations exceeding 1.0 M were necessary to eliminate enzyme activity and this thermophilic 2° ADH was also highly active in ethanol concentrations greater than 2.0 M (data not shown). Therefore, the extremely high thermal and chemical stabilty of *T. ethanolicus* 2° ADH activity, demonstrates that this is an extremely robust enzyme catalyst. As a chiral chemical catalyst, the effect of temperature on *T. ethanolicus* 2° ADH chirality has been shown to be reaction specific [39]. The effect of chemical denaturants on 2° ADH product chirality however, remains to be determined.

ACKNOWLEDGMENTS

148

REFERENCES

1   Whitesides, G. M. and Wong, C.-H. [1983] Enzymes as catalysts in organic synthesis. *Aldrichimica Acta* 16, 27-34

2   Rosazza, J. P. N. [1995] Biocatalysis, microbiology, and chemistry: The power of positive linking. *Amer. Soc. Microbiol. News* 61, 241-245

3   Meyer, H.-P. [1991] Microbiology in the contemporary organic chemical industry. *Biol. Ferm. Eng.* 8, 602-606

4   Wong, C.-H. [1989] Enzymatic catalysts in organic synthesis. *Science* 244, 1145-1152

5   Jones, J. B. [1986] Enzymes in organic synthesis. *Tetrahedron* 42, 3351-3403

6   Irwin, J. B., Lok, K. P., Huang K. W. C. and Jones, J. B. [1978] Enzymes in organic synthesis. Influence of substrate structure on rates of horse liver alcohol dehydrogenase-catalyzed oxidoreductions. *J. Chem. Soc. Perkin I* 12, 1636-1641

7   Bradshaw, C. W., Hummel, W. and Wong, C.-H. [1992] *Lactobacillus kefir* alcohol dehydrogenase: a useful catalyst for synthesis. *J. Org. Chem.* 57, 1532-1536

8   Bradshaw, C. W., Fu, H., Shen, G.-J. and Wong, C.-H. [1992] A *Pseudomonas sp.* alcohol dehydrogenase with broad substrate specificity and unusual stereospecificity for organic synthesis. *J. Org. Chem.* 57, 1526-1532

9   Hummel, W. and Kukla, M. R. [1989] Dehydrogenases for the synthesis of chiral compounds. *Eur. J. Biochem.* 184, 1-13

10   May, S.W. and Padgette, S.R. [1983] Oxidoreductase enzymes in biotechnology: current status and future potential. *Bio/Technology* 1, 677-686

11   Faber, K. and Franssen, M. C. R. [1993] Prospects for the increased application of biocatalysts in organic transformations. *Trends Biotechnol.* 11, 461-470

12   Brändén, C.-I., Jörnvall, H., Eklund, H. and Fururgren, B. [1975] Alcohol dehydrogenases, in *The enzymes*, 3rd ed., vol. XI part A (P. D. Boyer, ed.) pp. 105-106. Academic Press, New York

13   Klibanov, M. [1983] Biotechnological potential of the enzyme hydrogenase. *Proc. Biochem.* 18, 13-23

14   Zong, M.-H., Fukui, T., Kawamoto, T. and Tanaka, A. [1991] Bioconversion of organosilicon compounds by horse liver alcohol dehydrogenase: the role of the silicon atom in enzymatic reactions. *Appl. Microbiol. Biotechnol.* 36, 40-43

15   Nagata, Y., Maeda, K. and Scopes, R. K. [1992] NADP-linked alcohol dehydrogenases from extreme thermophiles: Simple affinity purification schemes, and comparative properties of the enzymes from different strains. *Bioseparation* 2, 353-362

16   Keinan, E., Hafeli, E. K., Seth, K. K. and Lamed, R. L. [1986] Thermostable
      enzymes in organic synthesis. 2. Asymmetric reduction of ketones with alcohol
      dehydrogenase from *Thermoanaerobium brockii. J. Am. Chem. Soc.* **108**,
      162-169

17   Burdette, D. S. and Zeikus, J.G. [1994] Purification of acetaldehyde
      dehydrogenase and alcohol dehydrogenases from *Thermoanaerobacter ethanolicus*
      39E and characterization of the secondary-alcohol dehydrogenase (2° ADH) as a
      bifunctional alcohol dehydrogenase-acetyl-CoA reductive thioesterase. *Biochem. J.*
      302, 163-170

18   Zheng, C., Pham, V. T. and Phillips, R. S. [1992] Asymmetric reduction of
      ketoesters with alcohol dehydrogenase from Thermoanaerobacter ethanolicus.
      Bioorg. Med. Chem. Let. **2**, 619-622

19   Keinan, E., Seth, K.K. and Lamed, R. [1986] Organic synthesis with
      enzymes. 3. TBADH-catalyzed reduction of chloro ketones. Total synthesis of (+)-
      (S,S)-(*cis*-6-methyltetrahydropyran-2-yl)acetic acid: a civet constituent. *J. Am.
      Chem. Soc.* **108**, 3474-3480

20   Keinan, E., Seth, K.K., Lamed, R., Ghirlando, R. and Singh, S.P. [1990]
      thermostable enzyme in organic synthesis, 4. TBADH-catalyzed preparation of
      bifunctional chirons. Total synthesis of S-(+)-Z=tetradec-5-en-13-olide.
      Biocatalysis **4**, 1-15

21   Peretz, M. and Burstein, Y. (1989) Amino acid sequence of alcohol
      dehydrogenase from the thermophilic bacterium *Thermoanaerobium brockii*.
      Biochemistry **28**, 6549-6555

22   Lamed, R.J., Keinan, E. and Zeikus, J.G. [1981] Potential applications of
      an alcohol-aldehyde/ketone oxidoreductase from thermophilic bacteria. *Enz.
      Microbiol. Technol.* **3**, 144-148

23   Burdette, D.S., Vieille, C. and Zeikus, J.G. [1996] Cloning, expression, and
      biochemical characterization of the thermophilic secondary-alcohol dehydrogenase
      from *Thermoanaerobacter ethanolicus* 39E. *Biochem. J.*, in the press

24   Vieille, C., Burdette, D. S. and Zeikus, J. G. [1996] Thermozymes, in
      Biotechnology Annual Reviews, vol. 2, (M. R. El Geweley, ed.) pp. 1-83.
      Elsevier Press, Amsterdam, Neth.

25   Böhm, G. and Jaenicke, R. [1994] Relevance of sequence statistics for the
      properties of extremophilic proteins. *Int. J. Peptide Protein Res.* **43**, 97-106

26   Brooks, S. [1992] A simple computer program with statistical tests for the analysis
      of enzyme kinetics. *Biotechniques* **13**, 906 - 911

27   Chang, R. [1977] Physical chemistry with applications to biological systems.
      MacMillan Pub. Co., NY

28   Hermans, J. and Scheraga, H. A. [1961] Structural studies of ribonuclease V.
      Reversible changes of configuration. *J. Am. Chem. Soc.* **83**, 3283-3292

29   Tiktopulo, E. I. and Privalov, P. L. [1974] Heat denaturation of ribonuctease. *Biophys. Chem.* **1**, 349-357

30   Klump, H., Di Ruggiero, J., Kessl, M., Park, J.-B., Adams, M. W. W. and Robb, F. T. [1992] Glutamate dehydrogenase from the hyperthermophile *Pyrococcus furiosis J. Biol. Chem.* **267**, 22681-22685

31   Remington, R. D. and Schork, M. A. [1985] Statistics with applications to the biological and health sciences, Prentice-Hall, Engelwood Cliffs, NJ

32   Wrba, A., Schweiger, A., Schultes, V., Závodszky, P. and Jaenicke, R. [1990] Extremely thermostable D-glyceraldehyde-3-phosphate dehydrogenase from the eubacterium *Thermatoga matitima. Biochemistry* **29**, 7584-7592

33   Ismaiel, A. A., Zhu, C.-X., Colby, G. D. and Chen, J.-S. [1993] Purification and characterization of a primary-secondary alcohol dehydrogenase from two strains of *Clostridium beijerinckii. J. Bacteriol.* **175**, 5097-5105

33a   Lamed, R. J. and Zeikus, J. G. [1981] Novel NADP-linked alcohol aldehyde/ketone oxidoreductase in thermophilic ethanologenic bacteria. *Biochem J.* **195**, 183-190

34   Tomazic, S.J., and Klibanov, A.M. [1988] Mechanisms of irreversible thermal inactivation of *Bacillus* α-amylases. *J. Biol. Chem.* **263**, 3086-3091

35   Meng, M., Bagdasarian, M. and Zeikus, J. G. [1993] The role of active-site aromatic and polar residues in catalysis and substrate discrimination by xylose isomerase. *Biotechnology* **11**, 1157-1161

36   Jaenicke, R. [1991] Protein stability and molecular adaptation to extreme conditions. *Eur. J. Biochem.* **202**, 715-728

37   Fontana, A. [1990] How nature engineers protein (thermo) stability, in *Life under extreme conditions: Biochemical adaptation* (G. di Prisco, ed.) pp. 89-113. Springer-Verlag, Heidelberg.

38   Tchemajenko, V., Vieille, C., Burdette, D.S., and Zeikus, J.G. [1996] Physicobiochemical studies of xylose isomerases from thermophilic and mesophilic microorganisms. Manuscript in preparation

39   Phillips, R. S., Zheng, C., Pham, V. T., Andrade, F. A. C. and Andrade, M. A. C. (1994) Effects of temperature on the ctereochemistry of enzyme reactions. *Biocatalysis* **10**, 77-86

Chapter IV

Mutagenic and Biophysical Analysis of *Thermoanaerobacter ethanolicus* Secondary-Alcohol Dehydrogenase Activity

Prepared for submission to the Journal of Biological Chemistry

ABSTRACT

The *Thermoanaerobacter ethanolicus* 39E *adhB* gene encoding the secondary-alcohol dehydrobenase (2° Adh) was overexpressed in *Escherichia coli* (DH5α) at ~15% of total protein. Recombinant enzyme was purified in high yield (67%) by simple heat treatment at 85°C and ammonium sulfate precipitation. Purified site directed 2° Adh mutants Cys37 to Ser, His59 to Asn, Asp150 to Asn, Asp150 to Glu, and Asp150 to Cys were analyzed to test the peptide sequence comparison based predictions of amino acids responsible for putative catalytic Zn binding. X-ray absorption spectrometry confirmed the presence of a protein bound Zn atom (spectral transition at ~9662 eV) with a $ZnS_1(N-O)_{3-4}$ coordination sphere. Induction coupled plasma emmission spectrometry measured 0.48 ± 0.021 Zn atoms per wild type 2° Adh subunit. The Cys37 to Ser, His59 to Asn, and Asp150 to Asn mutant enzymes bound only 0.11, 0.13, and 0.33 Zn per subunit, respectively, suggesting that these residues were involved in Zn liganding. The Asp150 to Glu and Asp150 to Cys mutants retained 0.47 and 1.2 Zn atoms per subunit, suggesting that an electron rich sidechain moiety at this position preserves the bound Zn. All five mutant enzymes had ≤ 3% of wild type catalytic activity, indicating that the *T. ethanolicus* 2° Adh requires a catalytic Zn atom. Also, the His59 and Asp150 mutations altered 2° Adh affinity for propan-2-ol over a 140-fold range while the overall change in affinity for ethanol spanned a range of only 7-fold, supporting the importance of the metal in 2° ADH substrate binding. The lack of significant changes in cofactor affinity due to these catalytic Zn ligand mutations suggested that 2° ADH substrate and cofactor binding are structurally distinct. Altering Gly198 to Asp reduced the enzyme specific activity 2.7-fold, increased the Km_{app} for $NADP^+$ 225-fold, and decreased the Km_{app} for NAD^+ 3-fold, supporting the prediction that this 2° Adh binds cofactor in a common nicotinamide cofactor binding motif, a Rossmann fold. Therefore, these mutant enzyme data indicate that, unlike the liver 1° ADH, the NADP(H) linked *T. ethanolicus* 2° Adh binds its catalytic Zn atom using a

novel Cys-His-Asp motif, it does not bind a structural Zn atom, and it uses a Rossmann fold to bind nicotinamide cofactor with Gly198 significantly responsible for its NADP(H) specificity.

INTRODUCTION

Alcohol dehydrogenases (ADHs), central to prokaryotic and eukaryotic

metabolism, are also potential biocatalysts for chiral chemical production [Keinan *et al.*,

1986a; Keinan *et al.*, 1986b; Hummel, 1990; Keinan *et al.*, 1990, Bradshaw *et al.*; 1992a,

Bradshaw *et al.*;1992b]. The functionally and presumably structurally similar ADHs are

classified as primary or secondary based on their higher catalytic efficiencies toward

primary or secondary alcohols. These typically homodimeric or homotetrameric enzymes

are almost exclusively Zn containing metalloenzymes [Brändén *et al.*, 1975] that use

NAD(H) (EC 1.1.1.1), NADP(H) (EC 1.1.1.2), or both (EC 1.1.1.71) as cofactor. Fe

linked [Scopes, 1983] and ferredoxin F_{420}-dependent [Bleicher and Winter, 1991] 1°

ADHs have also been reported. The catalytic Zn's role in liver 1° ADH activity has been

proposed to involve binding directly to the substrate oxygen atom, facilitating transfer of a

hydride between the adjacent substrate carbon atom and the nicotinamide cofactor [Brändén

et al., 1975]. While 1° ADHs from diverse sources have been extensively characterized

[Brändén *et al.*, 1975], little is known about the molecular basis for 1° ADH substrate

specificity. Far less is is known about 2° ADH structure-function, with relatively few

reported studies on enzyme purification and characterization [Lamed and Zeikus, 1981;

Steinbüchel *et al.*, 1984; Bryant *et al.*, 1988; Ismaiel *et al.*, 1993; Burdette and Zeikus,

1994] and a complete lack of exact structural or site directed mutagenic information.

Because of the 2° ADHs' central metabolic roles [Lovitt *et al.*, 1984; Burdette and Zeikus

1994] and their potential biotechnological value [Lamed *et al.*, 1981], structure-function

analysis of catalysis and substrate specificity would contribute significantly to the basic

understanding of enzyme function and to the design of biocatalyst specificity.

The exact 3-dimensional structure of the 1° ADH from horse liver has been

determined [Eklund *et al.*, 1974], indicating that the catalytic Zn atom was bound by two

Cys and a His residue and that the nicotinamide cofactor was bound in a Rossman fold

[Brändén *et al.*, 1975; Rossmann and Argos, 1976]. Research to date has failed to identify ADH active site amino acid residues which are specifically responsible for substrate binding. Instead, substrate appears to bind the catalytic metal directly [Dunn and Hutchinson, 1973; Jacobs *et al.*, 1974; McFarland *et al.*, 1974]. The presumed similarity between 1° and 2° ADH structures based on shared catalytic properties provided the rationale for sequence based comparisons between these enzymes [Jendrossek *et al.*, 1988; Peretz and Burstein, 1989; Burdette *et al.*, 1996]. Peptide sequence alignments of thermophilic and mesophilic 1° and 2° ADHs were used to hypothesize that 2° ADHs similarly bind nicotinamide cofactor in a Rossman fold but that 2° ADHs lack a structural Zn binding loop and they use a unique Cys-His-Asp motif for catalytic metal liganding [Burdette *et al.*, 1996]. Chemical modification experiments have implicated Zn, Cys, and His in *Thermoanaerobium brockii* and *Thermoanaerobacter ethanolicus* 39E 2° ADHs catalysis [Lamed and Zeikus, 1981; Burdette *et al.*, 1996] but validation of the predicted Zn liganding motif has awaited assessment by biophysical analysis and site directed mutagenesis. Furthermore, the strong similarities among 2° ADH peptide sequences, kinetic parameters, subunit compositions, and molecular masses [Nagata *et al.*, 1992; Burdette *et al.*, 1996] suggest that the results of *T. ethanolicus* 2° ADH catalytic structure-function analysis will be generally applicable to other 2° ADHs.

The research reported here describes the construction of a recombinant overexpression and purification system providing high yields of homogeneous enzyme. The results of site directed mutagenesis to test sequence comparison based predictions of amino acid residues important for *T. ethanolicus* 39E 2° ADH catalysis are also described. Predictions of catalytic Zn binding ligands and the amino acid responsible for NADP(H) cofactor specificity are assessed from site directed mutation effects on enzyme Km_{app}, $Vmax_{app}$, and Zn binding. Finally, the mutant enzyme kinetic data on 2° ADH mutants and exact structural information on 1° ADH are used to create a hypothetical working model for

the 2° ADH active site and to distinguish key structure-function differences between 2° and 1° ADHs.

EXPERIMENTAL PROCEDURES

Chemicals and reagents - All chemicals were of at least reagent/molecular biology grade. Oligonucleotide synthesis and amino acid sequence analysis were performed by the Macromolecular Structure Facility (Department of Biochemistry, Michigan State University). The kanamycin resistance GenBlock (*Eco*RI) DNA cartridge used in expression vector construction was purchased from Pharmacia (Uppsala, Sweden) [Oka *et al.*, 1981]. DNA for sequencing was isolated using the Wizard Miniprep kit (Promega; Madison, WI).

Media and strains - *Escherichia coli* (DH5α) containing the 2° ADH recombinant plasmids were grown in rich complex medium (20 g l⁻¹ tryptone, 10 g l⁻¹ yeast extract, 5 g l⁻¹ NaCl) at 37°C in the presence of 25 µg ml⁻¹ kanamycin and 100 µg ml⁻¹ ampicillin.

Mutagenesis - All DNA manipulations were performed using established protocols [Sambrook *et al.*, 1989; Ausubel *et al.*, 1993]. Point mutations were introduced into the *adhB* gene by PCR [Ausubel *et al.*, 1993] using the *T. ethanolicus* 39E 2° ADH gene clonal plasmid pADHB25-kan [Burdette *et al.*, 1996] as template DNA. An oligonucleotide primer (KA4 N-end) was synthesized to bind the noncoding strand and included a *Kpn*I restriction enzyme site, the native *adhB* gene ribosome assembly site, and the initiation codon for the *adhB* gene (Table I). An oligonucleotide primer (KA4 C-end) was synthesized to bind the coding strand, it included the compliment of the *adhB* termination codon and an *Apa*I restriction enzyme site. Complimentary 30-45 base oligonucleotide primers that contained the mutated bases were used in conjunction with two KA4 end primers to amplify the N-terminal and C-terminal segments of the *adhB* gene. PCR

Table I

Oligonucleotide primers for PCR amplification of mutant adhB gene DNA

List of PCR primers used in the construction of 2° Adh overexpression plasmids and genes encoding mutant enzymes. The primers for the 5' and 3' end of the adhB gene are listed with the mutated bases underlined and the restriction sites in italics. Altered bases in the mutation encoding primers are underlined.

Mutation	Primer sequence
KA4	
5'-end	5'CGG*GGTACC*CCGTATTTTAGGAGGTGTTTAATGATGAAAGG3'
3'-end	5'CAGTCC*GGGCCC*TTATGCTAATATTACAACAGGTTTG3'
1Met1	
5'-end	5'CGG*GGTACC*CCGTATTTTAGGAGGTGTTTATTAATGAAAGG3'
3'-end	5'CAGTCC*GGGCCC*TTATGCTAATATTACAACAGGTTTG3'
Cys37 to Ser	
coding	5'GCTGTGGCCCCTTCCACTTCGGACATTCATACC3'
noncoding	5'GGTATGAATGTCCGAAGTGGAAGGGGC3'
His59 to Asn	
coding	5'CATGATACTCGGTAACGAAGCTGTA3'
noncoding	5'CTTCACCTACAGCTTCGTTACCGAG3'
Asp150 to Asn	
coding	5'GGAAGCTGCACTTATGATTCCCAATATGATGACCACTGG3'
noncoding	5'GCTCCGTGAAAACCAGTGGTCATCATATTGGGAATCATA3'
Asp150 to Cys	
coding	5'GGAAGCTGCACTTATGATTCCCTGTATGATGACCACTGG3'
noncoding	5'GCTCCGTGAAAACCAGTGGTCATCATACAGGGAATCATAAG3'
Asp150 to Glu	
coding	5'GGAAGCTGCACTTATGATTCCCGAAATGATGACCACTGG3'
noncoding	5'GCTCCGTGAAAACCAGTGGTCATCATTTCGGGAATCATA3'
Gly 198 to Asp	
coding	5'GAATTATTGCCGTAGACAGTAGACCAGTTTGTG3'
noncoding	5'CACAAACTGGTCTACTGTCTACGGCAATAATTC3'

syntheses of partial and complete mutated genes were performed using the Taqplus,

exonuclease containing polymerase (Strategene; La Jolla, CA). All clones were expressed

in pBluescriptII KS(+) with a kanamycin resistance cartridge introduced into the polylinker

*Eco*RI site. Mutations were verified by DNA sequencing using the method of Sanger *et al.*

(1977).

Enzyme purification - The recombinant enzyme was purified from *E. coli* (DH5α)

aerobically. The pelleted cells from batch cultures were resuspended (0.5 g wet wt. ml^{-1})

in buffer A (50 mM Tris:HCl [pH 8.0], 5 mM DTT, and 10 μM ZnCl$_2$) containing 3 μg

ml^{-1} lysozyme. The resuspended cells were incubated on ice for 30 min, frozen in N$_2$

(liq), thawed and centrifuged for 30 min at 15 000g. The clarified lysate was incubated at

85°C for 15 min, cooled on ice for 30 min, and then centrifuged for 30 min at 15 000g.

We have recently shown that *T. ethanolicus* 2° Adh displayed a half-life of 1.2 days at 80°C

(see chapter 3). Ammonium sulfate {50% (wt/vol)} was added to the supernatant and

stirred at 4°C for 30 min. The 2° ADH was recovered from the supernatant after

centrifugation for 30 min at 15 000g and precipitated with stirring at 4°C for 30 min in 70%

(wt/vol) ammonium sulfate. After centrifugation for 30 min at 15 000g, the purified 2°

ADH was resuspended in buffer A and stored at 4°C.

Zn binding and molecular mass determination - EXAFS determination of 2° ADH

Zn coordination was performed by Dr. Bob Scott (Univ. of Georgia; Athens, GA).

Enzyme samples were concentrated to 1-10 mg ml^{-1} in 50 mM Tris:HCl (pH 8.0) and

stored in N$_2$ (liq) prior to analysis. 2° ADH samples were dialyzed against 50 mM

NH$_4$HCO$_3$ (pH 8.0) containing 100 g l^{-1} Chelex resin (BioRad; Hercules, CA) for 12 hr to

remove unbound or advantitiously bound metals. The metal cleared samples were

concentrated to between 1.2 and 12 mg ml^{-1} using centricon 30 ultrafiltration units

(Amicon; Beverly, MA) for induction coupled plasma emission spectrometry (ICP)

analysis (Chemical Analysis Laboratory at the University of Georgia, Athens). Subunit

molecular mass values were determined by comparison with standards (BioRad; Hercules, CA) using SDS-PAGE (12% polyacrylamide).

Enzyme kinetics - The standard 2° ADH activity assay was defined as NADP+ reduction coupled to propan-2-ol oxidation at 60°C as previously described [Burdette and Zeikus, 1994]. The enzyme was incubated at 55°C for 15 min prior to activity determination unless otherwise indicated. Tris buffer pH was adjusted at 25°C to be pH 8.0 at 60°C (thermal correction factor = -0.031 ΔpH °C^{-1}). Assays to determine Km$_{app}$ and Vmax$_{app}$ were conducted at 60°C with substrate concentrations between 20xKm$_{app}$ and 0.2xKm$_{app}$. Kinetic parameters were calculated from nonlinear best fits of the data to the Michaelis-Menten equation using Kinzyme software [Brooks, 1992] on an IBM PC. Protein concentrations were measured using the bicinchoninic acid (BCA) procedure (Pierce; Rockford, IL).

RESULTS

Enzyme overexpression - *T. ethanolicus* 2° ADH was purified from *E. coli* (DH5α) harboring the pADHBKA4-kan clonal construct (Table II) yielding 67% recovery of recombinant enzyme expressed at ~15% of total soluble protein without induction (Fig. I). This construct contained the native ribosome assembly site and translational initiation codons but removed the native transcriptional promoter regions, placing the gene under transcriptional control of the pBluescriptKS(+) T3 promoter. The recombinant protein population included 56 ± 11% enzyme with a duplicated N-terminal methionine residue (MMKGFA...) while 44 ± 11% of the protein possessed a single methionine residue at the N-terminus. Alteration of the "ATGATG" translational initiation site to "TTAATG" in the pADHB1M1-kan expression plasmid eliminated the synthesis of 2° ADH with two N-terminal methionine residues but did not alter the level of enzyme expression compared to

Table II
Purification of the recombinant wild type 2° Adh expressed from plasmid pADHBKA4-kan

Purification procedure for recombinant *T. ethanolicus* 39E 2° Adh. Resuspended cells (0.5 g wet wt. ml^{-1} in 50 mM Tris:HCl pH 8.0, 5 mM dithiothreitol, and 3 µg ml^{-1} lysozyme) were incubated 30 min at 4° C, frozen in N$_2$ (liq), thawed, and clarified by centrifugation. This cell extract was incubated at 85 °C for 15 min, removed to ice for 30 min, and clarified by centrifugation. Purified protein was recovered in the centrifuged pellet from a 50% (wt./vol.) to 70% (wt./vol.) ammonium sulfate precipitation at 4 °C.

Step	Total protein	Total activity	Protein concentration	Specific activity	Purification	Yield
	mg	units	mg ml^{-1}	units mg^{-1}	fold	(%)
Cell extract	59	270	8.7	4.8		100
heat treatment	25	280	4.6	11	2.3	100
(NH$_4$)$_2$SO$_4$ precipitation	4.2	180	3.4	44	9.2	67

that seen for pADHBKA4-kan (Fig. I) or the enzyme specific activities toward ethanol and propan-2-ol (Table III). The Km$_{app}$ values toward propan-2-ol, ethanol, and NADP$^+$ were 1.1 ± 0.22 mM, 53 ± 9.0 mM, and 17 ± 2.6 µM, for protein expressed from pADHBKA4-kan and 0.87 ± 0.32 mM, 37 ± 4.2 mM, and 8.5 ± 1.4 µM for enzyme expressed from pADHB1M1-kan.

Catalytic Zn liganding - Chemical modification experiments have implicated Zn, Cys, and His in *T. ethanolicus* 2° ADH activity. Peptide alignments identified Cys37, His59, and Asp150 as potential catalytic Zn ligands [Burdette *et al.*, 1996]. EXAFS analysis of the native and wild type recombinant enzyme suggested the presence of a single type of specifically bound Zn atom in the enzyme (Fig. II). The transition at 9660-9665 eV in Fig. IIa is characteristic of Zn. The Fourier transformed data (Fig. IIb) is best fit by a coordination sphere of $ZnS_1(N,O)_{3-4}$. The sulfur signal is consistent with that of a cysteine residue. The optimal data fit also indicated a single imidazole nitrogen ligand, leaving 2-3 potential oxygen ligands which could be from an amino acid such as Asp, water, or a hydroxide ion.

Mutant *adhB* genes encoding amino acid substitutions at residues 37, 59, and 150 were constructed by PCR (Fig. III) to test the involvement of these amino acids in catalytic Zn binding. Mutant proteins were expressed at levels similar to the wild type enzyme and were purified to homogeniety as seen on SDS-PAGE (Fig IV). Sequencing the mutated section of the *adhB* gene confirmed the identity of each mutant. ICP analysis indicated that the enzyme expressed from the wild type gene (pADHBKA4-kan) and the single N-terminal Met mutant enzyme bound 0.47 and 0.50 Zn ions per subunit, respectively. The Cys37 to Ser and His59 to Asn mutant enzymes contained 0.11 and 0.13 Zn per subunit, respectively. Altering Asp150 to Glu preserved binding of 0.47 Zn per subunit while the Asp150 to Asn mutant bound 0.33 Zn per subunit. Changing the putative Asp150 ligand to Cys increased binding to 1.2 Zn per subunit. Finally,the Gly198 to Asp mutant enzyme also bound 0.55 Zn per subunit. Mutant and wild type enzymes treated for Zn analysis by

Figure 1. SDS-PAGE of wild type recombinant 2° Adhs. Lanes: 1, clone 1M1 cell extract; 2, clone 1M1 dialized ammunium sulfate precipitate; 3, clone KA4 cell extract; 4, clone KA4 post heat treatment; 5, clone KA4 dialized ammunium sulfate precipitate. The Mr values for the molecular mass markers flanking the sample lanes are indicated on the left. Proteins were stained with Coomassee brilliant blue R-250.
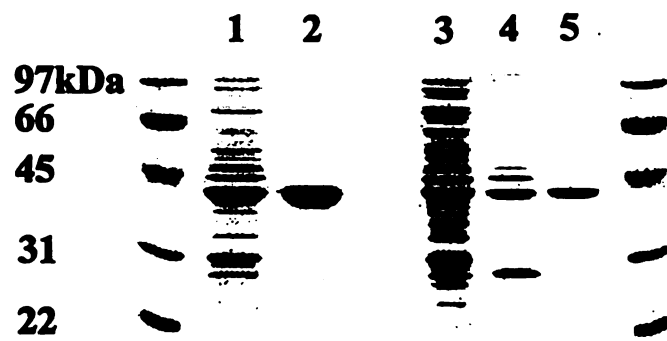
Table III

*Effect of site specific amino acid substitutions on T. ethanolicus 2° Adh activity*

Kinetic parameters for mutant enzyme catalysis. Enzyme initial rates were measured by the increase in absorbance due to $NAD(P)^+$ reduction at 60 °C (ε_{340nm} = 6.22 mM^{-1} cm^{-1}). Km_{app} and $Vmax_{app}$ values were determined by nonlinear curve fits to initial rate data from assays using $0.2 \times Km_{app}$ to $20 \times Km_{app}$ initial varied substrate concentrations with the cosubstrate at $10 \times Km_{app}$. Propan-2-ol and ethanol kinetic values were determined with $NADP^+$ cofactor except in the case of the G198D mutant which displayed higher $Vmax_{app}$ values using NAD^+ as the cofactor. $NADP^+$ kinetic values were determined with propan-2-ol as the cosubstrate in all cases. The approximate Km_{app} and $Vmax_{app}$ for ethanol oxidation by the C37S mutant were determined using only ethanol concentrations between 350 mM and 700 mM. Catalytic efficiency was defined as $Vmax_{app}/Km_{app}$.

Enzyme	propan-2-ol			Ethanol			NADP+		
	$Vmax_{app}$ $U\,mg^{-1}$	Km_{app} mM	Catalytic efficiency $ml\,min^{-1}\,mg^{-1}$	$Vmax_{app}$ $U\,mg^{-1}$	Km_{app} mM	Catalytic efficiency $ml\,min^{-1}\,mg^{-1}$	$Vmax_{app}$ $U\,mg^{-1}$	Km_{app} mM	Catalytic efficiency $ml\,min^{-1}\,mg^{-1}$
wt	68	1.1	$6.2e^{-2}$	19	53	$3.6e^{-4}$	72	0.016	4.5
C37S	0.48	6.9	$7.0e^{-5}$	a0.042	a1100	a$3.8e^{-8}$	0.48	0.0097	$4.9e^{-2}$
D150C	2.0	1.4	$1.4e^{-3}$	0.68	110	$6.2e^{-6}$	2.0	0.00094	2.1
D150E	0.20	32	$6.2e^{-6}$	0.028	180	$1.6e^{-7}$	0.19	0.011	$1.7e^{-2}$
D150N	0.60	0.20	$3.0e^{-3}$	0.044	230	$1.9e^{-7}$	0.64	0.037	$1.7e^{-2}$
H59N	0.26	6.6	$3.9e^{-5}$	0.017	38	$4.5e^{-7}$	0.26	0.0060	$4.3e^{-2}$
G198D	25	1.3	$1.9e^{-2}$	7.9	110	$7.2e^{-5}$	18	3.6	$5.0e^{-3}$

aApproximate values estimated from assays at 700 mM and 350 mM ethanol only.

Figure 2. *T. ethanolicus* 2° ADH EXAFS analysis. (A) Spectrograph of normalized signal intensity versus energy for native (—) and recombinant (- - -)2° Adhs. (B) Plot of Fourier transformed data versus distance for native (—) and recombinant (- - -)2° Adhs.

A
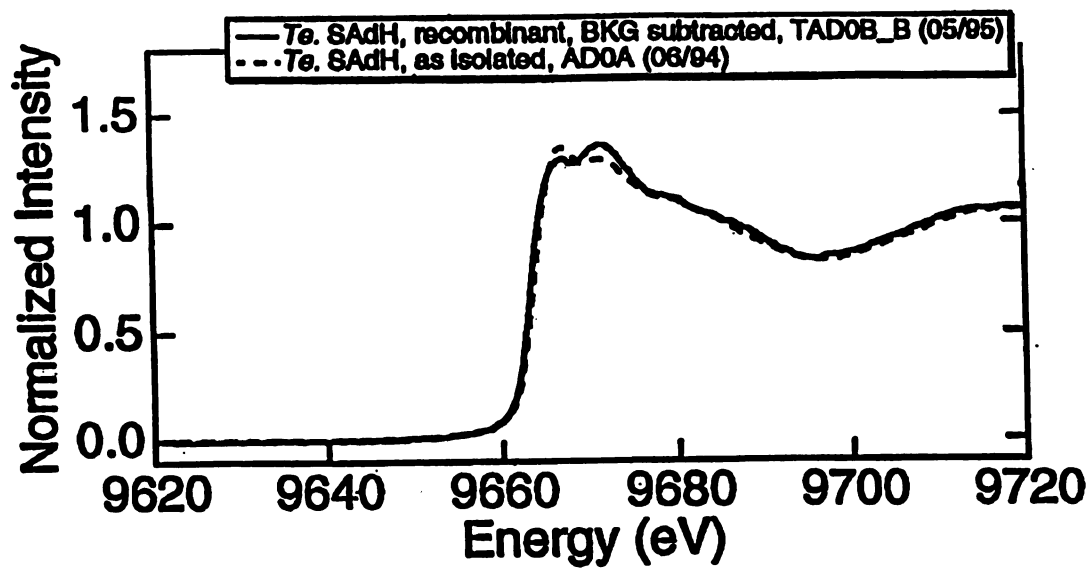
B
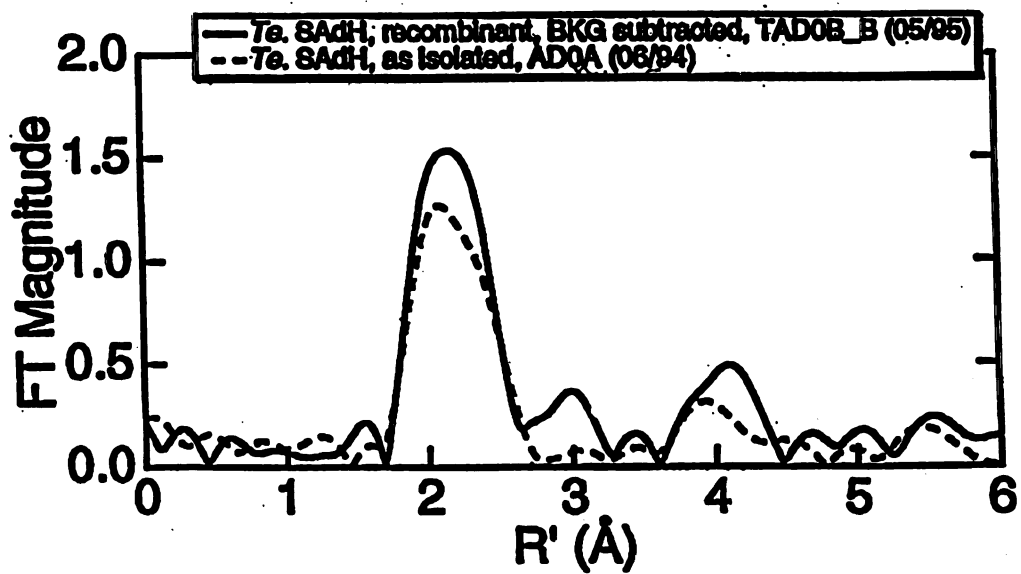
Figure 3. **PCR-based adhB gene site directed mutagenesis scheme.** The partial gene PCR products from reactions using overlapping coding- and noncoding-strand oligonucleotides encoding the mutation with one complimentary to the 5' or 3' end were combined, anealed, extended, and amplified in a second PCR reaction which yielded the complete mutant *adhB* gene.

Kpn I

- - - wild type template - - -

←mut.

mut. ←

- - - wild type template - - -

Apa I

Amplify (26 cycles)

Kpn I

mutant N-end

mutant C-end

Apa I

Aneal (50°C)

Kpn I

Apa I

+

Kpn I

Apa I

Extend (3 cycles)

Kpn I

Complete mutant

Apa I

Amplify (26 cycles)

Kpn I

Complete mutant

Apa I

169
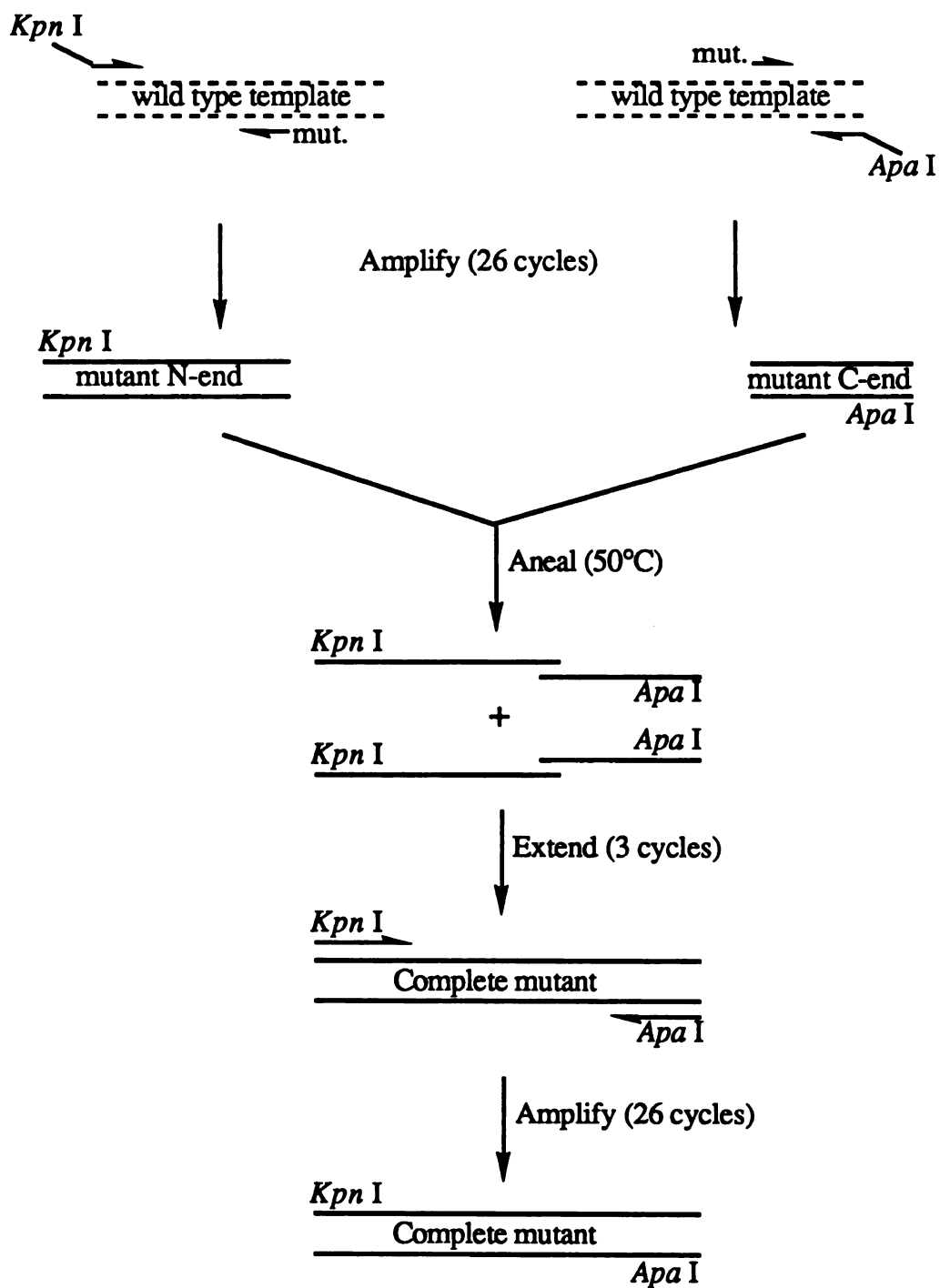
Figure 4. **SDS-PAGE of 2° ADH mutant proteins.** Lanes: 1, cell extract of *E. coli* DH5α containing pBluescriptII-KS (+) plus the kan^r cartridge; 2, C37S mutant cell extract; 3, purified C37S mutant; 4, purified D150C mutant; 5, purified D150E mutant; 6, purified G198D mutant; 7, purified H59N mutant. The Mr values for the molecular mass markers flanking the sample lanes are indicated on the left. Proteins were stained with Coomassee brilliant blue R-250.
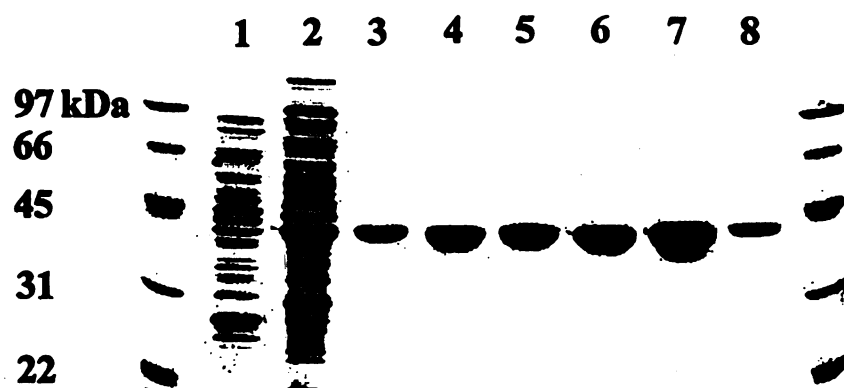
dialysis in buffer containing metal chelating resin were active and indicated no increase in activity upon Zn addition (data not shown).

Point mutations altering Cys37, His59, and Asp150 significantly reduced enzyme activity (Table III). The Cys37 to Ser mutation replaced the putative Zn ligand residue with one less than 0.4 Å shorter, but the Ser hydroxyl's ionization constant is 7 orders of magnitude smaller than that of the Cys sulfhydryl. This mutant retained less than 1% of the wild type enzyme activity toward both ethanol and propan-2-ol. Alteration of His59 to Asn eliminated the ionizable nitrogen atom and reduced enzyme activity to less than 0.5% of wild type. The Asp150 to Cys mutant which presumably mirrored the Cys-His-Cys motif for 1° ADH catalytic Zn binding, had the highest specific activity of the catalytic Zn mutants, retaining approximately 3% of the wild type enzyme activity. Replacement of the Asp carboxyl moiety with a sulfhydryl group preserved the strong negative ionic character but shortened the amino acid side chain by approximately 0.6 Å. Mutating Asp150 to Glu lengthened the side chain approximately 1.3 Å and reduced the specific activity to 0.2% of wild type. Both the Asp to Cys and Asp to Glu mutations retained the electronic characteristics for metal binding but mutation of Asp150 to Asn significantly reduced the polarity of this residue, maintaining the wild type residue geometry. The Asp150 to Asn mutant retained less than 3% of the wild type activity.

Cofactor and substrate specificity - All of these catalytic Zn liganding residue mutants had similarly high affinities for $NADP^+$ (Km_{app} values less than 40 μM) and, like the wild type enzyme, demonstrated very low activities using NAD^+. The Asp150 to Asn mutant Km_{app} value for $NADP^+$ was 37 μM with an unusually large standard deviation of 50%, making the difference between it and the wild type Km_{app} (16 μM) insignificant. The Asp150 to Cys mutant however, had an $NADP^+$ Km_{app} value near 1.0 μM which is 16-fold lower than that of the wild type enzyme.

Weringa and Hol, 1983, identified a specific amino acid position responsible for confering NAD(H) versus NADP(H) cofactor specificity in Rossman fold containing

proteins. Gly198 was predicted to occupy this position in the *T. ethanolicus* 2° ADH

sequence [Burdette *et al.*, 1996] and is consistent with the enzyme's ability to bind

NADP(H) since its small uncharged sidechain would accomodate the additional cofactor

ribose phosphate, unlike the Glu or Asp residues present in NAD(H) linked enzymes that

repel the negatively charged phosphate moiety. Based on Weringa and Hoi's hypothesis,

mutation of Gly198 to Asp would exclude NADP(H) from the cofactor binding site but

allow NAD(H) binding. The wild type enzyme had a 140-fold lower Km_{app} for $NADP^+$

than for NAD^+ and a catalytic efficiency for propan-2-ol oxidation 2400 fold greater using

$NADP^+$ than using NAD^+ (Table IV). Mutation of Gly198 to Asp increased the Km_{app} for

$NADP^+$ 225-fold and reduced the $Vmax_{app}$ using $NADP^+$ 4-fold. The mutant enzyme had

a 3-fold lower Km_{app} toward NAD^+ and a 5.5-fold higher propan-2-ol oxidation rate using

NAD^+ compared to wild type, giving the mutant a 6.7-fold lower catalytic efficiency for

$NADP^+$ linked propan-2-ol oxidation than for the NAD^+ dependent reaction.

Mutational effects on enzyme affinities for 1° versus 2° alcohol substrates and the

catalytic efficiencies for their oxidation indicate a role for these putative catalytic Zn ligands

in substrate specificity as well as catalytic rate enhancement. Mutation of the Cys37 to Ser

increased the Km_{app} 6-fold and 20-fold for propan-2-ol and ethanol, respectively (Table

III). Determining the Cys37 to Ser mutant enzyme Michaelis-Menten constant for ethanol

was complicated by the inability to measure enzyme initial rates using substrate

concentrations greater than 700 mM and less than 300 mM. Therefore, this Km_{app} is

presented for completeness but it is an approximate value and is not included in subsequent

data analysis. Alteration of His59 to Asn caused a 6-fold increase in Km_{app} toward

propan-2-ol and no significant change in the Km_{app} toward ethanol. The Asp150 to Cys

mutant Km_{app} for propan-2-ol was similar to the wild type value and that for ethanol was

increased only 2-fold. Mutating Asp150 to Glu caused a 3-fold increase in Km_{app} for

ethanol but a 29-fold increase in Km_{app} toward propan-2-ol. Finally, the Asp150 to Asn

mutant enzyme had a Km_{app} value toward propan-2-ol 5-fold lower than the wild type and

Table IV

Effect of the Gly198 to Asp mutation on T. ethanolicus 2° Adh nicotinamide cofactor preference

Kinetic parameters for mutant enzyme catalysis. Enzyme initial rates were measured by the increase in absorbance due to $NAD(P)^+$ reduction at 60 °C ($\epsilon_{340nm} = 6.22$ mM^{-1} cm^{-1}). Km_{app} and $Vmax_{app}$ values were determined by nonlinear curve fits to initial rate data from assays using $0.2xKm_{app}$ to $20xKm_{app}$ initial cofactor concentrations with the cosubstrate at $10xKm_{app}$. Propan-2-ol was used as the cosubstrate in all cases. Cofactor efficiency was defined as the ration of the catalytic efficiencies $(Vmax_{app}/Km_{app})$ for each cofactor form.

| Enzyme | NADP+ | | NAD+ | | Cofactor efficiency |
| | $Vmax_{app}$ | Km_{app} | $Vmax_{app}$ | Km_{app} | ratio |
	$U\,mg^{-1}$	mM	$U\,mg^{-1}$	mM	$NADP^+/NAD^+$
wt	72	0.016	4.4	2.3	2400
G198D	18	3.6	25	0.76	0.15

a Km_{app} value toward ethanol 4-fold greater than the wild type enzyme. The mutant enzyme catalytic efficiencies were lower than those for the wild type enzyme in all cases but the ratio of these values for propan-2-ol oxidation compared to ethanol oxidation varied from 4-fold lower than (the Asp150 to Glu mutant) to 10-fold higher than (the Asp150 to Asn mutant) that calculated for the wild type enzyme.

DISCUSSION

The research described here provides direct evidence for the molecular basis for specific amino acid residues in *T. ethanolicus* 39E 2° ADH catalysis. The presence of a single specifically bound Zn ion per enzyme subunit and the identities of at least two of the three potential liganding protein atoms were confirmed by EXAFS measurements. Induction coupled plasma emission spectrometry (ICP) analysis of wild type and mutant enzymes implicated Cys37, His59, and Asp150 in catalytic Zn binding. Kinetic analysis of mutant enzymes implicated Cys37, His59, and Asp150 in both catalytic rate enhancement and substrate specificity. Transformation of the highly NADP(H) specific *T. ethanolicus* 2° ADH into a NAD(H) preferring enzyme with the Gly198 to Asp mutation further validates the 1° to 2° ADH sequence comparisons, provides experimental evidence that this 2° ADH binds nicotinamide cofactor in a Rossmann fold, and makes the *T. ethanolicus* enzyme a more suitable as an industrial biocatalyst. Furthermore, construction of the overexpression system provides sufficient yield of purified enzyme for structural analysis and potential industrial scale biocatalyst production.

The original *T. ethanolicus* 39E 2° ADH clone was expressed in *E. coli* DH5α under the control of its native transcriptional promoter at ~2% of total protein [Burdette *et al.*, 1996]. Removal of the native transcriptional promoter region, placing the *adhB* gene

under control of the pBluescriptII KS(+) T3 promoter (the pADHBKA4-kan construct) increased recombinant protein expression to ~15% of total protein. This indicates that the native *T. ethanolicus* promoter is not as strong as the T3 promoter in *E. coli*. Thermophilic ADH purification using only 85°C treatment and ammonium sulfate precipitation reduced enzyme losses to 33% while also eliminating time consuming and costly chromatography steps. Protein expressed from the pADHBKA4-kan plasmid, like the original clone [Burdette *et al.*, 1996], consisted of a mixed recombinant enzyme population composed of protein with one or two N-terminal Met residues. However, recombinant protein with only a single N-terminal Met, like the mature enzyme isolated from *T. ethanolicus* 39E, was detected in the population expressed from pADHB1M1-kan. Thus, altering the "ATGATG" translational initiation site in the *adhB* gene to "TTAATG" resulted in the production of a homogeneous, native-like, recombinant enzyme population. The homogeneous and mixed recombinant enzyme populations were kinetically indistinguishable as previously reported for enzyme isolated from *T. ethanolicus* and the mixed recombinant protein [Burdette *et al.*, 1996]. This provides direct evidence that the additional Met did not significantly alter recombinant enzyme expression or function.

The prediction that Gly198 was in part responsible for *T. ethanolicus* 2° ADH NADP(H) specificity was based on the hypothesis that the cofactor was bound in a Rossmann fold [Burdette *et al.*, 1996]. Identification of a Rossmann fold consensus sequence in the 2° ADH peptide supported this hypothesis. However, the dramatic decrease in enzyme affinity for $NADP^+$ resulting from mutation of Gly198 to Asp is the first experimental evidence for the prediction that *T. ethanolicus* 39E 2° ADH binds nicotinamide cofactor in a Rossmann fold. The strong primary structural similarity between 2° ADHs argues that this is a common 2° ADH feature. Mutation of Gly198 to Asp only increased enzyme affinity for NAD^+ 3-fold, but this is still 48-fold lower affinity than the wild type enzyme for $NADP^+$. Therefore, the negatively charged Asp residue appears to repel the added $NADP^+$ phosphate moiety, consistent with theory [Weringa and

Hoi,1983], but does not significantly decrease the *T. ethanolicus* 2° ADH's high Km$_{app}$

toward NAD$^+$. While the Gly198 to Asp mutation increases the catalytic efficiency ratio

for NAD$^+$ reduction to NADP$^+$ reduction 16000-fold, the mutant enzyme Vmax$_{app}$ for

catalysis using either NAD$^+$ or NADP$^+$ is lower than the wild type enzyme's Vmax$_{app}$ for

NADP$^+$ dependent propan-2-ol oxidation, suggesting that mutant enzyme-cofactor

complex is suboptimal for turnover. The Gly198 to Asp mutant's Km$_{app}$ values for both

ethanol and propan-2-ol were similar to those of the wild type enzyme, suggesting that this

difference is not important to alcohol substrate affinity.

With few exceptions, ADHs require an enzyme bound Zn ion for catalysis. Two

Cys residues and a single His residue act as the catalytic Zn ligands in 1° ADHs [Bränden *et*

al., 1975]. Reactivation of *T. ethanolicus* 2° ADH after sulfhydryl modification was

enhanced by Zn suggesting that enzyme activity may be Zn dependent [Burdette *et al.*,

1996] however, confirmation of this observation requires more rigorous structural

analysis. The wild type recombinant enzymes (expressed from plasmids pADHBKA4-kan

and PADHB1M1-kan) were shown by ICP analysis to bind one Zn for every two subunits

suggesting that the 2° ADH possesses a single competent active site per dimer.

Determining the accuracy of this prediction however, requires exact enzyme 3-dimensional

structural information. The similar Zn content observed in the Gly198 to Asp mutant (0.55

Zn per subunit) suggests that this mutation does not interfere wth Zn binding and that the

Zn bindng site structure is retained. This analysis supports the sequence based prediction

that the 2° ADH subunits do not contain a Cys liganded structural Zn atom analogous to that

in the 1° ADHs [Peretz and Burstein, 1989; Burdette *et al.*, 1996]. The EXAFS results for

this enzyme identify a specifically bound Zn ion with sulfur, nitrogen, and oxygen ligands.

The coordination sphere {ZnS$_1$(N,O)$_{3-4}$} measured for the 2° ADH would be expected to

have a much larger number of sulfur ligands (5) if both a catalytic and a structural Zn (with

coordination sphere of ZnS$_4$) were present. Thus, the EXAFS data further support the

predicted lack of a 2° ADH structural Zn. Although calculations from the EXAFS spectra

indicated 1 or 2 possible sulfur atoms coordinated to the enzyme bound Zn, the best fit to the data is for a single sulfur ligand. The EXAFS data also indicated the presence of a single imidazole nitrogen ligand to the Zn atom. This result is consistent with peptide sequence alignments and chemical modification experiments that have implicated Cys and His residues in 2° ADH catalytic Zn binding [Burdette *et al.*, 1996]. Finally, the EXAFS analysis is is consistent with the peptide alignment-predicted Cys-His-Asp catalytic Zn liganding motif among 2° ADHs.

Site directed mutations of the predicted Cys and His residues to eliminate them as potential Zn ligands reduced enzyme activity to less than 1% of wild type. Also, the molar ratio of bound Zn per subunit for these mutants was 4-fold lower than for wild type enzyme. Therefore, both Cys37 and His59 are important to catalysis and to Zn binding. The protein samples prepared for ICP measurements displayed enzyme activity and 10 μM added Zn did not enhance catalytic rates in the standard assays. Zn contaminated assay components could have provided sufficient metal to reconstitute the mutant enzyme catalytic sites, explaining the observed activity. ICP measurements indicating that ~10% of the protein subunits bound metal after dialysis however, suggest that the low mutant catalytic activities (<1% of wild type enzyme activity using either substrate) displayed by these mutants did not result simply from the loss of Zn. The altered mutant enzyme affinities for substrates also support this conclusion. Mutation of Asp150 to Asn also reduced enzyme activity toward both ethanol and propan-2-ol to less than 1% of wild type indicating the importance of Asp150 in 2° ADH catalysis. Zn binding by the Asp150 to Asn mutant was one third less than wild type Zn binding. Furthermore, the mutant catalytic activity was significantly lower that expected by the mere loss of Zn in 33% of the protein population, arguing that this resudue, too, is directly involved in Zn binding and in catalysis. Replacing Asp150 with Glu alters only the carbonyl position, yet this eliminated all but 0.3% of the wild type enzyme activity. This mutant, like the wild type enzyme, bound 0.47 Zn ions per subunit so its low activity is also not due to the lack of bound Zn. If

Asp150 provides the third 2° ADH catalytic Zn ligand, then substitution of Cys for Asp150 should have replicated the 1° ADH Cys-His-Cys Zn liganding structure, although the Cys sidechain is slightly shorter that the Asp sidechain. This Asp150 to Cys mutant retained only about 3% of the wild type enzyme activity. Furthermore, this mutant enzyme bound 1.2 Zn atoms per subunit so reduced catalytic activity clearly did not result from loss of Zn. The enhancement of Zn binding by this mutant may be due to stronger metal binding or from the creation of a second competent Zn binding site per dimer. The effects of these three mutations at residue 150 argue that both the size and ionic character of the Asp sidechain are critical to enzyme function. Thus, changing the Zn binding site charactertistics affected enzyme catalysis, providing direct evidence that Cys37, His59, and Asp150 are Zn ligands and that, like 1° ADH activity, proper Zn binding is essential for 2° ADH catalysis. The effects of mutations at amino acid residues 37, 59, and 150 provide strong support for the sequence-comparison based predictions, but final confirmation of this novel Zn binding motif awaits crystallographic determination of the 2° ADH 3-dimensional structure.

Because the horse liver 1° ADH is so well characterized, it is an excellent comparative reference for predicting the structures of other ADHs. Reports of 3-dimensional structural similarity between 1° and 2° ADHs based on peptide sequence comparisons have suggested that although similarities exist, these enzyme groups may lack complete structural similarity [Peretz and Burstein, 1989; Burdette *et al.*, 1996]. The similarities and differences identified here between the horse liver 1° ADH and the *T. ethanolicus* 2° ADH are summarized in Fig 5a,b. Keinan *et al.*, 1986a, proposed an asymmetric 2 lobe *T. brockii* 2° ADH active site model based on its accomodation of substrates varying in length but its requirement for the site of dehydrogenation on the first to the third carbon atoms. The data presented here demonstrate that the *T. ethanolicus* 2° ADH is a Zn dependent, Rossmann fold containing enzyme that appears to use a novel Cys-His-Asp metal binding motif. Furthermore, kinetic analyses of site directed mutants

confirm the sequence-comparison based predictions of specific residues affecting cofactor selectivity and catalysis. Thus, we propose a hypothetical working model for 2° ADH structure-function studies (Fig. 5b). In addition, the previous discovery [Burdette and Zeikus, 1994] of a physiological acetylCoA linked reductive thioesterase activity for this *T. ethanolicus* enzyme suggested that the previously hypothesized long lobe [Keinan *et al.*; 1986a] may be open ended, allowing the coenzymeA portion of the substrate to remain outside the active site. The active site channel in the horse liver ADH has a short segment (Fig. 5a), one wall of which consists of the bound nicotinamide cofactor, and a long segment open at the far end, with the catalytic Zn atom at their junction [Eklund *et al.*, 1974]. This 1° ADH architecture is consistent with the 2° ADH active site requirements for a short lobe to accomodate the C1 to C3 substrate segment, but, it also allows very long moieties on the other end of the ketone.

The kinetic mechanisms proposed for both horse liver [Brändén *et al.*, 1975] and *T. brockii* [Pereira *et al.*, 1994] ADH catalysis require bound cofactor prior to solvent substrate binding. The similar *T. brockii* and liver ADH kinetic mechanisms suggest that cofactor binding is needed for solvent binding by both 1° and 2° ADHs. Analysis of the horse liver 3-dimensional structure [Eklund *et al.*, 1974] suggests that the cofactor may function to complete the active site wall. Although exact crystallographic evidence is needed to confirm this hypothesis, the similar kinetic mechanisms and predicted structural characteristics suggests that, like the liver 1° ADH active site, nicotinamide cofactor may complete 2° ADH catalytic pocket formation. Also, the His59 and Asp150 mutations altered 2° ADH affinity for propan-2-ol over a 140-fold range while the overall change in affinity for ethanol spanned a range of only 7-fold. The greater effects of these mutations on propan-2-ol than on ethanol affinity suggest that the catalytic Zn's position affects substrate specificity. The greater secondary alcohol sensitivity to zinc ligand mutations suggests that substrate binds the 2° ADH catalytic Zn atom, like in the liver 1° ADH [Brändén *et al.*, 1975], and that 2°-alcohol binding is more susceptible to steric hinderance

Figure 5. Comparison of 1° and 2° ADH active site models. (A) The liver 1°
ADH active site model, and (B) The hypothesized 2° ADH active site model incorporating
the mutagenic and biophysical information.

A.
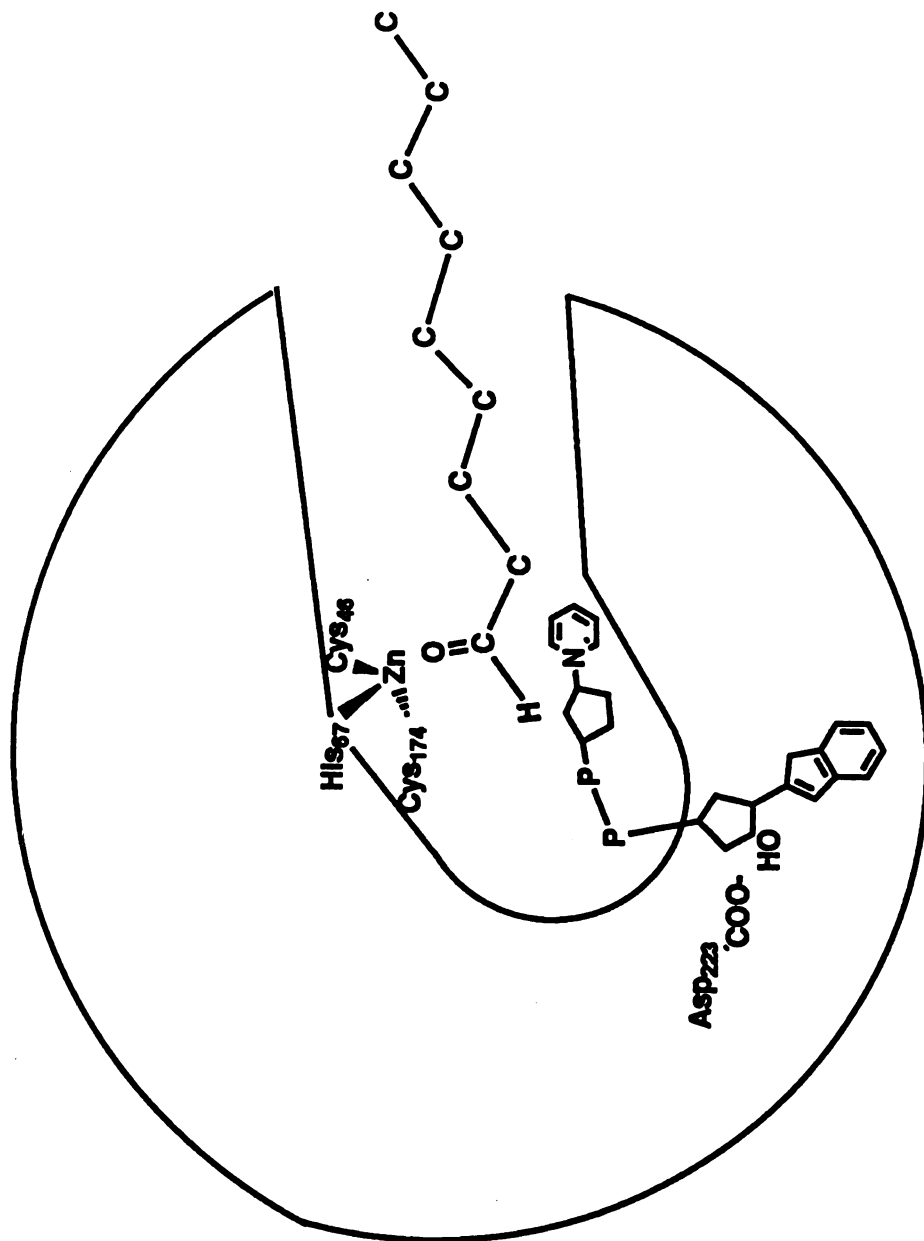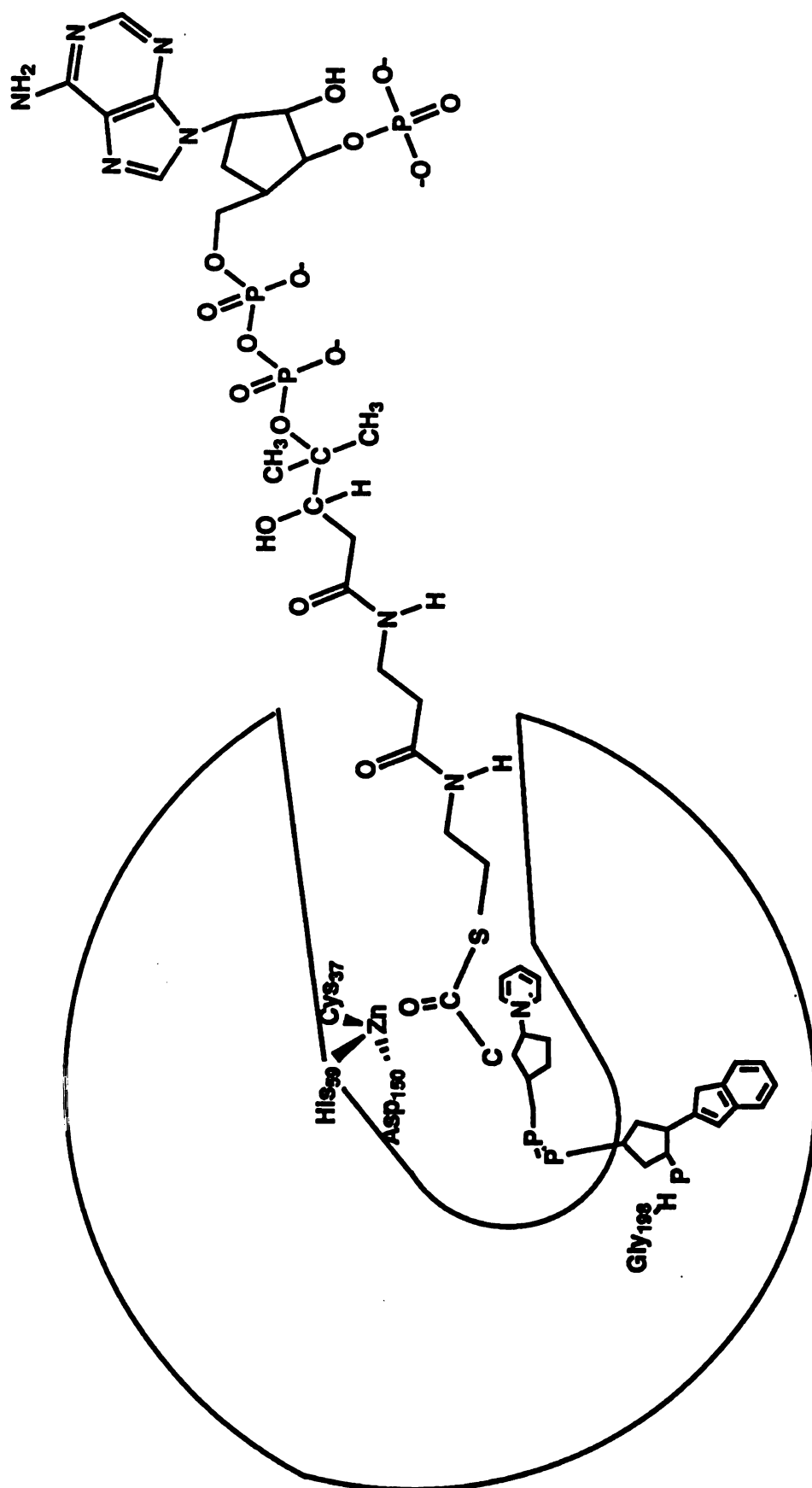
B.

in the catalytic pocket. The effect of mutations that alter the active site pocket depth on enzyme specificity must wait for determination of which residues line the bottom of the active site.

2° ADH have been employed in chiral compound synthesis and racemate resolution [Keinan *et al.*, 1986a; Keinan *et al.*, 1986b; Hummel, 1990; Keinan *et al.*, 1990; Bradshaw *et al.*; 1992a, Bradshaw *et al.*; 1992b, Zheng *et al.*, 1992]. Recombinant overexpression of this enzyme at ~15% of total protein followed by purification using only heat and ammonium sulfate precipitation made possible the economical production of large enzyme quantities, vital for structural analyses and for its potential utility as a commercial biocatalyst. Converting the strongly NADP(H) linked *T. ethanolicus* enzyme to one preferring NAD(H) through the Gly198 to Asp mutation created the first reported highly stable, thermophilic 2° ADH with higher catalytic efficiency using the more thermally stable NAD(H) cofactor without altering the substrate specificity. Thus, the Gly198 to Asp mutant enzyme retains the desirable 2° ADH catalytic properties but reduces the cost of potential industrial processes by switching enzyme preference to a cofactor form (NAD(H)) that degrades more slowly. The evidence presented here of 2° ADH active site structural similarity to 1° ADH active site structure has provided a revised thermophilic 2° ADH active site model based on comparison to the horse liver enzyme. This mutagenic evidence for 2° ADH-specific molecular structure-function relationships indicates that comparative analysis between these two enzyme classes may yield insights into the molecular determinants of both 1° and 2° ADH substrate specificities.

ACKNOWLEDGEMENTS

185

REFERENCES

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. and Struhl, K. (1993) Current Protocols in Molecular Biology (Janssen, K., ed.), Current Protocols, NY

Bleicher, K. and J. Winter (1991) Eur. J. Biochem. **200**, 43-51

Bradshaw, C. W., Fu, H., Shen, G.-J. and Wong, C.-H. (1992) J. Organic Chem. **57**, 1526-1532

Bradshaw, C. W., Hummel, W. and Wong, C.-H. (1992) J. Organic Chem. **57**, 1533-1536

Bränden, C.-I., Jörnvall, H., Eklund, H. and Furugren, B. (1975) in The enzymes, vol. XI part A (Boyer, P. D., ed.), pp. 103-190, Academic Press, NY

Brooks, S. (1992) Biotechniques **13**, 906-911

Bryant, F. O., Wiegel, J. and Ljungdahl, L. (1988) Appl. Env. Microbiol. **54**, 460-465

Burdette, D. S. and Zeikus, J. G. (1994) Biochem. J. **302**, 163-170

Burdette, D. S., Vieille, C. and J. G. Zeikus (1996) Biochem. J. **313**, in press

Dunn, M. F. and Hutchison, J. S. (1973) Biochemistry **12**:4882-4892

Eklund, H., Nordström, B., Zeppezauer, E., Söderlund, G., Ohlsson, I., Boiwe, T. and Bränden, C.-I. (1974) FEBS Lett. **44**, 200-204

Hummel, W. (1990) Appl. Microbiol. Biotech. **34**, 15-19

Ismaiel, A. A., Zhu, C.-X., Colby,G. D. and Chen, J.-S. (1993) J. Bacteriol. **175**, 5097-5105

Jacobs, J. W., McFarland, Wainer, I., Jeanmaier, D., Ham, C., Hamm, K., Wnuk, M., and Lam, M. (1974) Biochemistry **13**, 60-64

Jendrossek, D., Steinbuechel, A. and Schlegel, H. G. (1988) J. Bacteriol. **170**, 5248-5256

a Keinan, E., Hafeli, E. K., Seth, K. K. and Lamed, R. L. (1986) J. Am. Chem. Soc. **108**, 162-169

b Keinan, E., Hafeli, E. K., Seth, K. K. and Lamed, R. L. (1986) Ann. N. Y. Acad. Sci. **501**, 130-149

Keinan, E., Seth, K. K., Lamed, R. L., Ghirlando, R. and Singh, S. P. (1990) Biocatalysis **4**, 1-15

Lamed, R. J., Keinan, E. and Zeikus, J. G. (1981) Enz. Microbiol. Technol. **3**, 144-148

Lamed, R. J. and Zeikus, J. G. (1981) Biochem. J. **195**, 183-190

Lovitt, R. W., Longin, R. and Zeikus, J. G. (1984) Appl. Environ. Microbiol. **48**, 171-177

McFarland, Chu, Y.-H., and J. W. Jacobs (1974) Biochemistry **13**, 65-69

Nagata, N., Maeda, K. and Scopes, R. K. (1992) Bioseparation **2**, 353-362

Oka, A., Sugisaki, H. and Takanami., M. (1981) J. Mol. Biol. **147**, 217-226

Pereira, D. A., Gerson, F. P., and E. G. Oestreicher (1994) J. Biotech. **34**, 43-50

Peretz, M. and Burstein, Y. (1989) Biochemistry **28**, 6549-6555

Rossmann, M. G. and Argos, P. (1976) J. Mol. Biol. **105**, 75-95

Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecular cloning: A laboratory manual 2nd edition. (Nolan, C., ed.), Cold spring Harbor Press, NY

Sanger, F., Nicklen, S. and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. USA **74**, 5463-5467

Scopes, R. K. (1983) FEBS Lett. **156**, 303-306

Steinbüchel, A. and Schlegel, H. G. (1984) Eur. J. Biochem. **141**, 555-564

Wierenga, R. K. and Hol, G. J. (1983) Nature **302**, 842-844

Zheng, C, Pham, V. T., and Phillips, R. S. (1992) Bioorganic Medicinal Chem. **2**, 619-622

Chapter V

Conclusions

CONCLUSIONS

The *Thermoanaerobacter ethanolicus* 2° ADH is a thermophilic enzyme physiologically analogous to catabolic 1° ADHs. Both 1° and 2° ADHs have been used in analytical scale, industrially valuable chiral syntheses. The thermophilic ADHs have demonstrated generally higher stability than their mesophilic counterparts. Decades of research have yielded detailed information about liver 1° ADH 3-dimensional structure, kinetic mechanism, functional architecture, and biotechnological potential; yet, these studies have failed to identify specific amino acid residues responsible for substrate descrimination. Unanswered, this fundamental ADH structure-function question also prevents rational protein engineering approaches to optimizing their design for chiral catalysis. 2° ADHs are structurally and functionally similar to 1° ADHs, thus, they allow comparative analysis to identify the molecular determinants of 1° versus 2° alcohol substrate specificity. Isolating both mesophilic and thermophilic 1° and 2° ADHs similarly allows comparative analysis to identify the molecular strategies used in ADHs to thermostabilize the folded proteins. Finally, overexpressing the thermophilic 2° ADH cloned *adhB* gene, the only thermophilic 2° ADH clone reported, allows site directed mutagenesis to test thermal feature or substrate specificity hypotheses from comparative sequence analysis.

· Biochemical, biophysical, and molecular biological approaches were employed here to characterize the mechanisms responsible for *T. ethanolicus* 2° ADH thermal stability and catalysis. First, the *adhB* gene encoding the *T. ethanolicus* 39E 2° ADH was isolated, sequenced, and the corresponding 2° ADH peptide compared to other 1° And 2° ADH peptide sequences. Second, the recombinant protein was overexpressed, purified, and kinetically characterized. Third, site directed mutagenesis and biophysical analysis were used to validate the predictions of 2° ADH catalytic architecture based on peptide sequence comparison to 1° ADHs. Finally, enzyme thermostability and thermophilicity were characterized and their molecular biochemistry investigated.
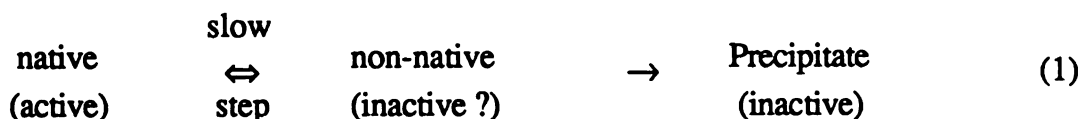
The similar catalytic and thermal properties of the recombinant 2° ADH to those of the catabolic enzyme purified from *T. ethanolicus* indicate that the isolated *adhB* gene encodes the catabolic *T. ethanolicus* 2° ADH. Calculations using the translated *T. ethanolicus adhB* sequence coincided with the amino acid composition and molecular mass experimentally determined for the native enzyme, arguing that the recombinant gene was properly expressed in *E. coli*. The recombinant enzyme's catalytic and thermal properties indicated that the enzyme 1° structural information was sufficient to properly fold the protein at both thermophilic and mesophilic temperatures. The deduced *T. ethanolicus* peptide sequence was extremely similar to that determined for the *T. brockii* enzyme by Edman degradation of the native enzyme. The *T. ethanolicus* 2° ADH amino acid sequence was also quite similar to the peptide deduced from the gene encoding the mesophilic *C. Beijerinckii* 2° ADH, arguing that these enzymes also share overall structural similarity. Finally, overexpression (~15% of total protein) of properly folded 2° ADH followed by purification using only heat treatment and $(NH_4)_2SO_4$ precipitation made production of sufficiently large enzyme quantities for structure-function or biotechnological utility analysis practical.

1° ADHs are either dimers or tetramers of identical or nearly identical 35 kDa to 45 kDa subunits. The characterized 2° ADHs are homotetramers of 36 kDa to 42 kDa peptides. Both classes of enzymes typically require a nicotinamide cofactor for catalysis and act through an ordered kinetic mechanism. Chemical modification experiments demonstrated the importance of Zn, cysteine residues, and histidine residues in *T. ethanolicus* 2° ADH catalytic activity, similar to that of 1° ADHs. EXAFS analysis identified a bound Zn atom with a Coordination sphere of $ZnS_1(NO)_{3-4}$ and ICP measurements indicated a single Zn atom per wild type enzyme dimer. Mutations of Cys37, His59, and Asp150 altered both enzyme catalytic activity and Zn binding. The mutagenesis results and the Zn coordination sphere atomic composition together support the comparative sequence prediction that this 2° ADH uses Cys37, His59, and Asp150 to

bind the catalytic Zn, a novel metal ligand structure compared to the Cys-His-Cys motif seen in 1° ADHs. The low level of bound Zn and the prediction of only a single Cys residue in the Zn coordination sphere also support the conclusion that the *T. ethanolicus* 2° ADH subunits, unlike 1° ADH subunits, do not bind a structural Zn atom liganded by 4 Cys residues. Converting the NADP(H) specific *T. ethanolicus* 2° ADH into a NAD(H) preferring enzyme by the Gly198 to Asp point mutation as predicted from comparison of the 2° ADH peptide to that of the horse liver 1° ADH. The dramatic reduction in mutant enzyme affinity for NADP$^+$ argues that both the sequence alignment and the assumption of a Rossmann fold for 2° ADH nicotinamide cofactor binding are correct. Furthermore, the experimental confirmation of both specific predicted structure-function similarities (eg., a catalytic Zn atom liganded by Cys and His residues) and differences (eg., the absence of a 2° ADH structural Zn and the Cys-His-Asp- for 2° ADH catalytic Zn binding versus the 1° ADH's Cys-His-Cys motif) between the enzyme classes, supports the assumption of overall 1° versus 2° ADH structural similarity and validates the continued use of 1° ADH structural data to direct future 2° ADH experiments.

T. ethanolicus 2° ADH catalyzed propan-2-ol oxidation was optimal near 90°C. The half-life of 2° ADH activity was more than an hour at 90°C. Arrhenius analysis of the 2° ADH thermoinactivation rate constants indicated a large activation energy (430 kJ mol^{-1}). The 2° ADH thermoinactivation rate is accelerated therefore, by factors influencing the large positive arrhenius constant (1.8×10^{60}). Precipitation, quantitated by spectrophotometric absorbance, occurred at >85°C. Enzyme unfolding in the absence of chemical denaturant was predicted - using both protein absorbance $T_{m[Gu]}$ and fluorescence measurements in the presence of GuHCl - to occur at higher temperatures (~115°C) than precipitation. Thermally precipitated protein did not redissolve in buffer nor did it display catalytic activity, indicating that thermoinactivation must precede or accompany precipitation. The pseudo-first order thermoinactivation rate determining step argued that this step was unimolecular with respect to protein and that it preceded multimolecular protein

precipitation. Scheme 1 summarizes the proposed thermophilic 2° ADH thermoinactivation process.

$$\begin{array}{ccccccc} & & \text{slow} & & & & \\ \text{native} & & \Leftrightarrow & \text{non-native} & \rightarrow & \text{Precipitate} & (1) \\ \text{(active)} & & \text{step} & \text{(inactive ?)} & & \text{(inactive)} & \end{array}$$

Propan-2-ol oxidation by enzyme preincubated at 55°C showed no statistically significant Arrhenius discontinuity. This temperature dependence of enzyme activity, consistent with Arrhenius theory, is completely described by the change in average substrate kinetic energy. Therefore, the discontinuous 2° ADH plots appear to result from a slow enzyme structural change from a highly active form to a less active form and that thermozyme Arrhenius plots discontinuities are not necessarily evidence of the role of reduced low temperature flexibility in the poor low temperature activity of thermozymes. The observed 2° ADH catalytic rate enhancement at low chemical denaturant concentrations required slightly greater GuHCl concentrations at higher temperatures contrary to the prediction of a catalytically significant effect of GuHCl on enzyme flexibility. The similar low effector concentration, enzyme rate enhancement was seen with added KCl, which is not a strong denaturant, indicated that the rate enhancement resulted from ionic interactions. These data argue that the observation of low temperature enzymatic rate enhancement by added chemical denaturants also may not indicate a role for optimal flexibility in the temperature dependence of thermozyme activity. The possible correlation between activity and flexibility therefore, requires direct measurements of both enzyme properties.

Chapter 6

Directions for Future Research

SEQUENCE BASED MUTAGENESIS AND KINETICS

The HCA alignment between thermophilic and mesophilic 1° and 2° ADHs (see chapter 2) predicted significant structural similarity despite dissimilar peptide sequences. These predicted catalytic structural similarities were supported by the results of mutagenic and biophysical analyses (see chapter 4). The predictive power of comparative sequence analysis was proposed as a tool to identify the molecular determinants of ADH substrate specificity. Furthermore, developing mutagenesis and active recombinant enzyme expression systems provide the means to test these structure function predictions. An extensive critical core residue comparison between the horse liver and *B. stearothermophilus* 1° ADHs and the *T. ethanolicus* and *C. beijerinckii* 2° ADHs indicated strong amino acid characteristic conservation among the peptides (see appendix A; table1). The conspicuous sequence differences between the 1° and 2° ADHs are listed in Table 1. These residues are mostly clustered around the horse liver enzyme active site based on the crystal structure (Fig. 1). Determining the effects of 1° versus 2° ADH comparison based *T. ethanolicus adhB* gene mutations on the 2° ADH substrate specificity and the effects complimentary mutations have on 1° ADH specificity will advance our understanding of the molecular biochemistry of ADH molecular biochemistry. This information may provide site directed mutagenesis based ADH specificity optimization for industrially relevant biotransformations.

The higher stability of NAD^+ versus $NADP^+$ makes it preferable for indusrtial applications. A single Gly to Asp mutation shifted the *T. ethanolicus* 2° ADH catalytic efficiency ratio from 2400-fold in favor of $NADP^+$ linked activity to 6.7-fold in favor of the NAD^+ dependent reaction. This mutation's effectiveness was strong evidence supporting the predicted 2° ADH Rossmann fold nicotinamide cofactor binding motif. The mutation reduced the maximal $NADP^+$ linked enzyme velocity 3-fold and increased the NAD^+ dependent reaction 5-fold creating a faster catalyst using NAD^+ but a generally

Table 1. Comparison of conserved 2° ADH amino acids with the corresponding 1° ADH residues deduced from HCA based sequence alignments

1° ADHs		2° ADHs	
Horse liver	*B. stearotherm.*	*C. beijerinckii*	*T. ethanolicus*
Aromatic-nonaromatic			
Val13	Val5	Trp15	Trp14
Val83	Leu77	Phe75	Phe75
Leu350	Val310	Phe328	Phe328
Phe93	Phe89	Thr86	Ile86
conformational			
Met40	Ile32	Pro31	Pro31
Ile45	Val47	Pro36	Pro36
Charge			
Thr59	Pro53	Asp51	Glu51
Leu171	Pro145	Met147	Met147
Phe176	Val151	Met152	Met152
Size			
Ala11	Ala3	Leu13	Val12
Ala12	Ala4	Gly14	Gly13
Leu14	Val6	Ile16	Ile15
Val36	Val28	Ala27	Ala27
Val58	Val51	Gly50	Gly50
Val63	Ile57	Met55	Met55
Ile90	Ile85	Val83	Val83
Ala183	Ala155	Gly158	Gly158
Val186	Val158	Leu161	Leu161
Val189	Ala161	Ile164	Ile164
Leu342	Ala300	Val314	Val314
Ile346	Ile307	Val320	Val320
Glu353	Glu314	Asp328	Asp328
Gly358	Val319	Ala333	Ala333
Arg368	Arg331	Lys346	Lys346

Figure 1. Position of the non-conservative amino acid differences between 1° and 2° ADHs relative to the horse liver 1° ADH catalytic site. The liver ADH structural coordinates were obtained from the Cambridge protein structure database (filename "pdbadh6.ent"). Residues identified from the catalytic domain (red) and in the cofactor binding domain (green) are shown in relation to the catalytic Zn atom (light blue), the His ligand to the catalytic Zn (dark blue), and the two catalytic Zn cys residues (yellow).
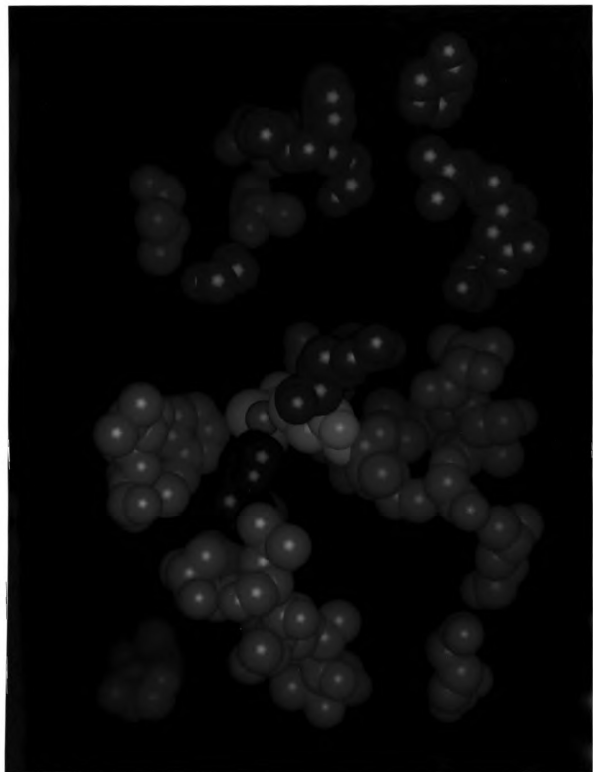
lower turnover enzyme. Also mutant enzyme affinity for NAD$^+$ increased only 3-fold while decreasing the 2° ADH affinity for NADP$^+$ 225-fold. This Gly198 to Asp mutant protein therefore, represents an incomplete transformation of nicotinamide cofactor preferences. The reduction potentials for NAD$^+$ and NADP$^+$ are identical, NAD$^+$ and NADP$^+$ structures differ only by a single phosphate group, the Rossmann fold structures that bind the cofactor forms differ only slightly, and NAD$^+$ linked 2° ADH with high turnover and high cofactor affinity exist in nature. These observations and the partial success already achieved argue that this NADP$^+$ linked 2° ADH may be successfully converted by further amino acid replacement into an improved functional NAD$^+$ dependent enzyme with high cofactor affinity. Sequence based comparative analysis of NAD$^+$ and NADP$^+$ linked ADHs (Appendix A; table 2) indicates nine further mutations that may complete this conversion (Table 2). These amino acid positions were identified based on their being consistently different between the NAD(H) linked and the NADP(H) dependent enzymes. The residue positions in the liver 3-dimensional structure (Fig. 1) indicate that they are also clustered near the catalyitc Zn atom. Assessing the potential contributions of these mutations on NAD(H) versus NADP(H) specificity requires additional kinetic analysis of the mutated enzymes.

The greater than 85% similarity between the *T. ethanolicus* and *C. beijerinckii* 2° ADH peptide sequences plus the vast difference in their reported thermostabilities makes the reciprocal site directed mutation of these proteins based on comparative sequence analysis a practicable and rational approach to understanding enzyme thermostability and thermophilicity. Of the 89 amino acid differences between these peptide sequences, only 12 were classified as nonconservative. Of these 12, nine were substituted Pro residues in the thermophilic *T. ethanolicus* enzyme (Table 3). Substituting Pro residues into protein loop regions have been shown to thermostabilize engineered proteins and to be important to natural protein thermostability (see chapter 1). A Proline-zipper hypothesis has been proposed to explain the enthalpic and entropic contributions of Pro residues to folded

Table 2. Comparison of conserved NADP+ dependent 2° ADH amino acids with the corresponding NAD+ dependent 1° ADH residues deduced from HCA based sequence alignments

Proposed site of interaction	aligned amino acid residues			
	Horse liver	B. stearotherm.	C. beijerinckii	T. ethanolicus
Adenine ring	Phe198 Ile224 Pro243 Thr274	Tyr171 Leu196 Pro215 Thr237	Ile173 Thr199 Tyr218 Gly244	Leu173 Thr199 Tyr281 Gly244
Adenosine ribose	Asn225 Lys228	Gly197 Lys200	Arg200 Cys203	Arg200 Cys203
Pyrophosphate	Arg47	His39	Thr38	Thr38
Nicotinamide ribose	Gly293	Gly261	His268	Phe268

Table 3. Added thermophilic 2° ADH proline residues compared to the mesophilic 2° ADH deduced from HCA based sequence alignments

Protein region description and proposed function	2° ADH	
	T. ethanolicus	*C. Beijerinckii*
Catalytic domain		
3-Pro in a 9-residue hydrophilic region	Pro22	Ala22
near Cys37(catalytic Zn ligand)	Pro24	Ser24
2-Pro in a 7-residue hydrophilic region	Pro313	Arg316
	Pro316	Val316
1-Pro in a 2-residue hydrophilic region	Pro347	Ala347
1-Pro in a 4-residue hydrophilic region	Pro149	Thr140
near Asp150 (catalytic Zn ligand)		
Rossmann fold		
in consensus motif - 1st turn of an α-helix	Pro177	Ala177
1-Pro in a 4-res. hydrophilic region	Pro222	His222
Dimer subunit interaction		
distorted helical region	Pro275	Leu275

thermozyme [1] stability. Importantly, in this thermophilic-mesophilic 2° ADH system, each Pro residue's contribution to enzyme thermal properties can be assessed by both the destabilizing effects due to thermophilic enzyme Pro removal and the stabilizing effects due to mesophilic 2° ADH Pro addition. This type of extensive reciprocal study using a set of multidomain, multisubunit enzymes has not been reported. Furthermore, the entensive sequence similarity between the mesophilic and thermophilic 2° ADHs reduces the possible number of thermal property determinants, allowing for potentially complete characterization of the structural determinants of these properties. The PCR based mutagenesis system developed here would also allow rapid construction of multiple mutations to assess the independence or cooperativity of these mutational effects. Therefore this system appears to be an excellent model to test the proline zipper hypothesis through biophysical measurement of protein thermal unfolding and thermoinactivation (eg., temperature or chemical denaturant dependent circular dichroism, differential scanning calorimetry, temperature or chemical denaturant dependent absorbance spectrophotometry or fluorimetry) (see chapter 3).

The thermophilic-mesophilic enzyme reciprocal mutagenesis experiments assess the contribution of each Pro to corresponding mesophilic peptide residue mutation to the overall folded protein stability. The specific thermophilicity and thermostability contributions of the added Pro residues as opposed to the replacement of the other amino acid must be quantitated by constructing a series of mesophilic and thermophilic protein mutants with a specific neutral (eg., Ser or Ala) residue at all Pro substitution positions and measuring their thermal properties (see chapter 3). The effect of each mutation on the enzyme thermal properties by itself, the effect due to interaction of the substituted residue with the rest of the protein, and the contribution of the individual mutation to the total difference in the thermal properties between the two 2° ADHs should be measured to identify potential mechanisms to account for the mutational effect. The difference between the completely neutral residue substituted thermophilic enzyme (thermophilic baseline

mutant) and the completely neutral residue substituted mesophilic enzyme (mesophilic baseline mutant) thermal properties reflects the contribution of other peptide factors in determining these thermal stability and activity properties. Next, subtracting the thermophilicity and thermostability values of each baseline mutant from the corresponding completely Pro substituted enzyme will quantitate the total stabilizing effect of these prolines on both 2° ADH folds. The difference between the mesophilic and thermophilic total stabilizing effects measures the interaction of these Pro mutations with the rest of the folded enzyme structure. The mutation specific and protein interactive effects of each Pro could be examined in this way. A large difference would indicate that the proline effects are strongly dependent on overall protein structure whereas a small difference would indicate independent thermal property contributions by these residues. The latter result would argue that Pro loop mutations are a predictable general thermostabilizing strategy. The neutral residue used in this study should mimic either general amino acid or Pro characteristics (aside from its side chain structural constraint). I propose that either Ala or Ser be used because they are small, reducing the possibility of disrupted local packing, they are uncharged but are often found at the protein surface, reducing the possibility of altering local 3-dimensional structure or salt bridging, and unlike Gly, both have residue geometric constraints typical of general amino acids. Furthermore, Ala's more Pro-like aliphatic sidechain makes it a closer analog and, since the substitutions are proposed in surface loops, the hypothesized α-helix stabilizing Ala characteristic should not be relevant.

Loop stabilization has also been proposed to enhance enzyme thermostability (see chapter 1). Determining the thermostability and thermophilicity effects of complimentary mutations to swap loop regions between the thermophilic and mesophilic 2° ADHs could specifically test this hypothesis. By identifying putative core regions the HCA representation also proposes the putative loop positions (those stretches of hydrophilic amino acids between the hydrophobic clusters). Like the Pro study, this work would evaluate both the role of this strategy in native enzyme stability and the protein engineering

potential for loop stabilization. Furthermore, the relative effects of Pro insertion into loops and loop stabilization strategies that do not involve prolines (i.e., added salt bridges) could be examined using this system. Substituting stabilized loops which do not contain Pro into both mesophilic and thermophilic 2° ADH loops that were shown to effect enzyme thermal properties due to the presence or absence of Pro residues would directly compare the relative efficacy Pro versus non-proline loop stabilization strategies.

STRUCTURAL ANALYSES

The research described in chapters 1 and 2 indicates that the poor low temperature activity of the thermophilic 2° ADH can be explained by the Arrhenius theory. Therefore the temperature dependence of enzyme activity can be described by changes in the average substrate kinetic energy alone. The potential influence of protein flexibility temperature dependence on the temperature dependence of enzyme activity however, has not been explicitly determined. Thermal property altering mutants of both the *T. ethanolicus* and the *C. beijerinckii* enzymes (eg., the Pro residue or loop substituted mutants described in the previous section) would provide an excellent system to examine this relationship using electron paramagnetic resonance spectrometry (EPR). The flexibility of immobilized liver 1° ADH in solvents has been determined from the spin label environment dependent EPR relaxation correlation times using a spin labelled compound coupled to Cys residues [1]. Enzyme flexibility has also been reported as the extent of deuterium exchange between the D_2O solvent and the protein (see chapter 1). Correlating the temperature dependence of flexibility and activity for such a system of mutants would test the hypothesis that enzymes require optimal flexibility for optimal activity. Furthermore, direct analysis of the relationship between protein flexibility and activity would examine whether the reported similar mesophilic and thermophilic protein flexibilities at their respective maximal activity temperatures are necessary for maximal activity or simply an indication that the folded structures are destabilizing.

2° ADH catalytic Zn binding was proposed to involve Cys, His, and Asp based on HCA ADH sequence comparisons, and chemical modification experiments (see chapter 2). The loss of catalytic activity and the predicted changes in protein Zn content for the mutant enzymes supported the hypothesis that Cys37, His59, and Asp150 provided the catalytic Zn ligands (see chapter 4). Also consistent with this hypothesis, EXAFS data identified a bound Zn atom with a coordination sphere containing 1 Cys sulfur, 1 imidazole nitrogen, and 3-4 other nitrogen or oxygen ligands. Similar EXAFS analysis of the Zn ligand mutant enzymes will provide relative signal intensities for the Zn coordinated atoms. Comparing the expected ligand specific EXAFS signals for each mutant to those observed will determine if the mutated residue directly altered the metal coordination sphere. These data would provide strong evidence to support or to refute the direct role of Cys37, His59, and Asp150 in catalytic Zn binding. This work is currently underway in our laboratory.

Constructing a recombinant *T. ethanolicus* 2° ADH overexpression system provided sufficient quantities of purified enzyme for crystallographic analysis. The exact 3-dimensional structure of this thermophilic 2° ADH is currently being determined using x-ray crystallography by our collaborator Dr. Bobby Arni. These data, the first 2° ADH exact 3-dimensional structure, would provide final confirmation of the overall 1° versus 2° ADH structural similarities. Comparing the liver 1° ADH and *T. ethanolicus* 2° ADH exact structures would also determine the validity of predictions based on the peptide sequence alignments and mutagenic analysis of catalytic Zn and cofactor binding. An exact 2° ADH structure would allow computer modeling to analyze the potential effects of proposed mutations and allow the identification of subtle mutations that may be structurally or functionally significant (ie., those based on the residue's position in the folded protein which do not require significantly altered amino acid properties which cannot be identified from the peptide sequence alone). Specifically, armed with both a 1° and a 2° ADH crystal structure residues influencing enzyme catalytic site geometry and altering substrate selectivity may be identified, allowing detailed analysis of the ADH substrate binding and

reaction mechanisms. Furthermore, the regions of subunit interaction would be identified, determining if this tetrameric enzyme is formed by the association of two 1° ADH-like dimers as currently believed. The number of structurally competent active sites in the holoprotein would also be determined. Crystallization of the 2° ADH with acetylCoA will also determine whether this physiological substrate is completely contained in the protein during catalysis or whether, as we propose, the long coenzyme A portion of the molecule is outside the protien during thioester reduction. If our hypothesis is correct, then the 2° ADH may catalyze the reduction of a wide range of very large ketone substrates but if the coenzyme is copletely retained in the protein active site then the 2° ADH substrate range would be expected to have a more limited size range. The crystallographic β values, the uncertainties in the atomic positions, for protein regions could be compared in the wild type enzyme x-ray structure and in thermal property altering mutant protein x-ray structures. Significantly altered regional β values would point to specific regions potentially responsible for folded enzyme structural integrity. Temperature dependent crystal structural analysis could directly measure differences in regional flexibilities within the protein (Petsco, personal communication). Comparing the measured flexibilities (residue β-values) and the corresponding enzyme activities could directly address the role of protein flexibility in enzyme activity.

BIOTECHNOLOGICAL UTILITY

Changing the cofactor specificity of the *T. ethanolicus* 2° ADH (eg., conversion from an $NADP^+$ to an NAD^+ linked enzyme), aside from yielding basic scientific protein structure function information, would construct a more valuable biocatalyst due to more desirable cofactor properties (eg., the greater thermal stability of NAD^+). Our collaborator, Dr. Robert Phillips, is examining the utility of this thermophilic 2° ADH in transforming ketones into chiral alcohols. Expanding the range of identified chiral reactions this 2° ADH performs and the effect of reaction conditions (eg., solvents and temperature) on both

activity and enantiospecificity also increases the likelyhood that it will be incorporated into a biotechnological production process. For very large substrates, solubility and the need for aggitation may be avoided by reversibly binding the substrate to a solid matrix, exposing it to enzyme in the presence of cofactor (similar to a chromatographic process), and releasing the product. This type of system would interface with the current diversimer organic synthesis technology used in drug development as the immobilized particulate enzyme could be separated from large or small scale batch reactants by centrifugation and it could be retained in a column for flow reacter systems. Furthermore, research identifying the exact molecular determinants of *T. ethanolicus* 2° ADH thermal and catalytic properties would allow tailoring the properties of this stable chiral oxidoreductase. The ability to modify this enzyme to perfom specific reactions under specific conditions identified by industry would make this a strong candidate for chiral industrial processes.

REFERENCES

1   Vieille, C., Burdette, D. S. and Zeikus, J. G. (1996) in Biotech. Ann. Rev., vol. 2, (M. R. El Geweley, ed.) pp. 1-83. Elsevier, Amsterdam, Neth.

2   Guinn, R. M., Skerker, P. S., Kavanaugh, P. and D. S. Clark (1991) Biotechnol. Bioeng. **37**, 303-308

APPENDIX A

Appendix A

Sequence Comparison Based Predictions of Important Catalytic Domain and Cofactor Binding Domain Amino Acids in 1° and 2° ADH Structures

Because of the relatively small number of exact 3-dimensional protein structures determined and the typically well conserved structural folds within each enzyme family, peptide sequence based comparisons are used to predict structurally analogous regions of these proteins. ADHs characteristically share significant 3-dimensional structural architecture but very little amino acid sequence similarity [1]. The horse liver 1° ADH x-ray structure has been determined and its amino acid structure-function characteristics extensively investigated [1]. HCA comparison of the thermophilic *T. ethanolicus* 39E 2° ADH, the thermophilic *B. stearothermophilus* 1° ADH, and the mesophilic *C. beijerinckii* and *A. eutrophus* 2° ADH peptides sequences to that of liver 1° ADH (Fig. 1) predicted a Rossmann fold for nicotinamide cofactor binding and specific amino acids associated with cofactor specificity and catalytic Zn liganding. The complete sequence alignments also predict corresponding residues to those that are part of identified liver ADH structural elements. Similarity among these structurally important residues, while not practically testable by point mutations, provides evidence in support of the folded-architectural similarity between the compared peptides. Tables 1 and 2 list the *T. ethanolicus* 2° ADH, *C. biejerinckii* 2° ADH, and *B. stearothermophilus* 1° ADH residues predicted to correspond to important residues identified in the horse liver enzyme. These tables also indicate the percent amino acid similarity and identity between these four peptides for specific sections of the horse liver ADH catalytic and cofactor binding structural domains.

Figure 1. HCA comparison of the *B. stearothermophilus* 1° ADH, the *T. ethanolicus* 2° ADH, and the *C. beijerinckii* 2° ADH amino Acid sequences to that of the horse liver 1° ADH. The proposed catalytic Zn liganding (O), structure Zn liganding (□), and Rossmann fold binding(□) residues are indicated. Symbols:✶, proline;◆, glycine;■, serine; □, threonine

T. ethanolicus 39E 2° Adh

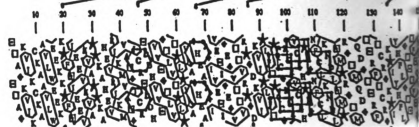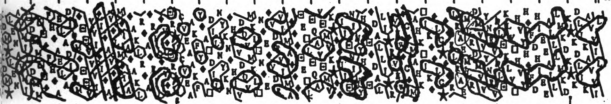C. beijerinckii 2° Adh
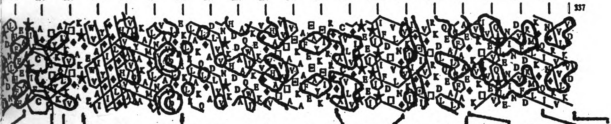
B. stearothermophilus 1° Adh

horse liver 1° Adh

Table 1. Comparison of aligned amino acids involved in ADH catalytic domain structure and function

Proposed role in liver ADH structure	Aligned peptide residues			
	Horse liver	*B. stearotherm.*	*C. beijerinckii*	*T. ethanolicus*
Active domain core residues	Ala11	Ala3	Leu13	Val12
	Ala12	Ala4	Gly14	Gly13
	Val13	Val5	Trp15	Trp14
	Leu14	Val6	Ile16	Ile15
	Phe21	Leu13	n.d.	n.d.
	Val26	N.D.	Val21	Ala21
	Val36	Val28	Ala27	Ala27
	Ile38	Val30	Val29	Val29
	Met40	Ile32	Pro31	Pro31
	Thr43	Cys45	Val34	Trp34
	Ile45	Val47	Pro36	Pro36
	Thr59	Pro53	Asp51	Glu51
	Val63	Ile57	Met55	Met55
	Ile64	Ile58	Ile56	Ile56
	Ala65	Pro59	Leu57	Leu57
	Ala69	Gly63	Ala61	Ala61
	Ala70	Val64	Val62	Val62
	Val73	Ile67	Val66	Val66
	Val80	Val74	Val72	Val72
	Val83	Leu77	Phe75	Phe75
	Val89	Val83	Ile82	Val82
	Ile90	Ile85	Val83	Val83
	Pro91	Pro86	Pro84	Pro84
	Val169	Ala143	Ala145	Ala145
	Leu171	Pro145	Met147	Met147
	Ile172	Ile146	Ile148	Ile148
	Gly173	Phe147	Thr149	Pro149
	Phe176	Val151	Met152	Met152
	Ala183	Ala155	Gly158	Gly158
	Val186	Val158	Leu161	Leu161
	Ala187	Thr159	Ala162	Ala162
	Val189	Ala161	Ile164	Ile164
Amino acid similarity (%)				
Horse liver	100(100)	38(84)	25(80)	25(80)
B. stearotherm.		100(100)	22(72)	25(75)
C. beijerinckii			100(100)	78(94)
T. ethanolicus				100(100)
Active site pocket residues	Ser48	Thr40	Ser39	Ser39
	Leu57	Pro50	Leu49	Ile49
	Val58	Val51	Gly50	Gly50
	Phe93	Phe89	Thr86	Ile86
	Phe110	Leu105	Leu93	Ser93
	Leu116	Tyr114	Phe99	Tyr99
	Thr178	Thr152	Ser154	Thr154

Amino acid similarity (%)				
Horse liver	100(100)	43(86)	28(78)	28(78)
B. stearotherm.		100(100)	14(86)	28(71)
C. beijerinckii			100(100)	78(100)
T. ethanolicus				100(100)

Catalytic Zn ligands	Cys46	Cys38	Cys37	Cys37
	His67	His62	His59	His59
	Cys174	Cys148	Asp150	Asp150
Amino acid similarity (%)				
Horse liver	100(100)	100(100)	67(67)	67(67)
B. stearotherm.		100(100)	67(67)	67(67)
C. beijerinckii			100(100)	100(100)
T. ethanolicus				100(100)

Structural Zn liganding residues	Cys97	Cys92	Cys85	n.d.
	Cys100	Cys95	n.d.	n.d.
	Cys103	Cys98	n.d.	n.d.
	Cys111	Cys106	n.d.	n.d.
Amino acid similarity (%)				
Horse liver	100(100)	100(100)	25(25)	------
B. stearotherm.		100(100)	25(25)	------
C. beijerinckii			100(100)	------
T. ethanolicus				------

Structural Gly residues	Gly66	Gly60	Gly60	Gly60
	Gly77	Gly71	Gly69	Gly69
	Gly86	Gly80	Gly78	Gly78
Amino acid similarity (%)				
Horse liver	100(100)	100(100)	100(100)	100(100)
B. stearotherm.		100(100)	100(100)	100(100)
C. beijerinckii			100(100)	100(100)
T. ethanolicus				100(100)

Overall amino acid similarity (%)				
Horse liver	100(100)	51(88)	32(76)	30(74)
B. stearotherm.		100(100)	29(25)	30(69)
C. beijerinckii			100(100)	79(94)
T. ethanolicus				100(100)

n.d.: no analogous residue was identified.

Table 2. Comaprison of aligned amino acids involved in ADH nicotinamide cofactor binding

Proposed role in liver ADH ADP- ribose binding	Aligned peptide residues			
	Horse liver	*B. stearotherm.*	*C. beijerinckii*	*T. ethanolicus*
Core residues	Ala183	Thr153	Ala159	Ala159
	Val186	Tyr154	Leu161	Leu161
	Ala187	Ala156	Ala162	Ala162
	Val189	Val159	Ile164	Ile164
	Cys195	Val168	Val170	Val170
	Val197	Ile170	Val172	Val172
	Leu200	Ile203	Ile175	Ile175
	Val203	Leu206	Val178	Val178
	Ala237	Ala209	Ala212	Ala212
	Phe264	Val231	aArg238	Ala238
	Val268	Val235	Ala242	Ala242
	Val288	Ala256	Ser263	Ala263
	Val290	Val258	Ile265	Val265
	Val292	Val260	Tyr267	Tyr267
	Ala317	Ala295	Gly239	Gly293
	Ala196	Ala169	Val171	Ala171
	Phe198	Tyr171	Ile173	Leu173
	Ile220	Val192	Ile195	Ile195
	Val222	Val194	Val197	Va197
	Ile250	His217	Ile263	Ile263
	Leu254	Ala221	Val227	Ile227
	Val262	Gly229	Val236	Val236
	Ala278	Ser250	Met255	Ile255
	Cys281	Arg252	Gly259	Gly259
	Cys282	Arg253	Gly260	Gly260
Amino acid similarity (%)				
Horse liver	100(100)	36(68)	36(68)	44(76)
B. stearotherm.		100(100)	20(60)	32(68)
C. beijerinckii			100(100)	76(88)
T. ethanolicus				100(100)
Adenine pocket residues	Phe198	Tyr171	Ile173	Leu173
	Val222	Val194	Val197	Val197
	Ile224	Leu196	Thr199	Thr199
	Pro243	Pro215	Tyr218	Tyr218
	Ile250	His217	Ile223	Ile223
	Ile269	Val236	Gly243	Gly243
	Thr274	Thr237	Gly244	Gly244
	Arg271	Ala238	Ser246	Ala246
Adenosine ribose binding residues	Asp223	Asp195	Gly198	Gly198
	Gly199	Gly172	Gly174	Gly174
	Ile269	Val236	Gly243	Gly243

	Asn225 Lys228	Gly197 Lys200	Arg200 Cys203	Arg200 Cys203
Pyrophosphate interacting	Arg47 Ile269	His39 Val236	Thr38 Gly243	Thr38 Gly243
Nicotinamide ribose biding	Gly293	Gly261	His268	Phe268
[b]Amino acid similarity (%)				
Horse liver	100(100)	50(78)	21(28)	21(28)
B. stearotherm.		100(100)	14(28)	21(36)
C. beijerinckii			100(100)	78(86)
T. ethanolicus				100(100)
Overall amino acid similarity (%)				
Horse liver	100(100)	41(72)	30(52)	35(57)
B. stearotherm.		100(100)	18(48)	28(56)
C. beijerinckii			100(100)	77(87)
T. ethanolicus				100(100)

[a]Possibly a sequencing error.
[b]Ile269 is only counted once in the calculation.

REFERENCES

1   Brändén, C.-I., Jörnvall, H., Eklund, H., and B. Fururgren (1975) Alcohol dehydrogenases. *In* The enzymes, 3rd ed., vol. XI part A" p. 105-106, (P. D. Boyer, ed.). Academic Press, New York.

.

APPENDIX B

Appendix B

Computer programs used to calculate (hyper)thermophilic versus mesophilic total amino acid composition and the theoretical Arrhenius data for thermophilic and mesophilic enzymes

Computer program to compile amino acid composition data from peptide sequences and to predict percent enzyme activity values at temperatures below the maximal temperature for mesophilic or thermophilic enzyme activity were written using Turbo Pascal 3.0 for IBM (Borland International; Scotts Valley, CA). Programs were compiled and executed using an IBM PC. Program "TOTALAA" calculates the percent amino acid composition of a peptide sequence input as a string of letters in an ascii format text file, and it calculates the average amino acid composition of a group of peptide sequence files input in succession (used to calculate the values reported in chapter 1; Table 3). The output is written to the default line printer and to the monitor. Program "Arrheniussimulation" calculates the expected precent catalytic activities within a temperature range below the temperature for optimal activity (Topt) specified by the low temperature (Told) in Kelvin (see chapter 1; Fig. 1). The calculation can be performed on enzymes with any reaction activation energy (Ea) in kJ/mol. The output data is written to an ascii format text file named "ARRSIM.DAT".

```
program TOTALAA;

VAR

COMPARE: ARRAY [1..20] OF RECORD

   LABELC: STRING[25];

   NGC: REAL;

   NPC: REAL;

   NAC: REAL;

   NVC: REAL;

   NLC: REAL;

   NIC: REAL;

   NFC: REAL;

   NWC: REAL;

   NYC: REAL;

   NHC: REAL;

   NCC: REAL;

   NMC: REAL;

   NSC: REAL;

   NTC: REAL;

   NKC: REAL;

   NRC: REAL;

   NDC: REAL;

   NEC: REAL;

   NNC: REAL;

   NQC: REAL;

   END;

FILEIN: TEXT;

NAMEIN,CATEGORY: STRING[25];
```

```
AA,PAUSE,YNT: STRING[1];

i,J,NG,NP,NA,NV,NL,NI,NF,NW,NY,NH,NC,NM,NS,NT,NK,NR,ND,NE,NN,NQ
,TOT: INTEGER;

NGT,NPT,NAT,NVT,NLT,NIT,NFT,NWT,NYT,NHT,NCT,NMT,NST,NTT,NKT,N
RT,NDT,NET,NNT,NQT: REAL;

SQNGT,SQNPT,SQNAT,SQNVT,SQNLT,SQNIT,SQNFT,SQNWT,SQNYT,SQNHT,
SQNCT,SQNMT,SQNST,SQNTT: REAL;

SQNKT,SQNRT,SQNDT,SQNET,SQNNT,SQNQT: REAL;

MGT,MPT,MAT,MVT,MLT,MIT,MFT,MWT,MYT,MHT,MCT,MMT,MST,MTT,MKT
,MRT,MDT,MET,MNT,MQT: REAL;

GTSD,PTSD,ATSD,VTSD,LTSD,ITSD,FTSD,WTSD,YTSD,HTSD,CTSD,MTSD,STS
D,TTSD,KTSD,RTSD,DTSD,ETSD,NTSD,QTSD: REAL;

TOTR: REAL;

YN,X: BOOLEAN;


PROCEDURE YESNO;
  VAR
    CK: BOOLEAN;
  BEGIN
  REPEAT
    IF YNT = 'Y' THEN
      BEGIN
        YN := TRUE;
        CK := TRUE;
      END
    ELSE IF YNT = 'y' THEN
      BEGIN
```

```
            YN := TRUE;

            CK := TRUE;

         END

      ELSE IF YNT = 'N' THEN

         BEGIN

            YN := FALSE;

            CK := TRUE;

         END

      ELSE IF YNT = 'n' THEN

         BEGIN

            YN := FALSE;

            CK := TRUE;

         END

      ELSE

         BEGIN

            CK := FALSE;

            WRITELN;

            WRITELN('PLEASE ENTER EITHER Y OR N');

            READ(YNT);

            WRITELN;

         END;

   UNTIL CK = TRUE;

   END;


FUNCTION RD(A:REAL): CHAR;

VAR

S,SO: STRING[18];
```

```
i,M,P: INTEGER;

Y,YY: CHAR;


BEGIN

   STR(A,S);

   FOR i:= 1 TO 18 DO

      BEGIN

         IF S[i] = '.' THEN

          BEGIN

             M := i + 3;

             P := i;

          END;

         IF S[i] = 'E' THEN

          BEGIN

             IF S[i+3] = '1' THEN

             BEGIN

              Y := S[P+1];

              S[P] := Y;

               S[P+1] :='.';

             END;

              IF S[i+3] = '3' THEN

             BEGIN

              Y := S[P-1];

               S[P-1] := '.';

              S[P] := Y;

             END;

              IF S[i+3] = '4' THEN
```

```
      BEGIN

        FOR i := 1 TO 18 DO

          BEGIN

          SO[i] := S[i];

          END;

          Y := SO[3];

            S[3] := '.';

            S[4] := '0';

          S[5] := SO[3];

          S[6] := SO[5];

          S[7] := SO[6];

          S[8] := SO[7];

          M := 8;

        END;

      END;

    END;

  FOR i := 1 TO M DO

    BEGIN

      WRITE(S[i]);

    END;

  RD := ' ';

END;


FUNCTION PRD(A:REAL): CHAR;

VAR

S: STRING[18];

i,M,P: INTEGER;
```

```
Y: CHAR;

BEGIN
    STR(A,S);
    FOR i:= 1 TO 18 DO
        BEGIN
            IF S[i] = '.' THEN
                BEGIN
                    M := i + 3;
                    P := i;
                END;
            IF S[i] = 'E' THEN
                BEGIN
                    IF S[i+3] = '1' THEN
                        BEGIN
                            Y := S[P+1];
                            S[P] := Y;
                            S[P+1] := '.';
                        END;
                END;
        END;
    FOR i := 1 TO M DO
        BEGIN
            WRITE(LST,S[i]);
        END;
    PRD := ' ';
END;
```

```
BEGIN

REPEAT

i := 1;

WRITELN('INPUT THE LABEL FOR THIS SET OF FILES (25 CHAR OR LESS)');

READ(CATEGORY);

WRITELN;

 REPEAT

   WRITELN('ENTER INPUT FILE NAME');

   READLN(NAMEIN);

   ASSIGN(FILEIN,NAMEIN);

   RESET(FILEIN);

   CLRSCR;

   NA := 0;

   NG := 0;

   NP := 0;

   NV := 0;

   NL := 0;

   NI := 0;

   NF := 0;

   NW := 0;

   NY := 0;

   NH := 0;

   NC := 0;

   NM := 0;

   NS := 0;

   NT := 0;
```

```pascal
NK := 0;

NR := 0;

ND := 0;

NE := 0;

NN := 0;

NQ := 0;

TOT := 0;

WHILE NOT EOF(FILEIN) DO

BEGIN

    READ(FILEIN,AA);

    IF UPCASE(AA)='A' THEN

    BEGIN

        NA := NA + 1;

    END

    ELSE IF  UPCASE(AA) = 'G' THEN

    BEGIN

        NG := NG + 1;

    END

    ELSE IF UPCASE(AA) = 'P' THEN

    BEGIN

        NP := NP + 1;

    END

    ELSE IF  UPCASE(AA) = 'V' THEN

    BEGIN

        NV := NV + 1;

    END

    ELSE IF UPCASE(AA) = 'L' THEN
```

```
BEGIN

   NL := NL + 1;

END

ELSE IF  UPCASE(AA) = 'I' THEN

BEGIN

   NI := NI + 1;

END

ELSE IF UPCASE(AA) = 'F' THEN

BEGIN

   NF := NF + 1;

END

ELSE IF  UPCASE(AA) = 'W' THEN

BEGIN

   NW := NW + 1;

END

ELSE IF UPCASE(AA) = 'Y' THEN

BEGIN

   NY := NY + 1;

END

ELSE IF  UPCASE(AA) = 'H' THEN

BEGIN

   NH := NH + 1;

END

ELSE IF UPCASE(AA) = 'C' THEN

BEGIN

   NC := NC + 1;

END
```

```
ELSE IF  UPCASE(AA) = 'M' THEN
BEGIN
   NM := NM + 1;
END
ELSE IF UPCASE(AA) = 'S' THEN
BEGIN
   NS := NS + 1;
END
ELSE IF  UPCASE(AA) = 'T' THEN
BEGIN
   NT := NT + 1;
END
ELSE IF UPCASE(AA) = 'K' THEN
BEGIN
   NK := NK + 1;
END
ELSE IF  UPCASE(AA) = 'R' THEN
BEGIN
   NR := NR + 1;
END
ELSE IF UPCASE(AA) = 'D' THEN
BEGIN
   ND := ND + 1;
END
ELSE IF  UPCASE(AA) = 'E' THEN
BEGIN
   NE := NE + 1;
```

```
    END

     ELSE IF UPCASE(AA) = 'N' THEN

     BEGIN

        NN := NN + 1;

     END

     ELSE IF UPCASE(AA) = 'Q' THEN

     BEGIN

        NQ := NQ + 1;

     END;

     TOT := TOT + 1;

     TOTR := TOT;

     WRITE(AA);

  END;

  WRITELN;

  CLOSE(FILEIN);

  WITH COMPARE[i] DO

     BEGIN

     LABELC := NAMEIN;

     NGC:= NG/TOTR;

     NPC:= NP/TOTR;

     NAC:= NA/TOTR;

     NVC:= NV/TOTR;

     NLC:= NL/TOTR;

     NIC:= NI/TOTR;

     NFC:= NF/TOTR;

     NWC:= NW/TOTR;

     NYC:= NY/TOTR;
```

```
    NHC:= NH/TOTR;

    NCC:= NC/TOTR;

    NMC:= NM/TOTR;

    NSC:= NS/TOTR;

    NTC:= NT/TOTR;

    NKC:= NK/TOTR;

    NRC:= NR/TOTR;

    NDC:= ND/TOTR;

    NEC:= NE/TOTR;

    NNC:= NN/TOTR;

    NQC:= NQ/TOTR;

    END;

WRITELN('PRESS ANY KEY TO CONTINUE');

READ(PAUSE);

CLRSCR;

X := FALSE;

IF X = TRUE THEN

BEGIN

WRITELN('GLY = ',NG);

WRITELN('PRO = ',NP);

WRITELN('ALA = ',NA);

WRITELN('VAL = ',NV);

WRITELN('LEU = ',NL);

WRITELN('ILE = ',NI);

WRITELN('PHE = ',NF);

WRITELN('TRP = ',NW);

WRITELN('TYR = ',NY);
```

```
WRITELN('HIS = ',NH);

WRITELN('CYS = ',NC);

WRITELN('MET = ',NM);

WRITELN('SER = ',NS);

WRITELN('THR = ',NT);

WRITELN('LYS = ',NK);

WRITELN('ARG = ',NR);

WRITELN('ASP = ',ND);

WRITELN('GLU = ',NE);

WRITELN('ASN = ',NN);

WRITELN('GLN = ',NQ);

WRITELN;

WRITELN('TOT = ',TOT);

READ(PAUSE);

END;

CLRSCR;

GOTOXY(1,1);

WRITE(NAMEIN);

GOTOXY(1,2);

WRITE('AMINO ACID');

GOTOXY(15,2);

WRITE('NUMBER');

GOTOXY(30,2);

WRITE ('PERCENT COMPOSITION');

GOTOXY(1,3);

WRITELN('GLY',' ',NG,' ',RD(NG/TOTR));

WRITELN('PRO',' ',NP,' ',RD(NP/TOTR));
```

```
WRITELN('ALA','          ',NA,'          ',RD(NA/TOTR));

WRITELN('VAL','          ',NV,'          ',RD(NV/TOTR));

WRITELN('LEU','          ',NL,'          ',RD(NL/TOTR));

WRITELN('ILE','          ',NI,'          ',RD(NI/TOTR));

WRITELN('PHE','          ',NF,'          ',RD(NF/TOTR));

WRITELN('TRP','          ',NW,'          ',RD(NW/TOTR));

WRITELN('TYR','          ',NY,'          ',RD(NY/TOTR));

WRITELN('HIS','          ',NH,'          ',RD(NH/TOTR));

WRITELN('CYS','          ',NC,'          ',RD(NC/TOTR));

WRITELN('MET','          ',NM,'          ',RD(NM/TOTR));

WRITELN('SER','          ',NS,'          ',RD(NS/TOTR));

WRITELN('THR','          ',NT,'          ',RD(NT/TOTR));

WRITELN('LYS','          ',NK,'          ',RD(NK/TOTR));

WRITELN('ARG','          ',NR,'          ',RD(NR/TOTR));

WRITELN('ASP','          ',ND,'          ',RD(ND/TOTR));

WRITELN('GLU','          ',NE,'          ',RD(NE/TOTR));

WRITELN('ASN','          ',NN,'          ',RD(NN/TOTR));

WRITELN('GLN','          ',NQ,'          ',RD(NQ/TOTR));

    WRITELN('H-PHOBIC','
',RD((NG+NA+NL+NI+NV+NP+NF+NW+NY)/TOTR));

    WRITELN('POLAR','                    ',RD((NS+NT+NH+NM+NC)/TOTR));

    WRITELN('CHARGED','                    ',
RD((NK+NR+NQ+NN+NE+ND)/TOTR));

  X := FALSE;

  IF X = TRUE THEN

  BEGIN

  WRITELN(LST,NAMEIN);
```

```
WRITELN(LST,'AMINO ACID','    ','NUMBER','         ','% COMP');

WRITELN(LST,'GLY','        ',NG,'        ',PRD(NG/TOTR));

WRITELN(LST,'PRO','        ',NP,'        ',PRD(NP/TOTR));

WRITELN(LST,'ALA','        ',NA,'        ',PRD(NA/TOTR));

WRITELN(LST,'VAL','        ',NV,'        ',PRD(NV/TOTR));

WRITELN(LST,'LEU','        ',NL,'        ',PRD(NL/TOTR));

WRITELN(LST,'ILE','        ',NI,'        ',PRD(NI/TOTR));

WRITELN(LST,'PHE','        ',NF,'        ',PRD(NF/TOTR));

WRITELN(LST,'TRP','        ',NW,'        ',PRD(NW/TOTR));

WRITELN(LST,'TYR','        ',NY,'        ',PRD(NY/TOTR));

WRITELN(LST,'HIS','        ',NH,'        ',PRD(NH/TOTR));

WRITELN(LST,'CYS','        ',NC,'        ',PRD(NC/TOTR));

WRITELN(LST,'MET','        ',NM,'        ',PRD(NM/TOTR));

WRITELN(LST,'SER','        ',NS,'        ',PRD(NS/TOTR));

WRITELN(LST,'THR','        ',NT,'        ',PRD(NT/TOTR));

WRITELN(LST,'LYS','        ',NK,'        ',PRD(NK/TOTR));

WRITELN(LST,'ARG','        ',NR,'        ',PRD(NR/TOTR));

WRITELN(LST,'ASP','        ',ND,'        ',PRD(ND/TOTR));

WRITELN(LST,'GLU','        ',NE,'        ',PRD(NE/TOTR));

WRITELN(LST,'ASN','        ',NN,'        ',PRD(NN/TOTR));

WRITELN(LST,'GLN','        ',NQ,'        ',PRD(NQ/TOTR));

WRITELN(LST,'H-PHOBIC','
',PRD((NG+NA+NL+NI+NV+NP+NF+NW+NY)/TOTR));

WRITELN(LST,'POLAR','
',PRD((NS+NT+NH+NM+NC)/TOTR));

WRITELN(LST,'CHARGED','
',PRD((NK+NR+NQ+NN+NE+ND)/TOTR));
```

```
END;

READ(PAUSE);

CLRSCR;

WRITELN('ADD ANOTHER SEQUENCE? (Y/N)');

READ(YNT);

WRITELN;

YESNO;

J := i;

i := i + 1;

UNTIL YN = FALSE;

WRITELN('NUMBER OF SEQUENCES = ',J);

READ(PAUSE);

    NAT := 0;

    NGT := 0;

    NPT := 0;

    NVT := 0;

    NLT := 0;

    NIT := 0;

    NFT := 0;

    NWT := 0;

    NYT := 0;

    NHT := 0;

    NCT := 0;

    NMT := 0;

    NST := 0;

    NTT := 0;

    NKT := 0;
```

```
        NRT := 0;

        NDT := 0;

        NET := 0;

        NNT := 0;

        NQT := 0;

        SQNAT := 0;

        SQNGT := 0;

        SQNPT := 0;

        SQNVT := 0;

        SQNLT := 0;

        SQNIT := 0;

        SQNFT := 0;

        SQNWT := 0;

        SQNYT := 0;

        SQNHT := 0;

        SQNCT := 0;

        SQNMT := 0;

        SQNST := 0;

        SQNTT := 0;

        SQNKT := 0;

        SQNRT := 0;

        SQNDT := 0;

        SQNET := 0;

        SQNNT := 0;

        SQNQT := 0;

   FOR i := 1 TO J DO

   BEGIN
```

```
WITH COMPARE[i] DO

  BEGIN

        NGT := NGT + NGC;

        SQNGT := SQNGT + SQR(NGC);

        NPT := NPT + NPC;

        SQNPT := SQNPT + SQR(NPC);

        NAT := NAT + NAC;

        SQNAT := SQNAT + SQR(NAC);

        NLT := NLT + NLC;

        SQNLT := SQNLT + SQR(NLC);

        NIT := NIT + NIC;

        SQNIT := SQNIT + SQR(NIC);

        NVT := NVT + NVC;

        SQNVT := SQNVT + SQR(NVC);

        NYT := NYT + NYC;

        SQNYT := SQNYT + SQR(NYC);

        NFT := NFT + NFC;

        SQNFT := SQNFT + SQR(NFC);

        NWT := NWT + NWC;

        SQNWT := SQNWT + SQR(NWC);

        NHT := NHT + NHC;

        SQNHT := SQNHT + SQR(NHC);

        NCT := NCT + NCC;

        SQNCT := SQNCT + SQR(NCC);

        NMT := NMT + NMC;

        SQNMT := SQNMT + SQR(NMC);

        NST := NST + NSC;
```

```
        SQNST := SQNST + SQR(NSC);

        NTT := NTT + NTC;

        SQNTT := SQNTT + SQR(NTC);

        NRT := NRT + NRC;

        SQNRT := SQNRT + SQR(NRC);

        NKT := NKT + NKC;

        SQNKT := SQNKT + SQR(NKC);

        NNT := NNT + NNC;

        SQNNT := SQNNT + SQR(NNC);

        NQT := NQT + NQC;

        SQNQT := SQNQT + SQR(NQC);

        NET := NET + NEC;

        SQNET := SQNET + SQR(NEC);

        NDT := NDT + NDC;

        SQNDT := SQNDT + SQR(NDC);

      END;

   END;

   MGT := NGT/J;

   MPT := NPT/J;

   MAT := NAT/J;

   MLT := NLT/J;

   MIT := NIT/J;

   MVT := NVT/J;

   MYT := NYT/J;

   MFT := NFT/J;

   MWT := NWT/J;

   MHT := NHT/J;
```

MCT := NCT/J;

MMT := NMT/J;

MST := NST/J;

MTT := NTT/J;

MRT := NRT/J;

MKT := NKT/J;

MNT := NNT/J;

MQT := NQT/J;

MET := NET/J;

MDT := NDT/J;

GTSD := SQRT((J*SQNGT - SQR(NGT))/(J*(J-1)));

PTSD := SQRT((J*SQNPT - SQR(NPT))/(J*(J-1)));

ATSD := SQRT((J*SQNAT - SQR(NAT))/(J*(J-1)));

LTSD := SQRT((J*SQNLT - SQR(NLT))/(J*(J-1)));

ITSD := SQRT((J*SQNIT - SQR(NIT))/(J*(J-1)));

VTSD := SQRT((J*SQNVT - SQR(NVT))/(J*(J-1)));

YTSD := SQRT((J*SQNYT - SQR(NYT))/(J*(J-1)));

FTSD := SQRT((J*SQNFT - SQR(NFT))/(J*(J-1)));

WTSD := SQRT((J*SQNWT - SQR(NWT))/(J*(J-1)));

HTSD := SQRT((J*SQNHT - SQR(NHT))/(J*(J-1)));

CTSD := SQRT((J*SQNCT - SQR(NCT))/(J*(J-1)));

MTSD := SQRT((J*SQNMT - SQR(NMT))/(J*(J-1)));

STSD := SQRT((J*SQNST - SQR(NST))/(J*(J-1)));

TTSD := SQRT((J*SQNTT - SQR(NTT))/(J*(J-1)));

RTSD := SQRT((J*SQNRT - SQR(NRT))/(J*(J-1)));

KTSD := SQRT((J*SQNKT - SQR(NKT))/(J*(J-1)));

NTSD := SQRT((J*SQNNT - SQR(NNT))/(J*(J-1)));

```
QTSD := SQRT((J*SQNQT - SQR(NQT))/(J*(J-1)));

ETSD := SQRT((J*SQNET - SQR(NET))/(J*(J-1)));

DTSD := SQRT((J*SQNDT - SQR(NDT))/(J*(J-1)));

CLRSCR;

    GOTOXY(60,2);

    WRITE(CATEGORY);

    GOTOXY(1,2);

    WRITE('AMINO ACID');

    GOTOXY(22,2);

    WRITE('MEAN');

    GOTOXY(42,2);

    WRITE ('STD DEV');

    GOTOXY(1,3);

    WRITELN('GLY','          ',RD(MGT),'          ',RD(GTSD));

    WRITELN('PRO','          ',RD(MPT),'          ',RD(PTSD));

    WRITELN('ALA','          ',RD(MAT),'          ',RD(ATSD));

    WRITELN('VAL','          ',RD(MVT),'          ',RD(VTSD));

    WRITELN('LEU','          ',RD(MLT),'          ',RD(LTSD));

    WRITELN('ILE','          ',RD(MIT),'          ',RD(ITSD));

    WRITELN('PHE','          ',RD(MFT),'          ',RD(FTSD));

    WRITELN('TRP','          ',RD(MWT),'          ',RD(WTSD));

    WRITELN('TYR','          ',RD(MYT),'          ',RD(YTSD));

    WRITELN('HIS','          ',RD(MHT),'          ',RD(HTSD));

    WRITELN('CYS','          ',RD(MCT),'          ',RD(CTSD));

    WRITELN('MET','          ',RD(MMT),'          ',RD(MTSD));

    WRITELN('SER','          ',RD(MST),'          ',RD(STSD));

    WRITELN('THR','          ',RD(MTT),'          ',RD(TTSD));
```

```
WRITELN('LYS','             ',RD(MKT),'           ',RD(KTSD));

WRITELN('ARG','             ',RD(MRT),'           ',RD(RTSD));

WRITELN('ASP','             ',RD(MDT),'           ',RD(DTSD));

WRITELN('GLU','             ',RD(MET),'           ',RD(ETSD));

WRITELN('ASN','             ',RD(MNT),'           ',RD(NTSD));

WRITELN('GLN','             ',RD(MQT),'           ',RD(QTSD));

    WRITELN('H-PHOBIC','
',RD(MGT+MAT+MLT+MIT+MVT+MPT+MFT+MWT+MYT));

    WRITELN('POLAR','         ',RD(MST+MTT+MHT+MMT+MCT));

    WRITELN('CHARGED','          ',RD(MKT+MRT+MQT+MNT+MET+MDT));

    GOTOXY(60,4);

    WRITE('FILES USED');

    FOR i := 1 TO J DO

    BEGIN

    WITH COMPARE[i] DO

    BEGIN

    GOTOXY(60,(i + 4));

    WRITE(LABELC);

    END;

    END;

READ(PAUSE);

CLRSCR;

WRITELN('ANOTHER SET OF FILES (Y/N)?');

READ(YNT);

WRITELN;

YESNO;

UNTIL YN = FALSE;
```

END.

program Arrheniussimulation;

VAR

DATOUT: TEXT;

X,Y,Z: STRING[20];

TOPT,EA,DT,T,FK,TOLD: REAL;

TF: BOOLEAN;

BEGIN

ASSIGN(DATOUT,'ARRSIM.DAT');

WRITELN('INPUT Topt (K)');

READ(TOPT);

WRITELN;

WRITELN('INPUT Ea (kJ/MOL)');

READ(EA);

WRITELN;

TOLD := 200;

TF := FALSE;

RESET(DATOUT);

APPEND(DATOUT);

WRITELN(DATOUT,'Topt = ',TOPT,' ','Ea = ',EA);

WRITELN(DATOUT);

REPEAT

DT := TOPT-TOLD;

FK := EXP((((EA*1000)/8.314)*((1/TOPT) - (1/TOLD))));

```
STR(TOLD,X);

STR(DT,Y);

STR(FK,Z);

WRITELN(DATOUT,X,Y,Z);

T := TOLD + 2;

IF TOLD >= (TOPT) THEN

      BEGIN

      TF := TRUE; END;

TOLD := T;

UNTIL TF=TRUE;

CLOSE(DATOUT);

END.
```

APPENDIX C

Appendix C

Construction and Kinetic Characterization of *Thermoanaerobacter ethanolicus* 39E 2° ADH Proline Deficient Mutants

INTRODUCTION

Expression of the genes encoding thermophilic and thermostable enzymes in mesophilic hosts typically yields recombinant enzymes with native-like thermal properties. Therefore, the molecular determinants of these characteristics appear to be in the amino acid sequences themselves. The general amino acid compositional similarity between mesophilic and thermophilic proteins further suggests that the positions of specific amino acids and not their number are critical to enzyme thermal stability. Finally, the significant energetic contributions of individual amino acids to the marginal stabilizing energy of folded protein structures indicates that, similar to proper protein folding, enzyme thermostability is due predominantly to energetic contributions from a subset of peptide amino acids. However, the identities and positions of these critical amino acid residues cannot yet be predicted even with exact 3-dimensional structure information.

Proline residues have been proposed to constrain protein surface loops and this is proposed to stabilize the interactions between the associated hydrophobic core elements [1]. This proline-zipper hypothesis argues that by preventing the disassociation of the surface most inter-core-element interactions, a constrained loop thermostabilize the folded protein. Loop engineering has been shown to thermostabilize protein folded structures (see chapter 1; prolines and loop regions). Furthermore, introducing Pro residues into turn regions has

thermostabilized both mesophilic and thermophilic proteins (see chapter 1; prolines and loop regions). The identification of thermophilic enzymes containing additional Pro residues compared to their mesophilic counterparts has suggesterd that this is also a protein thermostabilization strategy used in nature (see chapter 1; prolines and loop regions). The availability of functionally analogous thermophilic and mesophilic proteins with highly similar peptide sequences would allow comparative analysis to identify potentially critical Pro residues.

The recently cloned thermophilic *Thermoanaerobacter ethanolicus* 2° ADH amino acid sequence differs from that deduced for the mesophilic *Clostridium beijerinckii* 2° ADH from its cloned gene at 89 positions (see chapter 2). Of these, approximately 12 are nonconservative and 9 of these are Pro residues in putative loop regions of the more thermophilic enzyme. The overall sequence similarity among these 2 peptides is >85% (See chapter 2). This is significantly higher than the similarity seen between the *T. ethanolicus* 2° ADH and the horse liver 1° ADH (54%), yet mutagenic structure-function analysis suggested that they were also architecturally analogous (see chapter 4). Because the mesophilic *C. beijerinckii* 2° ADH shares greater peptide sequence identity to the thermophilic *T. ethanolicus* enzyme than the horse liver 1° ADH, the mesophilic *C. beijerinckii* 2° ADH is also predicted to share greater folded structural similarity with the thermophilic *T. ethanolicus* 2° ADH. Therefore, this is an excellent model system for mutagenic analysis to test the hypothesis that proline residues, constraining surface loops, can prevent hydrophobic core element separation and so thermostabilize proteins.

The data presented here examine the effect of mutating 3 individual *T. ethanolicus* 2° ADH added Pro residues to the corresponding *C. beijerinckii* enzyme residues predicted based on sequence alignments. The thermophilicities and thermostabilities of the wild type and proline mutant enzymes are reported.

MATERIALS AND METHODS

Chemicals and reagents

All chemicals were of at least reagent/molecular biology grade. Oligonucleotide synthesis and amino acid sequence analysis were performed by the Macromolecular Structure Facility (Department of Biochemistry, Michigan State University). The kanamycin resistance GenBlock (*Eco*RI) DNA cartridge used in expression vector construction was purchased from Pharmacia (Uppsala, Sweden). DNA for sequencing was isolated using the Wizard Miniprep kit (Promega; Madison, WI).

Media and strains

Escherichia coli (DH5α) containing the 2° ADH recombinant plasmids were grown in rich complex medium (20 g l⁻¹ tryptone, 10 g l⁻¹ yeast extract, 5 g l⁻¹ NaCl) at 37°C in the presence of 25 μg ml⁻¹ kanamycin and 100 μg ml⁻¹ ampicillin.

Mutagenesis

All DNA manipulations were performed using established protocols [2,3]. Point mutations were introduced into the *adhB* gene by PCR [2] using the *T. ethanolicus* 39E 2° ADH gene clonal plasmid pADHB25-kan (see chapter 2) as template DNA. An oligonucleotide primer (KA4 N-end) was synthesized to bind the noncoding strand and included a *Kpn*I restriction enzyme site, the native *adhB* gene ribosome assembly site, and the initiation codon for the *adhB* gene. An oligonucleotide primer (KA4 C-end) was synthesized to bind the coding strand, it included the compliment of the *adhB* termination codon and an *Apa*I restriction enzyme site. Complimentary 30-45 base oligonucleotide primers that contained the mutated bases were used in conjunction with 2 KA4 end primers to amplify the N-terminal and C-terminal segments of the *adhB* gene. PCR syntheses of partial and complete mutated genes were performed using the Taqplus, exonuclease containing polymerase (Strategene; La Jolla, CA). All clones were expressed in pBluescriptII KS(+) with a kanamycin resistance cartridge introduced into the polylinker *Eco*RI site. Mutations were verified by DNA sequencing using the method of Sanger *et al.* [4].

Enzyme kinetics

The standard 2° ADH activity assay was defined as NADP$^+$ reduction coupled to propan-2-ol oxidation at 60°C as previously described [5]. The enzyme was incubated at 55°C for 15 min prior to activity determination unless otherwise indicated. Tris buffer pH was adjusted at 25°C to be pH 8.0 at assay temperature (thermal correction factor = -0.031 ΔpH °C^{-1}). Protein concentrations were measured using the bicinchoninic acid (BCA) procedure (Pierce; Rockford, IL).
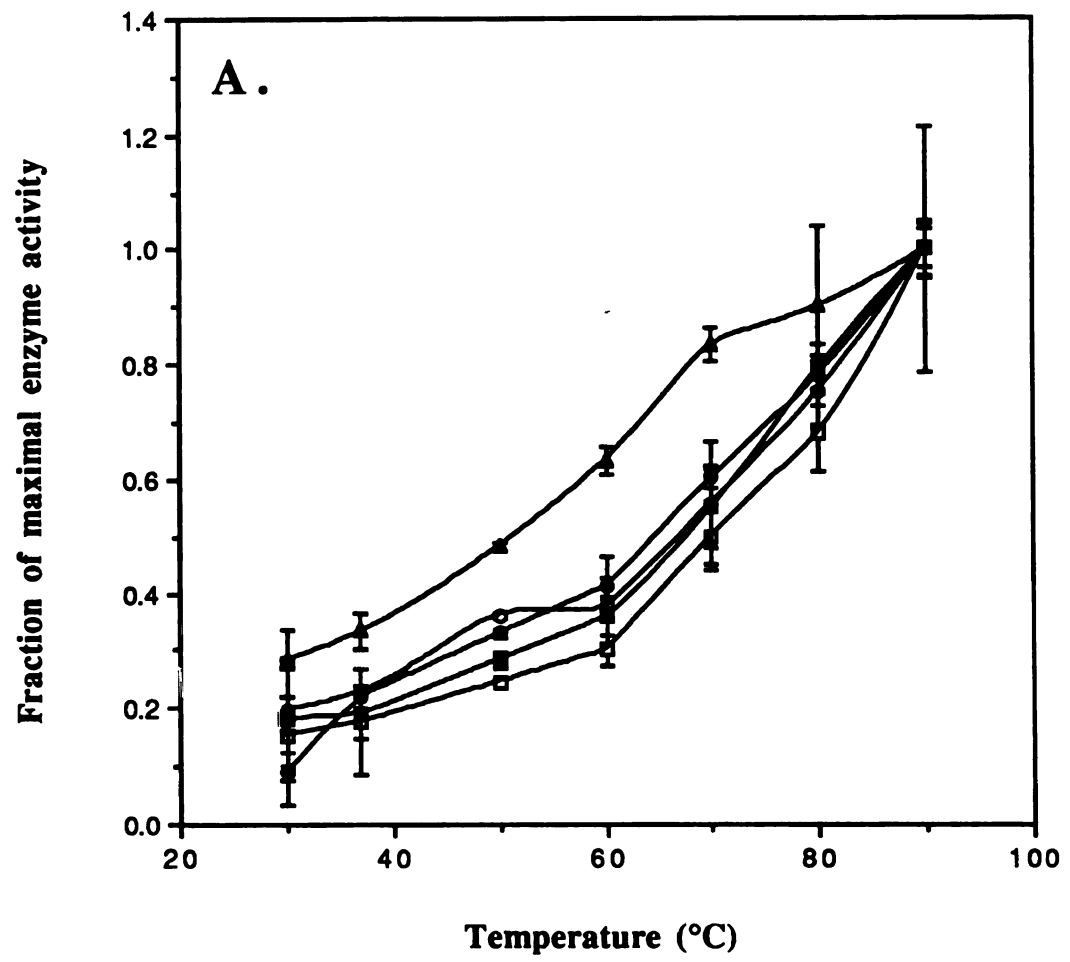
Thermostability

Recombinant 2° ADH thermostabilities were evaluated in *E. coli* cell extracts by timed incubation at the desired temperatures, followed by incubation for 30 min at 25°C. Incubations were performed in 100 µl PCR tubes (cat. #72.733.050, Sarstedt; Newton, NC) using 0.2 mg ml^{-1} protein in 100 µl of 50 mM Tris:HCl (pH 8.0). Activity was determined using the unfractionated samples.

RESULTS

The percent activities of the single and duplicated N-terminal Met containing wild type *T. ethanolicus* 2° ADHs and the Pro22 to Ala, Pro149 to Thr, and Pro222 to His mutants were compared at temperatures from 30°C to 90°C (Fig. 1A). The catalytic activities calculated for the cell extracts at 90°C ranged from 6.0 units per mg total protein for the Pro149 to Thr mutant to over 60 units per mg protein for the wild type enzymes. However, the percent activities increased similarly for all enzymes from 30°C to 90°C. The enzyme residual activities after timed incubations at 90°C were determined (Fig. 1B), indicating that thermoinactivation of both the wild type and mutant enzymes was not pseudo-first order in *E. coli* cell extracts. Also, the proline deficient mutant enzymes demonstrated greater thermostability than either of the wild type 2° ADHs.

**Figure 1. Thermophilicity and thermostability profiles for _T. ethanolicus_ 2°
ADH Pro residue mutants.** (A) The effect of temperature on propan-2-ol oxidation by
recombinant wild type (O), Single N-terminal Met (●), Pro24 to Ser (☐), Pro149 to Thr (▲),
and Pro222 to His (■), enzyme containing _E. coli_ DH5α cell extracts. (B) The effect of
timed incubations at 90°C on wild type (O), Single N-terminal Met (●), Pro24 to Ser (☐),
Pro149 to Thr (▲), and Pro222 to His (■), enzyme containing _E. coli_ DH5α cell extract
residual percent activities.
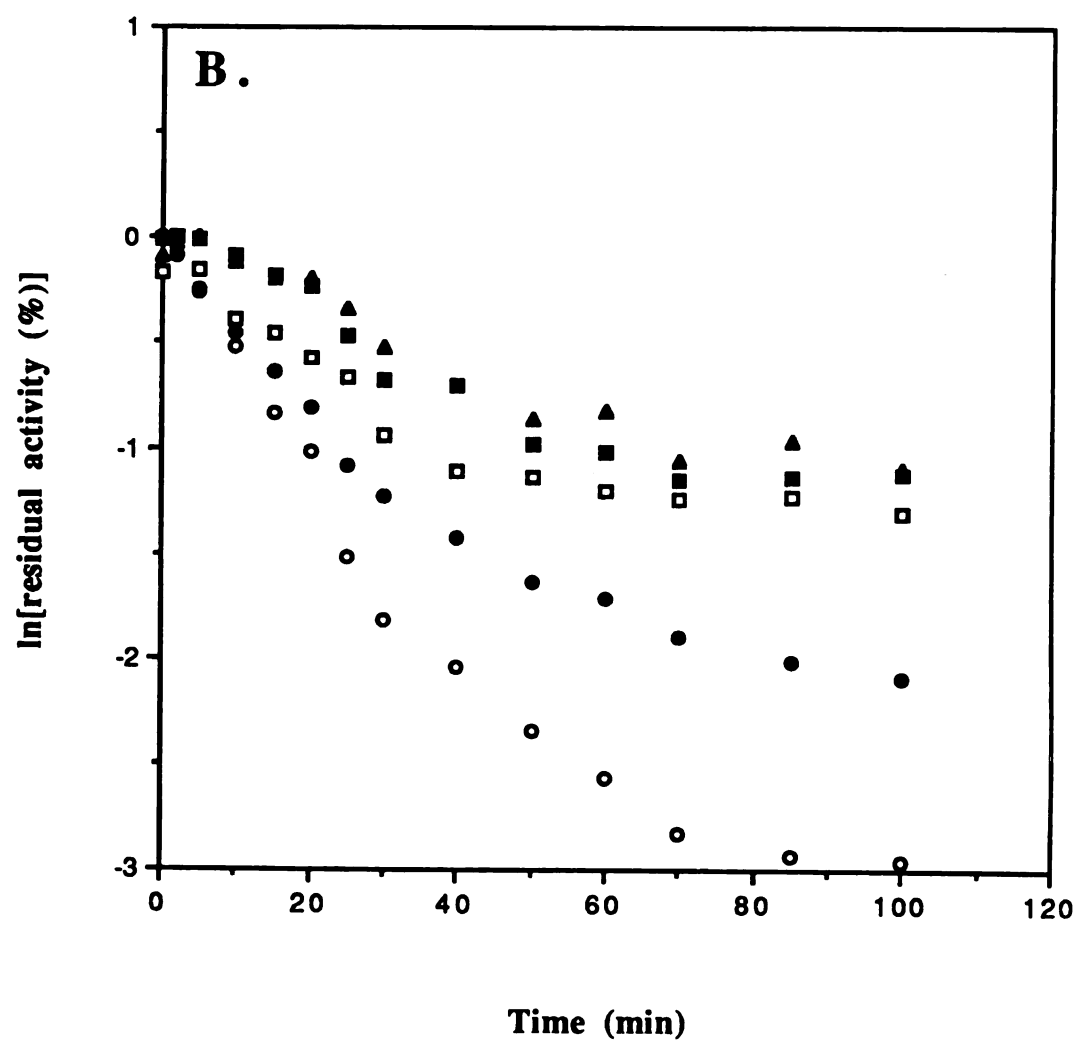
DISCUSSION

These preliminary data suggest that the 3 single Pro mutations tested do not significantly affect *T. ethanolicus* 2° ADH thermophilicity and thermostability. The wild type enzyme temperature activity profile is similar to that reported for purified enzyme (chapter 3). The temperature dependence of mutant and wild type enzyme activities were similarly except for the Pro149 to Thr mutant which may be inactivating at temperatures between 70°C and 90° C. The decreased slope for average percent activity versus temperature above 70°C compared to that below 70°C for this mutant is not consistent with wild type enzyme behavior and suggests loss of optimal active enzyme structure in at least part of the enzyme population. Furthermore, because the maximum Pro149 to Thr activity is less than that expected from extrapolating the lower temperature activity data to 90°C, dividing the lower temperature specific activities by the maximal activity (that at 90°C) to obtain the percent of maximum catalytic rate would artificially raise the percent activity values at lower temperatures. Recalculating the data using the temperature at 70°C as 100% activity indicates that all of the mutants and the wild type enzyme had similar temperature activity profiles below 70°C. Therefore replacing Pro 24 and Pro222 with the corresponding *C. beijerinckii* 2° ADH amino acids did not appear to significantly alter enzyme thermophilicity but replacing Pro149 may have reduced enzyme thermophilicity. However, activity determinations must be repeated using purified enzyme to confirm these findings.

The Pro residue deficient mutants all displayed similar thermostabilities which were greater than either wild type enzyme. The relationship between the logarithm of residual percent activity and incubation time however, was nonlinear for all 5 enzymes. Precipitate was seen in the thermoinactivated samples, consistent with data reported for the purified wild type enzyme. The failure of even the wild type enzyme data collected using cell extracts to fit a pseudo-first order rate equation suggests that the thermoinactivation slow step under these conditions differs from that for purified enzyme. The relatively rapid loss of 2° ADH activity in all 5 extracts compared to the purified enzyme further suggests that

purification stabilizes the 2° ADH against precipitation. While none of the Pro deficient mutants demonstrated a substantial loss of thermostability, the inconsistency of the cell extract data with that reported for the purified wild type enzyme makes analysis difficult. The entire set of nine added *T. ethanolicus* 2° ADH Pro residue mutations must be examined individually and together before drawing any conclusions but the preliminary evidence presented here using these 3 Pro mutants indicates that *T. ethanolicus* 2° ADH thermostability must be measured using purified enzyme.

REFERENCES

1. Vieille, C., Burdette, D. S. and Zeikus, J. G. [1996] Thermozymes, in *Biotechnology Annual Reviews, vol. 2*, (M. R. El Geweley, ed.) pp. 1-83. Elsevier Press, Amsterdam, Neth.

2. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. and Struhl, K. (1993) Current Protocols in Molecular Biology (Janssen, K., ed.), Current Protocols, NY

3. Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) Molecular cloning: A laboratory manual 2nd edition. (Nolan, C., ed.), Cold spring Harbor Press, NY

4. Sanger, F., Nicklen, S. and Coulson, A. R. (1977) Proc. Natl. Acad. Sci. USA **74**, 5463-5467

5. Burdette, D. S. and Zeikus, J. G. (1994) Biochem. J. **302**, 163-170