Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

Studies on Proteins Involved in Retinoid and Alcohol Metabolism

A thesis presented in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Biochemistry at Massey University

Treena Jayne Blythe Department of Biochemistry December 1997

Abstract

The primary biological role of the aldehyde dehydrogenase enzymes has long been a contentious issue. It was initially thought that the main function of these enzymes could be acetaldehyde metabolism; however, it seems unlikely that a large family of proteins evolved for this purpose. It has been suggested that an important function of aldehyde dehydrogenase enzymes may be in the metabolism of the vitamin A derivative, retinal. This thesis describes an investigation into the ability of human and sheep cytosolic aldehyde dehydrogenases to oxidise all-*trans* retinal, 9-*cis* retinal and CRBP-bound retinal under physiologically relevant conditions. A fluorescence-based assay following the production of NADH was employed, allowing the accurate measurement of low K_m data.

Firstly the ability of AlDH 1 to oxidise its putative biological ligands, free all-*trans* and 9-*cis* retinal, was demonstrated. It has been proposed that retinoids occur naturally as a 1:1 complex with the lipocalins cellular retinol binding protein (CRBP) and cellular retinoic acid binding protein (CRABP). If the sheep and human class 1 enzymes play a role in retinoid metabolism *in vivo*, it is likely that they will accept CRBP-bound retinal as a substrate. To investigate this possibility, recombinant CRBP was produced using an *E.coli* expression system. Using a spectrophotometric method, the purified recombinant CRBP was shown to bind all-*trans* but not 9-*cis* retinal, and using the same fluorescence-based assay as mentioned above, it was shown that both sheep and human AlDH 1 could accept CRBP-bound retinal as a substrate at physiologically relevant levels. *In vivo* studies into retinal oxidation were initiated using the retinoid-responsive human neuroblastoma cell-line SH-SY5Y. It was shown that AlDH 1 was expressed in this cell line by Western blotting, and that the cells were responsive to retinal in addition to retinoic acid, indicating that retinal was being converted to retinoic acid.

In addition, a novel, putative alcohol dehydrogenase was isolated, purified and partially characterised. The protein was purified using the techniques of subcellular fractionation by centrifugation, PEG precipitation, ion-exchange chromatography, preparative isoelectric focusing, hydrophobic interaction chromatography and gel purification. Elucidated characteristics of this protein include: subunit molecular weight 42-45 kDa, native molecular weight 42-45 kDa, isoelectric point 8.3-8.5, and activity with ethanol and other longer chain alcohols, but not with glucose, sorbitol or methanol. The protein

was blocked at the N-terminus, and cleavage and internal sequencing attempts yielded some sequence information. However, this information did not appear to match closely with any known protein sequence when submitted to a protein database, suggesting that the protein is novel.

From all available information, we propose that in sheep and humans, the enzyme responsible for retinal oxidation is the major cytosolic class 1 aldehyde dehydrogenase, as opposed to the situation in rats and mice, where specific retinal-oxidising aldehyde dehydrogenases exist and the major class 1 enzymes play a more important role in acetaldehyde metabolism.

Table of Contents

Table of Contents		i
List of Tables		vii
List of Figures		viii
List of Abbreviations		xi
Amino Acid Abbreviat	tions	xiii
Acknowledgements		xiv

Chapter 1: Introduction	1-1
1.1 The Retinoids	1-1
1.1.1 Functions of Retinoids	1-1
1.1.2 The Visual Cascade	1-3
1.1.3 The Role of Retinoic Acid in Development	
1.1.4 The Absorption, Storage, and Mobilisation of Die	tary
Retinoids	1-5
1.1.4.1 Plasma Retinoic Acid Circulation	
1.2 Intracellular Retinol Metabolism	1-9
1.2.1 The Conversion of Retinol to Retinal	1-11
1.2.1.1 The Alcohol Dehydrogenase Enzymes	1-12
1.2.1.2 The Short Chain Reductase/Dehydroger	lase
Enzymes	1-13
1.2.1.3 Other Retinol-Oxidising Enzymes	1-16
1.2.1.4 Summary	1-16
1.2.2 The Conversion of Retinal to Retinoic Acid	1-16
1.2.2.1 The Cytochrome P450 Enzymes	1-16
1.2.2.2 The Aldehyde Dehydrogenase Enzymes	1-17
1.2.2.3 The Retinal Dehydrogenase Enzymes	1-19
1.2.2.4 Other Retinal-Oxidising Enzymes	1-23
1.3 Cellular Retinoid-Binding Proteins	1-24
1.3.1 CRBP I	1-25
1.3.1.1 Ligands	1-26
1.3.1.2 Structure of CRBP I	
1.3.1.3 Ligand Binding	1-28

1.3.1.4 Function of CRBP I	1-31
1.3.1.5 Transfer of Ligand to Metabolic Enzymes	1-33
1.3.2 CRBP II	1-35
1.3.2.1 Ligands	1-35
1.3.2.2 Structure of CRBP II	1-35
1.3.2.3 Function of CRBP II	1-36
1.3.3 CRABP I and II	1-36
1.3.4 Plasma Retinol-Binding Protein (RBP)	1-37
1.3.5 Retinoid-Binding Proteins of the Visual System	1-38
1.3.6 Cis Isomers of Retinoids	1-39
1.4 Retinoic Acid Exerts it Effects Via Nuclear Receptors	1-40
1.4.1 Retinoic Acid Receptors (RARs)	1-43
1.4.2 Retinoid X Receptors (RXRs)	1-43
1.4.3 Retinoid Response Elements	1-44
1.4.4 Dimerisation of Retinoid Receptors	1-44
1.5 Vitamin A and Fetal Alcohol Syndrome (FAS)	1-45
1.6 The Aldehyde Dehydrogenases	1-47
1.7 The Aims of this Thesis	1-49

Chapter 2: Materials and Methods	
2.1 Materials	2-50
2.1.1 Retinoids	2-50
2.1.2 Molecular Biological Materials	2-50
2.1.3 Cell Culture Materials	2-51
2.1.4 CRBP Purification	2-51
2.1.5 Sheep Liver	2-51
2.1.6 Recombinant Human AlDH 1	2-51
2.1.7 Gels and Electrophoretic Transfer	2-51
2.1.8 Column Chromatography	2-52
2.1.9 General Protein Purification Materials	2-52
2.1.10 Protein Digestion	2-52
2.1.11 General Chemicals and Commonly Used Solutions	2-52
2.2 Methods	2-55
2.2.1 Maintenance of Retinoids	2-55
2.2.2 Spectrophotometry	2-55

2.2.3 Fluorimetry	2-56
2.2.3.1 Assays with CRBP	2-56
2.2.4 HPLC Based Assay - Extraction and Quantification	2-57
2.2.5 Recombinant DNA Techniques for CRBP Production	2-57
2.2.5.1 Transformation of E. coli	2-57
2.2.5.2 Large-Scale Preparation of Plasmid DNA	2-57
2.2.5.3 Digestion of Plasmid DNA with Cla I	2-57
2.2.5.4 TAE-Agarose Gel Electrophoresis of DNA	2-58
2.2.5.5 DNA Sequencing	2-58
2.2.6 Cytosolic Aldehyde Dehydrogenase Purification and	
Analysis	2-58
2.2.6.1 Purification of Sheep Liver AIDH 1	2-58
2.2.6.2 Protein and Activity Assays for Protein	
Purification	2-59
2.2.7 Isoelectric Focusing (IEF) Activity Assay	2-59
2.2.7.1 Transferral of Proteins Separated by IEF to	
Nitrocellulose	2-60
2.2.8 SDS-PAGE	2-60
2.2.9 Electroblotting	2-61
2.2.10 Electroelution of Proteins From Acrylamide Gels	2-62
2.2.11 Protein Cleavage	2-62
2.2.11.1 Cyanogen Bromide (Methionine-X)	2-63
2.2.11.2 Dilute Acid Cleavage (Aspartate-X)	2-63
2.2.11.3 Trypsin (Lysine-X and Arginine-X)	2-63
2.2.11.4 Thermolysin (X-Hydrophobic)	2-63
2.2.11.5 Leucine Aminopeptidase (N-Terminal Amino	
Acids)	2-63
2.2.11.6 Separation of Peptides Generated by Cleavage	2-64
2.2.12 Protein Sequencing	2-64
2.2.13 Amino Acid Analysis of Proteins	2-64
2.2.14 N-Terminal Deacylation	2-64
2.2.15 Western Blotting	2-65
2.2.16 Cell Culture.	2-66
2.2.16.1 Preparation of Cultured Cells for Western	
Blotting	2-66
2.2.16.2 Differentiation of Cultured Cells	2-66

2.2.16.3	DNA Staining2-6	67
----------	-----------------	----

Chapter 3: Purification of, and Kinetic Studies Using, Cytosolic Aldehyde	
Dehydrogenase	3-68
3.1 Introduction and Aims	3-68
3.2 Results and Discussion	3-69
3.2.1 Purification of sAlDH	3- 69
3.2.2 Method development for Kinetic Characterisation of	
Purified sAIDH	3-70
3.2.2.1 Development of an HPLC-Based Kinetic Assay	3-70
3.2.2.2 Reversed-Phase HPLC Retinoid Separation and	
Detection	3-73
3.2.2.3 Incubation of Enzyme Reaction Mixtures	3-73
3.2.2.4 The Extraction of Retinoids from Extraction	
Medium	3-75
3.2.2.5 Development of an Absorbance Spectroscopy	
Kinetic assay	3-80
3.2.2.6 Development of a Fluorescence Spectroscopy	
Kinetic assay	3-82
3.2.3 Kinetic Characterisation of sAlDH and hAlDH	3-84
3.2.3.1 NAD ⁺ -Dependent Oxidation of All-trans and	
9-cis Retinal by sAlDH and hAlDH	3-84
3.2.3.2 Inhibition of the Esterase Activity of sAlDH by	
Citral and Retinal	3-90
3.2.4 pH Optimum of NAD ⁺ -Dependent Retinal Oxidation	3-91

Chapter 4: Expression and Purification of CRBP and use of CRBP-Retinal

as a Sul	bstrate for AIDH 1	4-96
4.1 Introduct	ion and Aims	4-96
4.2 Results		
4.2.1	Expression and Purification of CRBP	4-96
4.2.2	The Binding of CRBP to Retinal	
4.2.3	CRBP-Retinal as a Substrate for Human and Sheep	
	AIDH 1	

iv

4.3 Discussion	
----------------	--

Chapter 5: A Comparison of the Major Aldehyde Dehydrogenases by

Multiple Sequence Alignment	5-112
5.1 Introduction	
5.2 Methods	5-112
5.3 Results	
5.3.1 Substrate Specificities of Class 1 and Class 2 AlDH	5-126
5.4 Conclusions	5-126

Chapter 6: In Vivo Studies on Retinal Dehydrogenation	6-128
6.1 Introduction and Aims	6-128
6.2 Results	6-128
6.2.1 Initial Characterisation of SH-SY5Y Cells	6-128
6.2.2 Effects of Retinal and Retinoic Acid on Cell Morphology	6-130
6.2.2.1 Introduction and Aims	6-130
6.2.2.2 Methods	6-130
6.2.2.3 Results and Discussion	6-132
6.3 Conclusion	6-147

Chapter 7: Purification of a Novel Protein with Alcohol Dehydrogenase

Activity	,	
Introduc	tion and Aims	7-148
Methods	and Results	7-148
7.2.1	Initial Protein Purification	
7.2.2	Ion-Exchange Chromatography	7-153
7.2.3	Dye-Ligand Chromatography	7-158
7.2.4	Affinity Chromatography	7-159
7.2.5	Preparative Isoelectric Focusing	
7.2.6	FPLC Hydrophobic Interaction Chromatography	
7.2.7	HPLC Size Exclusion Chromatography	7-170
7.2.8	Hydroxyapatite Chromatography	7-172
7.2.9	Gel Purification	7-179
	Activity Introduce Methods 7.2.1 7.2.2 7.2.3 7.2.4 7.2.5 7.2.6 7.2.7 7.2.8 7.2.8 7.2.9	Activity Introduction and Aims. Methods and Results 7.2.1 Initial Protein Purification 7.2.2 Ion-Exchange Chromatography. 7.2.3 Dye-Ligand Chromatography. 7.2.4 Affinity Chromatography. 7.2.5 Preparative Isoelectric Focusing 7.2.6 FPLC Hydrophobic Interaction Chromatography . 7.2.7 HPLC Size Exclusion Chromatography. 7.2.8 Hydroxyapatite Chromatography. 7.2.9 Gel Purification

7.2.10 Chemical and Enzymatic Cleavage	
7.2.11 Activity Assays	
7.2.12 Tissue Distribution	
7.3 Conclusions	

Chapter 8: Discussion and Future Work	
8.1 Discussion and Conclusions	8-190
8.2 Future Work	

APPENDIX	 7

List of Tables

Table 1.1: A Summary of the Characteristics of Four Major Retinol-	
Oxidising Enzymes	1-15
Table 1.2: A Summary of the Major Characteristics of Six Aldehyde	
Dehydrogenases Involved in Retinal Oxidation	1-22
Table 1.3: Genes Containing Retinoid Response Elements	1-42
Table 2.1: Company Details	2-50
Table 2.2: Light Absorbances of Selected Retinoids	2-55
Table 3.1: Typical Purification of Sheep Liver Aldehyde Dehydrogenase	3-71
Table 3.2: Relative Extraction of Retinol Acetate and Retinoic Acid	
Table 3.3: Retention Times for HPLC Separation of Retinoids	
Table 3.4: A Summary of Kinetic Parameters Determined for hAlDH1	
and sAlDH1	3- 89
Table 3.5: The Inhibition of the Hydrolysis of p -Nitrophenyl Pivalate by	
Retinal and Citral	
Table 4.1: N-terminal Sequencing of Expressed and Purified CRBP	4-101
Table 4.2: A Summary of Kinetic Parameters of sAlDH and hAlDH	
for CRBP-Retinal and Free All-trans Retinal	4-108
Table 5.1: Identification of Proteins Used in Multiple Sequence Alignment	5-113
Table 5.2: Secondary Structure Elements of Sheep AlDH1	5-121
Table 6.1: Quantification of SH-SY5Y Cell Morphology	6-138
Table 6.2: Quantification of SH-SY5Y Cell Morphology with Added	
Effectors	6-141
Table 7.1: Amino Acid Analysis of Unknown Protein	7-181
Table 7.2: Sequence Obtained From Tryptic Digest of Unknown Protein	
Table 7.3: Summary of Enzyme Activity	7-186

List of Figures

Figure 1.1: Structural Formula of β -Carotene and Naturally Occurring	
Retinoids	1-2
Figure 1.2: The Absorption and Storage of Dietary Retinoids	1-7
Figure 1.3: An Overview of Intracellular Retinoid Homeostasis	1-10
Figure 1.4: Dendrogram Showing Clustering Relationships Between	
Aldehyde Dehydrogenase Enzymes	1-21
Figure 1.5: Ribbon Diagram of CRBP I	1-27
Figure 1.6: Schematic Diagram Showing Structural Basis for Observed Binding	
Constants of CRBP and CRBP II for Retinol and Retinal	1-30
Figure 1.7: Overview of Mechanism of Retinoid Action	1-41
Figure 1.8: Ribbon Diagram of Sheep Aldehyde Dehydrogenase 1	1-48
Figure 2.1: Transferral of Proteins Separated by IEF to Nitrocellulose	2-60
Figure 3.1: SDS-PAGE and IEF Gel of Pure sAlDH1	3-72
Figure 3.2: HPLC Trace of Authentic Retinoids	3-74
Figure 3.3: Standard Curve for Retinoic Acid	3-78
Figure 3.4: Standard Curve for Retinol Acetate	3-79
Figure 3.5: Single Reaction Progress Curve of sAlDH-Catalysed Retinal	
Oxidation	3-81
Figure 3.6: The Conversion of Retinal to Retinoic Acid	3-84
Figure 3.7: V vs [S] Plot and Lineweaver-Burk Plot of hAlDH1 Catalysed	
All-trans Retinal Oxidation	3-85
Figure 3.8: V vs [S] Plot and Lineweaver-Burk Plot of hAlDH1 Catalysed	
9-cis Retinal Oxidation	3-86
Figure 3.9: V vs [S] Plot and Lineweaver-Burk Plot of sAlDH1 Catalysed	
All-trans Retinal Oxidation	3-87
Figure 3.10: V vs [S] Plot and Lineweaver-Burk Plot of sAlDH1 Catalysed	
9-cis Retinal Oxidation	3-88
Figure 3.11: Structural Formulae for All-trans Citral and All-trans Retinal	3-92
Figure 3.12: Inverse Plot of the Inhibition of sAlDH1 Esterase Activity by Citral	3-93
Figure 3.13: Inverse Plot of the Inhibition of sAlDH1 Esterase Activity by	
All-trans Retinal	3-94

Figure 3.14: pH Dependence of NAD+-Dependent Oxidation of All-trans	
Retinal by sAIDH1	95
Figure 4.1: Plasmid Map of pT7-7-CRBP	97
Figure 4.2: SDS-PAGE of Soluble <i>E. coli</i> Protein	99
Figure 4.3: SDS-PAGE of CRBP Purification	00
Figure 4.4: Absorbance Spectra of All- <i>trans</i> Retinal and CRBP-All- <i>trans</i> Retinal 4-10	03
Figure 4.5: Absorbance Spectra of 9-cis Retinal and CRBP-9-cis Retinal 4-1	04
Figure 4.6: V vs [S] Plot and Lineweaver-Burk Plot of hAIDH1 Catalysed	•••
CR BP-Retinal Oxidation 4-10	05
Figure 4.7. V vs [S] Plot and Lineweaver-Burk Plot of sAIDH1 Catalysed	00
CRBP-Retinal Oxidation 4-10	06
Figure 4.8. V vs [S] Plot and Lineweaver-Burk Plot of sAIDH1 Catalysed	00
CR BP-Retinal Oxidation 4-10	07
	07
Figure 5.1: Multiple Sequence Alignment of 32 Aldehyde Dehydrogenase	
Sequences	20
Figure 5.2: Dendrogram Showing Clustering Relationships Used by Pileup	
to Create Multiple Sequence Alignment	23
Figure 5.3: PlotSimilarity Output of Aldehyde Dehydrogenase Sequence	
Alignment	25
Figure 6.1: Western Blot of SH-SY5Y Cells	29
Figure 6.2: Outline of Treatment of SH-SY5Y Cells with Retinoids	31
Figure 6.3: SH-SY5Y Cells from Control Plate	33
Figure 6.4: SH-SY5Y Cells from Control + DMF Plate	34
Figure 6.5: SH-SY5Y Cells from All-trans Retinal Treated Plate	35
Figure 6.6: SH-SY5Y Cells from All-trans Retinoic Acid Treated Plate	36
Figure 6.7: Fluorescence Microscopy of SH-SY5Y Cells	37
Figure 6.8: Quantification of SH-SY5Y Cell Morphology	39
Figure 6.9: Outline of Treatment of SH-SY5Y Cells with Retinoids and	
Additional Effectors	40
Figure 6.10: Quantification of SH-SY5Y Cell Morphology with Added	
Effectors	42

ix

Figure 6.11: Two Possible Mechanisms for the Manifestation of Fetal Alcohol	
Syndrome	6-145
Figure 7.1: IEF Gel of Crude Liver Extract	7-149
Figure 7.1C: IEF Gel of Crude Liver Extract	7-150
Figure 7.2: IEF Gel of Crude Liver Extract Using Different Substrates	7-152
Figure 7.3: Initial Purification of Novel Protein	7-154
Figure 7.4: Ion-exchange Resins	7-155
Figure 7.5: Preparative Isoelectric Focusing	7-163
Figure 7.6: IEF Gel of Preparative Isoelectric Focusing Fractions	7-164
Figure 7.7: IEF Gel of FPLC Hydrophobic Interaction Chromatography	
Fractions	7-166
Figure 7.8: IEF Gel of FPLC Hydrophobic Interaction Chromatography	
Fractions (Protein Stain)	7-167
Figure 7.9: SDS-PAGE of Purified Protein	7-168
Figure 7.10: Western Blot of Pure Novel Protein and AlDH1	7-169
Figure 7.11: Developed Purification Scheme for Putative Alcohol	
Dehydrogenase	7-171
Figure 7.12: HPLC Size Exclusion Chromatography of Partially Purified	
Protein (Superose-12 Column)	7-173
Figure 7.13: SDS-PAGE of HPLC Size Exclusion Chromatography Fractions	
(Superose-12)	7-174
Figure 7.14: HPLC Size Exclusion Chromatography of Partially Purified	
Protein (Superdex-75 Column)	7-175
Figure 7.15: SDS-PAGE of HPLC Size Exclusion Chromatography Fractions	
(Superdex-75)	7-176
Figure 7.16: Elution Profile of Size Exclusion Standards (Superdex-75 Column)	7-177
Figure 7.17: Calibration Curve Constructed From Superdex-75 Elution Profile	7-178
Figure 7.18: Schematic Diagram of Gel Purification	7-180
Figure 7.19: SDS-PAGE of Gel Purification of Novel Protein	7-182
Figure 7.20: IEF Gel of Crude Extracts From Different Tissues	7-187

x

List of Abbreviations

A ₂₈₀	. Absorbance at 280 nm		
AlDH	Aldehyde dehydrogenase		
ADH	Alcohol dehydrogenase		
AMP	Ampicillin		
1,8-ANS	1-Anilinonaphthalene 8-sulfonic acid		
ARAT	. AcylCoA-retinol acyltransferase		
Bis-Tris	.Bis[2-hydroxyethyl]iminotris[hydroxymethyl]methane		
BSA	. Bovine serum albumin		
CAPS	. 3-[Cyclohexylamino]-1-propane sulfonic acid		
cDNA	Complementary DNA		
CNBr	. Cyanogen Bromide		
CRABP	Cellular retinoic acid-binding protein		
CRBP	Cellular retinol-binding protein		
C-terminal	. Carboxy-terminal		
DEAE	Diethylaminoethyl		
DMF	. Dimethylformamide		
DMSO	Dimethyl sulphoxide		
DNA	. Deoxyribonucleic acid		
DNA-BD	DNA-binding domain		
DR	Direct repeat		
DTT	Dithiothreitol		
EDTA	Ethylenediamine tetra-acetic acid		
FAE	Fetal alcohol effects		
FAS	. Fetal alcohol syndrome		
FPLC	. Fast protein liquid chromatography		
HEPES	N-2-hydroxyethylpiperazine-N'-ethanesulphonic acid		
HPLC	High performance liquid chromatography		
IEF	Isoelectric focusing		
iLBP	Intracellular lipid binding protein family		
IPTG	Isopropyl-y-D-thiogalactopyranoside		
kDa	. kiloDalton		
K _m	Substrate concentration at half maximum reaction rate		
LB	Luria Broth		

LBD	Ligand-binding domain		
LRAT	Lecithin-retinol acyltransferase		
λmax	Wavelength of maximum absorbance		
mRNA	Messenger RNA		
NAD ⁺	Nicotinamide adenine dinucleotide		
NADH	. Nicotinamide adenine dinucleotide (reduced form)		
N-terminal	. Amino-terminal		
PBS	Phosphate buffered saline		
PEG	Polyethylene glycol		
pI	Isoelectric point		
PMSF	Phenylmethylsulphonyl fluoride		
PPAR	Peroxisome proliferator-activated receptor		
PVDF	Polyvinyldifluoride		
RA	Retinoic acid		
RALDH (or RalDH)	Retinal-specific aldehyde dehydrogenase		
RAR	Retinoic acid receptor		
RARE	Retinoic acid response element		
RBP	Retinol-binding protein		
RoDH	Retinol-specific alcohol dehydrogenase		
RXR	Retinoid X receptor		
RXRE	Retinoid X response element		
[S]	Substrate concentration		
SDR	Short-chain dehydrogenase-reductase family		
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel		
	electrophoresis		
TR	Thyroid hormone receptor		
Tris	Tris(hydroxymethyl)aminoethane		
UV-vis	Ultraviolet-visible		
VDR	Vitamin D receptor		
V	Rate of reaction		
V _{max}	Maximum rate of reaction		
v/v	volume/volume		
w/v	weight/volume		
w/w	weight/weight		

Amino Acid	Abbreviation	One-letter Abbreviation	
Alanine	Ala	Α	
Asparagine or aspartic acid	Asx	В	
Cysteine	Cys	С	
Aspartic acid	Asp	D	
Glutamic acid	Glu	E	
Phenylalanine	Phe	F	
Glycine	Gly	G	
Histidine	His	Н	
Isoleucine	Ile	Ι	
Lysine	Lys	K	
Leucine	Leu	L	
Methionine	Met	M	
Asparagine	Asn	N	
Proline	Pro	Р	
Glutamine	Gln	Q	
Arginine	Arg	R	
Serine	Ser	S	
Threonine	Thr	Т	
Valine	Val	V	
Tryptophan	Тгр	W	
Tyrosine	Tyr	Y	
Glutamine or glutamic acid	Glx	Z	

Amino Acid Abbreviations

Acknowledgements

I would really like to thank my supervisor Assoc. Prof. (!!!) Kathy Kitson for her support throughout the last 4 years. I have really appreciated the many times she was able to help me financially and practically, and for giving me freedom to pursue many different avenues. I would also like to thank my other supervisors Assoc. Prof. Mike Hardman, Dr. Mark Grimes, and Assoc. Prof. Trevor Kitson for their advice and guidance. For travel and scholarship funding I am grateful to the Neurological Foundation, The Royal Society and the Massy Doctoral Committee. I would like to thank PNMRF and the Neurological Foundation for running costs, and finally Lottery Health and Lottery Science for equipment grants. I have really appreciated the help of many other members of the former Department of Biochemistry for their help. Thanks to Dr. Shaun Lott for help with molecular biology and sequence analysis, Dr. Gill Norris for help with protein work, Dr. Gretchen McCaffery for help with cell studies, Dr. Stan Moore for help with structural work, Jo Mudford for protein sequencing, Debbie Frumau for amino acid analysis, and Dr. Mark Patchett and Carole Flyger for heaps of stuff. I would also like to thank Rochelle Ramsay, Dick Poll, Prof. Ted Baker and Heather Baker, and Dr. Dave Harding for allowing me to use their equipment.

Personal Thanks:

Without a doubt I would never have made it through the last 4 years without some of the great people I have met. Most of all to Shaun, who has been so helpful at work, and without who the most successful part of my project would not have been possible. I will also be always grateful to Shaun for helping me through 1997 when I was ill for so long, grumpy and had no money. Thanks also to my friends Rose, Isobel, Rich, Andrew, Stan, Jakki, Wendy, Rick, Kath and Megan, Jeroen, Michelle, Rochelle, Catherine, Neil, Ross, Max and Ruth, Rish, Mish, Bec, Matt, Scott, Brendon - choice mates, maaaate. I would also like to say go you good things Hannah, Katharine, Sian, Clare, Caroline, Serena, Joanne, Joanna, Michelle, Rachel and Zoë, I hope you all had as much fun as I did! Thanks to my surrogate family Rex and Fran (thanks for the critiques Rex), Warren and Barb, Karen, Chris and Mike. Thanks also to Katrina and Leigh, the best friends a girl could have, Jo, Titi, Paul, Matai, Matt and the boys. Finally and mostly, I would like to thank my excellent family especially Mum and Dad for financial support and love and being such fun and cool parents. Also to Todd, Nan, Pop, Nana, Grandad (I hope you see this), Jenny, Phil, Tony and Kris for being the best family ever.

1. Chapter One: Introduction

1.1 The Retinoids

Retinoids are naturally occurring and synthetic polyisoprenoid lipids which are structurally related to retinol (vitamin A) (Figure 1.1). All structural analogues of retinol are termed retinoids whether or not they exhibit any retinoid-like activity. Biologically active retinoids are pleiotropic compounds with a wide range of functions in vertebrate embryonic growth and development (Gudas, 1994; Hoffman & Eichele, 1994; Maden, 1994) and maintenance of adult epithelial tissues (Gudas et al., 1994; Rosenthal et al., 1994; Wolbach & Howe, 1925). The most well characterised process with which vitamin A is involved is the visual system. The first recognition of the effect of vitamin A on the visual cascade was recorded by the ancient Greeks. Hippocrates noted a condition termed nyktalopia (night blindness) in c.400 BC, to which the noted prescribed cure was to eat large servings of raw beef liver - an excellent source of vitamin A (Wolf, 1996). "Minimal qualitative factors" in the rat diet necessary for growth and survival were identified in c.1881 (Hopkins, 1906; Lunin, 1881) and termed 'fat soluble factor A' (McCollum & Davis, 1913; Osborne & Mendel, 1913). An attempt to isolate this factor vielded an active oil (Osborne & Mendel, 1914), the yellow colour of which was proposed to be merely associated with the dietary factor, and which was then converted to the colourless active compound (Steenbock & Gross, 1920). It was not until the 1930's that the yellow compound was identified as β -carotene, the precursor of colourless vitamin A (Moore, 1930). During this time it was also discovered that the conversion of β -carotene to retinol occurred within the body, the structure of β -carotene was elucidated and recommended daily allowances identified (Karrer et al., 1931).

1.1.1 Functions of Retinoids

Vitamin A plays vital roles in both embryonic formation and various adult processes. A deficiency in vitamin A may lead to blindness, disease and in extreme cases, death. In most retinoid dependent processes, retinoic acid is the biologically active compound, with the exception of the visual system and the reproduction requirement. Retinoic acid is a potent modulator of development that affects multiple organs and structures (Morriss-Kay, 1993; Richman, 1992; Summerbell & Maden, 1990). During the formation of vertebrate embryos, retinoic acid programs positional information in limb generation and is crucial to nervous system, cardiovascular, urogenital and cranial facial



Figure 1.1: Structural Formula of β-Carotene and Naturally Occurring Retinoids

development (Hoffman & Eichele, 1994, and refs. therein). In mature vertebrates, retinoic acid is essential for processes such as normal hematopoiesis, bone remodelling and maintenance of epithelial integrity (Goss & McBurney, 1992; Wolf, 1984). The specific molecular processes with which retinoic acid is involved are: 1) inhibition of cell proliferation, 2) induction of cell differentiation, 3) induction of programmed cell death, 4) immunomodulation, 5) inhibition of angiogenesis, and 6) influence on oncogene expression (Bollag, 1996). The manner in which retinoic acid exerts its effects is through binding nuclear receptors, and affecting the transcription of certain target genes (De Luca, 1991) (see section 1.4). In relation to these functions, vitamin A therapy has now been successfully used to treat both dermatological conditions (acne, psoriasis, and other keratinising disorders) (Fisher & Voorhees, 1996 and refs. therein), and certain cancers (acute promyelocytic leukemia, and combination therapy on particular cancerous and pre-cancerous diseases) (De Luca, 1993; Gudas, 1994; Lotan, 1995; McBurney *et al.*, 1993 and refs. therein).

1.1.2 The Visual Cascade

The visual system is one of the best known signal transduction processes, and is one in which retinoids play an integral part. However, in this event, 11-*cis* retinal and not retinoic acid is the biologically active retinoid. The irreversibility of the oxidation of retinal to retinoic acid means that retinoic acid will not support this function; however, both retinol and retinal will. The ability to detect the intensity of the visual stimulus (light), and its wavelength (colour), originates in the absorption of a photon by 11-*cis* retinal or a derivative. 11-*cis* retinal is complexed with a protein called opsin to form rhodopsin, and in this form 11-*cis* retinal (and in some cases, derivatives of 11-*cis* retinal) is the chromophore of all known visual pigments. The derivatives of retinal mentioned may be derived from all-*trans* retinol, 3,4-didehydroretinol or 3-hydroxyretinol (Saari *et al.*, 1994). In addition, there are a number of specific retinoid binding proteins found in the eye, the functions of which are not certain as yet (see section 1.3.5).

1.1.3 The Role of Retinoic Acid in Development

A large volume of evidence exists which implicates retinoic acid as playing a principle role in the specification of spatial information in the developing embryo (see Hoffman & Eichele, 1994 and refs. therein for this section). Retinoic acid is involved in growth, differentiation and cellular rearrangement, as well as in pattern generation in embryo formation. The generation of patterns is a mechanism which is distinct from the other

processes — for example all digits in the hand consist of the same cell types, which have undergone identical growth and differentiation; however, the thumb is obviously different from the index finger. One of the giant steps in modern developmental biology has been the identification of particular regulatory molecules that control the formation of body pattern. It is accepted that retinoids influence limb development, anteroposterior (head to tail or head to foot) patterning, and regenerating amphibian limbs, but their relative importance in these roles has yet to be confirmed.

On the basis of considerable research, a mechanism for the action of retinoic acid (RA) on the limb bud has been formulated. In the early limb bud, a gradient of RA is established. The position of a cell in this gradient will determine the level of RA to which it is exposed. The detection of the local RA concentration by the retinoic acid receptors in the cell regulates the expression of downstream genes, which in turn define cell identities in the limb bud mesenchyme. This may provide the basis for digit patterning along the anteroposterior limb axis. It is likely that other cues are involved in this process such as tissue interactions or other growth factors. An indication of the importance of RA, however, is that local application of RA can completely recapitulate limb morphogenesis (Hoffman & Eichele, 1994).

Exposure of embryos to high doses of RA causes changes in the overall body plan resulting in malformations. The observation that RA can alter the expression of homeobox genes in cultured cells (Colberg-Poley et al., 1985) provided some clues as to the mechanism by which RA is involved in embryonic patterning. Homeobox genes are pivotal genes involved in axis formation, and it seems that the effects of RA are at least partly due to alterations in expression patterns of homeobox genes. Current evidence for the involvement of RA in the organisation of the primary body axis of vertebrates includes: 1) exogenously applied RA induces specific transformations in the identity of mesoderm and ectoderm derived structures, e.g. vertebrae, 2) these transformations occur at the time when affected structures are specified, and changes are preceded by changes in Hox gene patterning, 3) early embryos contain endogenous retinoids and express retinoid receptors and binding proteins, 4) retinoic acid induces Hox genes in vivo and in vitro, and 5) 'organiser regions' involved in axial pattern formation e.g. Hensen's node, are high points of retinoic acid synthesis and contain elevated levels of endogenous retinoids.

1.1.4 The Absorption, Storage, and Mobilisation of Dietary Retinoids

Dietary sources of naturally occurring retinoids are β -carotene from plant tissue and retinyl esters found in animal tissue. Retinyl esters are converted into retinol in the intestinal lumen, and absorbed along with β -carotene by intestinal enterocytes. While it has been shown that β -carotene is directly internalised by enterocytes (El-Gorab *et al.*, 1975; Hollander & Ruble, 1978), the hydrolysis of retinyl esters to retinol in the intestinal lumen suggested that a specific transporter for retinol may exist to transport it across the intestinal wall (Dew & Ong, 1994). The identification of large amounts of cellular retinol binding protein (II) (CRBP II) (see section 1.3.2) in the intestinal mucosa led to the proposal of a mechanism for the uptake of dietary retinoids involving CRBP II. CRBP II binds to all-*trans* retinol and all-*trans* retinal with an approximately equal strength, and thus is thought to play a central role in the processing of dietary retinoids in the mucosal cells (Kakkad & Ong, 1988; Levin *et al.*, 1987; Ong *et al.*, 1991; Ong & Page, 1987; Quick & Ong, 1990). Simple diffusion of retinol across the enterocyte bilayer, driven by the presence of large amounts of CRBP II present in the cell, is an alternative uptake scenario.

Once in the enterocytes, β -carotene is cleaved by either 1) a dioxygenase to form two molecules of retinaldehyde (retinal), or 2) excentric oxidative cleavage to form retinal via the formation of β -apocarotenals (Blaner & Olsen, 1994 and refs. therein). The retinal formed from cleavage of β -carotene and bound to CRBP II is then able to be reduced to retinol by a retinal reductase also found in the intestinal mucosa (Kakkad & Ong, 1988). The retinol produced from β -carotene, and the retinol absorbed as such (from retinyl ester hydrolysis in the intestinal lumen) is then esterified with palmitic or stearic acid for storage purposes (Ong et al., 1987). Two enzymes are primarily responsible for The most important of these is termed lecithin-retinol acyltransferase esterification. (LRAT) (Quick & Ong, 1990), and this has recently been shown to preferentially esterify retinol bound to CRBP II over free retinol. Due to the abundance of CRBP II, under normal conditions most of the retinol ingested would be bound to CRBP II, which means LRAT would be expected to play the major role in the esterification process. The other enzyme which has been shown to esterify free retinol is acyl coenzyme A:retinol acyltransferase (ARAT) (Helgerud et al., 1982; Helgerud et al., 1983; Yost et al., 1988). It is believed the esterification of retinol by ARAT would only occur under conditions with excess retinol, and that this would provide a temporary 'store' of retinol.

The retinyl esters are then packaged into nascent chylomicrons which are subsequently released into the lymphatic system. Here the chylomicrons undergo lipolysis, giving rise to chylomicron remnants. Chylomicron remnants containing retinyl esters are largely taken up by hepatic cells, although a small amount reaches other tissues, most notably bone marrow (Hussain *et al.*, 1989).

There are two types of hepatic cell involved with retinoid storage - the stellate cells and the parenchymal cells. Of total body vitamin A, 50-80 % is stored in mammalian stellate cells, making them the major storage cells of retinyl esters. While the stellate cells are the main storage site, parenchymal cells are responsible for the uptake of the chylomicron remnants. The parenchymal cells are also the major site of synthesis and storage of plasma retinol binding protein (RBP) (see section 1.3.4), which binds to retinol and facilitates its transport throughout the body. Upon internalisation of retinoid by hepatic parenchymal cells, some retinol may be delivered directly to stored RBP and put into circulation to meet the body's immediate retinoid requirements. From the major storage site in the liver, retinol is also mobilised and transported to other retinoid requiring tissues, where it may be utilised immediately, or stored for future use. The mobilisation of stored retinoid requires the hydrolysis of retinyl esters. Two enzymes have been identified which are involved in the hydrolysis of retinyl esters - 1) a bile-salt dependent hydrolase, and 2) a bile-salt independent hydrolase (Blaner & Olsen, 1994 and refs. therein). RBP is the sole protein involved in the transport of mobilised retinol. Retinol binds to RBP in a 1:1 ratio, and RBP-retinol travels complexed with the protein transthyretin which minimises losses by glomerular filtration.

It is still unclear how retinol enters target cells. The existence of a specific cell surface receptor involved in receptor mediated endocytosis of RBP has been suggested, as has the existence of a channel through which retinol may enter the cell. Recognition between circulating RBP-retinol and the specific receptor is postulated to initiate transfer of the retinoid across the membrane and into the cytosol of the target cell. It is also possible that such a receptor may recognise a cellular retinol binding protein on the intracellular side of the membrane. Support for these models includes the observation of specific binding of RBP (with and without transthyretin) to cellular membranes (Bavik *et al.*, 1991; Bok & Heller, 1976; Heller, 1975; Rask & Peterson, 1976; Shingleton *et al.*, 1989; Sivaprasadarao & Findlay, 1988), internalisation of RBP (Blomhoff *et al.*, 1985; Gjoen *et al.*, 1987; Senoo *et al.*, 1990; Vieira & Schneider, 1993). An alternative hypothesis



Figure 1.2: The Absorption and Storage of Dietary Retinoids

β-Carotene and retinol are absorbed across the intestinal brush border membrane, where both are converted into retinyl esters. In the retinyl ester form, retinol is transported via a chylomicron complex in the blood to the liver, where it is taken up and stored until required. Adapted from Ross, 1993a, p323

Abbreviations: β -C - β -carotene; REH - retinyl ester hydrolase; RDH - retinol dehydrogenase; LRAT - lecithin-retinol:acyl-coA transferase; RE - retinyl esters; CM - chylomicron; ADH - alcohol dehydrogenase; RA - retinoic acid; CRBP - cellular retinol-binding protein

1-7

does not require a specific receptor for retinol uptake. It has been demonstrated that retinol is able to 'flip' from one side of the lipid bilayer to the other. Retinol delivered to the membrane by RBP may rapidly equilibrate between the inner and outer leaflets of the bilayer. The amount of apo-CRBP inside the cell would then in part determine the amount of retinol entering the cell by binding to retinol, and effectively 'pulling the equilibrium across' (see Blaner & Olsen, 1994 and refs. therein for review). It is quite possible that both of these mechanisms exist, and are cell and/or tissue dependent. Once inside the target cell, retinol can either be metabolised, or esterified and stored. Thus retinoid-sensitive cells may have their own store of retinyl esters, which are maintained by the major retinol store in the liver (stellate cells). For an overview of retinoid absorption, storage and initial metabolism see Figure 1.2.

1.1.4.1 Plasma Retinoic Acid Circulation

Although retinol is the major circulating retinoid in human plasma, low levels of retinoic acid (4-14 nM), corresponding to 0.2-0.7 % of total plasma retinoid levels are also present (De Leenheer *et al.*, 1982; Eckhoff & Nau, 1990). This can be derived from dietary sources or from endogenous tissue metabolism of retinol. In contrast to retinol, retinoic acid circulates bound to serum albumin, and does not bind to plasma retinol binding protein (Lehman *et al.*, 1972; Smith *et al.*, 1973). At physiological pH, it has been demonstrated that uncharged retinoic acid can rapidly and spontaneously cross cell membranes, and in this manner gain entry to cells (Noy, 1992a; Noy, 1992b).

In addition to all-*trans* retinoic acid, the related compounds 13-*cis* retinoic acid and 13-*cis*-4-oxoretinoic acid have been shown to be present at significant levels in human plasma. The circulating concentrations of these retinoids in normal human plasma have been estimated to be - 4.4 nM (all-*trans*-), 5.4 nM (13-*cis*-), and 11.7 nM (13-*cis*-4-oxo-) (Eckhoff & Nau, 1990; Tang & Russell, 1990). It has also been demonstrated that circulating levels of all-*trans* and 13-*cis* retinoic acid rise transiently in response to oral retinyl palmitate doses, returning to basal levels within 20 days. Plasma 13-*cis*-4-oxoretinoic acid levels also rise, but the rise is sustained (Eckhoff *et al.*, 1991; Tang & Russell, 1990). This work indicates that there is a lot less known about circulating retinoic acid in comparison to the circulation and uptake of retinol. However, there is significant circulating retinoic acid, and the contribution this makes to overall retinoid signalling mechanisms, though this is likely to be minor relative to retinol, should be taken into account.

1.2 Intracellular Retinol Metabolism

In all retinoid-sensitive cells except ocular cells, the carboxylic acid retinoic acid is now thought to be the most potently active physiological retinoid (De Luca, 1991; Napoli et al., 1991; Ross, 1993b; Williams & Napoli, 1985). Although the body has a basic requirement for vitamin A, excess retinoic acid is teratogenic (Morriss-Kay & Sokolova, 1996; Wilson et al., 1953), and toxic postnatally. Retinoic acid excess or deficiency at crucial stages or specific areas in embryonic development result in a range of malformations and toxicity (see section 1.5). Therefore, due to the requirement for exact spatio-temporal retinoic acid concentration, it is obvious that the production and availability of retinoic acid to responsive cells is likely to be tightly controlled. The understanding of the regulation of retinoic acid biosynthesis from its precursor retinol is of vital importance when studying retinoid-dependent processes. Retinoic acid supports all vitamin A dependent processes with the exception of the reproductive and vision requirements (Dowling & Wald, 1960; Eskild & Hansson, 1994; Thompson et al., 1964) (see section 1.1.1). The machinery to produce retinoic acid from internalised retinol is present in every retinoic acid sensitive cell type (Napoli, 1986, Napoli & Race, 1987, Williams & Napoli, 1985) (Figure 1.3).

It is generally accepted that retinoic acid is formed from retinol by a two-step oxidative process with retinal as an intermediate. The oxidation of retinol to retinal and then to retinoic acid in this way is analogous to the oxidation of ethanol to acetic acid via acetaldehyde. The amount of retinoic acid available in the cell is of vital importance. At present the mechanisms of control which are involved in the biosynthesis of retinoic acid are unclear. There are a number of enzymes which have been shown to oxidise free retinol or free retinal in vitro, but the physiological relevance of these enzymes in relation to one another is not known. The presence of the intracellular retinoid binding proteins, (the CRBPs and the cellular retinoic acid binding proteins (CRABPs)), lends a further possible level of control and specificity to this process (see section 1.3). The proposal that holo-CRBP (i.e. CRBP bound to its ligand all-trans retinol or retinal) is the actual physiological substrate for retinoid oxidation, may answer the question of biological relevance of the retinoid-oxidising enzymes. In this scenario, only those enzymes which recognise holo-CRBP would play a role in retinoic acid biosynthesis. Other enzymes which may be able to metabolise free retinoids would normally play only a minor role, but perhaps a greater role in situations of hypervitaminosis A. There has also been an alternative suggestion that the CRBPs are not involved in retinoic acid production at all



Figure 1.3: An Overview of Intracellular Retinoid Homeostasis

Retinoids can be taken up into target cells as retinol, β-carotene, or retinoic acid. In the cytoplasm, conversion between free retinol, retinal and retinoic acid or CR(A)BP-bound retinol, retinal and retinoic acid can occur. Excess retinol may be stored, and retinoic acid may exert an effect in the nucleus, or be metabolised for excretion. (Adapted from Duester, 1996, p12223)

Abbreviations: SDR - short-chain reductase/dehydrogenase family; ADH - alcohol dehydrogenase; AlDH - aldehyde dehydrogenase; RAR - retinoic acid receptor; RXR - retinoid X receptor; Cyt-P450 - cytochrome P450 family.

(Duester, 1996), but are only involved in directing the storage of retinol as retinyl esters (see sections 1.3.1.4 and 1.3.2.3). In this hypothesis there would be many enzymes contributing to retinoid reduction and oxidation intracellularly, and no advantage would be conferred on those enzymes recognising holo-CRBP. There would also be less control of retinoic acid synthesis if this was the case, and the distribution and amount of CRABPs in the cell could be critical, especially in retinoid-sensitive cells (see section 1.3.3). The dissection and understanding of this mechanism is proving to be a contentious and complex issue, which is further complicated when looking at the mechanism by which retinoic acid exerts its effects (see section 1.4).

1.2.1 The Conversion of Retinol to Retinal

The conversion of retinol to retinal is the rate limiting step in retinoic acid biosynthesis (Blaner & Olsen, 1994; Chen *et al.*, 1995; Kim *et al.*, 1992; Napoli, 1986). This reaction is reversible, which means the reduction of retinal to retinol may also occur. A number of enzymes from both the alcohol dehydrogenase (ADH) family and the short chain dehydrogenase reductase (SDR) family have been shown to be involved in this conversion. Both of these families consist of several evolutionarily related forms, some of which use retinoids as substrates. Each family has an ancient origin, and it is likely that ancient forms involved in the metabolism of other compounds have diverged to give rise to more recent forms, some of which may use retinol and retinal (Hempel *et al.*, 1993; Jornvall *et al.*, 1995; Nelson *et al.*, 1993; Persson *et al.*, 1995). However, the relative importance of these enzymes, and therefore the biological role they play, has proved to be a controversial issue.

The ADH family contains 7 evolutionarily related classes, each class exhibiting unique tissue distribution and properties (Duester *et al.*, 1995; Jornvall *et al.*, 1995; Kedishvili *et al.*, 1997a; Kedishvili *et al.*, 1997b). Human ADH classes I, II, and IV have been shown to oxidise all-*trans*-, 9-*cis*-, and 13-*cis*- retinol, class III will not accept retinoids as substrates (Boleda *et al.*, 1993; Yang *et al.*, 1994), and classes V and VI have not been examined. The SDR family is evolutionarily related to the ADH family, both families sharing a similar coenzyme-binding domain. The main differences between the families are - 1) the ADHs have larger subunit molecular weights (~40 kDa compared to 25-35 kDa for the SDR - hence 'short-chain'), and 2) ADHs are zinc-dependent and SDR's have no metal requirement (Duester, 1996 for review). Recently, three new members of the SDR family which are able to utilise retinoids as substrates have been identified and cloned (Chai *et al.*, 1995a; Chai *et al.*, 1996; Chai *et al.*, 1995b). These

enzymes are located in the microsomal fractions of various tissues, and have been named retinol dehydrogenase (RoDH) I, II, and III. A cytosolic retinol oxidising activity distinct from the known cytosolic ADHs has also been observed (Ottonello *et al.*, 1993). These three retinol-oxidising (and retinal-reducing) groups of enzymes have been identified as potentially playing an important physiological role in the production of retinoic acid. However, their relative roles *in vivo* have yet to be fully elucidated.

1.2.1.1 The Alcohol Dehydrogenase Enzymes

A number of efforts have been made to address the controversial issue of biological relevance among retinol-oxidising enzymes. The case for the ADHs - in particular classes I and IV - playing a relatively major role in cellular retinol oxidation is put forward by Duester and colleagues (see Duester, 1996 for review). In this hypothesis, the role of the CRBP proteins is confined to facilitating storage and mobilisation of retinol, with no involvement in metabolism. Observations supporting this scenario include a comparison of catalytic efficiencies of ADH IV and RoDH I (Duester, 1996). Pure ADH IV showed a catalytic efficiency (k_{cat}/K_m) with cofactor NAD⁺ that was around 100 times greater (more efficient) for retinol oxidation than that for retinal reduction (Duester, 1996; Yang *et al.*, 1994). For RoDH I (data from Boerman & Napoli, 1995a; Boerman & Napoli, 1995b), the comparable calculations indicated a five-fold increased efficiency for retinal reduction over retinol oxidation than RoDH I (Duester, 1996).

In addition, RoDH I only uses NADP(H) as a cofactor (Boerman & Napoli, 1995a; Napoli *et al.*, 1992; Posch *et al.*, 1991), with ADH IV preferring NAD(H) (Boleda *et al.*, 1993; Connor & Smit, 1987; Kedishvili *et al.*, 1995; Yang *et al.*, 1994). The cytosolic NAD⁺/NADH ratio in aerobically growing cells is ~1000, compared to ~0.01 for NADP⁺/NADPH (Veech *et al.*, 1969). This means the reduction of retinal by the NADP(H)-preferring RoDH enzyme is more likely to occur, and the reverse is true for ADH IV (Duester, 1996). In addition, RoDH only accepts all-*trans* retinoids (Boerman & Napoli, 1995b), while ADH I and IV both utilise *cis* isomers (Boleda *et al.*, 1993; Yang *et al.*, 1994). The role of *cis*-isomers of retinoids (see section 1.3.6) indicates it is likely that the ability to interconvert 9-*cis*, 13-*cis* and 3,4-didehydro- retinoids between their alcohol, aldehyde and acid forms, or simply to more polar metabolites for excretion, is necessary.

Inhibitor studies also indicate an important role for ADH isoenzymes. A classic ADH inhibitor, 4-methylpyrazole, inhibited retinoic acid synthesis from retinol *in vivo* in mouse embryos (Collins *et al.*, 1992). Citral (another ADH inhibitor), eliminated retinol oxidation *in vitro* in rat embryo homogenates (Chen *et al.*, 1995), and ethanol inhibited ADH IV-catalysed retinol oxidation *in vitro* (Julia *et al.*, 1986; Kedishvili *et al.*, 1997a; Yin *et al.*, 1997). None of these compounds are inhibitors of microsomal RoDH activity (Chai *et al.*, 1995b). Final evidence for ADH I and/or IV playing an important role in retinol metabolism comes from tissue distribution studies. The spatio-temporal expression of class I and IV ADH mRNA has been correlated to the appearance of retinoic acid in mouse embryos (Ang *et al.*, 1996a; Deltour *et al.*, 1996), and in adult vertebrates, ADH I and IV are detected in retinoic acid sensitive epithelia (Ang *et al.*, 1996b).

1.2.1.2 The Short Chain Reductase/Dehydrogenase Enzymes

There is also an abundance of evidence in support of the microsomal RoDH activity being the biologically significant activity in retinol-retinal interconversion. A major difference in these two scenarios is the involvement of the CRBPs (see section 1.3.1.4). In this hypothesis CRBP I plays an important role in directing the metabolism of retinol by protecting retinol/retinal from the aqueous environment in which it is unstable and poorly soluble, protecting retinol/retinal from random enzymatic or non-enzymatic conversions, and conferring specificity by restricting access to those enzymes able to recognise CRBP. It is well documented that microsomal RoDHs recognise holo-CRBP as a substrate in addition to free retinoids (Boerman & Napoli, 1995b; Napoli et al., 1992; Posch et al., 1991), but the ability of ADH IV to utilise holo-CRBP has not been examined. If the CRBP proteins do play a central role in retinoid homeostasis, then ADH must be able to accept holo-CRBP as a substrate to be considered relevant. It has been estimated that the major source of retinol in CRBP-expressing cells occurs as retinol bound to CRBP, and the estimated concentrations are 5 µM (holo-CRBP) and 2 μM (apo-CRBP) (Bashor et al., 1973; Boerman & Napoli, 1996; Harrison et al., 1987; Napoli, 1993; Napoli et al., 1995; Napoli et al., 1992; Ong, 1982; Ottonello et al., 1993; Yost et al., 1988), and a free retinol concentration of 50-500 fold less (assuming only 10 % excess of CRBP over retinol) (Napoli, 1993).

Several experiments have indicated the requirement for both microsomal and cytosolic subcellular compartments in retinoic acid production from retinol (Kim *et al.*, 1992; Napoli, 1993; Napoli *et al.*, 1992; Posch *et al.*, 1991; Posch *et al.*, 1992). From these

results, a mechanism for retinoic acid production involving first retinol oxidation by microsomal enzyme(s), and second, retinal oxidation by cytosolic components was proposed. Evidence to the contrary includes the ability of cytosol alone to produce retinoic acid from retinol in ADH negative deermice (Posch & Napoli, 1992), and the identification of a cytosolic RoDH activity in calf liver (Ottonello et al., 1993). The microsomal RoDHs have a distinctive tissue distribution. In the adult rat, RNase protection assays identified RoDH isozymes in the liver, and to a lesser extent in the lung, kidney, testis and brain, while RoDH I and II expression was also observed in day 12.5 and 13.5 rat embryos (Chai et al., 1995a; Chai et al., 1996; Chai et al., 1995b; Napoli et al., 1995). This expression coincided with the expression of CRBP (Napoli et al., 1995). An investigation into the relative roles of microsomes and cytosol in retinol oxidation revealed that microsomes provided the major share of retinal for retinoic acid synthesis (Boerman & Napoli, 1996). Using estimated physiological concentrations of holo- (5 μ M) and apo- (2 μ M) CRBP, it was shown that 93 % of total retinal produced came from the microsomal fraction. In addition, the rate of microsomal retinal production was ~45 times faster than cytosolic retinal production. Inhibition of the cytosolic retinol dehydrogenase by ethanol in a subcellular fraction containing both cytosol and microsomes did not decrease the amount of retinoic acid produced - which shows that either the cytosol contributes little overall, or the drop in cytosolic contribution can be compensated for by the microsomal fraction. It was also seen that the microsomal activity (unlike the cytosolic) was insensitive to increased amounts of apo-CRBP. This is further evidence that holo-CRBP is a substrate (Boerman & Napoli, 1996). A chemical cross-linking experiment showed that retinol-CRBP cross-linked to a microsomal protein of approximately the same size as RoDH (Boerman & Napoli, 1995a). In addition to these experiments, other reasons why it is unlikely for ADH I and/or IV to be the main retinol oxidising system have been put forward (Napoli, 1996b). These are: 1) the relatively high K_m of ADH's for retinol (20-200 μ M), 2) lower K_{ms} for other substrates, 3) limited tissue distribution of ADH's, 4) the requirement for control in retinoic acid production, and 5) the existence of holo-CRBP as the major in vivo substrate.

It is also interesting that the group proposing the importance of the microsomal RoDHs previously published a paper citing evidence in opposition to this theory (Napoli & Race, 1987). Retinoic acid synthesis from retinol was observed in rat liver and kidney cytosol and was NAD⁺ but not NADP⁺ dependent. Ethanol (100 mM) inhibited retinoic acid production from retinol in the liver (54 %) and in the kidney (30 %), and

4-methylpyrazole was an inhibitor of retinoic acid synthesis, especially in liver. In addition, it was found that the CR(A)BPs were not obligatory players in this scenario. These observations not only contradict subsequent experiments performed by this group, they also go some way to supporting the case for the involvement of the ADHs, especially ADH IV, in physiological retinoic acid production.

Table 1.1: A Summary of the Characteristics of Four Major Retinol Oxidising Enzymes

MAJOR ENZYMES INVOLVED IN RETINOL METABOLISM				
	Class I ADH	Class IV ADH	RoDH (M)	RoDH(C)
$K_{\rm m}$ (free retinol)	45-290 μM ¹	$31 \mu\text{M}^1$	$4-10 \ \mu M^{2.4}$	$7.1 \mu M (NAD)^3$
κ_m (bound retinol) Catalytic efficiency (retinol) k_{cat}/K_m	ND ND	$\frac{ND}{1900 \text{ min}^{-1} \text{ mM}^{-1}}$ (human) ⁵	$1.6 - 1.9 \ \mu M^{-1}$ 17 min ⁻¹ mM ⁻¹	3.5 μM (NAD) ⁵ ND
Catalytic efficiency (retinal)	ND	1200 min ⁻¹ mM ⁻¹ (human) ⁵	93 min ⁻¹ mM ^{-1 5}	ND
Tissue Distribution	Various mouse embryonic tissues, adult epithelia, liver, kidney, intestine, lung, stomach ^{6.7.8,9}	Various mouse embryonic tissues, adult epithelia,stomach 6.7.9	Various rat embryonic tissues, liver, testes, lung, kidney brain ⁸⁻¹	Calf liver, testis, kidney, retinal pigment epithelia ³
Inhibitors	Ethanol, 4- methylpyrazole ⁵	Ethanol, 4- methylpyrazole ⁵	Ethanol ×, ketoconazole \times^2	ND
Activity with holo- CRBP	ND	ND	YES ²	YES ³
Preferred cofactor	NAD ⁺⁵	NAD ⁺⁵	NADP ^{+2,4}	NADP ⁺ /NAD ⁺³
Activity with <i>cis</i> isomers	YES⁵	YES ⁵	NO ⁷	ND

Key: \times = does not inhibit, ND = not determined, M = microsomal, C = cytosolic

References: ¹ Yang *et al.*, 1994; ² Posch *et al.*, 1991; ³ Ottonello *et al.*, 1993; ⁴ Napoli *et al.*, 1992; ⁵ Duester, 1996; ⁶ Ang *et al.*, 1996a; ⁷ Boerman & Napoli, 1995b, ⁸ Jornvall *et al.*, 1995, ⁹ Duester *et al.*, 1995

1.2.1.3 Other Retinol Oxidising Enzymes

An additional cytosolic retinol-oxidising enzyme distinct from the ADHs has been identified (Ottonello *et al.*, 1993). This enzyme is able to use both free retinol and CRBP-retinol as low K_m substrates (see Table 1.1), is inhibited by apo-CRBP but not by other retinoid-binding proteins, and uses both NAD⁺ and NADP⁺ as cofactors.

1.2.1.4 Summary

The issue of relative importance of these retinol-oxidising enzymes remains unresolved. The high K_m values of the ADHs for retinol would seem to discount them from playing a physiological role as the available free retinol has been estimated at <100 nM (Napoli, 1993). However, overestimations may have occurred when calculating ADH K_m values, due to the presence of detergents and other agents to solubilise the relatively insoluble retinol. The issue of structural isomers seems to be important - ADHs but not RoDHs can accept *cis* isomers, although the recent identification of a 9-*cis*-specific alcohol dehydrogenase which doesn't oxidise all-*trans* retinol (Mertz *et al.*, 1997) may support the proposal that the RoDHs are the major enzymes. A final important issue which remains to be resolved is the role that CRBP plays in retinoid metabolism.

1.2.2 The Conversion of Retinal to Retinoic Acid

The second step in the biosynthesis of retinoic acid from its precursor retinol is the irreversible oxidation of retinal. Two main families of enzymes are involved in this reaction — the aldehyde dehydrogenases, and cytochrome P450.

1.2.2.1 The Cytochrome P450 Enzymes

Of the multiple forms of cytochrome P450 identified to date, only two are able to efficiently oxidise retinal to retinoic acid — P450 1A1 and P450 1A2 (Raner *et al.*, 1996; Roberts *et al.*, 1992), while all other known forms have negligible retinal-oxidising activity (Raner *et al.*, 1996). The active forms can utilise all-*trans*, 9-*cis*, and 13-*cis* retinal; however, P450 1A1 is by far the more active, and the expression of the P450 1A1 gene has been demonstrated in mouse embryos as early as day 7 of gestation (Kimura *et al.*, 1987). A more likely role *in vivo* for the cytochromes P450, however, lies in the metabolism of retinoids to more polar metabolites, as many P450 isoenzymes have the ability to derivatise the carbon at position 4 in the ionone ring (Duester, 1996). The ability of the P450 enzymes to convert retinoic acid to such metabolites as retinoyl β -glucuronide, 5,6-epoxyretinoic acid, 4-hydroxyretinoic acid, 4-oxoretinoic acid and

1-16

3,4-didehydroretinoic acid has been well established (Eckhoff *et al.*, 1991; Heyman *et al.*, 1992; Levin *et al.*, 1992b; McCormick *et al.*, 1978; Skare *et al.*, 1982; Tang & Russell, 1990; Thaller & Eichele, 1990). Some of these metabolites may be active in retinoic acid function (e.g. 3,4-didehydroretinoic acid, 4-oxoretinoic acid), whereas many are likely to be catabolic products destined for excretion. A likely role for the cytochrome P450 enzymes is in the direct metabolism of retinoic acid. Retinoic acid is bound to CRABP (Fiorella & Napoli, 1991), though it still is not certain what the precise role of CRABP is in this process (see section 1.3.3).

1.2.2.2 The Aldehyde Dehydrogenase Enzymes

In general, members of the aldehyde dehydrogenase (AlDH) family exhibit a broad substrate specificity, and can oxidise various biogenic and xenobiotic aliphatic and aromatic aldehydes (Ambroziak & Pietruszko, 1991; Goedde & Agarwal, 1990; Hsu & Chang, 1991; Sladek *et al.*, 1989; Yoshida *et al.*, 1991). Only a limited number of identified AlDH enzymes are able to oxidise retinal — human and sheep AlDH 1 (Klyosov *et al.*, 1996; Yoshida *et al.*, 1992; this thesis), mouse AHD-2 (Dockham *et al.*, 1992; Lee *et al.*, 1991), and a number of novel enzymes isolated from rat and mouse tissues. These include rat kidney RALDH-1 (Bhat *et al.*, 1995; Labrecque *et al.*, 1993; Labrecque *et al.*, 1997; Labrecque *et al.*, 1995), mouse RALDH-2 (Zhao *et al.*, 1996), rat liver RalDH1 (el Akawi & Napoli, 1994; Posch *et al.*, 1992), and rat testis RalDH2 (Wang *et al.*, 1996).

A number of observations link class 1 AlDH to retinoic acid production in vertebrate animals. High endogenous levels of retinoic acid are found in the retina of the embryonic mouse, an area in which AlDH 1 activity is also localised (McCaffery *et al.*, 1992). In zebrafish embryos, AlDH 1 activity co-localises with retinoic acid levels (Marsh-Armstrong *et al.*, 1994), and AlDH 1 was shown to be expressed in the basal forebrain, early in the development of mouse embryos (McCaffery & Drager, 1994). Early work showed that enzyme-catalysed retinal oxidation is restricted to the cytosolic fraction of mouse liver (Lee *et al.*, 1991). This activity is 90 % NAD⁺ dependent, and 90 % of the NAD⁺ dependent activity was identified as AHD-2 — the mouse equivalent of human AlDH 1. The remaining portion of the NAD⁺ dependent activity was shown to be due to AHD-7 and xanthine oxidase (dehydrogenase form). The K_m values of these three enzymes for retinal were determined to be 0.7 μ M (AHD-2), 0.6 μ M (AHD-7), and 0.9 μ M (xanthine oxidase). The non-NAD⁺ dependent activity is likely to be due to aldehyde oxidase, as it is largely inhibited by the aldehyde oxidase inhibitor pyridoxal. However, the major retinal-oxidising activity of mouse liver was identified to be a class 1 cytosolic dehydrogenase. These experiments agreed with similar work using rat liver (Leo *et al.*, 1989), rat kidney (Napoli & Race, 1987), and human keratinocytes (Siegenthaler, 1990). In an experiment using human liver cytosol, in contrast to the majority of studies which use murine tissues and subcellular fractions, all aldehyde dehydrogenase activities were chromatographically isolated, using various aldehydes as substrates for activity (Dockham *et al.*, 1992). From the seven distinct aldehyde dehydrogenase activities separated, only one could oxidise retinal - AIDH 1.

Although the studies mentioned above all identify AlDH 1 as playing a major role in retinal oxidation, a number of papers, while identifying AlDH 1 activity as important, found other enzyme activities to play a significant role (Bhat et al., 1988b; Chen et al., 1995; Kishore & Boutwell, 1980; Leo et al., 1989; Napoli & Race, 1988). It is likely that these contrasting observations stem from differences in experimental approach; species specific differences, and tissue and cell specific differences. For example, the studies of (Chen et al., 1995) focused on rat conceptal homogenates. Homogenates from days 10.5, 11.5 and 12.5 of gestation each exhibited enzyme activity for retinoic acid production. The production of retinoic acid was reduced by 40-45 %, both by omission of NAD⁺, and replacement of NAD⁺ by NADH or NADPH. In addition, NAD⁺ was a more efficient cofactor than NADP⁺. Retinoic acid biosynthesis from retinol was inhibited by the alcohol and aldehyde dehydrogenase inhibitor citral, but not by high concentrations of azide, 4-methylpyrazole (ADH 1 inhibitor), or metapyrone (cytochrome P450 inhibitor). The biosynthesis of retinoic acid from retinal was inhibited by citral, but not metapyrone. From these observations, a two step pathway for retinoic acid synthesis in rat embryos was proposed, whereby the conversion of retinol to retinal is catalysed by an NAD⁺/NADP⁺ dependent retinol dehydrogenase and the conversion of retinal to retinoic acid is catalysed by an NAD⁺/NADP⁺ dependent retinal dehydrogenase and a retinal oxidase.

There is a strong case for cytosolic aldehyde dehydrogenase being the major enzyme involved in retinal oxidation in the human. It was previously thought that the principle role of AlDH 1 was the oxidation of acetaldehyde derived from the oxidation of ingested ethanol (Greenfield & Pietruszko, 1977). However, the reported K_m of AlDH 1 for acetaldehyde varies from 22 to 483 μ M (Rashkovetsky *et al.*, 1994), while typical physiological concentrations of acetaldehyde are only 0.4 - 2.5 μ M (Klyosov *et al.*,
1996). Reported K_m values for acetaldehyde oxidation by another major AlDH found in liver, mitochondrial AlDH 2, range from <0.1 μ M to 9 μ M, making it highly likely that this isozyme has acetaldehyde oxidation as its main function (Rashkovetsky *et al.*, 1994). The ability of AlDH 1 to oxidise retinal has been well documented (Dockham *et al.*, 1992; Klyosov *et al.*, 1996; Rashkovetsky *et al.*, 1994; Yoshida *et al.*, 1992; Yoshida *et al.*, 1993). In addition, AlDH 1 is the only identified liver enzyme with retinaloxidising ability in humans, as AlDH 2 is unable to accept retinal as a substrate (Yoshida *et al.*, 1992). The reported K_m values of AlDH 1 for all-*trans* retinal range from 0.06 μ M (Yoshida *et al.*, 1993) to 1.1 μ M (Klyosov *et al.*, 1996), and the catalytic efficiency for retinal is 100-500 times greater than for acetaldehyde (Klyosov *et al.*, 1996; Yoshida *et al.*, 1993). This means that, with an estimated biological concentration of retinal of ~0.1 μ M, retinal is an excellent candidate for the primary physiological substrate of AlDH 1.

1.2.2.3 The Retinal Dehydrogenase Enzymes

Recently, four retinal-specific dehydrogenases from rat and mouse tissues have been identified — RalDH1 (el Akawi & Napoli, 1994; Posch *et al.*, 1992), RalDH2 (Wang *et al.*, 1996), RALDH-1 (Bhat *et al.*, 1995; Labrecque *et al.*, 1993; Labrecque *et al.*, 1995), and RALDH-2 (Zhao *et al.*, 1996). The isolation of RalDH1 from rat liver cytosol was the first reported retinal-specific aldehyde dehydrogenase (Posch *et al.*, 1992). RalDH1 is tetrameric (214 kDa native MW), with a monomer size of 55 kDa estimated by SDS-PAGE, and a basic isoelectric point (8.3). NAD⁺ dependent RalDH1 accepts as substrates free all-*trans* retinal and 9-*cis* retinal, all-*trans* retinal bound to CRBP, and retinal generated from retinol by microsomes, but does not utilise 13-*cis* retinal (el Akawi & Napoli, 1994; Posch *et al.*, 1992). Allosteric kinetics were observed for all substrates, and substrate inhibition was seen at retinal concentrations greater then 6 μ M (Posch *et al.*, 1992). A chromatographically similar enzyme activity to RalDH1 was observed in both rat kidney cytosol, and rat testis cytosol (Posch *et al.*, 1992).

A further novel retinal-oxidising enzyme (RALDH-1) was isolated from the cytosol of rat kidney (Labrecque *et al.*, 1993). RALDH-1 has a native molecular weight of 140 kDa as determined by size exclusion chromatography, a subunit molecular weight of 53 kDa by SDS-PAGE, and a basic isoelectric point of 8.5 (Labrecque *et al.*, 1993). Aldehyde dehydrogenases are typically dimeric (class 3) or tetrameric (classes 1 and 2); however, the quaternary structure of RALDH-1 cannot be estimated from these data. The method used to estimate the native molecular weight has its limitations; in particular,

the effective Stokes radius of reference and test proteins has a great bearing on how they run through gel filtration columns. In this case no mention of any salt in the gel filtration buffers was made, the absence of which may increase random hydrophobic and ionic interactions by some proteins with the column matrix. It is likely therefore that this enzyme is in fact, tetrameric. RALDH-1 utilises acetaldehyde, free all-trans retinal, 9-cis retinal, and 11-cis retinal, but not 13-cis retinal as substrates (Labrecque et al., 1993; Labrecque et al., 1995). Michaelis-Menten kinetics were observed for all substrates, and substrate inhibition was seen at retinal concentrations greater than 20 µM. The activity of this enzyme with CRBP-retinal has not been tested, the authors stating that CRBP does not bind retinal. Although this was originally thought to be the case (MacDonald & Ong, 1987; Ong & Chytil, 1975), it is now accepted that CRBP does bind all-trans retinal (see section 1.3.1.1). The gene encoding RALDH-1 has been cloned from a rat kidney cDNA library (Bhat et al., 1995). The full length isolated cDNA contains a 2315 bp open reading frame, encoding a deduced protein of 501 amino acids. Northern blotting experiments identified this gene to be expressed largely in kidney, followed by lungs, testis, intestine, stomach and trachea, with only a small amount in liver. The absence of retinal-oxidising activity and constitutive AlDH activity from rat liver had been noted previously (Bhat et al., 1995; Bhat et al., 1988b; Lindahl & Evces, 1984; Tottmar et al., 1973).

Attempts to clone RalDH1 from a rat liver library were initially confounded by the presence of more abundant AlDH enzymes such as the rat phenobarbital-inducible enzyme (Wang *et al.*, 1996). Recently, the cDNA encoding this enzyme has been cloned using oligonucleotides designed to peptides purified during the initial purification (Penzes *et al.*, 1997). The expressed and purified protein showed similar characteristics to the protein originally purified from rat liver (see Table 1.2). While screening a rat testis library for RalDH1, a previously unknown aldehyde dehydrogenase was cloned, which was named RalDH2 (Wang *et al.*, 1996). Expressed and purified protein recognised free all-*trans* retinal, and retinal bound to CRBP as substrates, but not acetaldehyde, and the preferred cofactor was NAD⁺. Michaelis-Menten kinetics were observed for all substrates. Northern blotting experiments showed that RalDH2 was expressed solely in testis, but RNase protection assays identified small amounts (<10 % of that in testis) in lung, brain, heart, liver and kidney.

A fourth retinal-oxidising enzyme has been isolated (Zhao et al., 1996), which is thought to correspond to one of two previously identified enzyme activities (McCaffery &



Figure 1.4: Dendrogram Showing Clustering Relationships Between Aldehyde Dehydrogenase Enzymes

32 aldehyde dehydrogenase sequences were aligned using the GCG program 'pileup'. The above diagram is a pictorial representation of the clustering relationships used in the construction of the sequence alignment (see Chapter 5). The diagram represents similarity between primary amino acid sequences. The retinal oxidising enzymes are - the retinal specific enzymes RALDH-1 and -2, RalDH1 and 2, and the class 1 enzymes H1, S1, R1, and M1 (human, sheep, rat and mouse). For references and full descriptions see Chapter 5

1-21

MAJOR ENZYMES INVOLVED IN RETINAL OXIDATION						
	RALDH-1 1,2,3	RALDH-2	RalDH1 ^{5,6}	RalDH2 ⁷	AHD-2 ^{8,13}	AlDH 1 9,10,11,12,14
Native MW	140 kDa	ND	214 kDa	ND		240 kDa
Subunit MW	53 kDa	ND	55 kDa	~50 kDa		54 kDa
pI	8.5	ND	8.3	ND	7.96-8.25	5.2
Organism	Rat	Mouse	Rat	Rat	Mouse	Human, Sheep, Cow, Horse etc.
K _m AT- retinal	8-10 μM	ND	0.76 ± 0.35 μM, 1.6 μM	0.7 μM	0.7 μM	0.05 μM - 1.1 μM
<i>cis</i> -isomer activity	9-cis, 11-cis	None	9-cis	ND	ND	9-cis, 13- cis
Cofactor	NAD^+	NAD^+	NAD^+	NAD⁺/ NADP⁺	NAD^+	NAD^+
Kinetic Mechanism	Michaelis- Menten	ND	Allosteric	Michaelis- Menten	Michaelis- Menten	Michaelis- Menten
CRBP-retinal activity	ND	ND	0.13 μM	0.2 μM	ND	0.2 - 1 μM
Tissue Distribution (underlining denotes main location)	<u>Kidnev</u> , lung, testis, stomach, intestine, trachea, liver	Day 9 mouse embryo, adult mouse testes, pia	Liver, kidney, testes	<u>Testis</u> , lung, brain, heart, liver, kidney	<u>Liver</u> , lung, testis	Liver, brain, lung kidney, heart, spleen, stomach

Table 1.2: A Summary of the Characteristics of Six Aldehyde Dehydrogenases Involved in Retinal Oxidation

References:

¹Labrecque *et al.*, 1993; ²Labrecque *et al.*, 1995; ³Bhat *et al.*, 1995; ⁴Zhao *et al.*, 1996; ⁵Posch *et al.*, 1992; ⁶ el Akawi & Napoli, 1994; ⁷Wang *et al.*, 1996; ⁸Lee *et al.*, 1991; ⁹Yoshida *et al.*, 1991; ¹⁰Rashkovetsky *et al.*, 1994; ¹¹Klyosov *et al.*, 1996, ¹²This Thesis, 1997; ¹³Manthey *et al.*, 1990; ¹⁴Harada *et al.*, 1980.

Drager, 1993; McCaffery & Drager, 1995; McCaffery *et al.*, 1992). The gene encoding this enzyme has been cloned from P19 teratocarcinoma cells, and has been named RALDH-2 (Zhao *et al.*, 1996). No substrates for RALDH-2 apart from retinal have been identified using standard zymography. It has been shown that RALDH-2 is upregulated by the presence of its product (retinoic acid) in P19 cells, which is postulated to occur at the transcriptional level. To study retinoic acid synthesis, this

group uses a retinoic acid responsive reporter cell line, which consists of F9 teratocarcinoma cells transfected with a retinoic acid response element driving β -galactosidase expression. Traditional protein purification and enzyme characterisation have not been carried out.

The four retinal specific enzymes have been isolated from murine and rat tissues. Due to the fact that the groups involved do not refer to each other's work, it is difficult to understand the overall situation in these animals, and indeed to relate the situation to that which occurs in humans. Each enzyme is reported to be a novel enzyme, in its own class, but the relation to each other is not known. In an attempt to rationalise the data, the sequences of all known retinal-oxidising aldehyde dehydrogenase enzymes were aligned, and a dendrogram constructed to show the relationship of these enzymes to each other and to other members of the aldehyde dehydrogenase family (see Figures 1.4, 5.2, Chapter 5). It was found that RalDH1 and RALDH-1 were identical except in two positions which were thought likely to be attributable to sequencing errors, as the substitutions occurred at highly conserved positions. In addition RalDH2 and RALDH-2 were also shown to be identical bar one substitution. This substitution was a conservative one, an isoleucine in one enzyme and a valine in the other. It also seems likely that these two enzymes are identical.

As human and sheep cytosolic AlDH 1 are not considered significant in acetaldehyde metabolism (Klyosov, 1996; Klyosov *et al.*, 1996; Rashkovetsky *et al.*, 1994), it is possible that these isoenzymes may have evolved for the purpose of retinal oxidation in humans and sheep. In the rat, however, the main cytosolic enzyme has a lower K_m for acetaldehyde making it more likely to be involved in acetaldehyde oxidation *in vivo* (Klyosov *et al.*, 1996), hence the need for specific retinal-oxidising enzymes.

1.2.2.4 Other Retinal-Oxidising Enzymes

Two other classes of enzyme potentially play a role in retinal oxidation. Both belong to a broad family of molybdenum-containing proteins. They are the aldehyde oxidases and the xanthine oxidases (Calzi *et al.*, 1995). The aldehyde oxidases are flavoproteins which oxidise various xenobiotic substrates under aerobic conditions without a requirement for NAD⁺ or NADP⁺ (Calzi *et al.*, 1995). Typically they are homodimeric, with a subunit MW of ~150 kDa. An enzyme activity which catalysed the formation of retinoic acid from retinal was observed in rat kidney cytosol in 1988 (Bhat *et al.*, 1988a). On the basis of cofactor studies and inhibition studies using the aldehyde oxidase inhibitor 4'-(9-acridinylamino)methanesulfon-anisidide, this activity was identified as an aldehyde oxidase. Due to the relatively high K_m for all-*trans* retinal (125 μ M), it is unlikely that this enzyme plays a role *in vivo*. Further studies on rabbit liver cytosol showed that an NAD⁺ independent activity was responsible for ~40 % of retinoic acid production in this subcellular fraction (Huang & Ichikawa, 1994). This enzyme activity was again shown to belong to an aldehyde oxidase for which, however, the calculated K_m value was much lower (8 μ M). It is still unlikely with an estimated cellular retinal concentration of ~100 nM that this enzyme plays a large role. Confirmation of the identity and characteristics of the rabbit hepatocyte enzyme came with its purification in 1994 (Tsujita *et al.*, 1994). The purified protein is a homodimer of two 135 kDa subunits, with a K_m for all-*trans* retinal of 8 μ M, for 9-*cis* retinal of 13 μ M, and for 13-*cis* retinal of 0.9 mM.

The aldehyde oxidases may play a role in retinal metabolism, but their high K_m values seem to preclude this. To date, retinal oxidase activity has been identified in rat tissues (Bhat *et al.*, 1988b, Hupert *et al*, 1991), mouse liver (Lee *et al.*, 1991), and the above mentioned rabbit hepatocyte activity. It would be interesting to evaluate the activity of these enzymes with CRBP-retinal, which may prove to be important in terms of physiological relevance.

Xanthine oxidase (dehydrogenase form) is also a flavometalloprotein with similar substrate specificities to the aldehyde oxidases and generally a wider tissue distribution (Lee *et al.*, 1991). This protein has also been identified as possessing the ability to convert retinal to retinoic acid (Lee *et al.*, 1991) with a low K_m (0.9 μ M). These two enzymes (aldehyde and xanthine oxidases) may play an important role in certain tissues, possibly those in which other retinal-oxidising enzymes are not present, or present at low levels (Lee *et al.*, 1991).

1.3 Cellular Retinoid-Binding Proteins

In the search for nuclear retinoid receptor proteins, a number of cytosolic retinoid binding proteins were discovered. These small proteins (~15 kDa) are members of the larger intracellular lipid-binding protein family (iLBP's), which also includes the fatty acid-binding proteins. To date, four cytosolic retinoid-binding proteins have been isolated and characterised, though it is possible that more exist. There are two cellular retinol-binding proteins (CRBP and CRBP II), and two cellular retinoic acid-binding

proteins (CRABP I and CRABP II). The genes for all four proteins have been cloned, and the structures have been determined by X-ray crystallography. Although the structures are well understood, the functions of these proteins are less well understood. The cellular retinoid binding proteins facilitate the transport of the hydrophobic retinoid molecules, by sequestering them from the aqueous environment. A number of enzymes have been shown to recognise free retinoids as substrates, and activity of these enzymes may lead to random interconversion of free retinoids *in vivo*. By binding to their particular ligand(s), the cellular retinoid binding proteins are proposed to lend some stability and specificity to retinoid homeostatic mechanisms, by restricting enzyme access to the retinoids. In addition, this sequestration would restrict partitioning of retinoids into membranes and cellular compartments, which would have a deleterious effect on cellular integrity.

1.3.1 CRBP I

CRBP I was first identified in rat and human cytosol by sucrose gradient centrifugation (Bashor et al., 1973) during the search for retinoid receptor proteins analogous to those The molecular weight of the retinol-binding protein found for steroid hormones. identified (~15 kDa), made it unlikely that this was the retinoid equivalent to the steroid hormone nuclear receptors. However, its retinol binding property warranted further investigation. The protein was first purified from a number of sources in 1978 - rat liver (Ong & Chytil, 1978), rat testis (Ong & Chytil, 1978; Ross et al., 1978), and bovine retina (Saari et al., 1978). Each source yielded a single polypeptide chain of ~15 kDa. The cloning of human CRBP (Colantuoni et al., 1985) revealed a protein sequence of 134 amino acids, while protein sequencing methods applied to rat CRBP (Sundelin et al., 1985) also yielded an 134 amino acid protein. The rat and human CRBP protein sequences show 96 % identity, differing only at 5 positions. In rat, CRBP protein is found in most tissues, with the highest expression in liver, kidney and proximal epididymis (Kato et al., 1985; Porter et al., 1983). The corresponding mRNA levels agree well with the levels of expressed protein (Levin et al., 1987; Rajan et al., 1990). However, in human tissues the highest levels of CRBP are found in the ovary, followed by the testis and the liver (Fex & Johannesson, 1984; Ong & Page, 1986). CRBP I has been identified in every vertebrate tested although there are no invertebrate or prokaryotic homologues known as yet. The isoelectric point of CRBP purified from rat testis was shown to be 4.8 - 4.9 (Ross et al., 1978).

1.3.1.1 Ligands

The main physiological ligand of CRBP is all-trans retinol (Ong et al., 1976; Saari et al., 1982). It was originally thought that CRBP was unable to bind all-trans retinal (Bashor et al., 1973; MacDonald & Ong, 1987; Ong & Chytil, 1975), but it is now generally accepted that CRBP binds both all-trans retinol and all-trans retinal with a stoichiometry of 1:1 (Levin et al., 1988; Li & Norris, 1996; Li et al., 1991; Napoli, 1996b). CRBP also forms high affinity complexes with 13-cis retinol and 3-dehydro-retinol, but does not bind to 9-cis or 11-cis retinol, or to retinoic acid (MacDonald & Ong, 1987). The elucidation of the tertiary structure of CRBP (Cowan et al., 1993; Newcomer et al., 1981) enables us to understand the basis for this selectivity in binding (see section 1.3.1.2, and Figure 1.5). Spectrofluorometric techniques were used to study the binding characteristics of recombinant CRBP (Levin et al., 1988). The excitation, emission, and absorption spectra of all-trans retinol are significantly changed upon binding of apo-CRBP, and vice-versa, and these differences were used to quantitate retinol binding. The apparent K_d values obtained for a single binding site using data both from monitoring all-trans retinol fluorescence and from monitoring protein fluorescence were similar (26 ± 20 nM and 19 ± 8 nM, respectively) (Levin *et al.*, 1988). The calculated stoichiometry of binding at saturation was 1:1. The same methods were used to calculate the binding of all-trans retinal to CRBP. In contrast to previous data using competitive binding studies (Ong & Chytil, 1975) which showed that molar excesses of all-trans retinal could not displace all-trans retinol from holo-CRBP, these direct binding studies indicated that CRBP could bind all-trans retinal. Monitoring changes in protein fluorescence only (as retinal, unlike retinol, has no intrinsic fluorescence), an apparent $K_{\rm d}$ for a single binding site was found to be 50 ± 3 nM (Levin *et al.*, 1988). Spectral analysis confirmed this interaction. The wavelength of maximum absorbance of retinal complexed with CRBP was red-shifted approximately 33 nm from its spectrum in water (from 374 nm to 407 nm).

An alternative method employed to study the interaction of retinal and retinol with CRBP was comparative ¹⁹F NMR studies (Li *et al.*, 1991). Incorporation of 6-fluorotryptophan (6-FTrp) into CRBP allows differentiation between the resonances of 6-FTrp derivatives of apo-CRBP and CRBP complexed with both retinol and retinal. The apparent dissociation constants for CRBP complexes measured using this method were: CRBP-retinol, $K'_d = \sim 10$ nM and CRBP-retinal, $K'_d = \sim 50$ nM.



Figure 1.5: Ribbon Diagram of CRBP I

This diagram was constructed using the program TURBO-FRODO (Cambillau *et al*, 1996) from the PDB entry CRB1 (Cowan *et al.*, 1993). α -Helices are coloured red, β -strands blue, and loops yellow, while all-*trans* retinol is shown in green.

1.3.1.2 Structure of CRBP I

The first crystals of CRBP were grown in 1981 (Newcomer *et al.*, 1981), and the structure of CRBP in complex with all-*trans* retinol was solved to 2.1Å resolution in 1993 (Cowan *et al.*, 1993). The large time period (12 years) between obtaining crystals and publishing the structure, indicates that this structure was exceedingly difficult to determine. Crystals were grown by hanging drop vapour diffusion of rat liver CRBP protein. The crystals were orthorhombic with a space group $P2_12_12_1$. The structure was solved using molecular replacement from P2 myelin protein, and the final R-factor was 18.8%.

The basic structure of CRBP is a single domain orthogonal β -barrel (Figure 15). Although primary sequence identity between iLBP family members is fairly low - as little as 12 % identity between CRBP I and an insect fatty acid binding protein - the basic structural framework is highly conserved within the family. The structure consists of 10 anti-parallel β -strands organised into 2 anti-parallel β -sheets which are folded over to create the 'barrel'. Each strand is connected to the next by a short reverse turn, except strands A and B which are connected by a helix-turn-helix motif.

1.3.1.3 Ligand Binding

The ligand fits into the centre of the barrel, lying along the barrel axis, with the functional group in the core of the protein (Figure 1.5). Interactions in the form of hydrogen bonds, salt links and hydrophobic interactions are made both with internal protein side chains (from sheets and helices), and with buried solvent molecules. There is no connection between the internal closed cavity and the external solvent.

There are usually 2 entrances to the internal cavity created by an orthogonal barrel structure. However in CRBP I, both of these are closed off - one by α -helices and the other by side chain residues. This raises an interesting question as to how the ligand gains entry to and is released from the protein. The structures of the apo- and the holoproteins show very few differences (Winter *et al.*, 1993), and it is not obviously apparent how retinoids enter into the binding cavity. A current theory is that time dependent fluctuations in the conformation of the structure exist in solution. It is likely that one or more of these conformations may be more open which may allow ligand access. A candidate for the flexible part of the structure which could facilitate access is the helix-turn-helix motif between sheets A and B, functioning as a 'helical cap' (Jamison *et al.*,

1994). Favourable van der Waals contacts between the retinoid β -ionone ring and the helices exist in the holoprotein which may stabilise this structure. However in the absence of ligand, and without these interactions, the helical cap may be able to move enough to allow the ligand to enter. A similar process has been described for two lipases (Brzozowski *et al.*, 1991; Tilbeurgh *et al.*, 1993).

An examination of the tertiary structures of apo- and holo- CRBP I gives a structural basis for the observed differences in affinity for all-trans retinol and retinal. There are a number of general interactions between the binding protein and the ligand, which would be expected to occur whether retinol or retinal was the ligand. A positive charge on the side chain of lysine 40 interacts with the isoprene tail of the retinoid. There are eight water molecules in the internal cavity which participate in hydrogen bonding and create a cavity with a shape complementary to the ligand. However, there is one interaction at the tail end of the retinoid which occurs with all-trans retinol as the ligand but not with retinal (Figure 1.6). The carbonyl oxygen on the side chain of glutamine 108 (Gln 108) forms a hydrogen-bond with the alcohol moiety of retinol. At the same time, there is a stabilising interaction between an amino group hydrogen (Gln 108) and the phenylalanine ring of phenylalanine 4 (Phe 4). When retinal is the ligand, the most likely situation is one where the same favourable electrostatic interaction between Phe 4 and Gln 108 occurs. This leaves no hydrogen donors in the proximity which are free to form an interaction with the oxygen from the aldehyde carbonyl group of retinal. Alternatively, Gln 108 could rotate, providing a hydrogen donor from the amino side-chain which would stabilise the binding of retinal. The hydrogen bonding geometry between Gln 108 and the carbonyl oxygen of retinal adopted in this scenario is sub-optimal, and this interaction would also erase the favourable electrostatic interaction with Phe 4. These observations largely account for the different dissociation constants observed for CRBP I binding to all-trans retinol and all-trans retinal.

The inability of CRBP I to bind retinoic acid, and *cis*-isomers of retinol and retinal may be explained fairly straightforwardly. The binding of retinoic acid would introduce another negative charge into the barrel, leaving no space in the cavity for extra water molecules for solvation. This is considered very energetically unfavourable. The overall shape of the 9-*cis* or 11-*cis* isomers of retinal and retinol (Figure 1.1) compared to the all-*trans* conformation is such that it may be sterically impossible for the molecule to gain access to the highly restricted site, and for the molecule to fit in the highly defined site if it could gain access.



KEY: _____ electrostatic interaction hydrogen bonding

Figure 1.6: Schematic Diagram Showing Structural Basis for Observed Binding Constants of CRBP and CRBP II for Retinal and Retinol

Diagram adapted from Banaszak et al., 1994, p137

1-30

1.3.1.4 Function of CRBP I

Recent research has indicated that the function of CRBP I and II (see also section 1.3.2.3) is to direct retinoids to different sites of intracellular metabolism (Boerman & Napoli, 1996; Fiorella & Napoli, 1991; Kakkad & Ong, 1988; Napoli *et al.*, 1995; Napoli *et al.*, 1991; Napoli & Race, 1987; Ong & Page, 1987; Ottonello *et al.*, 1993; Posch *et al.*, 1991; Posch *et al.*, 1992; Saari *et al.*, 1982). In many retinoid responsive cells the concentration of CRBP exceeds the concentration of free unesterified retinol (Bashor *et al.*, 1973; Harrison *et al.*, 1987; Ong, 1982; Yost *et al.*, 1988). The total concentration of CRBP in liver cells has been estimated at ~7 μ M, while the total concentrations of retinol and retinal have been calculated at 5 μ M and 0.1 μ M, respectively (Harrison *et al.*, 1987; Napoli *et al.*, 1992). The physiological concentrations, along with the low dissociation constants of the binding proteins for their ligands, indicate that intracellular retinoids are likely to be bound to the CRBPs *in vivo*. This being the case, the most abundant and available substrate for retinoic acid biosynthesis would be holo-CRBP (Bashor *et al.*, 1973; Ong, 1982).

A number of studies have produced evidence in support of the hypothesis that CRBP in both its apo- and holo- forms plays a vital role in directing retinoid metabolism. In 1991 it was demonstrated that holo-CRBP was a substrate for retinal synthesis from microsomal liver, kidney, lung and testis preparations (Posch *et al.*, 1991). Four major observations provided evidence that the retinal synthesis was from retinol-CRBP, rather than from free retinol in equilibrium with holo-CRBP. These were: 1) the rate of synthesis from holo-CRBP exceeded the rate observed from the concentration of free retinol in equilibrium with holo-CRBP, 2) with holo-CRBP as a substrate, the rate was 3-fold greater when NADP⁺ was used as a cofactor compared to NAD⁺, while with free retinol, NAD⁺ and NADP⁺ dependent rates were equal; 3) an increase in the concentration of apo-CRBP which would decrease the concentration of free retinol in equilibrium did not affect the rate; and 4) an increase in holo-CRBP while maintaining a constant free retinol concentration increased the rate.

In 1990, it was demonstrated that microsomes cannot convert retinal produced from holo-CRBP into retinoic acid at high rates in the presence of apo-CRBP (Napoli & Race, 1990). The previous two studies together indicate a low microsomal retinal dehydrogenase activity, a high microsomal retinol dehydrogenase activity, and support binding of retinal to CRBP - an originally controversial issue. These observations were

further supported when Posch and co-workers observed a rat liver microsomal NADP⁺dependent enzyme activity which catalysed the oxidation of CRBP-retinol to retinal, and a cytosolic dehydrogenase activity converting retinal generated from the microsomal activity into retinoic acid (Napoli *et al.*, 1992; Posch & Napoli, 1992). A two step reaction requiring both microsomal and cytosolic subcellular fractions was proposed. The cytosolic enzyme activity was later isolated and purified, and shown to use free retinal or retinal-CRBP as a substrate (Posch *et al.*, 1992) (see section 1.2.2.3). In contrast to the above studies, a calf liver cytosolic retinol dehydrogenase enzyme activity which oxidised retinol-CRBP to retinal with an exogenous supply of NAD⁺ or NADP⁺ was observed (Ottonello *et al.*, 1993). This activity was distinct from previously identified cytosolic retinol oxidising activities, which were isolated using free retinol as a substrate (Kim *et al.*, 1992; Napoli & Race, 1987), and was the first cytosolic CRBPretinol oxidising activity identified.

In addition to the role played by CRBP in directing the biosynthesis of retinoic acid by sequestering retinol/retinal from multiple cytosolic and microsomal activities shown to utilise free retinoids, apo-CRBP itself is thought to be important in this process. Apo-CRBP has been demonstrated to be a specific inhibitor of retinoic acid synthesis from CRBP-retinol by the cytosolic dehydrogenase activity (Ottonello et al., 1993), but does not inhibit the microsomal activity (Boerman & Napoli, 1996). The presence of significant amounts of apo-CRBP would be likely to indicate low circulating retinol. In this situation, apo-CRBP stimulates hydrolysis of endogenous retinyl esters by cholateindependent retinyl ester hydrolase activity in microsomes (Boerman & Napoli, 1991), freeing stored retinol for subsequent conversion to retinoic acid. The rate of release of retinol has a Michaelis-Menten dependence on the concentration of apo-CRBP, with $2.6\,\mu\text{M}$ apo-CRBP producing half-maximal rate stimulation. No more than 10 % of endogenous retinyl esters were depleted, and no more than 4 % of apo-CRBP was converted into holo-CRBP at the apo-CRBP concentrations that produced a maximum rate of retinyl ester hydrolysis. Finally, it has been shown that a microsomal LRAT uses holo-CRBP to synthesise retinyl esters (Ong et al., 1988; Yost et al., 1988). This activity is also inhibited by apo-CRBP. Here again, in times of low retinol levels, available retinol is required to maintain retinoic acid synthesis, so apo-CRBP levels are increased and esterification of retinol for storage is inhibited by apo-CRBP. These observations point to the amounts of, and ratios of, apo- and holo-CRBP playing an important role in directing retinoids to their metabolic fate.

In contrast, it has also been proposed that the major role of CRBP I lies solely in the storage of retinoids, and that it plays no role in retinoid metabolism (Duester, 1996). The catalytic efficiency of the major microsomal retinol dehydrogenase, RoDH 1, has been estimated in both the direction of retinal reduction (93 min⁻¹ mM^{-1}), and retinol oxidation (17 min⁻¹ mM⁻¹) (Duester, 1996). It can be seen that this enzyme appears to be approximately five times more efficient in the direction of retinal reduction. Supporting this calculation is the fact that the microsomal retinol dehydrogenase has been shown to use NADP(H) as a cofactor in preference to NAD(H) (Napoli, 1996a; Napoli et al., 1995; Napoli et al., 1992; Posch et al., 1991; Posch et al., 1992). Because of the high physiological ratio of NADPH to NADP⁺ (Veech et al., 1969), retinal reduction would again be expected to be the preferred direction. The specificity of the microsomal enzyme for all-trans isomers is another reason put forward for the primary role of CRBP being one of storage. Retinoid turnover should be able to accommodate any isomeric retinoid forms which may arise; however, storage systems need only be concerned with all-trans isomers (see section 1.1.4). It is also known that CRBP-retinol is a substrate for retinyl ester formation. It is thus suggested that CRBP I along with the microsomal retinol dehydrogenase play a role in the liver analogous to that played by CRBP II and retinal reductase in the intestine (Duester, 1996) (see Section 1.3.2.3).

It seems generally accepted that CRBP I plays an important role in the storage of retinol in the liver. However, what is not clear is what role, if any, CRBP I plays in the biosynthesis of retinoic acid from retinol. What is obvious though is the complexity of retinoid homeostasis, and the number of enzymes that may potentially play an important role.

1.3.1.5 Transfer of Ligand to Metabolic Enzymes

The manner in which the ligand is transferred from the binding proteins to the metabolic enzymes is unresolved. It is clear that there must be a 'release trigger', which initiates opening of the helical cap (which is closed in holo-CRBP) and release of the ligand enabling either direct binding to the enzyme or binding after dissociation into the aqueous media to occur. This trigger may be a protein-protein interaction between CRBP and the enzyme. The next step in the process is not clear. There is a body of evidence suggesting that retinoid intermediates pass through the metabolic pathway by direct transfer (Bernhard, 1988; Clegg, 1984; Srivastava & Bernhard, 1986), rather than by diffusion through the aqueous phase. If the direct transfer model is correct, this would raise a number of questions. Firstly, the binding protein, ligand and enzyme may

form a complex such that the product is subsequently released from the binding protein after metabolism. If this is the case, the rate of reaction for CRBP-ligand would be expected to be a lot slower than that for free retinoid. The hydroxyl or aldehyde group of retinol/al is known to be buried deep in the centre of the CRBP molecule. The molecule would have to be rotated at some stage in order for the metabolising enzyme to have access to the active group. This is unlikely to happen in this situation due to space constraints imposed by the CRBP binding site. The second hypothesis involves an interaction between the binding protein and the enzyme, the transferral of the ligand to the active site of the enzyme, and release of the product from the metabolising enzyme. This is a more likely situation, although the retinoid is still likely to need rotation, which would have to be facilitated by the enzyme. Again, the rate of reaction would be expected to decrease markedly compared to that for free retinoid. A third transfer mechanism would be dissociation of the ligand into the aqueous medium. Once more, this would presumably involve a protein-protein interaction to initiate the opening of the helical cap and release of ligand, and is still likely to result in a slower rate of reaction due to the necessity of dissociation of the ligand from the binding protein before the enzyme can gain access.

There is a lot of indirect evidence for the direct transfer hypothesis involving transferral of the ligand to the enzyme. The apparent K_m for the cytosolic NAD⁺ dependent holo-CRBP reaction observed by Ottonello and co-workers was $0.84 \pm 0.12 \mu M$, compared to $7.1 \pm 0.3 \,\mu\text{M}$ for free retinol. The V_{max} for the CRBP-retinol reaction, however was 3-15 times lower than that for free retinol (Ottonello et al., 1993). However, these data do fit all the proposed models. The high level of conservation of amino acids found on the exteriors of both CRBPs and CRABPs also supports the existence of a protein recognition site. Conservation of primary structure is seen even though neither tertiary structure conservation nor ligand specificity compels such surface conservation (Newcomer, 1995; Ong et al., 1994). If a critical protein interaction with the surface of the binding proteins existed, this may have driven the conservation of exterior sequence. Computer graphics modelling of a CRABP with a cytochrome P450 showed a possible model of transfer of retinoic acid via a bimolecular complex with juxtaposed ligand portals (Thompson et al., 1995). This supports the general protein recognition/ interaction scenario with a specific model that may be representative of similar mechanisms; for example, the transferral of retinol and retinal between CRBP and their respective metabolic enzymes.

1.3.2 CRBP II

A CRBP similar to, but clearly distinct from, CRBP I, has been identified and purified from rat pups (Ong, 1984). This protein (termed CRBP II) showed 56 % identity with CRBP I, and existed in two forms, one of which was N-terminally acetylated. Comparison of CRBP II purified from other species revealed ~91 % identity across species. CRBP II is also cytosolic, and is found only in the enterocytes of the small intestine postnatally (Levin *et al.*, 1987; Li *et al.*, 1986) where it makes up 0.4-1 % of total soluble intestinal protein - around 1000 fold more than CRBP I (Ong, 1984). Perinatally, CRBP II is also found in the liver; however, mRNA levels in the liver decline rapidly after birth (Li *et al.*, 1986; Ong *et al.*, 1991). In comparison to CRBP I, which has an extremely wide distribution, CRBP II expression is very restricted.

1.3.2.1 Ligands

Unlike CRBP I, CRBP II binds all-trans retinol and all-trans retinal with an approximately equal affinity (Levin *et al.*, 1988; MacDonald & Ong, 1987). CRBP II also binds 13-*cis* isomers, but not 9-*cis* or 11-*cis* isomers, or retinoic acid. The dissociation constant of CRBP II for all-trans retinal is ~50-90 nM - similar to that of CRBP I. Studies using fluorometric titration have indicated similar dissociation constants of both binding proteins for all-trans retinol - 10-20 nM for CRBP I (Levin *et al.*, 1988; Ong & Chytil, 1978), and 10-40 nM for CRBP II (Levin *et al.*, 1988; MacDonald & Ong, 1987; Ong, 1987). However, a comprehensive study using 6-fluorotryptophan substitution (see section 1.3.1.1) suggested that the affinity of CRBP I for all-trans retinol may be ~100 fold greater than that of CRBP II.

1.3.2.2 Structure of CRBP П

The overall tertiary structure of CRBP II is essentially similar to that of CRBP I and the other iLBPs, as determined in both apo- and holo- states (Winter *et al.*, 1993). One main difference is the substitution of a glutamine at position 4 (CRBP II) for the phenylalanine in CRBP I (Figure 1.6). This accounts for the approximately equal affinity with which CRBP II binds all-*trans* retinol and all-*trans* retinal. Gln-4 and Gln-108 could both function as either a hydrogen bond donor or acceptor; thus CRBP II could theoretically bind all-*trans* retinol and all-*trans* retinal equally well. The binding of either retinol or retinal to CRBP II gives a sub-optimal hydrogen bonding geometry. This accounts for both CRBP I having a greater affinity for retinol than CRBP II and the approximately equal affinity of both proteins for retinal.

1.3.2.3 Function of CRBP II

Considering the restricted location of CRBP II to the small intestine where it makes up a quantitatively major portion of the total protein, it seems certain that CRBP II functions in the absorption and storage of dietary retinoids. β -Carotene is oxidatively cleaved upon ingestion into two molecules of retinal, which may then be reduced to retinol. Both retinol and retinal are thus available for absorption. CRBP II can bind both retinol and retinal, and it is accepted that, due to the distribution and amounts present, the major function for CRBP II is in the dietary uptake and subsequent storage of these compounds.

CRBP II-retinol is recognised as a substrate by the enzyme LRAT (lecithin-retinol acyltransferase) (Herr & Ong, 1992), which uses phosphatidyl choline as an acyl source to esterify retinol for storage. However, the enzyme ARAT (acyl-CoA-retinol acyltransferase), which esterifies free retinol with long-chain fatty acids (Helgerud *et al.*, 1982), is unable to recognise retinol bound to CRBP II (Ong *et al.*, 1987), but can esterify free retinol. As the abundance of CRBP II makes it unlikely that there will be significant free retinol/retinal, ARAT is unlikely to be physiologically relevant, except in situations of excess retinoid ingestion where there may be some free retinol. CRBP II-retinal has been shown to be a substrate for reduction to retinol by microsomal preparations from the rat small intestine mucosa (Kakkad & Ong, 1988). However, retinol or retinal bound to CRBP II was not oxidised to any extent. This study confirms the role of CRBP II to be in the uptake and esterification of dietary retinoids, rather than in their subsequent metabolism.

1.3.3 CRABP I and II

Two other proteins involved in the binding of retinoids have been identified. These were first detected in 1974 (Sani & Hill, 1974), but were not recognised as being distinct from CRBP (which was first isolated in 1973), until 1988 when what is now known as cellular retinoic acid binding protein (II) (CRABP II) was isolated (Bailey & Siu, 1988). Both CRABP I and II bind all-*trans* retinoic acid tightly, 9-*cis* retinoic acid weakly, and do not bind 13-*cis* retinoic acid, retinol or retinal. Only CRABP I has been shown to bind retinoic acid *in vivo*. The apparent dissociation constants for the binding of all-*trans* retinoic acid were determined in one study to be approximately equal for both proteins ->0.4 nM and ~2 nM for I and II respectively (Norris *et al.*, 1994). However, the apparent K'₄ of CRABP II was seen to be considerably higher in a different study -~64 nM (Bailey & Siu, 1988). The CRABP proteins are 136 amino acids in length, and are very similar with respect to amino acid sequence, with human CRABP I and II having 72 % identity. CRABP I is very highly conserved across species, with mouse, bovine and rat CRABP I being 100 % identical, the human form showing only 1 amino acid difference, and human and mouse CRABP II sharing 94 % identity. As with the CRBPs, the tissue distribution of CRABP I is wide, while that of CRABP II is restricted to the skin and the corpus luteum in the adult animal (reviewed in Li & Norris, 1996). Both proteins however, are expressed strongly during embryogenesis.

The biological function for the CRABP proteins remains unclear. It was originally hypothesised that the CRABP would bind to retinoic acid and facilitate or direct its transport to the nuclear receptors (Takase et al., 1986). However there is conflicting evidence as to whether the CRABPs are able to enter the nucleus in vivo (Bucco et al., 1995: Donovan et al., 1995; Zetterstrom et al., 1994). There is no identifiable nuclear localisation signal on the CRABPs, though the protein is small enough to allow passive diffusion through nuclear pores. Also, the temporal and spatial patterns of expression of the CRABPs are different from those of the nuclear retinoid receptors (Dolle et al., 1989; Ruberte et al., 1993). One current idea is that the CRABPs are expressed in highly retinoid-sensitive cells, and act to protect the cell from a potentially harmful rise in cytosolic retinoic acid concentration by binding the retinoic acid and restricting its availability to the nuclear receptors (Donovan et al., 1995; Maden & Holder, 1991; Maden et al., 1988; Ruberte et al., 1993; Ruberte et al., 1992). Another possible function is the catabolism of retinoic acid into polar metabolites (e.g. 4-oxoretinoic acid), which may be controlled by CRABP. It has been shown that CRABP I-retinoic acid may serve as a substrate for enzymes which convert retinoic acid into metabolites (Boylan & Gudas, 1992; Fiorella & Napoli, 1991; Ong, 1994). It seems likely that CRABP plays an important role in controlling the availability of retinoic acid, though the precise nature of this role remains unclear. Gene knockout studies using mice lacking both CRABP I and CRABP II have shown that these mice are essentially normal (Lampron et al., 1995). This study argues against the CRABPs playing an important role in retinoid metabolism and signalling. Instead, the idea that the CRABPs may bind to retinoic acid during vitamin A deficiency to maintain retinoid signalling was suggested.

1.3.4 Plasma Retinol-Binding Protein (RBP)

The plasma retinol-binding protein (RBP) is a retinol binding protein distinct from the CRBPs. It contains one binding site for retinol, and functions to transport retinol in the serum fraction of blood from its storage site in the liver, to target cells. RBP-retinol is

transported in a 1:1 complex with transthyretin (prealbumin) to minimise loss by glomerular filtration.

RBP consists of a single 182 amino acid polypeptide chain with a molecular weight of 21 000 Da (Kanai et al., 1968). The elucidation of the primary amino acid sequences showed sequence conservation across species, with rat and human showing 86 % sequence identity, and rabbit and human 92 %. The protein was crystallised in the early eighties (Newcomer et al., 1984b; Ottonello et al., 1993), and the first tertiary structure published soon after (Newcomer et al., 1984a). The basic structure consists of a β -barrel core made up of 8 antiparallel β -strands organised into 2 orthogonal β -sheets, an N-terminal coil, a C-terminal coil and an α -helix. This is the basic structural organisation of a family of proteins which bind small hydrophobic molecules, and includes RBP, β -lactoglobulin, apolipoprotein D and odorant-binding protein. This family of lipid-binding proteins is different from the family to which the CR(A)BPs belong. Apart from the general function and cellular localisation, a striking difference between RBP and the CR(A)BPs is seen in the orientation of ligand binding, which is likely to be related to the function. Retinol bound to RBP is oriented with the β -ionone ring deep in the centre of the structure, and the isoprene tail and the functional group stretching out towards the surface. The functional group is located deep in the barrel in the CR(A)BPs.

1.3.5 Retinoid-Binding Proteins of the Visual System

The uniqueness of the visual system lends itself to the existence of specific proteins involved in certain processes. In particular, the importance of 11-*cis* retinoids in the eye has led to the isolation of a number of proteins involved in their sequestration and metabolism. The existence of these proteins is noted to give a complete overview of retinoids and associated proteins, although this group of proteins is not thought to play a significant role in the processes studied in this thesis.

Cellular retinal-binding protein (CRALBP) is an intracellular lipid-binding protein whose major ligands are the aldehyde forms of vitamin A. This protein shows is not homologous to any of the other cellular retinoid binding proteins, or indeed to any other known protein. It has been detected in the neural retina, retinal pigment epithelial cells and the pineal gland, but in no other tissues examined. CRALBP is thought to function as a retinal binding protein in the visual system (Saari & Bredberg, 1988; Saari *et al.*, 1982).

Three types of soluble retinal-binding protein (RALBP) have been found in invertebrate retinae, two from honeybee retina and one from squid (Ozaki *et al.*, 1994). The structure of RALBP appears to be distinct from other isolated retinoid binding proteins, and is postulated to play a role in the transport of 11-*cis* and all-*trans* retinals, and in the maintenance of sensitivity in photoreceptor cells (Ozaki *et al.*, 1994).

Interphotoreceptor-binding protein (IRBP) is a high molecular weight protein located in the interphotoreceptor matrix of the retina, whose postulated function is to transport retinoids between the retinal pigment epithelial cells and photoreceptor cells (Ozaki *et al.*, 1994).

1.3.6 Cis Isomers of Retinoids

The importance of the retinoid geometric isomer 9-*cis* retinoic acid is best illustrated by its preferential binding of the RXRs (see section 1.4). Two important questions arise from this observation - how is 9-*cis* retinoic acid generated in the body, and how is its production controlled? Possible methods by which 9-*cis* retinoic acid is produced are either by isomerisation (enzymatic or non-enzymatic), or by a similar pathway to the production of all-*trans* retinoic acid from a biological precursor. The second possibility is supported by the isolation of a 9-*cis*-specific alcohol dehydrogenase (SDR family) from a human breast tissue library. Northern analysis identified RNA encoding this protein primarily in testis, kidney and mammary gland (Mertz *et al.*, 1997).

The isomerisation of all-*trans* retinoic acid to 9-*cis* and 13-*cis* retinoic acid has been observed *in vivo* and *in vitro* (Nagao & Olson, 1994; Urbach & Rando, 1994), but the method by which isomerisation occurs is not known. However, this isomerisation was observed to be temperature dependent, and to vary between cell types, indicating enzymatic catalysis (Mangelsdorf *et al.*, 1994). Significant isomerisation of all-*trans* retinol or retinal was not observed in work performed with bovine liver microsomes (Urbach & Rando, 1994). A likely method of 9-*cis* retinoic acid production comes from the bioavailability of 9-*cis*- β -carotene, which is found naturally in foods (Chandler & Schwartz, 1987; Khachik *et al.*, 1992). It has also been estimated that 9-*cis*- β -carotene may account for up to 25 % of total β -carotene in human liver (Stahl *et al.*, 1993). Experiments performed using rat liver and intestine preparations showed an enzymatic conversion of all-*trans* or 13-*cis*- β -carotenes to all-*trans* retinal (Nagao & Olson, 1994). The production of 9-*cis* retinal in this manner would necessitate an enzyme activity to

oxidise the aldehyde to a carboxylic acid. A number of enzymes are able to catalyse this reaction, among them AlDH 1 (see section1.2.2). Furthermore, if this was the correct physiological scenario, it would be likely that 9-*cis* retinoic acid production would be controlled in a manner such as the all-*trans* situation. It is possible that a family of binding proteins related to the CRBPs, but which are specific for *cis*-retinoid isomers, may exist.

1.4 Retinoic Acid Exerts its Effects Via Nuclear Receptors

The identification of nuclear receptors which bind retinoic acid with high affinity revealed a mode of action whereby the pleiotropy of retinoids could begin to be understood (Figure 1.7). The characterisation of steroid and thyroid hormone receptors revealed common structural motifs and functional properties (Evans, 1988). Identification of a highly conserved cysteine-rich DNA-binding region was exploited in the screening for related gene products. Using this strategy, a number of genes were isolated which coded for proteins belonging to the steroid/thyroid hormone receptor superfamily, and for which the putative ligand was yet to be identified (Chambon, 1996; Giguere, 1990; Giguere, 1994; Giguere & Evans, 1990; Mangelsdorf et al., 1994; and refs. therein). These gene products, together with those identified by transcription factor cloning and analysis of *Drosophila* developmental pathways are referred to as orphan receptors. Manufacturing chimeric constructs consisting of orphan receptors containing the DNA-binding domain of known steroid hormone receptors, and then transfecting the chimera into a cell line along with a reporter gene which responds to the known steroid receptor, provides a system whereby putative ligands can be tested for a positive response (Giguere & Evans, 1990). The use of this technique with the putative ligand retinoic acid (which invoked a potent response), led to the identification of the first retinoic acid receptor (RAR), a member of the steroid/thyroid hormone receptor superfamily (Giguere et al., 1987; Petkovich et al., 1987).

It is in the high affinity binding of retinoic acid to the RARs, and the subsequent binding of the receptor-ligand complex to defined sequences of DNA in the promoter regions of target genes, whereby retinoic acid exerts its effects (Chambon, 1996; Giguere, 1990; Giguere, 1994; Mangelsdorf *et al.*, 1994; and refs. therein). Genes isolated from many organisms contain retinoid response elements (see Table 1.3). The protein products of these genes exhibit a wide variety of functions, yet many of them are involved themselves in retinoid homeostasis, providing another layer of control in the overall retinoid



Figure 1.7: Overview of Mechanism of Retinoid Action

Retinoids enter the nucleus and bind various nuclear receptors. Binding of ligand induces dimerisation of nuclear receptors via dimerisation domains in the LBD and DBD, and susequent binding to response elements (e.g. RARE's) on target genes. This affects the transcription of the gene. Dimer combinations are shown in part A. Diagram adapted from Chambon, 1996, p944.

Abbreviations: RARE - retinoic acid response element; DBD - DNA-binding domain; DD - dimerisation domain; LBD - ligand binding domain; ATRA - all-*trans* retinoic acid; VDR - vitamin D receptor; T₃R - thyroid hormone receptor; N - amino-terminal domain; Orphan - orphan steroid receptor.

RARE Containing Genes		
DR 5 e.g. AGGTCAXXXXAGGTCA		
mRAR β 2, α 2	mouse RAR isoforms $\beta 2$, $\alpha 2$	
hRAR β 2, α 2, γ 2	human RAR isoforms $\beta 2$, $\alpha 2$, $\gamma 2$	
mCP-H	mouse complement factor H	
hCMV-IEP	human cytomegalovirus immediate early promoter	
hMGP	human matrix Gla protein	
mHox-1.6		
hADH3	human alcohol dehydrogenase 3 (class 1)	
DR 2 e.g. AGTT	CAXXAGTTCA	
mCRBP1	mouse CRBP 1	
rCRBP1	rat CRBP1	
hApoAl	human apolipoprotein A l	
mCRABPII-1	mouse CRABPII-1	
Palindrome e.g. AGGTCAXXXTGACCT		
bGH	bovine growth hormone	
hOST	human osteocalcin	
xVitA2	xenopus vitellogenin A2	
Complex e.g. TGACCTXXXXXXGGGTAAAGGTGA		
rGH	rat growth hormone	
mLamβl	mouse laminin β l	
hOXY	human oxytocin	
hMCAD	human medium chain acyl-coA debydrogenase	
RXRE Cont	aining Genes	
DR1 e.g. AGGO	GCAXXAGGTCA	
rCRBPII	rat CRBPI	
cOVAL	chicken ovalbumin	
rAOX	rat acyl-coA oxidase	
hApoAl	human apolipoprotein Al	
mCRABPII	mouse CRABPII	
mMHCI	mouse major histocompatibility complex 1	
rPEPCK	rat phosphoenolpyruvate carboxykinase	
hHBV	human hepatitis B virus	
Complex e.g. TGACCT(8)GGGTAAAGGTGA		
mMCAD	mouse medium chain acyl-coA dehydrogenase	

Table 1.3 Genes Containing Retinoid Response Elements¹

Table adapted from Mangelsdorf et al., 1994 pp328-329

¹See Mangelsdorf et al., 1994, and refs. therein for referencing

signalling system. It is postulated that upon hormone binding, the receptor undergoes an allosteric change which allows DNA binding; however, the mechanism of the process by which transcription is altered is unknown. Most members of the steroid/thyroid hormone receptor superfamily are nuclear whether bound to ligand or not, but it remains controversial whether unliganded receptors are free or bound to DNA. Also unknown is the manner in which retinoic acid is transported to the nucleus. The CRABPs have been hypothesised to play a role in intracellular retinoic acid transport (Donovan *et al.*, 1995), but this issue also remains unsolved.

1.4.1 Retinoic Acid Receptors (RARs)

Three genes encoding highly related RARs have been identified from mouse and human - RARα (Giguere et al., 1987; Petkovich et al., 1987), RARβ (Benbrook et al., 1988; Brand et al., 1988), and RARy (Ishikawa et al., 1990; Krust et al., 1989; Zelent et al., 1989). Each gene gives rise to multiple isoforms by the use of alternative splicing of exons, translation initiation at an internal CUG codon, and differential usage of two promoters (Leid et al., 1992a). RARs show greater than 95 % identity in the DNAbinding domain between species and between subtypes within a species, and 85 % identity in the ligand-binding domains. The RARs consist of six regions (A-F) based on homology within the steroid receptor superfamily. These regions include a variable length non-conserved N-terminal domain (A), a central DNA-binding domain (DNA-BD) (C), and a large C-terminal ligand-binding domain (LBD) (E). The LBD and DNA-BD are linked by a short hinge region (D), that may function in nuclear translocation. The highly conserved DNA-BD contains two zinc-finger motifs which allow target DNA sequence recognition, and are involved in the activation of target genes (Chambon, 1996). Besides ligand binding, the LBD has additional features - a ligand inducible transactivation function, a silencing activity involved in transcriptional repression, and a dimerisation surface. The amino-terminus is also involved in transactivation, and may contribute to the specificity of RAR action as the N-terminal is the most variable region. The RARs are able to bind all-trans retinoic acid and 9-cis retinoic acid with high affinity — $K_d = 1-5$ nM (Ishikawa et al., 1990; Yang et al., 1991), and 0.2-0.7 nM (Allenby et al., 1993) respectively.

1.4.2 Retinoid X Receptors (RXRs)

The identification of an orphan receptor showing slight activation by all-*trans* retinoic acid led to the hypothesis that this receptor (named RXR) was a new class of retinoic acid receptor whose natural ligand may be an isomer of retinoic acid (Mangelsdorf *et al.*,

1990). Independent researchers then identified this isomer to be 9-*cis* retinoic acid (Heyman *et al.*, 1992; Levin *et al.*, 1992a). Three RXR genes have been identified in mouse, as well as homologues in human, rat, chicken and frog (see Giguere, 1994 and Chambon, 1996 for review). Each gene (α , β , γ) can give rise to two isoforms, which, like the RARs, differ in the N-terminal region. The physiological ligand for the RXRs is 9-*cis* retinoic acid, which binds and activates with high affinity (K_d = 10-15 nM); however, all-*trans* retinoic acid does not activate RXRs efficiently (Heyman *et al.*, 1992; Levin *et al.*, 1992b; Mangelsdorf *et al.*, 1990). RARs and RXRs show lower sequence similarity than would be expected for two members of the same family which can both be activated by the same ligand - in fact 62 % identity for the DNA-BD, and only 27 % identity for the LBD (Mangelsdorf *et al.*, 1994). The RARs show more similarity with thyroid hormone receptors (TR) than RXRs.

1.4.3 Retinoid Response Elements

The specific sequence of DNA in the promoter regions of target genes (response elements), to which liganded RARs/RXRs recognise and bind, is based on a minimal half-site consensus — (A/G)G(G/T)TCA. This half-site is organised into a variety of specific elements using direct repeat, palindrome, and complex motifs (Naar *et al.*, 1991; Umesono & Evans, 1989; Umesono *et al.*, 1988; Umesono *et al.*, 1991). The retinoic acid response element (RARE) which is the most abundant and gives the strongest response is a direct repeat of the half site consensus separated by five nucleotides (DR5). DR2, palindromic and random complex RAREs have also been identified *in vivo* and *in vitro*; however, these are less abundant and less responsive, and in the case of the palindromic and random sequences, require overexpression of the RAR for activity. The RXR response elements (RXREs) are almost exclusively DR1 (Mangelsdorf *et al.*, 1994; Mangelsdorf *et al.*, 1991). Negative RAREs have also been identified, which are involved in the downregulation of certain genes (Mangelsdorf *et al.*, 1991; Nicholson *et al.*, 1990; Schule *et al.*, 1991). It is thought that these negative effects occur via RAR interfering with other positive acting transcription factors.

1.4.4 Dimerisation of Retinoid Receptors

Subsequent studies of the binding of liganded RARs to RAREs, showed that RARs alone did not bind efficiently to these sequences, but an 'accessory factor' found in nuclear extracts of a variety of cells greatly enhanced binding (Leid *et al.*, 1992b; Yu *et al.*, 1991). This factor was identified to be an RXR (Kliewer *et al.*, 1992; Leid *et al.*, 1992b; Yu *et al.*, 1991). It has since been shown that RXRs form heterodimers in solution with

RAR, TR, vitamin D receptor (VDR), and peroxisome proliferator-activated receptor (PPAR), hence the description of RXRs as 'promiscuous dimerisation partners'. Heterodimers bind more efficiently than homodimers to response elements, and binding is more selective. The high affinity binding of heterodimers results from dimerisation in solution through the LBD dimerisation surface, and upon DNA binding, from the formation of a second dimerisation interface through the DNA-BD which fits the length of the spacer between half sites (Figure 1.7).

The transmission of the retinoid signal to target genes is preferentially relayed by RXR/RAR heterodimers (Chambon, 1996; Kastner *et al.*, 1994). However, the role of RXR homodimers (which form in solution and bind to DR1 RXREs) in retinoid signalling and the role of RXR as a heterodimeric partner for other steroid superfamily members remains to be clarified. The presence of a dimerisation sequence in the LBD also raises the question of the role of ligand in heterodimer formation and signalling. Depending on its heterodimeric accessory, RXR may be a silent partner whose primary role is in specifying binding, or it may be a transcriptionally active ligand responsive partner (Leblanc & Stunnenberg, 1995; Mangelsdorf *et al.*, 1995). So though the role of 9-*cis* retinoic acid remains unsettled, it is likely to play an substantial role in retinoid signalling, and in steroid signalling in general.

Some of the diversity seen in retinoid functions can be attributed to aspects of retinoid receptors and their interaction with response elements. There are at least 14 RAR/RXR isoforms, two identified ligands, and 48 possible heterodimer combinations (though due to tissue and cell specific expression, not all combinations will always be possible). The next level of complexity lies in the response elements. These can be direct repeats separated by 1-5 nucleotides, palindromic repeats, or random sequences. A final area of variation lies in the asymmetry of heterodimer binding to response element, i.e. which partner binds at the 5' half-site, and which at the 3' half-site, which may be fixed or varied.

1.5 Vitamin A and Fetal Alcohol Syndrome (FAS)

Fetal alcohol syndrome arises due to excessive prenatal maternal ethanol ingestion (Clarren & Smith, 1978; Jones *et al.*, 1973; Lemoine *et al.*, 1968). The syndrome is characterised by craniofacial defects of the eyes, upper lip and jaw, growth retardation, and central nervous system dysfunction. The world-wide incidence of full FAS is

estimated to be 1.9 per 1000 live births, and for partial FAS (known as fetal alcohol effects or FAE), 3-5 per 1000 live births (Luke, 1990). Among alcoholic women, the estimated incidence rises to 25 per 1000 (FAS), and 90 per 1000 (FAE), making FAS the leading cause of mental retardation in the western world (Streissguth *et al.*, 1986). Symptoms observed in FAS-affected individuals vary widely, both in the severity of the malformation, and in the type of malformation seen. The molecular basis for ethanol-induced teratogenicity is unknown. Current ideas as to how ethanol causes FAS are: 1) direct teratogenic effects of ethanol or acetaldehyde on the fetus, 2) altered maternal or placental physiology, and/or 3) nutritional alterations caused by ethanol ingestion (DeJonge & Zachman, 1995). A nutritional compound proposed to be involved in ethanol-induced teratogenicity is vitamin A (DeJonge & Zachman, 1995; Deltour *et al.*, 1996; Duester, 1991; Duester, 1994; Grummer *et al.*, 1993; Grummer & Zachman, 1995; Morriss-Kay & Sokolova, 1996; Pullarkat, 1991).

The importance of retinoic acid during embryogenesis in basic processes such as cell differentiation, cellular rearrangement and pattern formation, has been well documented (see section 1.1.3). Many defects associated with vitamin A deficiency or toxicity are similar to those observed in FAS-affected individuals; for example, congenital heart defects seen in FAS are similar to those found in vitamin A teratogenesis. (DeJonge & Zachman, 1995). From these general observations, the search for a molecular basis linking FAS to alterations in vitamin A homeostasis was initiated. It is known that fetal and adult vitamin A status is affected by ethanol ingestion (Grummer et al., 1993; Grummer & Zachman, 1990; Lieber, 1991). More specifically, maternal ethanol ingestion has been identified to decrease fetal liver retinol levels, increase fetal lung and kidney retinol and retinyl palmitate stores, increase fetal brain retinoids, and increase fetal brain CRABP (Grummer et al., 1993). Further investigation showed no difference in maternal serum retinol levels, an increase in CRBP in fetal brain and whole embryo, a decrease in RAR β only in whole embryo, and an increase in RAR γ in fetal brain in ethanol-treated rats compared to control (Grummer & Zachman, 1995). It has long been known that ethanol inhibits ADH-catalysed retinol oxidation (Julia et al., 1986; Leo et al., 1987; Van Thiel et al., 1974). This is a proposed mechanism for the manifestation of FAS (Deltour et al., 1996; Duester, 1991; Duester, 1994). It has recently been shown that high concentrations of ethanol decreased retinoic acid levels in mouse embryos at stages most sensitive to ethanol induced cranial defects. Class IV ADH mRNA was also expressed at the same time (Deltour et al., 1996). Although not direct evidence, these experiments support the hypothesis of ethanol induced alterations in retinoic acid production, availability, and/or homeostasis being a mechanism for FAS manifestation.

It is possible that AlDH 1 may also play a role in this scenario. The ability of acetaldehyde to inhibit aldehyde dehydrogenase-catalysed retinal oxidation has not been investigated. The ability of AlDH 1 to oxidise free and CRBP-bound retinal, as well as at least the 9-*cis* geometric isomer indicates a primary role for AlDH 1 in retinoic acid homeostasis. In addition, AlDH 1 is the only aldehyde dehydrogenase isoenzyme with the ability to oxidise retinal that has been found in humans, and it is also able to oxidise acetaldehyde. Further studies may reveal a role for AlDH 1 in FAS.

1.6 The Aldehyde Dehydrogenases

The aldehyde dehydrogenases (aldehyde:NAD⁺ oxidoreductase, EC 1.2.1.3), are a group of enzymes which catalyse the oxidation of various aliphatic and aromatic aldehydes to their corresponding acids (Goedde & Agarwal, 1990; Pietruszko, 1983; Sladek *et al.*, 1989). These enzymes are considered to play a general role in detoxification. Specifically, AlDHs have been shown to metabolise acetaldehyde derived from ethanol, toxic aldehydes from food and lipid peroxidation (Harrington *et al.*, 1987; Jakoby & Ziegler, 1990; Mitchell & Petersen, 1987; Parrilla *et al.*, 1974), retinoids, and aldehydes derived from biogenic amines and neurotransmitters (Ambroziak & Pietruszko, 1991; Ambroziak & Pietruszko, 1993; Yoshida *et al.*, 1992; Yoshida *et al.*, 1993). On the basis of various properties, the AlDHs can be classified into 3 main classes (Lindahl & Hempel, 1991). These properties include physicochemical and enzymatic properties, tissue/subcellular distribution, and sequence identities (Yoshida *et al.*, 1991).

Class 1 enzymes are homotetramers, composed of 54 kDa monomers. They are cytosolic, with a fairly broad substrate specificity. Included in this class are the major human liver cytosolic isoform (hAlDH 1), and the homologues from rat (rAlDH 1), sheep (sAlDH 1) (Figure 1.8), mouse (mAHD 2), and others (note: nomenclature for mouse aldehyde dehydrogenases differs from that of the other enzymes). Based upon multiple sequence alignment (Figure 5.1), the recently isolated retinal-oxidising dehydrogenases are likely to be classified as class 1 also. Class 2 enzymes are also homotetrameric with a similar monomer size to the class 1 enzymes. These are found in the mitochondria, and are thought to be the major isoform involved in acetaldehyde metabolism, with a K_m for acetaldehyde in the low micromolar range. A mutation in this



Figure 1.8: Ribbon Diagram of Sheep Aldehyde Dehydrogenase 1.

This diagram was constructed from coordinates determined by S. Moore (Massey University, NZ) using the program TURBO-FRODO (Cambillau *et al*, 1996). α-Helices are coloured red, β-strands blue, and loops yellow, while all-*trans* retinal is shown in green. All 4 subunits are shown (labelled A-D) each containing a retinal molecule.

isoform (Glu to Lys at position 487) is involved in the 'flushing syndrome', or adverse reaction to alcohol consumption suffered by a large percentage (\sim 50 %) of Asian people (Harada *et al.*, 1981; Ikawa *et al.*, 1983). The class 3 enzymes are dimeric, with a subunit molecular weight of 50 kDa, and include corneal AlDH and a tumour-associated AlDH.

The hypothesis for the general mechanism of AlDH 1 catalysed NAD⁺-dependent oxidation of aldehydes is supported by studying the recently solved structure of this enzyme (S. Moore, personal communication) (Figure 1.8). The proposed kinetic mechanism is nucleophilic attack by Cys 302 on the carbonyl of the aldehyde substrate to form an acyl-intermediate, followed by hydride transfer to Glu 399. Deacylation then occurs, producing the acid, NADH and regenerated enzyme (Sheikh *et al.*, 1997).

(mitochondrial AIDH 2)

1.7 The Aims of this Thesis

The primary aim of this dissertation was to investigate the role of the major cytosolic aldehyde dehydrogenase in the oxidation of retinal for retinoic acid biosynthesis. The analysis of AlDH 1 function was studied as follows:

- 1. The purification of sheep liver cytosolic AlDH 1, and characterisation of its interaction with all-*trans* and 9-*cis* retinal;
- 2. The expression and purification of human CRBP I, and binding of purified CRBP I to all-*trans*, but not 9-*cis* retinal;
- 3. The interaction of CRBP-retinal with purified sheep liver AlDH 1 and recombinant human AlDH 1;
- 4. Initial studies using cultured neuronal cells to investigate in vivo retinal metabolism;
- 5. An alignment of all known aldehyde dehydrogenase sequences with particular reference to those involved in retinal oxidation, and an overview of recent literature concerning aldehyde dehydrogenase enzymes;

In addition, a sheep liver enzyme initially identified as an aldehyde dehydrogenase with retinal oxidising ability, which was subsequently determined to be a novel alcohol dehydrogenase, was purified and partially characterised.

2. Chapter Two: Materials and Methods

2.1 Materials

Company	Full Name	Address
Sigma	Sigma Chemical Company	St. Louis, MO, USA
New England Biolabs	New England Biolabs	Beverly, MA, USA
Amersham	Amersham Life Science International plc	Little Chalfont, Buckinghamshire, UK
USB	United States Biochemical	Cleveland, Ohio, USA
Biorad	Biorad Laboratories	Auckland, NZ
BDH	British Drug Houses	Poole, Dorset, UK
Boehringer	Boehringer Mannheim	Auckland, NZ
Mannheim	Biochemica	
Gibco BRL	Life Technologies Gibco BRL	Auckland, NZ
Pharmacia	Amrad Pharmacia Biotech	Auckland, NZ

Table 2.1: Company Details

2.1.1 Retinoids

All-*trans* retinal, 9-*cis* retinal, all-*trans* retinoic acid, 9-*cis* retinoic acid, retinol acetate, and all-*trans* retinol were all purchased from Sigma.

2.1.2 Molecular Biological Materials

pT7-7-CRBP (Cellular retinol-binding protein) was a kind gift from Prof J Findlay, (Leeds University, UK). *E. coli* strain BL21 (DE3) was a generous gift from Dr J.S.Lott (Massey University, NZ). Restriction enzymes were purchased from New England Biolabs, Boehringer Mannheim, or Gibco BRL. Ready mixed Luria Broth and Agar were purchased from Gibco BRL. Ampicillin was purchased from Intermed Scientific. IPTG was purchased from Biorad.

2.1.3 Cell Culture Materials

Fetal Bovine Serum was purchased from Gibco BRL. RPM-1640 media was purchased from Sigma. SH-SY5Y cells were obtained by Dr M Grimes (Massey University, NZ) from M. Israel (Department of Neurology, University of California, San Francisco).

2.1.4 CRBP Purification

HiTrap® 1 ml and 5 ml chelating columns were purchased pre-packed from Pharmacia Biotech. Imidazole and NiCl₂ were purchased from BDH. Phenylmethylsulfonyl fluoride (PMSF) was purchased from Sigma.

2.1.5 Sheep Liver

Sheep livers were obtained fresh off the production line from Lakeview Abattoirs, Hokio Beach Road, Levin, NZ.

2.1.6 Recombinant Human AlDH 1

Recombinant human AIDH 1 protein was a gift from Jane Wyatt and Erin Loughnane (Massey University).

2.1.7 Gels and Electrophoretic Transfer

Isoelectric focusing gels - Ampholine® PAGplate pH 3.5-9.5 were purchased prepoured from Pharmacia Biotech. Tricine was purchased from Sigma. Hybond nitrocellulose membranes were purchased from Amersham. Immobilon P PVDF membrane was purchased from Millipore. Molecular size markers for SDS-PAGE were purchased from Sigma (SDS-7s). ECL detection chemicals and secondary antibodies were purchased from Amersham. Alternative secondary antibodies conjugated to alkaline phosphatase were purchased from Sigma. Primary antibodies to sheep AlDH 1 were a kind gift from Dr K Jones (Massey University). Phenazine methosulphate (PMS) and nitroblue tetrazolium (NBT) were purchased from Sigma. β -Nicotinamide adenine dinucleotide was purchased from Sigma.

2.1.8 Column Chromatography

All reagents used for HPLC were HPLC grade, filtered through 0.22 micron filters purchased from Millipore. Size standards for gel filtration were purchased from Sigma. β -Lactoglobulin, α -lactalbumin and lactoferrin were kind gifts from Dr G Norris (Massey University, NZ).

2.1.9 General Protein Purification Materials

Ultrafiltration membranes were purchased from Amicon. Ampholytes (pH 7-9) and Ultrodex resin for preparative isoelectric focusing were purchased from Pharmacia Biotech. DEAE Iontosorb ion exchange resin was a kind gift from Dr D Harding (Massey University, NZ). *p*-Hydroxyacetophenone affinity resin was prepared by G Freeman and Dr T Kitson (Massey University, NZ). PEG 8000 was purchased from Sigma. HTP Hydroxyapatite was purchased from Biorad.

2.1.10 Protein Digestion

Trypsin, chymotrypsin, cyanogen bromide, thermolysin and leucine aminopeptidase were purchased from Sigma. N-acylaminoacyl-peptide hydrolase was purchased from Boehringer Mannheim.

2.1.11 General Chemicals and Commonly Used Solutions

All general chemicals used were of the highest grade available, and were purchased from Sigma, BDH, Riedel-de Haën or USB. Pure water was produced using a Barnstead Nanopure filtration apparatus.

SDS-PAGE S	Solutions
------------	-----------

Stacking gel:	5 % (w/v) (37.5:1) acrylamide:bisacrylamide,
	125 mM Tris.Cl pH 6.8, 0.1 % (w/v) SDS
Resolving gel:	10-15 % (w/v) (37.5:1) acrylamide:bisacrylamide,
	375 mM Tris.Cl pH 8.8, 0.1 % (w/v) SDS
Running Buffer:	25 mM Tris, 200 mM glycine, 0.1 % (w/v) SDS, pH 8.5

Sample buffer (2×):	62.5 mM Tris.Cl pH 6.8, 3 % (w/v) SDS, 700 mM
	β -mercaptoethanol, 10 % (v/v) glycerol, 0.01 % (w/v)
	bromophenol blue
Coomassie Blue stain:	0.1 % coomassie blue R250, 30 % (v/v) methanol,
	10 % (v/v) acetic acid
Destain:	30% (v/v) methanol, $10%$ (v/v) acetic acid
Tricine gel solutions	
Stacking gel:	Final = 4 % T, 3 % C, from 48:1.5
	acrylamide:bisacrylamide, 0.7 M Tris.Cl, 0.1 % SDS
Spacer gel:	Final = 10 % T, 3 % C, from 48:1.5
	acrylamide:bisacrylamide, 1 M Tris.Cl, 0.1 % SDS
Resolving gel:	Final = 16.5 % T, 6 % C, from 46.5:3
	acrylamide:bisacrylamide, 1 M Tris.Cl, 0.1 % SDS
Anode buffer:	0.2 M Tris.Cl pH 8.9
Cathode buffer:	0.1 M Tris, 0.1 M tricine, 0.1 % SDS pH 8.25

Silver stain solutions

Farmer's Reagent:	12 mM sodium thiosulphate, 4.5 mM potassium
	ferricyanide, 4.5 mM sodium carbonate
Developer:	2.5 % (w/v) sodium carbonate, 0.02 % (w/v)
	formaldehyde

Isoelectric focusing activity assay solutions

Anode solution:	1 M H ₃ PO ₄
Cathode solution:	1 M HCl
For activity assay: 0.1 M so	odium tetrapyrophosphate buffer pH 9.0, 0.7 mM NAD^+ ,
2.5 μM PMS, 3 μM NBT, 10	0-200 mM acetaldehyde or ~50 μM retinal.

Phosphate Buffer for General Use

Mix 50 mM Na₂HPO₄ and 50 mM NaH₂PO₄ to give 1150 mM phosphate buffer pH 7.4.

Sheep liver AIDH protein purification buffers

Buffer 1:	4~mM potassium phosphate pH 7.4, 0.25 M sucrose
Buffer 2:	10 mM Bis-Tris, 5 mM NaCl, pH 6.2
Buffer 3:	As Buffer 2, except pH 5.5

2-53

Buffer 4:	25 mM sodium acetate, pH 4.8
Buffer 5:	50 mM potassium phosphate, 15 mM NaCl, pH 7.4
Buffer 6:	50 mM potassium phosphate pH 7.4, 3 M ammonium
	sulphate
Buffer 7:	50 mM potassium phosphate pH 7.4

In addition, all buffers contained 0.3 mM EDTA, and 0.5 mM DTT was added to each buffer immediately prior to using.

Electroblotting solutions

10 mM 3-[cyclohexylamino]-1-propane sulfonic acid (CAPS), 10 % (v/v) methanol, pH 11.

Phosphate buffered saline (PBS)

8 g/l NaCl, 0.2 g/l KCl, 1.44 g/l Na₂HPO₄, 0.24 g/l KH₂PO₄, pH 7.2-7.4

TAE Electrophoresis buffer

40 mM Tris-acetate, 1 mM EDTA pH 8

CRBP purification buffers

Load and wash buffer:	$20 \text{ mM} \text{ Na}_2\text{HPO}_4, 0.5 \text{ M} \text{ NaCl}, 0.1 \text{ M} \text{ imidazole}, \text{pH} 7.2$
Elution buffer:	As above, except 0.3 M imidazole.
Charging buffer:	0.1 M NiCl ₂

Western blotting solutions

Quenching solution:	2 % BSA in PBS
Washing solution:	0.05 % Tween in PBS
Primary antibody solution:	1:2000 anti-AlDH antibody, in 1 % BSA in PBS
Secondary antibody:	Appropriate dilution of secondary antibody in 1 $\%$ BSA in
	PBS
Substrate solution (1):	30 mg 4-chloro-1-naphthol in 10 ml ice-cold methanol,
	supplemented immediately before use with 30 μ l
	30 % H ₂ O ₂ in 50 ml PBS
Substrate solution (2):	As per ECL directions (Amersham).
2.2 Methods

2.2.1 Maintenance of Retinoids

Retinoids were dissolved in dimethyl formamide or ethanol and stored at -70°C in the dark. All retinoid manipulations were performed in dim light. The concentration of retinoid solutions was confirmed by using absorbance at wavelengths with calculated ε values.

Table 2.2: Light Absorbances of Selected Retinoids

Compound	Solvent	λmax	3
All- <i>trans</i> retinal	Ethanol	383	42 880
9-Cis retinal	Ethanol	373	36 100

(Furr et al., 1994)

2.2.2 Spectrophotometry

Spectrophotometric assays were carried out on a Varian Cary Double Beam UV-vis spectrophotometer. All assays were carried out in 50 mM phosphate buffer, pH 7.4 at 25 °C unless otherwise stated.

For routine cytosolic AlDH assays, the production of NADH was followed at 340 nm, using 3 mM NAD⁺, 500 μ M acetaldehyde, and 2-100 μ l of enzyme containing fraction. To check that the identified activity was due to cytosolic and not mitochondrial AlDH 1, disulfiram (final concentration ~33 μ M) was added 1-2 minutes into the assay. The addition of disulfiram immediately inactivates cytosolic but not mitochondrial AlDH.

For assaying the esterase activity of AlDH 1, the production of *p*-nitrophenoxide at 400 nm was monitored spectrophotometrically. The concentrations of the substrate *p*-nitrophenyl pivalate ranged from 12.5 μ M to 100 μ M, added in a small volume of ethanol. Assays were performed at 25°C using 50 mM phosphate buffer, pH 7.4.

For looking at the NAD⁺-dependent oxidation of retinal by sAlDH over a range of pH values, the disappearance of retinal at 400 nm was followed. For these assays, all-*trans* retinal was dissolved in dimethylformamide to give a concentration of 10 mM, then diluted to 1 mM in 70 % aqueous methanol. When retinal was included in the test cuvette, an equal amount of retinal was included in the reference cell. Assays were run for 5 minutes at 25°C. For pH values of 4.5-5.5, 50 mM acetic acid/sodium acetate buffer was used, for pH 5.9-8.0, 50 mM sodium phosphate buffer was used, and for pH values 8.5-11.0, sodium pyrophosphate buffer was used.

2.2.3 Fluorimetry

Fluorimetric analyses were carried out using a Perkin-Elmer LS50B luminescence spectrophotometer. All buffers were filtered through a 0.22 micron filter prior to use. Four-sided quartz cuvettes were acid washed prior to use in concentrated nitric acid.

For all kinetic assays following the production of NADH, the excitation wavelength was 340 nm and the emission wavelength was 457 nm. Slit widths for excitation and emission were set at 15 nm. A blank assay was first performed as a control. This assay consisted of 0.2 mM NAD⁺, 50 mM Na₂HPO₄/NaH₂PO₄ pH 7.4, and 1-5 μ g sheep or human AlDH. The reaction is initiated with the addition of 20 μ l of solvent, either DMF or ethanol. This reaction was monitored at the stated wavelengths. This blank assay usually had a very small rate, which was subtracted from the test rate. Additional blank assays were performed with the same components, except one omitted NAD⁺, and the other was done without AlDH. These blanks consistently showed a zero rate. The test assay was performed in exactly the same manner, except the appropriate concentration of retinal in ethanol or DMF was added to initiate the reaction.

2.2.3.1 Assays with CRBP

A blank run was performed with all assay components - buffer, enzyme, CRBP, and NAD^+ using solvent (ethanol or DMF) only. This blank rate was subtracted from the test rate which contained retinal in the same solvent. For the test assay, CRBP and retinal were pre-incubated for 15 minutes in the fluorimeter chamber prior to initiating the reaction by adding NAD⁺ and enzyme. After subtraction of the blank rate, the steady state rates were calculated, and averages of duplicate or triplicate assays determined.

Kinetic constants were calculated using the program ENZFITTER (Leatherbarrow, 1987).

2.2.4 HPLC Based Assay - Extraction and Quantification

The assay components for this method were similar to those described in section 2.2.3. The total volume was 500 μ l. Assays were carried out in foil-wrapped glass vials at 37°C. Following incubation for 10-20 minutes, retinoids were extracted and measured by HPLC. See Chapter Three for details on methods tested.

The separation and detection of standard and extracted retinoids was performed using a C18 column (Pharmacia) attached to a Spectra Physics HPLC pump and gradient mixer, and a Spectra Physics SP8490 UV-vis detection unit. A Spectra-Physics integrator was used to measure peak area and construct standard curves.

2.2.5 Recombinant DNA Techniques for CRBP Production

2.2.5.1 Transformation of E. coli

Plasmid vectors were introduced into *E. coli* cells which had previously been rendered competent by the method of Inoue *et al.*, (1990). 100 μ l of competent cell suspension was incubated with 1 μ l of plasmid DNA in solution for 10 minutes on ice. The transformed cells were then spread onto LB-agar plates (1.5 % (w/v) agar in LB broth) with ampicillin selection (50 μ g/ml), and incubated at 37°C for 12-16 hours.

2.2.5.2 Large-Scale Preparation of Plasmid DNA

Large-scale preparation of plasmid DNA was achieved using the Wizard[™] Plus Maxipreps system as described in the manufacturers instructions (Promega, 1996).

2.2.5.3 Digestion of Plasmid DNA with Cla I

Plasmid (~1 µg) was digested using 1-5 units of the restriction endonuclease Cla I in a total volume of 20 µl of 1× enzyme buffer, as recommended by the manufacturer, for 2 hours at 37°C.

2.2.5.4 TAE-Agarose Gel Electrophoresis of DNA

Digested and undigested DNA were analysed using agarose gel electrophoresis in $1 \times TAE$ buffer at 3-5 V/cm, essentially as described in Sambrook *et al.*, (1989) using 0.2 % bromophenol blue in 50 % glycerol as $6 \times$ loading buffer. Agarose gels (0.9 %) containing 0.5 µg/ml ethidium bromide were used, which enabled the DNA to be visualised by illumination with ultraviolet light (302 nm). Sizes were estimated by comparison with known standards, usually *Hin* dIII-digested λ phage DNA (Gibco BRL), loaded to give a total of 1.5 µg of DNA per lane.

2.2.5.5 DNA Sequencing

DNA sequencing was carried out using the dideoxy chain-termination method (Sanger *et al.*, 1977) on an automated Perkin-Elmer ABI Prism 377 DNA sequencer. Double stranded plasmid DNA was used as template.

2.2.6 Cytosolic Aldehyde Dehydrogenase Purification and Analysis

2.2.6.1 Purification of Sheep Liver AIDH 1

1 kg of fresh sheep liver was chopped into small pieces and homogenised in a Waring blender with 2 l of buffer (sucrose phosphate). The resulting homogenate was spun at 9000 rpm for 15 minutes in a Sorvall GSA centrifuge head. The supernatant was decanted and spun immediately in Sorvall GS3 heads at 13 000 rpm for 30 minutes to pellet the mitochondria. Polyethylene glycol (MW 8000) was then added to the supernatant to a final concentration of 10 % (w/v) over 20 minutes. The solution was left to equilibrate for a further 15 minutes, and then spun at 9000 rpm for 30 minutes. To the resulting supernatant, further PEG 8000 was added over 20 minutes, to give a final concentration of 20 % (w/v). This was left to equilibrate for a further 15 minutes, and then spun at 9000 rpm for 30 minutes. The solution was left to equilibrate for a further 16 minutes. The supernatant from this spin was discarded. The pellet was then redissolved in a small amount of buffer 2 (100-200 ml).

This solution was then applied to a DEAE-Iontosorb ion exchange column, preequilibrated in buffer 2 (pH 6.2). After the red hemoglobin containing solution had washed through the column, the column was washed overnight with 5 1 of buffer 3 (pH 5.5), to remove any class 2 AlDH (pI ~ 5.9) which may still be present. AlDH 1 bound to the column was then eluted by passing buffer 4 (pH 4.8) through the column, and collecting 5-10 ml fractions. The fractions were assayed for protein concentration and AlDH activity (see below), and active fractions pooled. The pooled fractions were then dialysed overnight through 2 or 3 changes of buffer into buffer 5.

The pooled dialysed fractions were then loaded onto an affinity resin pre-equilibrated with buffer 5. The affinity resin consists of p-hydroxyacetophenone (an inhibitor of the esterase activity of AlDH 1) bound to an inert matrix. The column was then washed extensively with buffer 5 to remove all non-specific proteins, checked by monitoring the absorbance at 280 nm of the column elute. A gradient of 0-200 mM p-hydroxyacetophenone was then run using either a manual gradient mixer, or an Econosystem automated gradient mixer. 5-10 ml fractions were collected and assayed for protein concentration and AlDH 1 activity, and the active fractions pooled. The pooled fractions were dialysed against 3 M ammonium sulphate overnight to precipitate the protein. The precipitate was then redissolved in a small amount of buffer 5, and final concentration and activity readings taken.

2.2.6.2 Protein and Activity Assays for Protein Purification

Approximate protein concentrations were determined by measuring the absorbance of an appropriately diluted solution at 280 nm on a Varian Cary 1 double beam UV-vis spectrophotometer. For pure AlDH this was then compared with the known A_{280} for AlDH of 1.13 for a 1 mg/ml solution. The activity was measured by following the production of NADH at 340 nm using an appropriately diluted solution. Components of the assay were 50 mM phosphate buffer pH 7.4, 3 mM NAD⁺, 2-100 µl crude enzyme solution, and 500 µM acetaldehyde. The complete removal of AlDH 2 was checked for by adding disulfiram (final concentration 33 µM), which inactivates AlDH 1 immediately, but not AlDH 2. A residual activity of about 3 % in the presence of disulfiram is indicative of cytosolic AlDH 1 free of mitochondrial AlDH 2.

2.2.7 Isoelectric Focusing (IEF) Activity Assay

IEF of proteins was performed on an LKB2117 Multiphor horizontal gel electrophoresis box connected to an LKB2103 power supply and an LKB 2209 multitemperature thermostatic circulator. Pure protein (~100 μ g) was loaded onto a pre-poured IEF gel (range 3.5-9.5), and electrophoresed for 1.5 hours at 30 W constant power, and at a constant temperature of 8-10°C. The focused gel was then rinsed 3 times in sodium pyrophosphate buffer, and incubated with NBT, PMS, NAD^+ , and substrate (concentrations as described in section 2.1.10), in the dark, until bands of activity could be visualised. Once bands of activity were fully developed, the stain was fixed in 5 % acetic acid.

2.2.7.1 Transferral of Proteins Separated by IEF to Nitrocellulose

Proteins separated by IEF as described (section 2.2.7) were transferred by capillary action to nitrocellulose membrane. Nitrocellulose membrane (Hybond-N, Amersham) was pre-wetted with sodium pyrophosphate buffer, and placed onto the electrophoresed gel. Whatman 3MM paper was placed either side of the gel and membrane sandwich. Paper towels (5-10 cm high) and weights (500 g) were placed on top of the sandwich and 3MM paper (see diagram below). Proteins were transferred to the membrane over 2-3 hours, and then stained for activity as described (section 2.2.7)



Figure 2.1: Transferral of Proteins Separated by IEF to Nitrocellulose

2.2.8 SDS-PAGE

Proteins were separated both analytically and preparatively using sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE), using a Hoeffer Mighty Small vertical slab minigel apparatus. Gels were poured and run essentially as described by

Yuen *et al.*, (1986). Routinely, 15 % (w/v) 37.5:1 acrylamide:bisacrylamide resolving gels and 5 % (w/v) 19:1 acrylamide:bisacrylamide stacking gels were used.

After electrophoresis, gels were stained using either Coomassie blue or silver. Gels were routinely fixed and stained by soaking in 2 % (w/v) Coomassie Brilliant Blue R-250 dissolved in 10 % (v/v) acetic acid and 30 % (v/v) methanol, and destained by soaking in several changes of 10 % (v/v) acetic acid and 30 % methanol. Molecular masses of proteins were estimated by comparison with known standards. SDS-7 markers (Sigma) were routinely used in 15 % acrylamide gels.

Silver staining was carried out as follows:

After electrophoresis, the gel was soaked in 30 % (v/v) ethanol, 10 % (v/v) acetic acid for 3 hours. If the gel had been previously stained with Coomassie blue, it was then soaked in 10 % (v/v) ethanol for 3 hours. The gel was then soaked in the following solutions:

10 % (v/v) ethanol, 5 minutes Milli-Q H₂O, 3×5 minutes Farmer's Reagent, 1 minute Milli-Q H₂O, 3×5 minutes 0.1 % (w/v) AgNO₃, 30 minutes Milli-Q H₂O, 3×5 minutes Developer, brief wash then soak until staining of desired intensity 1 % acetic acid, 5 minutes Farmer's Reagent 10-30 seconds to reduce background if necessary Milli-Q H₂O, 3×5 minutes

Gels stained by either method were either air dried between cellulose acetate membranes (BioRad) or vacuum dried (80°C, 2 hours) onto 3MM paper (Whatman) after soaking in 2 % (v/v) glycerol in 30 % (w/v) methanol. The inclusion of glycerol reduces gel shatter on drying.

2.2.9 Electroblotting

Proteins were electroblotted from acrylamide gels as follows:

After electrophoresis, the gel was soaked in transfer buffer (10 mM CAPS pH 11, 10 % (v/v) methanol) for 5 minutes, before being assembled in an electroblot sandwich. The sandwich consisted of the gel and membrane (PVDF or nitrocellulose) between sheets of 3MM paper soaked in transfer buffer as illustrated below:



Proteins were transferred at a constant current of 250 mA for 2 hours.

2.2.10 Electroelution of Proteins from Acrylamide Gels

A preparative SDS gel was prepared using 1.5 mm spacers and a one-well well former. Gels were prepared as described (2.2.8), and were generally 12 % (w/v). After running, the gel was stained and destained as usual. The band of interest was then excised from the gel and placed in dialysis tubing containing a small volume of 50 mM Tris.Cl pH 8, containing 0.1 % (w/v) SDS. Electroelution of the protein from the gel was carried out in the same buffer at a constant current of 50 mA for 48 hours at 4°C. After electroelution, the polarity of the current was briefly reversed to remove any protein that may have adhered to the dialysis tubing. The eluted protein was concentrated under vacuum.

2.2.11 Protein Cleavage

Methods used for protein cleavage are all adapted from methods outlined in Allen, (1981), Smith, (1994a) and Smith, (1994b).

2.2.11.1 Cyanogen Bromide (Methionine-X)

Pure protein was dissolved in water to a final concentration of 1-5 mg/ml. One volume of ammonium bicarbonate (0.4 M, pH 8.4), and 1-5 % (v/v) β -mercaptoethanol were added. After passing nitrogen gas over the solution, the solution was incubated at room temperature for 18 hours. The solution was then dried down under vacuum, and redissolved in formic acid to 1-5 mg/ml. Water was added to bring the final concentration of formic acid to 70 % (v/v), and excess crystalline cyanogen bromide added (1-2 crystals). This was incubated at room temperature for 24 hours, then dried down under vacuum, or lyophilised.

2.2.11.2 Dilute Acid Cleavage (Aspartate-X)

The protein was dissolved to 1-2 mg/ml in 30 mM HCl in a glass hydrolysis tube. The hydrolysis tube was sealed under vacuum and incubated at 108°C for 2 hours. The tube was then opened, the sample diluted with water, and the solution lyophilised.

2.2.11.3 Trypsin (Lysine-X and Arginine-X)

The protein was diluted to ~ 2 mg/ml in water. An equal volume of ammonium bicarbonate solution (0.4 M pH 8.4) was added. A number of tests showed good cleavage to occur after a 5 minute incubation at room temperature. Digestion was terminated by adding 1 mM PMSF.

2.2.11.4 Thermolysin (X-Hydrophobic)

To 20 μ l of ~2-5 mg/ml protein was added 5 μ l of ammonium bicarbonate solution (0.4 M pH 8.4), 11 μ l H₂O, and 4 μ l 1 mM CaCl₂. A small amount of solid thermolysin was added and the solution incubated at 45°C for 1 hour. 1 mM EDTA was added to stop the reaction, and the solution was then dried down under vacuum.

2.2.11.5 Leucine Aminopeptidase (N-terminal Amino Acids)

To 20 μ l of 2-5 mg/ml protein was added 80 μ l 0.1 M Tris.Cl containing 2.5 mM MgCl₂. The solution was incubated at 37°C for 10 minutes, frozen for 30 minutes at -70°C, and dried down under vacuum.

2.2.11.6 Separation of Peptides Generated by Cleavage

Two methods were commonly used to separate peptides generated by cleavage: tricine gels, and HPLC based separation.

3-layer tricine gels were constructed as outlined in section 2.1.11, from the method of Schagger & von Jagow, (1987). After running the peptide mixture through the 3-layer gel, the gel was electroblotted onto PVDF membrane for 2 hours at 250 mA in 10 mM CAPS. The PVDF membrane was then briefly stained with Coomassie blue solution, and destained in methanol until bands could be seen. Bands of interest were cut out, and frozen at -20°C until sequencing could be carried out.

To separate mixtures of peptides by HPLC, the mixture was applied to a C18 column pre-equilibrated in 0.1 % TFA in water. A gradient was then run over 20-30 minutes to 100 % acetonitrile, separated peptides were trapped as they eluted, and sequenced.

2.2.12 Protein Sequencing

Pure proteins and peptides generated by protein cleavage were sequenced on an Applied Biosystems 476A Automated Protein Sequencer, using sequential Edman degradation. Sequencing was performed by Ms. Jo Mudford

2.2.13 Amino Acid Analysis of Proteins

The amino acid composition of pure protein was determined by total acid hydrolysis in 6 M HCl under vacuum, and then applying the amino acid mixture to a Pharmacia LKB Alpha Plus Amino Acid Analyser. Amino acid analysis was carried out by Ms. Debbie Frumau.

2.2.14 N-Terminal Deacylation

The method used to deacylate purified protein with a blocked N-terminus was that described by Hirano *et al.*, (1992). Essentially, this method involves cleaving the purified protein and blocking the generated internal N-termini with phenylisothiocyanate. The original N-terminal residue is then removed with the enzyme N-acylaminoacyl-peptide hydrolase, and the mixture applied to the automated sequencer as usual. The

only free N-terminus should be the actual N-terminus. The cleavage is necessary as the deacylating enzyme does not work efficiently on peptides longer than 10-20 residues.

Pure protein was run on an SDS gel, and electroblotted onto PVDF membrane as previously described. The stained band was cut out and treated with 0.5 % (w/v)polyvinylpyrrolidone in 100 mM acetic acid at 37°C for 30 minutes to block the unbound protein-binding site. The membrane was then washed with water 10 times, and digested with trypsin in 0.1 M ammonium bicarbonate pH 8 containing 10 % (v/v) acetonitrile at 37°C for 24 hours with shaking. The digestion buffer containing tryptic peptides was then transferred to an microfuge tube, and the membrane washed by vortexing with 100 µl H₂O and combined with the peptides. The buffer was evaporated to dryness, and 100 μ l of 50 % (v/v) pyridine and 10 μ l phenylisothiocyanate were added to bind with free but not blocked N-terminal amino acids of the tryptic peptides. The mixture was purged with nitrogen gas for 20 seconds, and incubated for 1 hour at 60°C. One ml of benzene/ethylacetate (1:1 (v/v)), was added, and the solution vortexed and centrifuged at $3000 \times g$ for 1 minute. The supernatant containing reaction by-products and excess reagents was discarded, and this washing procedure was repeated three times. The sample was then evaporated to dryness under vacuum. Pre-made performic acid (100 ml) was then added to the dried sample and kept at 0°C for 1 hour, then evaporated to dryness under vacuum. Filtered water (100 µl) was added and the drying was repeated. A solution of the peptides in 100 µl of 0.2 M phosphate buffer pH 7.2 containing 1 mM DTT and one unit of N-acylaminoacyl-peptide hydrolase dissolved in 50 µl of the same buffer was then incubated at 37°C for 12 hours to remove the N^{α} -acetylated amino acid. The sample was then sequenced as usual. The only peptide with a free N-terminal available for sequencing is the N-terminal peptide.

2.2.15 Western Blotting

The protein sample was run on an SDS gel as described. The proteins were then transferred to nitrocellulose membrane by electrophoretic transfer as described previously. After transfer, the filter was placed in a suitable container and incubated in 2 % BSA in PBS for 60 minutes at room temperature with shaking. The primary (anti-AlDH 1) antibody was then appropriately diluted in 1 % BSA in PBS, and the filter incubated in this solution overnight at room temperature. The filter was washed for 3×20 minutes in 0.05 % Tween in PBS, and then incubated with appropriately diluted

secondary antibody in 1 % BSA in PBS for 2 hours at room temperature with shaking. The filter was washed as before, and then the substrate reaction was performed as detailed previously either with ECL reagents, or 4-chloro-1-naphthol. When ECL reagents were used, bands were visualised by exposure to X-ray film.

2.2.16 Cell Culture

Human neuroblastoma SH-SY5Y cells were grown in RPMI 1640 medium (5 % CO₂; 37° C), supplemented with 10 % (v/v) fetal calf serum. Fresh media was provided 3-4 times per week. Cells reached confluence at 7-10 days, and at this stage were passed. Passage of cells involved triturating cells with PBS (see Materials) and pelleting cells at 100 × g for 3 minutes. After pelleting, cells were diluted 1:4 with RPMI 1640 medium and replated, or used in subsequent experiments.

2.2.16.1 Preparation of Cultured Cells for Western Blotting

After trituration, cells suspended in PBS containing protease inhibitors (see Materials) were subjected to 3 rapid cycles of freeze-thawing in liquid nitrogen to crack open the cells. The cell debris was then spun down at 10 000 \times g for 15 minutes. The supernatant was treated as described in section 2.15.

2.2.16.2 Differentiation of Cultured Cells

Subconfluent monolayers of SH-SY5Y cells were induced to differentiate by the addition of 1 μ M retinoic acid or retinal in ethanol, such that the total ethanol did not exceed 0.05 % (v/v). As far as possible, manipulations involving retinoids were performed under dim light. Cells were plated onto plastic culture plates containing a number of sterile glass microscope coverslips. At pre-determined timepoints, a coverslip was removed from the plate, placed upside down onto a sterile microscope slide with a drop of glycerol, and viewed on a Zeiss Axioskop microscope. In addition, cells were stained with a DNA staining agent, Hoechst 33342 as described below, which was used to facilitate cell counting. Cells were viewed by fluorescence of stained nuclei, and by Phase Contrast microscopy to look at cell processes. A differentiated cell was one with a neurite of length at least 1.5 × that of the cell body. The total percentage of differentiated cells was calculated by examining 100-300 random cells for differentiation characteristics. Randomness was ensured by selecting a number of cell populations in different places on the slide without looking down the microscope. Cell counts were performed using an Improved Neubauer Cytometer.

2.2.16.3 DNA Staining

Hoechst 33342 working solution was prepared by diluting the stock solution (1 mg/ml) to 10 µg/ml in PBS. 100 µl of this solution was applied to cells grown on a coverslip, and left to incubate for 30 minutes at room temperature. The coverslip was then turned upside down onto a sterile microscope slide, and viewed as described.

3. Chapter Three: Purification of, and Kinetic Studies Using, Cytosolic Aldehyde Dehydrogenase

3.1 Introduction and Aims

The primary aim of the work in this section is to ascertain whether cytosolic aldehyde dehydrogenase (AlDH 1) from sheep and human liver showed kinetic parameters for retinal oxidation which would suggest that the enzyme could metabolise this substrate at the concentrations present *in vivo*. In addition, the comparison of AlDH 1 from the two species was designed to investigate the validity of the sheep enzyme as a useful model for the human enzyme in the absence of recombinant human protein. The sheep enzyme is a much closer to human AlDH 1 in sequence and in previously determined kinetic parameters than the commonly used rodent enzymes.

To assess the ability of AlDH 1 to oxidise retinaldehyde, it was first necessary to isolate adequate amounts of pure enzyme to use in kinetic studies. The raw material used was fresh sheep liver. Pure sheep liver AlDH 1 (sAlDH 1) was characterised in terms of:

- 1) pH optimum of NAD⁺-dependent retinal oxidation;
- 2) kinetic parameters of NAD⁺-dependent all-*trans* retinal oxidation;
- 3) kinetic parameters of NAD⁺-dependent 9-*cis* retinal oxidation;
- 4) the inhibition of the esterase activity of AlDH 1 by all-*trans* retinal and citral.

In addition, a quantity of recombinant human AlDH 1 (hAlDH 1) which had been expressed and purified by J. Wyatt and E. Loughnane (Massey University) was obtained. Kinetic parameters for NAD^{*}-dependent retinal oxidation were also determined for this enzyme, and compared with those for sAlDH 1.

The accurate measurement of kinetic constants for the dehydrogenation of retinal is difficult. Reported V_{max} values are low, and K_m values extremely low, which means to obtain a precise K_m by a conventional method such as a 6×6 plot, exceedingly low rates would need to be accurately measured. There are a number of methods suitable for determining kinetic constants. A spectrophotometric method is commonly used if one assay component has a distinct absorption maximum. A spectrophotometric method is potentially quick, inexpensive and relatively easy; however in some cases it may not afford the sensitivity required to accurately measure kinetic constants. In

spectrophotometry, the signal is measured indirectly by measuring the difference between the incident and transmitted beams, so the small decrease in intensity of a very large signal measured in spectrophotometry results in a large loss of sensitivity (Guilbault, 1990). The sensitivity of a spectrophotometer is not adequate to measure the required levels of retinoids in this case.

A fluorimetric-based method may be used to increase sensitivity by up to 1000-fold (Guilbault, 1990). This is because in fluorescence, the emitted radiation is measured directly and can be increased or decreased by altering the intensity of excitation. Fluorimetry can not always be used in place of a spectrophotometric method though, because only 10 % of compounds that absorb light emit absorbed energy as fluoresence (Guilbault, 1990). If the assay to be studied does contain a fluorescent compound, because of the relatively few fluorescent compounds that exist, the likelihood of the assay system containing another fluorescent compound is low. So, fluorimetry is potentially a sensitive and specific means of obtaining kinetic data.

A third commonly used method to be considered is using a direct quantification of enzymatically derived products by high pressure liquid chromatography (HPLC) (Bridges, 1990; Chen *et al.*, 1994; Landers, 1990; Landers & Olson, 1988; Napoli, 1986; Napoli & Race, 1990). This method is highly sensitive and has the advantage of measuring the substance of interest directly, rather than by following a cofactor or coupled reaction component as often happens when using spectrophotometry or fluorimetry.

3.2 Results and Discussion

3.2.1 Purification of sAIDH

The purification of sAlDH 1 from approximately 1 kg of fresh sheep liver was carried out as described in Materials and Methods (2.2.6.1). A typical enzyme preparation gave up to a 2500-fold purification as measured by specific activity. There are two major forms of aldehyde dehydrogenase present in liver; mitochondrial (AlDH 2) and cytosolic (AlDH 1). Both forms have activity with acetaldehyde as a substrate, and are of almost identical molecular weight (MacGibbon *et al.*, 1979). These isoforms will run at the same place on sodium-dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE). Therefore, the purity of sAlDH 1 was not only assessed by SDS-PAGE, but

also by isoelectric focusing (IEF), and the use of a specific inhibitor of cytosolic AlDH 1, disulfiram. The isoelectric focusing which was carried out routinely was vital to ensure no other isoforms of AlDH, which may appear as a single band on SDS-PAGE, were co-purified along with AlDH 1. The isoelectric points of AlDH 2 and AlDH 1 (~5.9 and ~5.2 respectively) allow the two isoforms to be separated and identified easily using IEF. The use of disulfiram, which rapidly inactivates cytosolic AlDH 1, but only relatively slowly inactivates AlDH 2, also aided in checking purity. Throughout the purification procedure, AlDH 1 activity is followed using a spectroscopic assay (section 2.2.2). The

addition of disulfiram to this assay as described (section 2.2.2) allows activity due to

cytosolic and that from mitochondrial AlDH to be distinguished.

Data from one purification is shown in Table 3.1. The final specific activity in this instance was lower than usual, due primarily to inactivation of the enzyme rather than the presence of other protein impurities. Other preparations typically yielded pure AlDH with a specific activity of between 50 and 150 nmoles/min/mg. The isolated protein was shown to be essentially pure by SDS-PAGE and IEF (Figure 3.1). The minor bands apparent with heavy loading of IEF and SDS gels did not show aldehyde dehydrogenase activity, but may have been proteolytic products due to traces of protease activity in the preparation, as their presence tended to increase with time of storage at 4°C.

3.2.2 Method Development for Kinetic Characterisation of Purified sAIDH

3.2.2.1 Development of an HPLC-Based Kinetic Assay

A frequently used method of measuring kinetic constants involving retinoids is a time based assay where the enzymatic reaction products are extracted and directly measured by HPLC (Bridges, 1990; Chen *et al.*, 1994; Landers, 1990; Landers & Olson, 1988; Napoli, 1986; Napoli & Race, 1990). There are a number of advantages of using an such an assay, which include the direct measuring of enzymatic reaction product, the availability of HPLC equipment and columns, low limits of detection, and the ability to separate most retinoids and their isomers. The separation of retinoids by HPLC may be facilitated by either adsorption chromatography ('normal phase') where less polar components elute earlier than more polar compounds, or partition chromatography ('reversed phase'), where polar compounds elute early. Reversed phase chromatography is able to separate different classes of retinoid quickly and easily, while normal phase gives greater separation of isomers from the same class (Furr *et al.*, 1994).

Fraction	Volume (ml)	Total Protein (mg)	Protein Concentration (mg/ml)	Activity (µmole/min/ ml)	Specific Activity (nmole/min/ mg)
Homogenate 1	2030	287854	141.8	0.013	0.063
Homogenate 2	1770	200647	113.4	0.015	0.070
PEG 1 Supernatant	1710	44460	26.0	0.020	0.096

56.5

13.3

15.9

0.111

0.085

0.750

1.96

6.4

47.2

PEG 2 Pellet

Pooled

DEAE Pure Protein 298

279

20

16837

3711

318

Table 3.1: Typical Purification of Sheep Liver Aldehyde Dehydrogenase

An important part of this method is the extraction of retinoids from the aqueous assay medium. There are a number of published methods for extraction, all of which involve extraction of retinoids using an organic solvent and phase separation (Bridges, 1990; Chen et al., 1995; Eckhoff & Nau, 1990; Furr et al., 1994; Landers, 1990; McCormick et al., 1978; Napoli & Race, 1990; Noll, 1996). Reports of extraction efficiency vary from method to method. It is important to include an internal standard in the assay, which is co-extracted with the products of the reaction, and can be used to determine the extraction efficiency in each case. This allows comparisons to be made between duplicates or triplicates of a single experiment, and comparisons between different experiments. To quantify the reaction products, a standard curve can be constructed for each experiment using standard amounts of the retinoids of interest (including the internal standard), and the amounts extracted from the assays compared to the standard curve. After adjustments have been made for efficiency of extraction, accurate results should be obtained. The adjustment for extraction efficiency is reliant on the extraction of each of the retinoids of interest remaining constant relative to the internal standard. If this is not the case, no adjustments can be made, and hence no accurate estimate of the amount of extracted retinoid calculated.



Figure 3.1: SDS-PAGE and IEF Gel of Pure sAIDH 1

Gel A: IEF gel, and blot of purified sAlDH 1. The gel and blot were run as described in section 2.2.7.

- Lane 1,2: IEF gel of purified sAlDH 1
- Lane 3-5: Blot of IEF gel of purified sAlDH1

<u>Gel B</u>: SDS-gel of purified sAlDH 1 stained with Coomassie blue as described in section 2.2.8

Lane 1: Biorad molecular weight markers Lane 2-4: Purified sAlDH 1 Considerable time and effort was invested in developing a direct HPLC-based assay system, based on published methods. The reversed-phase HPLC separation and detection of retinoids was successfully set up as detailed below (3.2.2.2), and an example of a trace for the separation of standard retinoids is shown in Figure 3.2. However, after extensive trials, none of the extraction procedures used proved to be reproducibly consistent, making results impossible to determine with any acceptable accuracy.

3.2.2.2 Reversed-Phase HPLC Retinoid Separation and Detection

It was decided to develop a system for the measurement of retinoids using reversedphase HPLC. The system used was a Waters HPLC gradient and pump, connected to a Spectrum integrator, and a Waters UV-Vis detection apparatus. The column used was C18 substituted silica matrix, dimensions 250 mm \times 4.6 mm, with a 10 μ M pore size (VYDAC 218TP). A number of different mobile phases were tested, including:

1) Acetonitrile: H_2O ratios between 70:30, and 95:5, with 0.1 M sodium acetate (isocratic).

2) Buffer A - 10 % v/v tetrahydrofuran, 52 % v/v acetonitrile, 50 mM ammonium acetate pH 7; and Buffer B - 100% acetonitrile. Run isocratically at 60 % A and 40 % B (Chen *et al.*, 1995).

3) Methanol:30 mM ammonium phosphate pH 7.4, ratios between 80:20, and 95:5 (Martini & Murray, 1994).

4) Methanol:0.01 M ammonium acetate ratios between 80:20 and 90:10 (Cavazzini *et al.*, 1996).

The detection of retinoids was carried out spectrophotometrically at 350 nm. Flow rates varied between 1.0-1.5 ml/min. While all of these solvent systems gave adequate results, mobile phase 3 using a ratio of 87.5:12.5 (methanol:30 mM ammonium phosphate pH 7.4) was used for the subsequent analysis.

3.2.2.3 Incubation of Enzyme Reaction Mixtures

Assays were carried out in 20 mM Hepes, pH 8.5 with 150 mM KCl, 1 mM EDTA, 1 mM DTT, and 1-5 μ g enzyme in a total volume of 500 μ l in glass vials. Retinoids were dissolved in DMF or ethanol and added to the reaction mixture so solvent concentration did not exceed 0.4 %. 2 mM NAD⁺ was added to initiate the reaction. All reactions were done under dim yellow light. Incubations were carried out at 37°C for 20 min. Just prior to extraction, an appropriate amount of the internal standard (retinol acetate) was added.



Figure 3.2: HPLC Trace of Authentic Retinoids

Known amounts of all-*trans* retinoic acid, all-*trans* retinol acetate and all-*trans* retinal were applied to a C18 column as described in section 3.2.2.2

3.2.2.4 The Extraction of Retinoids from Assay Medium

The first method of extraction involved two sequential extractions, which separated neutral retinoids from retinoic acid (Napoli & Race, 1990). The reaction was firstly quenched with 0.025 M KOH/ethanol to raise the pH to 12. Neutral retinoids were then extracted with two aliquots of 2.5 ml hexane. After each addition of 2.5 ml hexane, the solution was vortexed for 5 min, then the organic layer aspirated off and discarded. The pH was lowered to below 2 with HCl, and retinoic acid was extracted with two aliquots of 2.5 ml hexane. The samples were vortexed for 5 min for each extraction, and the upper hexane layer aspirated off as before, and transferred to a clean dry tube. The hexane was then evaporated under a stream of nitrogen, and the sample redissolved in the HPLC solvent.

A second method of extraction utilised the same assay components as those detailed above. After the 20 minute incubation, the internal standard was added, 1 ml of ice-cold isopropanol was added to the incubation vial and the solution vortexed for 1 min. The mixture was then centrifuged at 16 000 \times g for 30 min, and the supernatant containing retinoids removed for analysis.

The third extraction method used assay conditions as described above, except a buffer containing 0.1 M potassium phosphate and 1 mM EDTA, and a final reaction volume of 400 μ l was used. After addition of internal standard, the reaction was stopped by adding 1 volume (400 μ l) of ice cold ethanol. The retinoids were then extracted by adding 2 ml ethyl acetate and vortexing for 5 min. The organic phase was transferred to a clean tube, dried over anhydrous sodium sulphate, evaporated under nitrogen and redissolved in HPLC mobile phase.

Due to the general use of an HPLC assay for determining kinetic parameters with retinoids as substrates, the attempts described above were made to set up such a system to enable direct comparison with the work of other groups (Bhat *et al.*, 1995; Bhat *et al.*, 1996; Chen *et al.*, 1995; Chen *et al.*, 1996; Chen *et al.*, 1996; Chen *et al.*, 1996; Chen *et al.*, 1994; el Akawi & Napoli, 1994; Napoli *et al.*, 1995; Napoli *et al.*, 1991; Napoli & Race, 1990; Posch *et al.*, 1992; Labrecque *et al.*, 1993; Labrecque *et al.*, 1997; Labrecque *et al.*, 1995). Initial characterisation of this HPLC enzyme assay was carried out using each of the above extraction methods. The method which gave the highest extraction

Co-extraction number	Component	% Extraction ¹	Relative extraction (retinoic acid:retinol acetate)
1	Retinoic Acid	83.1%	1.45
	Retinol Acetate	57.3 %	
2	Retinoic Acid	57.1 %	1.22
	Retinol Acetate	46.8 %	
3	Retinoic Acid	67.7 %	1.38
	Retinol Acetate	48.9 %	
4	Retinoic Acid	52.5 %	1.18
	Retinol Acetate	44.5 %	
5	Retinoic Acid	79.9 %	1.47
	Retinol Acetate	54.3 %	
6	Retinoic Acid	76.0 %	1.25
	Retinol Acetate	61 %	
. 7	Retinoic Acid	88.8 %	2.27
	Retinol Acetate	39.2 %	
8	Retinoic Acid	114 %	1.73
	Retinol Acetate	66 %	

Table 3.2: Relative Extraction of Retinol Acetate and Retinoic Acid

¹Calculated from the known amount of sample added to reaction mixture

percentages was method three, and so the subsequent method development utilised this method.

For an accurate assay, the important parameter is the reproducibility of extraction of the internal standard compared to each of the retinoid species of interest. It can be seen from Table 3.2 that the relative extraction efficiencies of the product retinoic acid, and the internal standard retinol acetate, varied widely from a ratio of 2.27 to 1.18 (retinoic acid:retinol acetate). It can also be seen that the actual extraction of compounds from the assay mixture varied from 39.2 % to 88.8 %. Although an extraction efficiency of 95-100 % would have been ideal, and indeed seems to be the efficiency reported by some other groups, a low extraction efficiency does not immediately invalidate this method. If the relative extraction of the product of interest and the internal standard is consistent, this can be used to adjust the obtained results to give quantifiable answers. The internal standard is added to the assay after the reaction has occurred, and immediately prior to retinoid extraction. In addition, the nature of the standard is such that the components

of the assay will not alter it in any way. Therefore, knowing the amount of internal standard added to the assay, and knowing the ratio of co-extractability of the internal standard and the product, a simple calculation would yield a reliable answer.

In this case neither the actual extraction nor the co-extraction of the compounds were consistent to the degree necessary to obtain reliable answers. A number of different methods were used (see 3.2.2.4), and though some gave better results than others, none were satisfactory for the intended purpose. Reasons for the disparity between my results and those reported by other groups are not clear. All precautions were taken to ensure maximum extractibility such as the use of glass incubation vessels, working under dim yellow light, and treating each separate incubation in a consistent manner. Communications with a researcher who had had previous success with this type of methodology, also did not identify any possible causes for the inconsistencies (Martini, R; Personal Communication). However it was noted that in the experiments performed by this group, high concentrations of retinal were required (10-200 μ M) in order to get readable signals (Martini & Murray, 1994). These concentrations were not appropriate for the purpose of the present study, and the solubility of retinal at such high concentrations must also be questioned.

However, the detection of retinoids by reversed-phase HPLC, and quantification of retinoids using authentic standards and an integrator, was successful. Authentic retinoids applied to the column gave consistent retention times (Figure 3.4 and Table 3.3) to allow the identification of extracted retinoids. The standard curves obtained were accurate and linear over the amounts of retinoids used (Figure 3.3, 3.4). In addition the high degree of sensitivity required to visualise the small quantities of retinoids which are produced by the enzyme assays was obtained by the developed system.

Although the HPLC retinoid retention times and quantification by standard curves were well established (see Figure 3.2, 3.3 and 3.4 and Table 3.3), the variation in relative extractions (Table 3.2) meant that this method did not have the accuracy required to establish kinetic parameters.



Figure 3.3: Standard Curve for Retinoic Acid

Standard curves are constructed from peak area output using known amounts of standards (see section 3.2.2.3 and 3.2.2.4), which elute from the C18 column at an observed retention time (see Figure 3.2)



Figure 3.4: Standard Curve for Retinol Acetate

Standard curves are constructed from peak area output using known amounts of standards (see section 3.2.2.3 and 3.2.2.4), which elute from the C18 column at an observed retention time (see Figure 3.2)

Component	Retention Time (min) (n=50)	
All-trans Retinoic Acid	3.9 ± 0.19	
All-trans Retinal	10.4 ± 0.5	
Retinol Acetate	20 ± 1	

Table 3.3: Retention Times for HPLC Separation of Retinoids

3.2.2.5 Development of an Absorbance Spectroscopy Kinetic Assay

An alternative method of determining kinetic parameters is by spectroscopy. If a component of the reaction being studied absorbs light of a particular wavelength, absorbance spectroscopy is a convenient, useful method for determining kinetic parameters. However, although absorbance spectroscopy was used to give estimations of kinetic parameters, this method was not considered to be sensitive enough to determine accurate K_m values for isomers of retinal. Three components of the reaction have a maximum absorbance at similar wavelengths - 340 nm, 350 nm and 383 nm (NADH, all-*trans* retinoic acid, and all-*trans* retinal respectively). To avoid interference between overlapping absorbance spectra, the disappearance of all-trans retinal at 400 nm can be monitored. At 400 nm, there is no absorbance by either of the products, so the only absorbing compound is all-*trans* retinal, and even though this is not the λ_{max} of all-*trans* retinal, the absorbance at this wavelength is still great enough to be measured.

Initial estimates of kinetic parameters of the NAD⁻-dependent conversion of retinal to retinoic acid by sAlDH 1 were calculated by following the disappearance of retinal at 400 nm, and analysing the resultant data by the single reaction progress curve method (Yun & Suelter, 1977). To use the single reaction progress curve method for obtaining kinetic parameters, the velocity of the reaction at several different substrate concentrations throughout the course of a complete reaction is determined. The velocities are then treated by standard methods for dealing with Michaelis-Menten data, in this case Lineweaver-Burk analysis. After determining the rate at a number of substrate concentrations, the data was analysed using the microcomputer program ENZFITTER (Leatherbarrow, 1987) (Figure 3.5). The estimated K_m from at least 4 independent experiments was determined to be $0.24 \pm 0.08 \,\mu$ M.



Figure 3.5: Single Reaction Progress Curve of sAIDH-catalysed All-trans Retinal Oxidation

The reaction was carried out at 25°C at pH 7.4. Kinetic parameters were calculated from this curve as described (Yun & Suelter, 1977)

The results obtained using this method were considered to be an approximation only. Because of the low K_m , rates which need to be measured in order to obtain a more accurate result are exceedingly low, and fall out of the sensitivity range of the spectrophotometer, so only approximate data can be obtained using absorbance spectrophotometry.

3.2.2.6 Development of a Fluorescence Spectroscopy Kinetic Assay

An alternative method for obtaining kinetic data is using fluorescence spectroscopy. Molecules absorb light photons, after which they undergo a transition to a higher energy level. The amount of light absorbed is equal in energy to the difference between the two energy states. During the time the molecule spends in the exited state, some vibrational energy is lost by collision with other molecules. If all energy is not dissipated by collisions, the electron returns to the ground electronic state with the emission of energy as fluorescence (Guilbault, 1990). Fluorimetry is possible in this case, as one of the assay products (NADH) emits fluorescence at 457 nm when excited at 340 nm. In addition, the other assay components show negligible fluorescence. Retinol and its esters are intensely fluorescent, with an excitation maximum at 325 nm and an emission maximum at 480 nm, but other retinoids have little or no appreciable fluorescence (Furr *et al.*, 1994).

There are two main advantages of using fluorescence, 1) only 10 % of those compounds which absorb light, emit energy as fluorescence lowering the likelihood of other compounds in the assay interfering with the signal, and 2) fluorescence provides up to an order of magnitude more sensitivity than absorbance (Guilbault, 1990). Other advantages to using fluorescence spectroscopy are its specificity and large linear range of analysis. The specificity is due in part to the fact that two wavelengths are used which makes it unlikely that two compounds in the same system will interfere with each other's signal. The linear range of analysis in fluorescence is often over 6-7 orders of magnitude (with minimal intrinsic error), while a spectrophotometer typically gives 2-3 orders of magnitude (Guilbault, 1990). So it can be seen that fluorescence spectroscopy may be a good option for measuring the NAD⁺ dependent oxidation of retinal.

There are also some advantages to using this method over the HPLC based assay. The obvious advantage is the ease of this method, and if a fluorimeter is available, the cost is very low. A major advantage when studying kinetics is that the time-course of each assay can be examined. This allows the linearity of the rate to be checked each time, and

also other features of the assay (such as a burst) to be identified. With the HPLC based assay, the incubation time of the assay is important. The incubation must be stopped while the assay is in its linear or steady-state phase. If not, the kinetic data calculated may not accurately reflect the actual situation. When using the HPLC assay, it is likely that shorter incubation times for lower substrate concentrations would be needed, as at the low concentrations needed to accurately calculate a K_m , (i.e. concentrations below the K_m — the low nanomolar range), the substrate will be used up very quickly. Fluorimetry, therefore is a good alternative to the conventional HPLC based assay, and has a number of important advantages. The presence of a compound in the assay which emits absorbed energy as fluorescence allows the progress of a reaction to be monitored over time in a similar manner to absorbance spectroscopy.

The Perkin-Elmer LS50B Luminescence Spectrophotometer is a highly sensitive instrument which allows the measurement of chemiluminescence, phophorescence, fluorescence and bioluminescence at controlled temperatures. Because of the sensitivity of fluorescence spectroscopy, there are a number of practical considerations that must be taken into consideration. Only 4-sided quartz cuvettes may be used, as emission must be measured at right angles to the light beam used for excitation. In particular, since particles of dust interfere greatly with readings, all glassware must be acid washed, and all solutions filtered prior to use. Initial scans of pure NADH in milli-Q filtered water and 50 mM phosphate buffer pH 7.4, showed that the maximum excitation wavelength for NADH fluorescence was 340 nm, and the maximum emission wavelength was 457 nm. These were the values used in all subsequent experiments. Because of the low reaction rates to be measured, the slit widths for both excitation and emission were set at 15 nm to ensure the maximum signal was seen. All assays were run over 5 minutes initially, however, at lower substrate concentrations, all the substrate was metabolised within 30-60 seconds. The steady state rate was calculated for each reaction, with exceedingly good agreement between duplicate and triplicate reactions. Data was analysed using the microcomputer program ENZFITTER (Leatherbarrow, 1987).

The described fluorescence spectroscopy kinetic assay was then used in the kinetic characterisation of sAlDH 1 and hAlDH 1 as described in section 3.2.3.

3-83

3.2.3 Kinetic Characterisation of sAlDH and hAlDH

As indicated in the section on method development above, fluorimetric analysis had advantages over both spectrophotometric and HPLC methods for kinetic studies, and was used to obtain the results presented in this section.

3.2.3.1 NAD⁺-Dependent Oxidation of All-*trans* and 9-*cis* Retinal by sAlDH and hAlDH

All assays were monitored by following the production of NADH fluorimetrically (see above). The stoichiometry of retinal dehydrogenation is as follows:





The production of NADH is therefore equal to the production of retinoic acid, and so the appearance of NADH measured fluorimetrically can be used to calculate the rate of retinoic acid production. An apparent K_m was calculated for NAD⁺ with respect to all-*trans* and 9-*cis* retinal, and for all-*trans* and 9-*cis* retinal with respect to NAD⁺. Data points were collected at concentrations above and below the observed K_m . Each data point was an average of duplicates or triplicates from 2-4 different experiments. The K_m was calculated by fitting the data points to the Michaelis-Menten equation (V vs. [S] plot) using the microcomputer program ENZFITTER (Leatherbarrow, 1987). The V vs. [S] plot was then transformed into a Lineweaver-Burk plot (1/V vs. 1/[S]) by the same program (Figures 3.7 - 3.10). Substrate inhibition by the aldehyde was observed at concentrations above 5 μ M, so the high concentrations were not used in the calculation of kinetic parameters. A summary of kinetic parameters obtained using these methods is shown in Table 3.4



Figure 3.7: V vs [S] Plot and Lineweaver-Burk Plot of hAlDH1 Catalysed All-trans Retinal Oxidation



Figure 3.8: V vs [S] Plot and Lineweaver-Burk Plot of hAlDH1 Catalysed 9-cis Retinal Oxidation



Figure 3.9: V vs [S] Plot and Lineweaver-Burk Plot of sAlDH1 Catalysed All-trans Retinal Oxidation



Figure 3.10: V vs [S] Plot and Lineweaver-Burk Plot of sAIDH1 Catalysed 9-cis Retinal Oxidation

	sAlDH	hAlDH
K _m (all- <i>trans</i> retinal)	$0.140 \pm 0.013 \ \mu M$	$0.058 \pm 0.0082 \ \mu M$
K _m (9-cis retinal)	$0.140 \pm 0.011 \mu M$	$0.060 \pm 0.0085 \ \mu M$
K_{m} (NAD ⁺ - 9- <i>cis</i>)	$1.30 \pm 0.10 \ \mu M$	ND
K_{m} (NAD ⁺ - all- <i>trans</i>)	$1.90 \pm 0.20 \ \mu M$	ND

 Table 3.4: A Summary of Kinetic Parameters Determined for hAlDH 1

 and sAlDH 1

These results, obtained using fluorimetry, agree with those estimated using spectrophotometry (Yoshida et al., 1992, this thesis, section 3.2.2.5). Prior to this work, the only estimate of a K_m of human AlDH for all-trans retinal (Yoshida et al., 1992), was calculated using a single reaction progress curve (Yun & Suelter, 1977), and no work had been done using 9-cis retinal. While this method may have been adequate, (and indeed proved to give similar results in our hands to those given by the more accurate method of fluorimetry), it was considered desirable to calculate the K_m more accurately, and look at the reaction with another substrate of biological importance — 9-cis retinal. The fluorimetric assay using different retinal concentrations above and below the observed K_m , while holding NAD⁺ at a constant saturating concentration, gives a more accurate measure of K_m than the method using a single reaction at one retinal concentration only. Recent results using a complex double-substrate spectrophotometric method with hAlDH 1 (Klyosov, 1996), give a much higher estimate for the K_m of retinal (1.1 μ M compared with the value of 0.058 μ M in this study). We believe that the fluorimetric method developed in this study is a direct and highly accurate method for the measurement of K_m values for retinal oxidation, and is probably providing a more correct and physiologically relevant value than that reported in the study by Klyosov (1996).

The data in this study indicate that AIDH 1 may play an important role in retinoid metabolism *in vivo*, as the concentration of cytosolic retinal has been estimated at around 0.1 μ M (Yoshida *et al.*, 1992). As yet, no other aldehyde dehydrogenase isoenzymes with the ability to oxidise retinal have been identified in organisms other than the rat and mouse (see section 1.2.2). There are other enzymes which have been shown to oxidise retinal - the cytochrome P450, xanthine dehydrogenase and aldehyde oxidase enzymes

(see section 1.2.2). However, due to the high K_m values for retinal oxidation (8-125 μ M (Bhat *et al.*, 1988a; Huang & Ichikawa, 1994; Tsujita *et al.*, 1994), 0.9 μ M (Lee *et al.*, 1991), 14 μ M (Raner *et al.*, 1996) for aldehyde oxidase, xanthine oxidase and cytochrome P450 respectively), and other proposed roles for these enzymes (for example the role of P450 enzymes in retinoic acid metabolism), it seems likely that AlDH 1 may play an important role in retinoic acid production. In addition, the fact that AlDH 1 does not seem to be as important in acetaldehyde oxidation as the mitochondrial AlDH 2 also points towards its ability to oxidise retinal as being its most important physiological function.

3.2.3.2 Inhibition of the Esterase Activity of sAlDH by Citral and Retinal

The aim of these experiments was to examine how well citral (a mixture of *cis* and *trans* 3,7-dimethyl-2,6-octadienal), and all-*trans* retinal bind to the active site of AlDH 1 in the absence of NAD⁻ (Figure 3.11). This was achieved by studying the hydrolysis of the AlDH 1 esterase substrate *p*-nitrophenyl pivalate (*p*-nitrophenyl trimethylacetate), and the extent to which this activity was inhibited by the said aldehyde compounds. The kinetic assays were carried out spectrophotometrically by following the production of *p*-nitrophenoxide at 400 nm as stated in section 2.2.2. The resulting data were analysed using ENZFITTER (Leatherbarrow, 1987). Inhibition can best be seen by examining Lineweaver-Burk plots with and without inhibitors (Figure 3.12 and 3.13). Results summarising the inhibition of *p*-nitrophenoxide formation by citral and retinal are shown in Table 3.5.

The inhibition of the esterase activity of AlDH 1 by citral and retinal was studied in order to determine if these compounds would bind to the enzyme in the absence of NAD⁺. The elucidated kinetic mechanism for AlDH 1 with aldehyde as a substrate is an ordered bi-bi mechanism, where the binding of NAD⁺ is postulated to cause a conformational change in the enzyme to allow binding of the aldehyde (Blackwell *et al.*, 1989). The hypothesis that the main function of AlDH 1 was retinal oxidation led to the idea that the active site may have a high affinity for retinal binding, even in the absence of NAD⁺. The *trans* form of citral is a structural analogue of retinal (see Figure 3.11), and inhibition of the hydrolysis of *p*-nitrophenyl pivalate by this compound was also investigated. Both compounds competitively inhibited this esterase activity, with K_i values of approximately 21 μ M (citral), and approximately 11 μ M (retinal) (Table 3.5). While both compounds appear to bind AlDH 1 in the absence of NAD⁺, the fact that the K_i for retinal inhibition is not exceptionally low, and the fact that retinal binds with a similar affinity to citral,
Table 3.5: The Inhibition of the Hydrolysis of p-Nitrophenyl pivalate by Retinal and Citral

Concentration of Inhibitor		K _i	
Retinal	6.7 μM	~11 µM	
Citral	53 μM	~21 µM	

(Reported in Kitson et al, 1995).

seems to indicate that the active site is not designed to bind retinal with a high affinity in the absence of NAD⁺. In fact, the postulated NAD⁺-induced conformational change prior to aldehyde binding is supported by this data, and by the low K_m values for retinal obtained in the dehydrogenase reaction with NAD⁺ present (50-140 nM) (Table 3.4).

3.2.4 pH Optimum of NAD⁺-Dependent Retinal Oxidation

The effect of pH on retinal dehydrogenation was investigated using three different buffer systems covering a pH range of 4.5-11. A similar pH profile for this reaction was obtained from four separate experiments, using averages of duplicate assays. Assays were performed spectrophotometrically using 2 μ M retinal as stated in section 2.2.2. The pH dependence of this reaction is shown in Figure 3.12. It can be seen that the pH optimum for this reaction is at pH 9.5-10. Previous studies using propionaldehyde as a substrate identified maximum enzyme activity to be at pH 7 (Buckley *et al.*, 1991). For studying the kinetics of the reaction with retinal as substrate, the pH of 7.4 was routinely used in all assays. This is because physiological pH is 7.4, and so reactions studied at this pH would more likely represent the situation *in vivo*. The rates at pH 7.4 were adequate enough to see the reaction when using low amounts of substrate.

CHO All-trans retinal CHO

All-trans citral

Figure 3.11 Structural Formulae for All trans-Citral and All-trans Retinal



Figure 3.12: Inverse Plot of the Inhibition of sAIDH1 Esterase Activity by Citral



Figure 3.13: Inverse Plot of the Inhibition of sAIDH1 Esterase Activity by All-trans Retinal



Figure 3.14: pH Dependence of NAD⁺-Dependent Oxidation of All-trans Retinal by sAIDH1

4. Chapter Four: Expression and Purification of CRBP and use of CRBP-Retinal as a Substrate for AIDH 1

4.1 Introduction and Aims

It has been proposed that cytosolic retinal exists primarily as a 1:1 complex with the lipocalin cellular retinol-binding protein (CRBP) I (see sections 1.2.2 and 1.3.1.4). If this hypothesis is correct, enzymes which play an important role in retinoic acid synthesis are likely to accept retinal bound to CRBP as a substrate. The proposal that cytosolic AlDH 1 plays a significant role in retinoic acid biosynthesis made it necessary to see if this enzyme would accept CRBP-retinal as a substrate. The aim of this section was therefore to express and purify CRBP protein from CRBP cDNA, and to use the purified CRBP I to study the kinetics of AlDH 1-catalysed CRBP-retinal oxidation.

4.2 Results

4.2.1 Expression and Purification of CRBP

A plasmid containing the gene encoding human CRBP I was obtained (pT7-7-CRBP) (J. Findlay, Leeds University). This vector contains the ampicillin resistance gene, and the CRBP gene under the control of a T7 promoter (see Figure 4.1). The C-terminus of the CRBP cDNA had been engineered to encode a 'His-tag' of 6 consecutive histidine residues to enable affinity purification of expressed protein using a chelating resin. E. coli strain BL21 (DE3) cells (containing a λ phage lysogen incorporating T7 RNA polymerase under the control of a lac promoter), were transformed with the vector pT7-7-CRBP, and cells containing the plasmid were selected by growing the transformed cells on LB agar plates containing $50 \,\mu g/ml$ ampicillin. A single colony was picked and a small liquid LB culture seeded with the colony and grown at 37°C for about 16 hours, again using ampicillin selection. From this, a 100 ml culture was seeded, and grown to an OD_{600} of 0.5-0.6. Overexpression of CRBP was then induced by the addition of 0.5 mM isopropyl- γ -D-thiogalactopyranoside (IPTG), which binds to and releases the lac repressor from its promoter, allowing production of T7 RNA polymerase, which in turn induces the expression of CRBP. After incubating for a further 3 hours, the cells were harvested and pelleted by centrifugation at $4000 \times g$ for 20 min. The pelleted cells were



Figure 4.1: Plasmid Map of pT7-7-CRBP

This plasmid was obtained from Professor J. Findlay (Leeds University, UK). The map shows the insertion of human CRBP cDNA between Xba I and Hin dIII sites. Also shown is the ampicillin resistance gene, origin of replication, T7 promotor, and ribosome binding site. resuspended in PBS containing 1 mM PMSF, and the cells lysed with a Virsonic sonicator $(3 \times 30$ seconds, power level 4, 10 mm probe diameter at tip). This mixture was then centrifuged at 13 000 \times g for 30 minutes to pellet insoluble material and the contents of the cytosol were separated by SDS-PAGE (Figure 4.2). A comparison of induced and uninduced cells showed that CRBP was being overexpressed; a strong protein band could be seen at around 18 kDa. There was a small amount of 'leakage' in the system, i.e. some CRBP was being produced in the uninduced cells. The expression system was then optimised in terms of cell density at induction, length of induction, concentration of IPTG, and temperature. Optimal expression was seen with 0.5 mM IPTG, cells grown at 37°C, and 3 hours induction time. The cell density at induction did not seem critical to high levels of expression, with $OD_{600} = 0.6-1.0$ giving identical results. The expression of this plasmid was extremely robust however, and varying any parameter did not produce serious deleterious effects. After characterisation of the expression system, a glycerol stock was made of the parent culture, and this stock was used to seed further expression experiments. This ensured that the same clonal line was used in every preparation. A large-scale preparation of pT7-7-CRBP was made (see section 2.2.5.2), and the CRBP gene was sequenced, and was shown to be identical to the published human CRBP nucleotide sequence.

The purification of expressed CRBP was facilitated by the C-terminal 6-His tag. HiTrap® (Pharmacia) chelating columns loaded with Ni²⁺ ions were used, which bind to imidazole rings including histidine residues in proteins. After charging the column with nickel chloride, the column was equilibrated in 0.02 M sodium phosphate buffer pH 7.2, containing 100 mM imidazole. The sample (cytosolic preparation from cells expressing CRBP) was then loaded on to the column and 5 column volumes of equilibration buffer were washed through to remove unbound proteins. Obviously, any proteins with surface histidine residues will also have an affinity for the column, and proteins with more surface histidine residues will have a greater affinity for the column. The concentration of imidazole in the equilibration and wash buffer must be optimised in order to exclude other proteins which may otherwise bind to the column and compromise the purification. After a number of trials, it was seen that 100 mM imidazole was sufficient to exclude all protein binding, except that of the 6-His tagged CRBP. The protein was eluted using the same buffer, containing 300 mM imidazole. The protein was typically eluted within 1-2 column volumes, and was pure as shown by silver staining of SDS-gels (see Figure 4.3). After elution, fractions containing the protein (identified by SDS- PAGE) were pooled and dialysed at 4°C with three changes of buffer (21) to remove the imidazole. The pure



Figure 4.2: SDS-PAGE of Soluble E.coli Protein

E.coli transformed with pT7-7-CRBP and grown under the conditions described in the text, were then harvested and the cytosolic contents run on an SDS-gel as described in Materials and Methods. 'Induced' lanes show the induction of CRBP production upon the addition of 0.5 mM IPTG.

Lane:

- 1,8: Size Ladder (Sigma 7)
- 2,4,6: Uninduced cytosolic contents
- 3,5,7: Induced cytosolic contents



LANE: 1 2 3 4 5 6 7 8



Figure 4.3: SDS-PAGE of CRBP Purification

Gel A: Purification of CRBP using 6-His tag as described in section 4.2.1. Pure CRBP is seen in lanes 7 and 8.

Gel B: Elution fractions from chelating column purification stained with silver stain as described in Materials and Methods. Pure CRBP is seen in lanes 3 and 6.

Gel A Lane:

Size Ladder (Sigma 7)

1:

- 2,3,4: Proteins not bound to column at 100 mM imidazole
- 5-8: Succesive fractions eluted with 300 mM imidazole

Gel B Lane:

1,7: Size Ladder (Sigma 7)2-6: Fractions eluted with

300 mM imidazole

protein was then stored at 4°C until needed. The identity of the expressed purified protein was confirmed by electrophoretic transfer onto PVDF membrane as described (section 2.2.9), and N-terminal protein sequencing by an Applied Biosystems 476A automated Edman degradation (section 2.2.12). Results from this procedure identified the first 20 amino acids as identical to human CRBP (Table 4.1).

TAVIE 4.1. IN-LEI MIMAI SEQUENCINZ VI L'ADIESSEU ANU I UTITEU UND	Table -	4.1:	N-terminal	Sequencin	g of	Expressed	and]	Purified	CRB
---	---------	------	------------	-----------	------	-----------	-------	----------	-----

Residue Number	Residue predicted by Sequencer	CRBP sequence (SWISSPROT)
1	Met	Met
2	Pro	Pro
3	Val	Val
4	Asp	Asp
5	Phe	Phe
6	Thr	Thr
7	Gly	Gly
8	Tyr	Tyr
9	Ттр	Ттр
10	Lys	Lys
11	Met	Met
12	Leu	Leu
13	Val	Val
14	Asn	Asn
15	Glu	Glu
16	Asn	Asn
17	Phe	Phe
18	Glu	Glu
19	Glu	Glu
20	Tyr	Tyr

4.2.2 The Binding of CRBP to Retinal

CRBP has been shown to bind all-*trans* retinal with a K_d of 50-100 nM (see section 1.3.1.1). For the purpose of this study it was necessary to ascertain whether the recombinant CRBP actually bound retinal. Two approaches were used to determine if CRBP was binding retinal. The first was a simple measurement of absorbance spectra. The absorbance maximum of retinal has been shown to be red shifted by about 30 nm when bound to CRBP, as compared to free retinal (Levin *et al.*, 1988). This is a reflection of the environment in which retinal is found. A simple test for binding was to look for a λ_{max} shift of ~30 nm for bound retinal compared to free. This was carried out on each day that kinetic data were measured using CRBP (Figure 4.4). As a control, the same test was performed with 9-cis retinal (Figure 4.5). While a typical experiment yielded a λ_{max} shift of ~25-30 nm for all-*trans* retinal bound to CRBP, a shift of only 3-4 nm was seen with 9-cis retinal.

An additional method used to confirm the binding of all-trans retinal to CRBP involved the use of a fluorescent probe, 1-anilinonaphthalene 8-sulfonic acid (1,8-ANS). This probe has been used in a general assay for assessing the binding of ligands to members of the intracellular lipid-binding protein family (iLBP family) (Kane & Bernlohr, 1996). Firstly, the binding of 1,8-ANS by CRBP was studied as described (Kane & Bernlohr, 1996). 1,8-ANS has no appreciable fluorescence in aqueous solution, but upon binding to CRBP (or other iLBP), a considerable enhancement of fluorescence is seen. The strength of binding of the probe to CRBP was then examined by looking at the increase of fluorescence as increasing amounts of CRBP were added to a constant amount of 1,8-ANS, and as CRBP was added in excess, the plateau in fluorescence. Finally, all-trans retinal was used in competition assays with previously bound 1,8-ANS. These experiments were not performed to re-quantify previously established binding parameters (Li et al., 1991). The purpose of this experiment was to show that all-trans retinal was binding to CRBP by using a different method from the absorbance spectra shift observed upon addition of CRBP to all-trans retinal. The fluorescent signal produced by CRBP and 1,8-ANS was quenched upon the addition of increasing amounts of all-trans retinal. This indicates that the retinal is binding to the CRBP and displacing the 1,8-ANS, thus decreasing the intensity of the fluorescent signal produced only when 1,8-ANS binds to the iLBP.

4.2.3 CRBP-Retinal as a Substrate for Human and Sheep AIDH 1

The recombinant CRBP I was used to investigate the ability of AlDH 1 to accept retinal bound to CRBP as a substrate. The same fluorimetric method was used to study this reaction as had been previously applied to examine the reaction of AlDH 1 with free retinal (see Chapter Three, and section 2.2.3). The main difference from the aforementioned method was the pre-incubation of retinal and CRBP in reaction buffer for 15 minutes in the dark prior to initiating the reaction, to ensure that binding of substrate to the CRBP was at equilibrium. Data for the oxidation of CRBP-retinal by hAlDH 1 was calculated using a 2:1 ratio (CRBP:retinal) (Figure 4.6). Data for the oxidation of CRBP-retinal by sAlDH 1 was calculated using a 2:1 ratio (CRBP:retinal) (Figure 4.7), and a 3.5:1 ratio (Figure 4.8). Kinetic parameters calculated using these



Figure 4.4: Absorbance Spectra of All-trans Retinal and CRBP-All-trans Retinal

Spectra were collected on a Cary UV-vis double-beam spectrophotometer. A ratio of CRBP:Retinal of 2:1 was used. Absorbance maxima are noted.



<u>Figure 4.5: Absorbance Spectra of 9-cis Retinal and</u> <u>CRBP-9-cis Retinal</u>

Spectra were collected on a Cary UV-vis double-beam spectrophotometer. A ratio of CRBP:Retinal of 2:1 was used. Absorbance maxima are noted.



Figure 4.6: V vs [S] Plot and Lineweaver-Burk Plot of hAlDH1 Catalysed CRBP-Retinal Oxidation

Data were fitted using the microcomputer program ENZFITTER (Leatherbarrow, 1987). CRBP was in a 2-molar excess over all-*trans* retinal. Kinetic parameters derived from these curves are shown in Table 4.2



Figure 4.7: V vs [S] Plot and Lineweaver-Burk Plot of sAlDH1 Catalysed CRBP-Retinal Oxidation

Data were fitted using the microcomputer program ENZFITTER (Leatherbarrow, 1987). CRBP was in a 2-molar excess over all-*trans* retinal. Kinetic parameters derived from these curves are shown in Table 4.2



Figure 4.8: V vs [S] Plot and Lineweaver-Burk Plot of sAIDH1 Catalysed CRBP-Retinal Oxidation

Data were fitted using the microcomputer program ENZFITTER (Leatherbarrow, 1987). CRBP was in a 3.5-molar excess over all-*trans* retinal. Kinetic parameters derived from these curves are the same as those calculated using a 2-molar excess shown in Table 4.2 curves are summarised in Table 4.2. No differences were seen in kinetic parameters between these ratios. This additional evidence supports other observations showing that AlDH 1 uses holo-CRBP as a substrate.

Table 4.2: A Summary of Kinetic Parameters of sAIDH and hAIDH for CRBP-Retinal and Free All-trans Retinal

	sAlDH	hAlDH
K _m (CRBP-retinal)	$0.92 \pm 0.20 \ \mu M$	0.216±0.035 μM
K _m (Free Retinal)	$0.14 \pm 0.013 \ \mu M$	$0.058 \pm 0.0082 \ \mu M$
V _{max} (CRBP-Retinal)	32.18 ± 3.26 F/min/µl	3.443 ± 0.14 F/min/µl
V _{max} (Free Retinal)	88.34 ± 4.33 F/min/µl	9.919 ± 0.453 F min/µl

4.3 Discussion

The results in Table 4.2 show that both sAlDH and recombinant hAlDH recognise CRBP bound to all-*trans* retinal as a substrate. This result has important implications in our understanding of the production of retinoic acid, and the regulation of its synthesis.

The optimisation of expression and purification parameters such as induction time, temperature and IPTG concentration, and nickel column wash and elution buffer imidazole concentration, allowed the production of large quantities of pure recombinant CRBP protein very quickly. Although it is now generally accepted that CRBP I binds to all-*trans* retinal with a K_d of 50-100 nM, it was vitally important to determine if the recombinant CRBP was actually binding retinal. To check this, two methods were used. One involved all-*trans* retinal competing for binding with a fluorescent probe. The replacement of the probe 1,8-ANS by all-*trans* retinal in the binding site of CRBP decreases the intensity of the original fluorescent signal, indicating that binding of the ligand to the binding protein is occurring. The results obtained in this study using 1,8-ANS, were similar to those published by (Kane & Bernlohr, 1996), and although this experiment yielded no novel information, the results supported retinal binding. A second method was a simple retinal absorbance spectrum, run in the presence and absence of

CRBP. Previous results using this method gave a red shift in the λ_{max} of ~30 nm when CRBP was present, compared to the λ_{max} without CRBP (Levin *et al.*, 1988). This shift was also seen using our CRBP preparation. Absorbance spectra were run prior to the start of each experiment, and the experiment was not carried out unless the required absorbance shift of 30 nm was seen upon addition of CRBP to all-*trans* retinal. This was absolutely necessary, as degradation of CRBP over time, or after undergoing even one freeze-thaw cycle, often left protein which was unable to bind retinal as efficiently. A smaller λ_{max} shift was seen upon CRBP binding in such cases, and these samples of CRBP were not used for kinetic measurements.. A suitable control for this experiment was to collect absorbance spectra in the same manner using 9-*cis* retinal in the place of all-*trans*. It has been shown that CRBP I does not bind to 9-*cis* retinal, and indeed only a small λ_{max} shift of 3-4 nm was seen using 9-*cis* retinal.

The fluorimetric assay which was developed to study the NAD⁺ dependent oxidation of retinal by AlDH 1 was also used in looking at the reaction with CRBP-retinal. As stated, the only difference when using CRBP was a 15 minute incubation prior to initiating the reaction to ensure binding of retinal to CRBP. To ensure adequate retinal binding, the assays were performed with a two-fold molar excess of CRBP over retinal. Calculations performed using this ratio of CRBP to retinal, and using the highest estimated K_d of CRBP for retinal (100 nM) (Li et al., 1991), showed that over the range of concentrations used for this experiment, the concentration of free retinal in equilibrium with the bound would remain essentially constant ranging from 0.028 μ M to 0.084 μ M (see Appendix 1). In comparison, the concentration of bound retinal would range from $0.022 \mu M$ to $4.90 \mu M$. This means that if AlDH 1 could not use retinal bound to CRBP as a substrate, the observed rates would be due only to the small amount of free retinal available. These rates would be extremely low, the substrate would be used up quickly, and over the range of concentrations used, the rates would be indistinguishable from each other. Therefore it is immediately obvious if recognition between CRBP-retinal and AIDH 1 is occurring. The assay itself also provides additional confirmation that all-trans retinal does bind to CRBP. If retinal was not binding, the reaction rates would be expected to be exactly the same as those measured for free retinal.

Both sheep and recombinant human enzymes showed reaction rates that varied over the range of concentrations used, and were different from those seen for the free substrate. These data are consistent with AlDH 1 using CRBP-retinal as a substrate. The exact significance of this will remain uncertain until the role of CRBP in retinoid metabolism is

clarified. However, the fact that AlDH 1 has now been shown to utilise 9-*cis* retinal, and all-*trans* retinal, in both free and bound forms, clearly indicates that regardless of the physiological function of CRBP, this enzyme is highly likely to play a major role in human and sheep retinoic acid biosynthesis.

A comparison of the K_m values for free and bound all-*trans* retinal reveals that both hAlDH and sAlDH have a higher K_m for the bound substrate relative to the free. Even though the K_m for bound substrate is higher than for free, hAlDH has a similar K_m for CRBP-bound retinal compared to other retinal dehydrogenases for which values have been reported, and sAlDH has a K_m 4-7 fold higher. The K_m for free retinal for both enzymes is 5-200 fold lower than that of other identified retinal-oxidising enzymes (see Table 1.2). The difference in K_m for bound retinal which was observed between the sheep and human enzymes may arise for a number of reasons. The most obvious reason lies with the species from which the CRBP was obtained. The CRBP used is human, which may mean that the human AlDH 1 enzyme has a more specific recognition of target residues or structural units in CRBP that are important in ligand release. The observed differences may simply be species derived, reflecting the differing retinoid requirements for humans and sheep.

The use of CRBP-retinal as a substrate also influences the V_{max} of the reaction. As previously stated (sections 1.3.1.3 and 1.3.1.5), the CRBP ligand (retinal or retinol) is bound in an orientation such that the aldehyde or alcohol functional group lies in the centre of the structure (see Figure 1.5). In addition, it is not obvious from examining the structure how the ligand enters and leaves the binding site. The two portals by which ligands can enter the binding protein which are usually found in a β -barrel such as CRBP, are blocked in both the apo- and holo-CRBP structures. One entrance is blocked by the N-terminus of the polypeptide chain, and the other is blocked by the helix-turn-helix motif joining two β -strands. Regardless of the mechanism of ligand entry and transfer to metabolic enzymes, the retinoid must be rotated approximately 180° in order for the functional group to enter the active site of the enzyme. The overall process of ligand release and subsequent metabolism may then occur as follows: the first step is recognition of CRBP by an enzyme which induces or initiates some conformational change in CRBP. This in turn leads to ligand release. Once released, the ligand may be either directly transferred to the enzyme via a complex formed between holo-CRBP and the enzyme, or released into the cytosol where it may be immediately bound by the enzyme which is already in close proximity. If direct transfer is the correct model, the ligand will need to rotate before binding in the active site of the enzyme.

Whichever of these mechanisms is correct, it is obvious that the rate of reaction will be slower for CRBP-retinal than for free retinal. This is reflected in these experiments. The V_{max} values are 2-3 fold higher for free retinal compared to those for CRBP-retinal (see Table 4.2). It is possible that the lower reaction rates observed with CRBP present are an additional method of regulating retinoic acid production. Allowing a lower, constant, amount of retinoic acid to be produced over a longer period of time may be more desirable for the cell than producing the same amount of retinoic acid very quickly, especially in times where excess vitamin A is available. This would also be dependent on the amount of retinal produced by retinol and alcohol dehydrogenases, which is probably the rate determining step in retinoic acid production. Again, in times of excess retinol availability, the slow controlled production of retinoic acid by the suggested interaction between CRBP and aldehyde dehydrogenase may provide an additional layer of control.

5. Chapter Five: A Comparison of the Major Aldehyde Dehydrogenases by Multiple Sequence Alignment

5.1 Introduction

The major focus of this thesis is the role that AlDH 1 and other retinal-oxidising AlDH enzymes play in retinal metabolism *in vivo*. A variety of factors have led to the proposal that, in sheep and humans, the major cytosolic class 1 AlDH plays a primary role in retinoid metabolism. In addition, it seems that the situation may be different in rats and mice. In order to analyse the relationship between aldehyde dehydrogenases from different species, a multiple sequence alignment of 32 aldehyde dehydrogenase protein sequences from classes 1, 2, and 3 was undertaken, including all known retinal-oxidising AlDH's. Information from this sequence alignment may be useful in studying the relationship between retinal-oxidising aldehyde dehydrogenase proteins and also in determining possible physiological functions.

5.2 Methods

Thirty two aldehyde dehydrogenase sequences were identified in the protein databases, including sequences representing classes 1, 2, and 3, and also a number of recently identified AlDH's not yet assigned to an existing or new class (see Table 5.1). The sequences were identified using the database search programs ENTREZ (at The National Centre for Biocomputing Information, NCBI) and DELPHOS (University of Leeds, University College of London). These programs search source databases such as SWISS-PROT, PIR (at Johns Hopkins), GenBank translations, DDBJ (DNA Database of Japan at GenomeNet), PRF (Protein Research Foundation of Japan), and the nonredundant protein sequence database OWL (Bleasby & Wooton, 1990). Sequence information was then entered into the GCG suite of programs for the manipulation of biological sequence data (Genetics Computer Group, v 9.0, Madison, Wisconsin, USA). The sequences were aligned using the program 'Pileup'. Pileup creates a multiple sequence alignment using a clustering approach based on a simplification of the progressive alignment method of Feng and Doolittle (Feng & Doolittle, 1987). All sequences are initially compared to one another and the two most similar sequences used to form the first pairwise alignment. Subsequent sequences or clusters of sequences are

Table 5.1: Identification of Proteins Used in Multiple Sequence Alignment

Name	Identification	Databank Accession Number	Number of amino acids	Reference (See below)
BA1	Bovine Class 1 from amacrine cells (cyc)	L36128 (GenBank)	501	l
Hol	Horse Class 1	P15437 (SwissProt)	500	2
E1	E.Coli Class 1	P23883 (SwissProt)	494	3
YCI	Yeast Class 1 (S.Cerivisiae)	U56604 (GenBank)	501	4
C1	Chicken Class 1	P27463 (SwissProt)	509	5
Ml	Mouse Class 1	P24549 (SwissProt)	500	6
R1	Rat Class 1	P13601 (SwissProt)	500	7
S1	Sheep liver Class 1	P51977 (SwissProt)	500	8
111	Human liver Class 1	P00352 (SwissProt)	500	9
RALDH-1	Rat kidney retinal-specific Class 1, type 1	P51647 (SwissProt)	501	10
RalDH1	Rat liver retinal-specific Class 1, type 1	U79118 (GenBank)	501	11
RALDH-2	Mouse retinal-specific, type 2	X99273 (GenBank)	499	12
RalDH2	Ratretinal-specific, type 2	U60063 (GenBank)	499	13
B2	Bovine liver Class 2	P20000 (SwissProt)	520	14
14o2	Horse Class 2	P12762 (SwissProt)	500	15
YMI	Yeast Class 2 (S. Cerivisiae)	U56605 (GenBank)	519	16
M2	Mouse Class 2	P47738 (SwissProt)	519	17
R2	Rat liver Class 2	P11884 (SwissProt)	519	18
H2	Human Class 2	P05091 (SwissProt)	517	19
CRY	Short eared elephant shrew crystallin (eye)	U03906 (GenBank)	501	20
H6	Human saliva precursor	A55684 (PIR)	512	21
HX	Human stomach	M63967 (GenBank)	453	22
YM3	Yeast mitochondrial 3 precursor	P40047 (SwissProt)	519	23
M3	Mouse Class 3 (dioxin inducible)	P47739 (SwissProt)	453	24
R3	Rat Class 3 (Tumour-associated)	P11883 (SwissProt)	452	25
EI3	Human Class 3	P30838 (SwissProt)	453	26
MM	Mouse liver microsomal (dioxin inducible)	U14390 (GenBank)	484	27
RM	Rat liver microsomal	A41028 (PIR)	484	28
H4	Human 4-aminobutyraldehyde oxidising	P49189 (SwissProt)	462	29
· H9	Human γ-aminobutyraldehyde oxidising	P49189 (SwissProt)	493	30
H7	Human stomach AlDH7	P43353 (SwissProt)	468	31
H10	Human liver fatty aldehyde dehydrogenase	U46689 (GenBauk)	485	32

Table 5.1 (Continued): References

.

¹ Saari *et al.*, 1995; ² von Bahr-Lindstrom *et al.*, 1984; ³ Heim & Strehler, 1991; ⁴ Weiner *et al.*, Direct Submission (cytosolic) ⁵ Godbout, 1992; ⁶ Rongnoparut & Weaver, 1991; ⁷ Dunn *et al.*, 1989; ⁸ Stayner & Tweedie, 1995; ⁹ Hsu *et al.*, 1989; ¹⁰ Bhat *et al.*, 1995; ¹¹ Penzes *et al.*, 1997; ¹² Zhao *et al.*, 1996; ¹³ Wang *et al.*, 1996; ¹⁴ Guan & Weiner, 1990; ¹⁵ Johannson *et al.*, 1988; ¹⁶ Weiner *et al.*, Direct Submission (mitochondrial) ¹⁷ Chang & Yoshida, 1994; ¹⁸ Farres *et al.*, 1989; ¹⁹ Hsu *et al.*, 1988; ²⁰ Graham *et al.*, 1996; ²¹ Hsu *et al.*, 1994a; ²² Hsu & Chang, 1991; ²³ Dietrich & Davis, 1994; ²⁴ Vasiliou *et al.*, 1993; ²⁵ Jones *et al.*, 1988; ²⁶ Schuuring *et al.*, 1992; ²⁷ Nebert *et al.*, 1994; ²⁸ Miyauchi *et al.*, 1991; ²⁹ Pietruszko, 1993; ³⁰ Lin *et al.*, 1996; ³¹ Hsu *et al.*, 1994b; ³² Chang & Yoshida, 1997.

5-114

then compared to the initial cluster by a simple extension of the pairwise alignment of two individual sequences. In this way the alignment is built up, progressively including the less similar sequences until all sequences have been included in the final alignment. This pairwise clustering procedure is computationally swift, but does not necessarily produce the optimal multiple alignment. The initial alignment needs to be checked manually to ensure that it is sensible, both with respect to primary sequence and to secondary structure elements if known. The alignment can be viewed and edited using the GCG's multiple sequence alignment editor 'lineup'. However, manual sequence checking is greatly aided by the use of a colour alignment editor such as XALIGN Version 1.0 (Attwood & Perkins) or CINEMA (Attwood et al., 1997). Different colours are assigned to each amino acid type: acidic residues are red, hydrophobic aliphatic residues are grey, aromatic residues are mauve, glycine and proline are orange, cysteine is yellow, basic residues are blue and polar residues are green. The colour coding indicates residues of a similar chemical or structural nature, and greatly aids manual assessing of computer-generated sequence alignments, especially when dealing with large numbers of sequences. The sequence alignment was first manually edited with respect to primary sequence data.

A useful verification of the correctness of a multiple alignment made on a sequence basis only, is to compare the alignment to a known structure if one is available. The crystal structure of sheep liver AlDH 1 has recently been solved to 2.3 Å resolution (S. Moore, manuscript in preparation). Using this structure, the secondary structure elements (α -helices, β -sheets, turns etc.) were identified, and the primary sequence relating to each region of secondary structure was determined (Table 5.2). The initial sequence alignment was then checked against the regions of secondary structure, to ensure that no gaps had been inserted within these elements. Careful examination of the structure showed that the sequence alignment conformed with the structural elements, and no corrections to the alignment with regards to structural elements were necessary. The resulting sequence alignment is displayed in Figure 5.1.

Pileup creates a pictorial representation of the clustering relationships used in creating the sequence alignment called a dendrogram (Figure 5.2). The clustering strategy used to construct the dendrogram is called UPGMA - unweighted pair-group method using arithmetic averages (Sneath & Sokal, 1973). Distances along the vertical axis of the dendrogram are proportional to differences between sequences, while the horizontal distances are arbitrary. It must be stressed that the dendrogram is not a phylogenetic

Figure 5.1: Multiple Sequence Alignment of 32 Aldehvde Dehvdrogenase Sequences (Pages 115-120)

The alignment was created using Pileup and XALIGN as stated in section 5.2. All amino acid numbering relates to sheep AlDH 1 (S1). The sequence is labelled every 50 residues (normal text) to enable particular residues or sequences to be identified. In addition, residues referred to in the text are labelled (bold text).

<u>KEY</u>

Residue Type:

Residue Location:

red	acidic	page 116residues 1-115	
blue	basic	page 117residues 116-254	4
grey	hydrophobic aliphatic	page 118residues 255-386	5
green	. polar	page 119residues 387-500	С
mauve	aromatic	page 120C-terminus of M	M,
orange	glycine and proline	RM, H7, H10	
yellow	cysteine		

sı ↓

BA1	MSSSAMPDVPAPLTNLOFKYTKIFINNEWHSSVSCKKFPVFNPAFF
HO1	SSSGTPDLPVLLTDLKFQYTKIFINNEWHDSVSGKKFPVFNPATE
E1	MNFHHLAYWQDKALSLAIENRLFINGEYTAAAENETFETVDPVTQ
YC1	MTKLHFDTAEPVKITLPNGLTYEQPTGLFINNKFMKAQDGKTYPVEDPSTE
C1	MKKQGSPSNPAPVLPALPEPLKULKIKYTKIFINNEWHDSVSGKKFEVFNPANE
M1	MSSPAQPRVPAPLADLKIQHTKIFINNEWHNSVSGKKFPVLNPATE
R1	SSPAQPAVPAPLANLKIQHTKIFINNEWHNSLNGKKFPVINPATE
S1	MSSSAMPUVPAPLTNLQFKYTKIFINNEWHSSVSGKKFPVFNPATE
H1	SGIP DEPUTY ALT DET TO THE PUTY AND SUSCE
RALDH-1	
RALDH-2	MASTALIA PAPENDALA TATINA AND A CONTRACT AND A CONT
RALDH2	MASLOLLPSPTPNLEIKYTKIRINNEWONSESGRVEPVCNPATG
B2	MLRAVALAAARLGPROGRRLLSAATOAVPTPNOOPEVLYNOIFINNEWHDAVSKKTFPTVNPSTC
HO2	AAAATQAVPAPNQQPEVFYNQIFINNEWHDAVSKKTFPTVNPSTG
YM1	-MLSRTRAAAPNSRIFTRSLLRLYSQAPLRVPITLPNGFTYEQPTGLFTNGEFVASKOKKTFDVINPSNE
M2	MLRAALTTVRRGPRLSRLLSAAATSAVPAPNHQPEVFCNQIFINNEWHDAVSRKTFPTVNPSTG
R2	MLRAALSTARRGPRLSRLLSAAATSAVPAPNQQPEVFCNQIFINNEWHDAVSKKTFPTVNPSTG
H2	MLRAAARFGPRLGRRLLSAAATOAVPAPNOOPEVFCNOIFINNEWHDAVSRKTFPTVNPSTG
CRY	MSSSGMPDLPAPLTNIKIQHTKLFINNEWHDSVSGKTFPVFNPATE
н6	MATANGAVENGOPDGKPPALPRPIRNLEVKFTKIFINNEWHESKSGKKFATCNPSTR
HX	MLRFLAPRLLSLQGRTALYSSAAALPSPILNPDIPYNQLFINNEWQDAVSKKTFPTVNPTTG
YM3	-MLSRTHAAAPNSRIFTRSLLRLYSQAPLRVPITLPNGFTYEQPTGLFTNGEFVASKQKKTFTVINPSNE
M3	
R3	
MM	
PM	
H4	PATG
Н9	MSTGTFVVSOPLNYRGGAAGAGGRSGTEKAFEPATG
H7	
H10	
	E 50 R 99 L 100
	↓ ↓
BA1	EKLCEVEEGDKEDVDKAVKAVRQAFQIGSPWRTMDASERGRLLNKLADLIERDHLLLATMEAMNGGKLFS
BA1 HO1	EKLCEVEEGDKEDVDKAVKAVRQAFQIGSPWRTMDASERGRLLNKLADLIERDHLLLATMEAMNGGKLFS EKLCEVEEGDKEDVNKAVAAARQAFQIGSPWRTMDASERGRLLYKLADLVERDRLILATMESMNGGKLFS
BA1 HO1 E1	EKLCEVEEGDKEDVDKAVKAVROAFOIGSPWRTMDASERGRLLNKLADLIERDHLLLATMEAMNGGKLFS EKLCEVEEGDKEDVNKAVAAAROAFOIGSPWRTMDASERGRLLYKLADLVERDRLILATMESMNGGKLFS APLAKIARGKSVIIDRAMSAARGVFERGD-WSLSSPAKRKAVLNKLADLMEAHAEELALLETLDTGKPIR
BA1 H01 E1 YC1	EKLCEVEEGDKEDVDKAVKAVRQAFQIGSPWRTMDASERGRLLNKLADLIERDHLLLATMEAMNGGKLFS EKLCEVEEGDKETVNKAVAAARQAFQIGSPWRTMDASERGRLLYKLADLVERDRLILATMESMNGGKLFS APLAKIARGKSVIIDRAMSAARGVFERGJ-WSLSSPAKRKAVLNKLADLMEAHAEELALLETLJTGKPIR NTVCEVSSATTEIVEYAIECADRAFH-DTEWATQDPRERGRLLSKLADELESQIDLVSSIEALDNGKTLA
BA1 HO1 E1 YC1 C1	EKLCEVEEGDKEDVDKAVKAVRQAFQIGSPWRTMDASERGRLLNKLADLIERDHLLLATMEAMNGGKLFS EKLCEVEEGDKEDVDKAVKAVRQAFQIGSPWRTMDASERGRLLYKLADLVERDRLILATMESMNGGKLFS APLAKIARGKSVIIDRAMSAARGVFERGJ-WSLSSPAKRKAVLNKLADLMEAHAEELALLETLDTGKPIR NTVCEVSSATTEIVEYAIECADRAFH-DTEWATODPRERGRLLSKLADELESQIDLVSSIEADNGKTLA EKICEVAEGDKAIIDKAVKAARKAFELGSPWRTMLASERGRLLNKLADLVERDRLTLATMEAIDGGKLFS
BA1 HO1 E1 YC1 C1 M1	EKLCEVEEGDKEDVDKAVKAVRQAFQIGSPWRTMDASERGRLLNKLADLIERDHLLLATMEAMNGGKLFS EKLCEVEEGDKEDVNKAVAAARQAFQIGSPWRTMDASERGRLLYKLADLVERDRLILATMESMNGGKLFS APLAKIARGKSVIIDRAMSAARGVFERGJ-WSLSSPAKRKAVLNKLADLMEAHAEELALLETLDTGKPIR NTVCEVSSATTEIVEYAIECADRAFH-DTEWATQDPRERGRLLSKLADELESQIDLVSSIEALDNGKTLA EKICEVAEGDKAIIDKAVKAARKAFELGSPWRTMLASERGRLLNKLADLVERDRLTLATMEAIDGGKLFS EVICHVEEGDKAIVOKAVKAARQAFQIGSPWRTMLASERGRLLNKLADLMEADRLTLATMEALNGGKVFA
BA1 HO1 E1 YC1 C1 M1 R1	EKLCEVEEGDKEDVDKAVKAVRQAFQIGSPWRTMDASERGRLLNKLADLIERDHLLLATMEAMNGGKLFS EKLCEVEEGDKEDVDKAVKAVRQAFQIGSPWRTMDASERGRLLYKLADLVERDRLILATMESMNGGKLFS APLAKIARGKSVIIDRAMSAARGVFERG-WSLSSPAKRKAVLNKLADLWERDRLILATMESMNGGKLFS NTVCEVSSATTEIVEYAIECADRAFH-DTEWATQDPRERGRLLSKLADELESQIDLVSSIEALDNGKTLA EKICEVAEGDKAIIDKAVKAARKAFELGSPWRTMDASERGRLLNKLADLVERDRLILATMEALDGGKLFS EVICHVEEGDKAIVKAARKAFELGSPWRTMDASERGRLLNKLADLWERDRLLLATMEALNGGKVFA EVICHVEEGDKAIVKAARKAARAFIGIGSPWRTMDASERGCLLNKLADLMERDRLLLATMEALNGGKVFA
BA1 HO1 E1 YC1 C1 M1 R1 S1	EKLCEVEEGDKETVDKAVKAVRQAFQIGSPWRTMDASERGRLLNKLADLIERDHLLLATMEAMNGGKLFS EKLCEVEEGDKETVNKAVAAARQAFQIGSPWRTMDASERGRLLYKLADLVERDRLILATMEAMNGGKLFS APLAKIARGKSVIIDRAMSAARGVFERG-WSLSSPAKRKAVLNKLADLWERDRLILATMEAMNGGKLFS NTVCEVSSATTEIVSYAIECADRAFH-DTEWATQDPRERGRLLSKLADELESQIDLVSSIEALDNGKTLA EKICEVAEGDKADIDKAVKAARKAFELGSPWRTMDASERGRLLNKLADLVERDRLTLATMEATDGGKLFS EVICHVEEGDKADVDKAVKAARQAFQIGSPWRTMDASERGCLLNKLADLMERDRLLLATMEALNGGKVFA EVICHVEEGDKADVDKAVKAARQAFQIGSPWRTMDASERGCLLNKLADLMERDRLLLATMEANNGGKIFT EKLCEVEEGDKADVDKAVKAARQAFQIGSPWRTMDASERGCLLNKLADLTERDRLLLATMEAMNGGKIFT
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1	EKLCEVEEGDKETVDKAVKAVROAFOIGSPWRTMDASERGRLLNKLADLIERDHLLLATMEAMNGGKLFS EKLCEVEEGDKETVNKAVAAAROAFOIGSPWRTMDASERGRLLYKLADLVERDRLILATMEAMNGGKLFS APLAKIARGKSVIIDRAMSAARGVFERG-WSLSSPAKKAVLNKLADLWERDRLILATMESMNGGKLFS NTVCEVSSATTEIVSYAIECADRAFH-DTEWATODPRERGRLLSKLADELESOIDLVSSIEALDNGKTLA EKICEVAEGDKADIDKAVKAARKAFELGSPWRTMDASERGRLLNKLADLVERDRLTLATMEATDGGKLFS EVICHVEEGDKADVDKAVKAAROAFOIGSPWRTMDASERGCLLNKLADLMERDRLLLATMEANNGGKLFS EKICEVEEGDKADVDKAVKAAROAFOIGSPWRTMDASERGCLLNKLADLMERDRLLLATMEANNGGKLFS EKICEVEEGDKADVDKAVKAAROAFOIGSPWRTMDASERGCLLNKLADLIERDRLLLATMEAMNGGKLFS EKICEVEEGDKEDVDKAVKAAROAFOIGSPWRTMDASERGRLLNKLADLIERDRLLLATMEAMNGGKLFS EKICEVEEGDKEDVDKAVKAAROAFOIGSPWRTMDASERGRLLNKLADLIERDRLLLATMEAMNGGKLFS
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1	EKLCEVEEGDKETVDKAVKAVROAFOIGSPWRTMDASERGRLLNKLADLIERDHLLLATMEAMNGGKLFS EKLCEVEEGDKETVNKAVAAAROAFOIGSPWRTMDASERGRLLVKLADLVERDRLILATMEAMNGGKLFS APLAKIARGKSVIIDRAMSAARGVFERG-WSLSSPAKKAVLNKLADLVERDRLILATMESMNGGKLFS NTVCEVSSATTEIVSYAIECADRAFH-DTEWATODPRERGRLLSKLADELESQIDLVSSIEALDNGKTLA EKICEVAEGDKADIDKAVKAARAFELGSPWRTMDASERGRLLNKLADLVERDRLTLATMEATDGGKLFS EVICHVEEGDKADVDKAVKAAROAFOIGSPWRTMDASERGCLLNKLADLMERDRLLLATMEANNGGKLFS EKICEVEEGDKADVDKAVKAAROAFOIGSPWRTMDASERGCLLNKLADLMERDRLLLATMEAMNGGKLFS EKICEVEEGDKADVDKAVKAAROAFOIGSPWRTMDASERGRLLNKLADLIERDRLLLATMEAMNGGKLFS EKICEVEEGDKEDVDKAVKAAROAFOIGSPWRTMDASERGRLLNKLADLIERDRLLLATMEAMNGGKLFS EVICHVEEGDKADVDKAVKAAROAFOIGSPWRTMDASERGRLLNKLADLIERDRLLLATMEAMNGGKLFS EKICEVEEGDKEDVDKAVKAAROAFOIGSPWRTMDASERGRLLNKLADLIERDRLLLATMEAMNGGKLFS EVICHVEEGDKADVDKAVKAAROAFOIGSPWRTMDASERGRLLNKLADLIERDRLLLATMEAMNGGKLFS
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH1 PAIDH 2	EKLCEVEEGDKETVDKAVKAVROAFOIGSPWRTMDASERGRLLNKLAD LIERDHLLLATMEAMNGGKLFS EKLCEVEEGDKETVNKAVAAAROAFOIGSPWRTMDASERGRLLYKLAD LVERDRLILATMEAMNGGKLFS APLAKIARGKSVIIDRAMSAARGVFERG-WSLSSPAKKKAVLNKLAD LWERDRLILATMEAMNGGKLFS APLAKIARGKSVIIDRAMSAARGVFERG-WSLSSPAKKKAVLNKLAD LWERDRLILATMEAMNGGKLFS APLAKIARGKSVIIDRAMSAARGVFERG-WSLSSPAKKKAVLNKLAD LWERDRLILATMEAMNGGKLFS APLAKIARGKSVIIDRAWSAARGAFELGSPWRTMDASERGRLLNKLAD LWERDRLILATMEATDGGKLFS EVICHVEEGDKADVDKAVKAAROAFOIGSPWRTMDASERGRLLNKLAD LVERDRLILATMEATDGGKLFS EVICHVEEGDKADVDKAVKAAROAFOIGSPWRTMDASERGCLLNKLAD LMERDRLLLATMEAMNGGKLFS EKLCEVEEGDKEDVDKAVKAAROAFOIGSPWRTMDASERGRLLNKLAD LIERDRLLLATMEAMNGGKLFS EVICHVEEGDKADVDKAVKAAROAFOIGSPWRTMDASERGRLLNKLAD LIERDRLLLATMEAMNGGKLFS EKLCEVEEGDKEDVDKAVKAAROAFOIGSPWRTMDASERGRLLNKLAD LIERDRLLLATMEAMNGGKLFS EVICHVEEGDKADVDKAVKAAROAFOIGSPWRTMDASERGRLLNKLAD LIERDRLLLATMEAMNGGKLFS EVICHVEEGDKADVDKAVKAAROAFOIGSPWRTMDASERGRLLNKLAD LMERDCLLLATIEATNGGKVFA EVICHVEEGDKADVDKAVKAAROAFOIGSPWRTMDASERGRLLNKLAD LMERDCLLLATIEATNGGKVFA
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH-2	EKLCEVEEGDKETVDKAVKAVRQAFQIGSPWRTMDASERGRLLNKLADLIERDHLLLATMEAMNGGKLFS EKLCEVEEGDKETVNKAVAAARQAFQIGSPWRTMDASERGRLLVKLADLVERDRLILATMESMNGGKLFS APLAKIARGKSVIIDRAMSAARGVFERG-WSLSSPAKKAVLNKLADLVERDRLILATMESMNGGKLFS APLAKIARGKSVIIDRAMSAARGVFERG-WSLSSPAKKAVLNKLADLMEAHAEELALLETLDTGKPIR NTVCEVSSATTEIVSYAIECADRAFH-DTEWATQDPRERGRLLSKLADELESQIDLVSSIEALDNGKTLA EKICEVAEGDKADINKAVKAARAFELGSPWRTMDASERGRLLNKLADLVERDRLILATMEAIDGGKLFS EVICHVEEGDKADVDKAVKAARQAFQIGSPWRTMDASERGCLLNKLADLMERDRLLLATMEALNGGKVFA EVICHVEEGDKADVDKAVKAARQAFQIGSPWRTMDASERGCLLNKLADLMERDRLLLATMEANNGGKLFS EKLCEVEEGDKEDVDKAVKAARQAFQIGSPWRTMDASERGRLLNKLADLIERDRLLLATMEAMNGGKLFS EVICHVEEGDKADVDKAVKAARQAFQIGSPWRTMDASERGRLLNKLADLIERDRLLLATMEAMNGGKLFS EVICHVEEGDKADVDKAVKAARQAFQIGSPWRTMDASERGRLLNKLADLIERDRLLLATMEAMNGGKLFS EVICHVEEGDKADVDKAVKAARQAFQIGSPWRTMDASERGRLLNKLADLIERDRLLLATMEAMNGGKLFS EVICHVEEGDKADVDKAVKAARQAFQIGSPWRTMDASERGRLLNKLADLIERDRLLLATMEAMNGGKLFS EVICHVEEGDKADVDKAVKAARQAFQIGSPWRTMDASERGRLLNKLADLIERDRLLLATMEAMNGGKLFS EVICHVEEGDKADVDKAVKAARQAFQIGSPWRTMDASERGRLLNKLADLIERDRLLLATMESMNGGKLFS EVICHVEEGDKADVDKAVKAARQAFGIGSPWRTMDASERGRLLNKLADLMERDCLLLATIEAINGGKVFA EVICHVEEGDKADVDKAVKAARQAFGIGSPWRTMDASERGRLLNKLADLMERDCLLLATIEAINGGKVFA
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2	SKLCEVEEGDKETVDKAVKAVRQAFQIGSPWRTMDASERGRLLNKLAD LIERDHLLLATMEAMNGGKLFS SKLCEVEEGDKETVNKAVAAARQAFQIGSPWRTMDASERGRLLNKLAD LVERDRLILATMESMNGGKLFS APLAK IARGKSVIIDRAMSAARGVFERG - WSLSSPAKKAVLNKLAD LVERDRLILATMESMNGGKLFS APLAK IARGKSVIIDRAMSAARGVFERG - WSLSSPAKKAVLNKLAD LVERDRLILATMESMNGGKLFS APLAK IARGKSVIIDRAMSAARGVFERG - WSLSSPAKKAVLNKLAD LVERDRLILATMESINGGKLFS APLAK IARGKSVIIDRAMSAARGVFERG - WSLSSPAKKAVLNKLAD LVERDRLILATMESINGGKLFS APLAK IARGKSVIIDRAMSAARGVFERG - WSLSSPAKKAVLNKLAD LVERDRLILATMESINGGKLFS EVICEVEEGDKAD VOKAVKAARQAFQIGSPWRTMDASERGRLLNKLAD LVERDRLILATMEALNGGKVFA EVICEVEEGDKAD VOKAVKAARQAFQIGSPWRTMDASERGCLLNKLAD LMERDRULLATMESMNGGKLFS EVICEVEEGDKED VOKAVKAARQAFQIGSPWRTMDASERGRLLNKLAD LIERDRULLATMESMNGGKLFS EELCQVEEGDKED VOKAVKAARQAFQIGSPWRTMDASERGRLLNKLAD LIERDRULLATMESMNGGKLFS EVICEVEEGDKAD VOKAVKAARQAFQIGSPWRTMDASERGRULVKLAD LIERDRULLATMESMNGGKLFS EVICEVEEGDKAD VOKAVKAARQAFQIGSPWRTMDASERGRULVKLAD LIERDRULLATMESMNGGKLFS EVICEVEEGDKAD VOKAVKAARQAFQIGSPWRTMDASERGRULVKLAD LIERDRULLATMESMNGGKLFS EVICEVEEGDKAD VOKAVKAARQAFQIGSPWRTMDASERGRULVKLAD LIERDRULLATMESMNGGKLFS EVICEVEEGDKAD VOKAVKAARQAFQIGSPWRTMDASERGRULVKLAD LIERDRULLATMESMNGGKLFS EVICEVEEGDKAD VOKAVKAARQAFGIGSPWRTMDASERGRULVKLAD LIERDRULLATMESMNGGKLFS EVICEVEEGDKAD VOKAVKAARQAFGIGSPWRTMDASERGRULVKLAD LIERDRULLATMESMNGGKLFS EVICEVEEGDKAD VOKAVKAARQAFGIGSPWRTMDASERGRULVKLAD LMERDCULLATIEAINGGKVFA EVICEVEEGDKAD VOKAVKAARQAFGIGSPWRTMDASERGRULVKLAD LWERDRULLATMESINGGKFFL EVICEVEEDAKAV JAKAARAFFOIGSPWRTMDASERGRULVKLAD LVERDRATLATMESINGGKFFL EVICEVEEDAKAV JAKAARAFFOIGSPWRTMDASERGRULVKLAD LVERDRATLATMESINGGKFFL EVICEVEEDAKAV JIDKAVKAARAFFOIGSPWRTMDASERGRULVKLAD LVERDRATLATMESINGGKFFL EVICEVEEDAKAV JIDKAVKAARAFFOIGSPWRTMDASERGRULVKLAD LVERDRATLATMESINGGKFFL EVICEVEEDAKAV JIDKAVKAARAFFOIGSPWRTMDASERGRULVKLAD LVERDRATLATMESINGGKFFL
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2	EKLCEVEEGDKETVDKAVKAVROAFOIGSPWRTMDASERGRLLNKLAD LIERDHLLLATMEAMNGGKLFS BKLCEVEEGDKETVNKAVAAAROAFOIGSPWRTMDASERGRLLVKLAD LVERDRLILATMESMNGGKLFS APLAK IARGKSVIIDRAMSAARGVFERG - WSLSSPAKKAVLNKLAD LVERDRLILATMESMNGGKLFS APLAK IARGKSVIIDRAMSAARGVFERG - WSLSSPAKKAVLNKLAD LVERDRLILATMESMNGGKLFS APLAK IARGKSVIIDRAMSAARGVFERG - WSLSSPAKKAVLNKLAD LVERDRLILATMESMNGGKLFS APLAK IARGKSVIIDRAMSAARGVFERG - WSLSSPAKKAVLNKLAD LVERDRLILATMESMNGGKLFS APLAK IARGKSVIIDRAMSAARGVFERG - WSLSSPAKRKAVLNKLAD LVERDRLILATMESMNGGKLFS EVICEVEEGDKAD VDKAVKAAROAFOIGSPWRTMDASERGRLLNKLAD LVERDRLILATMEALNGGKVFA EVICEVEEGDKAD VDKAVKAAROAFOIGSPWRTMDASERGRLLNKLAD LMERDRULLATMESMNGGKLFS EELCOVEEGDKED VDKAVKAAROAFOIGSPWRTMDASERGRLLNKLAD LIERDRULLATMESMNGGKLFS EVICEVEEGDKED VDKAVKAAROAFOIGSPWRTMDASERGRLLNKLAD LIERDRULLATMESMNGGKLFS EVICEVEEGDKED VDKAVKAAROAFOIGSPWRTMDASERGRLLNKLAD LIERDRULLATMESMNGGKLFS EVICEVEEGDKED VDKAVKAAROAFOIGSPWRTMDASERGRULNKLAD LIERDRULLATMESMNGGKLFS EVICEVEEDKAD VDKAVKAAROAFOIGSPWRTMDASERGRULNKLAD LIERDRULLATIEAINGGKFFL EVICEVEEDKAD VDKAVKAAROAFOIGSPWRTMDASERGRUNKLAD LIERDRULLATIEAINGGKFFL EVICEVEEDKAD VDKAVKAAROAFOIGSPWRTMDASERGRUNKLAD LIERDRULATIEAINGGKFFL EVICEVEEDKAD VDKAVKAARAAFOLGSPWRTMDASERGRUNKLAD LIERDRULATIEAINGSKFFL EVICEVEEDKAD VDRAVKAARAAFOLGSPWRTMDASERGRUNKLAD LIERDRULATIEATHESLNGGKFFL EVICEVEEDKAD VDRAVKAARAAFOLGSPWRTMDASERGRUNKLAD LIERDRULATIEATUESLNGGKFFL EVICEVEEDKAD VDRAVKAARAAFOLGSPWRTMDASERGRUNKLAD LIERDRULAATIETTELDNGKYFI
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH-2 B2 HO2 YM1	SKLCEVEEGDKETVDKAVKAVROAFOIGSPWRTMDASERGRLLNKLAD LIERDHLLLATMEAMNGGKLFS SKLCEVEEGDKETVNKAVAAAROAFOIGSPWRTMDASERGRLLYKLAD LVERDRLILATMESMNGGKLFS APLAK IARGKSVIIDRAMSAARGVFERG - WELSSPAKKKAVLNKLAD LVERDRLILATMESMNGGKLFS APLAK IARGKSVIIDRAMSAARGVFERG - WELSSPAKKKAVLNKLAD LVERDRLILATMESMNGGKLFS APLAK IARGKSVIIDRAMSAARGVFERG - WELSSPAKKKAVLNKLAD LVERDRLILATMESMNGGKLFS AVICEVEEGDKAD VOYAIECADRAFH - DTEWATOD PRERGRLLSKLADELESQIDLVSSIEALDNGKTLA SKICEVAEGDKAD VOYAVKAARAAFELGSPWRTMDASERGRLLNKLAD LVERDRLILATMEAL OGGKLFS EVICHVEEGDKAD VOKAVKAAROAFOIGSPWRTMDASERGCLLNKLAD LMERDRLLLATMEALNGGKVFA EVICHVEEGDKAD VOKAVKAAROAFOIGSPWRTMDASERGCLLNKLAD LMERDRULLATMEAMNGGKLFS SELCOVEEGDKED VDKAVKAAROAFOIGSPWRTMDASERGRLLNKLAD LIERDRULLATMEAMNGGKLFS SELCOVEEGDKED VDKAVKAAROAFOIGSPWRTMDASERGRLLNKLAD LIERDRULLATMEAMNGGKLFS SEVICHVEEGDKAD VDKAVKAAROAFOIGSPWRTMDASERGRLLNKLAD LIERDRULLATMEAMNGGKLFS SEVICHVEEGDKAD VDKAVKAAROAFOIGSPWRTMDASERGRLLNKLAD LIERDRULLATMEAMNGGKLFS SEVICHVEEGDKAD VDKAVKAAROAFOIGSPWRTMDASERGRLLNKLAD LIERDRULLATMESMNGGKLFS SEVICHVEEGDKAD VDKAVKAAROAFOIGSPWRTMDASERGRLLNKLAD LIERDRULLATMESMNGGKLFS SEVICHVEEGDKAD VDKAVKAAROAFOIGSPWRTMDASERGRLLNKLAD LIERDRULLATMESMNGGKLFS SEVICHVEEGDKAD VDKAVKAAROAFOIGSPWRTMDASERGRLLNKLAD LIERDRULLATMESMNGGKLFS SEVICHVEEGDKAD VDKAVKAAROAFOIGSPWRTMDASERGRLLNKLAD LIERDRULLATMESMNGGKLFS SEVICHVEEGDKAD VDKAVKAAROAFOIGSPWRTMDASERGRLLNKLAD LIERDRULLATIEANNGGKLFS SEVICHVEEGDKAD VDKAVKAAROAFOIGSPWRTMDASERGRLLNKLAD LIERDRULLATIEANNGGKLFS SEVICHVEEGDKAD VDKAVKAAROAFOIGSPWRTMDASERGRLLNKLAD LIERDRULLATIEANNGGKLFS SEVICHVEEGDKAD VDKAVKAAROAFOIGSPWRTMDASERGRLLNKLAD LIERDRULLATIEANNGGKLFS SEVICHVEEGDKAD VDKAVKAAROAFOIGSPWRTMDASERGRLLNKLAD LIERDRULATIEATUESLNGGKPFL SEVICHVEEGDKAD VDRAVKAARAAFOLGSPWRTMDASERGRLLNKLAD LIERDRULATIEATUESLNGGKPFL SEVICEVAAGDKEDVDRAVKAARAAFOLGSPWRTMDASERGRLLNRLAD LIERDRULAALETLDNGKPYV SKITTVYKAMEDIVDEAVAALKKLLK-RSVYCRAGVRAKALFNLAD LVEKHOETLAAIESMDOKSL-
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2 YM1 M2	EKLCEVEEGDKEDVDKAVKAVRQAFQIGSPWRTMDASERGRLLNKLADLIERDHLLATMEAMNGGKLFS EKLCEVEEGDKEDVDKAVKAVRQAFQIGSPWRTMDASERGRLLVKLADLVERDRLILATMEAMNGGKLFS APLAKIARGKSVDIDRAMSAARGVFERGD-WSLSSPAKRKAVLNKLADLVERDRLILATMESMNGGKLFS APLAKIARGKSVDIDRAMSAARGVFERGD-WSLSSPAKRKAVLNKLADLVERDRLILATMESMNGGKLFS APLAKIARGKSVDIDRAMSAARGVFERGD-WSLSSPAKRKAVLNKLADLVERDRLILATMESMNGGKLFS EVICEVSSATTEDVEYAIECADRAFH-DTEWATQDPRERGRLLSKLADELESQIDLVSSIEALDDKTLA EKICEVAEGDKADIDKAVKAARKAFELGSPWRTMDASERGRLLNKLADLVERDRLILATMEAIDGGKLFS EVICHVEEGDKADVDKAVKAARQAFQIGSPWRTMDASERGRLLNKLADLVERDRLLLATMEAINGGKVFA EVICHVEEGDKADVDKAVKAARQAFQIGSPWRTMDASERGRLLNKLADLMERDRLLLATMEAMNGGKLFS EELCQVEEGDKEDVDKAVKAARQAFQIGSPWRTMDASERGRLLNKLADLIERDRLLLATMEAMNGGKLFS EVICHVEEGDKADVDKAVKAARQAFQIGSPWRTMDASERGRLLNKLADLIERDRLLLATMEAMNGGKLFS EVICHVEEGDKADVDKAVKAARQAFQIGSPWRTMDASERGRLLNKLADLIERDRLLLATMESMNGGKLFS EVICHVEEGDKADVDKAVKAARQAFQIGSPWRTMDASERGRLLNKLADLIERDRLLLATMESMNGGKLFS EVICHVEEGDKADVDKAVKAARQAFQIGSPWRTMDASERGRLLNKLADLWERDRLLLATMESMNGGKLFS EVICHVEEGDKADVDKAVKAARQAFQIGSPWRTMDASERGRLLNKLADLWERDRLLLATMESMNGGKLFS EVICHVEEGDKADVDKAVKAARQAFQIGSPWRTMDASERGRLLNKLADLWERDRLLLATMESMNGGKLFS EVICHVEEGDKADVDKAVKAARQAFQIGSPWRTMDASERGRLLNKLADLWERDRLLLATMESNGGKFFL EVICHVEEGDKADVDKAVKAARQAFQIGSPWRTMDASERGRLLNKLADLWERDRLLLATMESLNGGKFFL EVICHVEEGDKADVDKAVKAARQAFQIGSPWRTMDASERGRLLDKLADLVERDRATLATMESLNGGKFFL EVICHVEEGDKADVDKAVKAARAAFGIGSPWRTMDASERGRLLDKLADLVERDRATLATMESLNGGKFFL EVICHVEEGDKADVDRAVKAARAAFGIGSPWRTMDASERGRLLDKLADLVERDRATLATMESLNGGKFFL EVICHVEEGDKADVDRAVKAARAAFGIGSPWRTMDASERGRLLDKLADLVERDRATLATMESLNGGKFFL EVICHVEEGDKADVDRAVKAARAAFGIGSPWRTMASERGRLLDKLADLVERDRATLATMESLNGGKFFL EVICEVAAGDKEDVDRAVKAARAAFGIGSPWRTMASERGRLLDKLADLVERDRATLATMESLNGGKFFL EVICEVAAGDKEDVDRAVKAARAAFGIGSPWRTMASERGRLLDKLADLVERDRATLATMESLNGGKFFL EVICEVAAGDKEDVDRAVKAARAAFGIGSPWRTMASERGRLLDKLADLVERDRATLAALETLDNGKYVE EKITTVYKAMEDIVDRAVAALKKLK-RSVYCRAGVRAKALFNLADLVEKHQETLAALETLDNGKYVE
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH-2 B2 HO2 YM1 M2 R2	EKLCEVEEGIKEIVJKAVKAVROAFOIGSPWRTMIASERGRLLNKLADLIERDRLLLATMEAMNGGKLFS EKLCEVEEGIKEIVJKAVKAAROAFOIGSPWRTMIASERGRLLYKLADLIERDRLLLATMEAMNGGKLFS EKLCEVEEGIKEIVJKAVKAAROAFOIGSPWRTMIASERGRLLYKLADLVERDRLLLATMEAMNGGKLFS APLAKIARGKSVIIDRAMSAARGVFERGD-WELSSPAKKAVLNKLADLMEAHAEELALETLDTGKPIR NTVCEVSSATTEDVEYAIECADRAFH-DTEWATODPREGRLJSKLADELESQIILVSSIEALDDGKTLA SKICEVAEGIKADVDKAVKAAROAFOIGSPWRTMIASERGRLLNKLALIVERDRLLLATMEAIDGGKVFA EVICHVEEGIKADVDKAVKAAROAFOIGSPWRTMIASERGRLLNKLALIVERDRLLLATMEAINGGKVFA EVICHVEEGIKADVDKAVKAAROAFOIGSPWRTMIASERGRLLNKLADLIERDRLLLATMEAMNGGKLFS SELCOVEEGIKADVDKAVKAAROAFOIGSPWRTMIASERGRLLNKLADLIERDRLLLATMEAMNGGKLFS EVICHVEEGIKADVDKAVKAAROAFOIGSPWRTMIASERGRLLNKLADLIERDRLLLATMEAMNGGKLFS SELCOVEEGIKADVDKAVKAAROAFOIGSPWRTMIASERGRLLNKLADLIERDRLLLATMEAMNGGKLFS EVICHVEEGIKADVDKAVKAAROAFOIGSPWRTMIASERGRLLNKLADLIERDRLLLATMESMNGGKLFS SVICHVEEGIKADVDKAVKAAROAFOIGSPWRTMIASERGRLLNKLADLIERDRLLLATMESMNGGKLFS EVICHVEEGIKADVDKAVKAAROAFOIGSPWRTMIASERGRLLNKLADLIERDRLLLATMESMNGGKLFS SVICHVEEGIKADVDKAVKAAROAFOIGSPWRTMIASERGRLLNKLADLIERDRLLLATMESLNGGKVFA EVICHVEEGIKADVDKAVKAAROAFOIGSPWRTMIASERGRLLNKLADLVERDRATLATMESLNGGKVFA EVICHVEEGIKADVDKAVKAAROAFOIGSPWRTMIASERGRLLNKLADLVERDRATLATMESLNGGKVFA EVICHVEEGIKADVDRAVKAAROAFOIGSPWRTMIASERGRLLNKLADLVERDRATLATMESLNGGKVFA EVICHVEEGIKADVDRAVKAAROAFOIGSPWRTMIASERGRLLNKLADLVERDRATLATMESLNGGKVFA EVICHVEEGIKADVDRAVKAAROAFOIGSPWRTMIASERGRLLNKLADLVERDRATLATMESLNGGKVFA EVICUVAAGIKEDVDRAVKAARAAFOLGSPWRMIASERGRLLNKLADLIERDRTYLAALETLDNGKPYV EKITTVYKAMEDOVDEAVAALKKLK-RSVYCRAGVRAKALFNLADLIERDRTYLAALETLDNGKPYV EKITTVYKAMEDOVDEAVAALKKLKK-RSVYCRAGVRAKALFNLADLIERDRTYLAALETLDNGKPYV EKITTVYKAMEDOVDEAVAALKKLKK-RSVYCRAGVRAKALFNLADLIERDRTYLAALETLDNGKPYV EVICQVAAGNKEDVKAVKAARAAFOLGSPWRMIASDRGRLLYRLADLIERDRTYLAALETLDNGKPYV
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2 YM1 M2 R2 H2	Image: Construction of the second
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH-2 B2 HO2 YM1 M2 R2 H2 CRY	Image: Construction of the system of the
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2 YM1 M2 R2 H2 CRY H6	EXICE VEEGXET VOKAVKAVROAFOIGS PWRTMIASERGELLNKLAD LI ERDHLLATMEAMNGGKLFS EXICE VEEGXET VOKAVKAVROAFOIGS PWRTMIASERGELLYKLAD LVCERDELILATMESMNGGKLFS APLAKTARGKS VI IDRAMSAAROAFOIGS PWRTMIASERGELLYKLAD LVCERDELILATMESMNGGKLFS APLAKTARGKS VI IDRAMSAAROAFOIGS PWRTMIASERGELLYKLAD LOSOI LVSSI EALDNGKTLA EXICE VESATTED VEYAI CADRAFH - DTEWATOD PRERGELLSKLADE LOSOI LVSSI EALDNGKTLA EXICE VESATTED VEYAI CADRAFH - DTEWATOD PRERGELLSKLADE LOSOI LVSSI EALDNGKTLA EXICE VESATTED VEYAI CADRAFH - DTEWATOD PRERGELLSKLADE LOSOI LVSSI EALDNGKTLA EXICE VESATTED VEYAI CADRAFH - DTEWATOD PRERGELLSKLADE LOSOI LVSSI EALDNGKTLA EXICE VESATTED VEYAI CADRAFH - DTEWATOD PRERGELLSKLADE LOSOI LVSSI EALDNGKTLA EXICE VESATTED VEYAI CADRAFH - DTEWATOD PRERGELLSKLADE LOSOI LVSSI EALDNGKTEA EXICE VESATTED VEYAI CADRAFH - DTEWATOD PRERGELLSKLADE LOSOI LVSSI EALDNGKTFA EVIC HVEEG CKAD VOKAVKAAROAFOIGS PWRTMIASERGELLNKLAD LVERDRULLATMESMNGGKLFS ELCOVERG KED VOKAVKAAROAFOIGS PWRTMIASERGELLNKLAD LIERDRULLATMESMNGGKLFS ELCOVERG KED VOKAVKAAROAFOIG SPWRTMIASERGELLNKLAD LIERDRULLATI EAINGGKVFA EVIC HVEEG CKAD VOKAVKAAROAFOIG SPWRTMIASERGELLNKLAD LIMER DRULLATI EAINGGKVFA EVIC HVEEG CKAD VOKAVKAAROAFOIG SPWRTMIASERGELLNKLAD LMERDRULLATI EAINGGKVFA EVIC HVEEG CKAD VOKAVKAAROAFOIG SPWRTMIASERGELLNKLAD LMERDRULLATI EAINGGKVFA EVIC HVEEG CKAD VOKAVKAAROAFOIG SPWRTMIASERGELLNKLAD LWERDRATLAAMESLNGGKPFL EVVC EVQEADKVI I KAVGAARDAFOLGS PWRTMIASERGELLNKLAD LWERDRATLAAMESLNGGKPFL EVVC EVQEADKVI I KAVGAARDAFOLGS PWRRMIASERGELLNKLAD LVERDRATLAAMESLNGGKPFL VIC OVAAGUKED VOKAVKAARAAFOLGS PWRRMIASERGELLNKLAD LIERDRTYLAALETLDNGK PYV EXIT TVYKAMEDIV DEAVAALKKLLK - RSVYCRAC - VRAKALFNLAD LVERDRATLAALETLDNGK PYV EXIT TVYKAMEDIV DEAVAALKKLLK - RSVYCRAC - VRAKALFNLAD LVERDRTYLAALETLDNGK PYV EXIT CVAAGUKED VOKAVKAARAAFOLGS PWRRMIASDRGRLLYRLAD LIERDRTYLAALETLDNGK PYV EXIT CVAAGUKED VOKAVKAARAAFOLGS PWRRMIASDRGRLLYRLAD LIERDRTYLAALETLDNGK PYV EXIT CVAAGUKED VOKAVKAARAAFOLGS PWRRMIASDRGRLLYRLAD LIERDRTYLAALETLDNGK PYV EXIT CVAAGUKED VOKAVKAARAFOLGS PW
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2 YM1 M2 R2 HO2 YM1 M2 R2 H2 CRY H6 HX	EXICE VEEGOKED VDKAVKAVROAFOIGS PWRTMIASERGELLNKLADLIERDHLLATMEAMNGGKLFS EXICE VEEGOKED VDKAVKAVROAFOIGS PWRTMIASERGELLYKLADLVEEDELIATMEAMNGGKLFS APLAKIARGKSVDIDRAMSAARGVFERGO-WELSS PAKRKAVLNKLADLWEAHAEELALLETLDTGKPIR NTVCEVSSATTEDVEYAIECADRAFH-DTEWATODPRERGELLSKLADELESOIDLVSSIEALDNGKTLA EXICE VASGUKADIDKAVKAARGAFOIGS PWRTMIASERGELLNKLADLWEAHAEELALLETLDTGKPIR NTVCEVSSATTEDVEYAIECADRAFH-DTEWATODPRERGELLSKLADELESOIDLVSSIEALDNGKTLA EXICE VASGUKADIDKAVKAARGAFOIGS PWRTMIASERGELLNKLADLWERDRVLLATMEALNGGKVFA EVICHVEEGOKADVDKAVKAAROAFOIGS PWRTMIASERGELLNKLADLMERDRVLLATMEAMNGGKLFS EVICHVEEGOKADVDKAVKAAROAFOIGS PWRTMIASERGELLNKLADLMERDRVLLATMEAMNGGKLFS EVICHVEEGOKADVDKAVKAAROAFOIGS PWRTMIASERGELLNKLADLMERDRVLLATMESMNGGKLFS EVICHVEEGOKADVDKAVKAAROAFOIGS PWRTMIASERGELLNKLADLMERDRVLLATMESMNGGKLFS EVICHVEEGOKADVDKAVKAAROAFOIGS PWRTMIASERGELLNKLADLMERDRVLLATMESMNGGKLFS EVICHVEEGOKADVDKAVKAAROAFOIGS PWRTMIASERGELLNKLADLMERDRLLLATMESMNGGKLFS EVICHVEEGOKADVDKAVKAAROAFOIGS PWRTMIASERGELLNKLADLMERDRLLLATMESMNGGKLFS EVICHVEEGOKADVDKAVKAAROAFOIGS PWRTMIASERGELLNKLADLMERDRLLLATMESMNGGKLFS EVICHVEEGOKADVDKAVKAAROAFOIGS PWRTMIASERGELLNKLADLMERDRLLLATIEANNGGKVFA EVICHVEEGOKADVDKAVKAAROAFOIGS PWRTMIASERGELLNKLADLMERDRLLLATIEANNGGKVFA EVICHVEEGOKADVDKAVKAAROAFOIGS PWRTMIASERGELLNKLADLVERDRATLATMESLNGGKPFL EVICHVEEGOKADVDKAVKAARAAFOLGS PWRTMIASERGELLDKLADLVERDRATLATMESLNGGKPFL EVICUVEAGKAAVDRAVKAARAAFOLGS PWRTMIASERGELLDKLADLVERDRATLATMESLNGGKPFL EVICUVEAGKAAVDRAVKAARAAFOLGS PWRTMIASERGELLNKLADLVERDRATLATMESLNGGKPFYV EXICTVEAGOKEDVDKAVKAARAAFOLGS PWRTMIASDRGELLNKLADLVERDRATLAALETLDNGKPYV EXICTVEAGOKEDVDKAVKAARAAFOLGS PWRTMIASDRGELLNKLADLIERDRTYLAALETLDNGKPYV EXICTVEAGOKEDVDKAVKAARAAFOLGS PWRTMIASDRGELLNKLADLIERDRTYLAALETLDNGKPYV EXICTVEAGOKEDVDKAVKAARAAFOLGS PWRTMIASDRGELLNKLADLIERDRTYLAALETLDNGKPYV EXICTVEAGOKEDVDKAVKAARAAFOLGS PWRTMIASDRGELLNKLADLIERDRTYLAALETLDNGKPYV EXICTVEAGOKEDVDKAVKAARAAFOLGS PWRTMIASDRGELLNKLADLIERDRTYLAALETLDNGKPYV EXICTVEAGOKEDVDKAVKAARAAFOLGS PWRTMIASDRGELLNKLADLIERDRTYLAALETLDNGKPYV E
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2 YM1 M2 R2 HO2 YM1 M2 R2 H2 CRY H6 HX YM3	Image: Construction of the second
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2 YM1 M2 R2 HO2 YM1 M2 R2 H2 CRY H6 HX YM3 M3	Image: Construction of the second
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2 YM1 M2 R2 HO2 YM1 H2 R2 H2 H2 R2 H2 H2 R2 H2 H2 R2 H2 H2 R2 H2 H2 R2 H2 H2 R2 H2 H2 R2 H2 H2 R2 H2 H2 H2 H2 R2 H2 H2 H2 H2 H2 H2 H2 H2 H2 H2 H2 H2 H2	Image: Construction of the second of the
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2 YM1 M2 R2 HO2 YM1 M2 R2 HO2 YM1 M2 R2 HO2 YM1 M3 R3 H3 H3	Image: Contract of the second seco
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2 YM1 M2 R2 HO2 YM1 M2 R2 H2 CRY H6 HX YM3 M3 R3 H3 MM PM	V EKLCEVERCOKEDVDKAVKAVROAFOICSPWRTM AS ERGRLLNKLALLI ERDHLLLATM AMNGCKLFS EKLCEVERCOKEDVDKAVKAVROAFOICSPWRTM AS ERGRLLYKLADLVERDRLLATMESMNGCKLFS APLAKIARGSVI I DRAMSAARGYERCO-WSLSSPAKRKAVLMKLADLM ARAFELALLETLDTCKFIR NTVCEVSSATTEDVEYAI CADRAFH - TEWATOD PRENGRLLSKLADELESGI I LVSI I ALDNOKKLFS APLAKIARGSVI I DRAMSAARGYERCO-WSLSSPAKRKAVLMKLADLM ARAFELALLETLDTCKFIR NTVCEVSSATTEDVEYAI CADRAFH - TEWATOD PRENGRLLSKLADELESGI I LVSI I ALDNOKKLFS EVICHVERCOKADVDRAVKAAROAFOICSPWRTM AS ERGRLLNKLADLMERDRULLATMESMNACKIFS EVICHVERCOKADVDRAVKAAROAFOICSPWRTM AS ERGRLLNKLADLMERDRULLATMESMNACKIFS ELCOVELCOKEDVDRAVKAAROAFOICSPWRTM AS ERGRLLNKLADLI ERDRULLATMESMNACKIFS ELCOVELCOKEDVDRAVKAAROAFOICSPWRTM AS ERGRLLNKLADLMERDRULLATMESMNACKIFS ELCOVELCOKEDVDRAVKAAROAFOICSPWRTM AS ERGRLLNKLADLMERDRULLATMESMNACKIFS ELCOVELCOKEDVDRAVKAAROAFOICSPWRTM AS ERGRLLNKLADLMERDRULLATMESMNACKIFS ELCOVELCOKEDVDRAVKAAROAFOICSPWRTM AS ERGRLLNKLADLMERDRULLATMESMNACKIFS EVICHVELGOKADVDRAVKAAROAFOICSPWRTM AS ERGRLLNKLADLMERDRULLATIESINGGKVFA VICHVELGOKADVDRAVKAAROAFOICSPWRTM AS ERGRLLDKLADLWERDCLLLATIEAINGGKVFA EVICHVELGOKADVDRAVKAAROAFOICSPWRTM AS ERGRLLDKLADLWERDRALLATIESLNGGKPFL EVICEVOLADKVI DRAVAARAAFOICSPWRTM AS ERGRLLDKLADLWERDRALLATHESLNGGKPFL VICHVELGOKADVDRAVKAARAAFOLGSPWRRM AS ERGRLLDKLADLVERDRATLATMESLNGGKPFL VICHVELGOKADVDRAVKAARAAFOLGSPWRRM AS ERGRLLDKLADLVERDRATLATMESLNGGKPFL VICOVAGKEDVDRAVKAARAAFOLGSPWRRM AS ERGRLLDKLADLVERDRATLATHESLNGGKPFL VICOVAGKEDVDRAVKAARAAFOLGSPWRRM AS DRGRLLDKLADLIERDRIVLAALETLDNGKPYV VKICOVAGKEDVDRAVKAARAAFOLGSPWRRM AS DRGRLLYKLADLIERDRIVLAALETLDNGKPYV VKICOVAGKEDVDRAVKAARAAFOLGSPWRRM AS DRGRLLYKLADLIERDRIVLAALETLDNGKPYV VKICOVAGKEDVDRAVKAARAAFOLGSPWRRM AS DRGRLLYKLADLIERDRIVLAALETLDNGKPYV VKICOVAGKEDVDRAVKAARAAFOLGSPWRRM AS DRGRLLYKLADLIERDRIVLAALETLDNGKPYV VKICOVAGKKEDVDRAVKAARAAFOLGSPWRRM AS DRGRLLYKLADLIERDRIVLAALETLDNGKYFY EKITTYYKAMDDIVDAVAARAAFOLGSPWRRM AS DRGRLLYKLADLIERDRIVLAALETLDNGKYFY EKITTYKAMDDIVDAVAARAFAFOLGSPWRRM AS DRGRLLYKLADLIERDRIVLAALETLDNGKYFY EKITTYKAMDDIVDAVAA
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2 YM1 M2 R2 HO2 YM1 M2 R2 H2 CRY H6 HX YM3 M3 R3 H3 MM RM H4	EX LCE VEE CDK ELVDKAVKAVROAFOICS PWRTM AS ERGR LLVKLA LI ERDHLLLATM EAMNGCKLFS EK LCE VEE CDK ELVDKAVKAVROAFOICS PWRTM AS ERGR LLVKLA LVER BRLILALETL DTCKPIR DK LCE VEE CDK ELVDKAVKAVROAFOICS PWRTM AS ERGR LLVKLA LVER BRLILALETL DTCKPIR NTVCE VES ATTET VEYAI ECA BRAFF DTEWATOD PRERGRLDSKLA DELSOID LVS I EALDNCKTLA EK ICE VAS GDKA I DKAVKAARCAFOICS PWRTM AS ERGRLLNKLA LVER DRLLLATMEAI CGKLFS EVICHVEE CDKA VVKAVKAARCAFOICS PWRTM AS ERGRLLNKLA LVER DRLLLATMEAI CGKLFS EVICHVEE CDKA VVKAVKAARCAFOICS PWRTM AS ERGRLLNKLA LVER DRLLLATMEAINGCKLFS EVICHVEE CDKA VVKAVKAARCAFOICS PWRTM AS ERGRLLNKLA LMER DVLLATMEAN AGKIFT EKLCE VEE CDKE VVKAVKAARCAFOICS PWRTM AS ERGRLLNKLA LMER DVLLATMEANNGCKLFS ELCOVEE CDKA VVKAVKAARCAFOICS PWRTM AS ERGRLLNKLA LMER DVLLATMEANNGCKLFS ELCOVEE CDKA VVKAVKAARCAFOICS PWRTM AS ERGRLLNKLA LMER DVLLATMEANNGCKLFS ELCOVEE CDKA VVKAVKAARCAFOICS PWRTM AS ERGRLLNKLA LMER DVLLATMEANNGCKLFS EVICHVEE CDKA VVKAVKAARCAFOICS PWRTM AS ERGRLLNKLA LMER DVLLATMEANNGCKLFS EVICHVEE CDKA VVKAVKAARCAFOICS PWRTM AS ERGRLLNKLA LMER DRLLLATMESNNGCKLFS EVICHVEE CDKA VVKAVKAARCAFOICS PWRTM AS ERGRLLNKLA LMER DRLLLATMESNNGCKLFS EVICHVEE CDKA VVKAVKAARCAFOICS PWRTM AS ERGRLLNKLA LMER DRLLLATMESLNGCK PFL EVICHVEE CDKA VVKAVKAARCAFOICS PWRTM AS ERGRLLNKLA LVER DRATLATMESLNGCK PFL EVIC VOE ANKVIIKAVKAARAAFOICS PWRTM AS ERGRLLNKLA LVER DRATLATMESLNGCK PFL EVIC VVAACKKEUV KAVKAARAAFOICS PWRTM AS ERGRLLNKLA LU ERDRTYLAALETLDNCK PYV EKITTVYKAMEDIV FAVAARAAFOICS PWRTM AS ERGRLLNKLA LI ERDRTYLAALETLDNCK PYV EKIT EVEAAKEUV KAVKAARAAFOICS PWRTM AS ERGRLLNKLA LI ERD
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2 YM1 M2 R2 HO2 YM1 M2 R2 H2 CRY H6 HX YM3 M3 R3 H3 MM RM H4 H9	Image: Construction of the second
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2 YM1 M2 R2 HO2 YM1 M2 R2 H2 CRY H6 HX YM3 M3 R3 H3 MM RM H4 H9 H7	EXICE VDEGUKED VDRAVKAVROAFOIGS PWRTM AS ERGELLNILAL LI ERDHLLLAT MEAMNGGKLFS EXICE VDEGUKED VDRAVKAVROAFOIGS PWRTM AS ERGELLNILAL LI ERDHLLLAT MEAMNGGKLFS EXICE VDEGUKED VDRAVRAVAAR GAFOIGS PWRTM AS ERGELLNILAL DWEARA DE LALLET LOTGKPIR NTVCE VVS SATTED VDYAIEGA FARH-DDE NEWEL SATAULNILAD LIVER DRILLATING SANGGKLFS EVICEVDEGUKAD VDRAVKAAR GAFOIGS PWRTM AS ERGELLNILAT LIVER DRILLAT MEALNGGKLFS EVICEVDEGUKAD VDRAVKAAR GAFOIGS PWRTM AS ERGELLNILAT LIVER DRILLATING SANGGKLFS EVICEVDEGUKAD VDRAVKAAR GAFOIGS PWRTM AS ERGELLNILAT LIMER DRILLATING SANGGK LFS EVICEVDEGUKAD VDRAVKAAR GAFOIGS PWRTM AS ERGELLNILAT LIMER DRILLATING SANGGK LFS EVICEVDEGUKAD VDRAVKAAR GAFOIGS PWRTM AS ERGELLNILAT LIMER DRILLATING SANGGK LFS EVICEVDEGUKAD VDRAVKAAR GAFOIGS PWRTM AS ERGELLNILAT LIMER DRILLATING SANGGK LFS EVICEVDEGUKAD VDRAVKAAR GAFOIGS PWRTM AS ERGELLNILAT LIMER DRILLATING SANGGK PFL EVICE VDEADINTID KANDAAR LAFS LGG VWRRM AS ERGELLNILAT LIVER DRITIALAT LIDING SANGGK PFL EVICEVAS GARGET VDRAVKAAR RAAFOLGS PWRRM AS ERGELLNILAT LIVER DRITYLAAL ETLDNG KPYV KITTYYKAMEDD VDRAVAAR RAAFOLGS PWRRM AS DRGRILLNILAT LIVER DRITYLAAL ETLDNGK PYV VICEVAS GAKET VDRAVKAAR AAFOLGS PWRRM AS DRGRILLNILAT LIER DRITYLAAL ETLDNGK PYV VICEVAS GAKET VDRAVKAAR AAFOLGS PWRRM AS DRGRILLNILAT LIER DRITYLAAL ETLDNGK PYV VICEVAS GAKET VDRAVKAAR AAFOLGS PWRRM AS BRGRILLNILAT LIER DRITYLAAL ETLDNGK PYV VICEVAS GAKET VDRAVKAAR AAFOLGS PWRRM AS BRGRILLNILAT LIVERHOETLATIES STANDAKKS L-

	M 120	F 150	F 170 M 174 N 169 L 173 W 177
		↓	$\downarrow \downarrow \downarrow \downarrow \downarrow$
BA1	NAYLMELGGCIKTLRYCAG-WADKIQG	RTIPMDGNFFTYTRSEPV	GVCGQIIPWNFPLLMFLWKIGPALS
H01	NAYLMDLGGCLKTLRYCAG-WADKIQG	RTIPSDGNFFTYTRHEPV	CVCGQILPWNFPLLMFLWKIAPALS
EL VC1	HSLRDDIPGAARAIRWYAE-AIDKVYG	EVATTSSHELAMIVREPV	GVIAAIVPWNFPLLLTCWKLGPALA
C1	TAYLMDLGACIKTIRYCAG-WADKIHG	RTVPMDGNFFTFTRHEPV	GVCGOIIPWNFPLVMFIWKIAPALC
M1	NAYLSDLGGCIKALKYCAG-WADKIHG	QTIPSDGDIFTYTRREPI	GVCGQIIPWNFPMLMFIWKIGPALS
R1	HAYLLDTEVSIKALKYFAG-WADKING	QTIPSDGDVFTYTRREPI	GVCGQIIPWNGPLILFIWKIGAALS
SI H1	NAYLMDLGGCIKTLRYCAG-WADKIQG	RTIPMDGNEETYTRSEPV	GVCGQIIPWNFPLLMFLWKIGPALS
RALDH-1	NAYLSBLGGSIKALKYCAG-WADKIHG	OTIPSDGDIFTFTRREPI	GVCGOIIPWEFPLLMFIWKIGPALS
RALDH1	NAYLSDLGGSIKALKYCAG-WADKIHG	QTIPSDGDIFTFTRREPI	GVCGQIIPWNFPLLMFIWKIGPALS
RALDH-2	QAFYIDLQCVIKTLRYYAG-WADKIHG	MTIPVDGDYFTFTRHEPI	GVCGQIIPWNFPLLMFTWKIAPALC
RALDHZ B2	QAFYIDLOGVIKTLRYYAG-WADKING	MTIPVDGDYFTFTRHEPI	GVCGQIIPWNFPLLMPTWKIAPALC
HO2	ISYLVDLDMVLKCLRYYAG-WADKYHG	KTIPIDGDFFSYTRHEPV	GVCGOIIPWNFPLLMOAAKLGPALA
YM1	FCARGDVALVSKYLRSCGG-WADKIYG	NVI DTGKNHFTYSIKEPL	GVCGQIIPWNFPLLMWSWKIGPALA
M2	ISYLVDLDMVLKCLRYYAG-WADKYHG	KTIPIDGDFFSYTRHEPV	GVCGQIIPWNFPLLMQAWKLGPALA
RZ H2	ISYLVILDMVLKCLRYYAG-WADKYHG	KTIPIDGDFFSYTRHEPV	GVCGQIIPWNFPLLMQAWKLGPALA
CRY	SAYLMDLDYCIKVLRYCAG-WADKIOG	RTIPVDGEFFSYTRHEPI	GVCGOIFPWNAPMILLACKIGPALC
H6	HAFFIDLEGCIRTLRYFAG-WADKIOG	KTIPTDDNVVCFTRHEPI	GVCGAI TPWNFPLLMLVWKLAPALC
HX	ESYALDLDEVIKVYRYFAG-WADKWHG	KTIPMHGQHFCFTRHEPV	GVCGQIIPWNFPLVMQGWKLAPALA
M3	FCARGDVALVSKYLRSCGG-WADKIYG	CKEPT CENTRE CONTRACTOR	GVCGQIIPWNFPLLMWSWKIGPALA
R3	YEEVAHVLEELDTTIKELPDWAEDEPV	AKTROTOODDLYIHSEPI	GVVLVIGAWNYPFNLTIOPMVGAVA
нз	YEEVVYVLEEIEYMIQKLPEWAADEPV	EKTPOTOODELYIHSEPI	GVVLVIGTWNYPFNLTIQPMVGAIA
MM	SHEVITILGEIDFMLGNLPELASARPA	KKNLLTMMDEAYVQPEPI	GVVLIIGAWNYPFVLTMQPLVGAIA
H4	SHEVITILGEI FMLGNLPELASARPA	KKNLLTMMDEAYVQPEPI	GVVLIIGAWNYPPVLILQPLVGAIA
H9	EARLDIDISWOCLEYYAGLAASMAC	SEHIOLPGGSFGYTRREPI	GVCVGIGAWNYPFOIASWKSAPALA
н7	VSEVAL SOCEVTLAL RNLRAWMK DERV	PKNLATQL DSAFIRKEPH	GLVLIIAPWNYPLNLTLVPLVGALA
H10	COFUT TUT, CFT DEMT, ENT, DEWITTAKDI	TERNIT, TMT. DEA VTODODI	CINET TTOL DURING TO DI TOLT
mito	SAFAT ADOPT BUILDENDLENA EVILL	CUM & DEMONDATI ALAL	GVVDIIGAWNIPPVDIIGAIA
mito	T 200	OUR CHEMINEST AL ALL	G 250
BA1		AGF PPGVVNIVPGYGPTA	
BA1 HO1	CGNTVVVKPAEQTPLSALHVATLIKE	AGFPPGVVNIVPGYGPTAC AGFPPGVVNIVPGYGPTAC	G 250 G AAISSHMOVDKVAFTGSTEVGKLIK GAAISSHMOIDKVAFTGSTEVGKLIK
BA1 HO1 E1	T 200 CGNTVVVKPAEQTPLTALHMGSLIKE CGNTVVVKPAEQTPLSALHVATLIKE AGNSVILKPSEKSPLSAIRLAGLAKE	AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGYGPTA AGLPDGVLNVVTGFGHEA	G 250 G 250
BA1 HO1 E1 YC1 C1	CGNTVVVKPAEQTPLTALHMGSLIKE CGNTVVVKPAEQTPLSALHVATLIKE AGNSVILKPSEKSPLSAIRLAGLAKE MGNVCILKPAAVTPLNALYPASLCKK	AGFPPGVVNIVPGYGPTAG AGFPPGVVNIVPGYGPTAG AGLP GVLNVVTGFGHFAG VGIPAGVVNIVPGPGRTVG	G250 GAAISSHMOVDKVAFTGSTEVGKLIK GAAISSHMOIDKVAFTGSTEVGKLIK GALSRHNOIDAIAFTGSTEVGKLIK GALSRHNOIDAIAFTGSTEVGKLIK GALSRHNOIDAIAFTGSTEVGKSVA
BA1 HO1 E1 YC1 C1 M1	T 200 CCNTVVVK PAEQT PLTALHMGSLIK FA CCNTVVVK PAEQT PLSALHVATLIK FA AGNSVILK PSEK SPLSAIRLAGLAK FA MGNVCILK PAAVT PLNALYFASLCKK CGNTVVVK PAEQT PLSALYMGSLIK FA CGNTVVVK PAEQT PLSALYMGSLIK FA	AGFPPGVVNIVPGYGPTAG AGFPPGVVNIVPGYGPTAG AGLPDGVLNVVTGFGHEAG VGIPAGVVNIVPGPGRTVG AGFPPGVVNIVPGFGPTAG	G 250 G 250
BA1 HO1 E1 YC1 C1 M1 R1	T 200 CGNTVVVK PAEQT PLTALHMGSLIKER CGNTVVVK PAEQT PLSALHVATLIKER AGNSVILK PSEKS PLSAIRLAGLAKER MGNVCILK PAAVT PLNALYPASLCKK CGNTVVVK PAEQT PLSALYMGSLIKER CGNTVVVK PAEQT PLTALHLASLIKER CGNTVIVK PAEQT PLTALHLASLIKER	AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGYGPTA AGLP GVLNVVTGFGHEA VGIPAGVVN IVPGPGRTV AGFPPGVVN IVPGFGPTA AGFPPGVVN IVPGYGPTA AGFPPGVVNVVPGYGSTA	G 250 G 250
BA1 HO1 E1 YC1 C1 M1 R1 S1	T 200 CGNTVVVK PAEQT PLTALHMGSLIKEA CGNTVVVK PAEQT PLSALHVATLIKEA AGNSVILK PSEKS PLSAIRLAGLAKEA MGNVCILK PAAVT PLNALYPASLCKK CGNTVVVK PAEQT PLSALYMGSLIKEA CGNTVVVK PAEQT PLTALHLASLIKEA CGNTVVVK PAEQT PLTALHMGSLIKEA	AGFPPGVVN IVPGYGPTAG AGFPPGVVN IVPGYGPTAG AGLP GVLNVTGFGHEAG VGIPAGVVN IVPGPGRTVG AGFPPGVVN IVPGFGPTAG AGFPPGVVN IVPGYGPTAG AGFPPGVVN IVPGYGPTAG	G 250 G 250
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 BALDH-1	T 200 T 200 CGNTVVVK PAEQTPLTALHMGSLIKER CGNTVVVK PAEQTPLSALHVATLIKER AGNSVILKPSEKSPLSAIRLAGLAKER MGNVCILKPAAVTPLNALYPASLCKK CGNTVVVK PAEQTPLSALYMGSLIKER CGNTVVVK PAEQTPLTALHLASLIKER CGNTVVVK PAEQTPLTALHVASLIKER CGNTVVVK PAEQTPLTALHVASLIKER CGNTVVVK PAEQTPLTALHVASLIKER	AGFPPGVVNIVPGYGPTA AGFPPGVVNIVPGYGPTA AGLP GVLNVTGFGHFA VGIPAGVVNIVPGPGRTV AGFPPGVVNIVPGYGPTA AGFPPGVVNIVPGYGPTA AGFPPGVVNIVPGYGPTA AGFPPGVVNIVPGYGPTA	G 250 G 250
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH1	T 200 CGNTVVVK PAEQT PLTALHMGSLIKER CGNTVVVK PAEQT PLSALHVATLIKER AGNSVILK PSEKSPLSALHVATLIKER MGNVCILK PAAVT PLNALYFASLCKK CGNTVVVK PAEQT PLSALYMGSLIKER CGNTVVVK PAEQT PLTALHLASLIKER CGNTVVVK PAEQT PLTALHMGSLIKER CGNTVVVK PAEQT PLTALHMASLIKER CGNTVVVK PAEQT PLTALHMASLIKER CGNTVVVK PAEQT PLTALHMASLIKER	AGFPPGVVNIVPGYGPTA(AGFPPGVVNIVPGYGPTA(AGFPGVVNIVPGPGRTV(AGFPPGVVNIVPGPGRTV(AGFPPGVVNIVPGYGPTA(AGFPPGVVNIVPGYGPTA(AGFPPGVVNIVPGYGPTA(AGFPPGVVNIVPGYGPTA(AGFPPGVVNIVPGYGPTA(AGFPPGVVNIVPGYGPTA(AGFPPGVVNIVPGYGPTA)	G 250 G 250
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH1 RALDH-2	T 200 CGNTVVVK PAEQT PLTALHMGSLIK FA CGNTVVVK PAEQT PLSALHVATLIK FA AGNSVILK PSEKS PLSAI RLAGLAK FA MGNVC ILK PAAVT PLNALYFASLCKK CGNTVVVK PAEQT PLSALYMGSLIK FA CGNTVVVK PAEQT PLTALHLASLIK FA CGNTVVVK PAEQT PLTALHMGSLIK FA CGNTVVVK PAEQT PLTALHMASLIK FA CGNTVVVK PAEQT PLTALHMASLIK FA CGNTVVVK PAEQT PLTALHMASLIK FA	AGFPPGVVNIVPGYGPTA AGFPPGVVNIVPGYGPTA AGFPPGVVNIVPGFGHEA VGIPAGVVNIVPGFGHEA AGFPPGVVNIVPGYGPTA AGFPPGVVNIVPGYGPTA AGFPPGVVNIVPGYGPTA AGFPPGVVNIVPGYGPTA AGFPPGVVNIVPGYGPTA AGFPPGVVNIVPGYGPTA	G 250 GAAISSHMOVDKVAFTGSTEVGKLIK GAAISSHMOVDKVAFTGSTEVGKLIK GAAISSHMOIDKVAFTGSTEVGKLIK GALSRHNTIDAIAFTGSTEVGKLIK GAAISHMOIDKVAFTGSTEVGKLIK GAAISSHMOVDKVAFTGSTEVGKLIK GAAISSHMOVDKVAFTGSTEVGKLIK GAAISSHMOVDKVAFTGSTEVGKLIK GAAISSHMOVDKVAFTGSTQVGKLIK GAAISSHMOVDKVAFTGSTQVGKLIK GAAISSHMOVDKVAFTGSTQVGKLIK GAAISSHMOVDKVAFTGSTQVGKLIK
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 P22	T 200 CGNTVVVK PAEQTPLTALHMGSLIK HA CGNTVVVK PAEQTPLSALHVATLIK HA AGNSVILK PSEKSPLSALRUAGLAK HA MGNVCILK PAAVT PLNALYPASLCKK CGNTVVK PAEQTPLSALYMGSLIK HA CGNTVVK PAEQTPLTALHLASLIK HA CGNTVVK PAEQTPLTALHMGSLIK HA CGNTVVK PAEQTPLTALHMGSLIK HA CGNTVVK PAEQTPLTALHMASLIK HA	AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGYGPTA AGFPGVVN IVPGYGPTA AGFPPGVVN IVPGFGPTA AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGYGPTA	G 250 GAAISSHMUVDKVAFTGSTEVGKLIK GAAISSHMUVDKVAFTGSTEVGKLIK GAAISSHMUIDKVAFTGSTEVGKLIK GAAISHMUIDAIAFTGSTEVGKLIK GAAISHMUVDKVAFTGSTEVGKLIK GAAISSHMUVDKVAFTGSTEVGKLIK GAAISSHMUVDKVAFTGSTEVGKLIK GAAISSHMUVDKVAFTGSTEVGKLIK GAAISSHMUVDKVAFTGSTEVGKLIK GAAISSHMUVDKVAFTGSTEVGKLIK GAAISSHMUVDKVAFTGSTEVGKLIK GAAISSHMUVDKVAFTGSTEVGKLIK GAAISSHMUVDKVAFTGSTEVGKLIK GAAISSHMUVDKVAFTGSTEVGKLIK GAAISSHMUVDKVAFTGSTEVGKLIQ GAAIASHIGIDKIAFTGSTEVGKLIQ
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH-2 B2 HO2	T 200 CGNTVVVK PAEQT PLTALHMGSLIKEA CGNTVVVK PAEQT PLSALHVATLIKEA AGNSVILK PSEKS PLSAIRLAGLAKEA MGNVCILK PAAVT PLNALYPASLCKK CGNTVVVK PAEQT PLSALYMGSLIKEA CGNTVVVK PAEQT PLTALHLASLIKEA CGNTVVVK PAEQT PLTALHMGSLIKEA CGNTVVVK PAEQT PLTALHMGSLIKEA CGNTVVVK PAEQT PLTALHMASLIKEA CGNTVVVK PAEQT PLTALHMASLIKEA CGNTVVVK PAEQT PLTALHMASLIKEA CGNTVVVK PAEQT PLSALYMGALIKEA CGNTVVIK PAEQT PLSALYMGALIKAA CGNTVVIK PAEQT PLSALYMGALIKAA CGNTVVIK PAEQT PLSALYMGALIKAA CGNTVVIK PAEQT PLSALYMGALIKAA CGNTVVIK PAEQT PLSALYMGALIKAA CGNTVVIK PAEQT PLSALYMAANAA CGNTVVIK PAEQT PLSALYMAANAAA CGNTVVIK PAEQT PLSALYMAANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	AGFPPGVVN IVPGYGPTAG AGFPPGVVN IVPGYGPTAG AGFPPGVVN IVPGYGPTAG AGFPPGVVN IVPGFGPTAG AGFPPGVVN IVPGYGPTAG AGFPPGVVN IVPGYGPTAG AGFPPGVVN IVPGYGPTAG AGFPPGVVN IVPGYGPTAG AGFPPGVVN IVPGYGPTAG AGFPPGVVN IVPGYGPTAG AGFPPGVVN IVPGYGPTAG AGFPPGVVN IVPGYGPTAG AGFPPGVVN IVPGYGPTAG AGFPPGVVN IVPGYGPTAG	G 250 G
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH1 RALDH2 RALDH2 B2 HO2 YM1	T 200 CGNTVVVK PAEQT PLTALHMGSLIKER CGNTVVVK PAEQT PLSALHVATLIKER AGNSVILK PSEKS PLSAI RLAGLAKER MGNVCILK PAAVT PLNALYFASLCKK CGNTVVVK PAEQT PLSALYMGSLIKER CGNTVVVK PAEQT PLTALHLASLIKER CGNTVVVK PAEQT PLTALHMGSLIKER CGNTVVVK PAEQT PLTALHMASLIKER CGNTVVVK PAEQT PLTALHMASLIKER CGNTVVVK PAEQT PLTALHMASLIKER CGNTVVVK PAEQT PLTALHMASLIKER CGNTVVVK PAEQT PLTALHMASLIKER CGNTVVVK PAEQT PLSALYMGALIKER CGNTVVIK PAEQT PLSALYMGALIKER CGNTVVIK PAEQT PLSALYMGALIKER CGNTVVIK PAEQT PLTALYVANLIKER CGNTVVIK PAEQT PLSALYMGALIKER CGNTVVIK PAEQT PLSALYMALKER CGNTVVIK PAEQT PLSALYMANTKER CGNTVVIK PAEQT PLSALYMANTKER CGNTVNIK PAEQT PLSALY	AGFPPGVVNIVPGYGPTA(AGFPPGVVNIVPGYGPTA(AGFPGVVNIVPGYGPTA(AGFPGVVNIVPGPGRTV(AGFPPGVVNIVPGFGPTA(AGFPPGVVNIVPGYGPTA(AGFPPGVVNIVPGYGPTA(AGFPPGVVNIVPGYGPTA(AGFPPGVVNIVPGYGPTA(AGFPPGVVNIVPGYGPTA(AGFPPGVVNIPGYGPTA) AGFPPGVVNIPGYGPTA(AGFPPGVVNIPGFGPTA) AGFPPGVVNVPGFGPTA(AGFPPGVVNVPGFGPTA)	G 250 G
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2 YM1 M2	T 200 CGNTVVVK PAEQT PLTALHMGSLIK FA CGNTVVVK PAEQT PLSALHVATLIK FA AGNSVILK PSEKS PLSAI RLAGLAK FA MGNVC ILK PAAVT PLNALYFASLCKK CGNTVVVK PAEQT PLSALYMGSLIK FA CGNTVVVK PAEQT PLTALHLASLIK FA CGNTVVVK PAEQT PLTALHMGSLIK FA CGNTVVVK PAEQT PLTALHMASLIK FA CGNTVVVK PAEQT PLTALHMASLIK FA CGNTVVVK PAEQT PLTALHMASLIK FA CGNTVVK PAEQT PLTALHMASLIK FA CGNTVVK PAEQT PLTALHMASLIK FA CGNTVVK PAEQT PLTALHMASLIK FA CGNTVVK PAEQT PLTALYMALIK FA CGNTVVK PAEQT PLSALYMGALIK FA CGNTVVK PAEQT PLTALYVANLIK FA	AGFPPGVVNIVPGYGPTA AGFPPGVVNIVPGYGPTA AGFPPGVVNIVPGYGPTA AGFPPGVVNIVPGFGPTA AGFPPGVVNIVPGYGPTA AGFPPGVVNIVPGYGPTA AGFPPGVVNIVPGYGPTA AGFPPGVVNIVPGYGPTA AGFPPGVVNILPGYGPTA AGFPPGVVNILPGYGPTA AGFPPGVVNILPGYGPTA AGFPPGVVNIPGFGPTA AGFPPGVVNILPGFGPTA	G 250 G 250
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2 YM1 M2 R2 YW2	T 200 CGNTVVVK PAEQT PLTALHMGSLIK FA CGNTVVVK PAEQT PLSALHVATLIK FA AGNSVILK PSEKS PLSAL RUACLAK FA MGNVC ILK PAAVT PLNALYFASLCKK CGNTVVVK PAEQT PLSALYMGSLIK FA CGNTVVVK PAEQT PLTALHASLIK FA CGNTVVVK PAEQT PLTALHMGSLIK FA CGNTVVVK PAEQT PLTALHMGSLIK FA CGNTVVVK PAEQT PLTALHMASLIK FA CGNTVVK PAEQT PLTALHMASLIK FA CGNTVVK PAEQT PLTALHMASLIK FA CGNTVVK PAEQT PLTALYMGALIK FA CGNTVVK PAEQT PLTALYVANLIK FA CGNTVVK FAEQT PLTALYVANLIK FA CGNTVVK FAEQT PLTALYVANLIK FA CGNTVVK FAEQT PLTALYVANLIK FA CGNTVVK FAEQT PLTALYVANLIK FA	AGFPPGVVNIVPGYGPTAG AGFPPGVVNIVPGYGPTAG AGFPPGVVNIVPGYGPTAG AGFPPGVVNIVPGFGPTAG AGFPPGVVNIVPGYGPTAG AGFPPGVVNIVPGYGPTAG AGFPPGVVNIVPGYGPTAG AGFPPGVVNIVPGYGPTAG AGFPPGVVNIVPGYGPTAG AGFPPGVVNILPGYGPTA AGFPPGVVNILPGYGPTA AGFPPGVVNIPGFGPTA AGFPPGVVNIVPGFGPTA AGFPPGVVNIVPGFGPTA	G 250 GAAI SSHMD VDKVAFTGSTEVGKLIK GAAI SSHMD VDKVAFTGSTEVGKLIK GAAI SSHMD I DKVAFTGSTEVGKLIK GAAI SSHMD I DKVAFTGSTEVGKLIK GAAI SSHMD I DAIAFTGSTEVGKSVA GAAI SHMD I DKVAFTGSTEVGKLIK GAAI SSHMD VDKVAFTGSTEVGKLIK GAAI SSHMD VDKVAFTGSTEVGKLIK GAAI SSHMD VDKVAFTGSTEVGKLIK GAAI SSHMD VDKVAFTGSTEVGKLIK GAAI SSHMD VDKVAFTGSTEVGKLIK GAAI SSHMD VDKVAFTGSTEVGKLIK GAAI SSHMD VDKVAFTGSTEVGKLIQ GAAI ASHIGI DKIAFTGSTEVGKLIQ GAAI ASHEDVDKVAFTGSTEVGKLIQ GAAI ASHEDVDKVAFTGSTEVGHLIQ GAAI ASHEDVDKVAFTGSTEVGHLIQ GAAI ASHEDVDKVAFTGSTEVGHLIQ GAAI ASHEDVDKVAFTGSTEVGHLIQ GAAI ASHEDVDKVAFTGSTEVGHLIQ GAAI ASHEDVDKVAFTGSTEVGHLIQ GAAI ASHEDVDKVAFTGSTEVGHLIQ
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2 YM1 M2 R2 H2 CBY	T 200 CGNTVVVK PAEQT PLTALHMGSLIK HA CGNTVVVK PAEQT PLSALHVATLIK HA AGNSVILK PSEKS PLSALRUAGLAK HA MGNVC ILK PAAVT PLNALYPASLCKK CGNTVVK PAEQT PLSALMGSLIK HA CGNTVVK PAEQT PLTALHASLIK HA CGNTVVK PAEQT PLTALHASLIK HA CGNTVVK PAEQT PLTALHMGSLIK HA CGNTVVK PAEQT PLTALHMSLIK HA CGNTVVK PAEQT PLTALHMSLIK HA CGNTVVK PAEQT PLTALHMASLIK HA CGNTVVK PAEQT PLTALYMALIK HA CGNTVVK PAEQT PLTALYANLIK HA TGNVVMK VAEQT PLTALYANLIK HA TGNVVMK HA TGNVVMK HA TGNVVMK HA TGNVVMK HA TGNVVMK HA TGNVVMK HA TGNVVMK HA TGNVYMK HA TGNY TGN	AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGFGPTA AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGYGPTA AGFPPGVVN IPGYGPTA AGFPPGVVN IPGFGPTA AGFPPGVVN IVPGFGPTA AGFPPGVVN IVPGFGPTA AGFPPGVVN IVPGFGPTA AGFPPGVVN IVPGFGPTA	G 250 GAAI SSHMU VDKVAFTGSTEVGKLIK GAAI SSHMU VDKVAFTGSTEVGKLIK GAAI SSHMU I DAIAFTGSTEVGKLIK GAAI SSHMU I DAIAFTGSTEVGKLIK GAAI SHMU I DAIAFTGSTEVGKLIK GAAI SHMU I DAVAFTGSTEVGKLIK GAAI SHMU VDKVAFTGSTEVGKLIK GAAI SSHMU VDKVAFTGSTEVGKLIK GAAI SSHMU VDKVAFTGSTEVGKLIK GAAI SSHMU VDKVAFTGSTEVGKLIK GAAI SSHMU VDKVAFTGSTEVGKLIK GAAI SSHMU VDKVAFTGSTEVGKLIK GAAI SSHMU VDKVAFTGSTEVGKLIQ GAAIASHEU VDKVAFTGSTEVGKLIQ GAAIASHEU VDKVAFTGSTEVGKLIQ GAAIASHEU VDKVAFTGSTEVGHLIQ GAAIASHEU VDKVAFTGSTEVGHLIQ
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2 YM1 M2 R2 H2 CRY H6	T 200 CGNTVVVK PAEQT PLTALHMGSLIK FA CGNTVVVK PAEQT PLSALHVATLIK FA AGNSVILK PSEKS PLSALRLAGLAK FA MGNVC ILK PAAVT PLNALYPASLCKK CGNTVVVK PAEQT PLSALYMGSLIK FA CGNTVVVK PAEQT PLTALHLASLIK FA CGNTVVVK PAEQT PLTALHASLIK FA CGNTVVVK PAEQT PLTALHMGSLIK FA CGNTVVVK PAEQT PLTALHMASLIK FA CGNTVVVK PAEQT PLTALHMASLIK FA CGNTVVVK PAEQT PLTALHMASLIK FA CGNTVVVK PAEQT PLTALHMASLIK FA CGNTVVVK PAEQT PLTALYMALIK FA CGNTVVK PAEQT PLTALYMALIK FA CGNTVVK PAEQT PLTALYMALIK FA CGNTVVK PAEQT PLTALYVANLIK FA CGNTVVK PAEQT PLTALYVANLIK FA CGNTVVK PAEQT PLTALYVANLIK FA CGNTVVK PAEQT PLTALYVANLIK FA CGNTVVK FAEQT FLTALYVANL FA CGNTVVK FAEQT FALVY FALVYNK FA CGNTVVK FAEQT FAL	AGF PPGVVN IVPGYGPTAG AGF PPGVVN IVPGYGPTAG AGF PPGVVN IVPGYGPTAG AGF PPGVVN IVPGFGPTAG AGF PPGVVN IVPGYGPTAG AGF PPGVVN IVPGYGPTAG AGF PPGVVN IVPGYGPTAG AGF PPGVVN IVPGYGPTAG AGF PPGVVN IVPGYGPTAG AGF PPGVVN IVPGYGPTA AGF PPGVVN IVPGFGPTA AGF PPGVVN IVPGFGPTA	G 250 GAAI SSHMU VDKVAFTGSTEVGKLIK GAAI SSHMU VDKVAFTGSTEVGKLIK GAAI SSHMU I DAIAFTGSTEVGKLIK GAAI SSHMU I DAIAFTGSTEVGKLIK GAAI SHMU I DAIAFTGSTEVGKLIK GAAI SHMU I DAVAFTGSTEVGKLIK GAAI SSHMU VDKVAFTGSTEVGKLIK GAAI SSHMU VDKVAFTGSTEVGKLIK GAAI SSHMU VDKVAFTGSTEVGKLIK GAAI SSHMU VDKVAFTGSTEVGKLIK GAAI SSHMU VDKVAFTGSTEVGKLIK GAAI SSHMU VDKVAFTGSTEVGKLIK GAAI SSHMU VDKVAFTGSTEVGKLI GAAI SSHMU VDKVAFTGSTEVGKLIQ GAAI SSHMU VDKVAFTGSTEVGKLIQ GAAI SSHMU VDKVAFTGSTEVGKLIQ GAAI ASHEU VDKVAFTGSTEVGKLIQ GAAI ASHEU VDKVAFTGSTEVGHLIQ GAAI ASHEU VDKVAFTGSTEVGHLIQ GAAI ASHEU VDKVAFTGSTEVGHLIQ GAAI ASHEU VDKVAFTGSTEVGHLIQ GAAI ASHEU VDKVAFTGSTEVGHLIQ GAAI ASHEU VDKVAFTGSTEVGHLIQ GAAI SSHMU VDKVAFTGSTEVGHLIQ GAAI SSHMU VDKVAFTGSTEVGHLIQ GAAI SSHMU VDKVAFTGSTEVGHLIQ GAAI SSHMU VDKVAFTGSTEVGKUIQ GAAI SSHMU VDKVAFTGSTEVGKUIQ GAAI SSHMU VDKVAFTGSTEVGKUIQ
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2 YM1 M2 R2 H2 CRY H6 HX	T 200 CGNTVVVK PAEQT PLTALHMGSLIKER CGNTVVVK PAEQT PLSALHVATLIKER AGNSVILK PSEKS PLSAI RLAGLAKER MGNVCILK PAAVT PLNALYFASLCKK CGNTVVVK PAEQT PLSALYMGSLIKER CGNTVVVK PAEQT PLTALHLASLIKER CGNTVVVK PAEQT PLTALHMGSLIKER CGNTVVVK PAEQT PLTALHMASLIKER CGNTVVVK PAEQT PLTALHMASLIKER CGNTVVVK PAEQT PLTALHMASLIKER CGNTVVVK PAEQT PLTALHMASLIKER CGNTVVVK PAEQT PLTALHMASLIKER CGNTVVVK PAEQT PLTALHMASLIKER CGNTVVIK PAEQT PLSALYMGALIKER CGNTVVIK PAEQT PLSALYMGALIKER CGNTVVIK PAEQT PLTALYVANLIKER CGNTVVIK PAEQT PLTALYVANLIKER CGNTVVIK PAEQT PLTALYVANLIKER CGNTVVLK PAEQT PLTALYVANLIKER CGNTVVLK PAEQT PLTALYVANLIKER CGNTVVKK VAEQT PLTALYVANLIKER CGNTVVKK VAEQT PLTALYVANLIKER CGNTVVKK VAEQT PLTALYVANLIKER CGNTVVKK PAEQT PLTALYVANLIKER CGNTVVKKAR CGNTVKKAR	AGF PPGVVN IVPGYGPTAG AGF PPGVVN IVPGYGPTAG AGF PPGVVN IVPGYGPTAG AGF PPGVVN IVPGFGPTAG AGF PPGVVN IVPGYGPTAG AGF PPGVVN IVPGYGPTAG AGF PPGVVN IVPGYGPTAG AGF PPGVVN IVPGYGPTAG AGF PPGVVN IVPGYGPTA AGF PPGVVN IVPGYGPTA AGF PPGVVN IVPGFGPTA AGF PPGVVN IVPGFGPTA	G 250 G 200 G
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH1 RALDH2 B2 HO2 YM1 M2 R2 HO2 YM1 M2 R2 H2 CRY H6 HX YM3	T 200 CGNTVVVK PAEQT PLTALHMGSLIK FA CGNTVVVK PAEQT PLSALHVATLIK FA AGNSVILK PSEKS PLSAT RLAGLAK FA MGNVC ILK PAAVT PLNALYFASLCKK CGNTVVVK PAEQT PLSALYMGSLIK FA CGNTVVVK PAEQT PLTALHLASLIK FA CGNTVVK PAEQT PLTALHASLIK FA CGNTVVK PAEQT PLTALHMGSLIK FA CGNTVVK PAEQT PLTALHMASLIK FA CGNTVVK PAEQT PLTALYMALIK FA CGNTVVK PAEQT PLTALYMALIK FA CGNTVVK FAEQT PLTALYVANLIK FA	AGFPPGVVNIVPGYGPTAG AGFPPGVVNIVPGYGPTAG AGFPPGVVNIVPGYGPTAG AGFPPGVVNIVPGFGPTAG AGFPPGVVNIVPGYGPTAG AGFPPGVVNIVPGYGPTAG AGFPPGVVNIVPGYGPTAG AGFPPGVVNIVPGYGPTAG AGFPPGVVNIVPGYGPTAG AGFPPGVVNILPGYGPTA AGFPPGVVNILPGYGPTA AGFPPGVVNILPGGGPVA AGFPPGVVNIVPGFGPTA AGFPPGVVNIVPGFGPTA AGFPPGVVNIVPGFGPTA AGFPPGVVNIVPGFGPTA AGFPPGVVNIVPGFGPTA AGFPPGVVNIVPGFGPTA AGFPPGVVNIVPGFGPTA AGFPPGVVNIVPGFGPTA AGFPPGVVNIVPGFGPTA AGFPPGVVNIVPGFGPTA	G 250 GAAI SSHMD VDKVAFTGSTEVGKLIK GAAI SSHMD VDKVAFTGSTEVGKLIK GAAI SSHMD I DKVAFTGSTEVGKLIK GALSSHMD I DAIAFTGSTEVGKLIK GALSRHNT I DAIAFTGSTEVGKLIK GALSSHMD VDKVAFTGSTEVGKLIK GAAI SSHMD VDKVAFTGSTEVGKLIQ GAAI ASHED VDKVAFTGSTEVGKLIQ GAAI ASHED VDKVAFTGSTEVGKLIQ GAAI ASHED VDKVAFTGSTEVGHLIQ GAAI ASHED VDKVAFTGSTEVGHLIQ GAAI ASHED VDKVAFTGSTEVGHLIQ GAAI SSHMD VDKVAFTGSTEVGHLIQ GAAI SSHMD VDKVAFTGSTEVGHLIQ GAAI SSHMD VDKVAFTGSTEVGHLIQ GAAI SSHMD VDKVAFTGSTEVGHLIQ GAAI SSHMD VDKVAFTGSTEVGHLIQ GAAI SSHMD VDKVAFTGSTEVGKLIQ GAAI SSHMD VDKVAFTGSTEVGKLIQ GA
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2 YM1 M2 R2 HO2 YM1 M2 R2 H2 CRY H6 HX YM3 M3 R3	T 200 CGNTVVVK PAEQT PLTALHMGSLIK HA CGNTVVVK PAEQT PLSALHVATLIK HA AGNSVILK PSEKS PLSALRLAGLAK HA MGNVC ILK PAAVT PLNALYPASLCKK CGNTVVVK PAEQT PLSALYMGSLIK HA CGNTVVVK PAEQT PLSALYMGSLIK HA CGNTVVVK PAEQT PLTALHASLIK HA CGNTVVVK PAEQT PLTALHMGSLIK HA CGNTVVVK PAEQT PLTALHMGSLIK HA CGNTVVVK PAEQT PLTALHMASLIK HA CGNTVVVK PAEQT PLTALHMASLIK HA CGNTVVVK PAEQT PLTALHMASLIK HA CGNTVVVK PAEQT PLTALHMASLIK HA CGNTVVK PAEQT PLTALYMALIK HA TGNVVMK VAEQT PLTALYVANLIK HA TGNVVMK VAEQT PLTALYVANLIK HA TGNVVMK VAEQT PLTALYVANLIK HA TGNVVMK VAEQT PLTALYVANLIK HA CGNTVVK PAEQT PLTALYVANLIK HA TGNVVMK VAEQT PLTALYVANLIK HA TGNVVMK VAEQT PLTALYVANLIK HA TGNVVVMK VAEQT PLTALYVANLIK HA TGNVVMK VAEQT PLTALYVANLIK HA TGNVVVMK VAEQT PLTALYVANLIK HA TGNVVVK HA TGNVVK HA TGNVVK HA TGNVVK HA TGNVVK HA TGNVVK HA TGNVVK HA TGNVY HA TGNY HA TGNY HA TGNY HA TGNY H	AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGFGPTA AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGFGPTA AGFPPGVVN IVPGFGPTA	G 250 GAAI SSHMD VDKVAFTGSTEVGKLIK GAAI SSHMD VDKVAFTGSTEVGKLIK GAAI SSHMD I DKVAFTGSTEVGKLIK GALSRHNT I DAIAFTGSTEVGKLIK GALSRHNT I DAIAFTGSTEVGKLIK GALSSHMD VDKVAFTGSTEVGKLIK GAAI SHMD I DKVSFTGSTEVGKLIK GAAI SSHMD VDKVAFTGSTEVGKLIK GAAI SSHMD VDKVAFTGSTEVGKLIQ GAAI ASHED VDKVAFTGSTEVGKLIQ GAAI ASHED VDKVAFTGSTEVGKLIQ GAAI ASHED VDKVAFTGSTEVGHLIQ GAAI ASHED VDKVAFTGSTEVGHLIQ GAAI SSHMD VDKVAFTGSTEVGKLVK GAAI SSHMD VDKVAFTGSTATGFNIM TELLKEK FDHIMYTGSTAVGKIVM
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2 YM1 M2 R2 HO2 YM1 M2 R2 H2 CRY H6 HX YM3 M3 R3 H3	T 200 CGNTVVVK PAEQT PLTALHMGSLIK FA CGNTVVVK PAEQT PLSALHVATLIK FA AGNSVILK PSEKS PLSALRLAGLAK FA MGNVC ILK PAAVT PLNALYPASLCKK CGNTVVVK PAEQT PLSALYMGSLIK FA CGNTVVVK PAEQT PLSALYMGSLIK FA CGNTVVVK PAEQT PLTALHASLIK FA CGNTVVVK PAEQT PLTALHASLIK FA CGNTVVVK PAEQT PLTALHMGSLIK FA CGNTVVVK PAEQT PLTALHMASLIK FA CGNTVVVK PAEQT PLTALHMASLIK FA CGNTVVVK PAEQT PLTALHMASLIK FA CGNTVVVK PAEQT PLTALHMASLIK FA CGNTVVK PAEQT PLTALHMASLIK FA CGNTVVK PAEQT PLTALHMASLIK FA CGNTVVK PAEQT PLTALYMALIK FA CGNTVVK PAEQT PLTALYMALIK FA CGNTVVK PAEQT PLTALYMALIK FA CGNTVVK VAEQT PLTALYVANLIK FA TGNVVMK VAEQT PLTALYVANLIK FA TGNVVVMK VAEQT PLTALYVANLIK FA CGNTVVVK PAEQT PLTALYVANLIK FA CGNTVVK PAEQT PLTALYVANLIK FA CGNTVVK PAEQT PLTALYVANLIK FA CGNTVVK PAEQT PLTALYVANLIK FA CGNTVVK FAEQT PLTALYVANL FA CGNTVVK FAEQT PLTALYVANL FA CGNTVYNK FAEQT PLTA	AGFPPGVVN IVPGYGPTAG AGFPPGVVN IVPGYGPTAG AGFPPGVVN IVPGYGPTAG AGFPPGVVN IVPGFGPTAG AGFPPGVVN IVPGYGPTAG AGFPPGVVN IVPGYGPTAG AGFPPGVVN IVPGYGPTAG AGFPPGVVN IVPGYGPTAG AGFPPGVVN IVPGYGPTAG AGFPPGVVN IVPGYGPTAG AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGFGPTA AGFPPGVVN IVPGFGPTA	G 250 GAAI SSHMU VDKVAFTGSTEVGKLIK GAAI SSHMU VDKVAFTGSTEVGKLIK GAAI SSHMU I DAIAFTGSTEVGKLIK GAAI SSHMU I DAIAFTGSTEVGKLIK GAAI SHMU I DAIAFTGSTEVGKLIK GAAI SHMU I DAVAFTGSTEVGKLIK GAAI SHMU VDKVAFTGSTEVGKLIK GAAI SSHMU VDKVAFTGSTEVGKLIQ GAAI ASHEU VDKVAFTGSTEVGKLIQ GAAI ASHEU VDKVAFTGSTEVGKLIQ GAAI ASHEU VDKVAFTGSTEVGHLIQ GAAI ASHEU VDKVAFTGSTEVGHLIQ GAAI ASHEU VDKVAFTGSTEVGHLIQ GAAI ASHEU VDKVAFTGSTEVGHLIQ GAAI SSHMU VDKVAFTGSTEVGHLIQ GAAI SSHMU VDKVAFTGSTEVGHLIQ GAAI SSHMU VDKVAFTGSTEVGHLIQ GAAI SSHMU VDKVAFTGSTEVGHLIQ GAAI SSHMU VDKVAFTGSTEVGKLIK GAAI SSHMU VDKVAFTGSTEVGKLIQ GAAI SSHMU VDKVAFTGSTEVGKU VM TULK KN FDH IMYTGSTAVGKU VM TULK KN FDH IMYTGSTAVGKU VM TULK KN FDH IMYTG
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2 YM1 M2 R2 H2 CRY H6 HX YM3 M3 R3 H3 MM	T 200 CGNTVVVK PAEQT PLTALHMGSLIK EA CGNTVVVK PAEQT PLSALHVATLIK EA AGNSVILK PSEKS PLSALRLAGLAK EA MGNVC ILK PAAVT PLNALYPASLCKK CGNTVVVK PAEQT PLSALYMGSLIK EA CGNTVVVK PAEQT PLTALHLASLIK EA CGNTVVVK PAEQT PLTALHASLIK EA CGNTVVVK PAEQT PLTALHMGSLIK EA CGNTVVVK PAEQT PLTALHMGSLIK EA CGNTVVVK PAEQT PLTALHMASLIK EA CGNTVVVK PAEQT PLTALYMALIK EA CGNTVVK PAEQT PLTALYMALIK EA CGNTVVK PAEQT PLTALYVANLIK EA CGNTVVK PAEQT PLTALYVANLIK CGNTVVK PAEQT PLTALYVANLY CGNTVVK PAEQT PLTA	AGFPPGVVNIVPGYGPTAG AGFPPGVVNIVPGYGPTAG AGFPPGVVNIVPGFGFTAG AGFPPGVVNIVPGFGFTAG AGFPPGVVNIVPGYGPTAG AGFPPGVVNIVPGYGPTAG AGFPPGVVNIVPGYGPTAG AGFPPGVVNIVPGYGPTAG AGFPPGVVNIVPGYGPTAG AGFPPGVVNIVPGYGPTA AGFPPGVVNIVPGFGPTA AGFPPGVVNIVPGFGPTA AGFPPGVVNIVPGFGPTA AGFPPGVVNIVPGFGPTA AGFPPGVVNIVPGFGPTA AGFPPGVVNIVPGFGPTA AGFPPGVVNIVPGFGPTA AGFPPGVVNIVPGFGPTA AGFPPGVVNIVPGFGPTA AGFPPGVVNIVPGFGPTA AGFPPGVVNIVPGFGPTA AGFPPGVVNIVPGFGPTA AGFPPGVVNIVPGFGPTA AGFPPGVVNIVPGFGPTA AGFPPGVVNIVPGFGPTA AGFPPGVVNIVPGFGPTA AGFPPGVVNIVPGFGPTA AGFPPGVVNIVPGFGPTA AGFPPGVVNIVFGFGPTA AGFPPGVVNIVFGFGPTA AGFPPGVVNIVFGFGPTA AGFPPGVVNIVFGFGPTA AGFPPGVVNIVFGFGPTA AGFPPGVVNIVFGFGPTA AGFPPGVVNIVFGFGPTA AGFPPGVVNIVFGFGPTA AGFPFGVVNIVFGFGPTA AGFPFGVVNIVFGFGPTA AGFPFGVVNIVFGFGPTA AGFPFGVVNIVFGFGPTA AGFPFGVVNIFFGFGPTA AGFPFGVVNIFFGFGPTA AGFPFGVVNIFFGFGPTA AGFPFGVVNIFFGFGPTA AGFPFGVVNIFFGFGPTA AGFPFGVVNIFFGFGPTA AGFPFGVVNIFFGFGPTA AGFPFGVVNIFFGFGPTA AGFPFGVVNIFFGFGPTA AGFPFGVVNIFFGFGPTA AGFPFGVVNIFFGFGPTA AGFPFGVVNIFFGFGPTA AGFPFGVVNIFFGFGPTA AGFPFFFFFGVVNIFFGFGPTA AGFPFGVVNIFFGFGPTA AGFPFGVVNIFFGFGPTA AGFPFGVVNIFFGFGPTA AFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF	G 250 GAAI SSHMU VDKVAFTGSTEVGKLIK GAAI SSHMU VDKVAFTGSTEVGKLIK GAAI SSHMU I DAIAFTGSTEVGKLIK GAAI SSHMU I DAIAFTGSTEVGKLIK GAAI SHMU I DAIAFTGSTEVGKLIK GAAI SHMU I DAVSFTGSTEVGKLIK GAAI SHMU I DKVSFTGSTEVGKLIK GAAI SSHMU VDKVAFTGSTEVGKLIK GAAI SSHMU VDKVAFTGSTEVGKLIQ GAAI ASHEU VDKVAFTGSTEVGKLIQ GAAI ASHEU VDKVAFTGSTEVGKLIQ GAAI ASHEU VDKVAFTGSTEVGHLIQ GAAI ASHEU VDKVAFTGSTEVGHLIQ GAAI ASHEU VDKVAFTGSTEVGHLIQ GAAI ASHEU VDKVAFTGSTEVGHLIQ GAAI ASHEU VDKVAFTGSTEVGHLIQ GAAI SSHMU VDKVAFTGSTEVGKLIQ GAAI SSHMU VDKVAFTGSTEVGKIIQ GAAI SSHMU VDKVAFTGSTEVGKIIQ
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH-2 RALDH2 B2 HO2 YM1 M2 R2 HO2 YM1 M2 R2 H2 CRY H6 HX YM3 M3 R3 H3 MM RM	T 200 CGNTVVVK PAEQT PLTALHMGSLIK EA CGNTVVVK PAEQT PLSALHVATLIK EA AGNSVILK PSEKS PLSAI RLAGLAK EA MGNVC ILK PAAVT PLNALYFASLCKK CGNTVVVK PAEQT PLSALYMGSLIK EA CGNTVVVK PAEQT PLTALHLASLIK EA CGNTVVVK PAEQT PLTALHMGSLIK EA CGNTVVVK PAEQT PLTALHMGSLIK EA CGNTVVVK PAEQT PLTALHMASLIK EA CGNTVVIK PAEQT PLSALYMGALIK EA CGNTVVIK PAEQT PLSALYMGALIK EA CGNTVVIK PAEQT PLSALYMGALIK EA CGNTVVIK PAEQT PLSALYMGALIK EA CGNTVVIK PAEQT PLTALYVANLIK EA CGNTVVK PAEQT PLTALYVANLIK EA CGNTVVKK PAEQT PLTALYVANLIK EA CGNTVVKKA EQT PLTALYVANLIK EA CGNTVVKKA EQT PLTALYVANLIK EA CGNTVVKK PAEQT PLTALYVANLIK EA CGNTVVKKA EQT PLTALYVANLIK EA CGNTVVKKA EA CGNTVVKA EA CGNTVVKA EA CGNTVVKA EA CGNTVYKA EA CGNTVYKA EA CGNTVYKA EA CGNTVYKA EA CGNTVK	AGFPPGVVNIVPGYGPTAG AGFPPGVVNIVPGYGPTAG AGFPPGVVNIVPGFGPTAG AGFPPGVVNIVPGFGPTAG AGFPPGVVNIVPGYGPTAG AGFPPGVVNIVPGYGPTAG AGFPPGVVNIVPGYGPTAG AGFPPGVVNIVPGYGPTAG AGFPPGVVNIVPGYGPTA AGFPPGVVNIVPGYGPTA AGFPPGVVNIVPGYGPTA AGFPPGVVNIVPGFGPTA	G 250 GAAI SSHMD VDKVAFTGSTEVGKLIK GAAI SSHMD VDKVAFTGSTEVGKLIK GAAI SSHMD I DKVAFTGSTEVGKLIK GAAI SSHMD I DAIAFTGSTEVGKLIK GAAI SSHMD I DAIAFTGSTEVGKLIK GAAI SSHMD VDKVAFTGSTEVGKLIK GAAI SSHMD VDKVAFTGSTEVGKLIQ GAAI SSHMD VDKVAFTGSTEVGKLIQ GAAI SSHMD VDKVAFTGSTEVGKLIQ GAAI SSHMD VDKVAFTGSTEVGKLIQ GAAI SSHMD VDKVAFTGSTEVGKLIQ GAAI SSHMD VDKVAFTGSTEVGKLIQ GAAI SSHMD VDKVAFTGSTEVGHLIQ GAAI SSHMD VDKVAFTGSTEVGKLIK GAAI SSHMD VDKVAFTGSTEVGKLIQ GAAI SSHMD VDKVAFTGSTAVGKIVM TELLKEC - FDHIMTGSTAVGKIVM TELLK QR - FDHILYTGNTAVGKIVM
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2 YM1 M2 R2 HO2 YM1 M2 R2 H2 CRY H6 HX YM3 M3 R3 H3 MM RM H4 H9	T 200 CGNTVVVK PAEQT PLTALHMGSLIK HA CGNTVVVK PAEQT PLSALHVATLIK HA AGNSVILK PSEKS PLSAL RLAGLAK HA MGNVC ILK PAAVT PLNALYPASLCKK CGNTVVVK PAEQT PLSALYMGSLIK HA CGNTVVK PAEQT PLTALHASLIK HA CGNTVVK PAEQT PLTALHASLIK HA CGNTVVK PAEQT PLTALHMSLIK HA CGNTVVK PAEQT PLTALYMALIK HA CGNTVVK PAEQT PLTALYVANLIK HA CGNTVVK KAEQT PLTALYVANLIK HA TGNVVMKVAEQT PLTALYVANLIK HA TGNVVKKVAEQT PLALYANNIK HA TGNVVKKVAEQT PLTALYVANLIK HA TGNVVKKY HA TGNVVKKAEQT PLTALYVANLIK HA TGNVVKKAEQT PLTALYVANLIK HA TGNVVKKAEQT PLTALYVANLIK HA TGNVVKKAEQT PLTALYVANLIK HA TGNVVKKAEQT PLALYANNIK HA TGNVVKKAEQT PLALYANNIK HA TGNVVKKAEQT PLALYANNIK HA TGNVYK HA TGNYYK HA TGNYYK HA TGNYYK HA TGNYYK HA TGNYYK	AGFPPGVVNIVPGYGPTA AGFPPGVVNIVPGYGPTA AGFPPGVVNIVPGYGPTA AGFPPGVVNIVPGFGPTA AGFPPGVVNIVPGYGPTA AGFPPGVVNIVPGYGPTA AGFPPGVVNIVPGYGPTA AGFPPGVVNIVPGYGPTA AGFPPGVVNIVPGYGPTA AGFPPGVVNIVPGYGPTA AGFPPGVVNILPGYGPTA AGFPPGVVNILPGYGPTA AGFPPGVVNILPGYGPTA AGFPPGVVNIVPGFGPTA AGFPPGVVNIVPGGVPTT Y-LOQLYAIVNGGVFTT AGVPPGLFNVVQG-GAAT	G 250 GAAI SSHMIVDKVAFTGSTEVGKLIK GAAI SSHMI I DKVAFTGSTEVGKLIK GAAI SSHMI I DKVAFTGSTEVGKLIK GALSRHNI I DAIAFTGSTEVGKLIK GALSRHNI I DAIAFTGSTEVGKLIK GALSSHMIVDKVAFTGSTEVGKLIK GAAI SSHMIVDKVAFTGSTEVGKLIK GAAI SSHMIVDKVAFTGSTEVGKLIK GAAI SSHMIVDKVAFTGSTEVGKLIK GAAI SSHMIVDKVAFTGSTEVGKLIK GAAI SSHMIVDKVAFTGSTEVGKLIK GAAI SSHMIVDKVAFTGSTEVGKLIK GAAI SSHMIVDKVAFTGSTEVGKLIK GAAI SSHMIVDKVAFTGSTEVGKLIK GAAI SSHMIVDKVAFTGSTEVGKLIK GAAI SSHMIVDKVAFTGSTEVGKLIQ GAAI SSHMIVDKVAFTGSTEVGKLIQ GAAI SSHMIVDKVAFTGSTEVGKLIQ GAAI SSHMIVDKVAFTGSTEVGKLIQ GAAI SSHMIVDKVAFTGSTEVGKLIQ GAAI SSHMIVDKVAFTGSTEVGKLIQ GAAI SSHMIVDKVAFTGSTEVGKLIQ GAAI SSHMIVDKVAFTGSTEVGKLIQ GAAI SSHMIVDKVAFTGSTEVGHLIQ GAAI SSHMIVDKVAFTGSTEVGHLIQ GAAI SSHMIVDKVAFTGSTEVGHLIQ GAAI SSHMIVDKVAFTGSTEVGHLIQ GAAI SSHMIVDKVAFTGSTEVGHLIQ GAAI SSHMIVDKVAFTGSTEVGHLIQ GAAI SSHMIVDKVAFTGSTEVGHLIQ GAAI SSHMIVDKVAFTGSTEVGKLIQ GAAI SSHMIVDKVAFTGSTEVGKIQ GAAI SSHMIVDKVAFTGSTEVGKIQ GAAI SSHMIVDKVAFTGSTEVGKIQ GAAI SSHMIVDKVAFTGSTEVGKIQ GAAI SSHMIVDKVAFTGSTEVGKIQ GAAI SSHMIVDKVAFTGSTEVGKIQ GAAI SSHMIVDKVAFTGSTEVGKIQ GAAI SSHMIVDKVAFTGSTEVGKIQ GAAI SSHMIVDKVAFTGSTEVGKIQ GAAI SSHMICT GAAI SSHMICT GAAI SSHMICT GAAI SSHMICT GAAI SSHMICT GAAI SSHMICT GAAI SSHMICT GAAI SSHMICT GAAI SSHMICT GAAI S
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2 YM1 M2 R2 HO2 YM1 M2 R2 HO2 YM1 M2 R2 H2 CRY H6 HX YM3 M3 R3 H3 MM RM H4 H9 H7	T 200 CGNTVVVK PAEQT PLTALHMGSLIK FA CGNTVVVK PAEQT PLSALHVATLIK FA AGNSVILK PSEKS PLSALRLAGLAK FA MGNVC ILK PAAVT PLNALYPASLCKK CGNTVVK PAEQT PLSALMGSLIK FA CGNTVVK PAEQT PLTALHASLIK FA CGNTVVK PAEQT PLTALYANLIK FA CGNTVVK PAEQT PLTALYANLIK FA TGNVVMK VAEQT PLTALYANLIK FA TGNVVKK FA FA TGNVVKK FA TGNVVK FA TGNVVKK FA TGNVVKK FA TGNVVK FA TGNVVK FA TGNVK FA TGNVVK FA TGNVVK FA TGNVK FA TGNVK FA TGNVVK FA TGNVK FA TGNVK FA TG	AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGFGPTA AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGYGPTA AGFPPGVVN IVPGFGPTA AGFPPGVVN IVPGFGPTA AGFPFGVVN I I DO I I I I I I I I I I I I I I I I I I	G 250 GAALSSHMUVDKVAFTGSTEVGKLIK GAALSSHMUVDKVAFTGSTEVGKLIK GALSSHMUIDKVAFTGSTEVGKLIK GALSSHMUIDKVAFTGSTEVGKLIK GALSSHMUVDKVAFTGSTEVGKLIK GALSSHMUVDKVAFTGSTEVGKLIK GAALSSHMUVDKVAFTGSTEVGKLIK GAALSSHMUVDKVAFTGSTEVGKLIK GAALSSHMUVDKVAFTGSTEVGKLIK GAALSSHMUVDKVAFTGSTEVGKLIK GAALSSHMUVDKVAFTGSTEVGKLIK GAALSSHMUVDKVAFTGSTEVGKLIK GAALSSHMUVDKVAFTGSTEVGKLIK GAALSSHMUVDKVAFTGSTEVGKLIK GAALSSHMUVDKVAFTGSTEVGKLIK GAALSSHMUVDKVAFTGSTEVGKLIQ GAALSSHMUVDKVAFTGSTEVGKLIQ GAALSSHUVDKVAFTGSTEVGKLIQ GAALSSHUVDKVAFTGSTEVGKLIQ GAALSSHUVDKVAFTGSTEVGHLIQ GAALSSHUVDKVAFTGSTEVGHLIQ GAALSSHUVDKVAFTGSTEVGHLIQ GAALSSHUVDKVAFTGSTEVGHLIQ GAALSSHUVDKVAFTGSTEVGHLIQ GAALSSHUVDKVAFTGSTEVGHLIQ GAALSSHUVDKVAFTGSTEVGHLIQ GAALSSHUVDKVAFTGSTEVGHLIQ GAALSSHUVDKVAFTGSTEVGHLIQ GAALSSHUVDKVAFTGSTEVGHLIQ GAALSSHUVDKVAFTGSTEVGHLIQ GAALSSHUVDKVAFTGSTEVGKLVK GAALSSHUVDKVAFTGSTEVGKLVK GAALSSHUVDKVAFTGSTEVGKLVK GAALSSHUVDKVAFTGSTEVGKLVK GAALSSHUVDKVAFTGSTEVGKLVK GAALSSHUVDKVAFTGSTEVGKLVM TELLKER FDHIMYTGSTAVGKIVM TELLKER FDHIMYTGSTAVGKIVM TELLKER FDHILYTGNTAVGKIVM GOFLCOHPVAKVSFTGSVPTGMKIM GOLLEHR FDYIFFTGSPRVGKIVM

	E 268		C 302 Y 296 Q 300 I 303	
			* * **	
BA1	EAAGKSNLKRVSLELGGKSPC	IVFADA-DLDNAVEFAH	QCVFYHQGQCCIAASR	lfv <mark>ees</mark> iydefvrrsv
HO1	BAAGKSNLKRVTLELCGKSPF	IVFADA-DLETALEVTH	DALFYHQGQCCVAASR	LEVEESIYDEFVRRSV
YC1	VDSSESNIKKITI,ELCCKSAH	LVFDDA-NIKKTLPNLV	NGIFKNAGOICSSGSR	IYVOEGIYDELLAAFK
C1	EAAGKINLKRVILELGGKSPN	IIFADA-DLDEAAEFAH	IGLFYHQGQCC IAGS	IFVEEPIYDEFVRRSI
M1	EAAGKSNLKRVTLELGGKSPC	IVFADA-DLDIAVEFAH	H <mark>GVFYHQ</mark> GQ <mark>CC</mark> VAASF	IFVEESVYDEFVKRSV
R1	EAAGKSNLKRVTLELGGKSPC	IVFADA-DLDSAVEFAH	OGVFFHOGOICVAASH	LEVEESIYDEFVRRSV
H1	EAAGKSNLKRVTLELGGKSPC	IVLADA-DLDNAVEFAH	HGVFYHOGOCCIAASE	IFVEESIYDEFVRRSV
RALDH-1	EAAGKSNLKRVTLELGGKSPC	IVFADA-DLDIAVEFAH	HGVFYHQGQCCVAASF	IFVEESVYDEFVRKSV
RALDH1	EAAGK SNLKRVTLELGGK SPC	IVFADA-DLDIAVEFAH	H <mark>GVFYHQ</mark> GQ <mark>CC</mark> VAASI	IFVEESVYDEFVRKSV
RALDH-2	EAAGRSNLKRVTLELGGKSPN	IIFADA-DLDYAVEQAH	QGVFFNQGQCCTAGS	IFVEESIYEEFVKRSV
B2	VAAGK SNI KRVTTE I GGK SPA	TTMSDA - DMDWAVEGAH	FALFFNOGOCCCAGS	TEVOEDIYAEEVERSV
HO2	VAAGRSNLKKVTLELGGKSPN	NIIVSDA-DMDWAVEQAH	FALFFNQGQCCGAGS	TFVQEDVYAEFVERSV
YM1	KVAADT-VKKVTLELGGKSP	NIVFADA-DL <mark>DK</mark> AVKNIA	FGIFYNS <mark>GF</mark> VCCAGSI	IYIQDTVYEEVLQKLK
M2	VAAGSSNLKRVTLELGGKSP	NIIMSBA-DMDWAVEQAH	FALFFNQGQCCCAGS	TFVQENVYDEFVERSV
H2	VAAGSSNLKRVTLELGGKSPI	NTIMSDA-DMDWAVEQAH	FALFFNOGOCCCAGS	TEVOEDUDEEVERSV
CRY	EAAAKSNLKRVTLELGAKNPO	IVFADA-DLDSAVEFAH	QGVFTNQGQSCIAAS	LEVEETIYDEFVQRSV
H6	EAASRSNLKRVTLELGGKNPC	CIVCADA-DLDLAVECAH	QGVFFNQGQCCTAASI	RVFVBEQVYSEFVRRSV
HX	KAACDSNLKRVTLELGGKSPS	SIVLADA-DMEHAVEQCH	EALFENMGOCCCAGS	RTEVEESIYNEFLERTV
M3	AAAAK -HI PUTI BLOCKSPO	CYVDKDC-DLDVACRETA	WCKEMNSCOTCVAPD	TIC DETONETVERIK
R3	AAAAK-HLTPVTLELGGKSPO	CYVDKDC-DLDVACRRIA	WGKFMNSGOTCVAPD	TILCOPSIONOIVEKLK
нз	TAAAK-HLTPVTLELGGKSPO	CYVDKNC-DLDVACRRIA	WGKFMNSGQTCVAPD	YIL <mark>CDPSIQNQIVEKL</mark> K
MM	EAAAK-HLTPVTLELGGKSPC	CYIDRDC-DLDVACRRVA	WGKYMNCGQTCIAPD	YILCEASLQNQIVQKIK
HA	HAAAK-HLTPVTLELGGKSPC	CYIDRDC-DEDVACRRIT	MANEL TOCOVCCNCT	YILCEASSODOIVOKIK
Н9	EMSAK-GIKPVTLELGGKSP	LIIFSDC-DMNNAVKGAL	MANFLTQGQVCCNGT	RVFVQKEILDKFTEEVV
H7	TAAAK - HLTPVTLELGGKNP	CYVDDNC - DPQTVANRVA	WFRYFNAGQTCVAPD	YVLCSPEMQERLLPALQ
H10	EAAAK-HLTPVTLELGGKSP	CYIDKDC-DLDIVCRRI	WGKYMNCGQTCIAPD	YILCEASLONQIVWKIK
		Y 3501 353		
531	The Property of the Provide of the P			
BAL HO1	HRAKKYV-1.GNPL#PGVS0G	FOT DKEOXEKTPOPTER	SKKEGAKLECGGGPWG	NKGYBLOPUVE
nor	FRAKKYV-LCNPL PCVSOC	POT DE POVDETT.DT.TESC	KKEGAKLECCCCPWC	NKCYFTOPTUP
E1	ERAKKYV-LGNPLTPGVSQG OCAONWO-PGHPLDPATTMG	PQIDKEQYDKILDLIESC TLIDCAHADSVHSFIREC	KKEGAKLECGGGPWG	NKGYFIQPTVF LAAAIGPTIF
E1 YC1	ERAKKYV-LGNPLTPGVSQG QQAQNWQ-PGHPLDPATTMG AYLETEIKVGNPFDKANFQG	POIDKEQYDKILDLIESO TLIDCAHADSVHSFIREO AITNRQQFDTIMNYIDIO	KKEGAKLECGGGPWG ESKCOLLLDGRN-AG KKEGAKILTGCEKVG	NKGYFIQPTVF LAAAIGPTIF DKGYFIR PTVF
E1 YC1 C1	ERAKKYV-LGNPLTPGVSQG QQAQNWQ-PGHPLDPATTMG AYLETEIKVGNPFDKANFQG ERAKKYT-LGDPLLPGVQQG	POIDKEQYDKILDLIES TLIDCAHADSVHSFIRE AITNRQQFDTIMNYIDIC POIDKEQFQKILDLIES	KKEGÅKLECGGGPWG SESKGQLLLDGRN-AG SKREGAKILTGGEKVG SKREGAKLECGGGPWG	NKGYFIQPTVF LAAAIGPTIF DKGYFIRPTVF NKGYFIQPTVF
E1 YC1 C1 M1 R1	ERAKKYV-LCNPLTPGVSQG QQAQNWQ-PGHPLDPATTMG AYLETEIKVGNPFDKANFQG ERAKKYT-LGPLLPGVQQG ERAKKYV-LGNPLTPGINQG	POIDKEQYDKILDLIES TLIDCAHADSVHSFIRE AITNRQOFDTIMNYIDIC POIDKEQFQKILDLIES POIDKEQHDKILDLIES	KKEGAKLECGGGPWG SSKGQLLLDGRN-AG KKEGAKILTGCEKVG KKEGAKLECGGGPWG KKEGAKLECGGGRWG	NKGYFIQPTVF LAAAIGPTIF DKGYFIRPTVF NKGYFIQPTVF NKGFFVQPTVF
E1 YC1 C1 M1 R1 S1	ERAKKYV-LGNPLTPGVSOG OQAQNWQ-PGHPLDPATTMG AYLETEIKVGNPFDKANFOG ERAKKYT-LGDPLLPGVQOG ERAKKYV-LGNPLTPGINOG ERAKKYV-LGNPLDSGISOG ERAKKYV-LGNPLTPGVSOG	POIDKEQYDKILDLIES TLIDCAHADSVHSFIRE AITNROOFDTIMNYIDIC POIDKEOFOKILDLIES POIDKEOHDKILDLIES POIDKEOHAKILDLIES	KKEGAKLECGGGPWG SKKGQLLLDGRN-AG KKEGAKILEGGEKVG KKEGAKLECGGGPWG KKEGAKLECGGGRWG SKKEGAKLECGGGRWG	NKGYFIQPTVF LAAAIGPTIF DKGYFIRPTVF NKGYFIQPTVF NKGFFVQPTVF NKGFFVQPTVF NKGYFIQPTVF
E1 YC1 C1 M1 R1 S1 H1	ERAKKYV-LGNPLTPGVSOG OQAONWO-PGHPLDPATTMG AYLETEIKVGNPFDKANFOG ERAKKYT-LGDPLLPGVOOG ERAKKYV-LGNPLTPGINOG ERAKKYV-LGNPLTPGVSOG ERAKKYV-LGNPLTPGVSOG ERAKKYI-LGNPLTPGVTOG	POIDKEQYDKILDLIES TLIDCAHADSVHSFIRE AITNRQQFDTIMNYIDI POIDKEQFQKILDLIES POIDKEQHDKILDLIES POIDKEQHAKILDLIES POIDKEQYEKILDLIES	KKEGAKLECGGGPWG SKKEGLLLJGRN-AG KKEGAKILIGGEKVG KKEGAKLECGGGPWG KKEGAKLECGGGRWG SKKEGAKLECGGGRWG SKKEGAKLECGGGPWG	NKGYFIQPTVF LAAAIGPTIF DKGYFIQPTVF NKGYFIQPTVF NKGFFVQPTVF NKGFFVQPTVF NKGYFIQPTVF
E1 YC1 C1 M1 R1 S1 H1 RALDH-1	ERAKKYV-LGNPLTPGVSQG QQAQNWQ-PGHPLDPATTMG AYLETEIKVGNPFDKANFQG ERAKKYT-LGDPLLPGVQQG ERAKKYV-LGNPLTPGINQG ERAKKYV-LGNPLDSGISQG ERAKKYV-LGNPLTPGVSQG ERAKKYV-LGNPLTPGVTQG ERAKKYV-LGNPLTQGINQG	POIDKEQYDKILDLIES TLIDCAHADSVHSFIRE AITNRQQFDTIMNYIDI POIDKEQFQKILDLIES POIDKEQHDKILDLIES POIDKEQHAKILDLIES POIDKEQYEKILDLIES POIDKEQYDKILDLIES	KKEGAKLECGGGPWG ESKGQLLLDGRN-AG KKEGAKILTGGEKVG KKEGAKLECGGGPWG KKEGAKLECGGGRWG KKEGAKLECGGGPWG SKKEGAKLECGGGPWG SKKEGAKLECGGGPWG	NKGYFIQPTVF LAAAIGPTIF DKGYFIRPTVF NKGYFIQPTVF NKGFFVQPTVF NKGFFVQPTVF NKGYFVQPTVF NKGYFVQPTVF
E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH1 BALDH-2	ERAKKYV-LGNPLTPGVSQG QQAQNWQ-PGHPLDPATTMG AYLETEIKVGNPFDKANFQG ERAKKYT-LGPLLPGVQQG ERAKKYV-LGNPLTPGINQG ERAKKYV-LGNPLTPGVSQG ERAKKYL-LGNPLTPGVTQG ERAKKYV-LGNPLTQGINQG ERAKKYV-LGNPLTQGINQG	POIDKEQYDKILDLIES TLIDCAHADSVHSFIRE AITNROOFDTIMNYIDIO POIDKEOFOKILDLIES POIDKEOHDKILDLIES POIDKEOHAKILDLIES POIDKEOYDKILDLIES POIDKEOHDKILDLIES	KKEGAKLECGGGPWG SKKEGAKLECGGGPWG KKEGAKLECGGGPWG KKEGAKLECGGGPWG KKEGAKLECGGGPWG KKEGAKLECGGGPWG KKEGAKLECGGGPWG KKEGAKLECGGGRWG	NKGYFIQPTVF LAAIGPTIF DKGYFIQPTVF NKGYFIQPTVF NKGFFVQPTVF NKGFFVQPTVF NKGFFVQPTVF NKGFFVQPTVF NKGFFVQPTVF
E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH2 RALDH2	ERAKKYV-LGNPLTPGVSQG QQAQNWQ-PGHPLDPATTMG AYLETEIKVGNPFDKANFQG ERAKKYT-LGDPLLPGVQQG ERAKKYV-LGNPLTPGINQG ERAKKYV-LGNPLTPGVTQG ERAKKYV-LGNPLTPGVTQG ERAKKYV-LGNPLTQGINQG ERAKKYV-LGNPLTQGINQG ERAKKYV-LGNPLTQGINQG ERAKKYV-LGNPLTQGINQG ERAKRI-VGSPFDPTTEQG	POIDKEQYDKILDLIES TLIDCAHADSVHSFIRE AITNROOFDTIMNYIDI POIDKEQFOKILDLIES POIDKEQFOKILDLIES POIDKEQHAKILDLIES POIDKEQYEKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES	KKEGAKLECGGCFWG SKKEGAKLECGGCFWG KKEGAKLECGGCFWG KKEGAKLECGGCFWG KKEGAKLECGGCFWG SKKEGAKLECGGCFWG SKKEGAKLECGGCFWG SKKEGAKLECGGCFWG SKKEGAKLECGGCFWG SVAEGAKLECGGCFWG	NK GYFIQPTVF LAAAIGPTIF DK GYFIQPTVF NK GYFIQPTVF NK GFFVQPTVF NK GYFIQPTVF NK GYFVQPTVF NK GFFVQPTVF NK GFFVQPTVF RK GFFI PTVF RK GFFI PTVF
E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH2 RALDH2 B2	ERAKKYV-LGNPLTPGVSOG OQAQNWQ-PGHPLDPATTMG AYLETEIKVGNPFDKANFOG ERAKKYV-LGNPLTPGINOG ERAKKYV-LGNPLTPGINOG ERAKKYV-LGNPLTPGVSOG ERAKKYV-LGNPLTPGVTQG ERAKKYV-LGNPLTQGINOG ERAKKYV-LGNPLTQGINOG ERAKKYV-LGNPLTQGINOG ERAKRI-VGSPFDPTTEQG ARAKSRV-VGNPFDSRTEQG	POIDKEQYDKILDLIES POIDKEQFORINYIDI POIDKEQFORILDLIES POIDKEQFORILDLIES POIDKEQYEKILDLIES POIDKEQYEKILDLIES POIDKEQYEKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES	KKEGAKLECGGCFWG SKKEGAKLECGGGFWG SKKEGAKLECGGGFWG SKKEGAKLECGGGFWG SKKEGAKLECGGGFWG SKKEGAKLECGGGFWG SKKEGAKLECGGGFWG SKKEGAKLECGGGFWG SKKEGAKLECGGCFWG SKEGAKLECGGCFWG SKEGAKLECGGCFWG SKEGAKLECGGCFWG SKEGAKLECGGCFWG SKEGAKLECGGCFWG	NK GYFIQPTVF LAAAIGPTIF DK GYFIQPTVF NK GYFIQPTVF NK GFFVQPTVF NK GFFVQPTVF NK GYFVQPTVF NK GFFVQPTVF RK GFFIEPTVF RK GFFIEPTVF RK GYFIQPTVF
E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH2 B2 HO2	ERAKKYV-LGNPLTPGVSOG QQAQNWQ-PGHPLDPATTMG AYLETEIKVGNPFDKANFOG ERAKKYV-LGNPLTPGINOG ERAKKYV-LGNPLTPGINOG ERAKKYV-LGNPLTPGVSOG ERAKKYV-LGNPLTPGVTOG ERAKKYV-LGNPLTQGINOG ERAKKYV-LGNPLTQGINOG ERAKRI-VGSPFDPTTEQG ARAKSRV-VGNPFDSRTEQG	POIDKEQYDKILDLIES TLIDCAHADSVHSFIRE AITNRQQFDTIMNYIDI POIDKEQFQKILDLIES POIDKEQHDKILDLIES POIDKEQHAKILDLIES POIDKEQYEKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLISS POIDKEQHDKILDLISS POIDKEQHDKILDLISS POIDKEQHDKILDLISS POIDKEQHDKILSLISS	KKEGAKLECGGCRWG KKEGAKLECGGCRWG KKEGAKLECGGCRWG KKEGAKLECGGCRWG KKEGAKLECGGCRWG KKEGAKLECGGCRWG KKEGAKLECGGCRWG KKEGAKLECGGCRWG KKEGAKLECGGCRWG JVAEGAKLECGGCRGLG JVAEGAKLECGGCAAA JKEEGAKLECGGCAAA	NKGYFIQPTVF LAAAIGPTIF DKGYFIQPTVF NKGYFIQPTVF NKGFFVQPTVF NKGFFVQPTVF NKGYFVQPTVF NKGFFVQPTVF NKGFFIPTVF RKGFFIPTVF RKGFFIPTVF RKGYFIQPTVF DRGYFIQPTVF
E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2 YM1 M2	ERAKKYV-LGNPLTPGVSQG QQAQNWQ-PGHPLDPATTMG AYLETEIKVGNPFDKANFQG ERAKKYT-LGDPLLPGVQQG ERAKKYV-LGNPLTPGINQG ERAKKYV-LGNPLTPGVSQG ERAKKYV-LGNPLTPGVSQG ERAKKYV-LGNPLTQGINQG ERAKKYV-LGNPLTQGINQG ERAKKYV-LGNPLTQGINQG ERAKRI-VGSPFDPTTEQG ARAKRI-VGSPFDPTTEQG ARAKSV-VGNPFDSTEQG ARAKSV-VGNPFDSTEQG ARAKSV-VGNPFDSTEQG ARAKSV-VGNPFDSTEQG	POIDKEQYDKILDLIES TLIDCAHADSVHSFIRE AITNRQQFDTIMNYIDI POIDKEQFQKILDLIES POIDKEQFQKILDLIES POIDKEQHAKILDLIES POIDKEQYEKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIS POIDKEQHDKILDLIS POIDKEQHDKILDLIS POIDKEQHDKILDLIS POIDKEQHDKILDLIS POIDKEQHDKILDLIS POIDKEQHDKILDLIS POIDKEQHDKILDLIS POIDKEQHDKILDLIS POIDKEQHDKILDLIS POIDKEQHDKILDLIS POIDKEQHDKILD	KKEGAKLECGGCPWG KKEGAKLECGGCPWG KKEGAKLECGGCPWG KKEGAKLECGGCPWG KKEGAKLECGGCPWG KKEGAKLECGGCPWG KKEGAKLECGGCPWG KKEGAKLECGGCPWG KKEGAKLECGGCRWG SVAEGAKLECGGCRWG SVAEGAKLECGGCAAA SKEEGKLLCGGGAAA SKEEGAKLCCGGAAA SKEEGAKLCCGGAAA	NK GYFIQPTVF LAAAIGPTIF DK GYFIQPTVF NK GYFIQPTVF NK GFFVQPTVF NK GFFVQPTVF NK GYFVQPTVF NK GFFVQPTVF NK GFFVQPTVF RK GFFIPTVF RK GFFIPTVF RK GYFIQPTVF SK GYFIQPTVF
E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2 YM1 M2 R2	ERAKKYV-LGNPLTPGVSQG QQAQNWQ-PGHPLDPATTMG AYLETEIKVGNPFDKANFQG ERAKKYT-LGDPLLPGVQQG ERAKKYV-LGNPLTPGINQG ERAKKYV-LGNPLTPGVQG ERAKKYV-LGNPLTPGVQG ERAKKYV-LGNPLTQGINQG ERAKKYV-LGNPLTQGINQG ERAKKYV-LGNPLTQGINQG ERAKRI-VGSPFDPTTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG	POIDKEQYDKILDLIES TLIDCAHADSVHSFIRE AITNRQQFDTIMNYIDIC POIDKEQFQKILDLIES POIDKEQHDKILDLIES POIDKEQYEKILDLIES POIDKEQYDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHKILDIS POIDKEQHKLDS POVDETQFKKVLGYIKS POVDETQFKKILGYIKS POVDETQFKKILGYIKS	KKEGAKLECGGCAWG SKKEGAKLECGGCRWG KKEGAKLECGGCRWG KKEGAKLECGGCRWG KKEGAKLECGGCRWG SKKEGAKLECGGCRWG SKKEGAKLECGGCRWG SKKEGAKLECGGCRWG SKKEGAKLECGGCRWG SVAEGAKLECGGCRWG SVAEGAKLECGGCAAA SKEEGAKLECGGCAAA SKEEGAKLECGGCAAA	NKGYFIQPTVF LAAIGPTIF DKGYFIQPTVF NKGYFIQPTVF NKGFFVQPTVF NKGFFVQPTVF NKGYFIQPTVF NKGFFVQPTVF NKGFFVQPTVF NKGFFIPTVF NKGFFIPTVF NKGFFIPTVF NKGYFIQPTVF DRGYFIQPTVF DRGYFIQPTVF
E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH2 B2 HO2 YM1 M2 R2 H2	ERAKKYV-LGNPLTPGVSQG QQAQNWQ-PGHPLDPATTMG AYLETEIKVGNPFDKANFQG ERAKKYV-LGNPLTPGINQG ERAKKYV-LGNPLTPGINQG ERAKKYV-LGNPLTPGVSQG ERAKKYV-LGNPLTPGVTQG ERAKKYV-LGNPLTQGINQG ERAKKYV-LGNPLTQGINQG ERAKRI-VGSPFDPTTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG	PQIDKEQYDKILDLIES AITNRQQFDTIMNYIDI AITNRQQFDTIMNYIDI PQIDKEQFQKILDLIES PQIDKEQFQKILDLIES PQIDKEQHAKILDLIES PQIDKEQYEKILDLIES PQIDKEQYDKILDLIES PQIDKEQHDKILDLIES PQIDKEQHDKILDLIES PQIDKEQHDKILDLIES PQIDKEQHDKILDLISS PQIDKEQHDKILDLISS PQIDKEQHDKILDLISS PQUDETQFKKVLGYIKS AQTSDKQLHKILGYIKS PQVDETQFKKILGYIKS	KKEGAKLECGGGPWG SKKEGAKLECGGGPWG SKKEGAKLECGGGPWG SKKEGAKLECGGGPWG SKKEGAKLECGGGPWG SKKEGAKLECGGGPWG SKKEGAKLECGGGPWG SKKEGAKLECGGGRWG SVAEGAKLECGGGRWG SVAEGAKLECGGCAAA SKEEGAKLECGGAAA SKEEGAKLECGGAAA SKEEGAKLECGGAAA SKEEGAKLECGGAAA SQQEGAKLECGGGAAA	NK GYFIQPTVF LAAAIGPTIF DK GYFIQPTVF NK GYFIQPTVF NK GFFVQPTVF NK GFFVQPTVF NK GYFVQPTVF NK GFFVQPTVF RK GFFIEPTVF RK GFFIEPTVF RK GYFIQPTVF DR GYFIQPTVF DR GYFIQPTVF DR GYFIQPTVF DR GYFIQPTVF
E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2 YM1 M2 R2 H2 R2 H2 CRY	ERAKKYV-LGNPLTPGVSOG QQAQNWQ-PGHPLDPATTMG AYLETEIKVGNPFDKANFOG ERAKKYV-LGNPLTPGINOG ERAKKYV-LGNPLTPGINOG ERAKKYV-LGNPLTPGVSOG ERAKKYV-LGNPLTQGINOG ERAKKYV-LGNPLTQGINOG ERAKKYV-LGNPLTQGINOG ERAKKYV-LGNPLTQGINOG ERAKRI-VGSPFDPTTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG	PQIDKEQYDKILDLIES TLIDCAHADSVHSFIRE AITNRQQFDTIMNYIDI PQIDKEQFQKILDLIES PQIDKEQFQKILDLIES PQIDKEQHAKILDLIES PQIDKEQYEKILDLIES PQIDKEQYEKILDLIES PQIDKEQHDKILDLIES PQIDKEQHDKILDLIES PQIDKEQHDKILDLIES PQIDKEQHDKILDLIES PQIDKEQHDKILDLIES PQUDETQFKKVLGYIKS PQVDETQFKKILGYIKS PQVDETQFKKILGYIKS PQVDETQFKKILGYIKS	KKEGAKLECGGGPWG SKKEGAKLECGGGPWG SKKEGAKLECGGGPWG SKKEGAKLECGGGPWG SKKEGAKLECGGGRWG SKKEGAKLECGGGRWG SKKEGAKLECGGGRWG SKKEGAKLECGGGRWG SKKEGAKLECGGGRWG SKEEGAKLECGGGRAA SKEEGAKLECGGGAAA SKEEGAKLLCGGGAAA SQQEGAKLLCGGGAAA SKQGAKLECGGGPWG	NK GYFIQPTVF LAAAIGPTIF DK GYFIQPTVF NK GYFIQPTVF NK GFFVQPTVF NK GFFVQPTVF NK GFFVQPTVF NK GFFIPTVP RK GFFIPTVP RK GYFIQPTVF DR GYFIQPTVF DR GYFIQPTVF DR GYFIQPTVF NK GYFIQPTVF NK GYFIQPTVF NK GYFIQPTVF NK GYFIQPTVF DR GYFIQPTVF DR GYFIQPTVF DR GYFIQPTVF DR GYFIQPTVF DR GYFIQPTVF DR GYFIQPTVF
E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2 YM1 M2 R2 HO2 YM1 M2 R2 H2 CRY H6 HY	ERAKKYV-LGNPLTPGVSOG QQAQNWQ-PGHPLDPATTMG AYLETEIKVGNPFDKANFOG ERAKKYT-LGDPLLPGVQQG ERAKKYV-LGNPLTPGINOG ERAKKYV-LGNPLTPGVQG ERAKKYV-LGNPLTPGVQG ERAKKYV-LGNPLTQGINQG ERAKKYV-LGNPLTQGINQG ERAKKYV-LGNPLTQGINQG ERAKRI-VGSPFDPTTEQG ARAKSRV-VGNPFDSRTEQG	POIDKEQYDKILDLIES TLIDCAHADSVHSFIRE AITNRQQFDTIMNYIDI POIDKEQFQKILDLIES POIDKEQFQKILDLIES POIDKEQHAKILDLIES POIDKEQYEKILDLIES POIDKEQYEKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POVDETQFKKVLGYIKS POVDETQFKKILGYIKS POVDETQFK	KKEGAKLECGGAME	NK GYFIQPTVF LAAAIGPTIF DK GYFIQPTVF NK GYFIQPTVF NK GFFVQPTVF NK GFFVQPTVF NK GFFVQPTVF NK GFFVQPTVF NK GFFIPTVF RK GYFIQPTVF DR GYFIQPTVF
E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2 YM1 M2 R2 H2 CRY H6 HX YM3	ERAKKYV-LGNPLTPGVSOG QQAQNWQ-PGHPLDPATTMG AYLETEIKVGNPFDKANFOG ERAKKYV-LGNPLTPGINOG ERAKKYV-LGNPLTPGINOG ERAKKYV-LGNPLTPGVSOG BRAKKYV-LGNPLTPGVTOG ERAKKYV-LGNPLTQGINOG ERAKKYV-LGNPLTQGINOG ERAKKYV-LGNPLTQGINOG ERAKKYV-LGNPLTQGINOG ERAKKYV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG BRAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG BRAKSRV-V	POIDKEQYDKILDLIES TLIDCAHADSVHSFIRE AITNRQQFDTIMNYIDIO POIDKEQFQKILDLIES POIDKEQFQKILDLIES POIDKEQHAKILDLIES POIDKEQYDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POVDETQFKKUGYIKS POVDETQFKKILGYIKS POVDETQFK	KKEGAKLECGGAAA KKEGAKLECGGAAA KKEGAKLECGGAAA KKEGAKLECGGGRWG KKEGAKLECGGGRWG KKEGAKLECGGGRWG KKEGAKLECGGGRWG KKEGAKLECGGGRWG GAKLECGGAAA GAKLECGGAAA GQGEGAKLLCGGGAAA GQGEGAKLCCGGAAA GQGEGAKLCCGGAAA GQGEGAKLCCGGAAA GQGEGAKLCCGGAAA GQGEGAKLCCGGAAA	NKGYFIQPTVF LAAAIGPTIF DKGYFIQPTVF NKGYFIQPTVF NKGFFVQPTVF NKGFFVQPTVF NKGFFVQPTVF NKGFFVQPTVF NKGFFIPTVF RKGYFIQPTVF DRGYFIQPTVF DRGYFIQPTVF DRGYFIQPTVF DRGYFIQPTVF DRGYFIQPTVF DRGYFIQPTVF DRGYFIQPTVF DRGYFIQPTVF DR
E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2 YM1 M2 R2 H2 CRY H6 HX YM3 M3	ERAKKYV-LGNPLTPGVSQG QQAQNWQ-PGHPLDPATTMG AYLETEIKVGNPFDKANFQG ERAKKYT-LGDPLLPGVQQG ERAKKYV-LGNPLTPGINQG ERAKKYV-LGNPLTPGVQG ERAKKYV-LGNPLTPGVQG ERAKKYV-LGNPLTQGINQG ERAKKYV-LGNPLTQGINQG ERAKKYV-LGNPLTQGINQG ERAKRI-VGSPFDPTTEQG ARAKSV-VGNPFDSRTEQG	POIDKEQYDKILDLIES TLIDCAHADSVHSFIRE AITNRQQFDTIMNYIDIC POIDKEQFQKILDLIES POIDKEQFQKILDLIES POIDKEQHDKILDLIES POIDKEQYDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POVDETQFKKVLGYIKS POVDETQFKKILGYIKS POVDETQFKILGYIKS POVDETQFKILGYIKS POVDETQFKILGYIKS POVDETQFKILGYIKS POVDETQFKILGYIKS POVDETQFKILGYIKS POVDETQFKILGYIS POVDETQFKILGYIS POVDETQFKILGYIS POVDETQFKILGYIS POVDETQFKILGYIS POVDETQFKILGYIS POVDETQFKI POVDETQFKILGYIS POVDETQFKILGYIS POVDETQFKI POVDETQFKI POVDETQFKI POVDETQFKI POVDETQFKI POVDETQFKI POVDETQFKI POVDETQFKI POVDETQFKI POVDETQFKI POVDETQFKI POVDETQFKI POVDETQFKI POVDETQFKI POVDETQFKI POVDETQFKI POVDETQFKI POVDETQF	XKEGAKLECGGCPWG XKEGAKLECGGCPWG XKEGAKLECGGCPWG XKEGAKLECGGCPWG XKEGAKLECGGCPWG XKEGAKLECGGCPWG XKEGAKLECGGCPWG XKEGAKLECGGCPWG XKEGAKLECGGCRWG YAEGAKLECGGCAA GVAEGAKLECGGCAA GVAEGAKLCGGCAA GVAEGAKLCGGCAA GVAEGAKLCGGCAA GVAEGAKLCGGCAA GVAEGAKLCGGCAA GVAEGAKLCGGCAA GVAEGAKLCGGCAA GVAEGAKLCGGCAA GVAEGAKLCGGCAA GVAEGAKLCGGCAA GVAEGAKLCGGCAA GVAEGAKLCGGCAA GVAEGAKLCGGCAA	NK GYFIQPTVF LAAAIGPTIF DK GYFIQPTVF NK GYFIQPTVF NK GFFVQPTVF NK GFFVQPTVF NK GFFVQPTVF NK GFFVQPTVF NK GFFIPTVF NK GFFIPTVF NK GFFIPTVF DR GYFIQPTVF DR GYFIQPTVF DR GYFIQPTVF DR GYFIQPTVF DR GYFIQPTVF SK GYFIQPTVF NK GYFIQPTVF NK GYFIQPTVF SK GYFIQPTVF SK GYFIQPTVF SK GYFIQPTVF SK GYFVKPTVF SK GYFVKPTVF
E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH-2 B2 HO2 YM1 M2 R2 HO2 YM1 M2 R2 H2 CRY H6 HX YM3 M3 R3	ERAKKYV-LGNPLTPGVSQG QQAQNWQ-PGHPLDPATTMG AYLETEIKVGNPFDKANFQG ERAKKYT-LGDPLLPGVQQG ERAKKYV-LGNPLTPGINQG ERAKKYV-LGNPLTPGVSQG ERAKKYV-LGNPLTQGVQQG ERAKKYV-LGNPLTQGINQG ERAKKYV-LGNPLTQGINQG ERAKKYV-LGNPLTQGINQG ERAKRI-VGSPFDPTTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ERAKKYV-FGNPLTPGVNHG EYAKKRP-VGDPFDFVKTEQG EKAKQRK-VGNPFELDTQQG DYTESLK-VGDPFDEFVFQG KSLKDFYGEDAKQSRDYG	PQIDKEQYDKILDLIES TLIDCAHADSVHSFIRE AITNRQQFDTIMNYIDIO PQIDKEQFQKILDLIES PQIDKEQFQKILDLIES PQIDKEQHAKILDLIES PQIDKEQYEKILDLIES PQIDKEQYDKILDLIES PQIDKEQHDKILDLIES PQIDKEQHDKILDLIES PQIDKEQHDKILDLIES PQIDKEQHDKILDLIES PQIDKEQHDKILDLIES PQUDETQFKKVLGYIKS PQVDETQFKKILGYIKS PQVDETQFKKILGYIKS PQVDETQFKKILGYIKS PQVDETQFKKILGYIKS PQUDETQFKKILGYIKS PQUDETQFKKILGYIKS PQUDETQFKKILGYINT PQINKAQHNKIMELIES PQUDKEQFERVLGYICL PQIDKEQFERVLGYICL PQIDKEQFERVLGYICL PQUDETQFKKILGYINT PQINKAQHNKIMELIES PQIDKEQFERVLGYICL PQIDKEQFERVL	KKEGAKLEGGGPWG SKKEGAKLECGGGPWG SKKEGAKLECGGGPWG SKKEGAKLECGGGPWG SKKEGAKLECGGGRWG SKKEGAKLECGGGRWG SKKEGAKLECGGGRWG SKKEGAKLECGGGRWG SKKEGAKLECGGGRWG SKEGAKLECGGCRWG SKEGAKLECGGCAAA SKEGAKLECGGGAAA SKEGAKLECGGGAAA SQQEGAKLLCGGGAAA SKQQGAKLECGGGAAA SKQQGAKLECGGCAAA SKQQGAKLECGGCAAA SKQQGAKLECGGCAAA SKQQGAKLECGGCAAA SKQQGAKLECGGCAAA SKQQGAKLECGGCAAA SKQQGAKLECGGCAAA SKQQGAKLECGGCAAA SKQQGAKLECGGCAAA SKQQGAKLECGGCAAA SKQQGAKLECGGCAAA	NK GYFIQPTVF LAAAIGPTIF DK GYFIQPTVF NK GYFIQPTVF NK GFFVQPTVF NK GFFVQPTVF NK GFFVQPTVF NK GFFVQPTVF NK GFFIQPTVF RK GFFIQPTVF DR GYFIQPTVF DR GYFIQPTVF SK GYFIQPTVF SK GYFIQPTVF SK GYFIQPTVF SK GYFIXPTVF SK GYFVKPTVF PS RYIAPTII
E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2 YM1 M2 R2 HO2 YM1 M2 R2 H2 CRY H6 HX YM3 M3 R3 H3 MM	ERAKKYV-LGNPLTPGVSOG QQAQNWQ-PGHPLDPATTMG AYLETEIKVGNPFDKANFQG ERAKKYV-LGNPLTPGINQG ERAKKYV-LGNPLTPGINQG ERAKKYV-LGNPLTPGVQG ERAKKYV-LGNPLTQGIQG ERAKKYV-LGNPLTQGIQG ERAKKYV-LGNPLTQGIQG ERAKKYV-LGNPLTQGIQG ERAKKYV-LGNPLTQGIQG ERAKRI-VGSPFDPTTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ERAKKYV-FGNPLTPGVNHG EYAKKRP-VGDPFDVKTEQG EKAKQK-VGNPFELJTQQG OYTESLK-VGDPFDEVFQG KSLKDFYGEDAKQSRDYG KSLKEFYGEDAKQSRDYG	POIDKEQYDKILDLIES TLIDCAHADSVHSFIRE AITNRQQFDTIMNYIDIO POIDKEQFQKILDLIES POIDKEQFQKILDLIES POIDKEQHAKILDLIES POIDKEQYEKILDLIES POIDKEQYEKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POVDETQFKKVLGYIKS POVDETQFKKILGYIKS POVDETQFK	KKEGAKLECGGAAA SKKEGAKLECGGAAA SKKEGAKLECGGGRWG SKKEGAKLECGGGRWG SKKEGAKLECGGGRWG SKKEGAKLECGGGRWG SKKEGAKLECGGGRWG SKKEGAKLECGGGRWG SKKEGAKLECGGGRWG SKEGAKLECGGGRWG SKEGAKLECGGGAAA SKEGAKLECGGAAA	NK GYFIQPTVF LAAAIGPTIF DK GYFIQPTVF NK GYFIQPTVF NK GFFVQPTVF NK GFFVQPTVF NK GFFVQPTVF NK GFFVQPTVF NK GFFIPTVP RK GYFIQPTVF DR GYFIQPTVF DR GYFIQPTVF DR GYFIQPTVF DR GYFIQPTVF SK GYFIQPTVF DR GYFIQPTVF SK GYFIQPTVF DR GYFIQPTVF DR GYFIQPTVF SK GYFIQPTVF SK GYFIQPTVF SK GYFIQPTVF SK GYFIQPTVF SK GYFIQPTVF SK GYFIQPTVF SK GYFIQPTVF SK GYFIQPTVF SK GYFIQPTVF SS RYIAPTII AT RYIAPTII
E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2 YM1 M2 R2 HO2 YM1 M2 R2 H2 CRY H6 HX YM3 M3 R3 H3 MM RM	ERAKKYV-LGNPLTPGVSQG QQAQNWQ-PGHPLDPATTMG AYLETEIKVGNPFDKANFQG ERAKKYV-LGNPLTPGINQG ERAKKYV-LGNPLTPGINQG ERAKKYV-LGNPLTPGVQG ERAKKYV-LGNPLTPGVQG ERAKKYV-LGNPLTQGINQG ERAKKYV-LGNPLTQGINQG ERAKKYV-LGNPLTQGINQG ERAKRI-VGSPFDPTTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG CSAKKQRK-VGNPFELDTQGG YTESLK-VGDPFDEVFQG KSLKDFY-GEDAKQSHDYG KSLKDFY-GEDAKSRDYG FTVKDFY-GENKASPDYF	POIDKEQYDKILDLIES TLIDCAHADSVHSFIRE AITNRQQFDTIMNYIDI POIDKEQFQKILDLIES POIDKEQFQKILDLIES POIDKEQHAKILDLIES POIDKEQYEKILDLIES POIDKEQYEKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POVDETQFKKVLGYIKS POVDETQFKKILGYIKS POVDETQFKKILGYIKS POVDETQFKKILGYIKS POVDETQFKKILGYIKS POVDETQFKKILGYIKS POVDETQFKKILGYIKS POVDETQFKKILGYIKS POVDETQFKKILGYIKS POVDETQFKKILGYIKS POVDETQFKKILGYIKS POVDETQFKKILGYIKS POVDETQFKKILGYIKS POVDETQFKKILGYIKS POVDETQFKKILGYINT POINKAQHNKIMILIES POVDETQFKKILGYIKS POVDETQFKKILS POVDETQFK	XKEGAKLECGGCPWG SKKEGAKLECGGCPWG SKKEGAKLECGGCPWG SKKEGAKLECGGCPWG SKKEGAKLECGGCPWG SKKEGAKLECGGCPWG SKKEGAKLECGGCPWG SKKEGAKLECGGCRWG SKEGAKLECGGCRWG SKEGAKLECGGCRWG SKEGAKLECGGCAAA SKEGAKLECGGCAAA SKEGAKLECGGAAA SKEGAKLECGGAAA SKEGAKLECGGAAA SKEGAKLECGGAAA SKEGAKLECGGAAA SKEGAKLECGCAAA	NKGYFIQPTVF LAAAIGPTIF DKGYFIQPTVF NKGYFIQPTVF NKGFFVQPTVF NKGFFVQPTVF NKGFFVQPTVF NKGFFIPTVF NKGFFIPTVF RKGFFIPTVF DRGYFIQPTVF DRGYFIQPTVF DRGYFIQPTVF DRGYFIQPTVF DRGYFIQPTVF DRGYFIQPTVF DRGYFIQPTVF DRGYFIQPTVF DRGYFIQPTVF DRGYFIQPTVF DRGYFIQPTVF DRGYFIQPTVF DRGYFIQPTVF DRGYFIQPTVF DRGYFIQPTVF DR
E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2 YM1 M2 R2 H2 CRY H6 HX YM3 M3 R3 H3 MM RM H4	ERAKKYV-LGNPLTPGVSOG QQAQNWQ-PGHPLDPATTMG AYLETEIKVGNPFDKANFOG ERAKKYV-LGNPLTPGVQQG ERAKKYV-LGNPLTPGVQQG ERAKKYV-LGNPLTPGVQQG ERAKKYV-LGNPLTPGVQQG ERAKKYV-LGNPLTQGINQG ERAKKYV-LGNPLTQGINQG ERAKKYV-LGNPLTQGINQG ERAKKYV-LGNPLTQGINQG ERAKKYV-LGNPLTQGINQG ERAKKYV-UGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ERAKKYV-FGNPLTPGVNHG EYAKKRP-VGDPFDVKTEQG EXAKQK-VGNPFELDTQQG KSLKDFYGEDAKQSHDYG KSLKDFYGEDAKQSHDYG KSLKEFYGEDAKKSRDYG ETVKDFYGENKASPDYE	POIDKEQYDKILDLIES TLIDCAHADSVHSFIRE AITNRQQFDTIMNYIDIO POIDKEQFQKILDLIES POIDKEQFQKILDLIES POIDKEQHDKILDLIES POIDKEQYDKILDLIES POIDKEQYDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POVDETQFKKVLGYIKS POVDETQFKKILGYIKS POVDETQFKKILGYIKS POVDETQFKKILGYIKS POVDETQFKKILGYIKS POVDETQFKKILGYIKS POVDETQFKKILGYIKS POVDETQFKKILGYIKS POVDETQFKKILGYIKS POVDETQFKKILGYIKS POVDETQFKKILGYIKS POVDETQFKKILGYIKS POVDETQFKKILGYIKS POVDETQFKKILGYIKS POVDETQFKKILGYIS POVDETQFKKILGYIKS POVDETQFKKILGYIKS POVDETQFKKILGYING POVDETQFKKILGYIKS POVDETQFKKILGYING POVDETQFKKILGYING POVDETQFKKILGYING POVDETQFKKILGYING POVDETQFKKILGYIKS POVDETQFKI POVDETQFKKILGYIKS P	XKEGAKLECGGGPWG XKEGAKLECGGGPWG XKEGAKLECGGGPWG XKEGAKLECGGGPWG XKEGAKLECGGGPWG XKEGAKLECGGGPWG XKEGAKLECGGGPWG XKEGAKLECGGGPWG XKEGAKLECGGGRWG XKEGAKLECGGGRWG XAEGAKLECGGGAAA XEEGAKLECGGAAA XKEGAKLECGGAAA GQQEGAKLLCGGGAAA GQQEGAKLLCGGGAAA GQQEGAKLLCGGGAAA GQQEGAKLCGGGAAA GKEGAKLECGGAAA GQQEGAKLCGGGAAA GKEGAKLECGGAAA GKEGAKLECGGAAA GKEGAKLECGGAAA GKEGAKLECGGAAA GKEGAKLECGGAAA GKEGAKLECGGAAA GKEGAKLECGGAAA GKEGAKLECGGAAA C C C C C C C C C C C C C C C C C C	NKGYFIQPTVF LAAAIGPTIF DKGYFIQPTVF NKGYFIQPTVF NKGFFVQPTVF NKGFFVQPTVF NKGFFVQPTVF NKGFFVQPTVF NKGFFIPTVF RKGFFIPTVF RKGFFIPTVF DRGYFIQPTVF DRGYFIQPTVF DRGYFIQPTVF DRGYFIQPTVF DRGYFIQPTVF DRGYFIQPTVF SKGFFINPTVF SKGFFINPTVF SKGFFINPTVF SK
E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2 YM1 M2 R2 H2 CRY H6 HX YM3 M3 R3 H3 MM RM H4 H9	ERAKKYV-LGNPLTPGVSOG QQAQNWQ-PGHPLDPATTMG AYLETEIKVGNPFDKANFQG ERAKKYV-LGNPLTPGINQG ERAKKYV-LGNPLTPGINQG ERAKKYV-LGNPLTPGVSQG BRAKKYV-LGNPLTPGVSQG BRAKKYV-LGNPLTQGINQG ERAKKYV-LGNPLTQGINQG ERAKKYV-LGNPLTQGINQG ERAKRI-VGSPFDPTTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG BRAKSRV-VGNPFDSRTEQG CARAKSRV-VGNPFDSRTEQG CARAKSRV-VGNPFDSRTEQG CARAKSRV-VGNPFDSRTEQG CARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG CARAKSRV-VGNPFDSRTEQG CARAKSRV-VGNPFDSRTEQG CARAKSRV-GNPFDSRTEGG CARAKSRV-GNPFDSRTEGG CARAKSRV-GNPFDSRTEGG CARAKSRV-GNPFDSRTEGG CARAKSRV-GNPFC CARAKSRV-G	POIDKEQYDKILDLIES TLIDCAHADSVHSFIRE AITNRQQFDTIMNYIDIO POIDKEQFQKILDLIES POIDKEQFQKILDLIES POIDKEQHDKILDLIES POIDKEQYDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POVDETQFKKLGYIKS POVDETQFKKVGYIKS POVDETQFKKLGYIKS POVDETQFKKVGYIKS POVDETQFKY POVDETQFKKVGY POVDET	XKEGAKLECGGCPWG XKEGAKLECGGCPWG XKEGAKLECGGCPWG XKEGAKLECGGCPWG XKEGAKLECGGCPWG XKEGAKLECGGCPWG XKEGAKLECGGCPWG XKEGAKLECGGCPWG XKEGAKLECGGCRWG YAEGAKLECGGCAA YAEGAKLECGGCAA YAEGAKLECGGCAA YAEGAKLCGGCAA YAEGAKLCGGCAA YAEGAKLCGGCAA YAEGAKLCGGCAA YAEGAKLCGGCAA YAEGAKLCGGCAA YAEGAKLCGGCAA YAEGAKLCGGCAA YAEGAKLCGGCAA YAEGAKLCGGCAA YAEGAKLCGGCAA YAEGAKLCGGCAA YAEGAKLCGGCAA YAEGAKLCGGCAA YAEGAKLCGCCAA YAEGAKLCGCAA YAEGAKLCGCAA YAEGAKLCGCCGCAA	NKGYFIQPTVF LAAAIGPTIF DKGYFIQPTVF NKGYFIQPTVF NKGFFVQPTVF NKGFFVQPTVF NKGFFVQPTVF NKGFFVQPTVF NKGFFIPTVF RKGFFIPTVF RKGYFIQPTVF DRGYFIQPTVF DRGYFIQPTVF DRGYFIQPTVF DRGYFIQPTVF DRGYFIQPTVF DRGYFIQPTVF DRGYFIQPTVF DRGYFIQPTVF DRGYFIQPTVF DRGYFIQPTVF DRGYFIQPTVF DRGYFIQPTVF DR
E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH-2 B2 HO2 YM1 M2 R2 HO2 YM1 M2 R2 H2 CRY H6 HX YM3 M3 R3 H3 MM RM H4 H9 H7 H10	ERAKKYV-LGNPLTPGVSQG QQAQNWQ-PGHPLDPATTMG AYLETEIKVGNPFDKANFQG ERAKKYV-LGNPLTPGINQG ERAKKYV-LGNPLTPGINQG ERAKKYV-LGNPLTPGVQQG ERAKKYV-LGNPLTQGINQG ERAKKYV-LGNPLTQGINQG ERAKKYV-LGNPLTQGINQG ERAKKYV-LGNPLTQGINQG ERAKKYV-LGNPLTQGINQG ERAKKYV-LGNPLTQGINQG ERAKRI-VGSPFDPTTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ARAKSRV-VGNPFDSRTEQG ERAKKYV-FGNPLTPGVNHG EYAKKPP-VGDPFDVKTEQG EKAKQK-VGNPFELJTQGG OYTESLK-VGDPFDEVFQG KSLKDFY-GEDAKQSDYG KSLKDFY-GEDAKQSDYG KSLKEFY-GEDAKKSPYT FVKDFY-GENKASPYT KQTQRIK-IGDPLLEDTRMG STITRFY-GDDPQSSPNLG	POIDKEQYDKILDLIES TLIDCAHADSVHSFIRE AITNRQQFDTIMNYIDI POIDKEQFQKILDLIES POIDKEQFQKILDLIES POIDKEQHAKILDLIES POIDKEQYEKILDLIES POIDKEQYEKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POIDKEQHDKILDLIES POVDETQFKKVLGYIKS POVDETQFKKILGYKS POVDETQFKKILGYIKS POVDETQFKKILGY	XKEGAKLECGGCPWG SKKEGAKLECGGCPWG SKKEGAKLECGGCPWG SKKEGAKLECGGCPWG SKKEGAKLECGGCPWG SKKEGAKLECGGCPWG SKKEGAKLECGGCPWG SKKEGAKLECGGCPWG SKKEGAKLECGGCRWG SKEGAKLECGGCRWG SKEGAKLECGGCRWG SKEGAKLECGGCAA SKEGAKLECGGCAA SKEGAKLECGGCAA SKEGAKLECGGCAA SKEGAKLECGGCAA SKEGAKLECGGCAA SKEGAKLECGGCAA SKEGAKLECGGCAA SKEGAKLECGGCAA SKEGAKLECGGCAA SKEGAKLECGGCAA SKEGAKLECGGCAA SKEGAKLECGGCAA SKEGAKLECGGCAA SKEGAKLECGGCAA SKEGAKLECGGCAA SKEGAKLECGGCAA SKEGAKLECGGCAA SKEGAKLECGGCAA SKEGAKLECGCAA SKEGAKAA SKEGAKAA SKEGAKAA SKEGAKAA SKEGAKAA SKEGAKAA SKEGAKAA SKEGAKAA SKEGAKAA SKEGAKAA SKEGAAA SKEGAKAA SKEGAAA SKEGAKAA SKEGAAA SKEGAAA SKEGAAA SKEGAAA SKEGAAA SKEGAAA SKEGAAA SKEGAAA SKEGAAA SKEGAAA SKEGAAA SKEGAAA SKEGAAAA SKEGAAA SKEGAAAA SKEGAAAA SKEGAAAA SKEGAAAA SKEGAAAA SKEGAAAAA SKEGAAAAA SKEGAAAA SKEGAAAAAA SKEGAAAAAAAAAA SKEGAAAAAAAAAAAA SKEGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	NK GYFIQPTVF LAAAIGPTIF DK GYFIQPTVF NK GYFIQPTVF NK GFFVQPTVF NK GFFVQPTVF NK GFFVQPTVF NK GFFVQPTVF NK GFFIPTVP RK GFFIPTVF DR GYFIQPTVF DR

		1 400 L					T 450
BA1	SDUTDOMRIAKE	EIFGPVOOIM	KFKSLDD-V	IKRANNTE	TGLSAGIFTN	DIDKAITVS	ALOSCTVWVN-
HO1	SNVSDEMRIAKE	EIFGPVQQIM	KFKSL <mark>DD</mark> -V	IKRANNTT	YGLFAGSFTK	DLDKAITVSZ	ALQACTVWVN-
E1 VC1	VDVDPNASLSRE	EIFGPVLVVT	RFTS-EEQA	LOLANDS	YGLGAAVWTR	DLSRAHRMSH	RLKAGSVFVN-
C1	SNVTDDMRIAKE	EIFGPVOOIM	KFKTIDE-V	IKRANNT	YGLAAAVETK	DIDKALTFA	ALOAGTVWIN-
M1	SNUTDEMRIAKE	EIFGPVQQIM	KFKSVDD-V	IKRANNT	YGLAAGLETK	DLDKAITVS	ALQAGVVWVN-
R1	SNVTDEMRIAKE	EIFGPVQQIM	KFKSIDE-V	I KRANNT I	PYGLAAGVFTK	DLDRAITVS	ALQACTVWVN-
S1 H1	SOVTODMRIAKE	EIFGPVQQIM EIFGPVQQIM	KFKSLDD-V	TERANNT	YGLSAGIFTN	DIDKAITVS	ALQSGTVWVN-
RALDH-1	SNUTDEMRIAKE	EIFGPVQQIM	KFKSIDD-V	IKRANNT'	TYGLAAGVETK	DLDRAITVS	ALOAGVVWVN-
RALDH1	SNVTDEMRIAKE	EIFGPVQQIM	KFKSIDD-V	IKRANNT'	TYGLAAGVETE	DLDRAI TVS:	ALQAGVVWVN-
RALDH-2	SNVTDDMRIAKE	EIFGPVQEII	RFKTMDE-V	TERANNSI	FGLVAAVETN	INKALMVS:	AMOAGTVWIN-
B2	GDLODGMTTAKE	FIFGPVMOII	KEKSMEE-V	VGRANNSI	YGLAAAVET	DI.DKANYI.S	ALOAGTVWIN-
HO2	GDVQDGMTIAKE	EIFGPVMQII	KFKTIEE-V	VGRANNS	KYGLAAAVFT	DLDKANYLS	ALQAGTVWIN-
YM1	ADVKEDMRIVKE	EVFGPIVTVE	KFSTVDE-1	TAMANDS	QYGLAAGI HTM	DINKAVDVS	KRVKAGTVWIN-
M2 R2	GDVKDGMTIAKE	EIFGPVMOII	KFKTIEE-V	VYGRANDS	YGLAAAVFTI	DLDKANYLS	CALCACTVWIN-
H2	GDVODGMTIAKE	EIFGPVMOII	KFKTIEE-V	VGRANNS'	TYGLAAAVFT	DLDKANYLS	OALOAGTVWVN-
CRY	SNVTDDMRIAKE	EIFGPVQQIN	KFKSLDE-	VIKRANNT	YYGLVAGVF TI	DLDKAVTVS	SALQAGTVWVN-
H6	SEVTONMRIAKE	EIFGPVQPII	KFKSIEE-	VI KRANST	DYGLTAAVFTI	NLDKALKLA	SALESGTVWIN-
MAX YM3	GGVQDDMRIAKE	FURCETVEV	KFKKIEE-	UT AMANDS	RYGLAAAVFTI	RDLDKAMYFT	QALQACTVWVN-
M3	VUVDPOSPVMOE	EIFGPVMPIN	CVRSLDE-2	AIKFINOR	EKPLALYVESI	NNDKVIKKMI	AETSSGGVTAND
R3	VDVDPQSPVMQE	EIFGPVMPIN	/CVRSLEE-2	AIQFINQR	EK PLALYVF SI	NNEKVIKKMI.	AETSSGGVTAND
НЗ	TDVDPQSPVMQE	EIFGPVLPI	CVRSLEE-2	AIQFINOR	EKPLALYMES	SNDKVIKKMI	AETSS GGVAAND
RM	TEVEPNSKVMQE	ELFGPILPI	SVKNVEE-	AINFINDR	EKPLALYVFS	KNNKLIKRVI	DETSSGGVTGND
H4	TNCRDDMTCVK	EIFGPVMSI	SFDTEAF	VLERANDT	TFGLAAGVET	RDIORAHRVV	AELOAGTCFIN-
Н9	TNCRDDMTCVK	EIFGPVMSII	LSFDTEAE-	VLERANDT	TFGLAAGVFT	RDIQRAHR VV	AELQAGTCFIN-
H7	VDVQEMEPVMQ	FGPILPI	VNVQSLDE-	AIEFINRR	EKPLALYAFS	NSSQVVKRVL	TQTSSGGFCGND
		TTODTT DT	TTOT THE ATT FTO TO	A TRATE TRATE	DIFT DT AT SETTION	TRATING T T ST ST ST ST	DEM O O O OTTO ONTO
HIU	TOVOPKTKVMQ	EIFGPILPI	PVKNVDE-	AINFINER	EKPLALYVES	H <mark>NHK</mark> LI <mark>KR</mark> MI	DETSSGGVTGND
HIU	V 459 F 4	EIFGPILPI 165	<u>/PVKNVDE</u> -	AINFINER E 487 ↓	S 500	H <mark>NHK</mark> LIKRMI	DETSSGGVTGND
BA1		EIFGPILPI 165 GGFKMSGNG	VPVKNVDE-A	E 487	EKPLALYVFS S 500 TIKISQKNS-	H <mark>NHK</mark> LIKRMI	DETSSGGVTGND
BA1 HO1	V 459 F 4 V 459 F 4 V 459 F 4 V V CYSVVSAQCPE -CYSVVSAQCPE	EIFGPILPI 165 F <mark>GGFKMSGNG</mark> F <mark>GGFKMSGNG</mark>	VPVKNVDE- Relgeygfh Remgeygfh	E 487 E 487 EYTEVKTV EYTEVKTV	S 500 S 500 TIKISQKNS - TVKISQKNS -	H <mark>NHK</mark> LIKRMI	DETSSGGVTGND
BA1 HO1 E1 VC1	V 459 F ↓ 459 F ↓ V - CYSVVSAQCPI - CYGVVSAQCPI - NYNJGDMTVPI - NYNJGDMTVPI	IFGPILPI GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG	VPVKNVDE- Relceycfh Remceycfh Rdkslhale Rmceyvyh	AINFINER E 487 EYTEVKTV EYTEVKTV KFTELKTI	S 500 TIKISQKNS- TVKISQKNS- WISLEA	H <mark>NHK</mark> LIKRMI	DETSSGGVTGND
BA1 HO1 E1 YC1 C1	V 459 F	EIFGPILPI 165 GGFKMSGNG GGFKMSGNG GGFKMSGNG GGVKQSGYG GGFKMSGNG	VPVKNVDE- RELGEYGFH REMGEYGFH RDKSLHALE REMGEEVYH RELGEYGLO	AINFINER E 487 ETEVKTV ETEVKTV KFTELKTI ATEVKAV ETEVKAV	S 500 S 500 TIKISQKNS- TVKISQKNS- WISLEA RIKL TIKIPQKNS-	H NHK LIKRMI	DETSSGGVTGND
BA1 HO1 E1 YC1 C1 M1	CYSVVSAQCPE -CYSVVSAQCPE -CYGVVSAQCPE -NYNDGDMTVPE -TYNDFDSRVPE -CYSAFSAQCPE -CYIMLSAQCPE	FIFGPILPI GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG FGGFKMSGNG	VPVKNVDE-: RELCEYGFH REMCEYGFH RDKSLHALE REMCEEVYH RELCEYGLQ RELCEHGLY	E 487 E 487 E YTEVKTV E YTEVKTV KFTELKTI A YTEVKAV E YTEVKTV E YTELKTV	EKPLALYVFS S 500 ↓ TIKISQKNS- TVKISQKNS- WISLEA RIKL TIKIPQKNS- AMKISQKNS-	HNHKLIKRMI	DETSSGGVTGND
BA1 HO1 E1 YC1 C1 M1 R1	CYSVVSAQCPE -CYSVVSAQCPE -CYGVVSAQCPE -NYNDGDMTVPE -TYNDFDSRVPE -CYSAFSAQCPE -CYIMLSAQCPE -CYIMLSAQCPE	EIFGPILPI GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG	VPVKNUDE- RELGEYGFH RDKSLHALS RELGEYGFU RELGEYGLQ RELGEHGLY REMGEQGVY	E 487 E 487 E YTEVKTV E YTEVKTV KFTELKTI À YTEVKAV E YTEVKTV E YTELKTV E YTELKTV	EKPLALYVFS S 500 TIKISQKNS- TVKISQKNS- WISLEA RIKL TIKIPQKNS- AMKISQKNS- AMKISQKNS-	HNHKLIKRMI	DETSSGGVTGND
H10 BA1 H01 E1 YC1 C1 M1 R1 S1 H1	V 459 F V 459 F - CYSVUSAQCPH - CYGVUSAQCPH - NYNDGDMTVPH - TYNDFDSRVPH - CYSAFSAQCPH - CYSAFSAQCPH - CYLLLSVQCPH - CYLLLSVQCPH - CYGVUSAQCPH	GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG	VPVKNUDE-: Relgeygfh Roksihale Regeygfh Regeygly Reigehgiy Reigehgiy Reigeygfh Reigeygfh	AINFINER E 487 EYTEVKTV EYTEVKTV KFTELKTI AYTEVKAV EYTEVKTV EYTELKTV EYTELKTV EYTEVKTV	KPLALYVFS S 500 TIKISQKNS- TVKISQKNS- WISLEA RIKL TIKIPQKNS- AMKISQKNS- AMKISQKNS- TIKISQKNS- TIKISQKNS-	HNHKLIKRMI	DETSSGGVTGND
H10 BA1 H01 E1 YC1 C1 M1 R1 S1 H1 RALDH-1	V 459 F V 459 F - CYSVVSAQCPH - CYGVVSAQCPH - NYNDGDMTVPH - TYNDFDSRVPH - CYSAFSAQCPH - CYSAFSAQCPH - CYLTLSVQCPH - CYSVVSAQCPH - CYGVVSAQCPH - CYMILSAQCPH	GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG	VPVKNVDE-: RELGEYGFH RDKSLHALE REMGEEVYH RELGEYGLQ RELGEHGLY RELGEYGFH RELGEYGFH RELGEYGFH RELGEYGFH	AINFINER E 487 EYTEVKTV EYTEVKTV KFTELKTI AYTEVKAV EYTELKTV EYTELKTV EYTELKTV EYTEVKTV EYTEVKTV EYTELKTV	S 500 TIKISQKNS- TVKISQKNS- WISLEA RIKI TIKIPQKNS- AMKISQKNS- TIKISQKNS- TVKISQKNS- AMKISQKNS-	HNHKLIKRMI	DETSSGGVTGND
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH1	V 459 F - CYSVVSAQCPI - CYSVVSAQCPI - CYGVVSAQCPI - TYNDFDSRVPI - CYSAFSAQCPI - CYIMLSAQCPI - CYSVVSAQCPI - CYGVVSAQCPI - CYMILSAQCPI - CYMILSAQCPI	GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG	VPVKNUDE -: RELGEYGFH REMGEYGFH REMGEFVYH RELGEYGLQ RELGEHGLY RELGEYGFH RELGEYGFH RELGEHGLY RELGEHGLY	AINFINER E 487 E YTEVKTV E YTEVKTV KFTELKTI AYTEVKTV E YTELKTV E YTELKTV E YTEVKTV E YTEVKTV E YTEVKTV E YTEVKTV E YTELKTV	EKPLALYVFS S 500 TIKISQKNS- TVKISQKNS- WISLEA RIKL TIKIPQKNS- AMKISQKNS- TIKISQKNS- TIKISQKNS- TVKISQKNS- AMKISQKNS- AMKISQKNS-	HNHKLIKRMI	DETSSGGVTGND
HIO BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 BAIDH2	V 459 F V 459 F CYSVVSAQCPH -CYGVVSAQCPH -NYNDGDMTVPH -TYNDPDSRVPH -CYSAFSAQCPH -CYSAFSAQCPH -CYSAFSAQCPH -CYSVVSAQCPH -CYSVVSAQCPH -CYSVVSAQCPH -CYMILSAQCPH -CYMILSAQCPH -CYMILSAQCPH -CYMILSAQCPH	GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG	VPVKNUDE- RELGEYGFH REMGEYGFH RDKSLHALE REMGEEVYH RELGEYGLQ RELGEHGLY RELGEYGFH RELGEHGLY RELGEHGLY RELGEHGLY RELGEHGLY	AINFINER E 487 E YTEVKTV EYTEVKTV KFTELKTI AYTEVKAV EYTELKTV EYTELKTV EYTEVKTV EYTEVKTV EYTEVKTV EYTELKTV EYTELKTV	EKPLALYVFS S 500 TIKISQKNS- TVKISQKNS- WISLEA RIKL RIKL TIKIPQKNS- AMKISQKNS- TVKISQKNS- AMKISQKNS- AMKISQKNS- AMKISQKNS- TVKIPQKNS-	HNHKLIKRMI	DETSSCGVTGND
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH1 RALDH2 RALDH2 B2	V 459 F 4 V 459 C 4 V 459 C 4 V 459 C 4 V 50 C 4 V 5	GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG FGGFKMSGNG FGGFKMSGNG FGGFKMSGNG FGGFKMSGNG FGGFKMSGNG	VPVKNVDE-: RELGEYGFH RDKSLHALE REMGEYGFU RELGEYGLQ RELGEYGFU RELGEYGFU RELGEYGFH RELGEYGFH RELGEHGLY RELGEHGLY REMGEFGLR REMGEFGLR	AINFINER E 487 EYTEVKTV EYTEVKTV KFTELKTI AYTEVKAV EYTELKTV EYTELKTV EYTELKTV EYTELKTV EYTELKTV EYTELKTV EYTELKTV EYSEVKTV EYSEVKTV	EKPLALYVFS S 500 TIKISQKNS- TVKISQKNS- WISLEA RIKL RIKL RIKLSQKNS- AMKISQKNS- TKISQKNS- TVKISQKNS- AMKISQKNS- TVKISQKNS- TVKIPQKNS- TVKIPQKNS- TVKIPQKNS- TVKIPQKNS-	HNHKLIKRMI	DETSSGGVTGND
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2	V 459 F - CYSVVSAQCPH - CYGVVSAQCPH - CYGVVSAQCPH - NYNDGDMTVPH - TYNDFDSRVPH - CYSAFSAQCPH - CYSAFSAQCPH - CYSAFSAQCPH - CYSVVSAQCPH - CYSVVSAQCPH - CYSVVSAQCPH - CYMILSAQCPH - CYMILSAQCPH - CYMILSAQCPH - CYNALNAQSPH - CYNVFGAQSPH - CYDVFGAQSPH	GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG	VPVKNVDE- RELGEYGFH RDKSLHALE REMGEYVYH RELGEYGL RELGEHGLY RELGEYGFH RELGEHGLY RELGEHGLY RELGEHGLY RELGEHGLY RELGEYGLQ RELGEYGLQ	AINFINER E 487 EYTEVKTV EYTEVKTV KFTELKTI AYTEVKAV EYTELKTV EYTELKTV EYTEVKTV EYTELKTV EYTELKTV EYTELKTV EYTELKTV EYTELKTV EYSEVKTV AYTEVKTV	EKPLALYVFS S 500 TIKISQKNS- TVKISQKNS- WISLEA RIKL RIKL TIKIPQKNS- AMKISQKNS- TKISQKNS- TVKISQKNS- TVKIPQKNS- TVKIPQKNS- TVKIPQKNS- TVKIPQKNS- TVKIPQKNS-	HNHKLIKRMI	DETSSGGVTGND
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2 YM1	V 459 F - CYSVVSAQCPI - CYGVVSAQCPI - NYNDGDMTVPI - TYNDFDSRVPI - CYSAFSAQCPI - CYSAFSAQCPI - CYNLISAQCPI - CYNLISAQCPI - CYMILSAQCPI - CYMILSAQCPI - CYMILSAQCPI - CYMILSAQCPI - CYMILSAQCPI - CYNALNAQSPI - CYNVFGAQSPI - CYDVFGAQSPI - CYNVFGAQSPI - CYNVFGAQSPI - CYNVFGAQSPI	GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG FGGFKMSGNG FGGFKMSGNG FGGFKMSGNG FGGFKMSGNG	VPVKNVDE- RELGEYGFH RDKSLHALE REMGEEVYH RELGEYGLQ RELGEHGLY RELGEHGLY RELGEHGLY RELGEHGLY RELGEHGLY RELGEHGLY RELGEHGLY REMGEFGLR RELGEYGLQ RELGEYGLQ RELGEYGLQ REMGEALS	AINFINER E 487 EYTEVKTV EYTEVKTV KFTBLKTI AYTEVKAV EYTELKTV EYTELKTV EYTELKTV EYTELKTV EYTELKTV EYTELKTV EYTELKTV EYTELKTV EYTELKTV EYSEVKTV AYTEVKTV AYTEVKTV AYTEVKTV	KPLALYVFS S 500 TIKISQKNS- TVKISQKNS- WISLEA RIKL TIKIPQKNS- AMKISQKNS- TIKISQKNS- TVKISQKNS- TVKISQKNS- AMKISQKNS- TVKIPQKNS- TVKIPQKNS- TVKIPQKNS- TVKIPQKNS- TIKVPQKNS- TIKVPQKNS- TIKVPQKNS- TIKVPQKNS-	HNHKLIKRMI	DETSSGGVTGND
HIU BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2 YM1 M2 P2	V 459 F 4 - CYSVVSAQCPE - CYSVVSAQCPE - NYNDGDMTVPE - TYNDFDSRVPE - CYSAFSAQCPE - CYSAFSAQCPE - CYSAFSAQCPE - CYSVVSAQCPE - CYSVVSAQCPE - CYSVVSAQCPE - CYSVVSAQCPE - CYSVVSAQCPE - CYSVVSAQCPE - CYSVVSAQCPE - CYNLLSAQCPE - CYNLLSAQCPE - CYNLLSAQCPE - CYNLLSAQCPE - CYNLLSAQCPE - CYNLSAQCPE - C	GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG	VPVKNVDE- RELGEYGFH REMGEYGFH RDKSLHALS REMGEYGLQ RELGEYGLQ RELGEYGFH RELGEYGFH RELGEHGLY REMGEFGLR RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ	AINFINER E 487 E 487 E YTEVKTV EYTEVKTV KFTELKTI AYTEVKAV EYTEVKTV EYTEVKTV EYTEVKTV EYTEVKTV EYTEVKTV EYTEVKTV EYSEVKTV AYTEVKTV AYTEVKTV AYTEVKTV	EKPLALYVFS S 500 TIKISQKNS- TVKISQKNS- WISLEA RIKL RIKL RIKL TIKIPQKNS- AMKISQKNS- TKISQKNS- TVKISQKNS- TVKIPQKNS- TVKIPQKNS- TVKIPQKNS- TKVPQKNS- TKVPQKNS- TIKVPQKNS- TIKVPQKNS-	HNHKLIKRMI	
HIU BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH1 RALDH2 RALDH2 B2 HO2 YM1 M2 R2 H2	V 459 F 4 V 459 V 5 V 5 V 5 V 5 V 5 V 5 V 5 V 5	GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG GGFKMSGNG FGGFKMSGNG FGGFKMSGNG FGGFKMSGNG FGGFKMSGNG FGGFKMSGNG FGGFKMSGNG FGGFKMSGNG FGGFKMSGNG FGGFKMSGNG FGGFKMSGNG FGGFKMSGNG FGGFKMSGNG	VPVKNVDE-: RELGEYGFH RDKSLHALE REMGEEVYH RELGEYGLQ RELGEYGLQ RELGEYGFH RELGEYGFH RELGEYGFH RELGEYGFH RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ	AINFINER E 487 E 487 E YTEVKTV EYTEVKTV KFTELKTI AYTEVKTV EYTELKTV EYTELKTV EYTEVKTV EYTELKTV EYTELKTV EYTELKTV EYTELKTV EYTELKTV EYSEVKTV AYTEVKTV AYTEVKTV AYTEVKTV	EKPLÄLYVFS S 500 TIKISQKNS- TVKISQKNS- WISLEA RIKL TIKIPQKNS- AMKISQKNS- TKISQKNS- TKISQKNS- TVKISQKNS- TVKIPQKNS- TVKIPQKNS- TVKIPQKNS- TVKVPQKNS- TVKVPQKNS- TVKVPQKNS-	HNHKLIKRMI	DETSSGGVTGND
HIU BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2 YM1 M2 R2 H2 CRY	V 459 F. - CYSVVSAQCPH - CYSVVSAQCPH - CYGVVSAQCPH - NYNDGDMTVPH - TYNDFDSRVPH - CYSAFSAQCPH - CYIMLSAQCPH - CYIMLSAQCPH - CYMILSAQCPH - CYMILSAQCPH - CYMILSAQCPH - CYMILSAQCPH - CYNALNAQSPH - CYNVFGAQSPH - CYDVFGAQSPH - CYDVFGAQSPH - CYDVFGAQSPH - CYDVFGAQSPH - CYDVFGAQSPH - CYDVFGAQSPH	GGFKMS GNG GGFKMS GNG GGGFKMS GNG	VPVKNVDE-: RELGEYGFH RDKSLHALE REMGEYGFH RDKSLHALE REMGEYGLV RELGEYGL RELGEYGFH RELGEHGLY RELGEHGLY RELGEHGLY RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ	AINFINER E 487 EYTEVKTV EYTEVKTV KFTELKTI AYTEVKAV EYTEVKTV EYTELKTV EYTELKTV EYTELKTV EYTELKTV EYTELKTV EYTELKTV EYTEVKTV AYTEVKTV AYTEVKTV AYTEVKTV AYTEVKTV	EKPLALYVFS S 500 TIKISQKNS- TVKISQKNS- WISLEA RIKL RIKL RIKLSQKNS- AMKISQKNS- TKISQKNS- TVKISQKNS- TVKIPQKNS- TVKIPQKNS- TVKIPQKNS- TVKVPQKNS- TVKVPQKNS- TVKVPQKNS- TVKVPQKNS- TVKVPQKNS- TVKVPQKNS- TVKVPQKNS- TVKVPQKNS- TVKVPQKNS- TVKVPQKNS- TVKVPQKNS- TVKVPQKNS- TVKVPQKNS- TVKVPQKNS-	HNHKLIKRMI	
HIU BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2 YM1 M2 R2 H2 CRY H6 UW	V 459 F. - CYSVVSAQCPI - CYGVVSAQCPI - NYNDGDMTVPI - TYNDFDSRVPI - CYSAFSAQCPI - CYSAFSAQCPI - CYIMLSAQCPI - CYIMLSAQCPI - CYIMLSAQCPI - CYMILSAQCPI - CYMILSAQCPI - CYMILSAQCPI - CYMILSAQCPI - CYMALNAQSPI - CYDVFGAQSPI - CYDVFGAQSPI - CYDVFGAQSPI - CYDVFGAQSPI - CYDVFGAQSPI - CYDVFGAQSPI - CYDVFGAQSPI - CYDVFGAQSPI - CYNALNAQAPI	GGFKMS GNG GGFKMS GNG GGGFKMS GNG GGGFKMS GSG GGGFKMS GSG GGGFKMS GSG GGGFKMS GSG GGGFKMS GNG GGGFKMS GNG	VPVKNVDE- RELGEYGFH REMGEYGFH RDKSLHALE REMGEFVYH RELGEYGLQ RELGEHGLY RELGEHGLY RELGEHGLY RELGEHGLY RELGEHGLY RELGEHGLY RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ	AINFINER E 487 EYTEVKTV EYTEVKTV KFTELKTI AYTEVKAV EYTELKTV EYTELKTV EYTELKTV EYTELKTV EYTELKTV EYTELKTV EYTELKTV EYTEVKTV AYTEVKTV AYTEVKTV AYTEVKTV AYTEVKTV AYTEVKTV AYTEVKTV AYTEVKTV	EKPLÄLYVFS S 500 TIKISQKNS- TVKISQKNS- WISLEA RIKL RIKL AMKISQKNS- TIKIPQKNS- TKISQKNS- TVKISQKNS- TVKISQKNS- TVKIPQKNS- TVKIPQKNS- TVKIPQKNS- TVKIPQKNS- TVKIPQKNS- TVKVPX		DETSSGGVTGND
HIU BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2 YM1 M2 R2 H2 CRY H6 HX YM3	V 459 F. - CYSVVSAQCPH - CYSVVSAQCPH - CYGVVSAQCPH - TYNDFDSRVPH - TYNDFDSRVPH - CYSAFSAQCPH - CYSAFSAQCPH - CYSVVSAQCPH - CYSVFGAQSPH - CYNALVAQAPH - TYNNPHONUPH	GGFKMS GNG GGFKMS GNG FGGFKMS GSG FGGFKMS GSG FGGFKMS GNG FGGFKMS GNG FGGFKMS GNG FGGFKMS GNG FGGFKMS GNG FGGFKMS GNG	VPVKNVDE- RELGEYGFH REMGEYGFH REMGEYGFH RELGEYGLQ RELGEHGLY RELGEHGLY RELGEHGLY RELGEHGLY RELGEHGLY RELGEHGLY RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ	AINFINER E 487 EYTEVKTV EYTEVKTV KFTELKTI AYTEVKTV EYTELKTV EYTELKTV EYTELKTV EYTELKTV EYTELKTV EYTELKTV EYTEVKTV EYSEVKTV AYTEVKTV AYTEVKTV AYTEVKTV AYTEVKTV EYTEVKTV EYTEVKTV	S 500 TIKISQKNS- TVKISQKNS- TVKISQKNS- RIKI RIKI TIKIPQKNS- TKISQKNS- TKISQKNS- TKISQKNS- TKISQKNS- TVKIPQKNS- TVKIPQKNS- TVKIPQKNS- TVKIPQKNS- TVKIPQKNS- TVKVPQK		
HIU BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2 YM1 M2 R2 HO2 YM1 M2 R2 CRY H6 HX YM3 M3	V 459 F 4 - CYSVVSAQCPI - CYSVVSAQCPI - CYGVVSAQCPI - TYNDFDSRVPI - TYNDFDSRVPI - CYSAFSAQCPI - CYSAFSAQCPI - CYSVVSAQCPI - CYSVVSAQCPI - CYSVVSAQCPI - CYSVVSAQCPI - CYSVVSAQCPI - CYSVVSAQCPI - CYSVVSAQCPI - CYNILSAQCPI - CYNILSAQ	GGFKMS GNG GGFKMS GNG FGGFKMS GNG FGGFKMS GNG FGGFKMS GSG FGGFKMS GSG FGGFKMS GSG FGGFKMS GNG FGGFKMS GNG FGGFKMS GSG FGGFKMS GNG FGGFKMS GNG	VPVKNVDE- RELGEYGFH REMGEYGFH RDKSLHALE REMGEEVYH RELGEYGLQ RELGEYGLQ RELGEYGFH RELGEYGFH RELGEYGFH RELGEYGFU REMGEFGLR RELGEYGLQ	AINFINER E 487 E 487 E YTEVKTV EYTEVKTV EYTELKTI EYTEVKTV EYTELKTV EYTELKTV EYTELKTV EYTEVKTV EYTEVKTV EYTEVKTV AYTEVKTV AYTEVKTV AYTEVKTV EYTEVKTV AYTEVKTV EYTEVKTV AYTEVKTV EYTEVKTV EYTEVKTV	EKPLALYVFS S 500 TIKISQKNS- TVKISQKNS- WISLEA RIKL RIKLSQKNS- AMKISQKNS- TKISQKNS- TKISQKNS- TVKISQKNS- TVKIPQKNS- TVKIPQKNS- TVKIPQKNS- TVKVPQKNS-	HNHKLIKRMI	DETSSGGVTGND
BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH2 B2 HO2 YM1 M2 R2 HO2 YM1 M2 R2 H2 CRY H6 HX YM3 M3 R3	V 459 F. - CYSVVSAQCPH - CYSVVSAQCPH - CYGVVSAQCPH - NYNDGDMTVPH - TYNDFDSRVPH - CYSAFSAQCPH - CYIMLSAQCPH - CYIMLSAQCPH - CYIMLSAQCPH - CYMILSAQCPH - CYMILSAQCPH - CYMILSAQCPH - CYNALNAQSPH - CYNVFGAQSPH - CYDVFGAQSPH - CYDV	GGFKMS GNG GGFKMS GNG GGFKMS GNG GGFKMS GNG GGFKMS GNG GGFKMS GNG GGFKMS GNG GGFKMS GNG FGGFKMS GNG	VPVKNVDE- RELGEYGFH REMGEYGFH RDKSLHALE REMGEYGLY RELGEYGLY RELGEYGFH RELGEYGFH RELGEHGLY RELGEHGLY RELGEHGLY RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLA RELGEYGLA	AINFINER E 487 EYTEVKTV EYTEVKTV KFTELKTI AYTEVKAV EYTELKTV EYTELKTV EYTELKTV EYTELKTV EYTELKTV EYTEVKTV EYTEVKTV AYTEVKTV AYTEVKTV AYTEVKTV AYTEVKTV AYTEVKTV AYTEVKTV AYTEVKTV AYTEVKTV AYTEVKTV AYTEVKTV AYTEVKTV AYTEVKTV AYTEVKTV AYTEVKTV AYTEVKTV AYTEVKTV	EKPLALYVFS S 500 TIKISQKNS- TVKISQKNS- WISLEA RIKL RIKLSQKNS- AMKISQKNS- TIKISQKNS- TKISQKNS- TVKISQKNS- TVKIPQKNS- TVKIPQKNS- TVKVPQKNS- T	HNHKLIKRMI	DETSSGGVTGND
HIU BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2 YM1 M2 R2 HO2 YM1 M2 R2 H2 CRY H6 HX YM3 M3 R3 H3 H3	V 459 F - CYSVVSAQCPH - CYGVVSAQCPH - CYGVVSAQCPH - NYNDGDMTVPH - TYNDFDSRVPH - CYSAFSAQCPH - CYSAFSAQCPH - CYSAFSAQCPH - CYMLSAQCPH - CYMLSAQCPH - CYMLSAQCPH - CYMLSAQCPH - CYMLSAQCPH - CYMLSAQCPH - CYMLSAQCPH - CYMVFGAQSPH - CYMVFGAQSPH - CYDVFGAQSPH - CYNALYAQAPH - TYNIVTCHTPH - TYNNFHQNVPH VIVHITVPTLPH VIVHITUPTLPH	GGFKMS GNG GGFKMS GNG FGGFKMS GNG	VPVKNVDE- RELGEYGFH REMGEYGFH RDKSLHALE REMGEYVYH RELGEYGLYH RELGEYGLYH RELGEYGLY RELGEYGLQ RELGEYGLY	AINFINER E 487 EYTEVKTV EYTEVKTV KFTELKTI AYTEVKAV EYTELKTV EYTELKTV EYTELKTV EYTELKTV EYTELKTV EYTELKTV EYTEVKTV EYTEVKTV AYTEVKTV AYTEVKTV AYTEVKTV AYTEVKTV EYTEVKTV EYTEVKTV EYTEVKTV EYTEVKTV FYTEVKTV FYTEVKTV FYTEVKTV TFSHRRSC TFSHRRSC	EKPLALYVFS S 500 TIKISQKNS- TVKISQKNS- WISLEA RIKL RIKLSQKNS- AMKISQKNS- TIKISQKNS- TKISQKNS- TVKIPQKNS- TVKIPQKNS- TVKIPQKNS- TVKIPQKNS- TVKIPQKNS- TVKVPQKNS- T	HNHKLIKRMI	DETSSGGVTGND
HIU BA1 HO1 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 HO2 YM1 M2 R2 H2 CRY H6 HX YM3 M3 R3 H3 MM RM	V 459 F. - CYSVVSAQCPH - CYGVVSAQCPH - CYGVVSAQCPH - TYNDFDSRVPH - TYNDFDSRVPH - CYSAFSAQCPH - CYSAFSAQCPH - CYSVVSAQCPH - CYSVSAQCPH - CYSVFGAQSPH - CYSVFGAQSPH	GGFKMS GNG GGFKMS GNG FGGFKMS GNG	VPVKNVDE- RELGEYGFH REMGEYGFH RDKSLHALE REMGEFVYH RELGEYGLQ RELGEHGLY RELGEHGLY RELGEHGLY RELGEHGLY RELGEHGLY RELGEHGLY RELGEHGLY RELGEYGLQ	AINFINER E 487 EYTEVKTV EYTEVKTV KFTBLKTI AYTEVKAV EYTEVKAV EYTEVKTV EYTELKTV EYTELKTV EYTELKTV EYTELKTV EYTELKTV EYTEVKTV EYTEVKTV AYTEVKTV AYTEVKTV AYTEVKTV AYTEVKTV AYTEVKTV AYTEVKTV AYTEVKTV EYTEVKTV AYTEVKTV AYTEVKTV AYTEVKTV EYTEVKTV AYTEVKTV	EKPLÄLYVFS S 500 TIKISQKNS- TVKISQKNS- WISLEA RIKI RIKISQKNS- AMKISQKNS- TIKISQKNS- TVKISQKNS- TVKISQKNS- TVKISQKNS- TVKIPQKNS- TVKIPQKNS- TVKIPQKNS- TVKIPQKNS- TVKIPQKNS- TVKVPQKNS-	NKARYPPS PZ NKARYPPS PZ NKARYPPS PZ NKLRYPPS PZ	LETSSGGVTGND
H10 BA1 H01 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH2 B2 H02 YM1 M2 R2 H02 YM1 M2 R2 H2 CRY H6 HX YM3 M3 R3 H3 MM RM H4	V 459 F.4 - CYSVVSAQCPH - CYSVVSAQCPH - CYGVVSAQCPH - TYNDFDSRVPH - TYNDFDSRVPH - CYSAFSAQCPH - CYSAFSAQCPH - CYSVVSAQCPH - CYSVVSAQ	GGFKMS GNG GGFKMS GNG FGGFKMS GNG FGGFGS GNS GMG FGGFGS GNS GMG FGGVGNS GMG FGGVGNS GMG FGGVGNS GMG FGGVGNS GMG FGGVGNS GMG	VPVKNVDE- RELGEYGFH REMGEYGFH RDKSLHALE REMGEØYGFU RELGEYGLQ RELGEHGLY RELGEHGLY RELGEHGLY RELGEHGLY RELGEHGLY RELGEHGLY RELGEHGLY RELGEHGLY RELGEHGLY RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLY	AINFINER E 487 EYTEVKTV EYTEVKTV EYTEVKTV EYTELKTI AYTEVKTV EYTELKTV EYTELKTV EYTELKTV EYTELKTV EYTELKTV EYTEVKTV EYTEVKTV AYTEVKTV AYTEVKTV AYTEVKTV AYTEVKTV AYTEVKTV AYTEVKTV EYTEVKTV AYTEVKTV EYTEVKTV AYTEVKTV AYTEVKTV EYTEVKTV EYTEVKTV EYTEVKTV AYTEVKTV EYTEVKTV	EKPLALYVFS S 500 TIKISQKNS - TVKISQKNS - RIKL RIKL TIKIPQKNS - TIKISQKNS - TIKISQKNS - TIKISQKNS - TVKIPQKNS - TVKIPQKNS - TVKIPQKNS - TVKIPQKNS - TVKVPQKNS	HNHKLIKRMI NKARYPPSP HKARYPPSP LKVRYPPSP NKLRYPPNS	DETSSGGVTGND
H10 BA1 H01 E1 YC1 C1 M1 R1 S1 H1 RALDH-1 RALDH-1 RALDH-2 RALDH-2 RALDH2 B2 H02 YM1 M2 R2 H02 YM1 M2 R2 CRY H6 HX YM3 M3 R3 H3 MM RM H4 H9	V 459 F. - CYSVVSAQCPI - CYSVVSAQCPI - TYNDFDSRVPI - TYNDFDSRVPI - CYSAFSAQCPI - CYSAFSAQCPI - CYSAFSAQCPI - CYSVVSAQCPI - CYSVSAQCPI - NYNVSPVELI	GGFKMS GNG GGFKMS GNG FGGFKMS GNG FGGFK	VPVKNVDE- RELGEYGFH REMGEYGFH RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLG RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYGLQ RELGEYALA RELGEYALA RELGEYALA SYHGKKSFE SYHGKKSFE SYHGKSFI RYHGKYSFI RENGRVTIE RENGRVTIE	AINFINER E 487 EYTEVKTV EYTEVKTV KFTELKTI AYTEVKAV EYTEVKTV EYTELKTV EYTELKTV EYTELKTV EYTEVKTV EYTEVKTV EYTEVKTV AYTEVKTV	EKPLÄLYVES S 500 TIKISQKNS- TVKISQKNS- WISLEA RIKL RIKLSQKNS- AMKISQKNS- TIKISQKNS- TKISQKNS- TVKISQKNS- TVKIPQKNS- TVKIPQKNS- TVKIPQKNS- TVKVPQKNS- T	HNHKLIKRMI	DETSSGGVTGND

BA1	
HO1	
E1	
YC1	
C1	
M1	
R1	
S1	
H1	
RALDH-1	
RALDH1	
RALDH-2	
RALDH2	
B2	
HO2	
YM1	
M2	
R2	
H2	
CRY	
H6	
HX	
YM3	
M3	
R3	
H3	
MM	OFNKGRLGMLLFVCLVAVAAVIV
RM	QFNKGRLQLLLLVCLVAVAAVIVKDQL-
H4	
Н9	
H7	QGCSCTLL
H10	RFNKEKLGLLLLTFLGIVAAVLVKAEYY

<u>Table 5.2a: Secondary Structure Elements of Sheep AIDH 1</u> (N-terminal cofactor-binding domain)

Amino acids	Structural Element	Comments
F23 - H29	β-hairpin	
P38 - P42	β-hairpin	Followed by turn
E46 -E52	β-hairpin	Preceded by turn
K56 - A69	α-helix	
S82 - R97	α-helix	
R99 - N109	α -helix	
F114 - A133	α-helix	
R142 - T143	β -strand	First β -strand in major β -sheet
F150 - E157	β -strand	β -strand in major β -sheet
V161 - I165	β-strand	β -strand in major β -sheet
L172 - S184	α-helix	Kinked (P), but still 1 continuous helix
T 188 - K192	β -strand	β -strand in major β -sheet
L199 - E210	α-helix	1
V218 - V221	β-strand	β -strand in major β -sheet
A228 - S233	α-helix	
K240 - T244	β-strand	β -strand in major β -sheet
T247 - K259	a-helix	
R264 - E268	β-strand	β -strand in major β -sheet

Table 5.2b: Secondary Structure Elements of Sheep AIDH 1 (C-terminal catalytic domain)

Amino Acids	Structural Element	Comments
C275 - V277	β-strand	Followed by 2 3-10 turns as surface loops which reverse around, and then followed by a-helix L283
L283 - V294	α-helix	
F295 - Q300	α -helix	Followed by 3-10 helix
R307 - E311	β -strand	
Y315 - R325	α-helix	
K347 - K362	α-helix	
K366 - L367	β-strand	Small β-strand on a surface protrusion, followed by hairpin which makes no H- bonds.
T384 - S387	β-strand	
R394 - K397	α-helix	
V404 - F410	β-strand	
L413 - N421	a-helix	
S428 - F432	β-strand	
I436 - A445	α-helix	
T450 - V453	β-strand	
E487 - K494	β-strand	

tree, and can not be used to infer evolutionary information; it is simply a pictorial representation of the relationship between primary sequences.

5.3 Results

All amino acid numbering referred to in the text relates to the numbering of sheep AlDH1 (S1). The information contained in the multiple sequence alignment is best understood by studying the dendrogram (Figure 5.2). The overall sequence alignment appears consistent with existing classification based on functional data, with the three different classes clustering together, and the class 1 and 2 enzymes being more similar to each other than to the class 3 enzymes.

Of particular interest to this thesis are the enzymes with retinal-oxidising ability. It can be seen that RalDH1 and RALDH-1 are identical except for residues 99 (arginine in RalDH1, cysteine in RALDH-1) and 169 (glutamate in RalDH1, asparagine in RALDH-1). However, asparagine 169 is absolutely conserved across all other 31 sequences, suggesting that a sequencing error may have occurred in RalDH1. At position 99, all mammalian class 1 and 2 enzymes have an arginine, and this residue is likely to be the first in the third α -helix of the N-terminal domain. It is feasible that cysteine 99 in RALDH-1 is also an error of sequencing (an arginine to cysteine change is brought about by a single cytosine to guanine substitution at the nucleotide level). These two enzymes were both isolated from the rat, RalDH1 found primarily in liver and RALDH-1 in the kidney. However, despite the fact there are minute sequence differences, and likely to be little if any structural differences, different properties have been reported for these enzymes (see Table 1.2), most notably the reported K_m for all-trans retinal. For RalDH1, the authors reported a K_m of 0.76 \pm 0.35 μ M, and an allosteric kinetic mechanism, while for RALDH-1 the authors report a K_m of 8-10 µM and Michaelis-Menten kinetics. From the sequence and structural information, it seems likely that the same protein has been isolated from 2 different rat tissues. However, the different properties reported for these two proteins may lead us to question the methodology used to arrive at those conclusions. The type 1 RALDH enzymes show the closest similarity to the rat and mouse class 1 AIDH enzymes, and the next cluster these enzymes form is with bovine, ovine, equine and human AlDH's, indicating that these 8 enzymes are very similar. In addition the kinetic mechanisms reported for S1, H1, RALDH-2 and RalDH2 all show Michaelis-Menten dependence.



Figure 5.2: Dendrogram Showing Clustering Relationships Used by <u>Pileup to Create Multiple Sequence Alignment</u>

The next two enzymes to cluster with this group of 8 enzymes is the chicken class 1 enzyme (C1) followed by the short-eared elephant shrew crystallin (CRY). The other retinal-specific enzymes - RALDH-2 and RalDH2 join this group next, indicating that they are significantly different from the identified mammalian AlDH 1 enzymes and that they may play a different role. In fact, RALDH-2 does not accept acetaldehyde, citral, propanal and benzaldehyde as substrates, and has been found to be mainly expressed in the testis, with only low levels found in the liver, lung, brain, heart and kidney. The two proteins RALDH-2 (mouse) and RalDH2 (rat) are identical to each other with the exception of residue 353, which in RalDH2 is isoleucine and in RALDH-2 is a valine. This substitution is conserved, and it seems that these two proteins are also identical proteins isolated from different species.

Other information which can be drawn from the sequence alignment concerns newly cloned AlDH genes - AlDH6 (H6), AlDH7 (H7), AlDH9 (H9), AlDH10 (H10), and AlDHX (HX). AlDH8 (H8) has also been cloned (Hsu *et al.*, 1995), however this cDNA encodes only 385 amino acids which was considered to be only a partial transcript and so left out of the alignment. The dendrogram shows that H7 and H10 ('fatty aldehyde dehydrogenase') cluster with the class 3 enzymes, with H10 being similar to the microsomal class 3 enzymes, and H9 an outlier within the class 1 and 2 cluster, but more similar to the class 1 and 2 enzymes than to the class 3 enzymes. HX (human stomach) clusters with the class 2 enzymes, and H6 (human saliva) clusters most closely with the class 1 enzymes than the yeast and *E.coli* class 1 aldehyde dehydrogenases, but more similar to the class 1 and 2 enzymes than to the class 3 enzymes.

The running averages of the similarity among the sequences of a multiple sequence alignment can be plotted using a GCG program called Plotsimilarity (Figure 5.3). This program calculates the average similarity among all members of a group of aligned sequences at each position in the alignment using a user-specified sliding window of comparison. The average similarity across the entire alignment is plotted as a dotted line. From the plot we can see that the regions of highest similarity are between residues 155 - 315, and 395 - 415, and 475 - 495. Although these regions contain many conserved residues important for either structure or function (e.g. Cys 302, Glu 487, Glu 268), it must be noted that there are other regions of high similarity. For example, there are some highly conserved motifs in the N-terminal region of the mammalian class 1 and 2 enzymes (e.g. hydrophobic - Phe - Ile - Asn - Asn - Glu - Trp, from residues 22-28),



Similarity Score

Figure 5.3: PlotSimilarity Output of Aldehyde Dehydrogenase Sequence Alignment

The dotted line represents the average similarity across the entire alignment. The peaks above the dotted line indicate regions of higher than average similarity, and those below the dotted line lower than average similarity

5-125

5-126

but due to the fact that the class 3 and some of the newly identified enzymes do not start until sheep residue 80-100, this region does not score highly in Plotsimilarity. However, this kind of analysis is useful for identifying regions of high sequence conservation which may have functional significance.

5.3.1 Substrate Specificities of Class 1 and Class 2 AIDH

The recent solving of the crystal structures of a class 1 AlDH (S. Moore et al, manuscript in progress), and a class 2 AlDH (Steinmetz et al., 1997), can be used in addition with the above multiple sequence alignment, to explain the differing substrate specificities of these two classes of enzymes. In particular, the reasons for the inability of AlDH 2 enzymes to accept retinal as a substrate, while the AlDH 1 enzymes can oxidise retinal may be addressed. The tertiary structures of the two enzymes - in particular the substrate binding pockets - are very similar. However, there are two amino acid substitutions between the classes which may explain the differences in substrate specificities. Residues which form the binding pocket (sheep AlDH 1 numbering) are: Met 120, Phe 170, Leu 173, Met 174, Trp 177, Tyr 296, Ile 303, Val 459, and Phe 465. All of these residues except two (Ile 303 and Val 459) are highly conserved between class 1 and class 2 enzymes. However, position 459 in the class 1 enzymes (except E.coli, yeast and chicken enzymes) is always a small hydrophobic residue, e.g. valine in the sheep, while in the class 2 enzymes the equivalent amino acid is a phenylalanine. Position 303 in the class 1 enzymes is always a small hydrophobic residue, while in the class 2 enzymes it is generally a cysteine (except the horse class 2 (Ho2), where it is Gly). The introduction of a larger side chain (Phe) and/or a difference in charge in the binding pocket (with the substitution of a Cys for a small hydrophobic residue) may change the ability of the enzyme to make interactions with substrates and hence, may alter the ability of apparently similar enzymes to bind the large hydrophobic substrate retinal.

5.4 Conclusions

The multiple sequence alignment clarifies relationships between a large number of aldehyde dehydrogenases, of particular interest those with the ability to oxidise retinal. The four retinal-specific aldehyde dehydrogenases (RALDHs) recently isolated from rat and mouse tissues (Bhat *et al.*, 1995; Penzes *et al.*, 1997; Wang *et al.*, 1996; Zhao *et al.*, 1996) have been proposed to play a major role in retinal metabolism. However, due to the lack heretofore of comparison between these and other aldehyde dehydrogenases
with the ability to oxidise retinal, it has been difficult to determine the relative role played by each protein in retinoic acid production. The sequence alignment has shown that the four cloned RALDH genes (RalDH1, RALDH-1, RalDH2 and RALDH-2), in all likelihood encode only two distinct proteins. The type 1 RALDHs show high similarity with class 1 aldehyde dehydrogenases, some of which have been demonstrated to oxidise retinal (S1, H1, M1), and would be classified as class 1 enzymes on this basis. The type 2 RALDHs clustered most closely with the class 1 enzymes; however, they show significant sequence differences and may not be true class 1 enzymes. It has also been shown that type 2 RALDH enzymes do not oxidise some aldehyde substrates such as acetaldehyde, propanaldehyde and benzaldehyde (Wang *et al.*, 1996), and hence may be more retinal-specific enzymes than S1, H1, M1 and RALDH type 1 enzymes, and may have a different role.

In addition, the basis for the differing substrate specificities of the class 1 and 2 enzymes with respect to retinal has been proposed, based on the conservation of residues found in the substrate binding pocket. The elucidation of the crystal structure of sheep AlDH 1 (S. Moore, Massey University), and the recent publication of the structure of bovine AlDH 2 (Steinmetz *et al.*, 1997), identified residues important in substrate binding and forming the binding pocket (see section 5.3.1). On the basis of a comparison of the class 1 and 2 structures and multiple sequence alignment, we have proposed that two residues located in the substrate binding pocket are responsible for the ability of only class 1 aldehyde dehydrogenases to bind and oxidise retinal; Ile 303 and Val 459 (usually Cys 303 and Phe 459 in the class 2 enzymes). Further crystallographic modelling studies currently being carried out may clarify this important difference in substrate specificity (S. Moore, manuscript in progress). With the large number of isolated and sequenced mammalian AlDHs, an alignment such as this one is invaluable when looking at sequence and structural motifs, conserved residues, substrate specificities, relationships between enzymes, and in the classification of new AlDHs.

6. Chapter Six: *In Vivo* Studies on Retinal Dehydrogenation

6.1 Introduction and Aims

A human neuroblastoma cell line, SH-SY5Y (Biedler *et al.*, 1973), was used to attempt to study the *in vivo* production of retinoic acid from retinal. This cell line was chosen because it had previously been demonstrated to be responsive to retinoic acid (Pahlman *et al.*, 1984) and because of its availability. The SH-SY5Y cell line is a neuroblast- or N-type cell line derived from SK-N-SH cells. Culturing SH-SY5Y cells in the presence of all-*trans* retinoic acid induces morphological conversion of the cells into a neuronallike phenotype, with the extension of cell processes known as neurites. Because of the responsiveness of this cell line to retinoic acid, it was thought likely that enzymes involved in retinoic acid synthesis would be present. The main aim of this work was to develop a system using SH-SY5Y cells, whereby the effects of ethanol and acetaldehyde on the conversion of retinal to retinoic acid could be examined. This work could have potential importance with regards to the mechanisms by which fetal alcohol syndrome arises (see section 1.5).

6.2 Results

6.2.1 Initial Characterisation of SH-SY5Y Cells

The primary aim of the initial work was to establish whether or not AlDH 1 was expressed in the cell line. SH-SY5Y cells were cultured as described in section 2.2.16. After harvesting the cells, they were lysed and the cytosolic contents separated using SDS-PAGE as described in section 2.2.8. The proteins were electroblotted onto nitrocellulose membrane, and tested for the presence of AlDH 1 using antibodies raised to sheep AlDH 1 as described in section 2.15. Pure sheep AlDH protein was used as a positive control, and rat phaeocytochroma cells (PC12 cells, previously shown not to express AlDH 1 in preliminary experiments) treated in the same manner were used as a negative control. Results showed that AlDH 1 was expressed in this cell line (Figure 6.1).



Figure 6.1: Western Blot of SH-SY5Y Cells

SH-SY-5Y cells were harvested, cracked and run on a 15% SDS gel as described in Materials and Methods. Proteins were electroblotted onto nitrocellulose and probed using anti-rabbit sheep AlDH 1 antibodies. The blot was visualised using the ECL detection kit as described in Materials and Methods

<u>GEL 1</u>	(1/10	000 Ab	dilution)
--------------	-------	--------	-----------

- Lane 1: SH-SY5Y cells
- Lane 2: SH-SY5Y cells
- Lane 3: Pure Recombinant Human AlDH1
- Lane 4: Pure Sheep AlDH 1
- Lane 5: PC12 cells

- $\underline{\text{GEL 2}}$ (1/5000 Ab dilution)
- Lane 1: SH-SY5Y cells (+ retinoic acid 1µM)
- Lane 2: Pure Sheep AlDH 1
- Lane 3: Pure Recombinant
- Human AlDH1
- Lane 4: SH-SY5Y cells
- Lane 5: SH-SY5Y cells

6.2.2 Effects of Retinal and Retinoic Acid on Cell Morphology

6.2.2.1 Introduction and Aims

The next step after showing that AlDH 1 was present in SH-SY5Y cells, was to attempt to study the conversion of retinal to retinoic acid in the cells. The effect of retinoic acid on many cell types including SH-SY5Y is well documented (Haussler *et al.*, 1983 and refs. therein). Additions of nano- or micro-molar quantities of retinoic acid induces the cells to differentiate into a neuronal phenotype. Upon acquiring these characteristics, the cells extend elongated processes which typically give the cell a bipolar appearance, whereas undifferentiated cells have numerous shorter appendages. A typical method by which to determine if a cell has differentiated or not is to measure the length of the major process(s), and compare this to the length of the cell body. If the process was $1\frac{1}{2}$ -2 times the length of the cell body, then the cell was defined as differentiated.

In order to look at the production of retinoic acid from retinal, it was decided to see if any morphological changes similar to those seen upon the addition of retinoic acid were seen upon addition of retinal only. Similar changes would indicate that conversion of the added retinal to retinoic acid was occurring. There is no documented evidence that retinal has any effect on cell morphology on its own, so any changes that are seen are likely to be due to the production of retinoic acid.

6.2.2.2 Methods

Confluent cells were passed and divided equally onto four 10 cm diameter tissue culture plates. These plates contained 5 sterile glass microscope cover slips. The cells were then allowed to attach and grow as usual. After 48 hours, the media was changed and the effectors were added (Figure 6.2). Initially all-*trans* retinal and all-*trans* retinoic acid were tested, with two controls. One control was cells only with no additions, and the other consisted of cells to which just the solvent used to dissolve the retinoids was added. The solvent used was dimethylformamide. This was chosen over the typical solvent used in these studies, ethanol, because of the subsequent work we wished to pursue. Ultimately, the effect of ethanol and acetaldehyde on the production of retinoic acid was going to be studied, so it was decided not to use ethanol in order to simplify later experiments. The total amount of solvent added to the culture medium did not exceed 0.5% (v/v).

Just prior to the addition of effectors, one cover slip was removed from each plate, as a zero time point. The plates were then incubated as stated in section 2.2.16. One cover slip was removed every 24 hours for three subsequent time points for analysis. At 48 hours post-treatment, media was removed, the plate was gently washed with 5 ml media, and replaced with fresh media. Effectors were added as described earlier.

Cells were viewed using a Zeiss Axioskop microscope. To study the overall morphology of the cell, they were viewed by high magnification differential interference (DIC or Nomarsky) microscopy. The coverslip with cells attached was placed upside gown on a sterile microscope slide on a drop of glycerol. To count the nuclei, cells were stained with Hoechst 33342 dye as described (section 2.2.16.3) and viewed by fluorescence microscopy. The number of cells in a particular view was determined by counting fluorescently stained nuclei. Differentiation was assessed by cell morphology, and the percentage of differentiated cells was calculated.



Figure 6.2: Outline of Treatment of SH-SY5Y Cells With Retinoids

6.2.2.3 Results and Discussion

Results of adding effectors are shown in Table 6.1 and Figures 6.3-6.7. Differences between control and test plates were able to be seen by viewing changes in cell morphology by microscopy as described. Cells which had been treated with retinal or retinoic acid showed characteristics typical of differentiated cells. A large proportion (up to 1/3) of these cells were typically bipolar, with processes longer than 1½-2 times the length of the cell body (Figures 6.5 and 6.6). Compared to these cells, cells in the control plates with either no additions or the addition of vehicle only, were clumpier, rounder cells with numerous short processes (Figures 6.3 and 6.4). These observations were quantified by counting a random selection of cells from each plate at each time point, and calculating the number of differentiated cells. To ensure unbiased selection of fields, the slide was moved into random positions before counting. A large sample size of 200-300 cells were counted and assessed for differentiation characteristics in an effort of represent accurately the overall situation of the cells in each plate. Results are summarised in Table 6.1.

From this data it can be seen that the addition of all-*trans* retinal and all-*trans* retinoic acid at a final concentration of 1 μ M increases the proportion of differentiated cells compared to both control plates where no retinoid has been added. The proportion of differentiated cells is higher for retinoic acid treated plates compared to retinal. As previously stated, there is no documented evidence that retinal itself has any effect on cell morphology. We hypothesise that any changes seen upon addition of retinal are due to its conversion in to a compound such as retinoic acid that has a well defined effect on SH-SY5Y cells. We also know that at least one enzyme capable of oxidising retinal to retinoic acid is present in this cell line - AlDH 1. Alternatively, retinal may also be converted into retinol and stored, which would mitigate the retinoic acid signal. A time delay effect is also possible where the time taken for the conversion of retinal to retinoic acid delays the manifestation of the morphological changes seen, though the time periods used here may be too large for this to be a factor. These results suggest that retinal is being converted to retinoic acid in this cell line. If so, SH-SY5Y cells are suitable for studying the effects of ethanol and acetaldehyde on this reaction.

Preliminary analysis of the effects of ethanol, acetaldehyde and disulfiram on the production of retinoic acid in SH-SY5Y cells was also examined. These effectors were added to the plates using a similar experimental paradigm (Figure 6.9).



Figure 6.3: SH-SY5Y Cells from Control Plate

Representative views of typical cells using Nomarski microscopy. Magnification is 400X (photo A) and 630X (photo B) using an oil immersion lens. Note the clumped nature of the cells, with multiple processes.



Figure 6.4: SH-SY5Y Cells from Control + DMF Plate

Representative views of typical cells using Nomarski microscopy. Magnification is 400X (photo A) and 630X (photo B) using an oil immersion lens. Again note the clumped nature of the cells, with multiple processes. 6-135



Figure 6.5: SH-SY5Y Cells from All-trans Retinal Treated Plate

Representative views of typical cells using Nomarski microscopy. Magnification is 400X (photo A) and 630X (photo B) using an oil immersion lens. Note the bipolar nature of the cells, with few extended processes. The cells are less clumped compared to control plates.



Figure 6.6: SH-SY5Y Cells from All-trans Retinoic Acid Treated Plate

Representative views of typical cells using Nomarski microscopy. Magnification is 400X (photo A) and 630X (photo B) using an oil immersion lens. Note the bipolar nature of the cells, withfew extended processes. The cells are less clumped compared to control plates.



Figure 6.7: Fluorescence Microscopy of SH-SY5Y Cells

Representative views of cell nuclei stained with Hoecsht 33342 as described (section 2.2.16.2), and visualised by fluorescence microscopy. Magnification is 400X (photo A) using an oil immersion lens and 200X (photo B).

Plate	Time Point	Total Cells Counted	Number Differentiated	% Differentiated
1 (Control - no				
additives)	0 hours	243	0	0
1	24 hours	311	1	0.3 %
1	48 hours	312	7	2.2 %
2 (Control - 0.5 %				
(v/v) DMF added)	0 hours	278	2	0.7 %
2	24 hours	296	8	2.7 %
2	48 hours	271	12	4.4 %
3 (1 µM all- <i>trans</i>				
retinal)	0 hours	299	3	1.0 %
3	24 hours	258	56	21.7 %
3	48 hours	278	49	17.6 %
4 (1 μM all- <i>trans</i>				
retinoic acid)	0 hours	250	2	0.8 %
4	24 hours	301	100	33.2 %
4	48 hours	309	99	32 %

Table 6.1: Quantification of SH-SY5Y Cell Morphology

Results are from a single experiment, representative of 5 individual experiments. The averages of the 5 experiments, along with the standard deviation are seen in Figure 6.8.

Disulfiram is an immediate effective inhibitor of cytosolic (class 1) aldehyde dehydrogenase. If AlDH 1 is the primary enzyme involved in retinal metabolism as we have suggested (section 1.2.2, Chapters 3 and 4), then it is likely that addition of disulfiram *in vivo* would inhibit production of retinoic acid, and the morphological changes would not be seen. Results of these additions are shown in Table 6.2.

These preliminary experiments suggest that retinal is being oxidised to retinoic acid by SH-SY5Y cells. Results also suggest that retinol is being converted to retinoic acid. This is perhaps expected - it would be unlikely for components of the second half of a metabolic pathway to be present without those from the first half. However, the reaction of interest is the reaction catalysed by aldehyde dehydrogenase. It can also be seen that



Figure 6.8: Quantification of SH-SY5Y Cell Morphology

Note: Results shown in Figure 6.8 are the average of 5 experiments as detailed in Table 6.1. Abbreviation: DMF = dimethylformamide.



Figure 6.9: Outline of Treatment of SH-SY5Y Cells With Retinoids and Additional Effectors

Cells were plated onto 10 cm (diameter) plates containing 10 ml RPM1640 media and sterile glass coverslips, and left to adhere and grow for 24 hours. The zero time point was taken after 24 hours growth, before the addition of effectors. Effectors were added as described.

Plate	Time Point	Total Cells Counted	Number Differentiated	% Differentiated
$1 (1 \dots M all turns)$				
$1 (1 \mu N an-irans)$	0 hours	200	3	10%
	2.4 h aure	255	5	1.0 70
1	24 hours	238	36	21.7%
	48 hours	278	49	17.6%
$2(1 \mu\text{M all}\text{-trans})$		2 0 7		o - o (
retinol)	0 hours	287	2	0.7 %
2	24 hours	247	51	20.7 %
2	48 hours	260	70	26.9 %
3 (1 μM 9-cis				
retinal)	0 hours	277	3	1.1 %
3	24 hours	312	88	28.21 %
3	48 hours	225	40	17.8 %
4 (1 µM all-trans				
retinal $+$ 50 μ M				
acetaldehyde	0 hours	250	0	0 %
4*	24 hours	0	0	0
4*	48 hours	0	0	0
5 (1 µM all-trans				
retinal $+$ 50 μ M				
ethanol)	0 hours	282	3	1.1 %
5	24 hours	239	62	25.9%
5	48 hours	300	93	31 %
6(1 μM all- <i>trans</i>				
retinal $+10 \mu M$				
Disulfiram)	0 hours	302	2	0.7 %
6	24 hours	253	6	2.4%
6	48 hours	260	31	119%

Table 6.2: Quantification of SH-SY5Y Cell Morphology with Added Effectors

*The addition of acetaldehyde rendered the cells no longer viable.

Results are from a single experiment, representative of 2 independent experiments. The averages of the 2 experiments, along with the range of values are seen in Figure 6.10.

9-cis retinal is probably being converted to 9-cis retinoic acid, which also has been shown to have an effect on the differentiation of these cells (Han *et al.*, 1995; Kizaki *et al.*, 1993; Lovat *et al.*, 1994; Redfern *et al.*, 1995). This is also not surprising. The work in this chapter has shown that AlDH 1 is expressed in this cell line, and it was



Note: Results shown in Figure 6.10 are the average of 2 experiments as detailed in Tables 6.1 and 6.2. Abbreviations: AT = all-*trans*, 9C = 9-*cis*, ETOH = ethanol, DSFM = disulfiram, DMF = dimethylformamide.

previously shown that AlDH 1 is capable of converting 9-*cis* retinal to 9-*cis* retinoic acid (Chapter 3).

The addition of ethanol, acetaldehyde and disulfiram to plates containing all-trans retinal showed interesting results (Table 6.2). The addition of ethanol increased the percentage of differentiated cells slightly at 24 hours and by nearly 1.5-fold at 48 hours. The addition of acetaldehyde proved to be detrimental to the cells, and no quantification The addition of disulfiram greatly decreased the percentage of could be done. differentiated cells at 24 hours and 48 hours, although the difference between all-trans retinal treated cells and retinal + disulfiram treated cells was less at 48 hours. These very preliminary experiments seem to indicate that retinal is being oxidised to retinoic acid by AIDH 1 in SH-SY5Y cells. The addition of ethanol, acetaldehyde and disulfiram would provide preliminary information concerning the overall aim of this chapter, i.e., the effect that these compounds have on retinoic acid production. One theory as to how fetal alcohol syndrome may arise (see section 1.5), is a competitive inhibition of alcohol dehydrogenase catalysed retinal production by ethanol (Deltour et al., 1996; Duester, 1991; Duester, 1994). Ethanol is metabolised via a pathway analagous to retinol metabolism. Though it is an unresolved issue as to which enzymes capable of retinol oxidation *in vitro* actually catalyse the reaction *in vivo*, it is possible that a competitive inhibition may arise in some circumstances especially situations which give rise to fetal alcohol syndrome. For example in alcohol binge drinking, the concentration of blood alcohol may greatly exceed that which the primary alcohol oxidising enzymes can cope with. In such a case, enzymes which are normally involved in retinoid metabolism may be forced to oxidise the large excess of circulating ethanol. At certain times in fetal development when the controlled production of precise amounts of retinoic acid at specific loci is critical for correct fetal formation, if the enzymes are involved in oxidising large amounts of ethanol, the required retinoic acid may not be available.

It is also conceivable that the same scenario may exist for the aldehyde dehydrogenases. If large amounts of alcohol are ingested, high acetaldehyde levels may mean the recruitment of enzymes usually only involved in retinal oxidation to metabolise excess toxic acetaldehyde. The primary physiological substrate for human AIDH 1 has been proposed to be retinal. If acetaldehyde concentrations were significantly raised this enzyme may be required to oxidise the toxic compound acetaldehyde at the expense of retinal. However, the data shown in Table 6.1 and Table 6.2 shows that the addition of ethanol actually potentiates the effect that retinal has on cell morphology. If the

hypothesis presented by Deuster and colleagues is correct, then we would expect the addition of ethanol to decrease the production of retinoic acid and therefore, decrease the morphological changes seen. The most obvious reason why a decrease in differentiation may not be seen, is the time frame of the experiment. The addition of ethanol may cause a rapid increase in acetaldehyde concentration and the recruitment of retinal-oxidising enzymes for acetaldehyde oxidation initially; however, when the ethanol and acetaldehyde have been metabolised the enzymes resume retinoid oxidation and the same morphological changes are seen. One characteristic of FAS is the presence of incorrect amounts of retinoic acid at precise times during fetal development. The manner in which these experiments are being carried out does not allow for the monitoring of the production of retinoic acid over small periods of time in the way that the direct measurement of retinoic acid using HPLC would. So a decrease in retinoic acid production may be occurring initially, but due to the requirement of a long time frame for the experiment to allow morphological changes to take place, this may not be able to be observed.

The reasons as to why an increase in the percentage of differentiated cells was seen in the presence of ethanol is not easy to account for. One possible scenario is that a feedforward mechanism exists, such that in times of excess acetaldehyde, the production of AlDH protein is upregulated. The presence of more AlDH protein may lead to a relatively larger proportion of available retinal being converted into retinoic acid, as opposed to being converted to retinol for storage. This would lead to a larger percentage of cells differentiating as was observed. This also could account for FAS, as an excess of retinoic acid over that which is required during development is teratogenic (see Figure 6.11).

The addition of disulfiram inhibits cell differentiation after 24 hours by approximately 20 %, but after 48 hours differentiation is decreased by only 8 %. It is probable that the initial inhibition by disulfiram of AlDH 1 seen during the first 24 hours by the decrease in the percentage of differentiated cells, is temporary, which is shown by the subsequent increase in the differentiation of the cells.

To generate data which supports the hypothesis that AlDH 1 is oxidising retinal to retinoic acid and to study this reaction *in vivo* is very difficult. The first step would be to prove that AlDH 1 is of primary importance, or even participates in the *in vivo*



Figure 6.11: Two Possible Mechanisms for the Manifestation of Fetal Alcohol Syndrome

A: Binge drinking increases the concentration of ethanol and acetaldehyde in the body. Competitive inhibition by ethanol and acetaldehyde of retinal and retinoic acid production reduces the amount of retinoic acid available to target cells.

B: The increase of ethanol and acetaldehyde induces an increase in the amount of ADH and AlDH protein by a feed-forward mechanism. This leads to an increase in retinoic acid production which alters downstream effects in target cells.

Abbreviations: ADH - Alcohol dehydrogenase, AlDH - Aldehyde dehydrogenase, → - positive effect or increase → - negative effect or decrease production of retinoic acid. Western blotting experiments showed that AlDH 1 is expressed in the cell line, and morphological evidence indicates that it plays some role in retinoic acid production due to the fact that disulfiram reduced the percentage of differentiated cells. This evidence supports the above hypothesis, but is not sufficient to prove it. Disulfiram immediately inhibits AlDH 1 dehydrogenase activity completely, but also inhibits AlDH 2 (mitochondrial) activity after a period of time. A number of methods could be employed in an effort to prove the hypothesis. A more direct assay of retinoic acid production would be necessary, such as the HPLC based assay as described in Chapter 3. This assay has been used to measure *in vivo* retinoic acid production by a number of groups (see section 1.2.2) and would be vital for this work. As stated previously this method proved to be technically difficult. Had the assay been successful, further work to identify retinal-oxidising enzymes present in these cells could have been pursued.

The identification of enzymes present in the cell lines may have been determined by either performing further Western Blots if antibodies were available, or perhaps Northern analysis with appropriate oligonucleotide probes. The elucidation of which of the enzymes present are important in retinal oxidation might also be pursued by microinjecting antisense AlDH 1 RNA or oligonucleotides to reduce the amount of AlDH 1 protein produced by the cells by 80-98 % (Izant & Weintraub, 1984; Kim & Wold, 1985). It would then be possible to see how important AlDH 1 was in retinoic acid production. If AlDH 1 played a very important role, the injection of antisense AIDH 1 RNA would be expected to reduce the production of retinoic acid. It is quite possible that this type of experiment may yield no result even if AIDH 1 is the major retinoic acid synthesising enzyme in vivo. There may be other enzymes present in the cell line capable of oxidising retinal which, although they would not usually play such a large role, may act as a 'backup' mechanism to keep retinoic acid production constant. So even if AlDH 1 was the primary enzyme involved in retinal metabolism in vivo, results seen by performing an experiment such as the one described may not necessarily reflect this. It would be important to first establish which of the identified enzymes able to oxidise retinal (aldehyde oxidase, cytochrome P450 1A1 and 1A2, and xanthine oxidase) were present in these cell types.

Analysis of the effects of addition of acetaldehyde to plates treated with all-*trans* retinal provides an additional challenge. Acetaldehyde is likely to be metabolised by the cells. To see an effect of acetaldehyde on AlDH 1-catalysed retinoic acid production,

acetaldehyde levels must be maintained over the experimental time. In the case of examining morphological characteristics in preliminary experiments, this would mean maintaining a constant acetaldehyde concentration over 48-72 hours. In addition, the concentration of acetaldehyde must not exceed a reasonable physiological concentration. This would not be difficult when using direct measurement of retinoic acid production by HPLC based extraction and analysis over short time periods of up to 30 minutes.

6.3 Conclusion

Preliminary studies on in vivo retinal dehydrogenation were carried out. All-trans retinal and all-trans retinoic acid were shown to induce the differentiation of SH-SY5Y human neuroblastoma cells by studying morphological characteristics. Western analysis identified the presence of AlDH 1 in SH-SY5Y cells. Since retinal has never been shown to affect the differentiation of cells, it was considered likely that retinal was being converted into retinoic acid, which was subsequently inducing the differentiation seen. 9-cis retinal and all-trans retinol were also shown to induce differentiation indicating 1) the presence of a retinol oxidising enzyme(s), and 2) the presence of enzymes able to oxidise retinal isomers. The addition of ethanol to SH-SY5Y cells increased the percentage of differentiated cells by 4-12 %. The addition of disulfiram decreased the percentage of differentiated cells by 6-20 % compared to the percentage when only alltrans retinal was present. This indicates that AlDH 1 may play an important role in retinal dehydrogenation as disulfiram is an immediate potent inhibitor of AlDH 1. In addition AIDH 1 has been shown to oxidise the retinal isomer 9-cis retinal.

Despite the promising findings of these preliminary experiments, no firm conclusions can be drawn from the available data. However, the hypothesis that AlDH 1 plays a vital role in retinal oxidation is supported by the data collected.

7. Chapter Seven: Purification of a Novel Protein with Alcohol Dehydrogenase Activity

7.1 Introduction and Aims

During an isoelectric focusing zymography assay of a crude sheep liver homogenate, a band of strong dehydrogenase activity at pI ~8.3-8.5 was observed using both acetaldehyde and retinal as substrates. This band was originally thought to be an aldehyde dehydrogenase isozyme, in view of the facts that the only other observed bands of activity on the same gel were attributed to AlDH 1 and 2, and that the first retinal specific dehydrogenases with pI ~8.5 had recently been isolated from rat tissues (Labrecque et al., 1993; Posch et al., 1992). The initial aim of this work was to purify and study the novel aldehyde dehydrogenase with retinal oxidising ability. However, it was subsequently shown that the observed activity was due to trace amounts of ethanol and methanol present, and that the protein was actually an alcohol dehydrogenase. The project was pursued, due to the characteristics of the protein being novel for the alcohol dehydrogenase family. Sequence data obtained did not match any sequence data contained in the SwissProt protein databank, nor did the physical properties of the protein match any previously identified proteins. The novel putative alcohol dehydrogenase has therefore been purified and partially characterised.

7.2 Methods and Results

The initial isoelectric focusing gel showed five main bands of activity when stained using acetaldehyde as a substrate, and one band when retinal was substrate (Figure 7.1a, b). The 4 bands at the anodic end of the gel, are routinely seen when staining a fairly crude sample for aldehyde dehydrogenase activity. They correspond to AlDH 1 (one band), and AlDH 2 (multiple bands) as seen in Figure 7.1a, b, c. The band of activity toward the cathode appeared to be an isozyme of aldehyde dehydrogenase with a basic isoelectric point. Although all isozymes stained positively using acetaldehyde, the strongest signal observed when using retinal was produced by the basic pI dehydrogenase. From these data, it was concluded that the new AlDH used retinal more efficiently than AlDH 1, and could be a retinal-specific AlDH analogous to those recently isolated from rodent tissue. However, since it was known that AlDH 1 did use retinal as a substrate, but we could not see a convincing band, it was proposed that the





Figure 7.1: IEF Gel of Crude Liver Extract

A: IEF gel stained for activity B: IEF gel blotted onto nitrocellulose and stained for activity



Figure 7.1c: IEF Gel of Crude Liver Extract

IEF gel stained for activity as described in Materials and Methods. Clearly seen using both acetaldehyde and retinal as substrates are bands at the positions expected for AlDH1 and AlDH2, in addition to the band of new activity. hydrophobic retinal might not penetrate the gel very efficiently, thereby reducing the signal, or making it difficult to obtain a signal in some cases. To see if this was the case, proteins separated by isoelectric focusing were transferred by capillary action onto a nitrocellulose membrane, and the membrane was stained for activity in the same way as described for gels. The transferral of proteins from gel to membrane was successful, but no differences were seen between gel and membrane after staining (Figure 7.1a, b). However it can be seen in the blot (Figure 7.1b) that some diffusion of protein occurred, giving less discrete banding patterns.

The activity stain was repeated on the same crude homogenate, using a number of different aldehydes as substrates (Figure 7.2). Retinal, citral, p-carboxybenzaldehyde, and 3,4-diiodosalicylaldehyde, as well as retinal without NAD^+ , were used. From this were observed gel, positive results at а basic pI (~8.3-8.5) using 3,4-diiodosalicylaldehyde, retinal, and *p*-carboxybenzaldehyde, and negative results with citral, and retinal without NAD^+ . Positive results for AlDH 1 (toward anode) were observed with *p*-carboxybenzaldehyde and 3,4-diiodosalicylaldehyde only. These results also served as controls. The first control was that done with no NAD^+ , which was negative, and the second was the negative result obtained using citral. Thus the positive results obtained seemed to be NAD⁺-dependent and substrate-dependent.

From the above observations, and the recent isolation of rat retinal specific dehydrogenases, it was hypothesised that the band of activity with a pI of 8.3-8.5 was being produced by a previously unidentified retinal-preferring aldehyde dehydrogenase from sheep liver. The aim of further work at this stage was therefore to purify this protein using an approach based around the well established purification protocol for AlDH 1, to characterise the protein, to study its activity, and to compare it to the other aldehyde dehydrogenase capable of oxidising retinal, AlDH 1. However, later testing using the activity stain method revealed that the protein appeared to be using trace amounts of ethanol or methanol present in some preparations of NAD⁺ as a substrate instead of the aldehyde; i.e., appeared to be an alcohol dehydrogenase rather than an aldehyde dehydrogenase. The NAD⁺ used (ICN Pharmaceuticals (Cat. No. 100499) and Sigma (Grade AA-1)) contains ethanol and methanol as organic solvents (see ICN Pharmaceuticals Catalogue (Cat. No. 100499 and 100591) and Sigma Catalogue). This supported by later experiments (section 7.2.11) where a proposal was spectrophotometric assay using ethanol as a substrate showed a positive activity trace, while assays using aldehydes as substrates showed no activity at any time. This evidence



Figure 7.2 IEF Gel of Crude Liver Extract Using Different Substrates

IEF gel was run and stained for activity using the above substrates as described in Materials and Methods

seemed to conflict with the negative results which had been obtained using some aldehyde substrates, since NAD⁺ was always present in these assays (see Figure 7.2). Therefore, the proposal that the enzyme used NAD⁺ and alcohol present in the NAD⁺ preparation as substrates, would mean that all lanes should give positive results. In addition, the fact that this protein was the only 'alcohol dehydrogenase' identified in a crude liver extract was inexplicable, as there are a number of major alcohol dehydrogenase isozymes in liver which would have been expected to stain positively. Although the evidence obtained was conflicting, the stronger evidence pointed to the enzyme being an alcohol rather than an aldehyde dehydrogenase (e.g. the spectrophotometric assay, and subsequent positive IEF activity stains being obtained with no aldehyde substrate). The aim of this section of work therefore became to purify, identify and characterise a putative novel alcohol dehydrogenase.

7.2.1 Initial Protein Purification

The purification of this protein was followed using the described activity stain of proteins separated by isoelectric focusing gels (see Section 2.2.7). The advantages of using this method are its sensitivity, and the ability to distinguish between isozymes or enzymes which have the same activity, but a different isoelectric point.

The protein was shown to be localised to the same subcellular fraction as AlDH 1 in the cytosol, and was precipitated between PEG 8000 concentrations of 10 and 20 % (w/v). Therefore the initial steps taken to purify the new enzyme were identical to those used in the purification of AlDH 1 (Figure 7.3).

7.2.2 Ion-exchange Chromatography

The first approach taken to further purify this protein from the 10-20 % (w/v) PEG precipitate involved the use of ion-exchange chromatography. Since the isoelectric point was known, a method involving the interaction between a charged group attached to an inert matrix and the overall charge on the protein may provide a means to further purification. The resins used were carboxymethyl substituted Sephadex (CM-Sephadex), and diethylaminoethyl-Iontosorb (DEAE-Iontosorb) (Figure 7.4). All manipulations involving column chromatography were performed at 4°C, and all buffers contained 0.3 mM EDTA and 0.5 mM DTT, unless otherwise specified. The pellet obtained after precipitation of proteins with 20 % (w/v) PEG 8000, was then redissolved in a small volume of 10 mM Bis-Tris pH 6.2. This solution was applied to a DEAE-Iontosorb column (dimensions: 4 cm diameter \times 30 cm length) pre-equilibrated in the same buffer,





Discard supernatant Pellet containing protein

Figure 7.3 Initial Purification of Novel Protein

7-155



DEAE-Iontosorb

Figure 7.4: Ion-exchange Resins

and the column was washed until the red hemoglobin-containing component was eluted. When this fraction was tested for activity on an IEF gel, staining for activity with retinal and NAD⁺, the band of basic pI activity was found to be associated with the red-coloured component. At pH 6.2, proteins with a high pI would be positively charged, and not interact with the positively charged column, thus washing through as observed here. This step though, is a 'negative' purification step which does not involve specific targeting of the desired protein, e.g. by binding it to an inert matrix, and therefore does not result in such a good increase in purity as a positive step might.

Using the same column, a positive strategy using different pH buffers was employed. The 20 % PEG pellet was redissolved in a small amount of 20 mM tetrasodium pyrophosphate pH 9 with 5 mM NaCl, and tested for activity. The active fraction was applied to the DEAE-Iontosorb column which had been pre-equilibrated in the same buffer. The column was washed through with 2-3 l of buffer and fractions collected immediately to see if the protein was binding. The buffer was changed to 20 mM potassium phosphate buffer pH 7.4 with 5 mM NaCl, and fractions collected. Finally, the buffer was changed to 20 mM Bis-Tris pH 6.2 with 5 mM NaCl and fractions collected. Theoretically, the protein should bind to the column under initial conditions, and be eluted with the first buffer change along with other cytosolic proteins which precipitate between 10 and 20 % (w/v) PEG 8000, and which have an isoelectric point between 7.4 and 9. Remaining proteins would be expected to either remain bound to the column, or be eluted with the third buffer change if the isoelectric points are between 6.2 and 7.4. Fractions collected from each set of conditions were tested for activity using an IEF gel with activity stain. No activity was found in any elution fractions, indicating either that the protein was remaining bound to the column under all conditions used, or that the protein was losing its activity at some stage.

The possibility that the protein may require a metal ion for activity led us to repeat the above experiment without including the metal ion chelator EDTA in the buffers. However, the absence of EDTA did not alter the results observed.

A cation exchange column was used in an effort to bind the protein in a positive purification strategy. The column was CM-Sephadex, pre-equilibrated in 20 mM Bis-Tris pH 6.8. The 20 % PEG pellet was redissolved in the equilibration buffer and applied to the column. Five ml fractions were collected as the sample was loaded to check that the protein had bound, and the non-binding proteins were washed through with 5-10 column volumes of the same buffer. To elute the bound protein, a number of methods were employed. Firstly, a pH gradient of pH 6.8 to pH 9 was run. Using a manual gradient mixer, a gradient was set up to run from the equilibration buffer (pH 6.8) to 20 mM tetrasodium pyrophosphate pH 9.0. Again 5 ml fractions were collected while loading, they were assayed for activity using activity staining of proteins separated by isoelectric focusing. The gels showed no activity in the fractions eluted initially, indicating that the protein was either binding to the column, or was not bound but was no longer active. Assays of the eluted fractions were also all negative. In this situation the most likely scenario was that the protein had bound to the column, but had

lost its activity by the time it was eluted. This experiment was repeated a number of times, but no activity was ever found in any fractions after the loading of active protein.

The second method of elution used was to alter the ionic strength of the buffer. This method works by charged ions in the buffer competing with the charged protein for a limited number of oppositely charged groups attached to the matrix, and charged ions in the buffer binding to the oppositely charged residues on the protein thus occupying sites which could otherwise be attracted to the column. The protein was applied to the column as described previously. A gradient was then set up using a manual gradient mixer to run from 0.005 M NaCl to 0.8 M NaCl in the same buffer. The NaCl concentration was then held at 0.8 M for 10-20 fractions. Fractions were collected and assayed for activity using the described IEF gel activity stain (2.2.7). Again, using this method no activity was seen in fractions. The most likely conclusion again was that the protein had bound to the column, but was becoming inactivated during the washing and/or eluting processes.

A third method of elution tested was that of affinity elution. Affinity elution of proteins from ion-exchange columns uses a specific ligand which binds to the protein and may change its properties by a conformational change or a change in charge distribution such that the protein no longer binds to the column. This technique may also be used in conjunction with a small change in the ionic strength in the buffer. This combination may result in a more specific and effective elution of the protein (Scopes, 1987a). The ligand chosen was NADH. The active eluant from the negative DEAE-Iontosorb chromatography step was dialysed against three changes of potassium phosphate buffer pH 7.2 with 15 mM NaCl, and applied to a pre-equilibrated CM-Sephadex column. The column was washed extensively with the same buffer, and fractions were collected. The same buffer with 200 μ M NADH was then applied to the column, and fractions were collected. All fractions collected were tested for activity as previously described. No activity was seen in any fractions collected, though the solution applied to the CM-Sephadex was still active. Combining affinity elution with a change in pH was then tried. The procedure outline above was followed. After loading the column and extensive washing, the buffer was changed to 20 mM sodium pyrophosphate pH 8.6 containing 15 mM NaCl and 200 µM NADH. Fractions were collected as stated, and tested for activity. No activity was found.

Two other CM columns were used to ascertain whether the composition of the inert matrix had any effect on the purification process. These resins had the same charged group, but instead of a Sephadex matrix, a cellulose and a Sepharose matrix were used as the base to which the charged groups were attached. The Sephadex matrix which was used initially was a compressible resin. When using high salt concentrations such as those utilised in some elution techniques, this resin would compress to up to 1/3 of its original volume. The other resins did not exhibit this characteristic. However, the composition of the matrix did not appear to have any influence on this purification procedure, and no active protein could be eluted from any of these columns.

To check whether the protein was simply unstable after a certain length of time, one sample of redissolved 20 % (w/v) PEG 8000 precipitate was kept at 4°C and one at -20° C for the same length of time, and the activity tested as described. Both samples maintained their activity for at least 10 days. It was also possible that the dilution of the protein-containing sample which was occurring during the DEAE-Iontosorb step was decreasing the activity of the protein. To address this possibility, the diluted active eluant was concentrated 5-fold under pressure, in an Amicon stirred cell apparatus attached to a N₂ gas cylinder at 4°C, using an XM-100 membrane. The protein was active after this procedure.

7.2.3 Dye-Ligand Chromatography

The next method of purification used was dye-ligand chromatography. The principle of dye-ligand chromatography involves the similarity of the structures of certain coloured dyes to the structure of the nicotinamide ring contained in NAD(P)⁺ and NAD(P)H. Proteins which bind these compounds are good candidates for binding to coloured dyes which, when attached to a matrix, form columns useful for the purification of certain proteins. The dyes used here include Procion Yellow H-E4R (Scopes group 5), Procion Green H-4G (Scopes group 5), and Procion Blue R (Scopes group 3). The Scopes groups mentioned refer to the ability of the dye to bind protein from crude tissue extracts, group 5 dyes binding the most protein and group 1 the least (Scopes, 1987b). All dyes used in these experiments were bound to Sepharose resin. Columns of ~8 ml volume (1.5 cm diameter, 4.5 cm height) were packed and prepared by running 6 M urea in 0.1 M NaOH through, and then equilibrating in 20 mM Bis-Tris pH 6.2 containing 5 mM NaCl. All dye-ligand chromatography was performed at 4°C, and all buffers contained 0.3 mM EDTA and 0.5 mM DTT unless otherwise stated. The elution buffer for compounds binding to dye-ligand columns contained the nicotinamide-containing

compound NADH, to compete with the dye for binding of the adsorbed protein (affinity elution).

The procedure followed for dye-ligand chromatography was as follows: 2-3 ml of the solution which was active after the negative DEAE-Iontosorb step was loaded onto each dye column and washed extensively with the buffer in which the column was equilibrated. Five ml fractions were collected. The buffer was then changed to 20 mM Bis-Tris pH 6.2 with 5 mM NaCl and 300 µM NADH. The eluant was collected and all fractions collected were tested for activity using the IEF gel separated protein activity assay as described (2.2.7), and for protein content on an IEF gel with Coomassie blue staining. No activity or protein was seen in wash or elution fractions from any of the three dye columns used. To test whether enough protein was present to visualise, the procedure was repeated loading the columns to capacity. The loading procedure involved collecting fractions when loading, and testing the absorbance of these fractions at 280 nm (A₂₈₀). When the protein content of the loading eluant increased indicating that the column was loaded to capacity, the loading was stopped, and washing was started and continued until the A_{280} dropped to <0.05 absorbance units. Once the A_{280} dropped to this level, elution was initiated as described previously. While protein was seen in all fractions and the activity of interest was seen in wash fractions, no activity was seen in elution fractions. From the above experiments it was concluded that the amount of protein had been an initial problem in visualisation. However, loading the column to capacity with protein did not result in the elution of active protein. Again the protein was either still bound to the column, or becoming inactivated by the binding and/or elution processes.

This procedure was repeated a number of times with different initial samples. None were successful.

7.2.4 Affinity Chromatography

The original aim of this section of work involved the purification of an aldehyde dehydrogenase, before the realisation that the protein might, in fact, be an alcohol dehydrogenase. Therefore after the initial problems in purification, it was decided to try to use the affinity column used in the purification of AlDH 1 as a specific step for the purification of this protein. The affinity column (prepared by G Freeman and T Kitson, Massey University), is composed of *p*-hydroxyacetophenone molecules attached to a matrix (Ghenbot & Weiner, 1992). *p*-Hydroxyacetophenone is a competitive inhibitor of

the esterase activity of AlDH, and this resin has proven to be an excellent specific purification tool for the purification of sheep liver AlDH 1 (see section 2.2.6.1). Purification attempts using the affinity resin were carried out as follows: The affinity column was pre-equilibrated in 50 mM potassium phosphate buffer pH 7.4 with 15 mM NaCl. Active eluant from the DEAE column step (see section 7.2.2) was dialysed against 3×1 l changes of affinity column, and washed through. 10 ml fractions were collected from the time of sample application. These samples were tested for activity using IEF gel separated proteins stained for activity as previously described (section 2.2.7), to see if the protein was binding to the column. Analysis revealed that the protein was not binding to the column, but was coming straight through in the initial wash. This experiment was repeated 3 times with samples from different source livers, each with the same result.

7.2.5 Preparative Isoelectric Focusing

Preparative isoelectric focusing is a method of protein purification which separates proteins on the same basis as analytical isoelectric focusing, but on a preparative scale. Carrier ampholytes create a linear pH gradient through the gel as current is applied, and proteins travel through the gel until the pH in the gel is equal to the pI of the protein, at which point the protein remains and is concentrated. This technique therefore separates proteins on the basis of their overall charge. Preparative isoelectric focusing is useful for the early stage of preparation of many proteins, as it allows the fractionation of gram quantities of crude material with high resolution in a single run.

To prepare a gel for preparative isoelectric focusing, 75 ml distilled water was added to 5 ml Ampholine® Carrier Ampholytes pH range 7-9 (Pharmacia). Six 10.5 cm long electrode strips were soaked in the ampholyte mixture, and 3 placed at each end of the focusing tray. To prepare the sample, active eluant from the DEAE step (see section 7.2.2) was concentrated approximately 5-fold in an Amicon stirred cell apparatus using an XM-100 membrane at 4°C under pressure. 20 ml of this concentrated active sample was added to the ampholyte mixture. Ultrodex® gel (4 g) (Pharmacia) was added to the ampholyte-protein mixture to give a 4 % (w/v) gel slurry. This mixture in the beaker was weighed (1). After mixing gently, the slurry was poured into the focusing tray between the electrode strips. The beaker and remnants of ampholyte-protein mixture were weighed (2). The focusing tray containing the gel slurry was weighed (5). The tray was then placed on the cooling plate (4-6°C) and a fan mounted ~70 cm above the

tray to evaporate excess water from the suspension to the correct calculated level. The stated evaporation limit for the batch of gel was 33 %.

To work out the final weight for the focusing tray the following calculation was performed (numbers used are representative of a single experiment):

weight of gel slurry and beaker	181.25 g (1)
weight of beaker after pouring gel	82.92 g (2)
(1) - (2) = weight of gel slurry	98.33 g (3)

 $(3) \times 0.33$ (evaporation limit) = amount of water to be evaporated to reach correct limit

	32.4 g (4)
Initial weight of tray and slurry	252.59 g (5)
(5) - (4) = final weight of tray and slurry	<u>220.2 g</u>

Evaporation to the correct final weight took approximately 8 hours. When the correct weight was reached, the gel was connected to the LKB power supply, and proteins separated at 8 W constant power at 10°C for 14-16 hours. Separation was easily visualised by looking at the banding pattern produced after focusing, compared to a uniform red colour before focusing (Figure 7.5A). After separation, the fractionation grid was placed into the gel (Figure 7.5B). The grid separates the gel into 30 even sections, which can be removed separately from the tray. The pH of each fraction was then measured, and proteins separated from the gel by resuspending the gel in a small amount of 50 mM tetrasodium pyrophosphate buffer, pH 9, and passing the fraction over glass wool. The activity of each fraction was then measured by analytical IEF with activity staining, as described previously.

A good purification was achieved using this technique. Of primary importance, active protein was present after evaporation and focusing which took up to 24 hours (Figure 7.6). The only fraction in which activity was seen was typically the fraction nearest the cathode, which had a measured pH of \sim 8.8. The result obtained using preparative isoelectric focusing was \sim 2 ml of active protein, containing an enrichment of the protein of interest. At this stage, no further analysis was performed on this mixture, as it was deemed to be more important to retain as much active protein as possible for the next step of purification.

7-161

7.2.6 FPLC Hydrophobic Interaction Chromatography

To further purify this protein, hydrophobic interaction chromatography was the next step used. Separation using this technique is based on the differing surface hydrophobicity of proteins, such that proteins will interact with a phenyl group (or other hydrophobic group) attached to a matrix according to the amount of hydrophobic patches on their surface. Proteins bind to the column under conditions of high salt, and are eluted by decreasing the salt concentration of the buffer. Hydrophobic interaction chromatography is a non-specific technique, and does not usually give sharp separations (Scopes, 1987c). However the capacity is high and often good recoveries are achieved. The column used in this case was a phenyl-Superose column, dimensions 5×50 mm (Pharmacia), connected to a Pharmacia FPLC system (LCC-500), a Pharmacia single path UV-1 monitor, and a Kipp and Zonen BD-41 chart recorder. Proteins were identified by measuring the absorbance at 280 nm. The column was equilibrated by passing through 20 mM phosphate buffer pH 7.2, containing 1 M ammonium sulphate, for 25 minutes at 0.5 ml/min. After the column had been equilibrated, the protein solution was filtered through an 0.22 µM filter prior to injection. A linear gradient was then run from 1 M ammonium sulphate to 0 M ammonium sulphate over 60 min at 0.5 ml/min, and then held at 0 M for a further 30 min. Fractions were collected as proteins were eluted from the column by A₂₈₀ peaks on the chart recorder. Fractions were assayed for activity as previously described using the IEF activity assay, and in addition the protein content was assessed by staining an analytical IEF gel with Coomassie blue R250. The results show that the protein of interest was active, and protein staining of the IEF gel showed that the protein was pure (Figure 7.7, 7.8). To further check the purity of the protein, SDS-PAGE was performed on the sample as described previously (section 2.2.8) (Figure 7.9). Although a small impurity could be seen, the protein was estimated to be about 90 % pure. The monomeric molecular weight was estimated to be 40-45 kDa by comparison to SDS-7 protein size standards (Sigma).

Western blot analysis was performed on this sample using antibodies raised to sheep liver AlDH 1 as described in section 2.2.15 (Figure 7.10). The antibody did not interact with the purified protein, which helped to confirm our revised hypothesis that the isolated protein was not a member of the aldehyde dehydrogenase family, and supported the idea that it was in fact an alcohol dehydrogenase. A portion of the remainder of the purified protein was subjected to automated N-terminal sequence analysis as described in section 2.2.12. No sequence data were obtained. This is likely to have been due to one of two factors: 1) there was not enough sample present to sequence, or 2) the protein


Figure 7.5: Preparative Isoelectric Focusing

A: Gel slurry after focusing for 14-16 hours as described in the text, showing the characteristic banding patternB: The same focused slurry with the fractionation grid in place



Figure 7.6: IEF Gel of Preparative Isoelectric Focusing Fractions

Preparative IEF was run as described in section 7.2.5. The focused gel was then fractionated using the metal grid, the protein extracted from each section, and run on the analytical IEF gel above. The gel was stained for activity as described.

LANE:

- 1: Active before pIEF
- 2: pH 7.5 fraction
- 3: pH 7.6 fraction
- 4: pH 7.8 fraction
- 5: pH 8.0 fraction
- 6: pH 8.2 fraction
- 7: pH 8.4 fraction
- 8: pH 8.8 fraction

was N-terminally blocked. The remaining protein was cleaved using the chemical cyanogen bromide as described in section 2.2.11.1. Cyanogen bromide cleavage was chosen because it is likely to produce fewer fragments by cleaving at methionine residues than, say, by cleaving at arginine and lysine residues by trypsin. Peptides generated were separated by HPLC as described in section 2.2.11.6. Two large peaks were trapped, dried down completely under vacuum, and sent for automated sequencing as before. No sequence was obtained.

At this stage the main objective of this section of the project had been achieved. The putative novel protein had been purified. However, the identification of the protein was unclear. From the available data, isoelectric point (~8.5), subunit molecular weight (40-45 kDa), tissue and organism origin (sheep liver), western analysis (negative AlDH interaction), and activity (alcohol dehydrogenase), it seemed likely that the protein was a previously unidentified alcohol dehydrogenase isoenzyme. To confirm this hypothesis, the most important information needed was some unambiguous sequence which could be compared with other identified alcohol dehydrogenases. In addition the native molecular weight may provide some information as to the identity of the protein. In order to achieve these aims, more pure protein would need to be generated.

Multiple repetitions of the protocol outlined in Figure 7.11 did not yield any pure protein. This was puzzling, considering that the method had previously yielded pure The step which was not behaving as expected was the final stage of protein. hydrophobic interaction chromatography, despite the fact that the system and all components had not been altered from that which had previously been used. A major problem in the unsuccessful runs was that the protein did not seem to be binding to the column. In the first successful run, the activity was retained on the column, and was found in the final elution peak, indicating that the protein was quite hydrophobic. However, in subsequent runs the activity was found in the first peak to elute indicating that the protein was interacting only weakly or not at all with the column. Different conditions were tested to see if altering the pH or the ionic strength of the buffer system would allow binding of the targeted protein to the column in an effort to separate it from other contaminating proteins. None of these efforts were successful. It was necessary therefore to try alternative methods of obtaining pure protein after the preparative isoelectric focusing step.



Figure 7.7: IEF Gel of FPLC Hydrophobic Interaction Chromatography Fractions

IEF of fractions separated by phenyl superose hydrophobic interaction chromatography. Each protein peak was identified by the A280, and tested for the activity of interest. Activity was found in the fraction in lane 8.

LANE

- 1: Active sample before plEF
- 2: Pooled pIEF fraction
- 3 9: A₂₈₀ peaks from
- elution of phenyl superose column.
- 8: Active A₂₈₀ peak.



Figure 7.8: IEF Gel of FPLC Hydrophobic Interaction Chromatography Fractions (Protein Stain)

LANE:

1-5: A₂₈₀ peaks from elution of phenyl superose column stained with coomassie blue R250. Lane 3 contains the active fraction shown in lane 8 of Figure 7.7



Figure 7.9: SDS-PAGE of Purified Protein

Protein purified from hydrophobic interaction chromatography was run on an SDS gel and stained with coomassie blue R250

LANE:

- 1: Biorad molecular weight standards
- 2-4: Pure Cytosolic AlDH
- 5-8: Purified Novel Protein
- 9: Biorad molecular weight standards



LANE: 1 2 3 4

Figure 7.10: Western Blot of Pure Novel Protein and AlDH1

Western Blot performed as outlined in section 2.2.15, using 4-chloro-1 naphthol as substrate

LANE:

1,3: Pure Novel Protein (2 different Loadings)2,4: Pure AlDH 1 (2 different loadings)

7.2.7 HPLC Size Exclusion Chromatography

After preparative isoelectric focusing, there were only 5-10 proteins left in the mixture (depending on the success of this step) from which to isolate the desired protein. At this stage, size exclusion chromatography or gel filtration is a good option for further purification. This technique separates a mixture of proteins primarily on the basis of their Stokes radius, which in turn depends on the native molecular weight and on the overall shape of the protein. The columns used were Superose-12 and Superdex-75 (Pharmacia) under the control of a SMART HPLC (Pharmacia) system consisting of SMART manager software, uPrecision pump, variable volume gradient mixer, µSeparation unit, µFraction collector and µPeak monitor, connected to cooling apparatus. The Superdex 75 column (PC 3.2/30), has a 2.4 ml volume, 3.2×300 mm dimensions, 13 µM particle size and a molecular weight range from 3-70 kDa. The Superose 12 (PC 3.2/30) size exclusion column has a 2.4 ml volume, 3.2×300 mm dimensions, 10 µM particle size, and a molecular weight range from 1-300 kDa. The buffer used was 200 mM Hepes pH 7.2, 10 % (w/v) glycerol, 0.3 mM EDTA, 0.5 mM DTT, and 10 mM NaCl. Each run was at 40 µl/min, 10 °C, collecting 80 µlfractions, and proteins were identified by measuring the absorbance at 280 nm (A_{280}).

Size exclusion chromatography using a Superose-12 column (separation range 1-300 kDa) gave adequate separation of two major protein peaks (Figure 7.12), although the mixture of 5-10 proteins contained in the initial sample did not resolve very well (Figure 7.13, lane 9), indicating that the native (or effective) molecular weights of these proteins were too similar for this column to separate. The protein of interest also retained activity, which was consistently found in the first peak to elute from the column. SDS-PAGE of the sample after preparative isoelectric focusing and the peaks separated by gel filtration showed that the peak containing the activity, though symmetrical, contained a minimum of 2 bands, one at 40-45 kDa and one at 25-28 kDa (Figure 7.13). The band at 25-28 kDa often resolved as a doublet. The protein of interest was thought to be the 40-45 kDa band, due to the fact that the first successful purification had identified a band of this size (Figure 7.9). Further purification attempts concentrated on isolating this band..

To try to separate the proteins more effectively, a Superdex-75 column was used. This column has a separation range of 3-70 kDa, and a slightly larger particle size of 13 μ M, although the diameter and length dimensions were identical to the Superose-12. This column may separate a range of proteins within this molecular weight range better than

7-171



Figure 7.11: Developed Purification Scheme for Putative Alcohol Dehydrogenase

the Superose-12. However, this did not turn out to be the case. Elution profiles using both columns looked very similar (Figure 7.12, 7.14). Although activity was maintained throughout the process of gel filtration using both columns, separation of the protein of interest from the other proteins could not be achieved (Figure 7.13, 7.15). To estimate the molecular weight of the protein, the elution volume of the protein was compared to those of standards of known molecular weight using the Superdex-75 column (Figure 7.16). The buffer system used was 50 mM Hepes, pH 7.4, 20 mM NaCl, 0.5 mM EDTA, 10 % (v/v) glycerol. Standard proteins used were sheep liver AlDH 1 (MW 220 kDa), bovine lactoferrin (80 kDa), bovine serum albumin (BSA) (66 kDa), β -lactoglobulin (35 kDa), and α -lactalbumin (10-15 kDa). The estimate obtained from size exclusion chromatography (Figure 7.17) of the molecular weight of the unknown protein was 40-50 kDa, which is a very similar size to the band seen on SDS-PAGE.

From Figure 7.16 and SDS-PAGE data, it was concluded that the protein had a native molecular weight of 42-45 kDa, and a monomer size of 42-45 kDa, and therefore was a monomeric protein. This raised interesting questions. AlDH proteins are typically dimers or tetramers of 50-55 kDa subunits, while ADH proteins are typically dimers or tetramers of ~40 kDa subunits (Bosron *et al.*, 1993). The elucidated characteristics of the protein did not seem to fit into either of these categories, which made the isolation and identification of the protein vital, as it appeared to be a novel protein with an alcohol dehydrogenase activity. Monomeric ADH proteins have been identified in microbes, e.g. ethanol dehydrogenases from *Acetobacter aceti* and *A. polyoxogenes*, though all identified mammalian ADHs are dimeric (Reid and Fewson, 1994).

7.2.8 Hydroxyapatite Chromatography

Hydroxyapatite resin consists of an inert matrix to which calcium groups are attached. The bases of the attraction of proteins to hydroxyapatite resins are both non-specific attraction between protein positive charges and hydroxyapatite, and specific complexing of protein carboxyl groups with the calcium loci (Gorbunoff & Timasheff, 1984). Elution of bound proteins can be either by non-specific ion charge elution, or by specific displacement of protein groups from sites on the column to which they have complexed.

The initial buffer used was 10 mM NaH_2PO_4 , pH 7.0, containing 0.3 mM DTT and 0.5 mM EDTA, and to elute the proteins a series of increases in the phosphate ion concentration - 75 mM, 100 mM, 150 mM, 200 mM and 400 mM was used. Initially a batch trial was performed. Partially purified protein was added to 0.5 g (1.5 ml when



Figure 7.12: HPLC Size Exclusion Chromatography of Partially Purified Protein (Superose-12 Column)

The elution profile is created by continuously monitoring the A₂₈₀. The absorbance scale is approximately 0-0.5 absorbance units.



Figure 7.13: SDS-PAGE of HPLC Size Exclusion Chromatography Fractions (Superose-12)

These fractions relate to the Superose-12 size exclusion chromatography elution profile shown in Figure 7.12

LANE:

- 1: Fraction 9 (No activity)
- 2: Fraction 10 (Active)
- 3: Fraction 11 (Active)
- 4: Fraction 12 (No activity)
- 5: Fraction 13 (No activity)
- 6: Fraction 14 (No activity)
- 7: Fraction 15 (No activity)
- 8: Biorad molecular weight standards
- 9: Before Size Exclusion



Figure 7.14: HPLC Size Exclusion Chromatography of Partially Purified Protein (Superdex-75 Column)

The elution profile is created by continuously monitoring the A₂₈₀. The absorbance scale is approximately 0-0.5 absorbance units.





Figure 7.16: Elution Profile of Size Exclusion Standards (Superdex-75 Column)

The elution profile is created by continuously monitoring the A_{280} . The absorbance scale is approximately 0-0.5 absorbance units.



Figure 7.17: Calibration Curve Constructed From Superdex-75 Elution Profile

hydrated) HTP hydroxyapatite resin (Biorad) pre-equilibrated in initial buffer. After 10-15 min to allow binding, the mixture was centrifuged at $3000 \times G$ for 5 min in a Sorvall SM24 rotor. The supernatant was removed and kept for analysis. The procedure was repeated adding 1 ml of each increasing concentration phosphate buffer each time. The supernatants were analysed for content by SDS-PAGE and IEF gels with activity staining. Unfortunately no activity remained after this process. In addition, SDS-PAGE showed that none of the proteins appeared to bind to the resin. All Coomassie stained bands appeared in the supernatant from the 10 mM phosphate buffer wash.

7.2.9 Gel Purification

Because the initial successful purification of the novel protein had not been able to be repeated, and no other methods tried had worked, the next step was to purify the protein from an SDS gel and try to get some sequence information. With enough sequence information to 1) identify the protein as novel, and 2) place the protein in a known family of proteins, it was hoped that this information could be used to design oligonucleotides to screen a sheep liver cDNA library, and clone and express this protein for further characterisation.

Initially, because the molecular weight, isoelectric point, organism and tissue of origin, and an activity of the protein had been established, this information was entered into the SwissProt 2D electrophoresis database to see if any known protein matched these characteristics. This search yielded a number of potential matches from the alcohol dehydrogenase family with basic isoelectric points and similar subunit molecular weights including the class IV retinol dehydrogenase. However, these proteins are all dimeric, while it appears that the new protein is monomeric in its native state. So all available evidence still indicated that the protein was novel, and only a positive sequence identification would confirm this.

To purify the protein from SDS-PAGE, a 1.5 mm thick, 12 % (w/v) 37.5:1 acrylamide:bisacrylamide SDS gel with one large lane, and one lane for molecular weight standards was cast (section 2.2.10) and run as shown (Figure 7.18). After the gel had run, it was stained with Coomassie blue R250, as described in Materials and Methods. The band corresponding to 40-45 kDa was excised using a scalpel, and placed into dialysis tubing with 2-3 ml of 50 mM Tris-HCl pH 7.4 containing 0.1 % (w/v) SDS. The dialysis tubing containing the protein was placed into a horizontal electrophoresis tank,

filled with the same buffer. The protein was eluted from the gel by applying a 50 mA current for approximately 48 hours. The current was reversed for 5-10 min to remove any protein which may have adhered to the dialysis tubing. The protein-containing solution was then aspirated from the tube and concentrated under vacuum. An analytical SDS-PAGE gel showed that the protein was 95-100 % pure (Figure 7.19).



Figure 7.18: Schematic Diagram of Gel Purification

The SDS-PAGE is run as usual (see Materials and Methods). After staining, the band of correct size is identified using the protein ladder, and gently excised from the gel.

The first analyses carried out on the purified protein were another attempt at N-terminal sequencing and an amino acid analysis. To ensure that previous sequencing efforts had not failed due to insufficient quantity of protein, N-terminal sequencing was repeated. This was done by running a sample of the purified protein on a 15 % (w/v) 37.5:1 acrylamide: bisacrylamide SDS-gel. The gel was then electrotransferred onto PVDF membrane, as described in Materials and Methods, and sequenced. No sequence was obtained, which supported our initial hypothesis that the N-terminus was blocked. The amino acid analysis yielded the composition of amino acids of this protein (Table 7.1). From the resulting chromatogram of total amounts of each amino acid, the actual composition was calculated using the following formula:

Х

Moles of amino acid

Molecular weight of protein

Total moles in sample

Molecular weight of amino acid

Amino Acid	Molecular Weight (Da)	Residues (42 000 Da)	% of Total (42 000 Da)	Residues (45 000 Da)	% of Total (45 000 Da)
Ala	79	39	10.5	42	10.5
Gly	75	85	22.9	91	22.8
Val	117	32	8.6	34	8.5
Leu	131	25	6.7	27	6.8
Ile	131	13	3.5	14	3.5
Lys	147	18	4.9	19	4.8
Arg	175	9	2.4	10	2.5
Asx	132	24	6.5	26	6.5
Glx	146	28	7.5	30	7.5
His	156	6	1.6	6	1.5
Pro	115	17	4.6	19	4.8
Thr	119	19	5.1	20	5.0
Ser	105	37	10.0	40	10.0
Cys	122	4	1.1	5	1.3
Met	149	4	1.1	4	1.0
Phe	166	9	2.4	10	2.5
Tyr	182	2	0.5	2	0.5
Trp	204	0	0	0	0

Table 7.1: Amino Acid Analysis of Unknown Protein

The resulting amino acid composition was submitted to two database searching programs, ExPASy and PROPSEARCH, for matching with any other known protein. No close matches were found. The five closest matches made were: 1) fruit fly protein kinase sgg39, 2) hypothetical protein UL112, 3) protein kinase shaggy, zygotic, 4) fruit fly zeste-white 3-A, and 5) radish acetoacetyl-CoA-thiolase (all using 42 kDa molecular weight), and 1) sporozoite surface antigen, 2) procollagen EMF1-a, 3) neuroblast differentiation associated protein, 4) human AHNAK protein fragment, and 5) pecanex protein fragment (all using 45 kDa molecular weight). None of these have characteristics similar to the novel protein, and the closest match was less than 68 % reliable according to the program. Again, this search supports that this protein is a previously unidentified protein. However, to classify the protein as an alcohol dehydrogenase and confirm its uniqueness, some positive sequence information must be obtained.



Figure 7.19: SDS-PAGE of Gel Purification of Novel Protein

The protein of interest was gel purified by running on a 1.5mm 12% SDS gel (see Materials and Methods). After running the band of interest was excised and the protein eluted by applying an electric current as described.
(NB. The 'before gel purification' sample was not as clean as previous samples, e.g. in Figure 7.15)

LANE:

- 1: Before gel purification
- 2: After gel purification
- 3: Sigma molecular weight standards

7.2.10 Chemical and Enzymatic Cleavage

A number of methods were used for cleaving the purified protein into fragments suitable for the determination of internal sequence. The first method used was chemical cleavage using cyanogen bromide which should produce 5 fragments. This chemical cleaves on the C-terminal side of methionine residues, and as there are 4 methionine residues, complete cleavage will result in the production of 5 peptides. Five peptides should be easy to separate, and one or all peptides may then yield sequence information.

The cleavage was carried out as stated in section 2.2.11.1. The peptides generated were separated using an HPLC separation as outlined in section 2.2.11.6. HPLC separation of peptides generated by cyanogen bromide has an advantage that genuine peptide peaks may be seen as a close doublet due to approximately half the population being homoseryl lactone residue (cyclic) and the other half homoseryl residue. Peptide doublets of a good size were trapped as they eluted from the column, dried down under vacuum and sent for sequencing. No sequence was obtained from this method.

Enzymes used for cleavage were trypsin, chymotrypsin, thermolysin and leucine aminopeptidase. Under certain conditions these enzymes will cleave proteins at specific points, for example in a buffered solution trypsin cleaves on the C-terminal side of lysine and arginine residues. Though trypsin is a commonly used enzyme for producing fragments for sequencing, the amount of lysine and arginine residues contained in this protein (around 30), means that complete digestion would give about 30 peptides. These would probably be hard to separate, and this would hamper the process of obtaining sequence. However, a tryptic digest of limited duration may produce fewer peptides, by only allowing cleavage of the most accessible lysine and arginine residues. In this case, a limited digest of only a few minutes duration was used, compared to usual digestion times of up to 24 hours. The protein solution already contained SDS, which means that the protein would be denatured, and residues more accessible than in a native situation.

Tryptic digests were carried out as described in section 2.2.11.3 for 5 minutes on ice, and for 5 minutes, 15 minutes, 30 minutes and 1 hour at room temperature. The reactions were stopped by adding 1 mM PMSF, and the total mixture run on a 3-layer tricine gel as described in section 2.2.11.6. Tricine gels are SDS polyacrylamide gels made using tricine in place of glycine, which allows the separation of small proteins and

peptides as an alternative to HPLC separation (Schagger & von Jagow, 1987). The separated peptides were then electrotransferred onto a PVDF membrane, and the membrane was stained with Coomassie blue R250. Discrete bands were seen in the lanes containing both the 5 minute digestion on ice, and the 5 minute digestion at room temperature. The banding patterns were very similar in both lanes. Three bands which looked as though they contained enough material to sequence were isolated, and sent for sequencing. Although this was not successful the first time, a subsequent repeat of the above protocol yielded some sequence information. The sequence obtained showed two strong signals in each position (Table 7.2). This meant that either two peptides were resolving at the same place on the tricine gel, or two proteins had co-purified during the purification procedure, each one yielding a peptide of similar mobility upon tricine gel separation. It is likely that two peptides of similar size have run to exactly the same position on the tricine gel. If this is correct, then combinations of the sequence data should give the sequences of two peptides from the purified protein.

Position	First Choice	Second Choice	Third Choice
1	Thr		-
2	Gln	Val	Glu
3	Gly	Tyr	-
4	Asp	His	-
5	Glu	Val	-
6	Met	Phe	Lys
7	Asp	l Gly	-
8	Val	Gly	Glu

 Table 7.2: Sequence Obtained From Tryptic Digest of Unknown

 Protein

There are 288 possible peptides which could be made from combinations of the above amino acids. Each peptide was submitted to the database search program BLAST to assess any identity or similarity with protein sequences contained in the SwissProt database. As stated previously, two of these combinations should arise from cleavage of the purified protein or proteins, and if the protein was an already known protein or sheep homologue of a known protein from another organism, the peptides should have a high degree of similarity to this. Although many high matches were found, obviously 286 of these peptides do not derive from the purified protein, and are just hypothetical combinations. Some of the hypothetical combinations may be similar to the actual peptides, especially in cases where the peptides differ at one or two positions only. Similarity may be increased where the substituted amino acid is similar to the amino acid in the actual protein, e.g. Ile substituted for Leu etc. Therefore more than two peptides may be expected to show good similarity to the actual purified protein.

However, no proteins which were identified by this search engine to show similarity to one of the 288 peptides, shared other characteristics with the novel isolated protein, e.g. molecular weight, isoelectric point, activity, tissue distribution etc. Examples of close matches are: actin, adenylosuccinate synthetase, endothelin-1 receptor precursor, outer membrane usher protein and hypothetical protein H1140. All available information still pointed towards the purified protein as being a previously unidentified protein. Efforts were then concentrated on cleaving the gel purified protein with different enzymes, to try to generate new peptides. It was hoped that sequence information from one or more new peptides combined with sequence information from the tryptic digest, would confirm the novelty of the protein, and perhaps show some homology to the short chain reductase/dehydrogenase family.

Cleavage of the purified protein with chymotrypsin, thermolysin and leucine aminopeptidase was carried out as described in section 2.2.11. In addition, a limited acid cleavage was carried out as described in the same section. Peptide products were separated by 3-layer tricine gels and HPLC, and isolated peptides sent for sequencing. However, none of these methods yielded any further sequence information.

7.2.11 Activity Assays

The activity of the partially-purified protein was examined using a spectrophotometric assay in which the production of NADH was followed at 340 nm. Components of the assay included - NAD⁺ (5 mM), phosphate buffer (pH 7.4), 50 μ l of enzyme-containing fraction, and approximately 1 mM substrate. Substrates tested were all straight chain alcohols from one to six carbons in length, as well as glucose and sorbitol. Glucose and sorbitol were tested as substrates because during database searches using the sequence obtained from a tryptic digest, positive matches were found a number of times with a sorbitol dehydrogenase and a glucose dehydrogenase. These matches were the only positive matches which showed characteristics similar to those we found. However, the protein showed no activity with either of these compounds (Table 7.3). From this table it can be seen that this enzyme will oxidise a wide variety of alcohols, but the best

substrate by far is ethanol. It is interesting to note that methanol was not accepted as a substrate by this enzyme.

Substrate	Activity	Comment
Methanol	No	
Ethanol	Yes	Very good activity
Propanol	Yes	Low activity
Butanol	Yes	Low activity
Pentanol	Yes	Low activity
Hexanol	Yes	Low activity
Isopropanol	Yes	Low activity
Sorbitol	No	
Glucose	No	

Table 7.3: Summary of Enzyme Activity

7.2.12 Tissue Distribution

Using the IEF zymography activity assay described previously (section 2.2.7), crude extracts of various sheep organs were tested for the activity of interest. Briefly, 1-2 g of tissue was homogenised with 50 mM phosphate buffer pH 7.4 with 0.3 mM DTT and 0.5 mM EDTA, using a glass Potter-Elvetham homogeniser. This extract (20 μ l) was then placed onto the analytical isoelectric focusing gel and focused as described (2.2.7). The focused gel was stained for activity as described (2.2.7) using retinal and acetaldehyde as substrates. However, the real substrate was subsequently identified to be trace ethanol present in the NAD⁺ used (see section 7.2). Tissues examined were liver, kidney, heart, spleen, and brain. All tissues proved to be negative for the activity with the exception of liver (Figure 7.20).

7.3 Conclusions

The initial aim of this chapter was to isolate the protein with a basic isoelectric point responsible for a band of activity seen on an isoelectric focusing gel of a crude sheep liver extract. Initially it was thought that this activity was an aldehyde dehydrogenase activity, but upon further investigation, it was actually found to be an alcohol



Figure 7.20: IEF Gel of Crude Extracts From Different Tissues

Crude exctracts from each tissue were generated as described (section 7.2.12). Extracts were then run on an IEF gel and stained for activity using retinal as substrate

LANE:

- 1-3: Fractions from pIEF
- 4: Crude liver extract
- 5: Crude heart extract
- 6: Crude spleen extract
- 7: Crude kidney extract
- 8: Crude brain extract

dehydrogenase activity. There were a number of reasons for the initial identification. Firstly, the novel activity was seen when activity staining for cytosolic and mitochondrial liver AlDH was being carried out. The new band was the only unique band on the gel (Figure 7.1a, b, c), all other bands belonging to the class 1 and 2 AlDHs. Although the liver is a tissue in which alcohol dehydrogenase is usually found, no other ADH activity bands were identified on this gel. It seems surprising that no other alcohol dehydrogenases showed a positive stain except for this activity. In addition, the activity band appeared in the presence of some aldehyde substrates (acetaldehyde, p-carboxybenzaldehyde, 3,4-diiodosalicylaldehyde, retinal) but not with citral. In the assays with aldehyde substrate where no activity was observed, NAD^+ was present. The negative result initially led to the belief that the enzyme must be aldehyde-dependent, and obscured the fact that the NAD⁺ was contributing trace amounts of an alternative substrate, ethanol. It is not clear why negative results were obtained with some aldehyde substrates when ethanol-containing NAD⁺ was present, but the possible explanation is that the aldehyde acted as a product inhibitor of the enzyme which was in fact found to be an alcohol dehydrogenase. A second factor which reinforced the initial, wrong, assumption that the enzyme was an aldehyde dehydrogenase was that a new AlDH specific for retinal oxidation with a basic pI had recently been identified and purified from rat liver and kidney (Labrecque et al., 1993; Posch et al., 1992). The initial hypothesis was that the high pI activity seen on IEF gels was a sheep liver AIDH equivalent to one or both of these recently discovered proteins. This was exciting, and the new enzyme was considered an important protein to pursue, because the rodent and murine AIDH systems are quite different from those of other mammals such as humans, sheep, cows and horses. Differences are seen in nomenclature, tissue distribution and kinetic parameters.

For example, it has been suggested that human cytosolic AlDH 1 should not be considered significant in acetaldehyde metabolism (Klyosov, 1996; Klyosov *et al.*, 1996; Rashkovetsky *et al.*, 1994), which raises questions as to the identity of its physiological substrate. In the rat, however, the main cytosolic enzyme has a lower K_m for acetaldehyde (Klyosov, 1996) than the human enzyme, and may play a major role in the metabolism of acetaldehyde, hence the need for specific retinal-oxidising enzymes such as those which have recently been identified (Labrecque *et al.*, 1993; Posch *et al.*, 1992; Wang *et al.*, 1996; Zhao *et al.*, 1996). Initially it was thought that the activity identified may fulfil a similar role in sheep and perhaps in humans, i.e. a specific retinal-oxidising AlDH isoenzyme. This, however, did not turn out to be the case, as the enzyme was

subsequently determined to have ADH activity, and did not appear to oxidise aldehydes in spectrophotometric assays.

Despite a large volume of work being published on the retinal dehydrogenases, none have been found in human tissues. With the high level of sequence identity seen between retinal dehydrogenases and aldehyde dehydrogenases capable of oxidising retinal (see Chapter 5), and the cloning of the genes which code for these enzymes, it would seem likely that screening human libraries using oligonucleotides designed to regions of high sequence identity would result in the identification of such a retinal-specific protein, should one exist. This has not been the case yet, 5 years after the first enzyme was identified. We suggest that the properties of the major human AlDH 1 indicate that this isoenzyme may have evolved for the purpose of retinal oxidation in sheep and humans (see Chapter 5 and Chapter 1 for a full discussion).

In summary, a protein with alcohol dehydrogenase activity has been purified in native form and denatured form by subcellular fractionation, PEG precipitation (10-20 % (w/v)), exclusion from DEAE-iontosorb at pH 6.5, ultrafiltration, preparative isoelectric focusing, hydrophobic interaction chromatography, gel filtration (Superdex-75), and gel Characteristics of the protein which have been elucidated are native purification. molecular weight, 40-45 kDa; subunit molecular weight, 40-45 kDa; isoelectric point, 8.5 ± 0.2 ; organism and tissue of origin, sheep liver; subcellular location, cytosolic; and activity, NAD⁺-dependent activity with ethanol and slight activity with propanol, butanol, pentanol and hexanol, but no activity with methanol, glucose or sorbitol. In addition the protein is thought to be N-terminally blocked. Sequence information yielded no clue as to the identity of this protein. All information points to it being a novel alcohol dehydrogenase isozyme, but more positive sequence information is really needed to confirm this. Although the protein has not been positively identified during the course of this project, a potentially novel protein has been isolated, and it is hoped that further work will confirm its identity, and allow complete characterisation of this interesting protein.

8. Chapter 8: Discussion and Future Work

8.1 Discussion and Conclusions

The pleiotropic and potent retinoids are important in a multitude of biological processes including fetal development. The understanding of the manner in which retinoic acid (the most biologically active retinoid) exerts its effects, may be crucial in directing treatment in particular to fetal alcohol syndrome, diseases of the skin and certain cancers in which retinoic acid has already been shown to be a major effector. In addition to the understanding of the effects of retinoic acid, it is important to understand the regulation of the production and availability of retinoic acid in different cell and tissue types.

Many enzymes have been shown to be able to catalyse the production of retinoic acid from its biological precursors - vitamin A alcohol (retinol) and aldehyde (retinal) and β -carotene *in vitro*. A major question yet to be answered is which of these enzymes are physiologically relevant for retinoic acid production *in vivo*.

From the work presented in this thesis, along with information published by other groups, we propose that in sheep and humans the major cytosolic AlDH (AlDH 1) plays an important role in retinoic acid biosynthesis *in vivo*. In the work described in Chapter 3, purified sheep and human AlDH 1 were used to determine kinetic parameters for the oxidation of all-*trans* and 9-*cis* retinal to retinoic acid. The K_m values determined using the method of fluorimetric kinetic analysis are low (0.05 - 0.14 μ M). This means that at the estimated physiological concentration of retinal (0.1 μ M), cytosolic AlDH 1 would be catalytically efficient. In addition, a comparison with other enzymes which have been proposed to play a role in *in vivo* retinoic acid production (aldehyde oxidase, xanthine oxidase, cytochrome P450), shows that AlDH 1 has a lower K_m than all other enzymes except the retinal-specific dehydrogenases, which have a similar K_m. The K_m values for retinal of the other enzymes preclude these proteins from having a major role in retinoic acid at physiological concentrations.

In the research described in Chapter 4, recombinant cellular retinol-binding protein (CRBP) was expressed and purified from the vector pT7-7-CRBP. CRBP I is a small intracellular lipid-binding protein whose ligands are all-*trans* retinol and all-*trans* retinal.

Much work has suggested that CRBP plays a role in regulating retinoic acid production by sequestering the lipophilic retinol and retinal from the aqueous environment. It has also been proposed that CRBP bound to retinol or retinal is the major physiological substrate for metabolic enzymes including AlDH. It has been experimentally determined that a large proportion of cytosolic retinol (and retinal by comparison) is found in a complex with CRBP (Harrison *et al.*, 1987). If the role of CRBP in retinoid metabolism is as central as suggested in the above hypothesis, in order to play a role in retinoic acid production *in vivo*, metabolic enzymes would be required to accept retinal bound to CRBP as a substrate. This section of work also shows that the recombinant CRBP binds all-*trans* but not 9-*cis* retinal, and that sheep and human AlDH 1 accept CRBP-bound all-*trans* retinal as a substrate with a low K_m of 0.2 - 1.0 μ M.

However, it has also been proposed that the major role of CRBP does not lie in retinoic acid production, but in directing retinoids into storage (see section 1.3.2.3). If this scenario is true, then sheep and human AlDH still may play a major role *in vivo* as retinal-oxidising enzymes. Both have been shown to oxidise the biologically important retinoids free 9-*cis* and all-*trans* retinal with low K_m values. The work presented in this thesis has shown that AlDH 1 from sheep and humans can catalyse the conversion of free all-*trans*, free 9-*cis* and CRBP-bound retinal to the respective retinoic acid isomer at physiologically relevant concentrations.

Additional evidence supporting the idea that, in sheep and humans, the major cytosolic AlDH is the most important enzyme in retinal oxidation comes from sequence information presented in Chapter 5. AlDH 1 from sheep and humans are the only aldehyde dehydrogenase enzymes in these organisms with the ability to oxidise retinal (with the exception of eye AlDH enzymes which have different functions). The other major AlDH isoform found in human and sheep liver, AlDH 2, does not accept retinal as a substrate. Extensive screening of human genomic libraries (Hsu *et al.*, 1995), has failed to isolate human homologues of the retinal-specific enzymes found in rats and mice (Penzes *et al.*, 1997; Posch *et al.*, 1992; Wang *et al.*, 1996; Zhao *et al.*, 1996) although a number of new AlDH genes were discovered. Both AlDH 1 and 2 oxidise acetaldehyde, which was originally thought to be the biological substrate for these enzymes. However, it was subsequently hypothesised that AlDH 2 was the major acetaldehyde-oxidising enzyme (K_m for acetaldehyde of <0.1 μ M-9 μ M, compared to 22-483 μ M for AlDH 1) (Rashkovetsky *et al.*, 1994). Although AlDH 1 can oxidise acetaldehyde, its high K_m for acetaldehyde and its ability to oxidise retinal indicates that AlDH 1 would be biologically

important for acetaldehyde oxidation only in the presence of relatively high acetaldehyde concentrations which are not usually found in vivo, and its main substrate is probably retinal.

In addition, the multiple sequence alignment constructed (Figure 5.1) is the first comparison of the recently identified retinal-specific aldehyde dehydrogenase (RALDH) sequences with other ADHs, some of which can also oxidise retinal. This sequence alignment showed that the four RALDH genes isolated actually encode only 2 proteins, RALDH 1 and 2. We propose that in rats in mice, RALDH enzymes are the major retinal-oxidising enzymes, while the lower K_m of rat and mouse AlDH 1 for acetaldehyde means that both AlDH 1 and 2 may function in acetaldehyde oxidation as well as in the detoxification of other aldehydes.

Preliminary investigations were carried out into the role of AIDH 1 in retinoic acid production in vivo. Using the human neuroblastoma cell line SH-SY-5Y, which expresses ALDH 1, it was shown that exogenously added all-trans and 9-cis retinal was able to invoke morphological changes in the cells attributable to retinoic acid; i.e., that retinal is metabolised to retinoic acid in these cells. An initial decrease in cell differentiation was seen on the concurrent addition of all-trans retinal and disulfiram. This effect of disulfiram, a potent immediate inhibitor of AlDH 1, indicates that AlDH 1 may play an important role in retinoic acid production in vivo. The addition of ethanol and all-trans retinal at the same time was expected to show similar results to those seen when disulfiram was added, due to competitive inhibition by the product of ethanol oxidation, acetaldehyde, on retinal oxidation. However, the addition of ethanol potentiated the effect produced by all-trans retinal alone by approximately 30-50 %. From this result a new mechanism by which fetal alcohol syndrome may arise was proposed (Figure 6.11), whereby the increase in ethanol and acetaldehyde initiates an upregulation of ADH and AlDH protein, which subsequently increases retinoic acid production.

The last section of work presented in this thesis describes the purification of a novel protein with alcohol dehydrogenase activity. Although it was originally thought that this protein was an aldehyde dehydrogenase with the ability to oxidise retinal, this was shown not to be the case. Characteristics of the protein indicated that it may be a previously unidentified ADH, and as such justified further investigation. Active protein was purified using centrifugation, PEG precipitation, ion-exchange chromatography, preparative

isoelectric focusing, and hydrophobic interaction chromatography. Inactive protein was purified from a combination of these methods, with final purification from SDS gels, and used to obtain some internal sequence. However, unambiguous sequence which could positively identify this protein as an ADH and also indicate its novelty, remains elusive. It is hoped that work will be continued to confirm these conclusions.

8.2 Future Work

The work initiated in this thesis, which strongly suggests that AlDH 1 plays a primary role in retinoid metabolism in sheep and humans, provides a basis for further study in the area of enzymology and regulation of retinoic acid production.

We have proposed that A1DH 1 is the main retinal-oxidising enzyme in both sheep and humans, as opposed to the situation in rats and mice where retinal-specific enzymes exist. To prove this hypothesis, it would be necessary to ascertain whether human and homologues of the retinal-specific aldehyde sheep dehydrogenases exist. Oligonucleotide primers based on regions of strong homology between aldehyde dehydrogenase sequences (designed using the multiple sequence alignment created in this thesis) could be used to screen cDNA or genomic sheep and human libraries. Depending on the conditions and the design of the primers, a number of aldehyde dehydrogenase genes could be isolated. It may prove to be more difficult to determine which clones are previously identified aldehyde dehydrogenases, and which are novel aldehyde dehydrogenases that possess the ability to oxidise retinal. DNA sequencing may be used to aid in the identification of known aldehyde dehydrogenases, but depending on the number of positive clones obtained, this may be a large undertaking. Expression screening using antibodies raised to sheep AlDH 1 may prove useful in obtaining aldehyde dehydrogenase clones, however the success of this method may depend on the epitope to which the antibodies were raised. Alternatively, database information available from the human genome project could be screened, again using similar primers designed to areas of high homology in aldehyde dehydrogenases as mentioned above. Depending on the position of the gene within the genome, this method may prove to be technically less demanding, and yield a rapid result. The strategy outlined above has been employed successfully by Hsu, Yoshida and co-workers (Hsu et al., 1995), who isolated a number of new aldehyde dehydrogenase sequences using degenerate primers. None of these positive clones show similarity to any of the retinal-oxidising aldehyde dehydrogenases (see Figure 5.2), in fact H10 and H7 cluster most closely with class 3

enzymes, H6 is an outlier within the class 1 enzyme grouping, and H9 is less related to mammalian class 1 and 2 enzymes than *E.coli* and yeast AlDH 1 enzymes (see chapter 5). So, it seems likely that if a human retinal dehydrogenase enzyme exists, that it would have been isolated here. If the hypothesis proposed in this thesis is correct, then no retinal-specific enzyme from large mammals including sheep and humans would be expected to be found.

It has been shown that human and sheep AIDH 1 both accept CRBP-retinal as a substrate. The type of interaction between the metabolic enzyme and binding protein is unknown. It has been proposed that the ligand is transferred between the two proteins by a 'direct transfer mechanism', i.e., an interaction between the proteins allowing transferral of the ligand without the ligand first dissociating into solution. However it is possible that the release of ligand is triggered by a more transient interaction between the binding protein and metabolic enzyme, and then enters the active site due to the proximity of the enzyme (which just triggered the release of its substrate). The use of chemical cross-linking agents may determine whether or not the two proteins interact with each other, although conclusions drawn from such an experiment may be limited. The incubation of a cross-linking reagent with two purified proteins *in vitro* may produce a false positive. More realistic results may be seen using small amounts of protein, and a short incubation time, or alternatively using a cruder mix of proteins and seeing which proteins from the mixture interact. A more effective method of studying the interactions between proteins is using surface plasmon resonance (SPR). The use of SPR involves the immobilisation of one of the proteins of interest on the sensor chip in the machine, while the other protein in solution is passed over the immobilised protein. An interaction between the proteins results in an increase in the mass of immobilised protein, and a distortion of the angle of total internal reflection of the surface to which it is bound. Such a method may more accurately determine if there is a specific interaction between CRBP-retinal and AlDH. Modelling studies are currently being carried out to ascertain whether the surfaces of holo-CRBP and AlDH 1 are complementary in order for recognition to take place.

The recent elucidation of the tertiary and quaternary crystal structures of the three main classes of AlDH opens up new avenues of investigating the relationship between the structure and function of these enzymes, and in addition may allow the basis for different substrate specificities to be investigated. Of particular interest, are the differences in the substrate binding pockets of the major cytosolic and mitochondrial aldehyde dehydrogenases which allow retinal to be accepted as a substrate in the cytosolic but not the mitochondrial enzyme. It was suggested in Chapter 5 that the amino acid residues responsible are likely to be Ile 303 and Val 459, which in the class 2 enzymes are Cys and Phe respectively. To examine this hypothesis, residues 459 and 303 in human or sheep AlDH could be mutated by site-directed mutagenesis to the corresponding class 2 residues (Cys and Phe). Expressed and purified mutants could then be studied as described in Chapters 3 and 4, to determine if they exhibit kinetic parameters for the oxidation of retinal. An experiment that could be run in parallel with this would be to mutate the Phe and Cys residues in the mitochondrial enzyme to small hydrophobic amino acids. The similarity of the two classes of enzyme, especially around the active site, makes it very likely that the structural differences resulting in the different substrate specificities lie with these two residues located within the substrate binding pocket. Expressed and purified mitochondrial AlDH mutants (as described above), are likely to accept retinal as a substrate.

The successful development of an HPLC method such as that attempted in Chapter 3 would be vital to a number of further experiments. Obtaining reproducible results using an HPLC-based method would necessitate development of a system dedicated to retinoid extraction and analysis. Once such a system was set up satisfactorily, the neuronal cell culture work begun in this thesis (Chapter 6), could be continued. To be able to draw valid conclusions about the role of AlDH 1 in vivo, all enzymes with the ability to oxidise retinal which are expressed in the cell line used (in this case the human neuroblastoma cell line SH-SY5Y) must be identified. This may be achieved either by obtaining antibodies raised to each enzyme, and performing Western blots to see if the protein is expressed, or using oligonucleotides designed to unique parts of the sequence of each enzyme to determine if the RNA is present (Northern blotting). To determine the role that AIDH 1 plays in retinoic acid production in the cells, the enzyme activity needs to be 'knocked out', and the effect of the decrease in activity on retinoic acid production studied. One method of decreasing the amount of protein is to microinject antisense mRNA or oligonucleotides. The interaction of the antisense oligonucleotide with the sense strand results in RNA which cannot be translated. The consequence of this decrease in active protein on retinoic acid production could then be seen by using HPLC quantification. This procedure is, however, technically demanding, and may not result in definite conclusions. Microinjection is a delicate procedure, and the oligonucleotides need to be tagged with a fluorescent label to enable uptake efficiency to be determined. The procedure has been successful in a number of experiments, resulting in a reduction in gene expression and activity by 80-98 % (Izant & Weintraub, 1984; Kim & Wold, 1985; Kluess *et al*, 1993). This technique could be used to decrease the amount of aldehyde dehydrogenase protein (or other retinal-oxidising protein) in the cells. From the reduction in protein, a decrease in the production of retinoic acid may be observed, which may indicate that AlDH 1 plays a primary role in retinoic acid production. However, due to the importance of the production and control of retinoic acid, it is likely that compensatory mechanisms exist, i.e., a different protein with retinal-oxidising ability which does not usually play a large role in retinoic acid production, may be 'recruited' to maintain spatiotemporal retinoic acid concentrations.

Finally, the novel protein with alcohol dehydrogenase activity needs to be positively identified. Further purification attempts may yield pure active protein. The major problem in positively identifying the purified novel protein was obtaining unambiguous sequence data. Once more sequence information is obtained, and the protein can be confirmed as a novel alcohol dehydrogenase, a number of experiments may be initiated. Firstly, the protein could be cloned. A sheep liver cDNA library could be screened for the gene of the novel protein by using data from the obtained protein sequence to design oligonucleotides. A more efficient way may be to raise antibodies to the purified protein. To raise antibodies, active protein is not necessary, and protein purified by such a method as that successfully described in this thesis (Chapter 7 - gel purification) may be used. Expression cloning using antibodies could then be used to isolate the gene. This putative alcohol dehydrogenase is of particular interest as it appears to be active as a monomeric protein, which is an unusual property for a higher mammalian alcohol dehydrogenase.

Appendix 1: Calculation of free retinal concentration for the determination of kinetic constants when using CRBP-retinal as a substrate

 $R_{f} = Free retinal$ $P_{f} = Protein (CRBP)$ $R_{t} = Total retinal (R_{f} + [complex])$ $P_{t} = Total protein (P_{f} + [complex])$ $P_{f} = P_{t} - [complex]$ $= P_{t} - (R_{t} - R_{f})$

 $R_{f} + P_{f} \longleftarrow Complex CRBP-retinal$ $K_{d} = \frac{R_{f}P_{f}}{[complex]}$ $= \frac{R_{f}(P_{t} - R_{t} + R_{f})}{R_{t} - R_{f}}$ $K_{d}R_{t} - K_{d}R_{f} = R_{f}^{2} + R_{f}(P_{t} - R_{t})$ $0 = R_{f}^{2} + R_{f}(P_{t} - R_{t}) + R_{f}K_{d} - K_{d}R_{t}$ $= R_{f}^{2} + R_{f}(P_{t} - R_{t} + K_{d}) - K_{d}R_{t}$ This example uses 12 uM CRBP and 6 uM retinal (2:1 ratio)

If $P_t = 12$ $R_t = 6$ $K_d = 0.1$

This example uses 12 μ M CRBP and 6 μ M retinal (2:1 ratio as used in experimental work), and the highest estimated K_d of 100 nM (0.1 μ M).

$$R_{f} = \frac{-6.1 \pm \sqrt{6.1^{2} + 2.4}}{2}$$

$$R_{f} = \frac{-6.1 \pm \sqrt{37.21 + 2.4}}{2}$$

$$= \frac{-6.1 \pm 6.294}{2}$$

$$= 0.0968 \,\mu M$$

References:

Allen, G. (1981). In Sequencing of proteins and peptides, pp. 92-93, 278-281. Elsevier/North Holland, Amsterdam.

Allenby, G., Bocquel, M. T., Saunders, M., Kazmer, S., Speck, J., Rosenberger, M., Lovey, A., Kastner, P., Grippo, J. F., Chambon, P. & et al. (1993). Retinoic acid receptors and retinoid X receptors: interactions with endogenous retinoic acids. *Proceedings of the National Academy of Sciences of the United States of America* **90**(1), 30-4.

Ambroziak, W. & Pietruszko, R. (1991). Human aldehyde dehydrogenase. Activity with aldehyde metabolites of monoamines, diamines, and polyamines. *Journal of Biological Chemistry* **266**(20), 13011-8.

Ambroziak, W. & Pietruszko, R. (1993). Metabolic role of aldehyde dehydrogenase. *Advances in Experimental Medicine & Biology* **328**, 5-15.

Ang, H. L., Deltour, L., Hayamizu, T. F., Zgombic-Knight, M. & Duester, G. (1996a). Retinoic acid synthesis in mouse embryos during gastrulation and craniofacial development linked to class IV alcohol dehydrogenase gene expression. *Journal of Biological Chemistry* **271**(16), 9526-34.

Ang, H. L., Deltour, L., Zgombic-Knight, M., Wagner, M. A. & Duester, G. (1996b). Expression patterns of class I and class IV alcohol dehydrogenase genes in developing epithelia suggest a role for alcohol dehydrogenase in local retinoic acid synthesis. *Alcoholism, Clinical & Experimental Research* **20**(6), 1050-64.

Attwood, T.K., Payne, A.W.R., Michie, A.D., Parry-Smith, D.J. (1997). CINEMA - Colour Interactive Editor for Multiple Alignments.

Attwood, T.K., Perkins, D.N. XALIGN, Version 1.0.
Bailey, J. S. & Siu, C. H. (1988). Purification and partial characterisation of a novel binding protein for retinoic acid from neonatal rat. *Journal of Biological Chemistry* **263**(19), 9326-32.

Banaszak, L., Winter, N., Xu, Z., Bernlohr, D. A., Cowan, S. & Jones, T. A. (1994). Lipid-binding proteins: a family of fatty acid and retinoid transport proteins. *Advances in Protein Chemistry* **45**, 89-151.

Bashor, M. M., Toft, D. O. & Chytil, F. (1973). In vitro binding of retinol to rat-tissue components. Proceedings of the National Academy of Sciences of the United States of America 70(12), 3483-7.

Bavik, C. O., Eriksson, U., Allen, R. A. & Peterson, P. A. (1991). Identification and partial characterization of a retinal pigment epithelial membrane receptor for plasma retinol-binding protein. *Journal of Biological Chemistry* **266**(23), 14978-85.

Benbrook, D., Lernhardt, E. & Pfahl, M. (1988). A new retinoic acid receptor identified from a hepatocellular carcinoma. *Nature* **333**(6174), 669-72.

Bernhard, S. A. (1988). The intracellular equilibrium thermodynamic and steady-state concentrations of metabolites. *Cell Biophysics* **12**, 119-32.

Bhat, P. V., Poissant, L., Falardeau, P. & Lacroix, A. (1988a). Enzymatic oxidation of all-*trans* retinal to retinoic acid in rat tissues. *Biochemistry & Cell Biology* 66(7), 735-40.

Bhat, P. V., Poissant, L. & Lacroix, A. (1988b). Properties of retinal-oxidizing enzyme activity in rat kidney. *Biochimica et Biophysica Acta* **967**(2), 211-7.

Bhat, P. V., Labrecque, J., Boutin, J. M., Lacroix, A. & Yoshida, A. (1995). Cloning of a cDNA encoding rat aldehyde dehydrogenase with high activity for retinal oxidation. *Gene* **166**(2), 303-6.

Bhat, P. V., Poissant, L. & Wang, X. L. (1996). Purification and partial characterisation of bovine kidney aldehyde dehydrogenase able to oxidise retinal to retinoic acid. *Biochemistry and Cell Biology* 74(5), 695-700. Biedler, J. L., Helson, L. & Spengler, B. A. (1973). Morphology and growth, tumorigenicity, and cytogenetics of human neuroblastoma cells in continuous culture. *Cancer Research* **33**, 2643-2652.

Blackwell, L. F., Buckley, P. D. & MacGibbon, A. K. H. (1989). Aldehyde dehydrogenases - kinetic characterisation. In *Human Metabolism of Alcohol* (Crow, K. E. & Batt, R. D., eds.), Vol. II, pp. 89-104. CRC Press, Boca Raton, Florida.

Blaner, W. S. & Olsen, J. A. (1994). Retinol and Retinoic Acid Metabolism. In *The Retinoids: Biology, Chemistry, and Medicine*. 2nd Edition edit. (Sporn, M. B., Roberts, A. B. & Goodman, D. S., eds.), pp. 229-250. Raven Press, Ltd., New York.

Bleasby, A. J. & Wooton, J. L. (1990). Construction of validated, non-redundant composite protein sequence databases. *Protein Engineering* **3**, 153-179.

Blomhoff, R., Norum, K. R. & Berg, T. (1985). Hepatic uptake of [3H]retinol bound to the serum retinol binding protein involves both parenchymal and perisinusoidal stellate cells. *Journal of Biological Chemistry* **260**(25), 13571-5.

Boerman, M. H. & Napoli, J. L. (1991). Cholate-independent retinyl ester hydrolysis. Stimulation by Apo-cellular retinol-binding protein. *Journal of Biological Chemistry* **266**(33), 22273-8.

Boerman, M. H. & Napoli, J. L. (1995a). Characterization of a microsomal retinol dehydrogenase: a short-chain alcohol dehydrogenase with integral and peripheral membrane forms that interacts with holo-CRBP (type I). *Biochemistry* **34**(21), 7027-37.

Boerman, M. H. & Napoli, J. L. (1995b). Effects of sulfhydryl reagents, retinoids, and solubilization on the activity of microsomal retinol dehydrogenase. *Archives of Biochemistry & Biophysics* **321**(2), 434-41.

Boerman, M. H. & Napoli, J. L. (1996). Cellular retinol-binding protein-supported retinoic acid synthesis. Relative roles of microsomes and cytosol. *Journal of Biological Chemistry* 271(10), 5610-6.

Bok, D. & Heller, J. (1976). Transport of retinol from the blood to the retina: an autoradiographic study of the pigment epithelial cell surface receptor for plasma retinolbinding protein. *Experimental Eye Research* **22**(5), 395-402.

Boleda, M. D., Saubi, N., Farres, J. & Pares, X. (1993). Physiological substrates for rat alcohol dehydrogenase classes: aldehydes of lipid peroxidation, omega-hydroxyfatty acids, and retinoids. *Archives of Biochemistry & Biophysics* **307**(1), 85-90.

Bollag, W. (1996). The retinoid revolution. Overview. FASEB Journal 10(9), 938-9.

Boylan, J. F. & Gudas, L. J. (1992). The level of CRABP-I expression influences the amounts and types of all-*trans*-retinoic acid metabolites in F9 teratocarcinoma stem cells. *Journal of Biological Chemistry* **267**(30), 21486-91.

Bosron, W. F., Ehrig, T. & Li, T.-K. (1993). Genetic factors in alcohol metabolism and alcoholism. Seminars on Liver Disease 13, 126-135.

Brand, N., Petkovich, M., Krust, A., Chambon, P., de The, H., Marchio, A., Tiollais, P.
& Dejean, A. (1988). Identification of a second human retinoic acid receptor. *Nature* 332(6167), 850-3.

Bridges, C. D. (1990). High-performance liquid chromatography of retinoid isomers: an overview. *Methods in Enzymology* **189**, 60-9.

Brzozowski, A. M., Derewenda, U., Derewenda, Z. S., Dodson, G. G., Lawson, D. M., Turkenburg, J. P., Bjorkling, F., Huge-Jensen, B., Patkar, S. A. & Thim, L. (1991). A model for interfacial activation in lipases from the structure of a fungal lipase-inhibitor complex. *Nature* **351**(6326), 491-4.

Bucco, R. A., Melner, M. H., Gordon, D. S., Leers-Sucheta, S. & Ong, D. E. (1995). Inducible expression of cellular retinoic acid-binding protein II in rat ovary: gonadotropin regulation during luteal development. *Endocrinology* **136**(6), 2730-40.

Buckley, P. D., Motion, R. L., Blackwell, L. F. & Hill, J. P. (1991). pH effects on cytoplasmic aldehyde dehydrogenase from sheep liver. Advances in Experimental Medicine and Biology 284, 31-41.

Calzi, M. L., Raviolo, C., Ghibaudi, E., de Gioia, L., Salmona, M., Cazzaniga, G., Kurosaki, M., Terao, M. & Garattini, E. (1995). Purification, cDNA cloning, and tissue distribution of bovine liver aldehyde oxidase. *Journal of Biological Chemistry* **270**(52), 31037-45.

Cambillau, C., Roussel, A., Inisan, A-G., & Knoups-Mouthey, G. (1996). Turbo-Frodo Manual, Version 5.5b. University of Marseilles, France.

Cavazzini, D., Galdieri, M. & Ottonello, S. (1996). Retinoic acid synthesis in the somatic cells of rat seminiferous tubules. *Biochimica et Biophysica Acta* **1313**(2), 139-45.

Chai, X., Boerman, M. H., Zhai, Y. & Napoli, J. L. (1995a). Cloning of a cDNA for liver microsomal retinol dehydrogenase. A tissue-specific, short-chain alcohol dehydrogenase. *Journal of Biological Chemistry* **270**(8), 3900-4.

Chai, X., Zhai, Y., Popescu, G. & Napoli, J. L. (1995b). Cloning of a cDNA for a second retinol dehydrogenase type II. Expression of its mRNA relative to type I. *Journal of Biological Chemistry* 270(47), 28408-12.

Chai, X., Zhai, Y. & Napoli, J. L. (1996). Cloning of a rat cDNA encoding retinol dehydrogenase isozyme type III. *Gene* 169(2), 219-22.

Chambon, P. (1996). A decade of molecular biology of retinoic acid receptors. *FASEB Journal* **10**(9), 940-54.

Chandler, G. W. & Schwartz, S. J. (1987). HPLC separation of *cis-trans* carotene isomers in fresh and processed fruits and vegetables. *Journal of Food Science* **52**, 669-672.

Chang, C. & Yoshida, A. (1994). Cloning and characterisation of the gene encoding mouse mitochondrial aldehyde dehydrogenase. *Gene* **148**, 331-336.

Chang, C. & Yoshida, A. (1997). Human fatty aldehyde dehydrogenase gene (ALDH10): organisation and tissue-dependent expression. *Genomics* **40**, 80-85.

Chen, H., Namkung, M. J. & Juchau, M. R. (1995). Biotransformation of all-transretinol and all-trans-retinal to all-trans-retinoic acid in rat conceptal homogenates. *Biochemical Pharmacology* **50**(8), 1257-64.

Chen, H., Namkung, M. J. & Juchau, M. R. (1996). Effects of ethanol on Biotransformation of All-*Trans*-retinol and All-*Trans*-Retinal to All-*Trans*-Retinoic Acid in Rat Conceptal Cytosol. *Alcoholism: Clinical and Experimental Research* **20**(5), 942-947.

Chen, M., Achkar, C. & Gudas, L. J. (1994). Enzymatic conversion of retinaldehyde to retinoic acid by cloned murine cytosolic and mitochondrial aldehyde dehydrogenases. *Molecular Pharmacology* **46**(1), 88-96.

Clarren, S. K. & Smith, D. W. (1978). The fetal alcohol syndrome. New England Journal of Medicine 298(19), 1063-7.

Clegg, J. S. (1984). Properties and metabolism of the aqueous cytoplasm and its boundaries. *American Journal of Physiology* **246**(2 Pt 2), R133-51.

Colantuoni, V., Cortese, R., Nilsson, M., Lundvall, J., Bavik, C. O., Eriksson, U., Peterson, P. A. & Sundelin, J. (1985). Cloning and sequencing of a full length cDNA corresponding to human cellular retinol-binding protein. *Biochemical & Biophysical Research Communications* **130**(1), 431-9.

Colberg-Poley, A. M., Voss, S. D., Chowdhury, K. & Gruss, P. (1985). Structural analysis of murine genes containing homeo box sequences and their expression in embryonal carcinoma cells. *Nature* **314**(6013), 713-8.

Collins, M. D., Eckhoff, C., Chahoud, I., Bochert, G. & Nau, H. (1992). 4-Methylpyrazole partially ameliorated the teratogenicity of retinol and reduced the metabolic formation of all-*trans*-retinoic acid in the mouse. *Archives of Toxicology* **66**(9), 652-9.

Connor, M. J. & Smit, M. H. (1987). Terminal-group oxidation of retinol by mouse epidermis. Inhibition in vitro and in vivo. *Biochemical Journal* 244(2), 489-92.

Cowan, S. W., Newcomer, M. E. & Jones, T. A. (1993). Crystallographic studies on a family of cellular lipophilic transport proteins. Refinement of P2 myelin protein and the structure determination and refinement of cellular retinol-binding protein in complex with all-*trans*-retinol. *Journal of Molecular Biology* **230**(4), 1225-46.

De Leenheer, A. P., Lambert, W. E. & Claeys, I. (1982). All-*trans*-retinoic acid: measurement of reference values in human serum by high performance liquid chromatography. *Journal of Lipid Research* 23(9), 1362-7.

De Luca, L. M. (1991). Retinoids and their receptors in differentiation, embryogenesis, and neoplasia. *FASEB Journal* 5(14), 2924-33.

De Luca, L. M. (1993). Multiple mechanisms: the example of vitamin A. Basic Life Sciences 61, 17-25.

DeJonge, M. H. & Zachman, R. D. (1995). The effect of maternal ethanol ingestion on fetal rat heart vitamin A: a model for fetal alcohol syndrome. *Pediatric Research* 37(4 Pt 1), 418-23.

Deltour, L., Ang, H. L. & Duester, G. (1996). Ethanol inhibition of retinoic acid synthesis as a potential mechanism for fetal alcohol syndrome. *FASEB Journal* 10(9), 1050-7.

Dew, S. E. & Ong, D. E. (1994). Specificity of the retinol transporter of the rat small intestine brush border. *Biochemistry* **33**(40), 12340-5.

Dietrich, F. S. & Davis, R. W. (1994). Direct Submission (YM3).

Dockham, P. A., Lee, M. O. & Sladek, N. E. (1992). Identification of human liver aldehyde dehydrogenases that catalyze the oxidation of aldophosphamide and retinaldehyde. *Biochemical Pharmacology* **43**(11), 2453-69.

Dolle, P., Ruberte, E., Kastner, P., Petkovich, M., Stoner, C. M., Gudas, L. J. & Chambon, P. (1989). Differential expression of genes encoding alpha, beta and gamma retinoic acid receptors and CRABP in the developing limbs of the mouse. *Nature* **342**(6250), 702-5.

Donovan, M., Olofsson, B., Gustafson, A. L., Dencker, L. & Eriksson, U. (1995). The cellular retinoic acid binding proteins. *Journal of Steroid Biochemistry & Molecular Biology* **53**(1-6), 459-65.

Dowling, J. E. & Wald, G. (1960). The biological function of vitamin A acid. Proceedings of the National Academy of Sciences (USA) 46, 587-608.

Duester, G. (1991). A hypothetical mechanism for fetal alcohol syndrome involving ethanol inhibition of retinoic acid synthesis at the alcohol dehydrogenase step. *Alcoholism, Clinical & Experimental Research* **15**(3), 568-72.

Duester, G. (1994). Are ethanol-induced birth defects caused by functional retinoic acid deficiency? In *Vitamin A in Health and Disease* (Blomhoff, R., ed.), pp. 343-363. Marcel Dekker, New York.

Duester, G. (1996). Involvement of alcohol dehydrogenase, short-chain dehydrogenase/reductase, aldehyde dehydrogenase, and cytochrome P450 in the control of retinoid signalling by activation of retinoic acid synthesis. *Biochemistry* **35**(38), 12221-7.

Duester, G., Ang, H. L., Deltour, L., Foglio, M. H., Hayamizu, T. F. & Zgombic-Knight, M. (1995). Class I and class IV alcohol dehydrogenase (retinol dehydrogenase) gene expression in mouse embryos. *Advances in Experimental Medicine & Biology* **372**, 301-13.

Dunn, T. J., Koleske, A. J., Lindahl, R. & Pitot, H. C. (1989). Phenobarbital-inducible aldehyde dehydrogenase in the rat. cDNA sequence and regulation of the mRNA by phenobarbital in responsive rats. *Journal of Biological Chemistry* **264**, 13057-13065.

Eckhoff, C. & Nau, H. (1990). Identification and quantitation of all-trans- and 13-cisretinoic acid and 13-cis-4-oxoretinoic acid in human plasma. Journal of Lipid Research 31(8), 1445-54.

Eckhoff, C., Collins, M. D. & Nau, H. (1991). Human plasma all-*trans*-, 13-*cis*- and 13*cis*-4-oxoretinoic acid profiles during subchronic vitamin A supplementation: comparison to retinol and retinyl ester plasma levels. *Journal of Nutrition* **121**(7), 1016-25. el Akawi, Z. & Napoli, J. L. (1994). Rat liver cytosolic retinal dehydrogenase: comparison of 13-cis-, 9-cis-, and all-trans-retinal as substrates and effects of cellular retinoid-binding proteins and retinoic acid on activity. *Biochemistry* **33**(7), 1938-43.

El-Gorab, M. I., Underwood, B. A. & Loerch, J. D. (1975). The roles of bile salts in the uptake of beta-carotene and retinol by rat everted gut sacs. *Biochimica et Biophysica Acta* 401(2), 265-77.

Eskild, W. & Hansson, V. (1994). In *Basic Science and Clinical Aspects of Vitamin A* in *Health and Disease*. (Blomhoff, R., ed.), pp. 531-559. Marcel Dekker, New York.

Evans, R. M. (1988). The steroid and thyroid hormone receptor superfamily. Science 240(4854), 889-95.

Farres, J., Guan, K. L. & Weiner, H. (1989). Primary structure of rat and bovine liver mitochondrial aldehyde dehydrogenases deduced from cDNA sequences. *European Journal of Biochemistry* **180**, 67-74.

Feng D. F.& Doolittle R. F. (1987). Progressive sequence alignment as a prerequisite to correct phylogenetic trees. *Journal of Molecular Evolution* **25**, 351-360.

Fex, G. & Johannesson, G. (1984). Radioimmunological determination of cellular retinol-binding protein in human tissue extracts. *Cancer Research* **44**(7), 3029-32.

Fiorella, P. D. & Napoli, J. L. (1991). Expression of cellular retinoic acid binding protein (CRABP) in Escherichia coli. Characterization and evidence that holo-CRABP is a substrate in retinoic acid metabolism. *Journal of Biological Chemistry* **266**(25), 16572-9.

Fisher, G. J. & Voorhees, J. J. (1996). Molecular mechanisms of retinoid actions in skin. FASEB Journal 10(9), 1002-13.

Furr, H. C., Barua, A. B. & Olson, J. A. (1994). Analytical Methods. In *The Retinoids:* Biology, Chemistry, and Medicine. (Sporn, M. B., Roberts, A. B. & Goodman, D. S., eds.), pp. 179-209. Raven Press, Ltd, New York.

Ghenbot, G. & Weiner, H. (1992). Purification of liver aldehyde dehydrogenase by *p*-hydroxyacetophenone-sepharose affinity matrix and the coelution of chloramphenicol acetyl transferase from the same matrix with recombinantly expressed aldehyde dehydrogenase. *Protein Expression and Purification* **3**, 470-478.

Giguere, V. (1990). Structure and function of the nuclear receptor superfamily for steroid, thyroid hormone and retinoic acid. *Genetic Engineering* **12**, 183-200.

Giguere, V. (1994). Retinoic acid receptors and cellular retinoid binding proteins: complex interplay in retinoid signaling. *Endocrine Reviews* **15**(1), 61-79.

Giguere, V. & Evans, R. M. (1990). Identification of receptors for retinoids as members of the steroid and thyroid hormone receptor family. In *Retinoids. Part A: Molecular and Metabolic Aspects. Methods in Enzymology.* (Packer, L., ed.), Vol. 189, pp. 223-232. Academic Press, San Diego, CA.

Giguere, V., Ong, E. S., Segui, P. & Evans, R. M. (1987). Identification of a receptor for the morphogen retinoic acid. *Nature* **330**(6149), 624-9.

Gjoen, T., Bjerkelund, T., Blomhoff, H. K., Norum, K. R., Berg, T. & Blomhoff, R. (1987). Liver takes up retinol-binding protein from plasma. *Journal of Biological Chemistry* **262**(23), 10926-30.

Godbout, R. (1992). High levels of aldehyde dehydrogenase transcripts in the undifferentiated chick retina. *Exp. Eye Res.* 54, 297-305.

Goedde, H. W. & Agarwal, D. P. (1990). Pharmacogenetics of aldehyde dehydrogenase (ALDH). *Pharmacology & Therapeutics* **45**(3), 345-71.

Gorbunoff, M. J. & Timasheff, S. N. (1984). The interaction of proteins with hydroxyapatite. *Analytical Biochemistry* **136**, 440-445.

Goss, G. D. & McBurney, M. W. (1992). Physiological and clinical aspects of vitamin A and its metabolites. *Critical Reviews in Clinical Laboratory Sciences* **29**(3-4), 185-215.

Graham, C., Hodin, J. & Wistow, G. (1996). A retinaldehyde dehydrogenase as a structural protein in a mammalian eye lens. Gene recruitment of eta-crystallin. *Journal of Biological Chemistry* 271(26), 15623-15628.

Greenfield, N. J. & Pietruszko, R. (1977). Two aldehyde dehydrogenases from human liver. Isolation via affinity chromatography and characterization of the isozymes. *Biochimica et Biophysica Acta* **483**(1), 35-45.

Grummer, M. A. & Zachman, R. D. (1990). The effect of maternal ethanol ingestion on fetal vitamin A in the rat. *Pediatric Research* 28(3), 186-9.

Grummer, M. A. & Zachman, R. D. (1995). Prenatal ethanol consumption alters the expression of cellular retinol binding protein and retinoic acid receptor mRNA in fetal rat embryo and brain. *Alcoholism, Clinical & Experimental Research* **19**(6), 1376-81.

Grummer, M. A., Langhough, R. E. & Zachman, R. D. (1993). Maternal ethanol ingestion effects on fetal rat brain vitamin A as a model for fetal alcohol syndrome. *Alcoholism, Clinical & Experimental Research* 17(3), 592-7.

Guan, K. & Weiner, H. (1990). Sequence of the precursor of bovine liver mitochondrial aldehyde dehydrogenase as determined from its cDNA, its gene, and its functionality. *Archives of Biochemistry and Biophysics* 277, 351-360.

Gudas, L. J. (1994). Retinoids and vertebrate development. Journal of Biological Chemistry 269(22), 15399-402.

Gudas, L. J., Sporn, M. B. & Roberts, A. B. (1994). Cellular biology and biochemistry of the retinoids. In *The Retinoids: Biology, Chemistry, and Medicine.* 2nd edit. (Sporn, M. B., Roberts, A. B. & Goodman, D. S., eds.), pp. 443-520. Raven Press, Ltd, New York.

Guilbault, G. G. (1990). General Aspects of Luminescence Spectroscopy. In *Practical Fluorescence* 2nd Ed edit. (Guilbault, G. G., ed.). Marcel Dekker, Inc, New York.

Han, G., Chang, B., Connor, M. J. & Sidell, N. (1995). Enhanced potency of 9-cis versus all-trans-retinoic acid to induce the differentiation of human neuroblastoma cells. *Differentiation* **59**(1), 61-9.

Harada, S., Agarwal, D. P. & Goedde, H. W. (1980). Electrophoretic and biochemical studies of human aldehyde dehydrogenase isozymes in various tissues. *Life Sciences* 26, 1773-1780.

Harada, S., Agarwal, D. P. & Goedde, H. W. (1981). Aldehyde dehydrogenase deficiency as cause of facial flushing reaction to alcohol in Japanese [letter]. *Lancet* **2**(8253), 982.

Harrington, M. C., Henehan, G. T. & Tipton, K. F. (1987). The roles of human aldehyde dehydrogenase isoenzymes in ethanol metabolism. *Progress in Clinical & Biological Research* 232, 111-25.

Harrison, E. H., Blaner, W. S., Goodman, D. S. & Ross, A. C. (1987). Subcellular localization of retinoids, retinoid-binding proteins, and acyl-CoA:retinol acyltransferase in rat liver. *Journal of Lipid Research* 28(8), 973-81.

Haussler, M., Sidell, N., Kelly, M., Donaldson, C., Altman, A. & Mangelsdorf, D. (1983). Specific high-affinity binding and biologic action of retinoic acid in human neuroblastoma cell lines. *Proceedings of the National Academy of Sciences of the United States of America* **80**(18), 5525-9.

Heim, R. & Strehler, E. E. (1991). Cloning an *E.coli* gene encoding a protein remarkably similar to mammalian aldehyde dehdyrogenases. *Gene* **99**, 15-23.

Helgerud, P., Petersen, L. B. & Norum, K. R. (1982). Acyl CoA:retinol acyltransferase in rat small intestine: its activity and some properties of the enzymic reaction. *Journal of Lipid Research* 23(4), 609-18.

Helgerud, P., Petersen, L. B. & Norum, K. R. (1983). Retinol esterification by microsomes from the mucosa of human small intestine. Evidence for acyl-Coenzyme A retinol acyltransferase activity. *Journal of Clinical Investigation* **71**(3), 747-53.

Heller, J. (1975). Interactions of plasma retinol-binding protein with its receptor. Specific binding of bovine and human retinol-binding protein to pigment epithelium cells from bovine eyes. *Journal of Biological Chemistry* **250**(10), 3613-9.

Hempel, J., Nicholas, H. & Lindahl, R. (1993). Aldehyde dehydrogenases: widespread structural and functional diversity within a shared framework. *Protein Science* 2(11), 1890-900.

Herr, F. M. & Ong, D. E. (1992). Differential interaction of lecithin-retinol acyltransferase with cellular retinol binding proteins. *Biochemistry* **31**(29), 6748-55.

Heyman, R. A., Mangelsdorf, D. J., Dyck, J. A., Stein, R. B., Eichele, G., Evans, R. M. & Thaller, C. (1992). 9-cis retinoic acid is a high affinity ligand for the retinoid X receptor. Cell 68(2), 397-406.

Hirano, H., Komatsu, S., Takakura, H., Sakiyama, F. & Tsunasawa, S. (1992). Deblocking and Subsequent Microsequence Analysis of N^{α} -Blocked Proteins Electroblotted onto PVDF Membrane. *Journal of Biochemistry* 111, 754-757.

Hoffman, C. & Eichele, G. (1994). Retinoids in Development. In *The Retinoids: Biology, Chemistry, and Medicine.* 2nd Ed edit. (Sporn, M. B., Roberts, A. B. & Goodman, D. S., eds.), pp. 387-441. Raven Press, Ltd, New York.

Hollander, D. & Ruble, P. E., Jr. (1978). beta-carotene intestinal absorption: bile, fatty acid, pH, and flow rate effects on transport. *American Journal of Physiology* 235(6), E686-91.

Hopkins, F. G. (1906). The analyst and the medical man. Analyst 31, 385-404.

Hsu, L. C. & Chang, W. C. (1991). Cloning and characterization of a new functional human aldehyde dehydrogenase gene. *Journal of Biological Chemistry* **266**(19), 12257-65.

Hsu, L. C., Bendel, R. E. & Yoshida, A. (1988). Genomic structure of the human mitochondrial aldehyde dehydrogenase gene. *Genomics* 2(1), 57-65.

Hsu, L. C., Chang, W. C., Hiraoka, L. & Hsieh, C. L. (1994a). Molecular cloning, genome organisation, and chromosomal localisation of an additional human aldehyde dehydrogenase gene, ALDH6. *Genomics* **24**(2), 333-341.

Hsu, L. C., Chang, W. C., Lin, S. W. & Yoshida, A. (1995). Cloning and characterization of genes encoding four additional human aldehyde dehydrogenase isozymes. *Advances in Experimental Medicine & Biology* **372**, 159-68.

Hsu, L. C., Chang, W.-C. & Yoshida, A. (1989). Genomic structure of the human cytosolic aldehyde dehydrogenase gene. *Genomics* 5, 857-865.

Hsu, L. C., Chang, W. C. & Yoshida, A. (1994b). Cloning of a cDNA encoding human ALDH7, a new member of the aldehyde dehydrogenase family. *Gene* **151**(1-2), 285-9.

Huang, D. Y. & Ichikawa, Y. (1994). Two different enzymes are primarily responsible for retinoic acid synthesis in rabbit liver cytosol. *Biochemical & Biophysical Research Communications* 205(2), 1278-83.

Hupert, J., Mobarhan, S., Papa, V. M., & Lucchesi, D. J. (1991). In vitro formation of retinoic acid from retinal in rat liver. *Biochemistry and Cell Biology* **69**(8), 509-514.

Hussain, M. M., Mahley, R. W., Boyles, J. K., Fainaru, M., Brecht, W. J. & Lindquist, P. A. (1989). Chylomicron-chylomicron remnant clearance by liver and bone marrow in rabbits. Factors that modify tissue-specific uptake. *Journal of Biological Chemistry* **264**(16), 9571-82.

Ikawa, M., Impraim, C. C., Wang, G. & Yoshida, A. (1983). Isolation and characterization of aldehyde dehydrogenase isozymes from usual and atypical human livers. *Journal of Biological Chemistry* **258**(10), 6282-7.

Inoue, H., Nojima, H. & Okayama, H. (1990). High efficiency transformation of *E.coli* with plasmids. *Gene* 96, 23-28.

Ishikawa, T., Umesono, K., Mangelsdorf, D. J., Aburatani, H., Stanger, B. Z., Shibasaki, Y., Imawari, M., Evans, R. M. & Takaku, F. (1990). A functional retinoic acid receptor encoded by the gene on human chromosome 12. *Molecular Endocrinology* **4**(6), 837-44.

Izant, J. G. & Weintraub, H. (1984). Inhibition of thymidine kinase gene expression by anti-sense RNA: a molecular approach to genetic analysis. *Cell* **36**, 1007-15.

Jakoby, W. B. & Ziegler, D. M. (1990). The enzymes of detoxification. Journal of Biological Chemistry 265(34), 20715-8.

Jamison, R. S., Newcomer, M. E. & Ong, D. E. (1994). Cellular retinoid-binding proteins: limited proteolysis reveals a conformational change upon ligand binding. *Biochemistry* **33**(10), 2873-9.

Johansson, J., von Bahr-Lindstrom, H., Jeck, R., Woenckhaus, C. & Jornvall, H. (1988). Mitochondrial aldehyde dehydrogenase from horse liver. Correlations of the same species variants for both the cytosolic and mitochondrial forms of an enzyme. *European Journal of Biochemistry* **172**, 527-533.

Jones, D. E., Brennan, M. D., Hempel, J. & Lindahl, R. (1988). Proceedings of the New York Academy of Sciences (USA). 85, 1782-1786.

Jones, K. L., Smith, D. W. & Ulleland, C. N. (1973). Pattern of malformation in offspring of chronic alcoholic mothers. *Lancet* 1, 1267-1271.

Jornvall, H., Danielsson, O., Hjelmqvist, L., Persson, B. & Shafqat, J. (1995). The Alcohol Dehydrogenase System. Advances in Experimental Medicine and Biology 372, 281-294.

Julia, P., Farres, J. & Pares, X. (1986). Ocular alcohol dehydrogenase in the rat: regional distribution and kinetics of the ADH-1 isoenzyme with retinol and retinal. *Experimental Eye Research* **42**(4), 305-14.

Kakkad, B. P. & Ong, D. E. (1988). Reduction of retinaldehyde bound to cellular retinol-binding protein (type II) by microsomes from rat small intestine. *Journal of Biological Chemistry* **263**(26), 12916-9.

Kanai, M., Raz, A. & Goodman, D. S. (1968). Retinol-binding protein: the transport protein for vitamin A in human plasma. *Journal of Clinical Investigation* **47**(9), 2025-44.

Kane, C. D. & Bernlohr, D. A. (1996). A simple assay for intracellular lipid-binding proteins using displacement of 1-anilinonaphthalene 8-sulfonic acid. *Analytical Biochemistry* 233(2), 197-204.

Karrer, P., Morf, R. & Schoepp, K. (1931). Zur Kenntins des Vitamins A in Fischtranen. Helv. Chim. Acta. 14, 1431-1436.

Kastner, P., Chambon, P. & Leid, M. (1994). In Vitamin A in Health and Disease (Blomhoff, R., ed.), pp. 189-238. Marcel Dekker, Inc, New York.

Kato, M., Blaner, W. S., Mertz, J. R., Das, K., Kato, K. & Goodman, D. S. (1985). Influence of retinoid nutritional status on cellular retinol- and cellular retinoic acidbinding protein concentrations in various rat tissues. *Journal of Biological Chemistry* **260**(8), 4832-8.

Kedishvili, N. Y., Bosron, W. F., Stone, C. L., Hurley, T. D., Peggs, C. F., Thomasson, H. R., Popov, K. M., Carr, L. G., Edenberg, H. J. & Li, T. K. (1995). Expression and kinetic characterization of recombinant human stomach alcohol dehydrogenase. Activesite amino acid sequence explains substrate specificity compared with liver isozymes. *Journal of Biological Chemistry* **270**(8), 3625-30.

Kedishvili, N. Y., Gough, W. H., Chernoff, E. A. G., Hurley, T. D., Stone, C. L., Bowman, K. D., Popov, K. M., Bosron, W. F. & Li, T. K. (1997a). Cdna Sequence and Catalytic Properties Of a Chick Embryo Alcohol Dehydrogenase That Oxidizes Retinol and 3-Beta,5-Alpha-Hydroxysteroids. *Journal of Biological Chemistry* 272(11), 7494-7500.

Kedishvili, N. Y., Stone, C. L., Popov, K. M. & Chernoff, E. A. (1997b). Role of alcohol dehydrogenases in steroid and retinoid metabolism. *Advances in Experimental Medicine & Biology* **414**, 321-9.

Khachik, F., Beecher, G. R., Goli, M. B. & Lusby, W. R. (1992). Separation and quantitation of carotenoids in foods. *Methods in Enzymology* **213**, 347-359.

Kim, C. I., Leo, M. A. & Lieber, C. S. (1992). Retinol forms retinoic acid via retinal. Archives of Biochemistry & Biophysics 294(2), 388-93. Kim, S. K. & Wold, B. J. (1985). Stable reduction of thymidine kinase activity in cells expressing high levels of anti-sense RNA. *Cell* **42**, 129-38.

Kimura, S., Donovan, J. C. & Nebert, D. W. (1987). Expression of the mouse P(1)450 gene during differentiation without foreign chemical stimulation. *Journal of Experimental Pathology* **3**(1), 61-74.

Kishore, G. S. & Boutwell, R. K. (1980). Enzymatic oxidation and reduction of retinal by mouse epidermis. *Biochemical & Biophysical Research Communications* **94**(4), 1381-6.

Kitson, K. E., Blythe, T. J. & Kitson, T. M. (1995). Studies of the esterase activity of cytosolic aldehyde dehydrogenase using sterically hindered and cyclic substrates. *Advances in Experimental Medicine and Biology* **372**, 45-52.

Kizaki, M., Ikeda, Y., Tanosaki, R., Nakajima, H., Morikawa, M., Sakashita, A. & Koeffler, H. P. (1993). Effects of novel retinoic acid compound, 9-*cis*-retinoic acid, on proliferation, differentiation, and expression of retinoic acid receptor-alpha and retinoid X receptor-alpha RNA by HL-60 cells. *Blood* **82**(12), 3592-9.

Kliewer, S. A., Umesono, K., Mangelsdorf, D. J. & Evans, R. M. (1992). Retinoid X receptor interacts with nuclear receptors in retinoic acid, thyroid hormone and vitamin D3 signalling. *Nature* **355**(6359), 446-9.

Kluess, C., Scherübl, H., Hescheler, J., Schultz, G. & Wittig, B. (1993). Selectivity in signal transduction determined by gamma subunits of heterotrimeric G proteins. *Science* **259**, 832-34.

Klyosov, A. A. (1996). Kinetics and specificity of human liver aldehyde dehydrogenases toward aliphatic, aromatic, and fused polycyclic aldehydes. *Biochemistry* **35**(14), 4457-67.

Klyosov, A. A., Rashkovetsky, L. G., Tahir, M. K. & Keung, W. M. (1996). Possible role of liver cytosolic and mitochondrial aldehyde dehydrogenases in acetaldehyde metabolism. *Biochemistry* **35**(14), 4445-56.

Krust, A., Kastner, P., Petkovich, M., Zelent, A. & Chambon, P. (1989). A third human retinoic acid receptor, hRAR-gamma. *Proceedings of the National Academy of Sciences of the United States of America* **86**(14), 5310-4.

Labrecque, J., Bhat, P. V. & Lacroix, A. (1993). Purification and partial characterization of a rat kidney aldehyde dehydrogenase that oxidizes retinal to retinoic acid. *Biochemistry & Cell Biology* 71(1-2), 85-9.

Labrecque, J., Bhat, P. V. & Lacroix, A. (1997). Regulation of synthesis of all-trans and 9-cis retinoic acids - role of a new retinal dehydrogenase. [French]. *M S-Medecine Sciences* 13(3), 345-352.

Labrecque, J., Dumas, F., Lacroix, A. & Bhat, P. V. (1995). A novel isoenzyme of aldehyde dehydrogenase specifically involved in the biosynthesis of 9-cis and all-trans retinoic acid. *Biochemical Journal* **305**(Pt 2), 681-4.

Lampron, C., Rochette-Egly, C., Gorry, P., Dolle, P., Mark, M., Lufkin, T., LeMeur, M. & Chambon, P. (1995). Mice deficient in cellular retinoic acid binding protein II (CRABPII) or in both CRABPI and CRABPII are essentially normal. *Development* **121**(2), 539-48.

Landers, G. M. (1990). High-performance liquid chromatography of retinoid isomers. Methods in Enzymology 189, 70-80.

Landers, G. M. & Olson, J. A. (1988). Rapid, simultaneous determination of isomers of retinal, retinal oxime and retinol by high-performance liquid chromatography. *Journal of Chromatography* **438**(2), 383-92.

Leatherbarrow, R. J. (1987). Enzfitter Manual, Elsevier Science Publishers, Amsterdam.

Leblanc, B. P. & Stunnenberg, H. G. (1995). 9-cis retinoic acid signalling: changing partners causes some excitement. Genes & Development 9(15), 1811-6.

Lee, M. O., Manthey, C. L. & Sladek, N. E. (1991). Identification of mouse liver aldehyde dehydrogenases that catalyze the oxidation of retinaldehyde to retinoic acid. *Biochemical Pharmacology* **42**(6), 1279-85.

Lehman, E. D., Spivey, H. O., Thayer, R. H. & Nelson, E. C. (1972). The binding of retinoic acid to serum albumin in plasma. *Fed. Proc. (Fed. Am. Soc. Exp.Biol.)* 31, A672.

Leid, M., Kastner, P. & Chambon, P. (1992a). Multiplicity generates diversity in the retinoic acid signalling pathways. *Trends in Biochemical Sciences* 17(10), 427-33.

Leid, M., Kastner, P., Lyons, R., Nakshatri, H., Saunders, M., Zacharewski, T., Chen, J. Y., Staub, A., Garnier, J. M., Mader, S. & *et al.* (1992b). Purification, cloning, and RXR identity of the HeLa cell factor with which RAR or TR heterodimerizes to bind target sequences efficiently [published erratum appears in Cell 1992 Nov 27;71(5):following 886]. *Cell* 68(2), 377-95.

Lemoine, P., Harrousseau, H., Borteyru, J. P. & Menuet, J. (1968). Les enfants de parents alcooliques: Anomalies observees. A propos de 127 cas. *Quest Medical* 25, 476-482.

Leo, M. A., Kim, C. I. & Lieber, C. S. (1987). NAD⁺-dependent retinol dehydrogenase in liver microsomes. *Archives of Biochemistry & Biophysics* **259**(2), 241-9.

Leo, M. A., Kim, C. I., Lowe, N. & Lieber, C. S. (1989). Increased hepatic retinal dehydrogenase activity after phenobarbital and ethanol administration. *Biochemical Pharmacology* **38**(1), 97-103.

Levin, A. A., Sturzenbecker, L. J., Kazmer, S., Bosakowski, T., Huselton, C., Allenby, G., Speck, J., Kratzeisen, C., Rosenberger, M., Lovey, A. & *et al.* (1992a). 9-*cis* retinoic acid stereoisomer binds and activates the nuclear receptor RXR alpha. *Nature* **355**(6358), 359-61.

Levin, A. A., Sturzenbecker, L. J., Kazmer, S., Bosakowski, T., Huselton, C., Allenby, G., Speck, J., Kratzeisen, C., Rosenberger, M., Lovey, A. & et al. (1992b). A new pathway for vitamin A. Understanding the pleiotropic effects of retinoids. *Annals of the New York Academy of Sciences* 669, 70-85; discussion 85-6.

Levin, M. S., Li, E., Ong, D. E. & Gordon, J. I. (1987). Comparison of the tissuespecific expression and developmental regulation of two closely linked rodent genes encoding cytosolic retinol-binding proteins. Journal of Biological Chemistry 262(15), 7118-24.

Levin, M. S., Locke, B., Yang, N. C., Li, E. & Gordon, J. I. (1988). Comparison of the ligand binding properties of two homologous rat apocellular retinol-binding proteins expressed in Escherichia coli. *Journal of Biological Chemistry* **263**(33), 17715-23.

Li, E. & Norris, A. W. (1996). Structure/function of cytoplasmic vitamin A-binding proteins. *Annual Review of Nutrition* 16, 205-34.

Li, E., Demmer, L. A., Sweetser, D. A., Ong, D. E. & Gordon, J. I. (1986). Rat cellular retinol-binding protein II: use of a cloned cDNA to define its primary structure, tissue-specific expression, and developmental regulation. *Proceedings of the National Academy of Sciences of the United States of America* **83**(16), 5779-83.

Li, E., Qian, S. J., Winter, N. S., d'Avignon, A., Levin, M. S. & Gordon, J. I. (1991). Fluorine nuclear magnetic resonance analysis of the ligand binding properties of two homologous rat cellular retinol-binding proteins expressed in Escherichia coli. *Journal of Biological Chemistry* **266**(6), 3622-9.

Lieber, C. S. (1991). Alcohol, liver, and nutrition. Journal of American Colleges on Nutrition. 10, 602-632.

Lin, S. W., Chen, J. C., Hsu, L. C., Hsieh, C. L. & Yoshida, A. (1996). Human gammaaminobutyraldehyde dehydrogenase ALDH9: cDNA sequence, genomic organisation, polymorphism, chromosomal localisation, and tissue expression. *Genomics* **34**, 376-380.

Lindahl, R. & Evces, S. (1984). Comparative subcellular distribution of aldehyde dehydrogenase in rat, mouse and rabbit liver. *Biochemical Pharmacology* **33**(21), 3383-9.

Lindahl, R. & Hempel, J. (1991). Aldehyde dehydrogenases: what can be learned from a baker's dozen sequences? Advances in Experimental Medicine & Biology 284, 1-8.

Lotan, R. (1995). Retinoids and apoptosis: implications for cancer chemoprevention and therapy [editorial; comment]. *Journal of the National Cancer Institute* **87**(22), 1655-7.

Lovat, P. E., Lowis, S. P., Pearson, A. D., Malcolm, A. J. & Redfern, C. P. (1994). Concentration-dependent effects of 9-cis retinoic acid on neuroblastoma differentiation and proliferation *in vitro*. *Neuroscience Letters* **182**(1), 29-32.

Luke, B. (1990). The metabolic basis of the fetal alcohol syndrome. International Journal of Fertility 35(6), 333-7.

Lunin, G. (1881). Ueber die Bedeutung der anorganischen Salze bei der Ernahrung des Thieres. Hoppe-Seyler's Z. physiol. Chem 5, 311-326.

MacDonald, P. N. & Ong, D. E. (1987). Binding specificities of cellular retinol-binding protein and cellular retinol-binding protein, type II. *Journal of Biological Chemistry* **262**(22), 10550-6.

MacGibbon, A. K. H., Motion, R. L., Crow, K. E., Buckley, P. D. & Blackwell, L. F. (1979). Purification and properties of sheep-liver aldehyde dehydrogenases. *European Journal of Biochemistry* **96**, 585-595.

Maden, M. (1994). Vitamin A in embryonic development. *Nutrition Reviews* 52(2 Pt 2), S3-12.

Maden, M. & Holder, N. (1991). The involvement of retinoic acid in the development of the vertebrate central nervous system. *Development* Suppl(2), 87-94.

Maden, M., Ong, D. E., Summerbell, D. & Chytil, F. (1988). Spatial distribution of cellular protein binding to retinoic acid in the chick limb bud. *Nature* **335**(6192), 733-5.

Mangelsdorf, D. J., Ong, E. S., Dyck, J. A. & Evans, R. M. (1990). Nuclear receptor that identifies a novel retinoic acid response pathway. *Nature* **345**(6272), 224-9.

Mangelsdorf, D. J., Thummel, C., Beato, M., Herrlich, P., Schutz, G., Umesono, K., Blumberg, B., Kastner, P., Mark, M., Chambon, P. & et al. (1995). The nuclear receptor superfamily: the second decade. *Cell* **83**(6), 835-9.

Mangelsdorf, D. J., Umesono, K. & Evans, R. M. (1994). The Retinoid Receptors. In *The Retinoids: Biology, Chemistry, and Medicine* 2nd Ed edit. (Sporn, M. B., Roberts, A. B. & Goodman, D. S., eds.), pp. 319-349. Raven Press, Ltd, New York.

Mangelsdorf, D. J., Umesono, K., Kliewer, S. A., Borgmeyer, U., Ong, E. S. & Evans, R. M. (1991). A direct repeat in the cellular retinol-binding protein type II gene confers differential regulation by RXR and RAR. *Cell* 66(3), 555-61.

Manthey, C. L., Landkamer, G. J. & Sladek, N. E. (1990). Identification of the mouse aldehyde dehydrogenases important in aldophosphamide detoxification. *Cancer Research* **50**(16), 4991-5002.

Marsh-Armstrong, N., McCaffery, P., Gilbert, W., Dowling, J. E. & Dräger, U. C. (1994). Retinoic acid is necessary for development of the ventral retina in zebrafish. *Proceedings of the National Academy of Sciences of the United States of America* **91**(15), 7286-90.

Martini, R. & Murray, M. (1994). Retinal dehydrogenation and retinoic acid 4hydroxylation in rat hepatic microsomes: developmental studies and effect of foreign compounds on the activities. *Biochemical Pharmacology* **47**(5), 905-9.

McBurney, M. W., Costa, S. & Pratt, M. A. (1993). Retinoids and cancer: a basis for differentiation therapy. *Cancer Investigation* 11(5), 590-8.

McCaffery, P. & Dräger, U. C. (1993). Retinoic acid synthesis in the developing retina. Advances in Experimental Medicine & Biology **328**, 181-90.

McCaffery, P. & Dräger, U. C. (1994). High levels of a retinoic acid-generating dehydrogenase in the meso-telencephalic dopamine system. *Proceedings of the National Academy of Sciences of the United States of America* **91**(16), 7772-6.

McCaffery, P. & Dräger, U. C. (1995). Retinoic acid synthesizing enzymes in the embryonic and adult vertebrate. *Advances in Experimental Medicine & Biology* 372, 173-83.

McCaffery, P., Lee, M. O., Wagner, M. A., Sladek, N. E. & Dräger, U. C. (1992). Asymmetrical retinoic acid synthesis in the dorsoventral axis of the retina. *Development* 115(2), 371-82.

McCollum, E. V. & Davis, M. (1913). The necessity of certain lipids during growth. Journal of Biological Chemistry 15, 167-175.

McCormick, A. M., Napoli, J. L., Schnoes, H. K. & DeLuca, H. F. (1978). Isolation and identification of 5, 6-epoxyretinoic acid: a biologically active metabolite of retinoic acid. *Biochemistry* **17**(19), 4085-90.

Mertz, J. R., Shang, E. Y., Piantedosi, R., Wei, S. H., Wolgemuth, D. J. & Blaner, W. S. (1997). Identification and characterisation of a stereospecific human enzyme that catalyses 9-cis-retinol oxidation - a possible role in 9-cis-retinoic acid formulation. Journal of Biological Chemistry 272(18), 11744-9.

Mitchell, D. Y. & Petersen, D. R. (1987). The oxidation of alpha-beta unsaturated aldehyde products of lipid peroxidation by rat liver aldehyde dehydrogenases. *Toxicology* and Applied Pharmacology 87(3), 403-410.

Miyauchi, M., Masaki, R., Taketawi, S., Yamamoto, A., Akayama, M. & Tashiro, V. (1991). Molecular cloning, sequencing, and expression of cDNA for rat liver microsomal aldehyde dehydrogenase. *Journal of Biological Chemistry* **266**(29), 19536-19542.

Moore, T. (1930). Vitamin A and carotene. Biochemical Journal 24, 692-702.

Morriss-Kay, G. (1993). Retinoic acid and craniofacial development: molecules and morphogenesis. *Bioessays* 15(1), 9-15.

Morriss-Kay, G. M. & Sokolova, N. (1996). Embryonic development and pattern formation. *FASEB Journal* 10(9), 961-8.

Naar, A. M., Boutin, J. M., Lipkin, S. M., Yu, V. C., Holloway, J. M., Glass, C. K. & Rosenfeld, M. G. (1991). The orientation and spacing of core DNA-binding motifs dictate selective transcriptional responses to three nuclear receptors. *Cell* **65**(7), 1267-79.

Nagao, A. & Olson, J. A. (1994). Enzymatic formation of 9-cis, 13-cis, and all-trans retinals from isomers of beta-carotene [published erratum appears in FASEB J 1995 May;9(8):691]. FASEB Journal 8(12), 968-73.

Napoli, J. L. (1986). Quantification of physiological levels of retinoic acid. Methods in Enzymology 123, 112-24.

Napoli, J. L. (1993). Biosynthesis and metabolism of retinoic acid: roles of CRBP and CRABP in retinoic acid: roles of CRBP and CRABP in retinoic acid homeostasis. *Journal of Nutrition* **123**(2 Suppl), 362-6.

Napoli, J. L. (1996a). Biochemical pathways of retinoid transport, metabolism, and signal transduction. *Clinical Immunology & Immunopathology* **80**(3 Pt 2), S52-62.

Napoli, J. L. (1996b). Retinoic acid biosynthesis and metabolism. *FASEB Journal* 10(9), 993-1001.

Napoli, J. L. & Race, K. R. (1987). The biosynthesis of retinoic acid from retinol by rat tissues in vitro. Archives of Biochemistry & Biophysics 255(1), 95-101.

Napoli, J. L. & Race, K. R. (1988). Biogenesis of retinoic acid from beta-carotene. Differences between the metabolism of beta-carotene and retinal. *Journal of Biological Chemistry* **263**(33), 17372-7.

Napoli, J. L. & Race, K. R. (1990). Microsomes convert retinol and retinal into retinoic acid and interfere in the conversions catalyzed by cytosol. *Biochimica et Biophysica Acta* 1034(2), 228-32.

Napoli, J. L., Boerman, M. H., Chai, X., Zhai, Y. & Fiorella, P. D. (1995). Enzymes and binding proteins affecting retinoic acid concentrations. *Journal of Steroid Biochemistry* & *Molecular Biology* 53(1-6), 497-502.

Napoli, J. L., Posch, K. C. & Burns, R. D. (1992). Microsomal retinal synthesis: retinol vs. holo-CRBP as substrate and evaluation of NADP, NAD and NADPH as cofactors. *Biochimica et Biophysica Acta* **1120**(2), 183-6.

Napoli, J. L., Posch, K. P., Fiorella, P. D. & Boerman, M. H. (1991). Physiological occurrence, biosynthesis and metabolism of retinoic acid: evidence for roles of cellular retinol-binding protein (CRBP) and cellular retinoic acid-binding protein (CRABP) in the pathway of retinoic acid homeostasis. *Biomedicine & Pharmacotherapy* **45**(4-5), 131-43.

Nebert, E. W., Kozak, C. A., Lindahl, R. & Vasiliou, V. (1994). Mouse microsomal class 3 aldehyde dehydrogenase: AHD3 cDNA sequence, chromosomal mapping, and dioxin inducibility. *Toxicologist* 14, 410.

Nelson, D. R., Kamataki, T., Waxman, D. J., Guengerich, F. P., Estabrook, R. W., Feyereisen, R., Gonzalez, F. J., Coon, M. J., Gunsalus, I. C., Gotoh, O. & al, e. (1993). The P450 Superfamily: update on new sequences, gene mapping, accession numbers, early trivial names of enzymes, and nomenclature. *DNA and Cell Biology* **12**(1), 1-51.

Newcomer, M. E. (1995). Retinoid-binding proteins: structural determinants important for function. *FASEB Journal* 9(2), 229-39.

Newcomer, M. E., Jones, T. A., Aqvist, J., Sundelin, J., Eriksson, U., Rask, L. & Peterson, P. A. (1984a). The three-dimensional structure of retinol-binding protein. *EMBO Journal* 3(7), 1451-4.

Newcomer, M. E., Liljas, A., Eriksson, U., Sundelin, J., Rask, L. & Peterson, P. A. (1981). Crystallization of and preliminary X-ray data for an intracellular vitamin Abinding protein from rat liver. *Journal of Biological Chemistry* **256**(15), 8162-3.

Newcomer, M. E., Liljas, A., Sundelin, J., Rask, L. & Peterson, P. A. (1984b). Crystallization of and preliminary X-ray data for the plasma retinol-binding protein. Journal of Biological Chemistry 259(8), 5230-1.

Nicholson, R. C., Mader, S., Nagpal, S., Leid, M., Rochette-Egly, C. & Chambon, P. (1990). Negative regulation of the rat stromelysin gene promoter by retinoic acid is mediated by an AP1 binding site. *EMBO Journal* 9(13), 4443-54.

Noll, G. N. (1996). High-performance liquid chromatographic analysis of retinal and retinol isomers. *Journal of Chromatography A* 721(2), 247-59.

Norris, A. W., Cheng, L., Giguere, V., Rosenberger, M. & Li, E. (1994). Measurement of subnanomolar retinoic acid binding affinities for cellular retinoic acid binding proteins by fluorometric titration. *Biochimica et Biophysica Acta* **1209**(1), 10-8.

Noy, N. (1992a). The ionization behavior of retinoic acid in aqueous environments and bound to serum albumin. *Biochimica et Biophysica Acta* **1106**(1), 151-8.

Noy, N. (1992b). The ionization behavior of retinoic acid in lipid bilayers and in membranes. *Biochimica et Biophysica Acta* **1106**(1), 159-64.

Ong, D. E. (1982). Purification and partial characterization of cellular retinol-binding protein from human liver. *Cancer Research* 42(3), 1033-7.

Ong, D. E. (1984). A novel retinol-binding protein from rat. Purification and partial characterization. *Journal of Biological Chemistry* **259**(3), 1476-82.

Ong, D. E. (1987). Cellular retinoid-binding proteins. Archives of Dermatology 123(12), 1693-1695a.

Ong, D. E. (1994). Cellular transport and metabolism of vitamin A: roles of the cellular retinoid-binding proteins. *Nutrition Reviews* **52**(2 Pt 2), S24-31.

Ong, D. E. & Chytil, F. (1975). Specificity of cellular retinol-binding protein for compounds with vitamin A activity. *Nature* **255**(5503), 74-5.

Ong, D. E. & Chytil, F. (1978). Cellular retinol-binding protein from rat liver. Purification and characterization. *Journal of Biological Chemistry* **253**(3), 828-32.

Ong, D. E. & Page, D. L. (1986). Quantitation of cellular retinol-binding protein in human organs. *American Journal of Clinical Nutrition* **44**(3), 425-30.

Ong, D. E. & Page, D. L. (1987). Cellular retinol-binding protein (type two) is abundant in human small intestine. *Journal of Lipid Research* **28**(6), 739-45.

Ong, D. E., Kakkad, B. & MacDonald, P. N. (1987). Acyl-CoA-independent esterification of retinol bound to cellular retinol-binding protein (type II) by microsomes

from rat small intestine [published erratum appears in J Biol Chem 1987 May 25;262(15):7438]. Journal of Biological Chemistry 262(6), 2729-36.

Ong, D. E., Lucas, P. C., Kakkad, B. & Quick, T. C. (1991). Ontogeny of two vitamin A-metabolizing enzymes and two retinol-binding proteins present in the small intestine of the rat. *Journal of Lipid Research* **32**(9), 1521-7.

Ong, D. E., MacDonald, P. N. & Gubitosi, A. M. (1988). Esterification of retinol in rat liver. Possible participation by cellular retinol-binding protein and cellular retinol-binding protein II. *Journal of Biological Chemistry* **263**(12), 5789-96.

Ong, D. E., Newcomer, M. E. & Chytil, F. (1994). Cellular Retinoid-Binding Proteins. In *The Retinoids: Biology, Chemistry, and Medicine*. (Sporn, M. B., Roberts, A. B. & Goodman, D. S., eds.), pp. 283-317. Raven Press, Ltd., New York.

Ong, D. E., Tsai, C. H. & Chytil, F. (1976). Cellular retinol-binding protein and retinoic acid-binding protein in rat testes: effect of retinol depletion. *Journal of Nutrition* **106**(2), 204-11.

Osborne, T. B. & Mendel, L. B. (1913). The relation of growth to the chemical constituents of the diet. *Journal of Biological Chemistry* **15**, 311-326.

Osborne, T. B. & Mendel, L. B. (1914). The influence of cod liver oil and some other fats on growth. *Journal of Biological Chemistry* 17, 401-408.

Ottonello, S., Scita, G., Mantovani, G., Cavazzini, D. & Rossi, G. L. (1993). Retinol bound to cellular retinol-binding protein is a substrate for cytosolic retinoic acid synthesis. *Journal of Biological Chemistry* **268**(36), 27133-42.

Ozaki, K., Terakita, A., Ozaki, M., Hara, R., Hara, T., Hara-Nishimura, I., Mori, H. & Nishimura, M. (1994). Molecular characterization and functional expression of squid retinal-binding protein. A novel species of hydrophobic ligand-binding protein. *Journal of Biological Chemistry* **269**(5), 3838-45.

Pahlman, S., Ruusala, A.-I., Abrahamsson, L., Mattsson, M. E. K. & Esscher, T. (1984). Retinoic acid-induced differentiation of cultured human neuroblastoma cells: comparison with phorbolester-induced differentiation. *Cell Differentiation* 14, 135-144.

Parrilla, R., Okhawa, K., Lindros, K. O., Zimmerman, U.-I. P., Kobayashi, K. & Williamson, J. R. (1974). Functional compartmentation of acetaldehyde oxidation in rat liver. *Journal of Biological Chemistry* **249**, 4926.

Penzes, P., Wang, X. S., Sperkova, Z. & Napoli, J. L. (1997). Cloning of a rat cDNA encoding retinal dehydrogenase isozyme type I and its expression in E.coli. *Gene* **191**(2), 167-172.

Persson, B., Krook, M. & Jornvall, H. (1995). Short-chain dehydrogenases/reductases. Advances in Experimental Medicine & Biology 372, 383-95.

Petkovich, M., Brand, N. J., Krust, A. & Chambon, P. (1987). A human retinoic acid receptor which belongs to the family of nuclear receptors. *Nature* **330**(6147), 444-50.

Pietruszko, R. (1983). Aldehyde dehydrogenase isozymes. Isozymes: Current Topics in Biological Medical Research 8, 195-217.

Pietruszko, R. (1993). Direct Submission (h4). SWISSPROT accession P47895.

Porter, S. B., Fraker, L. D., Chytil, F. & Ong, D. E. (1983). Localization of cellular retinol-binding protein in several rat tissues. *Proceedings of the National Academy of Sciences of the United States of America* **80**(21), 6586-90.

Posch, K. C. & Napoli, J. L. (1992). Multiple retinoid dehydrogenases in testes cytosol from alcohol dehydrogenase negative or positive deermice. *Biochemical Pharmacology* **43**(10), 2296-8.

Posch, K. C., Boerman, M. H., Burns, R. D. & Napoli, J. L. (1991). Holocellular retinol binding protein as a substrate for microsomal retinal synthesis. *Biochemistry* **30**(25), 6224-30.

Posch, K. C., Burns, R. D. & Napoli, J. L. (1992). Biosynthesis of all-*trans*-retinoic acid from retinal. Recognition of retinal bound to cellular retinol binding protein (type I) as substrate by a purified cytosolic dehydrogenase. *Journal of Biological Chemistry* **267**(27), 19676-82.

Promega. (1996). DNA Purification. In Protocols and applications guide 3rd edit. (Doyle, K., ed.), pp. 80-82, U.S.A.

Pullarkat, R. K. (1991). Hypothesis: prenatal ethanol-induced birth defects and retinoic acid. *Alcoholism, Clinical & Experimental Research* 15(3), 565-7.

Quick, T. C. & Ong, D. E. (1990). Vitamin A metabolism in the human intestinal Caco-2 cell line. *Biochemistry* **29**(50), 11116-23.

Rajan, N., Blaner, W. S., Soprano, D. R., Suhara, A. & Goodman, D. S. (1990). Cellular retinol-binding protein messenger RNA levels in normal and retinoid-deficient rats. *Journal of Lipid Research* **31**(5), 821-9.

Raner, G. M., Vaz, A. D. & Coon, M. J. (1996). Metabolism of all-*trans*, 9-cis, and 13cis isomers of retinal by purified isozymes of microsomal cytochrome P450 and mechanism-based inhibition of retinoid oxidation by citral. *Molecular Pharmacology* **49**(3), 515-22.

Rashkovetsky, L. G., Maret, W. & Klyosov, A. A. (1994). Human liver aldehyde dehydrogenases: new method of purification of the major mitochondrial and cytosolic enzymes and re-evaluation of their kinetic properties. *Biochimica et Biophysica Acta* **1205**(2), 301-7.

Rask, L. & Peterson, P. A. (1976). *In vitro* uptake of vitamin A from the retinol-binding plasma protein to mucosal epithelial cells from the monkey's small intestine. *Journal of Biological Chemistry* **251**(20), 6360-6.

Redfern, C. P., Lovat, P. E., Malcolm, A. J. & Pearson, A. D. (1995). Gene expression and neuroblastoma cell differentiation in response to retinoic acid: differential effects of 9-cis and all-trans retinoic acid. European Journal of Cancer 31A(4), 486-94. Reid, M. F. & Fewson, C. A. (1994). Molecular characterization of microbial alcohol dehydrogenases. *Critical Reviews in Microbiology* **20**(1), 13-56.

Richman, J. M. (1992). The role of retinoids in normal and abnormal embryonic craniofacial morphogenesis. *Critical Reviews in Oral Biology & Medicine* 4(1), 93-109.

Roberts, E. S., Vaz, A. D. & Coon, M. J. (1992). Role of isozymes of rabbit microsomal cytochrome P-450 in the metabolism of retinoic acid, retinol, and retinal. *Molecular Pharmacology* **41**(2), 427-33.

Rongnoparut, P. & Weaver, S. (1991). Isolation and characterisation of a cytosolic aldehyde dehydrogenase-encoding cDNA from mouse liver. *Gene* 101, 261-265.

Rosenthal, D., Lancillotti, F., Darwiche, N., Sinha, R. & De Luca, L. M. (1994). Regulation of epithelial differentiation by retinoids. In *Vitamin A in Health and Disease* (Blomhoff, R., ed.), pp. 425-450. Marcel Dekker, Inc., New York.

Ross, A. C. (1993a). Cellular metabolism and activation of retinoids: roles of cellular retinoid-binding proteins. *FASEB Journal* 7(2), 317-27.

Ross, A. C. (1993b). Overview of retinoid metabolism. *Journal of Nutrition* **123**(2 Suppl), 346-50.

Ross, A. C., Takahashi, Y. I. & Goodman, D. S. (1978). The binding protein for retinol from rat testis cytosol. Isolation and partial characterization. *Journal of Biological Chemistry* **253**(18), 6591-8.

Ruberte, E., Friederich, V., Chambon, P. & Morriss-Kay, G. (1993). Retinoic acid receptors and cellular retinoid binding proteins. III. Their differential transcript distribution during mouse nervous system development. *Development* **118**(1), 267-82.

Ruberte, E., Friederich, V., Morriss-Kay, G. & Chambon, P. (1992). Differential distribution patterns of CRABP I and CRABP II transcripts during mouse embryogenesis. *Development* 115(4), 973-87.

Saari, J. C. & Bredberg, D. L. (1988). Purification of cellular retinaldehyde-binding protein from bovine retina and retinal pigment epithelium. *Experimental Eye Research* **46**(4), 569-78.

Saari, J. C., Bredberg, L. & Garwin, G. G. (1982). Identification of the endogenous retinoids associated with three cellular retinoid-binding proteins from bovine retina and retinal pigment epithelium. *Journal of Biological Chemistry* **257**(22), 13329-33.

Saari, J. C., Bredberg, D. L. & Noy, N. (1994). Control of substrate flow at a branch in the visual cycle. *Biochemistry* **33**(10), 3106-12.

Saari, J. C., Champer, R. J., Asson-Batres, M. A., Garwin, G. G., Huang, J., Crabb, J. W. & Milam, A. H. (1995). Characterization and localization of an aldehyde dehydrogenase to amacrine cells of bovine retina. *Visual Neuroscience* **12**(2), 263-72.

Saari, J. C., Futterman, S. & Bredberg, L. (1978). Cellular retinol- and retinoic acidbinding proteins of bovine retina. Purification and properties. *Journal of Biological Chemistry* 253(18), 6432-6.

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989). In *Molecular Cloning: A* Laboratory Manual (2nd Ed.). Cold Spring Harbor Laboratory Press

Sanger, F., Nicklen, S. & Coulsen, A. R. (1977). DNA Sequencing with chainterminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* 74, 5463.

Sani, B. P. & Hill, D. L. (1974). Retinoic acid: a binding protein in chick embryo metatarsal skin. *Biochemical & Biophysical Research Communications* 61(4), 1276-82.

Schagger, H. & von Jagow, G. (1987). Tricine-Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis for the Separation of Proteins in the Range from 1 to 100 kDa. *Analytical Biochemistry* **166**, 368-379.

Schule, R., Rangarajan, P., Yang, N., Kliewer, S., Ransone, L. J., Bolado, J., Verma, I. M. & Evans, R. M. (1991). Retinoic acid is a negative regulator of AP-1-responsive genes. *Proceedings of the National Academy of Sciences of the United States of America* **88**(14), 6092-6.

Schuuring, E. M., Verhoeven, E., Eckey, R., Vos, H. L. & Michalides, R. J. (1992). Direct Submission (H3).

Scopes, R. K. (1987a). Affinity Elution from Ion Exchangers and other Non-Specific Adsorbents. In *Protein Purification - Principles and Practice* 2nd edit., pp. 141-155. Springer-Verlag New York Inc., New York.

Scopes, R. K. (1987b). Dye Ligand Chromatography. In *Protein Purification -Principles and Practice* 2nd edit., pp. 141-155. Springer-Verlag New York Inc., New York.

Scopes, R. K. (1987c). Protein Purification: Principles and Practice. 2nd edit. Springer Advanced Texts in Chemistry (Cantor, C. R., Ed.), Springer-Verlang New York Inc., New York.

Senoo, H., Stang, E., Nilsson, A., Kindberg, G. M., Berg, T., Roos, N., Norum, K. R. & Blomhoff, R. (1990). Internalization of retinol-binding protein in parenchymal and stellate cells of rat liver. *Journal of Lipid Research* **31**(7), 1229-39.

Sheikh, S., Ni, L. & Weiner, H. (1997). Mutation of the conserved amino acids of mitochondria aldehyde dehydrogenase: role of the conserved residues in the mechanism of reaction. *Advances in Experimental Medicine and Biology* **414**, 195-200.

Shingleton, J. L., Skinner, M. K. & Ong, D. E. (1989). Characteristics of retinol accumulation from serum retinol-binding protein by cultured Sertoli cells. *Biochemistry* **28**(25), 9641-7.

Siegenthaler, G. (1990). Retinoic acid formation from retinol and retinal metabolism in epidermal cells. *Methods in Enzymology* **189**, 530-6.

Sivaprasadarao, A. & Findlay, J. B. (1988). The interaction of retinol-binding protein with its plasma-membrane receptor. *Biochemical Journal* **255**(2), 561-9.

Skare, K. L., Schnoes, H. K. & DeLuca, H. F. (1982). Biliary metabolites of all-transretinoic acid in the rat: isolation and identification of a novel polar metabolite. *Biochemistry* **21**(14), 3308-17. Sladek, N. E., Manthey, C. L., Maki, P. A., Zhang, Z. & Landkamer, G. J. (1989). Xenobiotic oxidation catalyzed by aldehyde dehydrogenases. *Drug Metabolism Reviews* **20**(2-4), 697-720.

Smith, B. J. (1994a). Chemical Cleavage of Proteins. In Methods in Molecular biology. Basic Protein and Peptide Protocols (Walker, J. M., ed.), Vol. 32, pp. 297-304. Humana Press, Inc, Totowa, New Jersey.

Smith, B. J. (1994b). Enzymatic Methods for Cleaving Proteins. In *Methods in Molecular Biology. Basic Protein and Peptide Protocols* (Walker, J. M., ed.), Vol. 32, pp. 289-296. Humana Press, Inc, Totowa, New Jersey.

Smith, J. E., Milch, P. O., Muto, Y. & Goodman, D. S. (1973). The plasma transport and metabolism of retinoic acid in the rat. *Biochemical Journal* **132**(4), 821-7.

Sneath, P. H. A. & Sokal, R. R. (1973). In *Numerical taxonomy*, pp. 230-234. WH Freeman and Co., San Francisco.

Srivastava, D. K. & Bernhard, S. A. (1986). Metabolite transfer via enzyme-enzyme complexes. *Science* 234(4780), 1081-6.

Stahl, W., Sunquist, A. R., Hanusch, M., Schwarz, W. & Sies, H. (1993). Separation of beta-carotene and lycopene geometrical isomers in biological samples. *Clinical Chemistry* **39**, 810-814.

Stayner, C. K. & Tweedie, J. W. (1995). Cloning and characterisation of the cDNA for sheep liver cytosolic aldehyde dehydrogenase. *Advances in Experimental Medicine & Biology* **372**, 61-6.

Steenbock, H. & Gross, E. G. (1920). Fat-soluble vitamine. Journal of Biological Chemistry 41, 149-162.

Steinmetz, C. G., Xie, P., Weiner, H. & Hurley, T. D. (1997). Structure of mitochondrial aldehyde dehydrogenase: the genetic component of ethanol aversion. *Structure* 5(5), 701-711.

Streissguth, A. P., Barr, H. M., Sampson, P. D., Parrish-Johnson, J. C., Kirchner, G. L. & Martin, D. C. (1986). Attention, distraction and reaction time at age 7 years and prenatal alcohol exposure. *Neurobehavioral Toxicology & Teratology* **8**(6), 717-25.

Summerbell, D. & Maden, M. (1990). Retinoic acid, a developmental signalling molecule. *Trends in Neurosciences* 13(4), 142-7.

Sundelin, J., Laurent, B. C., Anundi, H., Tragardh, L., Larhammar, D., Bjorck, L., Eriksson, U., Akerstrom, B., Jones, A., Newcomer, M. & et al. (1985). Amino acid sequence homologies between rabbit, rat, and human serum retinol-binding proteins. *Journal of Biological Chemistry* **260**(10), 6472-80.

Takase, S., Ong, D. E. & Chytil, F. (1986). Transfer of retinoic acid from its complex with cellular retinoic acid-binding protein to the nucleus. *Archives of Biochemistry & Biophysics* 247(2), 328-34.

Tang, G. W. & Russell, R. M. (1990). 13-cis-retinoic acid is an endogenous compound in human serum. Journal of Lipid Research 31(2), 175-82.

Thaller, C. & Eichele, G. (1990). Isolation of 3,4-didehydroretinoic acid, a novel morphogenetic signal in the chick wing bud [see comments]. *Nature* **345**(6278), 815-9.

Thompson, J. N., Howell, J. M. & Pitt, G. A. J. (1964). Proc. R. Soc. London Ser. B 159, 510-535.

Thompson, J. R., Bratt, J. M. & Banaszak, L. J. (1995). Crystal structure of cellular retinoic acid binding protein I shows increased access to the binding cavity due to formation of an intermolecular beta-sheet. *Journal of Molecular Biology* **252**(4), 433-46.

Tilbeurgh, H., Egloff, M. P., Martinez, C., Rugani, N., Verger, R. & Cambillau, C. (1993). Interfacial activation of the lipase pro-colipase complex by mixed micelles revealed by X-ray crystallography. *Nature* **362**, 814-820.

Tottmar, S. O., Pettersson, H. & Kiessling, K. H. (1973). The subcellular distribution and properties of aldehyde dehydrogenases in rat liver. *Biochemical Journal* 135(4), 577-86.

Tsujita, M., Tomita, S., Miura, S. & Ichikawa, Y. (1994). Characteristic properties of retinal oxidase (retinoic acid synthase) from rabbit hepatocytes. *Biochimica et Biophysica Acta* **1204**(1), 108-16.

Umesono, K. & Evans, R. M. (1989). Determinants of target gene specificity for steroid/thyroid hormone receptors. *Cell* 57(7), 1139-46.

Umesono, K., Giguere, V., Glass, C. K., Rosenfeld, M. G. & Evans, R. M. (1988). Retinoic acid and thyroid hormone induce gene expression through a common responsive element. *Nature* **336**(6196), 262-5.

Umesono, K., Murakami, K. K., Thompson, C. C. & Evans, R. M. (1991). Direct repeats as selective response elements for the thyroid hormone, retinoic acid, and vitamin D3 receptors. *Cell* 65(7), 1255-66.

Urbach, J. & Rando, R. R. (1994). Isomerization of all-trans-retinoic acid to 9-cisretinoic acid. Biochemical Journal 299(Pt 2), 459-65.

Van Thiel, D. H., Gavaler, J. & Lester, R. (1974). Ethanol inhibition of vitamin A metabolism in the testes: possible mechanism for sterility in alcoholics. *Science* **186**(4167), 941-2.

Vasiliou, V., Reuter, S. F., Kozak, C. A. & Nebert, D. W. (1993). Mouse dioxininducible cytosolic ALDH3: AHD4 cDNA sequence, genetic mapping, and differences in mRNA levels. *Pharmacogenetics* **3**, 281-290.

Veech, R. L., Eggleston, L. V. & Krebs, H. A. (1969). The redox state of free nicotinamide-adenine dinucleotide phosphate in the cytoplasm of rat liver. *Biochemical Journal* 115, 609-619.

Vieira, A. V. & Schneider, W. J. (1993). Transport and uptake of retinol during chicken oocyte growth. *Biochimica et Biophysica Acta* **1169**(3), 250-6.

Vieira, A. V., Sanders, E. J. & Schneider, W. J. (1995). Transport of serum transthyretin into chicken oocytes. A receptor-mediated mechanism. *Journal of Biological Chemistry* **270**(7), 2952-6.

von Bahr-Lindstrom, H., Hempel, J. & Jornvall, H. (1984). The cytoplasmic isoenzyme of horse liver aldehyde dehydrogenase. Relationship to the corresponding human isoenzyme. *Eur. J. Biochem.* 141, 37-42.

Wang, X., Penzes, P. & Napoli, J. L. (1996). Cloning of a cDNA encoding an aldehyde dehydrogenase and its expression in Escherichia coli. Recognition of retinal as substrate. *Journal of Biological Chemistry* **271**(27), 16288-93.

Weiner, H., Wang, X.-P. & Bai, Y. L. Direct Submission (cytosolic). GENBANK accession U56604.

Weiner, H., Wang, X.-P. & Bai, Y. L. Direct Submission (mitochondrial). GENBANK accession U56605.

Williams, J. B. & Napoli, J. L. (1985). Metabolism of retinoic acid and retinol during differentiation of F9 embryonal carcinoma cells. *Proceedings of the National Academy of Sciences of the United States of America* **82**(14), 4658-62.

Wilson, J. G., Roth, C. B. & Warkany, J. (1953). An analysis of the syndrome of malformations induced by maternal vitamin A deficiency. Effects of restoration of vitamin A at various times during gestation. *American Journal of Anatomy* **92**, 189-217.

Winter, N. S., Bratt, J. M. & Banaszak, L. J. (1993). Crystal structures of holo and apocellular retinol-binding protein II. *Journal of Molecular Biology* **230**(4), 1247-59.

Wolbach, S. B. & Howe, P. R. (1925). Journal of Experimental Medicine 42, 753-777. (as cited by Wolf, 1996).

Wolf, G. (1984). Multiple functions of vitamin A. Physiological Reviews 64, 873-937.

Wolf, G. (1996). A history of vitamin A and retinoids. FASEB Journal 10(9), 1102-7.

Yang, N., Schule, R., Mangelsdorf, D. J. & Evans, R. M. (1991). Characterization of DNA binding and retinoic acid binding properties of retinoic acid receptor. *Proceedings* of the National Academy of Sciences of the United States of America **88**(9), 3559-63.

Yang, Z. N., Davis, G. J., Hurley, T. D., Stone, C. L., Li, T. K. & Bosron, W. F. (1994). Catalytic efficiency of human alcohol dehydrogenases for retinol oxidation and retinal reduction. *Alcoholism, Clinical & Experimental Research* **18**(3), 587-91.

Yin, S. J., Liao, C. S., Wu, C. W., Li, T. T., Chen, L. L., Lai, C. L. & Tsao, T. Y. (1997). Human stomach alcohol and aldehyde dehydrogenases: comparison of expression pattern and activities in alimentary tract. *Gastroenterology* **112**(3), 766-75.

Yoshida, A., Hsu, L. C. & Dave, V. (1992). Retinal oxidation activity and biological role of human cytosolic aldehyde dehydrogenase. *Enzyme* **46**(4-5), 239-44.

Yoshida, A., Hsu, L. C. & Yanagawa, Y. (1993). Biological role of human cytosolic aldehyde dehydrogenase 1: hormonal response, retinal oxidation and implication in testicular feminization. *Advances in Experimental Medicine & Biology* **328**, 37-44.

Yoshida, A., Hsu, L. C. & Yasunami, M. (1991). Genetics of human alcoholmetabolizing enzymes. *Progress in Nucleic Acid Research & Molecular Biology* 40, 255-87.

Yost, R. W., Harrison, E. H. & Ross, A. C. (1988). Esterification by rat liver microsomes of retinol bound to cellular retinol-binding protein. *Journal of Biological Chemistry* **263**(35), 18693-701.

Yu, V. C., Delsert, C., Andersen, B., Holloway, J. M., Devary, O. V., Naar, A. M., Kim, S. Y., Boutin, J. M., Glass, C. K. & Rosenfeld, M. G. (1991). RXR beta: a coregulator that enhances binding of retinoic acid, thyroid hormone, and vitamin D receptors to their cognate response elements. *Cell* 67(6), 1251-66.

Yuen, S., Hunkapiller, M. W., Wilson, K. J., & Yuan, P. M. Applied Biosystems Ltd. (1986). ABI User Bulletin No. 25.
Yun, S. L. & Suelter, C. H. (1977). A simple method for calculating K_m and V from single enzyme reaction progress curve. *Biochimica et Biophysica acta* **480**, 1-13.

Zelent, A., Krust, A., Petkovich, M., Kastner, P. & Chambon, P. (1989). Cloning of murine alpha and beta retinoic acid receptors and a novel receptor gamma predominantly expressed in skin. *Nature* **339**(6227), 714-7.

Zetterstrom, R. H., Simon, A., Giacobini, M. M., Eriksson, U. & Olson, L. (1994). Localization of cellular retinoid-binding proteins suggests specific roles for retinoids in the adult central nervous system. *Neuroscience* **62**(3), 899-918.

Zhao, D., McCaffery, P., Ivins, K. J., Neve, R. L., Hogan, P., Chin, W. W. & Dräger, U. C. (1996). Molecular identification of a major retinoic-acid-synthesizing enzyme, a retinaldehyde-specific dehydrogenase. *European Journal of Biochemistry* **240**(1), 15-22.