Institut für Experimentelle Genetik
GSF-Forschungzentrum für Umwelt und Gesundheit,
Neuherberg



Identification and characterization of zebrafish 17beta-HSD type 1 and type 3: A comparative analysis of androgen/estrogen activity regulators

Rebekka Mindnich

Vollständiger Abdruck der von der Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzender:	Univ.- Prof. Dr. Bertold Hock
Prüfer der Dissertation:	1. Priv.-Doz. Dr. Jerzy Adamski 2. Univ.-Prof. Dr. Johannes Buchner 3. Univ.-Prof. Dr. Wolfgang Wurst

Die Dissertation wurde am 30.06.2004 bei der Technischen Universität München eingereicht und durch die Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt am 07.10. 2004 angenommen.

Table of contents

ABSTRACT **7**

ZUSAMMENFASSUNG **9**

ABBREVIATIONS **11**

1 INTRODUCTION **13**

1.1 THE AIM OF THIS STUDY **14**
1.2 FUNCTION AND SYNTHESIS OF ANDROGENS AND ESTROGENS **14**
1.2.1 MEMBERS OF THE STEROID HORMONE CLASS 14
1.2.1.1 The synthesis pathway of the steroid hormone classes 14
1.2.1.2 Tissues and compartments of steroid hormone synthesis 16
1.2.2 ANDROGENS AND THEIR FUNCTIONS 16
1.2.3 ESTROGENS AND THEIR FUNCTIONS 17
1.2.4 EVOLUTIONARY CONSIDERATIONS 18
1.2.4.1 The receptors 18
1.2.4.2 The metabolising enzymes 19
1.3 CHARACTERISTICS AND FUNCTIONS OF 17BETA-HSDS **19**
1.3.1 CLASSIFICATION AND FUNCTIONS 19
1.3.2 MEMBERS OF THE SHORT-CHAIN DEHYDROGENASE/ REDUCTASE (SDR) SUPERFAMILY 21
1.3.3 17BETA-HSD TYPE 1 23
1.3.4 17BETA-HSD TYPE 3 23
1.4 THE (ZEBRA) FISH MODEL ORGANISM **24**
1.4.1 FUNCTIONS OF ANDROGENS AND ESTROGENS 24
1.4.2 ENZYMES AND RECEPTORS OF STEROIDOGENESIS 25
1.4.3 ZEBRAFISH BIOLOGY 25

2 METHODS **29**

2.1 DNA **29**
2.1.1 ISOLATION AND PURIFICATION PROCEDURES 29
2.1.1.1 Isolation of genomic DNA from zebrafish tissue 29
2.1.1.2 Isolation of plasmid DNA from bacteria 30
2.1.1.3 Purification of linear dsDNA from solutions 30
2.1.1.4 Purification of dsDNA from gels 30
2.1.2 MEASUREMENT 30
2.1.2.1 Separation and monitoring by agarose gel electrophoresis 30
2.1.2.2 Measurement by optical density 31
2.1.3 CLONING STRATEGIES 31
2.1.3.1 TOPO-TA cloning 31
2.1.3.2 Cloning via restriction sites 31
2.1.3.3 Restriction site independent cloning 32
2.1.4 PCR-BASED METHODS 33
2.1.4.1 PCR (polymerase chain reaction) 33
2.1.4.2 Fusion-PCR 33
2.1.4.3 Secondary PCR/nested PCR 34
2.1.4.4 Sequencing 34
2.1.4.5 Site-directed mutagenesis 34
2.1.5 LIBRARY SCREEN 34
2.1.5.1 Probe labeling 35
2.1.5.2 Filter-screen and signal-detection 35
2.2 RNA **36**
2.2.1 ISOLATION FROM TISSUES 36
2.2.2 HANDLING AND MEASUREMENT 36
2.2.3 REVERSE TRANSCRIPTION INTO CDNA 36
2.3 PROTEIN CHEMISTRY **37**
2.3.1 PRODUCTION OF RECOMBINANT PROTEIN IN E. COLI 37
2.3.2 SDS-POLYACRYLAMIDE GEL ELECTROPHORESIS (SDS-PAGE) 38
2.3.3 WESTERN BLOT 39

2.3.4 MEASUREMENT OF SUBSTRATE SPECIFICITY 39
2.4 BIOINFORMATIC METHODS 40
2.4.1 *IN SILICO* SCREEN 40
2.4.2 *IN SILICO* NORTHERN BLOT 41
2.4.3 ASSEMBLIES AND GENE STRUCTURE 41
2.4.4 ALIGNMENTS AND PHYLOGENY 42
2.4.5 CALCULATION OF IDENTITY PLOTS 42
2.4.6 3D-MODELING OF PROTEIN STRUCTURE 42
2.4.7 IDENTIFICATION OF TERMINAL SIGNALLING MOTIFS AND GLOBULAR DOMAINS 43
2.5 WORK WITH ORGANISMS 43
2.5.1 WORKING WITH *E. COLI* 43
2.5.2 WORKING WITH ZEBRAFISH 44

3 MATERIALS, EQUIPMENT, PROGRAMS 45

3.1 MATERIALS 45
3.1.1 ANTIBODIES 45
3.1.2 BACTERIA 45
3.1.3 CDNA-LIBRARIES (COLONY MACROARRAYS FROM RZPD) 46
3.1.4 CHEMICALS (SPECIAL) 46
3.1.5 ENZYMES 46
3.1.6 KITS 47
3.1.7 OTHER MATERIAL 47
3.1.8 RADIOACTIVE MOLECULES 47
3.1.9 VECTORS 47
3.2 EQUIPMENT 48
3.3 PROGRAMS 48

4 RESULTS 51

INTRODUCTORY REMARKS 51

4.1 17BETA-HSD TYPE 1 52
4.1.1 THE CANDIDATE GENES 52
4.1.2 CHARACTERIZATION ON DNA LEVEL 53
4.1.2.1 Coding sequence 53
4.1.2.2 Exon structure 55
4.1.2.3 Gene and intron sizes 56
4.1.3 CHARACTERIZATION ON RNA LEVEL 57
4.1.3.1 Expression analysis by RT-PCR 57
4.1.3.2 Expression analysis by *in silico* Northern blot 59
4.1.4 CHARACTERIZATION ON PROTEIN LEVEL 60
4.1.4.1 1D-structure 60
4.1.4.1.1 Overall sequence identity 60
4.1.4.1.2 Phylogenetic analysis 61
4.1.4.1.3 Conserved motifs 63
4.1.4.1.4 Distribution of identical residues 66
4.1.4.2 3D-structure 68
4.1.4.2.1 Homology modeling 68
4.1.4.2.2 Expression of recombinant protein 73
4.1.4.2.3 Activity measurements and substrate specificity 76
4.2 17BETA-HSD TYPE 3 78
4.2.1 THE CANDIDATE GENES 78
4.2.2 CHARACTERIZATION ON DNA LEVEL 79
4.2.2.1 Coding sequence 79
4.2.2.2 Exon structure 81
4.2.2.3 Gene and intron sizes 82
4.2.3 CHARACTERIZATION ON RNA LEVEL 83
4.2.3.1 Expression analysis by RT-PCR 83
4.2.3.2 Expression analysis by *in silico* Northern blot 85
4.2.4 CHARACTERIZATION ON PROTEIN LEVEL 86
4.2.4.1 1D-structure 86
4.2.4.1.1 Overall sequence identity 86
4.2.4.1.2 Phylogenetic analysis 87

4.2.4.1.3 Conserved motifs ... 89
4.2.4.1.4 Distribution of identical residues ... 91
4.2.4.2 3D-structure ... 93
4.2.4.2.1 Expression of recombinant protein ... 93
4.2.4.2.2 Activity measurements and substrate specificity ... 95

5 DISCUSSION ... 97

5.1 17BETA-HSD TYPE 1: ZEBRAFISH VS. HUMAN VS. MOUSE ... 97
5.1.1 IN EVERY RESPECT HIGHLY CONSERVED ... 97
5.1.1.1 Expression might be absent during embryogenesis ... 97
5.1.1.2 Enhanced but not exclusive expression in tissues of female adults ... 98
5.1.1.3 Functional integrity: assessing the flexibility of SDR consensus motifs ... 99
5.1.1.4 Substrate specificity: dissecting the definitely important residues ... 100
5.1.2 WHY A FISH IS NO HUMAN IS NO MOUSE (GENE): SPECIATION EFFECTS ... 102
5.1.2.1 Additional regulation acquired in viviparous vs. oviparous organisms? -The very C-terminus of 17beta-HSD type 1 ... 102
5.1.2.2 The difference that makes the difference: 17beta-HSD type 1 in different organisms, sexes and tissues ... 105
5.1.2.3 Expansion of the understanding of 1D- and 3D-protein sequence features by integration of zebrafish data ... 106
5.1.3 THE SISTER GROUP: COMPARISON TO RETINOL DEHYDROGENASES ... 109
5.1.3.1 Sticking together: traces of common evolution ... 109
5.1.3.2 Functionally different and clearly non-similar to 17beta-HSD type 1 ... 111
5.1.3.3 The second duplication: two distinct prRDHs in (zebra) fish ... 112
5.2 17BETA-HSD TYPE 3: ZEBRAFISH VS. MAMMALS ... 115
5.2.1 17BETA-HSD TYPE 3 FROM ZEBRAFISH: HOMOLOG, ANALOG, ORTHOLOG? ... 115
5.2.1.1 On common ground: shared features of zebrafish and mammalian 17beta-HSD type 3 ... 115
5.2.1.2 Fish vs. mammals – different in more than one aspect. I: structure and activity ... 116
5.2.1.3 Fish vs. mammals – different in more than one aspect. II: expression ... 121
5.2.1.4 Fish vs. mammals – different in more than one aspect. III: function ... 124
5.2.2 THE DAWN OF A NEW SPAWN: COMPARISON TO 17BETA-HSD TYPE 12 AND DERIVED GROUPS ... 125
5.2.2.1 The legacy of features of fatty acid synthesis in 17beta-HSD type 3 ... 126
5.2.2.2 Redundancy or synergism: two types of 17beta-HSD type 12 in zebrafish ... 128

6 REFERENCES ... 131

7 APPENDIX ... 145

7.1 PUBLICATIONS ... 145
7.1.1 SCIENTIFIC PAPERS ... 145
7.1.2 SEQUENCES ... 145
7.1.3 SCIENTIFIC PRESENTATIONS ... 147
7.2 PRIMER ... 147
7.2.1 SEQUENCING ... 147
7.2.2 GENERATION OF PROBES ... 147
7.2.3 RT-PCR BASED CONSTRUCTION OF FULL-LENGTH CDS ... 148
7.2.4 EXPRESSION ANALYSIS VIA RT-PCR ... 148
7.2.5 CLONING INTO EXPRESSION VECTORS ... 149
7.2.6 SITE-DIRECTED MUTAGENESIS ... 149
7.3 CONSTRUCTS ... 150
7.3.1 CDNAS ... 150
7.3.2 EXPRESSION OF RECOMBINANT PROTEINS ... 150
7.4 PROTEIN SEQUENCES FOR PHYLOGENETIC ANALYSES ... 151
7.4.1 FROM DATABASES ... 151
7.4.2 GENERATED IN THIS PHD THESIS ... 152

DANKSAGUNG (ACKNOWLEDGEMENTS) ... 155

CURRICULUM VITAE ... 157

Abstract

17beta-hydroxysteroid dehydrogenase (17beta-HSD) type 1 and type 3 function as estrogen and androgen activity regulators, respectively, thereby playing key roles in sexual differentiation, pregnancy and sex-steroid associated types of cancers in mammals. Although in non-mammalian vertebrates similar functions of the respective sex-steroids have been reported, there is still only little information available on enzymes which metabolize and regulate these substances. In this work the identification and characterization of the zebrafish homologs of 17beta-HSD type 1 and type 3 is described. *In silico* screens of available EST and genomic databases yielded a set of candidate sequences, which were successively complemented to give rise to the full-length coding sequences. For 17beta-HSD type 1 and type 3 three putative zebrafish homologs each were obtained. To the respective human and mouse homologs they displayed a highly similar exon architecture and shared about 50-60% identical amino acids. Thorough subsequent analyses on DNA, RNA and protein level were carried out to identify the true homolog among the candidate sequences.

In this way, zebrafish 17beta-HSD type 1 (zfHSD 1) and two closely related, paralogous photoreceptor-associated retinol dehydrogenases (zfprRDH 1 and zfprRDH 2) were identified. Expression analyses during embryogenesis and in tissues of adult male and female fish revealed specific expression patterns for each zebrafish gene. In this respect, zfHSD 1 was highly similar to the mammalian 17beta-HSDs type 1 whereas from the two zfprRDHs only type 1 resembled an expression profile similar to its mammalian homolog and type 2 displayed widespread expression. Phylogenetic analyses supported the affiliation of the zebrafish proteins to the respective enzyme groups and suggested the duplicated zfprRDH to be fish-specific and absent in other vertebrates. These analyses identified a close evolutionary connection between 17beta-HSDs type 1 and retinol dehydrogenases, which was further highlighted by dissection of 17beta-HSDs type 1 and prRDHs on protein level. Sequence analyses revealed residues and conserved elements typical for each enzyme group. Furthermore, presence and integrity of structural and functional motifs of the short-chain dehydrogenase/reductase family suggested all three zebrafish enzymes to be functional and similar to their mammalian homologs. This aspect was more closely investigated by homology modeling, based on the crystal structure of the human enzyme. Several amino acids important in recognition and binding of cofactor and substrate, and essential for the catalytic mechanism were identified. Structural differences to the closely related zfprRDHs especially affecting the substrate binding part were outlined and one highly conserved amino acid close to the catalytic center suggested to function in retinoid-steroid discrimination. Activity measurements on recombinant zebrafish wild type zfHSD 1, zfprRDH 1 and zfprRDH 2 supported the notion that, in spite of their close relation, only the first of these enzymes was capable of catalyzing the reduction of estrone to estradiol, characteristic for all so far analyzed 17beta-HSDs type 1. Furthermore, mutation of the aforementioned amino acid hypothesized to be involved in substrate discrimination in all three enzymes did reduce catalytic activity in zfHSD 1 whereas the two zfprRDHs did not show an increased affinity towards steroids.

Concerning the 17beta-HSD type 3 candidates, extensive analyses revealed the zebrafish homolog (zfHSD 3) and two closely related paralogous forms of 17beta-HSD type 12. Phylogenetic analyses highlighted the close and complex evolution of both groups identifying 17beta-HSD type 12 as the ancient, fatty acid metabolising progenitor of vertebrate 17beta-HSDs type 3. Unlike in case of zfHSD 1, zfHSD 3 had an expression profile different to that of its mammalian homolog. It is discussed that this may be a leftover of its fatty acid synthesizing ancestor or might be caused by involvement of zfHSD 3 in several related functions that in mammals are taken over by additional enzymes.

Zusammenfassung

Die 17beta-Hydroxysteroid Dehydrogenasen (17beta-HSD) Typ 1 und Typ 3 fungieren als Regulatoren der Aktivität von Estrogenen bzw. Androgenen und spielen somit eine zentrale Rolle in der sexuellen Differenzierung, Schwangerschaft und in den Sexsteroid-assoziierten Formen von Krebs in Säugern. Obgleich in anderen Vertebraten über ähnliche Funktionen der jeweiligen Sexsteroide berichtet wurde, ist nur wenig über die entsprechenden Enzyme, die diese Substanzen metabolisieren und regulieren, bekannt. In dieser Arbeit ist die Identifizierung und Charakterisierung der Zebrafisch-Homologen von 17beta-HSD Typ 1 und Typ 3 beschrieben. *In silico* Screens auf den verfügbaren EST- und genomischen Datenbanken führten zu einer Anzahl an Kandidaten-Sequenzen, die sukzessive vervollständigt wurden, um die kompletten kodierenden Sequenzen zu erhalten. Für 17beta-HSD Typ 1 und Typ 3 wurden auf diese Weise je drei mögliche Zebrafisch-Homologe identifiziert. Diese wiesen im Vergleich zu den jeweiligen Maus- und Mensch-Homologen eine sehr ähnliche Exon-Struktur auf und zeigten auf Protein-Ebene eine Aminosäure-Identität von ca. 50-60%. Um das wahre Zebrafisch-Homologe unter diesen Kandidaten-Sequenzen zu identifizieren, wurden nachfolgend gründliche Untersuchungen auf DNA-, RNA- und Protein-Ebene durchgeführt.
Auf diese Weise konnte die Zebrafisch 17beta-HSD Typ 1 (zfHSD 1), sowie zwei hierzu nahe verwandte, paraloge Photorezeptor-assoziierte Retinol Dehydrogenasen (zfprRDH 1 und zfprRDH 2) identifiziert werden. Expressions-Analysen während der Embryogenese und in Geweben erwachsener, männlicher und weiblicher Fische ergaben spezifische Expressions-Muster für jedes Zebrafisch-Gen. Im Vergleich zu den jeweiligen Säuger-Homologen zeigte zfHSD 1 eine sehr hohe Ähnlichkeit, wohingegen nur Typ 1 der beiden zfprRDHs ein vergleichbares Expressions-Profil aufwies; zfprRDH 2 hingegen zeigte eine nahezu ubiquitäre Expression. Phylogenetische Analysen unterstützten die Zugehörigkeit der Zebrafisch-Proteine zu den jeweiligen Enzym-Gruppen und ließen vermuten, dass die beobachtete Duplizierung von zfprRDH Fisch-spezifisch und abwesend in anderen Vertebraten sei. Diese Untersuchungen identifizierten eine enge, evolutionäre Beziehung zwischen 17beta-HSDs Typ 1 und Retinol Dehydrogenasen, welche durch die eingehende Analyse dieser Enzyme auf Protein-Ebene noch hervorgehoben wurde. Sequenz-Analysen zeigten Aminosäuren und konservierte Elemente auf, die typisch für die jeweilige Enzymgruppe sind. Zudem ließ die Anwesenheit und Integrität von strukturellen und funktionellen Motiven der short-chain dehydrogenase/reductase Familie vermuten, dass alle drei Zebrafisch-Enzyme funktionell und in dieser Hinsicht ihren Säuger-Homologen ähnlich seien. Dieser Aspekt wurde noch eingehender durch Homologie-Modellierung, basierend auf der Kristallstruktur der menschlichen 17beta-HSD Typ 1, untersucht. Eine Anzahl von Aminosäuren, die wichtig für die Erkennung und Bindung von Cofaktor und Substrat und essentiell für den katalytischen Mechanismus sind, konnte identifiziert werden. Strukturelle Unterschiede zu den nahe verwandten zfprRDHs, die insbesondere den Substrat-Bindenden Teil betreffen, wurden herausgearbeitet; eine hoch konservierte Aminosäure wurde in unmittelbarer Nähe zum katalytischen Zentrum identifiziert und spekuliert, dass diese eine Rolle in der Unterscheidung zwischen Retinoid- und Steroid-Substraten spielen könnte. Aktivitätsmessungen mit den rekombinanten, wildtypischen Zebrafisch-Proteinen zfHSD 1, zfprRDH 1 und zfprRDH 2 unterstützten die Annahme, dass trotz der nahen Verwandtschaft nur das erste dieser Enzyme die Reduktion von Estron zu Estradiol, eine Reaktion die charakteristisch für alle bislang analysierten 17beta-HSDs Typ 1 ist, katalysierte. Des Weiteren führte die Mutation der zuvor erwähnten Aminosäure, evtl. beteiligt

an der Substrat-Unterscheidung, in allen drei Enzymen zu einer Reduktion der katalytischen Aktivität in zfHSD 1, während die beiden zfprRDHs keine erhöhte Affinität zu Steroiden zeigten.
Im Falle der 17beta-HSD Typ 3 Kandidaten führten die umfangreichen Analysen zur Identifizierung des Zebrafisch-Homologen (zfHSD 3), sowie zwei hierzu nahe verwandten, paralogen Formen der 17beta-HSD Typ 12. Phylogenetische Analysen enthüllten eine vielschichtige und komplexe Evolution beider Gruppen und identifizierten 17beta-HSD Typ 12 als den Fettsäuren metabolisierenden Vorläufer der 17beta-HSD Typ 3 der Vertebraten. Im Gegensatz zu zfHSD 1 wies das Expressions-Profil von zfHSD 3 deutliche Unterschiede zu dem des Säuger-Homologen auf. Dieser Befund wurde als ein Rudiment des Fettsäure metabolisierenden evolutiven Ursprungs diskutiert; alternativ wurde die Ausübung verschiedener, verwandter Funktionen von zfHSD 3, die in Säugern von zusätzlichen Enzymen ausgeführt werden, in Betracht gezogen.

Abbreviations

aa	amino acid
APS	ammoniumperoxodisulphate
BLAST	basic linear alignment search tool
bp	base pair
BSA	bovine serum albumin
B/W	blue/white
cDNA	complementary DNA
cds	coding sequence
C. elegans	Caenorhabditis elegans
CNS	central nervous system
DAB	diaminobenzidine
DHEA	dihydroxyepiandrosterone
DHT	dihydrotestosterone
DNA	deoxyribonucleic acid
dpc	days post coitum
NTP	nucleotide triphosphate
ds	double-stranded
E. coli	Escherichia coli
EDTA	ethylenediaminetetraacetic acid
ER	endoplasmic reticulum
ERGIC	ER-Golgi interacting cytoplasm
EST	expressed sequence tag
Fig	figure
GST	glutathione S-transferase
hpf	hours post fertilization
HPLC	high pressure liquid chromatography
HSD	hydroxysteroid dehydrogenase
IPTG	isopropyl-beta-D-thiogalactopyranoside
kDa	kilo Dalton
LB	Luria-Bertani
LCFA	long chain fatty acid
mRNA	messenger RNA
NAD	β-Nicotinamide adenosine dinucleotide
NADP	β-Nicotinamide adenosine dinucleotide 3'-phosphate
NADPH	β-Nicotinamide adenosine dinucleotide 3'-phosphate, reduced form
nt	nucleotide
PAGE	polyacrylamid gelelectrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
prRDH	photoreceptor-associated RDH
RDH	retinol dehydrogenase
RNA	ribonucleic acid
rpm	rotations per minute
RT-PCR	reverse transcriptase PCR
SDS	sodium dodecylsulfate
SDR	short-chain dehydrogenase/reductase
ss	single-stranded
Tab	table
TEMED	tetramethylethylenediamine
VLCFA	very long chain fatty acid
vs	versus
WGS	whole genome shotgun
X-Gal	5-bromo-4-chloro-3-indolyl-beta-D-Galactopyranoside

1 Introduction

In the light of evolution, sexual reproduction is one of the most powerful strategies to establish an adaptable yet stable life form. Quite a spectrum of possibilities has evolved to accomplish differentiation into two sexes accompanied by the split and subsequent new combination of the genetic material. In lower metazoans, sexual and asexual reproduction may be cyclically organized or influenced by environmental parameters. Furthermore, the specific sex can be acquired, lost, changed, or combined with the complementing sex in the same organism. The process of sexual differentiation can be determined genetically, by environmental factors or a combination of both. With increase of the complexity of the organism and a growing potential to actively interact with its environment, sex becomes more and more genetically determined as exemplified in the chordates.

Sex determination results in sexual differentiation apparent in gametogenesis, gonadogenesis, and neuronal and other somatic changes leading to the sexually dimorphic genders. Just like other adapting processes, this requires a complex interplay of the neuronal and endocrine system. In vertebrates, androgens and estrogens have been identified as the central hormonal mediators of vertebrate sexual development. They are embedded in a highly organized network of specific metabolic pathways, availability modes by storage, binding and serum clearance, and exertion of their activity via binding to their receptors, which elicits a number of genomic and non-genomic effects.

The fatal consequences of disturbing the balance of this hormonal system became most obvious when in the 1980s the endocrine disrupting potential of a number of chemically engineered substances was identified. Unnatural sex-ratios, malformations, impairment of reproduction and cancer were observed in wild-life. Infertility and elevated risk of sex-associated cancer was also monitored in humans that ingested an increased amount of the respective substances. Aside from these exogenous factors, similar effects can be observed due to genetic defects in the regulatory network. In addition, more and more data recently accumulate that suggest a strong influence of androgens and estrogens on the obvious sexually dimorphic appearance and severity of a variety of diseases, among them osteoporosis, Alzheimer's disease, schizophrenia, and obesity.

The unravelling and understanding of sex hormone action is a prerequisite to counter these negative effects on fertility, health and well-being. In addition to studies in humans, model systems allow for a more detailed investigation of processes not easily accessible let alone manipulable in men. In the work at hand, the zebrafish model was chosen to study the function of two key regulatory enzymes of androgen and estrogen activity. From an experimental point of view, this model system offers easy monitoring and manipulation of embryogenesis, rearing in high amounts, and well-established methods of genetics and molecular biology. From an evolutionary point of view, the zebrafish presents a highly organized vertebrate giving insights into regulatory processes and functions central to and conserved in this phylum. Furthermore, as an aquatic organism it is a model for endocrine disruptor function, and comparative studies can help to assess the risks and understand the mechanisms of potential endocrine disruptors.

1.1 The aim of this study

The last steps of androgen and estrogen formation include redox reactions at position C17 of the steroid scaffold, which are performed by the 17beta-hydroxysteroid dehydrogenases (17beta-HSDs). In mouse and human, 17beta-HSD type 1 is considered to be the main activator of estrogens while 17beta-HSD type 3 is the predominant enzyme catalyzing the formation of highly active testosterone from androstenedione. Although steroid hormones are common to vertebrates, and estrogens and androgens are supposed to play similar functions as well in non-mammalian species, data from the latter group of organisms are scarce and rather concentrate on the hormones themselves than on their metabolism and regulation. To get more insight into this field the study on hand investigated the zebrafish homologs of mammalian 17beta-HSD type 1 and type 3.

As the respective genes were unknown, the first aim of the study was their identification, and assessment of the true homologs among a set of candidate sequences. During this process closely related genes were identified for both, 17beta-HSD types 1 and 3, by careful phylogenetic analyses in a so far uncharacterized context.

The aim was also to identify modes of conservation typical for each enzyme type, comparing the fish and mammalian homologs as well as the respective 17beta-HSDs with their closely related genes. For this, a detailed analysis of the nucleic and amino acid sequence was carried out elucidating the gene and exon structures, and structural and functional motifs in the protein sequences.

To get insight into the function of these genes in non-mammalian organisms their expression profiles were analyzed during development and in tissues of adult fish of both sexes.

The substrate specificity was investigated upon cloning of expression constructs and generation of recombinant enzymes in *E. coli*.

1.2 Function and synthesis of androgens and estrogens

1.2.1 Members of the steroid hormone class

1.2.1.1 The synthesis pathway of the steroid hormone classes

Steroid hormones are all derived from the common precursor cholesterol. In subsequent steps of shortening the side chain at C17 of the steroid backbone and further modifications, the different steroid hormone classes are formed (Fig. 1). As a consequence, the steroid hormone groups appear in a sequential order and branch from common intermediate precursors.

According to the C-atoms in the scaffold, the steroids can be divided into three groups: 1) C21-steroids, including the progestogens and the two corticoid groups: mineralocorticoids and

glucocorticoids; 2) C19-steroids which comprise the androgens; 3) C18-steroids that are resembled by estrogens.

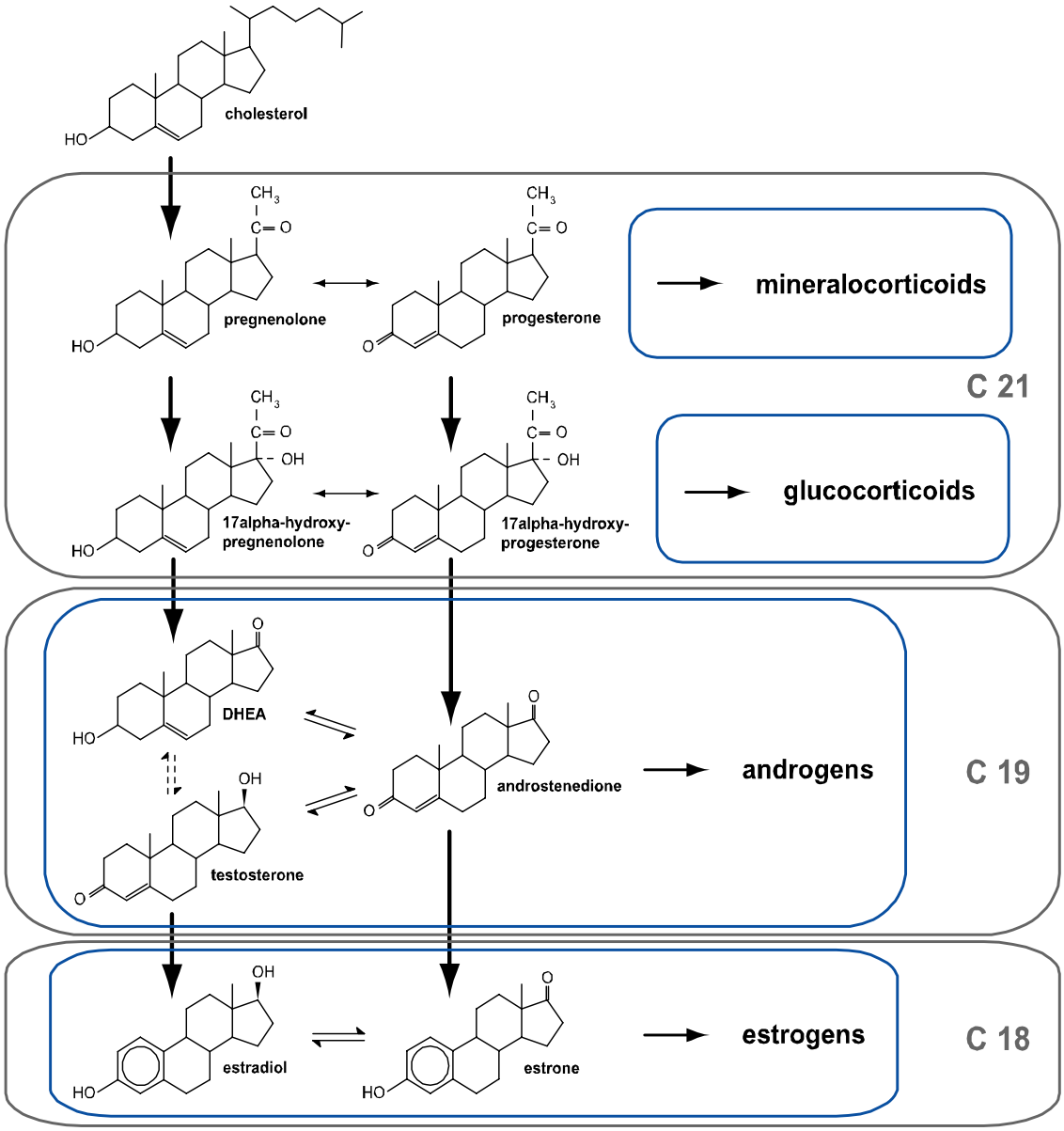


Fig. 1: Synthesis of the steroid hormone classes from cholesterol.
Groups with the same number of carbons in their backbone are boxed as are the four main steroid classes. Broken arrows indicate several steps of synthesis catalyzed by more than one enzyme.

Along the synthesis pathway of the vertebrate steroid hormones, progesterone is the first substance proven to be biologically highly active and which exerts its function through binding to specific receptors. In vertebrates, this hormone is involved in sexual behaviour and follicular development, and also in pregnancy in mammals. In non-mammalian vertebrates, it is as well involved in oogenesis and ovulation. Mineralocorticoids regulate the mineral metabolism while

glucocorticoids govern the glucose metabolism, effect protein biosynthesis, and exert an immunorepressive and anti-inflammatory effect. Finally, the so called sex hormones, androgens and estrogens, are considered to regulate sexual development and further sex-specific aspects in adults (see also chapters 1.2.2 and 1.2.3).

1.2.1.2 Tissues and compartments of steroid hormone synthesis

The steroid hormone synthesis as outlined above does not take place ubiquitously in the organism but is restricted to certain organs and tissues. Cholesterol can be acquired through the diet but in vertebrates is to a higher degree synthesised *de novo*. As this substance is also an essential part of the cell plasma and the mitochondrial membrane it is considered to be formed ubiquitously. During development though, a distinct expression pattern of the cholesterol synthesizing enzymes was detected (Laubner et al., 2003) which was in concert with another function of cholesterol, namely the modification of signalling molecules such as the hedgehog protein (Beachy et al., 1997). Cholesterol monooxygenase, the rate-limiting step in steroid synthesis, generates pregnenolone and is predominantly expressed in steroidogenic organs such as the adrenal, placenta, gonads and brain but it seems to be as well present at other sites like liver and intestine. The corticoids are specifically synthesised in the adrenal cortex: glucocorticoids are generated in the *zona fasciculata* and mineralocorticoids are formed in the *zona glomerulosa*. Synthesis of dehydroepiandrosterone and androstenedione, the precursors of sex hormones, might be restricted to steroidogenic tissues as the responsible enzyme, 17alpha-hydroxylase/17,20-lyase (Cyp17A1), is specifically expressed at the respective sites. Finally, P450aromatase (Cyp19A1), which catalyzes the formation of estrogens, is predominantly found in female specific tissues but is also present in the adrenal, the testis and other sites.
The steroids themselves are secreted to act on their target tissues. These could be cells in direct vicinity (paracrine action) or tissues further away only reachable via the blood stream (endocrine action). The steroid hormones can also be secreted and act back on the same cell from which they originated or might even act at the place of synthesis without being secreted (intracrine action)(Labrie, 1991). All these modes of action are conferred via binding to specific receptors and control of their expression is one of the key-steps in regulating responsiveness to hormonal signals. Furthermore, extensive metabolism of steroid hormones providing pre-receptor regulation in target tissues and inactivation at non-target sites is present. This on site synthesis of steroid hormones is defined by the complex expression pattern of regulatory enzymes such as the 17beta-HSDs (see also chapter 1.3) and can account for a great part of distinct hormone action apart from the “classical” pathway via synthesis in and subsequent secretion from endocrine organs as described above.

1.2.2 Androgens and their functions

DHEA and androstenedione are the first androgenic molecules in the steroid synthesis pathway, formed from the progestin precursors. The reactions take place in the testis and adrenal due to the presence of Cyp17A1 especially in these tissues. From here, the androgens can be

released into the serum and converted to active hormones wherever the necessary set of metabolising enzymes is present.
In mammals, DHEA, androstenedione, testosterone and 5α-dihydrotestosterone are considered the main active androgens, their affinity to the androgen receptor increasing in this sequence. Some by definition androgenic compounds can even bind to the estrogen receptor and hence, exert estrogenic activity (Poulin and Labrie, 1986; Weihua et al., 2002).
The androgens are specifically involved in development of male primary sexual characteristics during embryogenesis such as growth and differentiation of the male reproductive tract and outer genitalia (Matsumoto et al., 2003). They are essential for sexual maturation in puberty and spermatogenesis (Matsumoto et al., 2003) and furthermore, govern the development of male secondary sex characteristics such as the deepening of the voice and hair growth and bodily distribution in men (Andersson et al., 1996). Male specific differentiation processes in the central nervous system and brain, which amongst others affect sexual behavior and libido, are also dependent on androgens. Effects of this sex steroid concern sex specific skeletal and skeletal muscle growth and the distribution of adipose tissue (Matsumoto et al., 2003).
In non-mammalian vertebrates, similar physiological functions are controlled by androgens though the specific molecules might differ from those in mammals. Indeed, a huge set of androgens can be identified from a given organism (Borg et al., 1992; Borg, 1994), their physiological function mainly unknown. It is generally accepted that in fish 11-oxygenated androgens are the main players due to their high serum levels and functions in spermatogenesis (Miura et al., 1991) and male sexual behavior (Pall et al., 2002). But significant roles were also identified for the classical androgens (Loir, 1999) and it seems that different species developed their own composition and balance of metabolites to exert androgenic functions.

1.2.3 Estrogens and their functions

The most common vertebrate estrogens are estradiol and estrone, which are formed directly upon aromatization by P450 aromatase from testosterone and androstenedione, respectively. In vertebrates, the main sites of aromatization are the gonads and the brain, in accordance with the function of estrogens in reproduction. In non-mammalian organisms, estrogens govern the development of the ovaries and can also affect the testis of genetic males leading to feminization (Govoroun et al., 2001) or complete sex-reversal. In mammals, estrogen is obligatory for fertility as it is involved in normal folliculogenesis and maintenance of the somatic cells within the ovaries (Britt and Findlay, 2002). During pregnancy, another estrogen, estriol, is formed in high amounts from 16α-hydroxy-DHEA but its biological functions are unclear. In male mammalian reproduction, estrogens are essential for sustaining fertility by promotion of spermatogenesis (Eddy et al., 1996; Robertson et al., 1999).
There is not much known about estrogen functions in the brain of non-mammalian organisms but a general positive effect on neuronal proliferation, synaptogenesis and survival, as well as regulation of morphology has been observed. But exposure to very high levels of estrogens in zebrafish has also been shown to disturb early CNS development (Kishida et al., 2001). In mammals, estrogens are involved in female specific brain differentiation and are necessary for development of sexual behavior of both genders (Scordalakes et al., 2002; Bakker et al., 2003; Temple et al., 2003).

The P450 aromatase is also expressed in mammalian non-gonadal tissues, such as bone, skin and adipose tissue. In agreement with this, an involvement of estrogens in bone-turnover (Oz et al., 2000) and fat-metabolism (Jones et al., 2000; Nemoto et al., 2000) was observed.
The significant role of estrogens in reproduction may even be older than the evolution of vertebrates. This view is supported by identification of the respective metabolizing enzymes and receptors involved in estrogen signalling (see also chapter 1.2.4). Functioning of estrogens as sex steroids associated with gametogenesis, reproductive cycle, sexual differentiation and associated processes has been demonstrated in cnidarians (Tarrant et al., 1999; Pernet and Anctil, 2002; Twan et al., 2003), molluscs (Osada and Nomura, 1990; Osada et al., 2003), arthropoda (Keshan and Ray, 1998; Warrier et al., 2001) and echinodermata (Hines et al., 1992; Wasson et al., 2000).

1.2.4 Evolutionary considerations

1.2.4.1 The receptors

Phylogenetic studies suggest that the nuclear steroid receptors are unique to vertebrates (Laudet, 1997; Thornton, 2001). As a consequence of two subsequent genome duplications, considered to have taken place during vertebrate development, ancient estrogen (ER), corticoid (CR), and progesterone/androgen (PR/AR) receptor evolved from a common nuclear receptor progenitor. Following divergence of lampreys, each group split again to give rise to the known nuclear steroid receptors ERα, ERβ, GR (glucocorticoid receptor), MR (mineralocorticoid receptor), PR and AR (Thornton, 2001). In teleosts some steroid receptors underwent even more duplications leading to new functional receptors as was observed for two glucocorticoid (Bury et al., 2003; Greenwood et al., 2003), two androgen (Ikeuchi et al., 1999; Sperry and Thomas, 1999a; Sperry and Thomas, 1999b; Takeo and Yamashita, 1999), three estrogen (Hawkins et al., 2000), and two progestogen receptors (Ikeuchi et al., 2002). In all cases, the two isoforms likely take over specialised roles in the same physiological processes although their definite function still remains to be characterized.
Aside from their nuclear localization and exertion of genomic effects involving changes in gene transcription, these classical steroid receptors were demonstrated to initiate signalling cascades upon ligand binding alone (Kousteni et al., 2001). In addition, the observed nongenomic effects seem to be in a context with localization of the receptor to the plasma membrane (Pietras and Szego, 1977; Levin, 1999; Razandi et al., 1999).
Apart from these membrane bound classical receptors membrane-associated steroid receptors (MASR) have been identified. Their signalling mechanism is not yet clarified but at least some of them seem to act via G-protein coupling (Grazzini et al., 1998; Falkenstein and Wehling, 2000). In vertebrates, involvement of these MASR in oocyte maturation has been intensely studied (Hammes, 2003) but they might also be involved in non-sexual, somatic functions (Lieberherr and Grosse, 1994; Armen and Gay, 2000; Razandi et al., 2002). Interestingly, membrane-associated progesterone receptors form a highly conserved protein family (Zhu et al., 2003) with homologs present already in yeast (Hand et al., 2003). Receptors with an estrogen-binding potential and definite physiological role were also identified in invertebrates (Keshan and Ray,

1998; Di Cosmo et al., 2002; Paolucci et al., 2002; Osada et al., 2003), though their structural and genetic characteristics remain unclear.

1.2.4.2 The metabolising enzymes

The synthesis of sterols is an ancient pathway present already in yeast. Even bacteria are capable of forming polycyclic molecules, the hopanoids, from squalene, the common precursor of sterol synthesis. Cholesterol itself can be generated in fungi, plants and animals although this ability was secondarily lost in nematodes and arthropods that are cholesterol auxotroph. In non-vertebrate organisms, the potential to synthesise the same or similar steroids as identified in vertebrates has barely been investigated. At least concerning androgen and estrogen metabolism some data exist. A number of fungi is able to exert 17beta-HSD activity on androgenic and estrogenic compounds (Itagaki and Iwaya, 1988; Rizner et al., 1996; Rizner et al., 2001) and also 3beta-HSD activity in the interconversion of DHEA and androstenedione was observed (Vihko et al., 2002). Androgen metabolism has been demonstrated in invertebrates including echinoderms (Schoenmarkers and Voogt, 1980; Wasson et al., 1998) and mollusks (Le Guellec et al., 1987). As well, aromatase activity converting androgen precursors into estrogens was identified in the protochordate amphioxus (Callard et al., 1984) and mollusks (Le Guellec et al., 1987; Le Curieux-Belfond et al., 2001). Although these androgens and estrogens seem to play a role in reproduction of the respective invertebrate species, their function in protozoa may be different. Furthermore, 17beta-HSDs from fungi do not seem to be related to the vertebrate enzymes as was suggested by the phylogenetic analysis of the 17beta-HSD from *Cochliobolus lunatus* (Lanisnik Rizner et al., 2001a). No other gene for the corresponding activity in the aforementioned invertebrates has been identified so far, and the evolutionary context hence remains unclear.

1.3 Characteristics and functions of 17beta-HSDs

1.3.1 Classification and functions

The 17beta-HSDs are enzymes capable of catalyzing a redox reaction at the C17 of the steroid backbone with steric preference of a hydroxy group in beta-orientation (Fig. 2). By use of NADP or NAD as cofactors, 17beta-HSDs *in vivo* preferentially either reduce or oxidize their steroid substrate (for overview see Tab. 1). The identity of the oxygenated group at C17 has strong influence on the biological potential of the substrate: a hydroxy group leads to enhanced binding and thereby activation of the receptor whereas a keto group is less potent (Jungblut et al., 1981). As a consequence, 17beta-HSDs function as pre-receptor regulators of steroid activity.

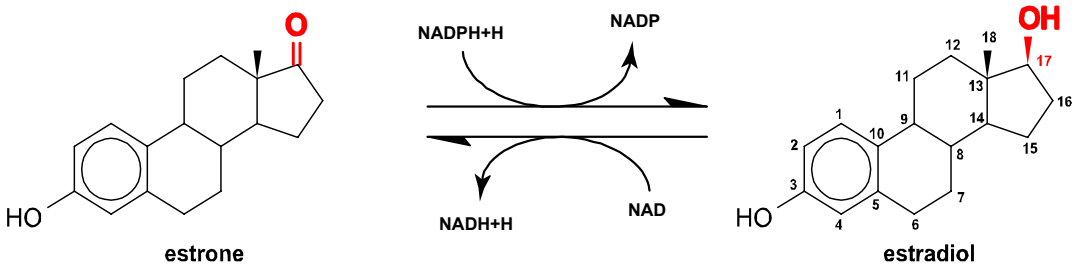


Fig. 2: Redox reaction at C17 of the steroid backbone as catalyzed by 17beta-HSDs.
The depicted interconversion of estrone and estradiol is an example to illustrate the stereo specific character of the reaction and the targeted position in the steroid. In general, NADPH is the preferred cofactor for 17beta-HSDs catalyzing the reduction whereas NAD functions in the oxidation reaction.

The first and up to now best characterized member of this enzyme family was cloned from human placenta libraries (Peltoketo et al., 1988; Luu The et al., 1989). This protein was named 17beta-HSD type 1 as due to enzymatic analysis, it was hypothesized that multiple forms of 17beta-HSDs might exist (Blomquist, 1995). Indeed, more and more proteins were discovered in the following time meeting the enzymatic requirements of a 17beta-HSD and they were numbered as type 2, type 3 and so on, in accordance to the sequence in which they were identified (Penning, 1997; Peltoketo et al., 1999; Adamski and Jakob, 2001; Mindnich et al., 2004).

So far, twelve different 17beta-HSDs have been annotated. The individual enzymes differ by their temporal and spatial expression pattern as well as their substrate specificities. Subsequent detailed analysis revealed that aside from the 17beta-HSD activity many members of this protein family also harbored additional enzymatic activities towards other positions of the steroid-backbone, fatty acids or retinoids. 17beta-HSDs do not form a gene family and quite likely emerged at different points during evolution, lacking a common ancestor. In this context it may not be surprising that for some 17beta-HSDs homologs do not exist in some closely related species as is the case for rat 17beta-HSDs type 6 (Biswas and Russell, 1997) and mouse type 9 (Su et al., 1999). Homologs for the functionally corresponding genes could up to now not be demonstrated in human.

In the wake of genome sequencing projects, more and more raw sequence data are accumulating that open new, bioinformatic ways to identify novel proteins or protein family members. This trend has also led to the annotation of human 17beta-HSD type 12, the most recent member of the family. Sequence similarity to 17beta-HSD type 3 might have prompted the authors to classify this enzyme as a 17beta-HSD, whereas at the same time another group demonstrated this sequence to be identical to the human 3-ketoacyl reductase, involved in fatty acid synthesis (Moon and Horton, 2003). So far, no activity of this enzyme towards hydroxy- or keto-groups at C17 of the steroid backbone or any other steroid has been proven. This example very well demonstrates the chances and risks of bioinformatic data analysis and that, at least in case of the 17beta-HSDs, thorough characterisation is still inevitable.

Tab. 1: Annotated 17beta-HSDs and their substrate and enzymatic characteristics.

17beta-HSD type	*In vivo* reaction	Main steroid specificity	Other enzymatic activities
1	reductase	estrone	20α-HSD
2	dehydrogenase	DHEA, estradiol, testosterone	20α-HSD, 3β-HSD
3	reductase	androstenedione	not known
4	dehydrogenase	estradiol	β-oxidation
5	reductase	DHEA, androstenedione	3α-HSD
6	dehydrogenase	DHT, estradiol	3α-HSD
7	reductase	estrone, zymosterone	3β-HSD, 3α-HSD
8	dehydrogenase	estradiol	not known
9	dehydrogenase	3α,17β-androstanediol	3α-HSD, retinol DH
10	dehydrogenase	DHT	β-oxidation, 3α-HSD
11	dehydrogenase	3α,17β-androstanediol	not known
12	reductase	not known	LCFA-, VLCFA-synthesis

DH: dehydrogenase; DHEA: dehydroepiandrosterone; DHT: dihydrotestosterone; HSD: hydroxysteroid dehydrogenase; LCFA: long chain fatty acid; VLCFA: very long chain fatty acid. For further information and references see also Mindnich et al. (2004).

Aside from the vertebrate enzymes, 17beta-HSD activity was identified in a variety of non-vertebrate organisms (see also chapter 1.2.4.2). The responsible proteins from some algae, fungi and bacteria were purified and analyzed but the corresponding genes are still elusive with the exception of 17β-HSDcl from *Cochliobolus lunatus*, which has been thoroughly characterized on amino and nucleic acid level (Rizner et al., 1996; Rizner et al., 2000; Lanisnik Rizner et al., 2001a; Lanisnik Rizner et al., 2001b; Kristan et al., 2003). 17β-HSDcl seems not to be related to the vertebrate 17beta-HSDs and rather might have evolved by convergent evolution. In contrast to the vertebrate 17beta-HSD family, it is not clear what functions the respective steroid compounds and enzymes play in these organisms. While in fungi and microbial organisms steroids might rather be catabolized in terms of a carbon-source (Talalay et al., 1952; Marcus and Talalay, 1956) or toxicant (Oppermann et al., 1996), this might be different in higher metazoans.

1.3.2 Members of the short-chain dehydrogenase/ reductase (SDR) superfamily

Although 17beta-HSDs do not belong to a gene family, all members with the exception of 17beta-HSD type 5 show the structural features characteristic for the SDR superfamily. At primary sequence level, identity among functionally different forms of SDRs is quite low, usually between 15 and 30% (Krozowski, 1994; Oppermann et al., 1999). Instead, 3D folds of the N-terminal part are very similar and a variety of motifs such as the Rossmann fold, the cofactor binding region, the residues of the catalytic triad and additional positions important in selection and positioning of the cofactor are common to all SDRs (Oppermann et al., 2003). Another common feature lies in their quaternary structure: with a few exceptions such as carbonyl reductase or 17beta-HSD type 10, active SDRs are present as either dimers or tetramers.

The superfamily can be split into five groups depending on the amino acid residues participating in the formation of the individual motifs and the length of the sequence forming the cofactor binding part of the enzyme (for overview see Fig. 3) (Kallberg et al., 2002). A second subdivision can be performed by taking specific patterns of the cofactor-binding residues into account (Persson et al., 2003).



Motifs in SDRs:

No.	Classical SDR	Extended SDR	Suggested function
(1)	TGxxxGhG	TgxxGhaG	Structural role in coenzyme binding region
(2)	Dhx[cp]	DhxD	Adenine ring binding of coenzyme
(3)	GxhDhhhNNAGh	[DE]xhhHxAA	Structural role in stabilising central β-sheet
(4)	hNhxG	hNhhGTxxhhc	Part of active site
(5)	GxhhxhSSh	hhhxSSxxhaG	Part of active site
(6)	Yx[AS][ST]K	Pyxx[AS]Kxxh[DE]	Part of active site
(7)	h[KR]h[NS]xhxPGxxxT	h[KR]xxNGP	Structural role, reaction direction

Fig. 3: SDR superfamily subgroups and typical motifs.
The superfamily consists of five distinct families with characteristic length and motif composition. In addition, Classical and Extended SDRs can be subdivided by their coenzyme-binding residue patterns (not shown). The table below lists the typical amino acids of seven structurally important and highly conserved SDR motifs as well as the suggested functions of these themes. Abbreviations for amino acid residues in one letter code and small letters: a: aromatic; c: charged; h: hydrophobic; p: polar; x: any residue.

SDRs have been identified in all life forms and with more than 3000 sequences known from the database set up one of the largest protein superfamilies. They comprise a broad range of substrate specificities being involved in amino acid, fatty acid, steroid, sugar, bile acid, retinol, and neurotransmitter metabolism, as well as synthesis of tropane alkaloids such as cocaine in plants.

1.3.3 17beta-HSD type 1

The first identified 17beta-HSD catalyzes the formation of a highly potential estrogen reducing estrone to estradiol by use of the cofactor NADPH. This enzymatic activity is also shared by the rodent, chick (Wajima et al., 1999) and eel (Kazeto et al., 2000b) homologs. In addition, 17beta-HSD type 1 to some extent converts DHEA to Δ5-androstene-3β,17β-diol (Dumont et al., 1992) and androstenedione to testosterone (Nokelainen et al., 1996; Puranen et al., 1997a). While in the human enzyme this androgenic activity is 715-fold lower than the estrogenic activity and hence, negligible, the rat isozyme displays a very similar catalytic efficiency for both reactions (Puranen et al., 1997a). Despite the partial acceptance of these other substrates, 17beta-HSD type 1 is considered to be the main activator of estrogens in all species, where it has been characterized so far. This is in line with the preponderant expression in ovaries and in humans also in placenta. The rodent enzyme on the other hand is absent from the placenta (Nokelainen et al., 1996; Akinola et al., 1998). In addition to these steroidogenic tissues, human 17beta-HSD type 1 is also found in other tissues such as breast epithelium, uterus, testis, brain and adipose tissue (Luu-The et al., 1990; Martel et al., 1992). During embryogenesis, 17beta-HSD type 1 mRNA was detected at various sites at very low levels (Mustonen et al., 1997; Takeyama et al., 2000).
Increased reduction of estrone to estradiol positively correlates with cell proliferation and hence, 17beta-HSD type 1 might play a role in estrogen-related cancers. Especially in human breast disorders, 17beta-HSD type 1 is expressed in proliferative diseases without atypia, atypical ductal hyperplasia, ductal carcinoma *in situ* and invasive ductal carcinoma. Results from the respective investigations indicate that breast carcinoma can effectively convert estrone to active estradiol and thus exert estrogenic actions on tumor cells through the estrogen receptor (Gunnarsson et al., 2001; Miyoshi et al., 2001; Suzuki et al., 2002; Vihko et al., 2002). In this context, 17beta-HSD type 1 seems to be a promising candidate for the development of specific inhibitors in the fight against estrogen-related cancers. Many studies investigate the potential of synthetic substances and as well phytoestrogens to modulate 17beta-HSD type 1 activity and that of other steroid converting enzymes (Le Bail et al., 2000; Krazeisen et al., 2001; Poirier, 2003). The elucidation of the 3D-structure by X-ray crystallography of 17beta-HSD type 1 complexed with estrone and NADPH (Ghosh et al., 1995; Azzi et al., 1996) boosted the understanding of substrate binding and the reaction mechanism. These findings were the base for further studies on substrate specificity investigating the potential of distinct residues and positions by use of chimeric enzymes (Puranen et al., 1997b).

1.3.4 17beta-HSD type 3

The reduction of androstenedione at C17 leads to the formation of testosterone and is catalyzed by 17beta-HSD type 3. Although this reaction can in addition be performed by 17beta-HSD type 5 and as well by rodent type 1, the type 3 enzyme is considered to be the main player in the generation of testosterone. In consistence with this assumption, 17beta-HSD type 3 has been shown to be predominantly expressed in testis (Geissler et al., 1994). Specific transcripts were also detected in the human brain (Stoffel-Wagner et al., 1999; Beyenburg et al., 2000). The influence of 17beta-HSD type 3 on levels of active androgen is highlighted in case of prostate cancer, where susceptibility might correlate with mutations in the 17beta-HSD type 3 gene

(Margiotti et al., 2002), and expression of the enzyme is indeed increased in prostate cancer compared to non-cancer tissues (Koh et al., 2002). Genetic defects in 17beta-HSD type 3 were also demonstrated to cause male hermaphroditism. In this disorder, affected males posses female external genitalia (Geissler et al., 1994), though structures derived from the mesonephric duct are normally developed. This phenotype hints at additional regulatory enzymes in androgen activity besides 17beta-HSD type 3, which is further sustained by the finding that affected male patients virilize at the time of expected puberty as the result of increases in serum testosterone (Andersson et al., 1996). Affected females, lacking a functional 17beta-HSD type 3, are asymptomatic (Rosler et al., 1996; Mendonca et al., 1999).

1.4 The (zebra) fish model organism

Fish model organisms are widely used to address questions of vertebrate evolution and various developmental aspects. That way, androgen and estrogen functions were identified in a number of teleosts as well as lungfish, coelacanths, sharks and rays, and the more primitive lampreys and hagfish. Furthermore, sexual reproduction is studied on fish to serve industrial purposes for improvement of farming and breeding as well as risk assessment of putative endocrine disruptors. Here, experimental findings are often still restricted to morphological observations or general changes in metabolism due to lack of knowledge of the molecular mechanisms and identity of the involved genes. This gap is now starting to close rapidly as can be seen by the accumulating data of fish genome and EST sequence data.

1.4.1 Functions of androgens and estrogens

In fish like in other vertebrates, androgens are supposed to be involved in male-specific and estrogens in female-specific aspects. The sex-steroids act on target tissues to regulate gametogenesis, reproduction, sexual phenotype and behavioural characteristics. Estradiol is required for differentiation of the gonads into ovaries. Its central function in this process is mirrored in the finding that inhibition of aromatase converts gonads of genetic females into functional testis (Guiguen et al., 1999; Bhandari et al., 2004) whereas the reverse effect can be observed by administration of estradiol or estrogenic substances to genetic males (Gimeno et al., 1996). During oogenesis, both testosterone and estradiol take over specific functions similar to those observed in mammals. In this process, testosterone is synthesized by the thecal cells and subsequently transformed into estradiol upon aromatization in the granulosa layer prior to plasma secretion (Redding and Patino, 1993; Cyr and Eales, 1996). Estradiol then triggers a series of steps including the production of vitellogenin, a necessary protein for oogenesis in oviparous organisms (Wahli, 1988).
Concerning fish reproductive endocrinology, androgen synthesis typically takes place in the Leydig cells and may include the typical mammalian hormones testosterone and androstenedione as well as their 11-oxygenated forms (Redding and Patino, 1993), which are considered to be fish-specific. Androgens are required for spermatogenesis (Miura et al., 1991;

Loir, 1999) but have also been studied in the context of male-specific behaviour. In the latter aspect, character and level of non-modified and 11-oxygenated androgens were found to be associated with mating (Brantley et al., 1993), nesting behaviour (Pall et al., 2002) and social status (Oliveira et al., 2002).

1.4.2 Enzymes and receptors of steroidogenesis

In the field of sex hormone metabolism and action, some of the central enzymes and receptors have recently been cloned and characterized from various fish: e.g. P450 cholesterol side chain cleavage enzyme (Lai et al., 1998; Hsu et al., 2002), 3beta-hydroxysteroid dehydrogenase (Sakai et al., 1994; Lai et al., 1998; Kazeto et al., 2003), 17alpha-hydroxylase/17,20-lyase (Trant, 1995; Sakai et al., 1992; Kazeto et al., 2000b), P450 aromatase (Ijiri et al., 2003), 11beta-hydroxylase (Kusakabe et al., 2002), androgen receptors (Ikeuchi et al., 1999; Sperry and Thomas, 1999a; Sperry and Thomas, 1999b; Takeo and Yamashita, 1999; Kim et al., 2002), and estrogen receptors (Xia et al., 1999; Hawkins et al., 2000; Kim et al., 2002). But pre-receptor regulators such as 17beta-HSDs remain elusive with the exception of eel 17beta-HSD type 1 (Kazeto et al., 2000a).
Analysis of the teleost genes revealed many duplications, which are absent in other vertebrates, for example duplicated steroid hormone receptors (see also chapter 1.2.4.1) and two functionally diverse genes for P450 aromatase (Chiang et al., 2001). For explanation of this phenomenon, it is assumed that the teleost genome underwent a complete duplication, which did not take place in other vertebrates (Aparicio, 2000). At first glance, this may complicate the understanding of the respective genes and functions, and as well comparison to their homologs e.g. in mammals. But then, studies on the two zebrafish P450 aromatases for example have demonstrated that the multiple functions exerted by one gene in mammals has been neatly split into two, now differently regulated genes in the fish (Chiang et al., 2001; Tong et al., 2001). This enables a much more detailed study of the diverse P450 aromatase functions than in the complex mammalian system and furthermore allows insight into gene evolution following duplication events.

1.4.3 Zebrafish biology

Various fish model systems are currently used, the choice mostly depending on the scientific question addressed but also influenced by availability, laboratory facility space and breeding characteristics. In these respects, zebrafish is a small fish, easy to keep and breed, which apart from a number of defined strains can also be obtained from pet shops. Further characteristics such as a short life cycle, rapid development of transparent embryos, high number of progeny, vertebrate body plan, and the possibility to produce mutations by using mutagenic agents in large-scale studies, are promoting the zebrafish as a model system in many diverse fields of bioscience research.

Embryogenesis and life cycle

A major focus of interest is still the rapid development of zebrafish embryos which is complemented about 3 days post fertilization (dpf) by hatching of a free-swimming larva (Fig. 4)(for further detailed information see also Kimmel et al. (1995)). Following the successive cleavages of the zygote leading to the blastula stage, gastrulation starts at about 6 hours post fertilization (hpf). With the end of gastrulation, head and tail, dorsal and ventral side, are defined.

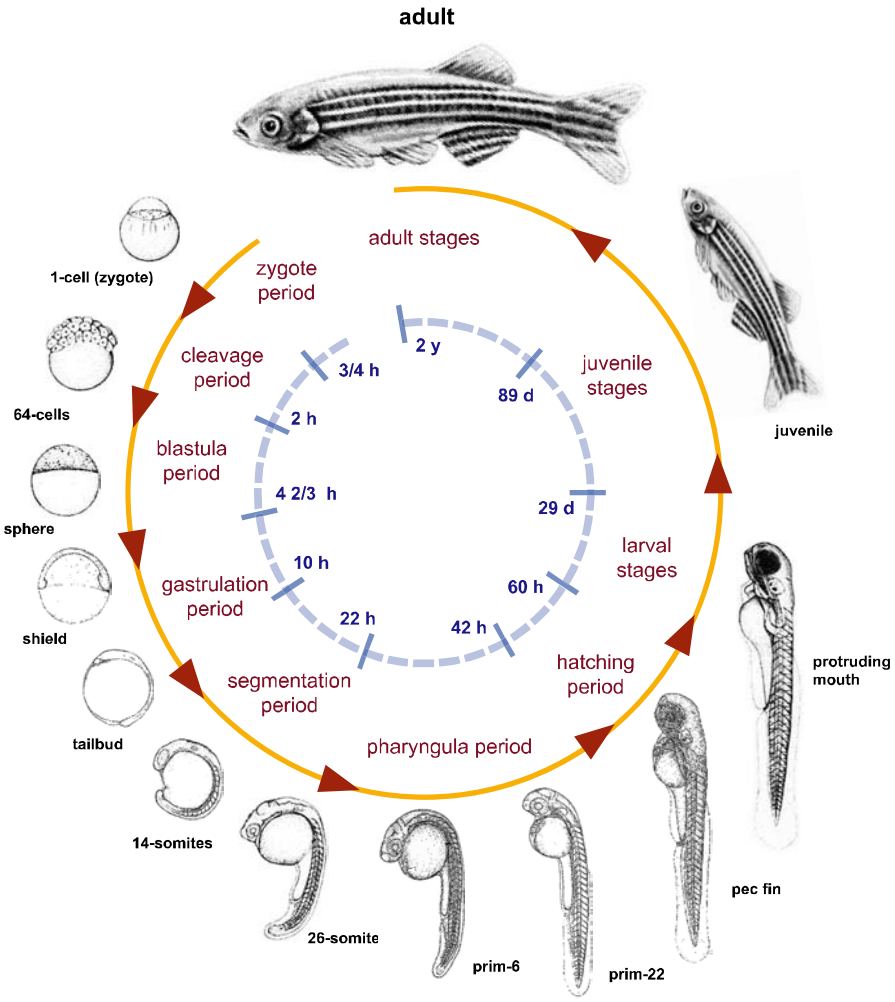


Fig. 4: Developmental stages in the life cycle of the zebrafish.
Embryogenesis of the zebrafish proceeds fast and is finished about 3 days post fertilization. After 30 days, the free-swimming larva has developed into the juvenile fish. Within the next 60 days the juveniles grow further and mature until they reach adulthood and fertility about 90 days after fertilization. The distinct developmental periods are outlined (red) in context with a time-scale (blue), that refers to the point of fertilization (0 h). Typical stages of development, and how they are termed, are depicted. Modified from Kimmel et al. (1995). h: hours: d: days; y: years.

During the next 12 hours, the segmentation period, somitogenesis, neurulation and early brain development, muscle formation and contraction, and primary organogenesis take place. The pharyngula period (at about 24-48 hpf) is characterized by fin formation, pigment cell differentiation, set-up of the circulatory system and onset of the heart-beat, appearance of the

liver and swim bladder primordia, and ongoing development of the gut. Finally, hatching takes place sporadically between 48-72 hpf during the proceeding development of the embryo. In the subsequent time the larva undergoes further growth and development changing gradually into the juvenile fish which reaches maturation at about three months post fertilization.

Germ cells and gonadogenesis

Aspects of reproduction have been widely studied in many commercially interesting fish such as trout or salmon. With growing interest in the molecular genetics of these processes though, data on germ cell development and gonadogenesis are also becoming available for zebrafish.

In fish, as in other vertebrates, the primordial germ cells (PGCs), the precursors of the gametes, are set aside from the somatic lineage early in development and arise at extragonadal locations (Nieuwkoop and Sutasurya, 1979). In the zebrafish embryo PGCs become morphologically distinct at the earliest at the 5-somite stage (Braat et al., 1999) but they can also be identified by expression of *vasa* RNA (Olsen et al., 1997; Yoon et al., 1997). This marker is already present in previtellogenic oocytes and allows to follow the formation of the PGCs at the 32-cell stage (Yoon et al., 1997), subsequent translocation to the dorsal aspect of the gut at around 3 dpf, and migration into the region of the prospective gonads at 4 dpf (Braat et al., 1999).

At about 10 days post hatch the differentiation of the gonads begins. As the zebrafish is considered to be a gonochoristic, undifferentiated species, both sexes now start to develop ovaries, which in males resembles a type of juvenile hermaphroditism (Takahashi 1977). The extent of apoptotic events and their timing in the subsequent gonadogenesis are characteristic for the differentiation into both sexes. While in female zebrafish development of the ovaries is characterized by elimination of a restricted amount of oocytes by apoptosis at an early stage of 15-19 days post-hatching (dph), this takes place at a later time-point (21-25 dph) in male-specific development and is here accompanied by testicular differentiation (Uchida et al., 2002). Differentiation of the gonads is finished at about 40 dph but their development proceeds up to 60 dph (Takahashi 1977). During this period gonadogenesis is still sensitive to developmental cues and can for example be completely reversed from male to female by administration of estrogenic compounds (Andersen et al., 2003).

2 Methods

2.1 DNA

2.1.1 Isolation and purification procedures

2.1.1.1 Isolation of genomic DNA from zebrafish tissue

Frozen intestine was used for the preparation of genomic DNA from zebrafish (for dissection see chapter 2.5.2). The tissue was thawed and digested by addition of 700 µl Tail Buffer K and agitation at 55°C overnight. The solution was then extracted twice in the following way: 700 µl Roti-Phenol/Chloroform (Roth) were added to the digested tissue solution and mixed for 1 hour at room temperature by overhead tumbling. After centrifugation in a table top centrifuge at 4°C for 10 minutes at 7000 rpm the upper phase was transferred to a fresh reaction tube. After the second extraction residual chloroform was removed by adding 700 µl chloroform/isoamylalcohol (24:1), shaking for 30 minutes and centrifugation as described above. The upper phase was transferred to a fresh reaction tube and the DNA precipitated by addition of 500 µl isopropanol, mixing and immediate centrifugation for 1 minute at 12000 rpm and 4°C. The pellet was washed in the same way with 500 µl 70% ethanol and afterwards air-dried. The genomic DNA was resuspended in 50 µl AE Buffer at 4°C for 1-3 days.

1x TE:	10 mM Tris-HCl, pH 7.4 1 mM EDTA, pH 8
AE Buffer:	5 mM Tris-HCl, pH 8.5
Proteinase K (Roche):	20 mg/ml in 0.5x TE, pH 8
Tail Buffer:	50 mM Tris, pH 8 50 mM EDTA 100 mM NaCl 0.5 % SDS
Tail Buffer K :	Tail Buffer : Proteinase K (40 :1)

2.1.1.2 Isolation of plasmid DNA from bacteria

Isolation of plasmid DNA from *E. coli* was carried out in two different scales: mini-preparation using 1-5 ml of an overnight culture and the NucleoSpin Plasmid Kit (Macherey&Nagel), and midi-preparation using 20 ml of an overnight culture and the NucleoBond PC 100 Kit (Macherey&Nagel). Bacteria were harvested by centrifugation at 4°C and the pellet resuspended in the appropriate buffer of the kit. DNA was isolated according to the manufacturer's protocol. In the mini-preparation procedure, DNA was eluted from the column with 50 µl AE buffer (provided with kit) or Ampuwa water. DNA from midi-preparations was resuspended in 100-200 µl Ampuwa water.

2.1.1.3 Purification of linear dsDNA from solutions

DNA purification from solutions containing enzymes, dNTPs, salts from buffers etc. was usually carried out by use of the PCR Purification Kit (Qiagen) according to manufacturer's protocol. For elution 30 µl Ampuwa water were applied.
Radioactive labeled DNA was purified by use of either the MicroSpin™ S-200 HR columns (Pharmacia Biotech) or the Nucleotide Removal Kit (Qiagen) according to manufacturers' protocol.

2.1.1.4 Purification of dsDNA from gels

The DNA was visualized in ethidiumbromid stained agarose gels under UV light and cut from the gel using a sharp scalpel. The DNA containing slice was transferred to reaction tubes and further processed by use of Gel Extraction Kit (Qiagen) as described in the manufacturer's protocol. Elution of the DNA was carried out with 30 µl Ampuwa water.

2.1.2 Measurement

2.1.2.1 Separation and monitoring by agarose gel electrophoresis

For qualitative measurement of DNA or subsequent isolation, DNA-solutions were subjected to agarose gel electrophoresis. Different w/v-percentages of agarose in 1x TBE were used for optimal analysis of the variously sized DNA fragments: 0.5% for dsDNA > 5 kb, 1% for dsDNA ~ 2-5 kb, 2% for dsDNA ~ 0.5-2 kb and 3% for dsDNA < 0.5 kb. For analysis, Loading Dye was added to the DNA solution, mixed, and loaded on a gel containing ethidium bromide (0.5 µg/ml). The DNA was separated in a constant electric field. DNA bands were monitored and photographed under UV illumination (λ=254nm).

10x TBE: 108 g Tris
55 g boric acid
9.3 g EDTA in 1l H_2O

6x Loading Dye: 15 % Ficoll 400 (Pharmacia)
0.25 % bromophenol blue
0.25 % xylene cyanol FF

2.1.2.2 Measurement by optical density

To monitor the quantity and quality of nucleic acid preparations, DNA (or RNA) containing solutions were diluted with Milli-Q water and the optical density of this dilution measured at λ=260 nm for quantification and additionally at λ=280 nm for qualification. The concentration of the nucleic acid solution was calculated according to the following formula:

OD_{260nm} x dilution factor x nucleic acid specific factor* = c [ng/µl]
*: for dsDNA: 50 ng/µl, for RNA: 40 ng/µl

The quality was assessed by the ratio value R (OD_{260nm}/OD_{280nm}) where R=1.8 for dsDNA and R=2.0 for RNA in an optimum case.

2.1.3 Cloning strategies

2.1.3.1 TOPO-TA cloning

For direct cloning of DNA fragments from PCR reactions without prior cleanup, the TOPO-TA cloning Kit (Invitrogen) was used. In this method, A-overhangs produced in the elongation step of the PCR by the Taq-polymerase are ligated to T-overhangs in the respective vector (TA-cloning). The yield of ligation products is enhanced by topoisomerase, attached to the vector's cloning sites (TOPO-cloning).
Following amplification of the DNA by PCR 4 µl of the reaction mixture were subjected to the ligation procedure into vector pCRII or pCR2.1 and transformed into chemocompetent TOP 10 cells as recommended by the manufacturer.

2.1.3.2 Cloning via restriction sites

Instead of TA-cloning (see above), DNA fragments were also inserted into vectors via restriction sites. Digestion of DNA fragment and vector with restriction enzymes yields complementary DNA ends which can then be used for ligation. In contrast to TA-cloning, utilization of two different restriction sites additionally allows for a site-directed insertion.

Restriction digestion
DNA fragments (inserts) and vectors were digested with the particular restriction enzymes. 1-10 µg DNA were digested in 20-50 µl reaction volume containing appropriate concentrations of buffer, BSA (if enzymes from NEB were used) and enzyme for 2 hours up to overnight at the optimum temperature for the enzyme's activity as reported by the manufacturer. The adequate amount of enzyme was calculated according to the assumption that 1 U of enzyme digests 1 µg DNA in 1 hour under optimum conditions. Following restriction the reaction was stopped by heat inactivation of the enzyme. Alternatively, the enzyme was removed by DNA purification using the PCR Purification Kit (Qiagen) (see also chapter 2.1.1.3).

Ligation
Inserts were cloned into the vector via restriction sites by T4-DNA-Ligase (NEB or MBI) activity. The 10 µl reaction mixture contained 100-200 ng vector and 5-10 times more moles of insert than vector in 1x T4-DNA-Ligase buffer. Ligation was carried out at room temperature for 2 hours up to overnight.

Transformation
For transformation of the ligation construct (see above) or plasmids from DNA preparations (see chapter 2.1.1.2) into *E. coli,* 3 µl of the ligation reaction were added to one reaction tube of chemocompetent cells (10-100 µl), left on ice for 30 minutes and heat-shocked in a 42°C-waterbath for 45 seconds. 300-900 µl LB-medium (for recipes see chapter 2.5.1) were added, the mixture incubated on a 37°C-shaker for 1 hour and an aliquot of the transformation reaction plated onto agar-plates containing the appropriate antibiotic for selection (for recipes see chapter 2.5.1).

2.1.3.3 Restriction site independent cloning

An alternative cloning method where the ligation itself is partially independent of the nature of the restriction sites was exploited in the context of this work. The idea is to increase the hybridization efficiency between vector and insert by elongation of the complementary sequence from a few base pairs (single strand overhang produced by restriction enzyme) to a stretch of up to 40 bp. As a side effect, ligation of any kind of DNA ends can be performed. The principle is as follows and schematically depicted in Fig. 5: a ligation primer is added in extent to vector and insert. Due to formation of homologous base pairs this primer spans the ligation site and adjoins the two ends of the complementary strand which can then be ligated. The result is a vector including the insert but consisting of only one intact circular strand and one with "loose ends", which can be repaired by the cells following transformation.
According to the specific sequences in vicinity of the ligation-sites of vector and insert, appropriate ligation primers of about 80 bp were designed. In principle, the ligation mixture was identical to that of a classical ligation (see above) but in addition contained the ligation primer. For ligation, vector (~ 100 ng), insert and ligation primer were mixed in a reaction tube in the molar ratio 1:3:7 respectively and water added up to a total volume of 8 µl. The mixture was heated to 95° for 10 minutes and afterwards cooled to room temperature for 10 minutes. 1 µl T4-DNA-Ligase buffer and 1 µl T4-DNA-Ligase were added and the mixture incubated for 3 hours at room temperature. The complete ligation reaction was then subjected to transformation and further processed as described above.

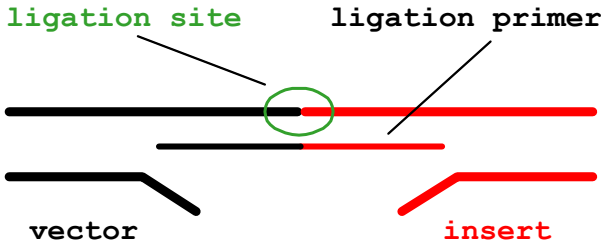


Fig. 5: Principle of restriction site independent cloning.
Only one of the two sites where the insert is to be ligated into the vector is depicted. The ligation primer will attach to the complementary sequence of both sites, displacing the second strand of vector and insert, respectively. As a result, the two ends on the opposite side will be adjoined and subsequently can be ligated by T4-DNA-Ligase.

2.1.4 PCR-based methods

2.1.4.1 PCR (polymerase chain reaction)

PCRs were conducted on DNA templates in 20, 50 or 100 µl formats containing 0.2 mM dNTP-mix, 1 µM forward- and reverse- primer each, and 0.5-2.5 U polymerase in 1x PCR Buffer. Usually, lab-made Taq-polymerase was used; in special cases the Taq was substituted by Herculase- or Pfu Turbo DNA-polymerase (both Stratagene) for better performance or improved accuracy. DNA templates included genomic DNA, plasmid DNA, linear dsDNA, cDNA, primary PCR (without prior cleanup) or bacterial culture.
The standard program on a Robo-Cycler PCR machine (Stratagene) implied an initial denaturing for 5' at 95°C followed by 35 cycles with 30'' at 95°C, 30'' at T_a , 1' per 1 kb at 72°C, where T_a (annealing temperature) is T_m - 5°C and T_m the melting temperature specific for the applied primer pair. Initial denaturing was elongated to 10 minutes if PCR was performed on bacterial culture; cycling number was decreased to 25 cycles if resulting PCR product was subjected to cloning into expression vectors and a final elongation step of 10 minutes at 72°C was added if TOPO-TA cloning was performed.

10x PCR Buffer: 100 mM Tris-HCl, pH 9.0
500 mM KCl
15 mM $MgCl_2$

2.1.4.2 Fusion-PCR

In some cases it was not possible to obtain a full-length PCR product in one step or a purchased plasmid contained only a truncated sequence. To generate a DNA fragment consisting of the full-length sequence Fusion-PCR was performed with the two templates

sharing a minimum overlap of at least 30 bp. A standard PCR (see above) was performed in 100 µl format containing equimolar amounts of both fragments if both were PCR fragments or double molar concentrations of PCR product compared to plasmid if a subcloned sequence was elongated.

2.1.4.3 Secondary PCR/nested PCR

Where a normal (primary) PCR did not yield a specific product or too low amounts, 1 µl of this reaction was subjected as template to a second PCR (secondary PCR). If possible, this secondary PCR was performed with one or both primers binding closer to the core of the primary PCR product (nested PCR).

2.1.4.4 Sequencing

To identify or verify sequences and to check for mutations, DNA fragments or plasmids were either sent to SequiServe company (Vaterstetten, Germany) or sequenced on the CEQ 2000 (Beckman) by the method of controlled termination of replication (Sanger dideoxy method) according to the manufacturer's protocol.

2.1.4.5 Site-directed mutagenesis

Insert-sequences of plasmids can be manipulated by site-directed mutagenesis changing, deleting, or adding a certain amount of nucleotides. For this procedure, a pair of complementary specific primers were designed containing the mutation in their core region and flanking it to both sides by ~15-20bp. In this work, site-directed mutagenesis was employed to exchange three neighboring nucleotides, thereby resulting in a changed amino acid at this position after translation.
The 50 µl PCR reaction with 2.5 U Pfu Turbo DNA polymerase (Stratagene) contained approximately 20 ng template plasmid, 0.2 mM dNTP-mix and 10 µM of each mutagenesis primer in 1x Pfu Buffer. After 3' denaturing at 95°C and 16 cycles with 30'' 95°C, 1' 55°C and 6' 68°C (about 1' per 1000bp), template DNA was digested by addition of 1 µl DpnI (10 U/µl, NEB) to the PCR reaction and incubation at 37°C for 1 hour. 3 µl of this mixture were subjected to transformation into JM107 chemocompetent cells as described in chapter 2.1.3.2.

2.1.5 Library screen

For identification of cDNA clones probably containing the full-length coding sequence of genes, a variety of filter libraries (for details on libraries see chapter 3.1.3) were screened with ^{32}P-labeled PCR-fragments derived from RT-PCR (for RT reaction see chapter 2.2.3).

2.1.5.1 Probe labeling

Depending on their length, probes were labeled by two different procedures:

Random primed labeling
dsDNA fragments ≥ 450 nt were labeled by use of α-^{32}P-dCTP and Prime-It RmT Random Primer Labeling Kit (Stratagene) according to the manufacturer's protocol. Probes of ~150-450 nt were labeled using the Megaprime DNA Labeling System (Amersham) according to manufacturer's protocol with a slight modification: to prevent generation of very short labeled fragments due to random labeling, the random primer mix (provided with the Kit) was substituted by gene specific forward and reverse primers added to the reaction mixture in a molar concentration about 2.5 times higher than that of the DNA template.

Terminal labeling
Alternatively to random primed labeling, small DNA-fragments (~150-300 nt) were 5'-end labeled by use of γ-^{32}P-dATP and T4 polynucleotide kinase (MBI) according to manufacturer's protocol.

2.1.5.2 Filter-screen and signal-detection

Dry library-filters were first moistened with water and subsequently with Church buffer that was pre-heated to 65°C. For prehybridization, filters were incubated in rotating hybridization tubes with 20 ml Church buffer at 65°C for at least 1 hour or up to overnight. The ^{32}P-labeled probe was added and hybridized overnight at 65°C. Stringency washes were carried out twice for 10 minutes each in a 65°C water bath with approximately 150 ml washing solution in agitated boxes. Filters were sealed in plastic wrap and transferred to film cassettes containing an intensifier screen. X-ray films were exposed to the radioactive membranes overnight or for up to one week at -80°C. Development of X-ray films revealed the coordinates of putative positive clones that were ordered from RZPD primary resource center (Berlin, Germany) according to manufacturer's protocol of evaluation.
In case of immediate re-use of filters for screening with a different probe, the membranes were stripped in strip buffer for 1 hour at 85°C before a new pre-hybridization was carried out.

Church buffer :	0.5 M sodium-phosphate buffer (pH 7.2) 7 % SDS 1 mM EDTA
Strip buffer:	5 mM sodium-phosphate buffer (pH 7.2) 0.1 % SDS
Washing solution:	40 mM sodium-phosphate buffer (pH 7.2) 0.1 % SDS

2.2 RNA

2.2.1 Isolation from tissues

RNA was prepared from zebrafish embryos and organs of adult male and female fish for expression analysis by RT-PCR. For other RT-PCR applications, RNA was isolated from a whole adult male fish.

Embryos
RNA extraction was performed using the RNeasy Mini Kit (Qiagen). Between 20 and 35 embryos (for staging of embryos see chapter 2.5.2) were homogenized in QBT buffer using a syringe and a 20G needle and subsequently processed according to the manufacturer's protocol.

Adult
For total RNA from adult zebrafish, one male fish (AB wild type strain) was homogenized in liquid nitrogen using a mortar and pestle. About 150 mg of the homogenized tissue was subjected to RNA-extraction using the RNeasy Midi Kit (Qiagen).
For preparation of RNA from zebrafish organs, several organs were pooled and homogenized in TRIzol (Invitrogen) using a rotor-stator (for dissection of organs see chapter 2.5.2). RNA was extracted by addition of 2/10 volume of chloroform, vortexing for 15 seconds, incubation for 3 minutes at room temperature and centrifugation for 15 minutes at 4°C at maximum speed in a table-top Eppendorf centrifuge. 0.53 volumes of ethanol were added to the supernatant while slowly vortexing. The solution was applied to a pre-equilibrated column of the RNeasy Midi Kit (Qiagen) and further processed according to the manufacturer's protocol.
Amount and quality of the total RNA were assessed by spectrophotometry and gel electrophoresis, respectively.

2.2.2 Handling and measurement

Following preparation of RNA, 20-40 U of Ribonuclease Inhibitor (MBI) were added (to prevent breakdown of labile RNA) and the RNA stored at -80°C. For quantification, a dilution of the RNA solution was measured via optical density as described in chapter 2.1.2.2. To survey the integrity of isolated RNA it was separated on a 1% agarose gel in TBE (as described for DNA in chapter 2.1.2.1).

2.2.3 Reverse transcription into cDNA

To avoid contamination by genomic DNA the RNA was treated with RQ1 RNase-free DNase I (Promega) according to the manufacturer's protocol; by application of the reaction mixture to the

RNeasy Mini Kit (Qiagen) in a clean-up procedure as described in the kit's manual, buffer and enzyme were removed. Transcription into cDNA was performed by use of RevertAid^{TM} First Strand cDNA Synthesis Kit (MBI) according to manufacturers' protocol. cDNA used in expression analysis was generated from 1 μg total RNA or the amount yielded from about 20 embryos (developmental stages shield to tailbud) with poly dT-primers in a total volume of 20 μl, from which 1 μl was then subjected to RT-PCR (for general PCR procedure see chapter 2.1.4.1)

2.3 Protein chemistry

2.3.1 Production of recombinant protein in E. coli

The full-length cds of the zebrafish genes was cloned into two different kinds of expression vectors resulting in N-terminally tagged fusion proteins: pGEX 2T PL2 vector (Leenders et al., 1996) attaching a GST and pQE 30, or alternatively pET 15b, enclosing a His_6-tag. Constructs were transformed into *E. coli* JM107 (for transformation procedure see chapter 2.1.3.2) and stored as glycerol-stocks (see also chapter 2.5.1) after verification of correct cloning by sequencing.
For production of recombinant protein, a 3 ml pre-culture was grown overnight at 37°C in LB-amp (for recipe of LB medium see chapter 2.5.1). 0.5 ml of this culture were transferred to 50 ml fresh LB-amp and bacteria were grown at 37°C for another 3-4 hours until OD_{600} reached 0.8-1.0. The culture was split evenly and to one aliquot IPTG was added for induction of protein expression at a final concentration of 1 mM; cultures were grown for another 3-4 hours. Bacteria were harvested by centrifugation, resuspended in lysis buffer and lysed by five freeze/thaw cycles using liquid nitrogen and hot water. The genomic DNA was digested by addition of 1 U endonuclease (Benzonase from Sigma) and $MgCl_2$ (5 mM final concentration). Samples were centrifuged to separate soluble and insoluble proteins. The supernatant was transferred to a fresh tube and the pellet fraction resuspended in an equal amount of lysis buffer.

Lysis buffer:	PBS + 0.1 mg/ml lysozyme + protease inhibitor mix
PBS:	10 mM sodium-phosphate buffer (pH 7.4) 150 mM NaCl
Protease inhibitor mix (1000x):	20 mg/ml Antipain 0.2 mg/ml Aprotinin 0.2 mg/ml Leupeptin

2.3.2 SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

Proteins of the fractions from isolation procedure (see previous chapter) were separated under denaturing conditions by SDS-PAGE using minigels (8.6x6.8cm), and subsequently visualized by Coomassie staining. 10 µl of each fraction was added to 10 µl 5x Laemmli, heated for 5 minutes at 95°C and separated on a Tricin-Gel (10 % resolving gel, 4 % stacking gel) in a constant electric field. For detection of the proteins the gel was immersed in Coomassie staining solution for 30 minutes and destained by washing with 10 % acetic acid at room temperature overnight or under heating for a few hours until background staining was removed. For documentation, gels were equilibrated in H_2O and dried between two sheets of cellophane.

AA:	acrylamide/bisacrylamide, 30 % (30:0.8)
Anode buffer:	0.2 M Tris-HCl, pH 8.9
Cathode buffer:	0.1 M Tris 0.1 M Tricin 0.1 % SDS
Coomassie staining solution:	200 ml methanol 5 ml acetic acid 295 ml H_2O 500 mg Coomassie blue G250 (filtrate)
Gel buffer:	3 M Tris-HCl, pH 8.45 0.3 % SDS
5x Laemmli buffer:	50 % glycerol 4 % SDS 0.1 % Coomassie blue G250 (filtrate) 0.2 M Tris-HCl, pH 6.8
Resolving gel (10 %):	3.3 ml AA 3 ml Gel buffer 1 ml H_2O 2.5 ml glycerol (50 %) 20 µl TEMED 50 µl APS
Stacking gel (4 %):	0.67 ml AA 0.67 ml Gel buffer 3.67 ml H_2O 10 µl TEMED 40 µl APS

2.3.3 Western blot

To visualize the fusion proteins via monoclonal antibodies against their specific tags, the separated protein fractions from the SDS gels were transferred onto a PVDF-membrane by semi-dry blotting. The gel was equilibrated in blotting buffer for 10 minutes, the membrane pre-wet with methanol and then immersed in blotting buffer; filter paper was soaked with blotting buffer. The blot was set up in the sequence “anode - filter paper – membrane – gel – filter paper – cathode” and the proteins were transferred onto PVDF in 30 minutes at 20 V. After blotting the membrane was blocked in PBS/5 % milk powder at RT for 30 min, rinsed with PBS and incubated overnight at 4°C with the primary antibody diluted 1:5000 in PBS/0.5 % milk powder (for primary antibodies see chapter 3.1.1). The membrane was rinsed three times for 5 minutes each in PBS and incubated for 2 hours at room temperature with the secondary antibody diluted 1:2000 in PBS/0.5 % milk powder (for secondary antibodies see chapter 3.1.1). After rinsing the membrane three times for 5 minutes each in PBS, 25 ml freshly prepared developing solution was added. Color development was stopped by removal of developing solution and rinsing the membrane with H_2O.

5x Developing buffer (pH 7.5):	0.225 g NaH_2PO_4 1.59 g Na_2HPO_4 11 g NaCl 8.5 g Imidazol 625 µl Tween-20 ad 250 ml with H_2O
Blotting buffer:	48 mM Tris 33 mM Tricin 1.3 mM SDS (10 %) 20 % methanol
Developing solution:	1x Developing Buffer 100 µl $CoCl_2$ (10 mg/ml) 1 mg Diaminobenzidine (DAB) 10 µl H_2O_2 ad 25 ml with H_2O
PBS:	10 mM sodium-phosphate buffer (pH 7.4) 150 mM NaCl

2.3.4 Measurement of substrate specificity

The aim of this work was to identify and characterize the zebrafish homologs of mammalian 17beta-HSDs type 1 and 3, which catalyze redox reactions at position 17 of estrone and estradiol (usually by 17beta-HSD type 1), as well as testosterone and androstenedione (usually by 17beta-HSD type 3). Therefore, these substances were tested as putative substrates for the recombinant zebrafish enzymes. The assay was performed with tritiated substances as described in Leenders et al. (1996).

Since the recombinant proteins were nearly always found in the pellet fraction, 20 µl of this fraction was added to the respective reaction mixture (see below). By the addition of NADPH in case of reduction, and addition of NAD in case of oxidation, the reaction was started and incubated with slight agitation for 30 min at 37°C. The reaction was stopped by addition of 100 µl stop-solution and samples afterwards preliminary purified on Strata C18-E columns (reverse phase extraction). Elution was carried out with two times 200 µl methanol.
The separation of the steroids in a 20 µl sample was performed through HPLC running an isocratic gradient H_2O/acetonitrile (52:48) on a reverse phase LUNA 5µm C18 (2) column (Phenomenex). The HPLC consisted of a Beckman Coulter system assembly composed of two HPLC pumps, a UV detector and was controlled with the 32Karat software (Beckman). The detection of the tritiated steroids proceeded with a HPLC radioactivity monitor after mixing with scintillation solution (Ready Flow III). The amount of the respective steroids was calculated through integration of the peaks in the HPLC spectra using the above mentioned software.
Each measurement was performed twice and in duplicates; errors were calculated as mean deviations.

Reaction mixtures:

for reduction:	430 µl sodium-phosphate buffer, 100 mM (pH 6.6) 20 µl bacterial lysate 0.5 µl substrate* 50 µl NADPH (5 mg/ml)
for oxidation:	430 µl sodium-phosphate buffer, 100 mM (pH 7.7) 20 µl bacterial lysate 0.5 µl substrate* 50 µl NAD (5 mg/ml)
*substrates:	estrone (2, 4, 6, 7-^{3}H (N)) (15nM) estradiol (6, 7- ^{3}H (N)) (21nM) androst-4-ene-3,17-dione (1, 2, 6, 7-^{3}H (N)) (13.5 nM) testosterone (1α, 2α-^{3}H (N)) (18.7 nM)
stop-solution:	0.21 M ascorbic acid in 1% acetic acid in methanol

2.4 Bioinformatic methods

2.4.1 *In silico* screen

Alternatively to methods of molecular biology, which use a labeled probe to screen cDNA or genomic libraries (compare chapter 2.1.5), searches for homologous sequences can also be performed *in silico*. For this, the probe is replaced by a DNA- or protein sequence that is compared to databases using a variety of algorithms. Here, the blast algorithms were employed

in search for putative zebrafish homologs and EST clones containing the full-length coding sequence or complementing known sequence parts.

On EST-databases
For the identification of zebrafish candidate genes, the mouse protein sequences were compared to the zebrafish EST database at NCBI (http://www.ncbi.nlm.nih.gov) using tblastn. EST sequences were chosen as putative homologs if either they were already annotated as similar to the respective 17beta-HSDs or had an alignment score >80 bits when their complete putative coding sequence was aligned to the respective 17beta-HSDs in the SwissProt database using blastx.

On genomic databases
Where screening EST-databases did not reveal putative zebrafish homologs, exons of mouse protein sequences were compared to the WGS database of *Takifugu rubripes* (fugu) at NCBI. Fugu-exons were checked by blastx against the SwissProt database and taken into account if the respective 17beta-HSD resembled the best match. Fugu-sequences were then compared against the WGS database of *Danio rerio* at NCBI to identify the homologous zebrafish exons.

2.4.2 *In silico* Northern blot

For analysis of RNA expression pattern complementing RT-PCR experiments and data from literature, *in silico* Northern blots were performed. For this, a given protein sequence was compared against the EST database of the organism of the sequence's origin at NCBI using tblastn. Results were screened for specific transcripts assuming that more than 90% sequence identity indicated congruence with the gene of interest. These ESTs were investigated for information about their origin concerning tissue, sex and developmental status; then they were counted according to these data. The resulting profile was indicative for the general level of expression as well as occurrence and amount of transcripts in certain tissues in relation to others.

2.4.3 Assemblies and gene structure

The assembly of coding sequences and construction of genomic sequences was conducted with AssemblyLIGN™ (Accelrys).
For identification of the gene structure, the full-length coding sequence of each zebrafish candidate gene was aligned against the zebrafish WGS database at NCBI using blastn. The resulting alignments revealed the exon-intron boundaries. Where necessary, boundaries were adjusted in accordance to the splice site consensus sequence (exon)-GT(intron)AG-(exon).

2.4.4 Alignments and phylogeny

Two protein or DNA sequences were compared to each other using either LALIGN (Pearson et al., 1997) at the Genestream Resource Center (http://www2.igh.cnrs.fr) or bl2seq at NCBI. For multiple sequence alignments, protein sequences were aligned by clustalW (Thompson et al., 1994). In any of these programs penalty gaps, word size and substitution matrix were usually kept at default values.
Multiple alignments, especially those calculated for phylogenetic analyses, were monitored and manually edited in BioEdit. The manual edition was necessary to check for sequence redundancy due to annotation errors. For phylogenetic analysis, data sets were created by retrieving related sequences from three different sources: (i) a psi-blast (Altschul et al., 1997) of the mouse protein sequences against the non-redundant protein database at NCBI; (ii) the BLink-link of the mouse protein entries in the NCBI database; (iii) a translated blast (tblastn) of the zebrafish 17beta-HSD candidates against EST-databases of *Ciona intestinalis*, *Caenorhabditis elegans* and *Drosophila melanogaster*. Sequences were processed in the same way as for direct comparisons and phylogenetic analyses subsequently conducted with *MEGA* version 2.1 ((Kumar et al., 2001) and references therein) using Neighborjoining (NJ) and unweighted pair group method with arithmetic mean (UPGMA) algorithms.

2.4.5 Calculation of Identity plots

Identity Plots were invented to view the distribution of identical residues in a multiple alignment along the linear sequence. As zebrafish and mammalian sequences shared a medium amount of these identical residues (about 50%), these often appeared in clusters outlining regions of higher conservation and therefore most likely important for structure and function of the enzyme.
First, a multiple alignment of the respective sequences was calculated with ClustalW. This was then manually transcribed into a binary code: 0 for non-identical positions, 1 for completely conserved positions. To simplify the method, only identical (and not conservative exchanges of) amino acids where taken into account. The identity was calculated in percent for windows of fifteen residues and plotted against the first position of this window.

2.4.6 3D-modeling of protein structure

To predict and analyze the 3-dimensional structure of some zebrafish proteins, models were calculated with SWISS-MODEL (Guex and Peitsch, 1997) including ProModII for model generation (Schwede et al., 2000) and Gromos96 (Stocker and van Gunsteren, 2000) for energy minimization. In this process, zebrafish linear amino acid sequences were first threaded onto a 3D model from the spdbv database (http://www.expasy.org/spdbv/), which had been selected due to sequence similarity, and this raw model sent to spdbv for refinements. The modeling involved loop and final energy minimization and was verified by PROVE (Pontius et al., 1996) and WHAT_CHECK (Hooft et al., 1996) packages. Homology models obtained in this

way were used to analyze cofactor and substrate binding sites and compare them with the known structures of the human protein.

2.4.7 Identification of terminal signalling motifs and globular domains

Comparison of the zebrafish proteins to other vertebrate homologs suggested the existence of putative signalling motifs in the N-terminal and C-terminal parts of the proteins. This hypothesis was addressed by comparison of the respective sequence parts against the database of eukaryotic linear motifs (ELM) at the ELM server (Puntervoll et al., 2003) (http://elm.eu.org/).

2.5 Work with organisms

2.5.1 Working with *E. coli*

Liquid culture

For plasmid preparation or protein production, *E. coli* was cultured in LB-medium at 37°C for a few hours or overnight in a shaking incubator. Except during transformation procedure, medium always contained either ampicillin or kanamycin as selective antibiotic, depending on the resistance gene carried by the bacteria's plasmid.

Agar plates

For selection of transformed bacteria or to obtain single clone colonies from glycerol stocks or agarose stabs, LB-agar plates containing the appropriate antibiotic were used. If vectors sustained blue-white screening, B/W-plates were preferred. In this process, vectors carry the gene for β-galactosidase which is synthesized upon IPTG induction and breaks up X-Gal. As a result, the respective colony is stained blue and denotes a bacterium carrying a vector without insert, as the β-galactosidase gene will be disrupted by successful cloning into the multiple cloning site.

Long-term storage

To store *E. coli* for longer periods, glycerol-stocks were prepared by mixing the liquid culture and sterile 80 % glycerol 1:1 in a storage tube, and kept at -80°C.

Antibiotics:	working concentration in media and plates:	ampicillin: 50 mg/l kanamycin: 50 mg/l

(added after cooling of the solution to at least 50°C)

B/W-plates: LB-plates (from storage): spread onto plate: 40 µl X-Gal (20 mg/ml in DMF)
8 µl IPTG (0.5 M)

LB-medium: 10 g BACTO Peptone
5 g Yeast Extract
10 g NaCl
in 1l H_2O (Milli-Q), autoclave

LB-plates: LB-medium + 10 g/l Bacto-agar, autoclave, pour plates, store at 4°C

2.5.2 Working with zebrafish

All work was carried out on AB wild type zebrafish from the zebrafish facility of the Institute of Developmental Genetics at the GSF-Research Center (Neuherberg, Germany). Adult fish as well as embryos were a kind gift of Dr. Laure Bally-Cuif.
Zebrafish embryos were staged in accordance to Kimmel et al. (1995). For RNA preparation, embryos were collected in 1.5 ml tubes and immediately frozen in liquid nitrogen after removal of medium.
Organs of adult AB wild type zebrafish (from 3 months up to 2 years) were dissected under the microscope, transferred to 1.5 ml tubes and frozen in liquid nitrogen. Composition of organ samples: brain (whole), gonads (total), skin (without scales), muscle (skeletal muscle), liver (mainly 1st and 3rd lobe), spleen (total), kidney (excluding most of the interrenal), heart (whole), intestine (total), eyes (whole).

3 Materials, Equipment, Programs

3.1 Materials

3.1.1 Antibodies

primary antibodies	
anti-GST	Zymed
anti-His-tag	Dianova
secondary antibodies	
peroxidase-coupled goat anti-mouse IgG	Dianova

3.1.2 Bacteria

The following *E. coli* strains were used for cloning and expression:

Top 10 (Invitrogen)
mcrA, delta (mrr-hsdRMS-mcrBC), phi 80delta lac delta M15, delta lacX74, deoR, recA1, araD139 delta (ara, leu), 7697, galU, galK, lambda^-, rpsL, endA1, mupG

JM107
endA1, gyrA96, thi, hsdR17, supE44, relA1, delta (lac-proAB) [F' traD36, proAB^+, lacI^q, lacZ delta M15]

DH5α (Life Technologies)
F'phi80d*lac*Z delta (*lac*ZYA-*arg*F)U169 *deo*R *rec*A1 *end*A1 *hsd*R17 (rk-, m k+) *pho*A *sup*E44 lambda-*thi*-1 *gyr*A96 *rel*A1/F' *pro*AB+ *lac*IqZdeltaM15 Tn10 (*tet*r)

E. coli strains provided by RZPD as hosts of ordered ESTs:

GeneHogs DH10B
F$^-$ *mcrA* Δ(*mrr-hsd*RMS-*mrc*BC) ϕ80*lac*ZDM15 Δ*lac*X74 *deo*R *rec*A1 *end*A1 *ara*D139 Δ(*ara-leu*)7697 *gal*U *gal*K λ *rps*L *nup*G

DH10B TonA
F$^-$ *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) ϕ80*lac*ZΔM15 Δ*lac*X74 *deo*R *rec*A1 *end*A1 *ara*D139 Δ(*ara, leu*)7697 *gal*U *gal*K λ^- *rps*L *nup*G *ton*A

XL1blue
recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F´ *proAB lacI*^{q} *Z*Δ*M15* Tn*10* (Tet^{r})]

XLOLR
Δ(*mcrA*)*183* Δ(*mcrCB-hsdSMR-mrr*)*173 endA1 thi-1 recA1 gyrA96 relA1 lac* [F´ *proAB lacI*^{q} *Z* Δ*M15* Tn*10* (Tetr)] Su^{-} (nonsuppressing) λ^{R} (lambda resistant)

3.1.3 cDNA-libraries (Colony Macroarrays from RZPD)

Zebrafish Brain cDNA (RZPD Lib.-No. 611)
Zebrafish Embryo (late somitogenesis) (RZPD Lib.-No. 524)
Zebrafish Liver cDNA (RZPD Lib.-No. 532)

3.1.4 Chemicals (special)

5-bromo-4-chloro-3-indolyl-beta-D-Galactopyranoside (X-Gal)	Biomol
Acrlyamide/Bisacrylamide (30%/0.8%)	Roth
Agarose	Biozym
Ammoniumperoxodisulphate (APS)	Merck
Ampicillin	Fluka
Ampuwa water	Fresenius
Bacto-agar	Difco
BACTO Peptone	Difco
Bovine serum albumin (BSA)	Biomol
Coomassie blue G250	Biomol
Diaminobenzidine (DAB)	Sigma
dNTPs	MBI Fermentas
Ethidium bromide	Sigma
Imidazol	Biomol
Isopropyl-beta-D-thiogalactopyranoside (IPTG)	DCL
Kanamycin	Sigma
Roti-phenol/chloroform	Roth
Sodium dodecylsulfate (SDS)	Merck
TEMED	Sigma
TRIzol	Invitrogen
Tween-20	Merck
Yeast extract	Nordwald

3.1.5 Enzymes

Endonuclease (Benzonase)	Sigma
Herculase DNA polymerase	Stratagene
Lysozyme	Merck
Proteinase K	Roche
Restriction enzymes	NEB or MBI Fermentas
Ribonuclease Inhibitor	MBI Fermentas
RQ1 RNase-free DNase I	Promega
Pfu Turbo DNA polymerase	Stratagene
T4-DNA-Ligase	NEB or MBI Fermentas
T4 polynucleotide kinase	MBI Fermentas

3.1.6 Kits

Gel Extraction Kit	Qiagen
Megaprime DNA Labeling System	Amersham
NucleoBond PC 100 Kit	Macherey&Nagel
Nucleotide Removal Kit	Qiagen
NucleoSpin Plasmid Kit	Macherey&Nagel
PCR Purification Kit	Qiagen
Prime-It RmT Random Primer Labeling Kit	Stratagene
RevertAid^{TM} First Strand cDNA Synthesis Kit	MBI Fermentas
RNeasy Midi Kit	Qiagen
RNeasy Mini Kit	Qiagen
TOPO-TA cloning Kit	Invitrogen

3.1.7 Other material

FluoroTrans® W Membrane (PVDF)	Pall
MicroSpin^{TM} S-200 HR columns	Amersham Pharmacia Biotech
Strata C18-E columns	Phenomenex
reverse phase LUNA 5u C18 (2) column	Phenomenex
XAR-5 (X-ray films)	Kodak

3.1.8 Radioactive molecules

^{32}P:

Redivue-α-^{32}P-dCTP	Amersham Pharmacia Biotech
Redivue-γ-^{32}P-dATP	Amersham Pharmacia Biotech

^{3}H:

androst-4-ene-3,17-dione (1, 2, 6, 7-^{3}H (N))	NEN
estrone (2, 4, 6, 7-^{3}H (N))	NEN
estradiol (6, 7- ^{3}H (N))	NEN
testosterone (1α, 2α-^{3}H (N))	NEN

3.1.9 Vectors

for cloning:

pCR2.1	Invitrogen
pCRII	Invitrogen
pQE 30	Qiagen
pET15b	Novagen

provided by RZPD as host vectors of ordered cDNAs:

pGEX 2T PL2	pT7T3D-PAC	pZipLox
pCMV-SPORT6	pAMP1	pBK-CMV
pME18S-FL3	pBS-SK	pSPORT1

3.2 Equipment

CEQ 2000 DNA Capillary Sequencer	Beckman
HPLC system assembly “32Karat Gold”	Beckman
Robo-Cycler	Stratagene
Mini-PROTEAN II - vertical electrophoresis cell	BioRad
Trans-Blot SD - Semidry Transfer Cell	BioRad
HPLC radioactivity monitor LB 506D	Berthold

3.3 Programs

AnalyzePPC (SoftGene)
Program for Macintosh system to administrate and analyze amino and nucleic acid sequences

AssemblyLIGN™ (Accelrys)
A tool for Macintosh system to facilitate the assembly of nucleotide sequences

BLAST-programs
BLAST = basic linear alignment search tool
Internet accessible tool for comparison of protein and nucleic acid sequences against databases or against each other (in case of bl2seq)

Employed BLAST-programs:
bl2seq: compares two sequences against each other using any of the given blast programs
blastn: nucleotides vs. nucleotides
blastp: proteins vs. proteins
blastx: translated nucleotides vs. translated nucleotides
psi-blast: iterative search of protein motifs vs. proteins
tblastn: translated nucleotide vs. proteins
http://www.ncbi.nlm.nih.gov

BioEdit
Free-download program (via internet) for the analysis of protein and nucleic acid sequences with stress on alignment and phylogenetics
http://www.mbio.ncsu.edu/BioEdit/bioedit.html

clustalW
Internet accessible tool for multiple sequence alignment of proteins or nucleic acids
http://www2.ebi.ac.uk/clustalw

ELM server
Free accessible internet server and database to investigate eukaryotic linear motifs in protein sequences
http://elm.eu.org/

LALIGN
Internet accessible tool to align two amino or nucleic acid sequences
http://www2.igh.cnrs.fr/bin/align-guess.cgi

MEGA (version 2.1)
Phylogenetic analysis of protein and nucleic acid sequences (free download from internet)
http://www.megasoftware.net

Swiss-Pdb Viewer
Free-download program (via internet) for visualization, analysis and modeling of protein 3D-structures
http://www..expasy.org/spdbv/

4 Results

Introductory remarks

All candidate genes were handled likewise analyzing them in comparison to the known mammalian genes and proteins but for each type of 17beta-HSD results are presented in a separate chapter. This arrangement was chosen to facilitate a better textual coherence though especially the introductory sections or references common to both types of 17beta-HSDs may be redundant. For each 17beta-HSD, results are presented in the chronology of DNA level, RNA level and finally protein level to allow a better overview and comparison although this might not always resemble the sequence in which results were obtained.

In **The candidate genes** a summary is given about how each single candidate sequence was identified, the process and change in naming them, clone and GeneBank identifiers and the composition of available clones concerning their coding area and untranslated regions.
Characterization on DNA level focuses on the gene structure. For all zebrafish genes the gene structure was investigated with emphasis on exon size and intron position. As intron size is a less conserved feature determination of exact sizes was of secondary interest. It was assumed that the coding nucleotide sequences of zebrafish in comparison to the mouse and human genes have diverged too much to hold any directly useful information, and therefore only alignments of the zebrafish cDNA sequences were investigated.
As the expression pattern may hint at the *in vivo* function of a gene, occurence of RNAs from the candidate genes was surveyed and results depicted in **Characterization on RNA level**. Differences between the sexes, developmental and adult stage as well as a number of organs was investigated.
Finally, **Characterization on protein level** looks at a variety of aspects connected with 1D and 3D protein structure of the zebrafish enzymes and analyzes them in comparison to the mammalian homologs. Protein sequence identity and characterization of conserved sites and motifs is monitored additionally to phylogenetic analyses, which place the zebrafish proteins into a capacious, evolutionary context. By laboratory means substrate specificity of the zebrafish proteins was investigated via expression of recombinant protein. In case of 17beta-HSD type 1, 3D-modelling was performed to assist in structure-function analysis.

4.1 17beta-HSD type 1

4.1.1 The candidate genes

An *in silico* screen of zebrafish EST-databases led to the identification of three putative candidate sequences for a zebrafish 17beta-HSD type 1 homolog. These were named due to the sequence in which they were discovered: zf 1.1, zf 1.2 and zf 1.3. Sequence complementation and detailed analysis (see following chapters) revealed the candidates to be 17beta-HSD type 1 (zf 1.2) and two photoreceptor associated retinol dehydrogenases, named type 1 (zf 1.3) and type 2 (zf 1.1). Here, they will be abbreviated and referred to as zfHSD 1, zfprRDH 1 and zfprRDH 2, respectively (for overview, see Tab. 2).

Tab. 2: Zebrafish candidates, identified and characterized in search of a homolog of mammalian 17beta-HSD type 1.

zebrafish homolog of	***acronym***	***original name***	***full-length clone***	***GeneBank***
17beta-HSD type 1	zfHSD 1	zf 1.2	LLKMp964D226Q2	AY306005
photoreceptor associated retinol dehydrogenase	zfprRDH 1	zf 1.3	IMAGp998K1110662Q3	AY306006
	zfprRDH 2	zf 1.1	zfHSD1.1-pGEX*	AY306007

The column “full-length clone” lists the names of clones that contain the full-length cds, which were either requested from RZPD or assembled molecularly (*) by laboratory means. Completely sequenced and analyzed sequences from this thesis were submitted to GeneBank for annotation.

The initial candidate sequences derived from *in silico* screen resembled only parts of the genes, and complete sequencing of requested clones was necessary to obtain the full-length coding sequence of each gene. In other cases, a clone containing the complete cds was not available and the respective construct had to be assembled by molecular means on the basis of database sequence data. As this procedure and its results were unique for each individual gene, they are outlined in the following paragraphs. The resulting clones, which were the base of all subsequent laboratory analyses, are depicted in a graphic overview in Fig. 6.

zfHSD 1

By its clone Id (IMAGE: 4725861), the clone LLKMp964D226Q2 was ordered from RZPD and its identity confirmed by PCR. Sequencing revealed an insert of 2543 bp, consisting of 68 bp 5’UTR, a full-length coding sequence of 888 bp and 1587 bp 3’UTR closing with a poly-A tail.

zfprRDH 1

The clone IMAGp998K1110662Q3 was identified via the candidate sequence's clone Id (IMAGE: 4789378) and obtained from RZPD. The insert of the clone comprised 2097 bp: 140 bp 5'UTR, the 954 bp full-length coding sequence and 1003 bp of 3'UTR including a poly-A tail.

zfprRDH 2

Clone MPMGp609M0716Q8 was requested from RZPD by the candidate sequence EST name (fb79a04.y1) and confirmed by PCR. The 1476 bp insert comprised 66 bp of non-coding sequence, an open reading frame of 837 bp and 573 bp 3'UTR ending on a poly-A tail. Another EST of zfprRDH 2 (GenBank: BG307645) overlapped in its 3'part with the insert of MPMGp609M0716Q8 but additionally elongated the 5' coding area by 26 codons and provided another 5'UTR. As this sequence contained more than half of exon 1 including the cofactor binding site, it was considered to be part of the full-length coding sequence of zfprRDH 2. As this clone was not available from RZPD, the respective but missing sequence piece in the 5' region was amplified by RT-PCR from zebrafish ovary total RNA. This fragment was joined to the insert of clone MPMGp609M0716Q8 via fusion PCR resulting in the 957 bp full-length coding sequence which was subcloned into pGEX vector.

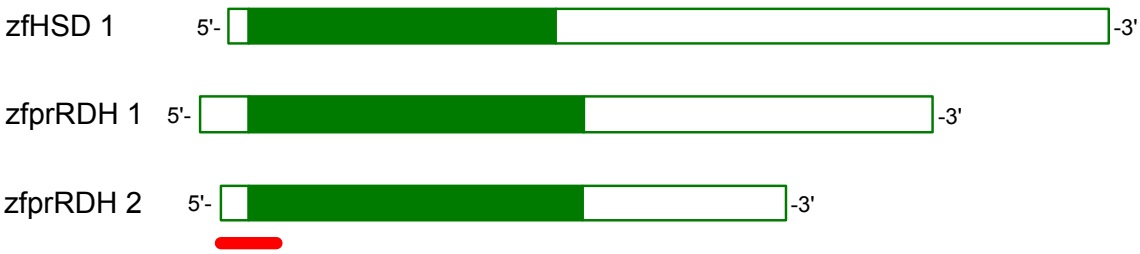


Fig. 6: cDNA structures of 17beta-HSD type 1 and the two prRDHs from zebrafish.
The coding sequence is depicted in green whereas the 5'- and 3'-UTRs are shown as white boxes. The red bar marks the area which was added to the original clone by fusion PCR (see description for zfprRDH 2 for details).

4.1.2 Characterization on DNA level

4.1.2.1 Coding sequence

As the three zebrafish candidate sequences appeared to be closely related to each other, their coding sequences were aligned and compared. The identity matrix (Tab. 3) revealed that less than 50% of the nucleotides were conserved between zfHSD 1 and each of the two zfprRDHs in all three sequences. In contrast, zfprRDH 1 and zfprRDH 2 share much higher conservation displaying 73% identical positions.

Tab. 3: Identity matrix of the coding sequences of the zebrafish 17beta-HSD type 1 candidate genes.

	zfHSD 1	***zfprRDH 1***	***zfprRDH 2***
zfHSD 1	100%		
zfprRDH 1	46%	100%	
zfprRDH 2	48%	73%	100%

The matrix was calculated with BioEdit following sequence alignment by ClustalW.

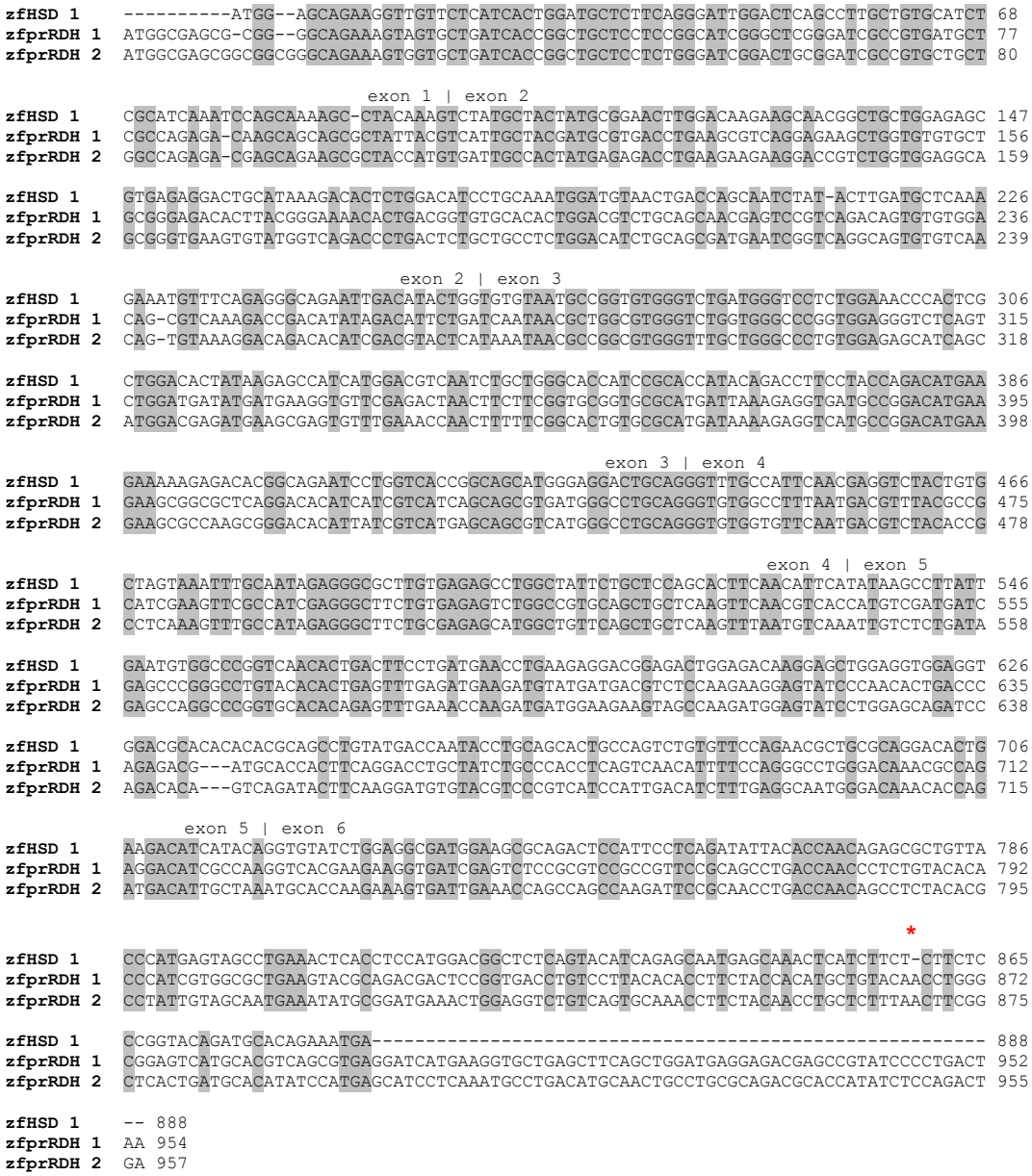


Fig. 7: Alignment of the full-length coding sequences of zfHSD 1 and the two zfprRDHs.
Nucleotides identical in all three sequences are shaded in gray. Dashes indicate gaps that were inserted to facilitate better alignment. Above the alignment, the sites where in the genomic sequence introns are introduced are outlined. The red asterisk marks the position in zfHSD 1 where by deletion a frame-shift might have occurred, leading to a shortened last exon in comparison to the two zfprRDHs. Alignment was performed by ClustalW using default word size and gap penalty values.

The sequence alignment (Fig. 7) shows that the number of identical nucleotides is significantly lower in the 3'-part, comprising about the last two-thirds of exon 5 and the complete exon 6. While the two prRDH-sequences share a similar overall length, zfHSD 1 is shorter by roughly 65 nucleotides. The alignment suggests that this shortening might have resulted from a frame-shift in the last exon leading to a premature stop at a conserved TGA.

4.1.2.2 Exon structure

Good indicators for the homology of two genes are the number and individual size of exons which are determined by the position of introns in the gene. The relative exons were identified as described in Methods (see chapter 2.4.2). From the three zebrafish genes as well as the homologous mouse and human genes the features of the exons are depicted in Tab. 4.

Both groups, HSDs and prRDHs, display a highly similar overall gene structure in which the coding sequence is partitioned into six exons of characteristic sizes (Tab. 4). When the zebrafish genes are compared to the respective murine and human genes it is apparent that exon size is very well conserved. Deviations occur only in certain exons and are restricted to an increase or decrease by exact triplets. Exons 3 and 4 have identical sizes in both 17beta-HSDs type 1 and the photoreceptor associated retinol dehydrogenases. In addition, the size of the first exon seems to be characteristic for the group of 17beta-HSDs whereas in case of the retinol dehydrogenases exon 5 is strictly conserved. The last exon displays the lowest degree of conservation, which is very prominent for the 17beta-HSDs, generally displaying a higher degree of variation.

Tab. 4: Exon sizes of the zebrafish candidate genes in comparison to the respective mouse and human homologs.

exon:	***1***	***2***	***3***	***4***	***5***	***6***
17beta-HSD type 1:						
zfHSD 1	97	162	180	94	187	168
mouse	~	+6	~	~	-12	+102
human	~	+6	~	~	-9	+153
photoreceptor associated retinol dehydrogenase:						
zfprRDH 1	106	162	180	94	184	228
zfprRDH 2	+3	~	~	~	~	~
mouse	-3	~	~	~	~	+3
human	-3	-3	~	~	~	-12

For the first sequence in each group, the individual exon size is given in nucleotides. The exon sizes of subsequent genes are depicted in comparison to the first sequence and deviations are given by +/- the respective number of nucleotides; ~ indicates concordant exon size.

The 17beta-HSD type 1 genes of zebrafish, mouse and human differ in nucleotide numbers of exons 2, 5 and 6. Whereas an elongation of exon 2 by six nucleotides seems to be characteristic for both mammals, changes in the last two exons are individual though still follow the same direction: compared to the zebrafish gene, exon 5 is shortened and exon 6 enlarged. The low size conservation in the last exon is especially apparent, suggesting a displacement of the stop-codon rather than insertion or deletion of triplets as was observed for other variations in exon size (see also chapter 4.1.4.1.3, Fig. 13).
In contrast, conservation of exon sizes in the photoreceptor associated retinol dehydrogenases is very strict. There are only few changes in exon architecture, which could be caused by speciation. Whereas the first exon is slightly shortened by three nucleotides in both mammals, a similar event has taken place for the second exon in the human gene alone. Variations in the last exon might have appeared independently leading to a slight increase by three nucleotides in mouse compared to the loss of twelve nucleotides in the human gene. The two zebrafish genes appear to be even more similar to each other than to their mammalian homologs. zfprRDH 2 displays a gene structure identical to zfprRDH 1 with merely a first exon increased by one triplet. This might suggest the origin of the two genes from a duplication in the prRDH gene in zebrafish.

4.1.2.3 Gene and intron sizes

The architecture of all zebrafish genes and the respective mouse and human homologs from this group were determined as described in Methods (see chapter 2.4.2), and a scheme of all genes was generated (see Fig. 8).

It becomes clear that overall gene size and especially intron size is conserved between mouse and human rather than between zebrafish and mammals. The gene size of the mammalian 17beta-HSDs type 1 as well as of the retinol dehydrogenases is quite similar and slightly smaller compared to the respective zebrafish genes. In addition, inside one species 17beta-HSDs are more condensed than the retinol dehydrogenases. It is also interesting to notice that the second zebrafish gene for the photoreceptor associated retinol dehydrogenase (zfprRDH 2) is about four times the size of zfprRDH 1 and hence, with roughly 30 kb, the largest gene in the whole group.

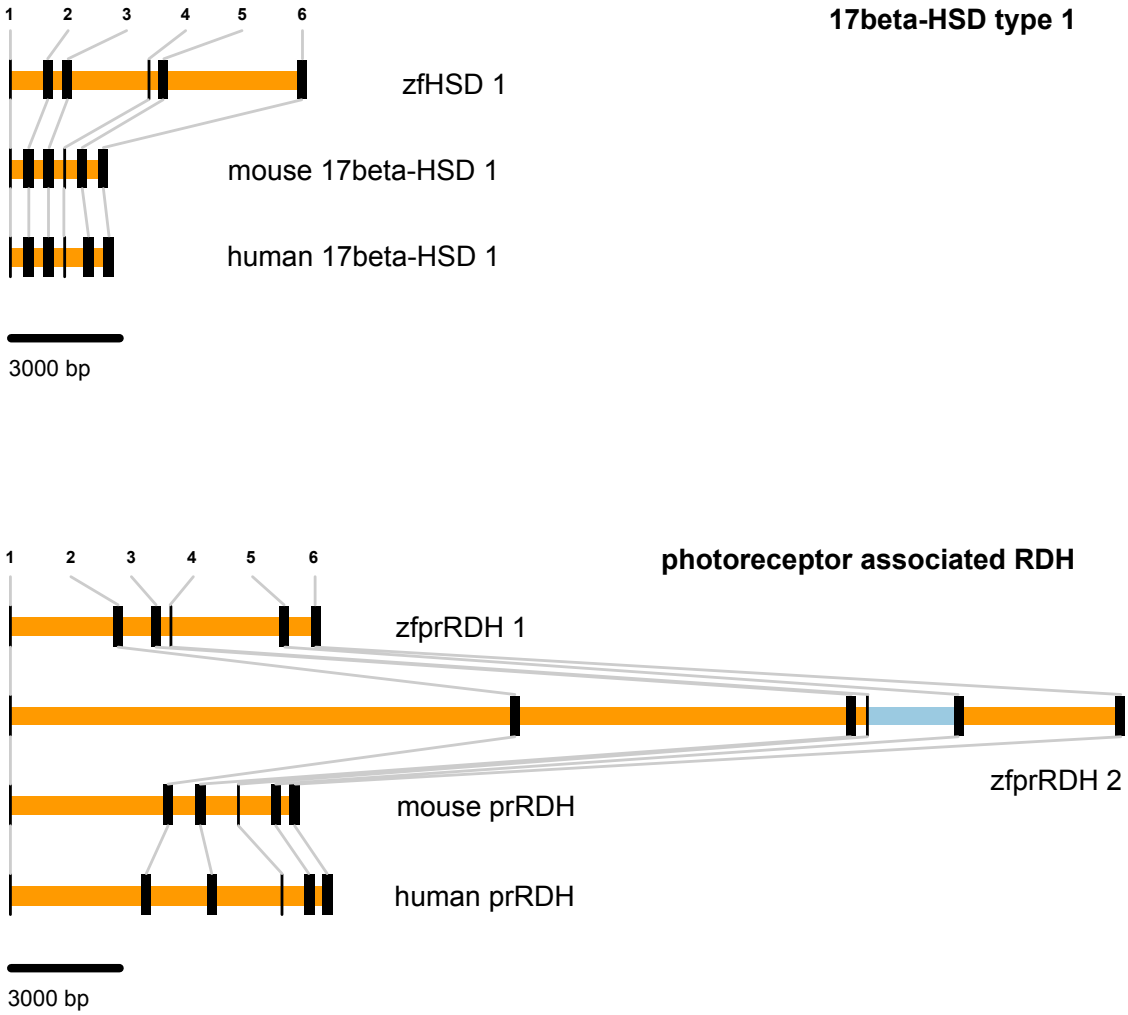


Fig. 8: Gene structure of zebrafish, mouse and human 17beta-HSD type 1 and photoreceptor associated retinol dehydrogenases.
Coding exons are shown as black bars; introns of confirmed size are in orange. The blue intron in zfprRDH is depicted in a minimum size measured from WGS sequences and their assemblies as no contig information was available to determine its exact size.

4.1.3 Characterization on RNA level

To characterize the expression of the various zebrafish candidate genes, RT-PCR was performed on organs of adult wild type fish and on several stages of embryonic development (for method, see chapters 2.2.3 and 2.5.2). In addition, data about gene expression were gathered by *in silico* Northern.

4.1.3.1 Expression analysis by RT-PCR

All three zebrafish genes display characteristic expression patterns during embryogenesis as shown in Fig. 9. While zfHSD 1 could not be detected in any of the stages from sphere up to 84 hpf, zfprRDH 2 was present throughout this period. Although the applied experimental setup is

not suited for quantitative conclusions, it appears as if expression enhanced during somitogenesis and remained on this level thereafter. Late somitogenesis is also the earliest developmental stage where zfprRDH 1 could be first detected by RT-PCR. From 14 somite stage onwards expression is present and may even increase slightly during subsequent embryogenesis.

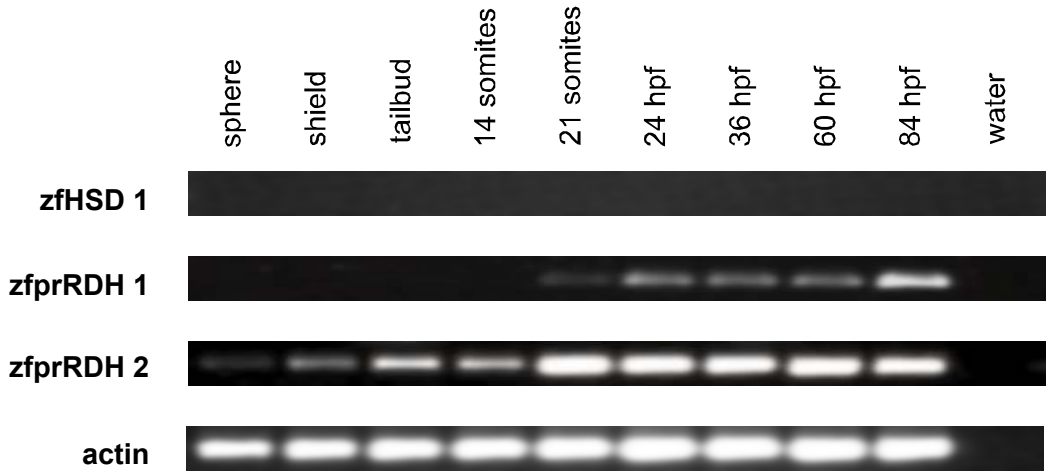


Fig. 9: Expression pattern of zebrafish 17beta-HSD type 1 and the two retinol dehydrogenases during embryonic development.
The presence of specific transcripts in total RNA preparation from developmental stages ranging from sphere (blastula stage) to 84 hpf (early larvae) was investigated as described in chapters 2.2.1 and 2.2.3. To monitor reliability of PCR, water controls were always included. To allow for comparison of signal strength, expression of actin was measured in the same way. hpf: hours post fertilization.

Specific expression profiles for the three zebrafish genes were also observed in the adult fish (Fig. 10). In case of zfHSD 1, no expression could be detected in male adult fish while in females a strong signal was produced by gonads in addition to some other female tissues such as heart, skin, muscle and eyes. zfprRDH 1 transcripts were exclusively detected in brain and eye in addition to male and female gonads; no sex-specific differences could be seen. zfprRDH 2 expression was readily detected in all monitored organs. In some organs, namely brain, muscle, liver, spleen and intestine, expression in male fish might be stronger as compared to female fish.

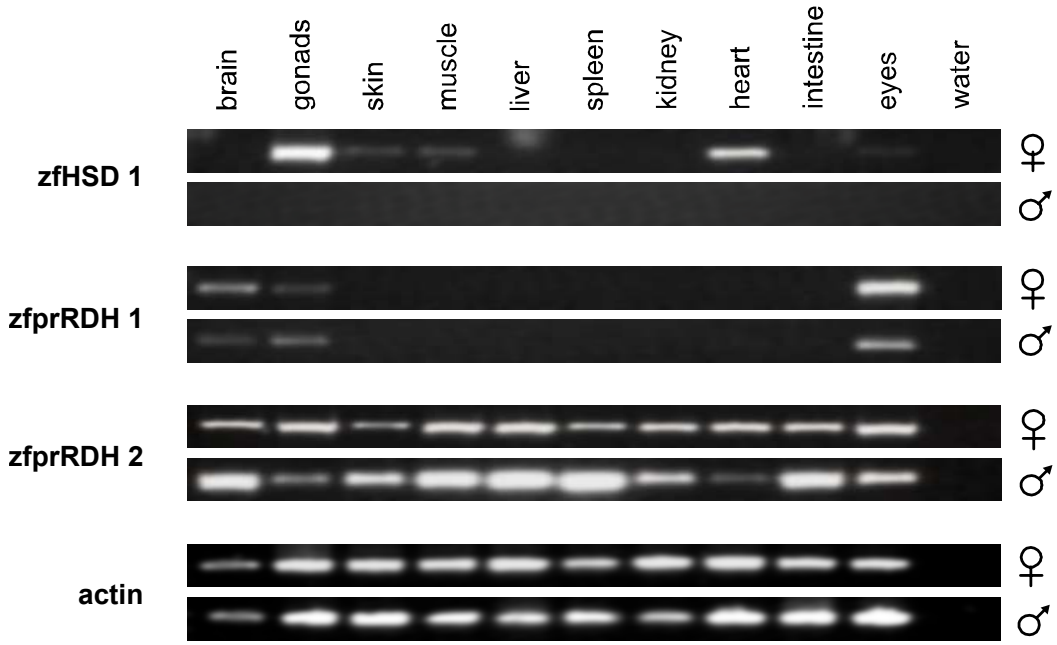


Fig. 10: Expression pattern of zebrafish 17beta-HSD type 1 and the two photoreceptor associated retinol dehydrogenases in tissues of adult male and female fish.
To monitor reliability of PCR, water controls were always included. To allow for comparison of signal strength, expression of actin was measured in the same way. For sample composition see chapter 2.5.2.

4.1.3.2 Expression analysis by *in silico* Northern blot

Database information are becoming more and more comprising not only for human and mouse but as well for zebrafish and other organisms. Therefore, *in silico* Northern blot (for method see chapter 2.4.2) was employed to corroborate and expand expression information about the three zebrafish genes. In addition to the source of the transcript (organ, tissue or developmental stage), these data give some information about the expression level by the general number of ESTs present in the databases.

Fig. 11 shows the analyzed data received from the *in silico* Northern. In general, only few ESTs of the three zebrafish genes could be identified, which is quite drastic for zfprRDH 1 (two transcripts); the highest number of ESTs was found for zfprRDH 2 (nine transcripts). This already indicates that zfprRDH 2 transcripts are more abundant than that of zfprRDH 1 while zfHSD 1 expression is intermediate in comparison to the two retinol dehydrogenases. As ESTs from all three genes predominantly stem from libraries prepared from whole fish or pooled tissue, the reason is probably either a moderately high but ubiquitous expression (zfprRDH 2) or a relative high expression in one or few organs (zfHSD 1, zfprRDH 1). These data very well reflect the situation identified by RT-PCR.

Another concurrence between *in silico* Northern and RT-PCR is found concerning the organs where zfprRDH 2 is expressed. Both methods indicate expression of this zebrafish gene in brain and eyes. In case of zfHSD 1, *in silico* Northern blot identified a specific EST in a library derived from kidney where by RT-PCR no transcript was detectable.

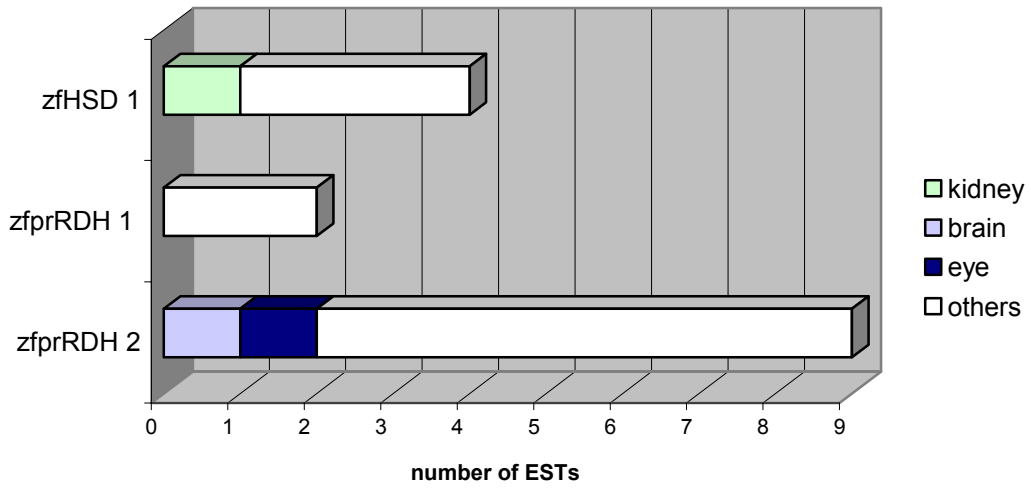


Fig. 11: Graphic analysis of *in silico* northern data retrieved for zfHSD 1 and the two zfRDHs.
For each gene the total number of EST hits is depicted. The composition of library sources is given in a color code. The group “others” includes all libraries from pooled or unspecified origins.

4.1.4 Characterization on protein level

4.1.4.1 1D-structure

4.1.4.1.1 Overall sequence identity

Upon completion of sequencing, the full-length cds of the candidate genes were translated by use of the genetic standard code, and the resulting amino acid sequences analyzed for overall identity.

Comparison of each full-length coding sequence with all other sequences of the set was carried out to establish an identity matrix as depicted in Tab. 5. This matrix points out the degree of similarity between the zebrafish and the mammalian sequences as well as between the two functionally different groups of HSDs and RDHs. In case of 17beta-HSD type 1, the zebrafish protein shows a slightly increased sequence identity to other proteins of this group (~50%) than compared to the retinol dehydrogenases (~40%), as should be expected for a true homolog. The murine and human 17beta-HSDs type 1 are less divergent, sharing 70% identical residues. Interestingly, this value is even higher for the two mammalian photoreceptor associated retinol dehydrogenases (77%) reflecting a generally more stringent conservation in this enzyme family. As was the case for zfHSD 1, both zebrafish proteins of the prRDH group share a higher degree of identity with their mammalian homologs (zfprRDH 1: 51%; zfprRDH 2: 48%) than with the 17beta-HSD type 1 family (zfprRDH 1 and 2: about 41%).

Tab. 5: Identity matrix of the proteins of zebrafish, mouse and human 17beta-HSD type 1 and photoreceptor associated retinol dehydrogenases (prRDHs).

17beta-HSD type 1:							
zfHSD 1	100						
mouse	51	100					
human	48	70	100				
prRDHs:							
zfprRDH 1	41	39	44	100			
zfprRDH 2	41	38	43	70	100		
mouse	37	36	40	49	46	100	
human	39	41	44	53	50	77	100
	zfHSD 1	mouse	human	zfprRDH 1	zfprRDH 2	mouse	human
	17beta-HSD type 1			**prRDHs**			

Pairwise alignments were performed with Blast2seq; identities are listed in percent.

Furthermore, a close relation between the two zebrafish prRDHs is apparent (70% identity). Sequence identity for zfprRDH 1 to the mammalian prRDHs is slightly higher than for zfprRDH 2. This might indicate zfprRDH 1 to be closer related to mammalian prRDHs but it is not clear whether this finding also reflects on functional similarity. That the two functionally diverse groups of 17beta-HSDs and prRDHs are still closely related is reflected in the comparison of non-homologous sequences. Here, a mean value of 40% sequence identity exists which is much higher as for SDRs that arose by convergent evolution not sharing a common ancestor.

4.1.4.1.2 Phylogenetic analysis

The evolutionary context, in which the zebrafish proteins are placed, was investigated by phylogenetic analysis. For this, a set of related sequences was collected as described in chapter 2.4.4. The phylogenetic analysis was considered to give the most reliable information about homology to the mammalian proteins, concerning evolution from a common ancestor. Results from this analysis were the base for classification of the zebrafish candidate genes.

A phylogenetic tree (Fig. 12) was calculated by means of Neighborjoining that shows a clear separation of the group of 17beta-HSDs type 1 and photoreceptor associated retinol dehydrogenases from all other sequences in this data set. The most distantly related group was considered to be formed by the microbial oxidoreductases and therefore the tree was rooted at this dichotomy. Sequences not belonging to HSDs, RDHs or outgroup form only two subgroups supported by bootstrap values: an “RDH similar group” and an “ancient group”. Both clusters contain sequences from *C. elegans* as well as vertebrates, suggesting an ancient origin and most likely distant relation to 17beta-HSD type 1 and prRDH. Since for none of the corresponding proteins/genes from these subgroups were data available in the literature or data

banks, concerning their character and function, a more detailed analysis including functional aspects was not possible.

Separation of 17beta-HSDs type 1 (orange) and photoreceptor associated retinol dehydrogenases (blue) in two groups is clearly supported by bootstrap values of 100%. This also means affiliation of zfHSD 1 to the former group whereas the two zebrafish prRDHs appear as paralogs of the latter group. Concerning the 17beta-HSDs type 1 family, the zebrafish sequence forms a separate subgroup together with the eel protein, which is located in direct vicinity to the chicken protein, followed by the mammalian enzymes. This sequence reflects the general evolution of vertebrates and is indicative for the good quality of the phylogenetic calculation.

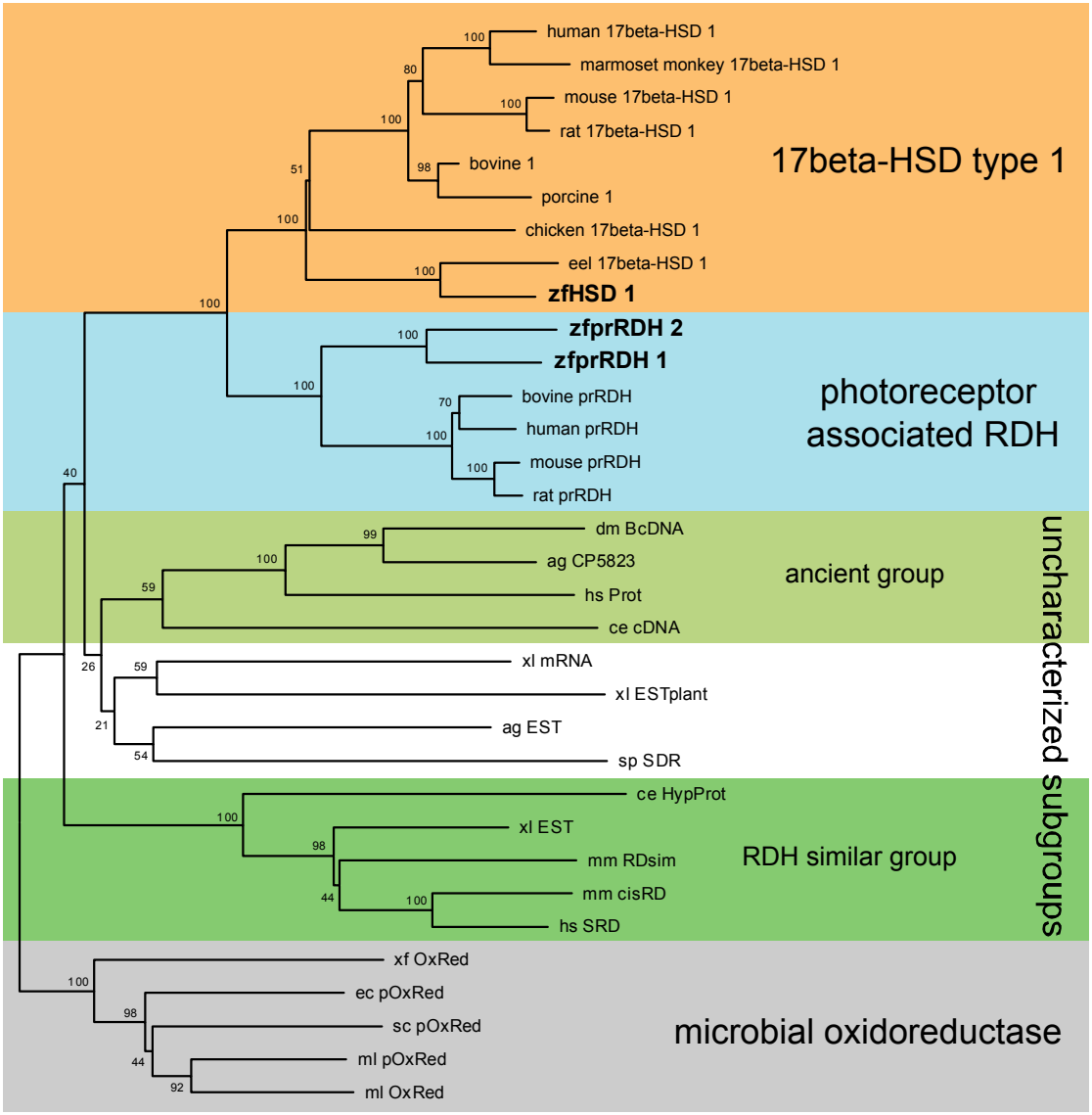


Fig. 12: Evolution of zebrafish 17beta-HSD type 1 and the two photoreceptor associated RDHs.
The family of 17beta-HSDs type 1 is marked in orange, the group of prRDHs in blue. Additionally, two uncharacterized subgroups of related sequences are shown in green. The tree was calculated with Neighborjoining algorithm and rooted, assuming microbial oxidoreductases to form the outgroup (gray). Bootstrap values of a 1000 pseudoreplicates are given in percent at the individual bifurcations of the tree. The protein sequences employed for this calculation are listed in the appendix (chapter 6.3).

Of the photoreceptor associated retinol dehydrogenases only few sequences from databases were available for phylogenetic calculations. Though both zebrafish proteins are assumed to be homologs of the mammalian prRDHs, it is not clear whether any of them is a functional ortholog. Location of both zebrafish sequences on a separate branch rather than affiliation of the one putative ortholog to the mammalian group seems to argue for a fish specific duplication. More prRDH sequences from non-mammalian organisms would be necessary for a detailed dissection.

4.1.4.1.3 Conserved motifs

All proteins analyzed in the context of this work belong to the superfamily of short chain dehydrogenases/reductases (SDRs). The presence and level of conservation of characteristic SDR motifs (see also chapter 1.3.2) was analyzed based on the alignment of the three zebrafish sequences together with their murine and human homologs (Fig. 13). These investigations were aimed at the assessment of functional integrity of the zebrafish proteins. In addition, amino acids that might have become typical for the function of either 17beta-HSDs type 1 or photoreceptor associated retinol dehydrogenases were explored.

Analysis of the alignment of all three zebrafish genes together with their respective human and mouse homologs revealed amino acids identical in both groups, 17beta-HSDs type 1 and prRDHs (Fig. 13). Mostly, these resemble structural and functional features common to all SDR family members. As could be expected for functionally diverse SDRs, the overall identity for the complete sequence set is with 23% relatively low. Furthermore, nearly 90% of the conserved residues are present in the N-terminal part of the protein, which forms the domain that conveys the binding of the cofactor. The seven SDR-motifs essential for structural stability, alignment and binding of the cofactor, and set up of the active site could be identified in all sequences.

SDR-motifs exerting alignment and binding of the cofactor

Motifs (1) and (2) both function in binding of the cofactor and are highly conserved. With the exception of threonine instead of serine in the two mammalian retinol dehydrogenases, the conservation of the central amino acids of motif (1) is even extended to a stretch comprising eleven completely conserved amino acids and demonstrating the close relation between the 17beta-HSDs type 1 and prRDHs. In motif (2), aspartic acid is present in all sequences, as is the subsequent hydrophobic residue in complete consistence with this theme.

Motif (3) plays a structural role in stabilising the central β-sheet. The most important part of this theme is the NAG, which indeed is present in all seven proteins. Apart from that, two non-conservative exchanges appear: (i) The first asparagine of NNAG is in all 17beta-HSDs type 1 substituted by a conserved cysteine, clearly separating them from the retinol dehydrogenases. (ii) Both zebrafish RDHs display an arginine at the position of the otherwise conserved glycine at the beginning of the motif. Interestingly, the mouse retinol dehydrogenase harbors an additional amino acid in this theme, increasing the number of hydrophobic residues between aspartic acid and the first asparagine from three to four.

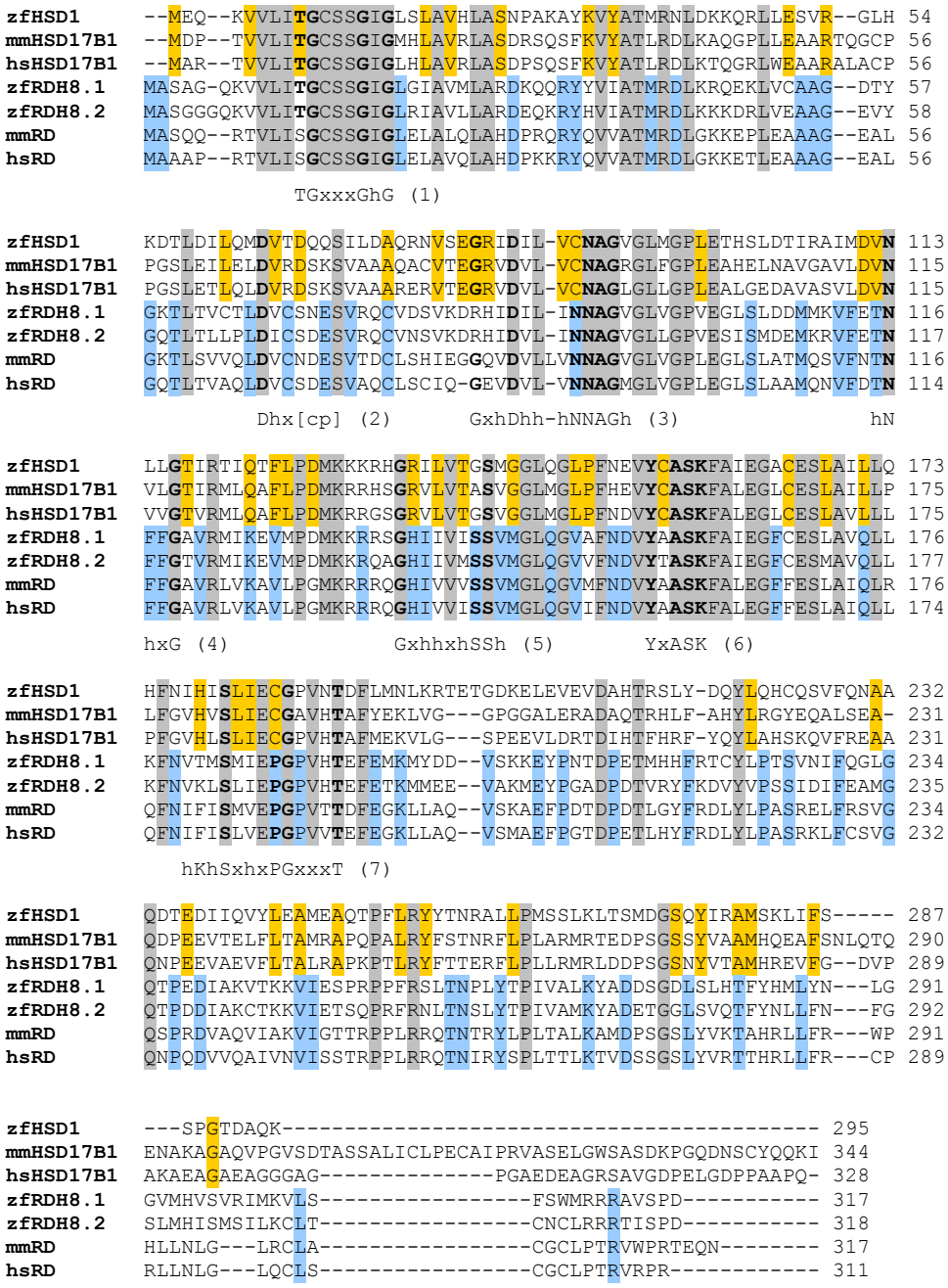


Fig. 13: Simultaneous comparison of SDR motifs in zebrafish and mammalian 17beta-HSDs type 1 and prRDHs proteins.
Residues identical in all sequences are shaded in gray. 100% conserved amino acids are marked in orange in the group of 17beta-HSDs type 1 and shaded in blue in the group of prRDHs. The consensus sequence of SDR motifs is given underneath the alignment; capital letters address the central, most conserved residues and small letters an extended version of the pattern with more flexible amino acid identity. Amino acids in the respective protein sequence consistent with central residues of the consensus are written in bold letters. Calculation of the alignment was performed with ClustalW; dashes indicate gaps that were introduced to facilitate better alignment. c: charged residue; h: hydrophobic residue; p: polar residue; x: any residue.

SDR-motifs involved in modeling of the active site

Motifs (4), (5) and (6) are part of the active site and therefore important for the stabilization of the reactive molecule parts in cofactor and substrate and the transfer of the hydride. The central

amino acids of motif (4), namely asparagine and glycine, are conserved in all sequences. Whereas the 17beta-HSDs also fit the extended version of this theme, the retinol dehydrogenases show a highly preserved displacement of the first hydrophobic residue by threonine. Although the first of the two serine residues in motif (5) is present only in the retinol dehydrogenases, all sequences fit this theme very well. In addition to some amino acids specifically conserved in one or the other group, a valine is common to all sequences shown here and might be typical for these closely related groups, as in general this residue can be any amino acid (compare also to motif (1) above). Tyrosine and lysine of motif (6) are completely conserved for all sequences, highlighting their role in transfer of the hydride ion in the catalytic triad. Furthermore, two more central residues (alanine and serine) are present and identical in the retinol dehydrogenases as well as the 17beta-HSDs.
Motif (7) is the last theme common to all SDRs, delineating the end of the primarily cofactor binding part of the enzyme. The main chain carbonyls of proline and glycine as well as the threonine interact with the cofactor. Whereas the first of these residues is conserved only in the retinol dehydrogenases, the other two amino acids appear to be identical for all sequences investigated here. Serine, another residue of the extended theme, is also conserved whereas the predicted lysine is present only in zfprRDH 2 and therefore might be less important for functional integrity. Common to all sequences here are a glutamic acid and a valine residue at positions where the pattern theoretically allows for any amino acid. This might again reflect the evolutionary close relation between the two groups similar to situations in motifs (1) and (5) (see above).

Conservation in 17beta-HSDs type 1 and prRDHs concerning their N- and C-terminal protein parts

When the alignment is split into the two different groups, hydroxysteroid and retinol dehydrogenases, additional identical amino acids typical for each family can be detected. The overall sequence identity between all family members rises to 43% in case of the 17beta-HSDs and to 45% for the RDHs. Residues now adding to the increased similarity are distributed all over the protein sequence but interestingly still seem to be in higher abundance in the N-terminal, mainly cofactor binding part (encompassing the sequence between motifs (1) and (7)). Before, about 90% of identical residues reflecting common SDR-themes were located in the N-terminal two-thirds of the protein due to the fact that these motifs solely influence cofactor binding and overall structure. Interestingly, the contribution of the group-specific amino acids differs between 17beta-HSDs and RDHs. In the first group 21% of residues in the cofactor binding part and 14% of residues in the substrate specific part (C-terminus starting from the end of motif (7)) are identical; in the latter group identities are 22% and 20%, respectively. This would mean that both groups show a similar increase in conservation in the predominantly cofactor binding part whereas the increase in specificity for the substrate binding part is about a third lower for the 17beta-HSDs.
It is known that the substrate specificity is not identical between human and mouse 17beta-HSDs and may lead to the observed lower conservation. If the investigation is narrowed down, now looking at a direct comparison between the zebrafish and each of the mammalian sequences, values for specific residues change to 26%/22% (cofactor-/substrate part) for comparison to human and to 27%/22% (cofactor-/substrate part) for comparison to mouse. Though the imbalance still remains, the difference is notably leveled out, decreasing from one-third to about one-seventh.

4.1.4.1.4 Distribution of identical residues

The investigation of the multiple alignment led to the emergence of several interlaced patterns of conservation. Though identical residues could readily be identified, the context of protein structure (reflected in the SDR-motifs), functionally different parts (cofactor vs. substrate binding) and the course of distribution of identical amino acids was difficult to interpret from this presentation. To better depict and investigate characteristics of conservation, the identity of individual amino acids was masked by transferring the alignment into a binary code, block-wise calculation of identity in percentage and plotting these values against the linear sequence (for method see chapter 2.4.4). The resulting Identity plots (Fig. 14) are analyzed in this chapter.

Progression of conservation: comparing general features typical for 17beta-HSDs type 1 and prRDHs

Both sets of Identity plots show unique patterns of identities that in themselves vary slightly, depending on the sequences selected for comparison. Although both 17beta-HSDs and retinol dehydrogenases display peaks of conservation where SDR-motifs are present, this mainly cofactor binding part shows dissimilar modes of conservation for the two enzyme groups.

In the 17beta-HSDs type 1, motifs (1) and (3) appear as rather isolated peaks of conservation. The area connecting these two themes is characterized by a deep rift of significantly lower sequence identity which reaches its minimum around positions 40-50. Although motif (2) is also a part of this region of relatively low conservation, no clear peak is detectable at the respective position. Instead, its position corresponds to a small plateau before percent amino acid identity rises to comprise motif (3). Subsequent to this theme percent identity drops again and reaches the second minimum of the predominantly cofactor binding part around position 105.

In contrast, the Identity plot of prRDHs displays only one stretch of low conservation located between motifs (2) and (3) where in 17beta-HSDs a steady increase in identity is apparent. In addition, two more peaks with about 50% identity can be discerned in the area between motifs (1) and (3), with the second peak comprising motif (2). This part of intermediate conservation corresponds to the rift of low conservation in the 17beta-HSDs.

The sequence part including motifs (4)-(7) conveys positioning of the cofactor and the substrate in the vicinity of the catalytic center and is much more similar between 17beta-HSDs type 1 and the prRDHs. In both groups, motifs (5) and (6) form a broad peak of sequence identity that reaches maximal conversation of about 80% (HSDs) to 90% (prRDHs). Motif (4) forms a slightly separate peak which is more or less discernable depending on the sequences which were compared. This is similar for motif (7): while in the 17beta-HSDs this SDR-motif is rather connected to the preceding sequence, indicating a consistent level of residue identity, this is for prRDHs only the case when zf 2 (prRDH 2) is left out of the alignment (see also comments below).

The sequence area subsequent to motif (7) is most important for substrate recognition and binding further away from the catalytic site. Both enzyme groups display divergent distributions of identical residues that might reflect their different substrate specificity. In the 17beta-HSDs, sequence identity falls off rapidly, reaching 0% identity when all three sequences are taken into account. In the following part, conservation recovers, showing 30% intermediate sequence identity. In contrast, prRDHs lack this strong and sudden drop in conservation subsequent to motif (7) and instead are marked by a rather slow and steady decrease up to position 275. The intermediate 30-40 % amino acid identity in prRDHs only slightly exceeds that in 17beta-HSDs.

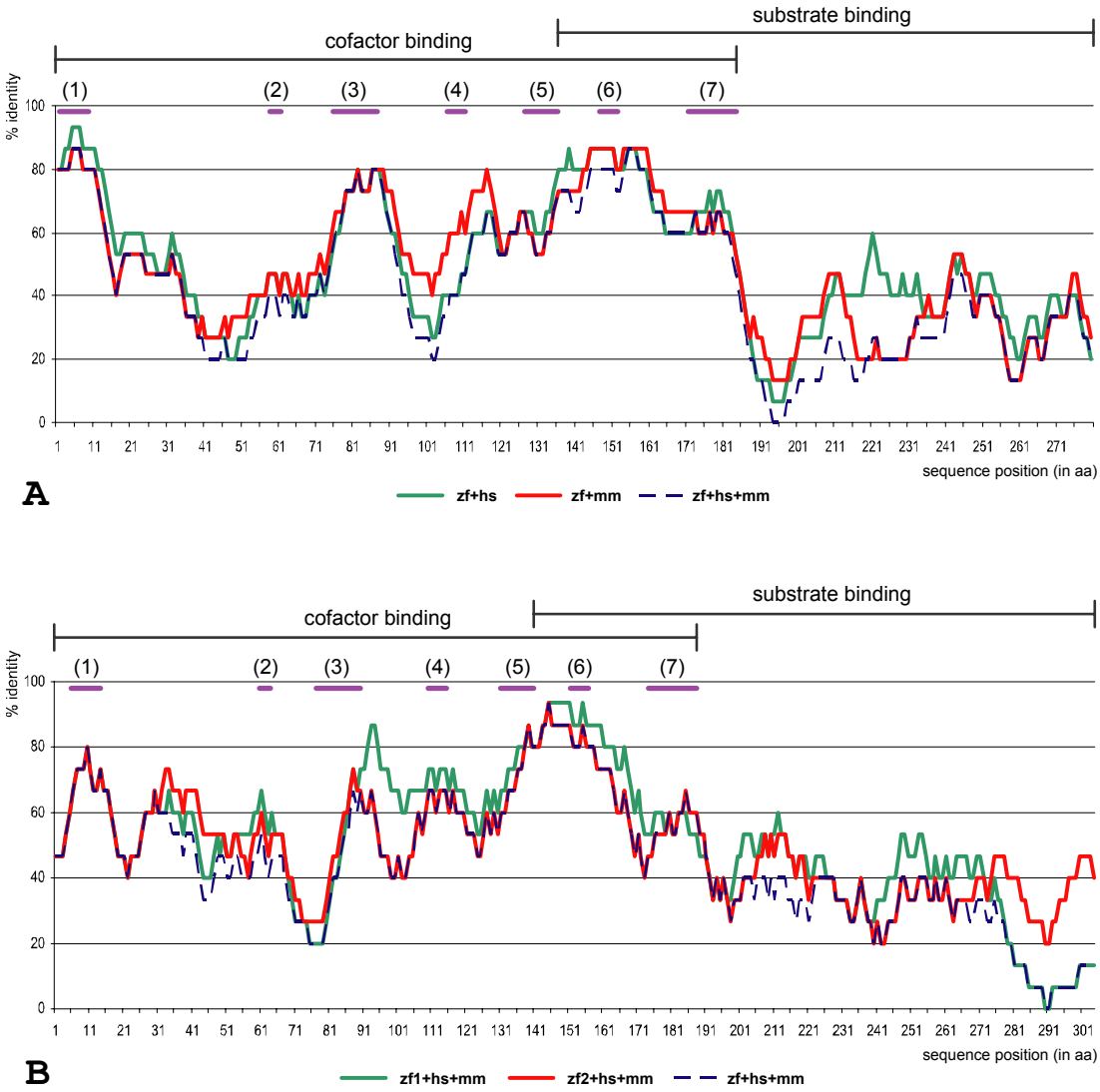


Fig. 14: Identity plots for the sequences of 17beta-HSDs type 1 (A) and photoreceptor associated retinol dehydrogenases (B).
The percent identity of a 15-residue block is plotted against the first amino acid of this block; as a result, plots are cut off fifteen residues before the end of the alignment. The SDR-motifs as described in the previous chapter are inserted in violet together with the respective motif number. Their position was shifted 7 aa positions to the left compared to the alignment in Fig. 13 to level out the slight distortion introduced by the 15-residue block calculation. The regions predominantly binding the cofactor and the substrate are outlined above each plot. For the cofactor binding region, borders were set in a way to mainly comprise the SDR motifs as described in the previous chapter. The substrate binding region was assumed to initiate at the end of motif (5) (for argumentation see chapter 4.1.4.2.1 and Fig. 17). For the group of 17beta-HSDs, the substrate binding region was presumed to stop at the end of the plot as sequences were cut off due to the decidedly shorter zebrafish protein (compare the alignment in Fig. 13). Due to lack of specific information regarding substrate binding in the photoreceptor associated retinol dehydrogenases, the substrate binding region was assumed to terminate at the end of the sequences. The caption points out the different sequences that were taken into account for the respective comparison. hs: homo sapiens; mm: mus musculus; zf: zebrafish; zf1: zfprRDH 1; zf2: zfprRDH 2.

Progression of conservation: integration and comparison of the zebrafish proteins to their mammalian homologs

By analysis of the distribution of residues identical in all enzymes of one group, characteristic regions generally important for functional integrity and at the same time different in 17beta-HSDs type 1 compared to prRDHs emerged. Pairwise sequence analysis in each group was

carried out to get more information about the putative functions of the zebrafish genes deduced from fitting a general characteristic pattern of residue identity rather than presence of specific amino acids at certain positions (as was investigated in the previous chapter).
In case of the 17beta-HSDs type 1, the sequence of zfHSD 1 was on the one hand compared to the human (green) and on the other hand to the mouse (red) homolog. Since the human and murine enzyme differ in their substrate specificity, the resulting Identity plots together with the previously carried out alignment of all three sequences should highlight the presumed function of zfHSD 1. In the mainly cofactor binding region only slight differences can be seen. While from positions 1-40 zfHSD 1 is closer to the human homolog, this is reverted in the following region and becomes more pronounced in the area from about aa 95-120. In the substrate binding area in direct vicinity to the catalytic site (aa 140-185) graphs do not seem to differ significantly. Instead major discrepancies appear in the enzyme part more distal to the catalytic site. In the area comprising positions 190-210 both pairwise alignments exceed that of the group alignment. This means that conserved residues in zebrafish and human are not the same as in zebrafish and mouse although a similar grade of identity in the monitored region is present. Also of high interest is the adjacent region from position 210-235: here, zebrafish is significantly more similar to human reaching a maximum of 60% identity whereas identity to the murine homolog drops to 20%. As this part of the 17beta-HSD type 1 is also responsible for estrogen-androgen discrimination, the zebrafish enzyme might rather resemble the substrate specificity of the human homolog.
Concerning the photoreceptor associated retinol dehydrogenases, it may be assumed that both mammalian proteins have similar cofactor and substrate binding characteristics whereas at least one of the zebrafish paralogs might have differed. Therefore, pairwise comparison of zfprRDH 1 (green) on the one hand and zfprRDH 2 (red) on the other hand to both mammalian sequences was carried out. In the cofactor as well as in the substrate binding region zfprRDH 1 closer resembles the contribution of identical residues seen in the mammalian homologs than zfprRDH 2. This is most significant in the complete area responsible for correct alignment and stabilization of the cofactor and substrate also including the catalytic site (aa 90-180). Position 240-275 might be specific for binding the distal parts of the substrate as residue identity of zfprRDH 1 to the mammalian homologs exceeds that of zfprRDH 2. Finally, in the very C-terminus of the prRDHs all graphs drop to a minimum that reaches 0% amino acid identity in the pairwise comparison of zfprRDH 1. Although this decline is also apparent in the pairwise comparison of zfprRDH 2 to the mammalian homologs, residue identity does here not fall below 20% and might therefore still indicate functional important residues. Interestingly, although zfprRDH 2 is altogether less similar to the mammalian sequences than zfprRDH 1, it displays significantly higher amino acid conservation in the most C-terminal part (position 275-300) compared to its paralog.

4.1.4.2 3D-structure

4.1.4.2.1 Homology modeling

As the 3D-structure of human 17beta-HSD type 1 complexed with NADPH and estradiol is known, it was taken as a template to model and assess the 3D-structure of the three zebrafish

proteins zfHSD 1, zfprRDH 1 and zfprRDH 2. The resulting homology models were investigated concerning differences in the complete structural alignment as well as in the characteristics of the cofactor and substrate binding pockets.

Complete structural alignment

The zebrafish and human proteins do not share identical chain lengths as was apparent after revelation of the complete coding sequence. In the sequence based alignment (chapter 4.1.4.1.3) this difference was accounted for by introduction of gaps. In a homology model, those deviations appear as loops, and in addition might be positioned at other sites of the primary sequence than in the sequence based alignment.

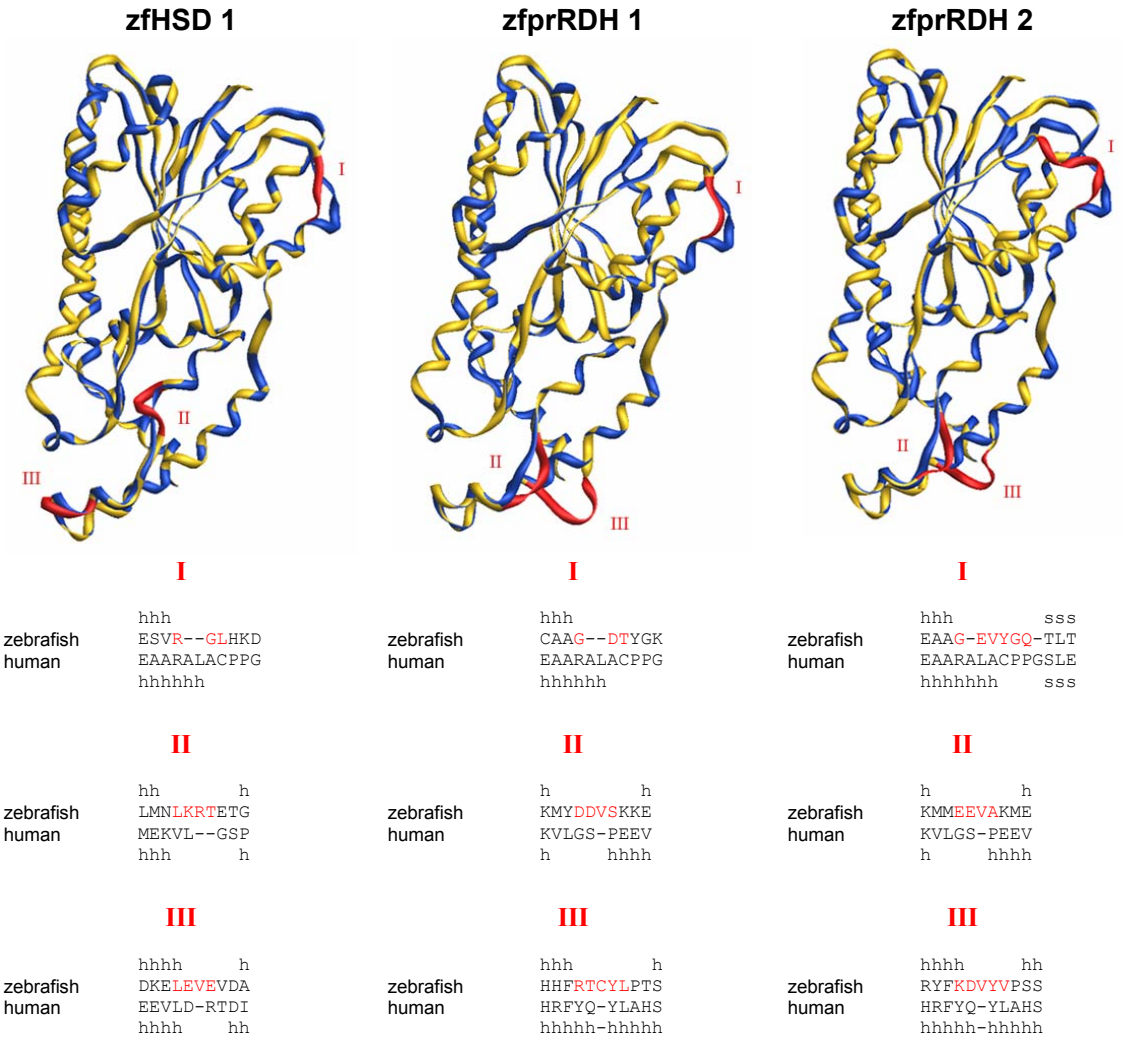
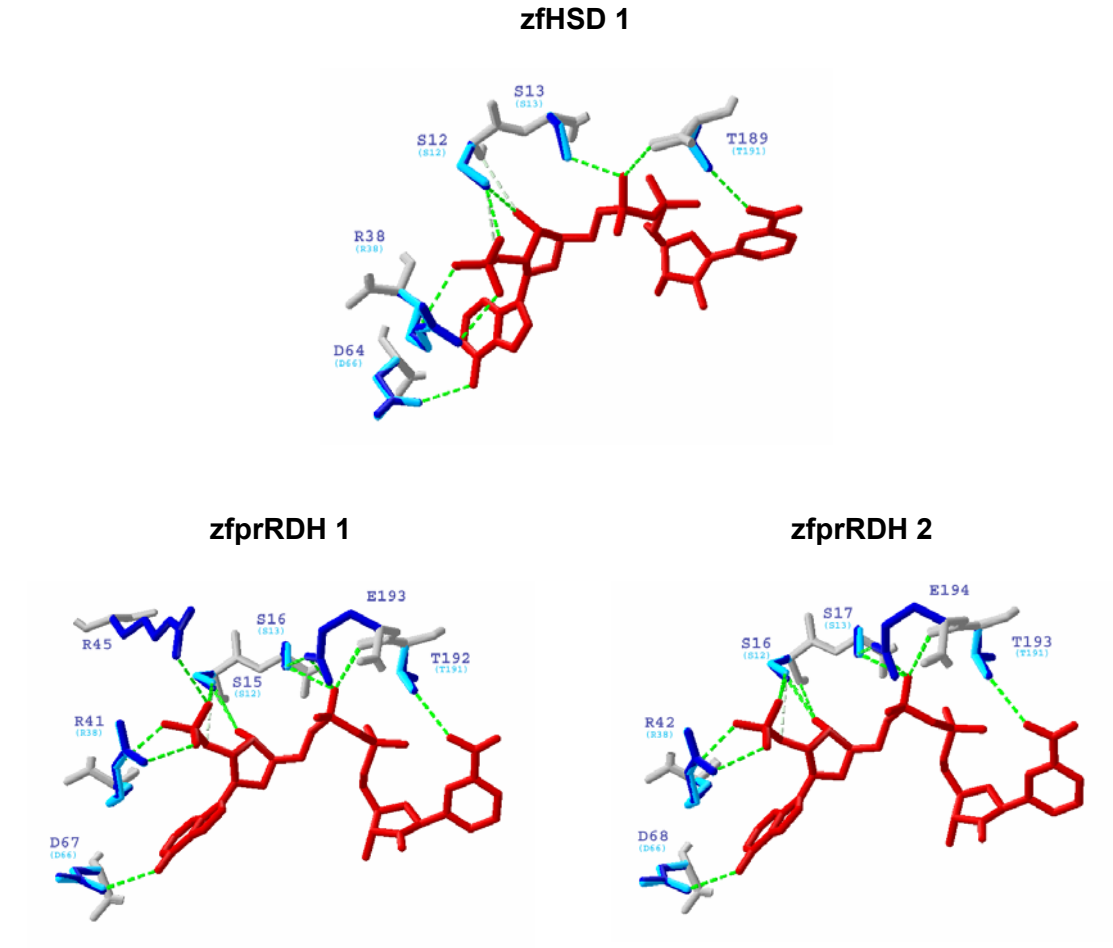


Fig. 15: Complete structural homology models of zfHSD 1 and the zfprRDHs in comparison to human 17beta-HSD type 1.
The structure of the human protein is depicted in blue and that of the respective zebrafish homology model in yellow. Areas in the zebrafish proteins displaying apparent structural differences to the human enzyme due to dissimilar chain length are marked in red and numbered with roman letters. The corresponding sequence alignments of these parts are given below each model. In addition, the affiliation of the amino acids to secondary structure elements is displayed below and above the alignment to facilitate a better overview of the involved structural consequences. h: α-helix; s: β-strand.

The complete structural alignments of the zebrafish proteins to human 17beta-HSD type 1 are shown in Fig. 14 and reveal three regions of deviations each (marked in red). In addition, the sequence alignments corresponding to these three areas are given below each model. In all three zebrafish proteins the first deviation (I) occurs at about the same position which is the loop connecting the αB-helix with the βC-strand. The lack of two residues compared to the human enzyme leads to a shortened αB-helix, but the subsequent structure (βC) does not seem to be affected. Since this area is not involved in binding of the cofactor, enzymatic function in this respect may not be altered in the zebrafish enzymes.
The second main structural difference (II) is located adjacent to the SDR motif (7) and affects the connection between the area predominantly involved in cofactor binding and the C-terminal part of the protein (for motifs and functionally different parts see also Fig. 13, chapter 4.1.4.1.3 and Fig. 14, chapter 4.1.4.1.4). Although zfHSD 1 harbors two additional amino acids in this region, the neighboring secondary structures are not apparently altered as only the preceding helix is shortened by one residue due to the extended loop. In contrast, the additional amino acid in the two zfprRDHs has strong influence on the subsequent structural element which is the αG-helix in the human 17beta-HSD type 1. This helix may still be present in the zebrafish RDHs but is shortened and starts later.
The third structural deviation (III) is caused by one additional residue in the zebrafish proteins influencing non-identical areas in the 17beta-HSD and the prRDHs. In zfHSD 1, the corresponding region is in close vicinity to deviation (II) and slightly extends the connecting loop between helices αG and αH. The two zebrafish prRDHs instead display a strongly altered structure compared to human 17beta-HSD type 1 affecting helix αH. Here, the additional amino acid leads to disruption of the αH-helix by formation of an internal loop, which is even more extended in zfprRDH 1. Since especially the part of αH downstream of this region conveys substrate binding in human 17beta-HSD type1, the observed conformational changes in the zfprRDHs may have consequences for their substrate specificity.

Cofactor and substrate binding areas

The homology models not only allow for comparison of the complete tertiary structure but as well for identification of specific residues at functionally important positions. For this, amino acids in a radius of 4 Å to cofactor and substrate were identified for the three zebrafish enzymes as well as the human 17beta-HSD type 1 and compared. In case of the cofactor binding site, residues forming hydrogen bonds to NADP were investigated as well.
Binding and positioning of NADP is enabled in great part by hydrogen bonds which are formed by the backbone and side chain of amino acids at certain positions (see Fig. 16). These are highly conserved and in case of the 17beta-HSDs type 1 and prRDHs comprise the two serines in motif (1), the aspartic acid in motif (2) and the threonine in motif (7) (for identification of SDR motifs see also chapter 4.1.4.1.3). In addition, a hydrogen bond is formed between a conserved arginine in the region between aa 35-40 and the phosphor group only present in NADP and not NAD. Hence, the three zebrafish enzymes seem to share with human 17beta-HSD type1 a preference for the NADP cofactor. Interestingly, the two zebrafish prRDHs harbor additional amino acids which may form hydrogen bonds with the cofactor: R45 and E193 in zfprRDH 1 and E194 in zfprRDH 2.
The positions of residues not directly interacting with the cofactor by hydrogen bonds but in 4 Å vicinity and hence, forming the cofactor binding pocket are completey identical between human 17beta-HSD type 1 and the three zebrafish enzymes. In all cases, these amino acids belong to SDR motifs highlighting their importance in functional integrity.



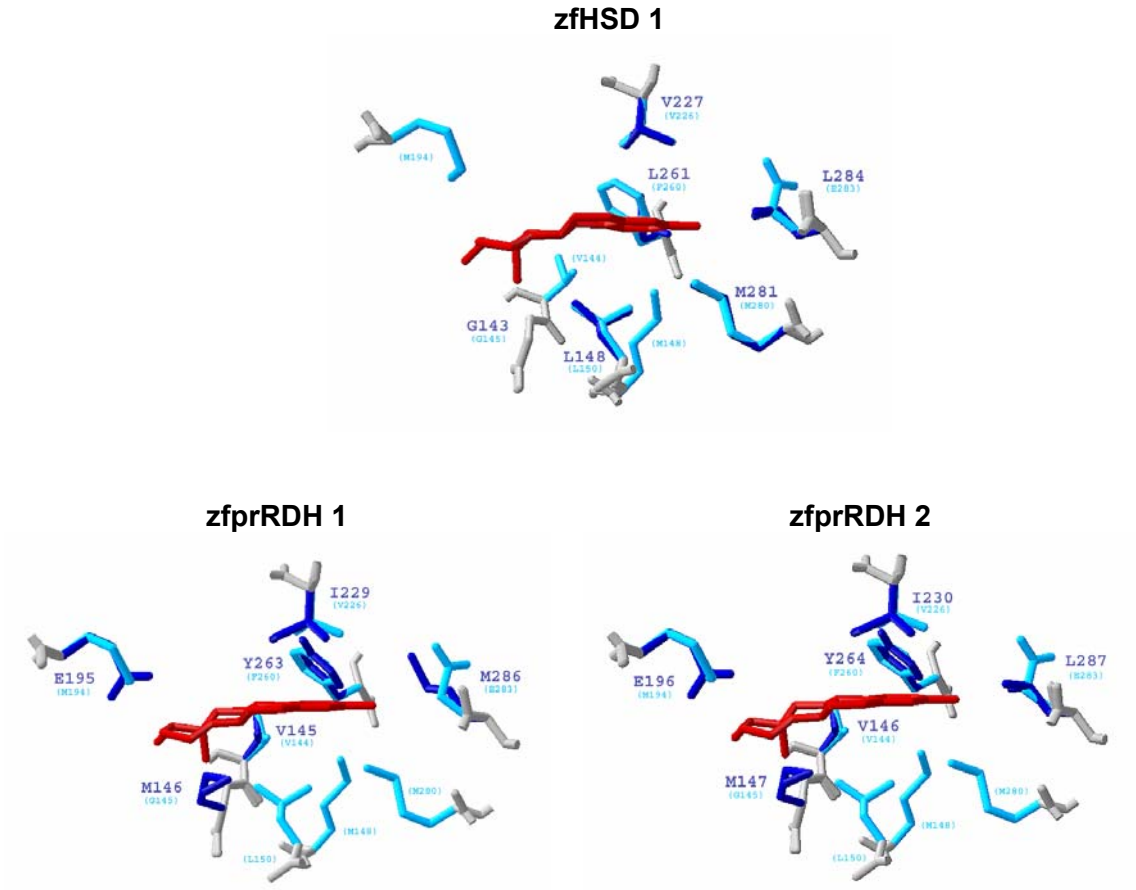

```
hsHSD1          --MAR--TVVLITGCSSGIGLHLAVRLASDPSQSFKVYATLRDLKTQGRLWEAARALACPPGSLETLQLDVRDSKSVAAA 76
zfHSD1          --MEQ--KVVLITGCSSGIGLSLAVHLASNPAKAYKVYATMRNLDKKQRLLESVR--GLHKDTLDILQMDVTDQQSILDA 74
zfprRDH 1       MASAG-QKVVLITGCSSGIGLGIAVMLARDKQQRYYVIATMRDLKRQEKLVCAAG--DTYGKTLTVCTLDVCSNESVRQC 77
zfprRDH 2       MASGGGQKVVLITGCSSGIGLRIAVLLARDEQKRYHVIATMRDLKKKDRLVEAAG--EVYGQTLTLLPLDICSDESVRQC 78
                              motif (1)                                                 motif (2)

hsHSD1          RERVTEGRVDVLVCNAGLGLLGPLEALGEDAVASVLDVNVVGTVRMLQAFLPDMKRRGSGRVLVTGSVGGLMGLPFNDVY 156
zfHSD1          QRNVSEGRIDILVCNAGVGLMGPLETHSLDTIRAIMDVNLLGTIRTIQTFLPDMKKKRHGRILVTGSMGGLQGLPFNEVY 154
zfprRDH 1       VDSVKDRHIDILINNAGVGLVGPVEGLSLDDMMKVFETNFFGAVRMIKEVMPDMKKRRSGHIIVISSVMGLQGVAFNDVY 157
zfprRDH 2       VNSVKDRHIDVLINNAGVGLLGPVESISMDEMKRVFETNFFGTVRMIKEVMPDMKKRQAGHIIVMSSVMGLQGVVFNDVY 158
                          motif (3)                 motif (4)                 motif (5)

hsHSD1          CASKFALEGLCESLAVLLLPFGVHLSLIECGPVHTAFMEKVLG---SPEEVLDRTDIHTFHRF-YQYLAHSKQVFREAAQ 232
zfHSD1          CASKFAIEGACESLAILLQHFNIHISLIECGPVNTDFLMNLKRTETGDKELEVEVDAHTRSLY-DQYLQHCQSVFQNAAQ 233
zfprRDH 1       AASKFAIEGFCESLAVQLLKFNVTMSMIEPGPVHTEFEMKMYDD--VSKKEYPNTDPETMHHFRTCYLPTSVNIFQGLGQ 235
zfprRDH 2       TASKFAIEGFCESMAVQLLKFNVKLSLIEPGPVHTEFETKMMEE--VAKMEYPGADPDTVRYFKDVYVPSSIDIFEAMGQ 236
                motif (6)                 motif (7)
```

Fig. 16: Identification and function of amino acids in the vicinity of the cofactor in human 17beta-HSD type 1 and the three zebrafish enzymes.
The position and identity of residues forming hydrogen bonds to the cofactor NADP (red) are depicted for each zebrafish homology model. The backbone of the proteins is presented in gray, the side chains of amino acids in dark blue for the zebrafish and in light blue for the human enzyme. Hydrogen bonds are shown as green dotted lines. The alignment below resembles part of the sequence based alignment (see also Fig. 13) and highlights the positions in the vicinity of the cofactor. Amino acids within 4 Å distance to NADP are written in blue; those additionally forming hydrogen bonds are shaded in the same color.

In 17beta-HSD type 1, the substrate binding pocket is less strictly organized than the cofactor binding part which is due to two reasons: (i) binding of the steroid substrate is rather performed through hydrophobic and steric interactions than through the formation of hydrogen bonds, (ii)

induced fit upon substrate binding reflects higher flexibility of the substrate binding part in comparison to the cofactor binding compartment. Therefore, this analysis is likely to reveal many important amino acids in the zebrafish enzymes but may not be comprising in identification of all residues potentially interacting with the substrate.
Amino acids within 4 Å of the estradiol were identified in all homology models (Fig. 17). Residues identical in the zebrafish proteins as well as the human enzyme are present only in the region which in addition exerts cofactor binding. Most of them are part of SDR motifs such as the conserved serine in motif (5) and the tyrosine in motif (6), which form a part of the catalytic triad, and also the conserved glycine and proline in motif (7). The asparagine preceding motif (6) is part of the highly conserved region connecting motifs (5) and (6) and is present in all four proteins. As these mentioned residues appear to be an inevitable part of general SDR functional integrity, they are not likely to influence substrate specificity. Hence, they were left out of the following analysis and omitted in Fig. 17, which depicts the amino acids involved in substrate interaction.
Of the remaining positions, four could be identified in all four proteins, corresponding to the human 17beta-HSD type 1 G145, V226, F260 and E283. Although these four residues are in vicinity to the substrate, they are not identical among the investigated enzymes. G145 is conserved in zfHSD 1 while both zfprRDHs have a methionine in this position. Similarly, V226 is also present in zfHSD 1 while the two zfprRDHs display isoleucines. As these two positions (aa 145 and 226) seem to be conserved but different in the two functionally dissimilar groups, they may be involved in specifically binding a steroid or retinoid substrate. F260 is not conserved between human and zebrafish 17beta-HSD type 1. In this position, both zebrafish prRDHs harbor a conserved tyrosine which therefore might play a role in substrate recognition in prRDHs. The last position within 4 Å of the estradiol in all four models (E283) shows no specific conservation in respect to enzyme function or organism.
The remaining positions could not be identified in all four proteins, which may be due to a slightly altered 3D-structure. Position M194 of human 17beta-HSD type 1 may be of interest concerning substrate specificity since here both zfprRDHs harbor a glutamic acid in contrast to the methionine. Human 17beta-HSD type 1 positions 150 and 280 could only be identified in the HSDs and not in the prRDHs. Furthermore, the respective amino acid is conserved in the human and zebrafish enzyme and therefore likely to have an impact on substrate specificity.

Aside from the calculation of hydrogen bonds, the Swiss-pdbViewer also allows for calculation of distances between residues thereby highlighting colliding amino acids being positioned too close to each other. In this context, the position of human G145 became even more interesting. In this area glycine is conserved in both the human and zebrafish 17beta-HSD type 1 where together with other residues it forms a pocket for the C18 methyl group of the estradiol. In contrast, the two prRDHs display a highly conserved methionine at this position which gets into steric conflict with the C18 methyl group and may impair the binding of an estrogen substrate in these retinol dehydrogenases.
The possibility that in the position of human aa 145 glycine is sterically necessary for estrogen binding while methionine is too bulky was further investigated by site-directed mutagenesis and the resulting effects on substrate specificity as described in the following chapters.




```
hsHSD1          KRRGSGRVLVTGSVGGLMGLIPFNDVYCASKFALEGLCESLAVLLLPFGVHLSLIECGPVHTAFMEKVLG---SPEEVLDR 210
zfHSD1          KKKRHGRILVTGSMGGLQGLIPFNEVYCASKFAIEGACESLAILLQHFNIHISLIECGPVNTDFLMNLKRTETGDKELEVE 208
zfprRDH 1       KKRRSGHIIVISSVMGLQGVAFNDVYAASKFAIEGFCESLAVQLLKFNVTMSMIEPGPVHTEFEMKMYDD--VSKKEYPN 211
zfprRDH 2       KKRQAGHIIVMSSVMGLQGVVFNDVYTASKFAIEGFCESMAVQLLKFNVKLSLIEPGPVHTEFETKMMEE--VAKMEYPG 212
                       motif (5)           motif (6)                  motif (7)

hsHSD1          TDIHTFHRF-YQYLAHSKQVFREAAQNPEEVAEVFLTALRAPKPTLRYFTTERFLPLLRMRLDDPSGSNYVTAMHREVFG 286
zfHSD1          VDAHTRSLY-DQYLQHCQSVFQNAAQDTEDIIQVYLEAMEAQTPFLRYYTNRALLPMSSLKLTSMDGSQYIRAMSKLIFS 287
zfprRDH 1       TDPETMHHFRTCYLPTSVNIFQGLGQTPEDIAKVTKKVIESPRPPFRSLTNPLYTPIVALKYADDSGDLSLHTFYHMLYN 289
zfprRDH 2       ADPDTVRYFKDVYVPSSIDIFEAMGQTPDDIAKCTKKVIETSQPRFRNLTNSLYTPIVAMKYADETGGLSVQTFYNLLFN 290
```

Fig. 17: Identity of amino acids in the vicinity of the in human 17beta-HSD type 1 and the three zebrafish enzymes.
The backbone of the proteins is presented in gray, the side chains in dark blue for zebrafish and light blue for human amino acids. In the model, only residues or positions within 4 Å of the substrate (estradiol in red) but non-identical in all four sequences are depicted. The sequence based alignment below highlights all positions calculated to be within a distance of 4 Å to the substrate: amino acids identical in all four sequences are written in blue, non-identicals are shaded in the same color.

4.1.4.2.2 Expression of recombinant protein

To enable investigation of their substrate specificities zfHSD 1 and the two zfprRDHs were cloned into expression vectors and transformed into *E. coli* (for method see chapter 2.3.1). Following verification of the clones by sequencing, production of recombinant protein was performed. After lysis of the bacteria the resulting fractions were analyzed via SDS-PAGE and Coomassie staining. Additionally, Western Blots were carried out to verify the results.

Wild type zebrafish enzymes

Fig. 18 depicts the expression of recombinant zfHSD 1, zfprRDH 1 and zfprRDH 2 under comparable conditions. Following SDS-PAGE and Coomassie staining, the GST fusion proteins are clearly visible in the fraction of pelleted proteins after induction in all three cases (arrow) at the expected sizes of about 59 kDa for zfHSD 1 and 61 kDa in case of both zfprRDHs.

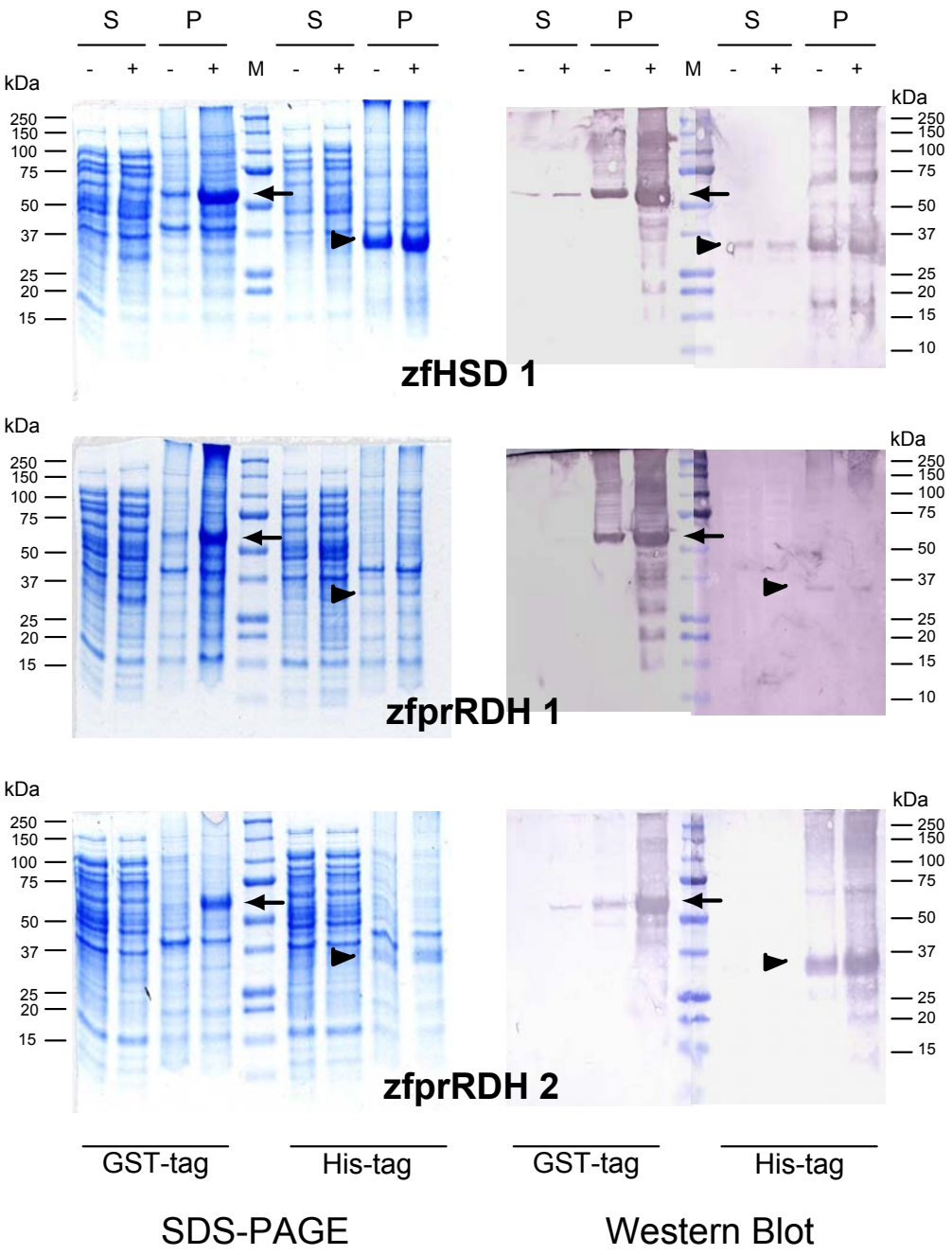


Fig. 18: Expression of recombinant zfHSD 1, zfprRDH 1 and zfprRDH 2.
The left side depicts Coomassie stained SDS-PAGE gels that show the various fractions after protein-induction for both vector-systems. The right side shows the result after blotting of identical gels and probing them with monoclonal antibodies against the respective tag (for method see chapter 2.3.3). Arrows point at the GST-fusion proteins; arrowheads mark the position where His_6-tagged protein would be expected. M: molecular size marker; P: pellet fraction; S: soluble fraction; - : without IPTG; + : with IPTG induction.

Expression seems to be especially high for zfHSD 1 and zfprRDH 1; here, protein also seems to be present in the same protein fraction without induction indicating impaired suppression control. Apart from that, no formation of GST fusion protein can be monitored without induction. The His_6-tag constructs are not clearly visible following Coomassie staining. In all cases there are no apparent differences between induced and non-induced fractions that would hint at expression of the recombinant protein. Although, at the expected size of about 35 kDa in case of the zfprRDH fusion-proteins, a weak band is visible in induced and non-induced fractions of pelleted proteins (arrowhead). A more prominent band is only in detectable case of His-tagged zfHSD 1 (~33 kDa).

The findings obtained from Coomassie stained gels are mirrored in the Western Blots performed with monoclonal antibodies against GST and His_6. In case of GST-fusion constructs, recombinant protein is readily detectable in all fractions of insoluble proteins after lysis. Induction with IPTG seems to increase protein expression in case of zfHSD 1 and zfprRDH 1 only slightly whereas in case of zfprRDH 2 formation of the recombinant protein was strongly enhanced. zfHSD 1 is the only construct where small amounts of protein is also present in solution (-S and +S fractions). The same is true for zfHSD 1 constructs carrying a His_6- instead of a GST-tag. Here, the fusion protein is detectable with anti-His_6 antibodies in all prepared fractions. The two retinol dehydrogenases expressed from pQE 30 vectors are present as insoluble proteins independent from induction with IPTG. But in comparison to zfHSD 1, expression levels are much lower; this is especially the case for zfprRDH 1, that is barely detectable.

Point mutated zebrafish enzymes

My analysis of the 3D-structure of the three zebrafish proteins by homology modeling in the previous chapter resulted in the postulation of a central role for a conserved glycine in 17beta-HSDs type 1 and methionine in prRDHs in substrate discrimination. To investigate this possibility the respective amino acid was changed to the specific residue of the functionally different enzyme in all three sequences by site-directed mutagenesis. The resulting mutants zfHSD 1-G143M, zfprRDH 1-M146G and zfprRDH 2-M147G were then subjected to expression and analysis similar to the wild type enzymes (see above). The expression of recombinant proteins after the single amino acid exchange was monitored by the same means as described for the wild type.

The results of expression of the mutated recombinant zebrafish proteins are shown in Fig. 19 and reflect a similar situation as was found for the wild type proteins: (i) formation of GST fusion proteins was visibly enhanced upon induction with IPTG and readily detectable already by Coomassie staining after SDS-PAGE, (ii) Western Blots mirror the gels and in addition demonstrate presence of small amounts of the recombinant proteins as well in uninduced and soluble fractions.

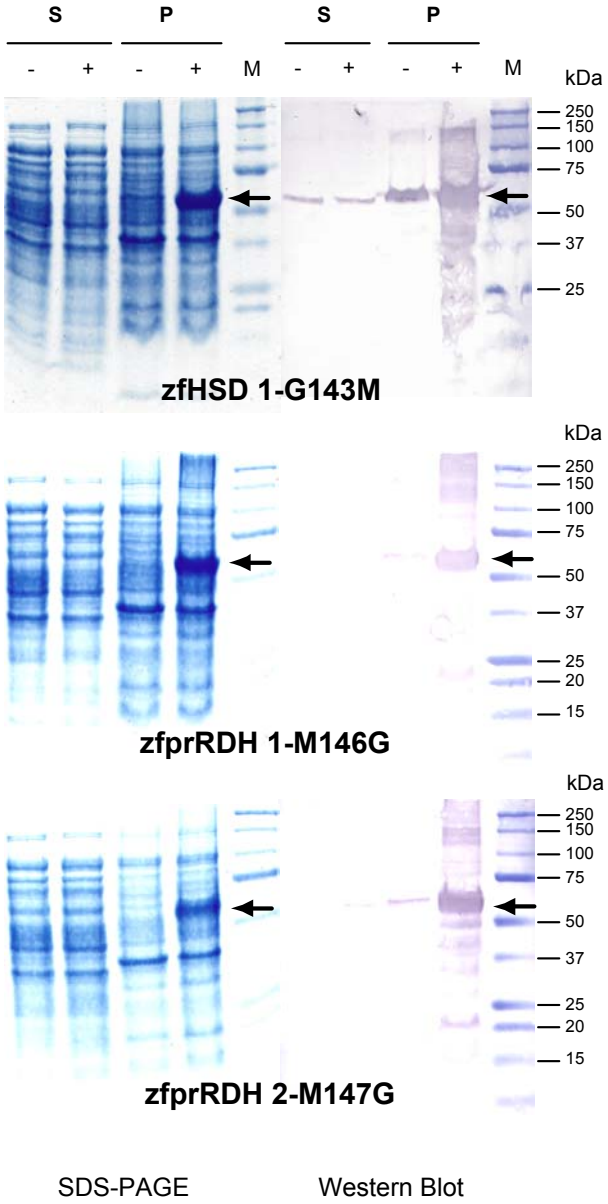


Fig. 19: Expression of the GST-tagged, point-mutated zebrafish proteins.
The left side depicts Coomassie stained SDS-PAGE gels that show the various fractions following lysis of bacteria. The right side shows the result after blotting of identical gels and probing them with monoclonal antibodies against the GST-tag (for method see chapter 2.3.3). Arrows point at the GST-fusion proteins. M: molecular weight marker; P: pellet fraction; S: soluble fraction; - : without IPTG; + : with IPTG induction.

4.1.4.2.3 Activity measurements and substrate specificity

Bacterial lysates containing the over-expressed zebrafish GST-fusion proteins (pellet fractions, see also chapter 4.1.4.2.2) were employed to investigate the potential of these enzymes to convert the substances typical to the corresponding mammalian enzymes. Aside from the wild type enzymes, the point mutated proteins (see chapter 4.1.4.2.2) were investigated in the same way to detect putative changes in substrate specificity. These *in vitro* assays (for method see chapter 2.3.4) focused on qualitative measurement rather than quantification.

A reaction characteristic for the known mammalian enzymes is the conversion of estrone to estradiol by use of NADPH as cofactor. This conversion is readily performed by the zebrafish 17beta-HSD type 1 homolog leading to the formation of nearly 100% product during 30 minutes incubation (Fig. 20A). In contrast, no product was formed when estrone was incubated with any of the zfRDHs under the same conditions. Activity of the mutant enzyme zfHSD 1-G143M is not significantly affected as under the same conditions product formation appeared to be only minimally reduced. In addition, traces of estradiol were formed by zfprRDH 1-M146G while wild type form of this enzyme did not catalyze this conversion. zfprRDH 2-M147G, like the wild type form of this enzyme, did not show any activity towards estrone.
In vitro mammalian 17beta-HSDs type 1 are also capable of performing the reverse reaction: oxidation of estradiol to estrone by addition of NAD as cofactor. Again zfHSD 1 catalyzed this reaction with high efficiency (Fig. 20B) while zfprRDH 1 may be very weakly active and zfprRDH 2 did not catalyze this reaction at all. Furthermore, the zfHSD 1-G143M showed severe impairment of the catalysis of this reaction. Interestingly, while the wild type zfHSD 1 displayed similar activity in both reaction directions, the point mutation in zfHSD 1 shows a significantly stronger effect selectively in the oxidation reaction.

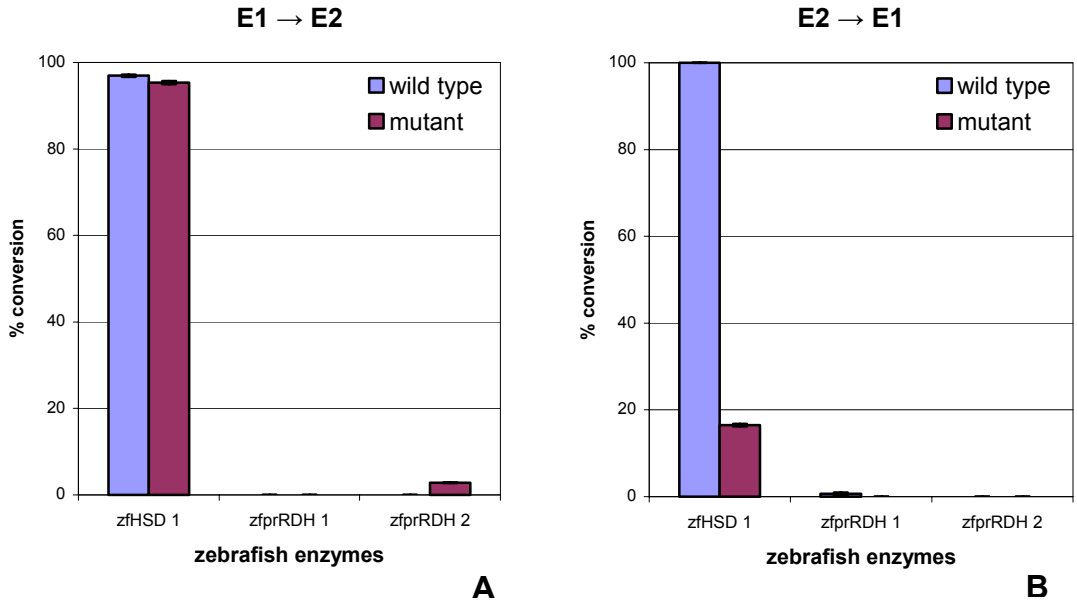


Fig. 20: Activity measurements of wild type and point mutated zebrafish GST fusion proteins.
Conversion of estrone to estradiol was measured for wild type (blue) and mutated (red) zebrafish proteins (**A**) under conditions described in chapter 2.3.4. The potential to catalyze the reverse reaction *in vitro* was investigated as well (**B**). E1: estrone; E2: estradiol.

As it is known that mouse and rat 17beta-HSD type 1 additionally catalyze the reduction of androstenedione to testosterone, these substances were also tested. Neither the wild type nor the mutated zebrafish enzymes were able to perform this conversion or the corresponding back-reaction.

4.2 17beta-HSD type 3

4.2.1 The candidate genes

Three candidate sequences for a zebrafish 17beta-HSD type 3 homolog, which were each represented by multiple, non-identical ESTs were identified by *in silico* screen of EST-databases. According to the chronology in which they were discovered they were named zf 3.1, zf 3.2 and zf 3.3. A fourth candidate was identified by comparing the mouse protein sequence against the whole genome shotgun (WGS) database and named zf 3.4. For this sequence, specific ESTs were found at a later time point (see also chapter 4.2.3.2), thereby independently verifying the presence and accuracy of the identified gene. Due to sequence complementation and detailed analyses, the candidates were found to be zebrafish 17beta-HSD type 3 (zf 3.4), 17beta-HSD type 12 A (zf 3.1) and B (zf 3.3). Here, these genes will be abbreviated and referred to as zfHSD 3, zfHSD 12A and zfHSD 12B (for overview, see Tab. 6); the respective complete cDNAs are shown in an overview in Fig. 21. Furthermore, analyses revealed one candidate sequence to be a yet uncharacterized putative steroid dehydrogenase (zf 3.2), which will not be subject of this thesis.

Tab 6: Zebrafish candidate genes that were identified as putative homologs of mammalian 17beta-HSD type 3.

zebrafish homolog of	*acronym*	*original name*	*full-length clone*	*GeneBank Acc.*
17beta-HSD type 3	zfHSD 3	zf 3.4	IMAGp998N0614301Q3	AY551081
17beta-HSD type 12	zfHSD 12A	zf 3.1	UCDMp611M09141Q10	AY551082
	zfHSD 12B	zf 3.3	zfHSD3.3AB-pGEX*	AY551080

In case of the 17beta-HSD type 12 two paralogous genes were discovered in zebrafish where in mammals only one gene is known. The column “full-length clone” lists the names of clones that contained the full-length cds, which were either requested from RZPD or molecularly reconstructed (*) by laboratory methods (see below). The candidates were completely sequenced and submitted to GeneBank following analysis.

zfHSD 3

The corresponding clones of two candidate EST sequences could be ordered from RZPD by their clone Id (IMAGE: 6793097, IMAGp998G1614299 and IMAGE: 6794023, IMAGp998N0614301). Sequencing demonstrated the latter to contain the full-length coding sequence whereas the first contained the same cds with some single nucleotide exchanges but lacked exon 2. Since the second exon harbors the cofactor binding site, only clone IMAGp998N0614301 was considered to resemble the complete-cds and was taken for further analysis. The 2172 bp insert of this clone contained 27 bp 5’UTR, 924 bp full-length cds and 1221 bp 3’UTR ending on a poly-A tail.

zfHSD 12A

Two clones were requested from RZPD via the candidate EST's GenBank accession number (AW128105, IMAGp998K148939Q2) and clone Id (IMAGE: 3725308, IMAGp998K058971Q2) but sequencing revealed them to contain the wrong inserts. Thus, a specific fragment amplified by RT-PCR was used as a probe to screen a zebrafish brain cDNA-library. The screen yielded two clones that were ordered from RZPD: UCDMp611C20127Q10 and UCDMp611M 09141Q10. Sequencing demonstrated that the first clone contained the 5' part of zfHSD 12A (starting at exon 3) and the latter contained the full-length cds. The 1982 bp insert of clone UCDMp611M09141Q10 is composed of the following parts: 111 bp 5'UTR, 960 bp full-length coding sequence and 911 bp 3'UTR closing with a poly-A tail.

zfHSD 12B

The clone WUSMp624N1615Q2 was obtained from RZPD via the candidate's EST name (fk28d08.x1) but found to contain the wrong insert after sequencing. Screening of zebrafish cDNA libraries from brain and liver with a specific probe amplified by RT-PCR did not yield any positive clones. As several ESTs for zfHSD 12B were present in the databases, the full-length coding sequence was constructed by an assembly of these EST sequences. To acquire this sequence for laboratory experiments two overlapping fragments were amplified by RT-PCR, combined by fusion PCR and subcloned yielding the full-length zfHSD 12B cds of 936 bp.



Fig. 21: cDNA structures of 17beta-HSD type 3 and the two 17beta-HSDs type 12 from zebrafish.
The coding sequence is depicted in green whereas the 5'- and 3'-UTRs are shown as white boxes. The blue bar indicates length and position of the probe employed for screening of the cDNA library. The two red bars represent the length and position of the two RT-PCR products that were assembled in a Fusion-PCR to yield the full-length cds.

4.2.2 Characterization on DNA level

4.2.2.1 Coding sequence

To investigate the relations between the coding sequences of all zebrafish 17beta-HSD type 3 candidates an alignment was performed (Fig. 22). In addition, an identity matrix was calculated (Tab. 7).

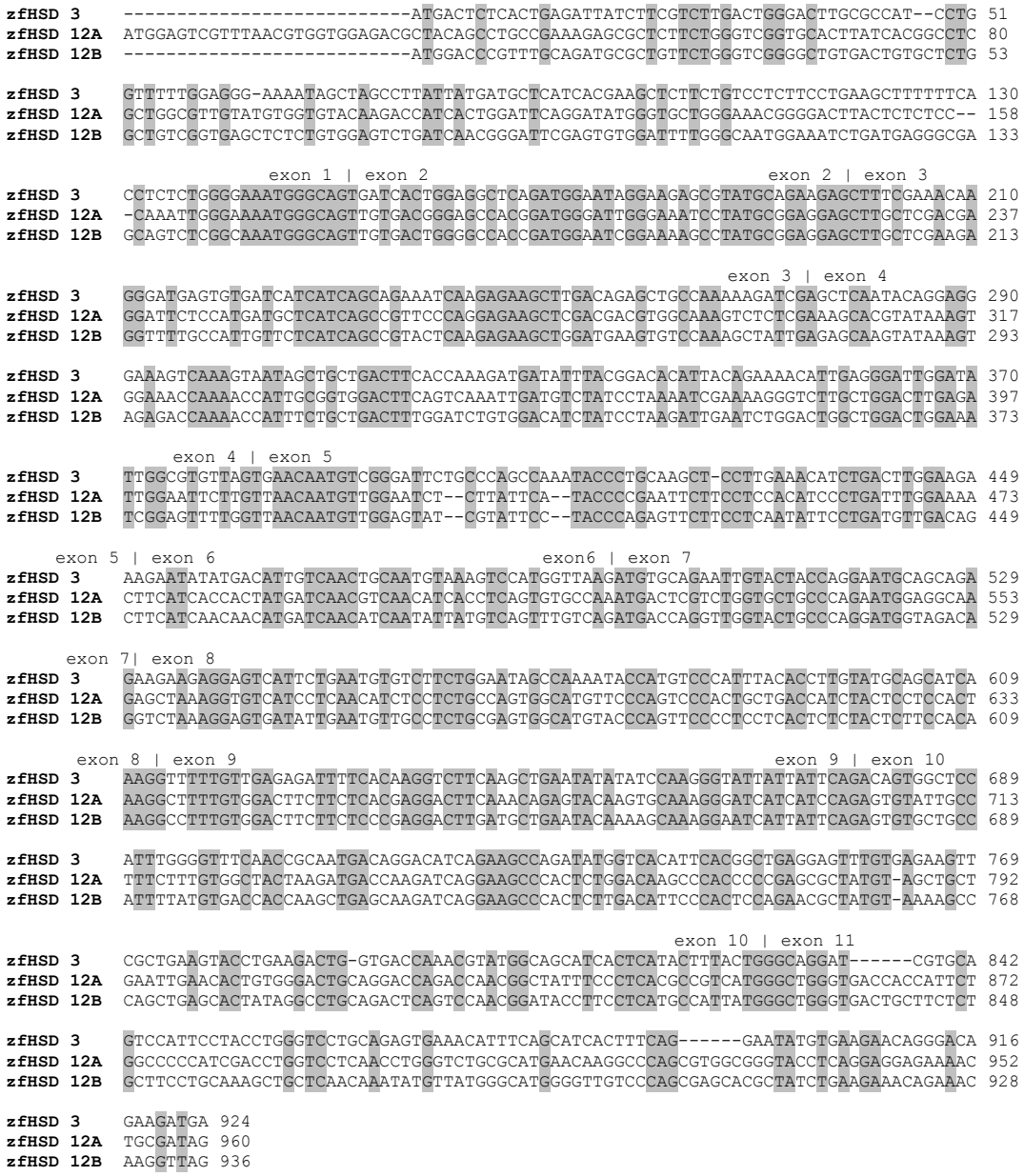


Fig. 22: Alignment of the full-length coding sequences of zfHSD 3 and the two types of 17beta-HSD type 12.
Nucleotides identical in all three sequences are shaded in gray. Dashes indicate gaps that were inserted to facilitate better alignment. Intron insertions as present in the genomic sequences are given above the respective sites. Alignment was performed with ClustalW.

Tab. 7: Identity matrix of the coding sequences of the zebrafish 17beta-HSD type 3 candidate genes.

	zfHSD 3	***zfHSD 12A***	***zfHSD 12B***
zfHSD 3	100%		
zfHSD 12A	49%	100%	
zfHSD 12B	51%	68%	100%

The matrix was calculated with BioEdit subsequent to alignment with ClustalW.

On this level about 50% of nucleotides seem to be conserved in all three sequences while zfHSD 12A and 12B share a significantly higher identity (68%). With the exception of some clusters and lowly conserved first and last exon, identical nucleotides seem to be evenly distributed throughout the full-length sequence. Interestingly, nucleotides in direct vicinity to an intron are conserved in 90% of cases.

4.2.2.2 Exon structure

A good indicator for the homology of two genes is the number and individual sizes of exons, which are determined by the position of introns in the gene. The respective exons were identified as described in Methods (see chapter 2.4.2). From this group of zebrafish genes as well as the homologous mouse and human genes the exon structures are shown in Tab. 8.

Tab. 8: Exon sizes of the three zebrafish candidate genes in comparison to the respective mouse and human homolog.

exon:	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*
17beta-HSD type 3:											
zfHSD 3	151	47	76	108	68	45	35	82	66	150	96
mouse	-9	~	~	~	~	-9	~	~	~	-3	+15
human	+3	~	~	~	~	-9	~	~	~	~	+15
17beta-HSD type 12:											
zfHSD 12A	178	47	76	108	65	45	35	82	66	150	102
zfHSD 12B	-24	~	~	~	~	~	~	~	~	~	~
mouse	-18	~	~	~	~	~	~	~	~	~	+3
human	-18	~	~	~	~	~	~	~	~	~	+3

For the first sequence of each group, the exact number of nucleotides comprised by the individual exon is given. The exon sizes of subsequent genes are depicted in comparison to the first sequence and deviations are given by +/- the respective number of nucleotides; ~ indicates concordant exon size.

All three zebrafish candidate genes as well as their murine and human homologs display the same highly conserved gene structure comprising eleven exons of characteristic sizes. The small variations that appear only in some specific exons and gene groups are restricted to exons 1, 5, 6, 10 and 11 and take place in a triplet-like manner. Moreover, the exact size of exon 5 seems to be characteristic for the respective gene group: 68 nt in 17beta-HSDs type 3 and 65 nt in 17beta-HSDs type 12. Possibly, exon 6 might be added to this pattern as its size seems to be strictly conserved with the exception of the zebrafish 17beta-HSD type 3 gene.
A higher degree of conservation in gene structure can be observed in the group of 17beta-HSD type 12. Here, the two zebrafish genes differ only in the first exon which is enlarged in zfHSD 12B by 24 nucleotides; both genes might have emerged from a gene duplication event. In comparison to zfHSD 12A the first exons of the mammalian genes are also smaller (by 18

nucleotides each) whereas the last exon is increased by three nucleotides. These changes might be characteristic for mammals as they are conserved to the mouse and human gene which show identical gene structures.
More variations take place in the group of 17beta-HSD type 3 genes. As was the case for 17beta-HSD type 12, the sizes of the first and last exon are not conserved between fish and mammals but furthermore the mouse and human gene also do not share an identical structure as was the case in 17beta-HSDs type 12. In comparison to zfHSD 3, the first exon of the mouse gene is shortened by nine nucleotides; the human's gene first exon instead is increased by three nucleotides. The last exon is longer in both mammalian genes: in mouse by fifteen, in human by nine nucleotides. Furthermore, exon 10 is shortened in the mouse gene compared to zebrafish or human. The only feature conserved for the mammalian genes and different from zebrafish is their shortened exon 6. While all other genes in Tab. 8, also including 17beta-HSDs type 12, have a 45bp-exon 6, mouse and human 17beta-HSD type 3 are the only ones that harbor a 36bp-exon 6. Therefore, this reduction in exon 6 might be typical for the mammalian 17beta-HSDs type 3.

4.2.2.3 Gene and intron sizes

The architecture of all zebrafish genes as well as the respective mouse and human homologs were investigated as described in Methods (see chapter 2.4.2). These data were used to generate a scheme of the gene structure considering intron and overall size (see Fig. 23). The respective mouse and human genes show a good conserved overall structure in both 17beta-HSD type 3 and type 12. In comparison to the mammalian genes the zebrafish genes are distinctly more condensed due to shorter intron sizes. There is also another difference between the two gene groups, which is apparent for all three species: the 17beta-HSDs type 12 are by far larger compared to 17beta-HSDs type 3. In zebrafish, this size difference accounts for a factor of about 1.5 but rises up to 4 times in the mammalian genes.

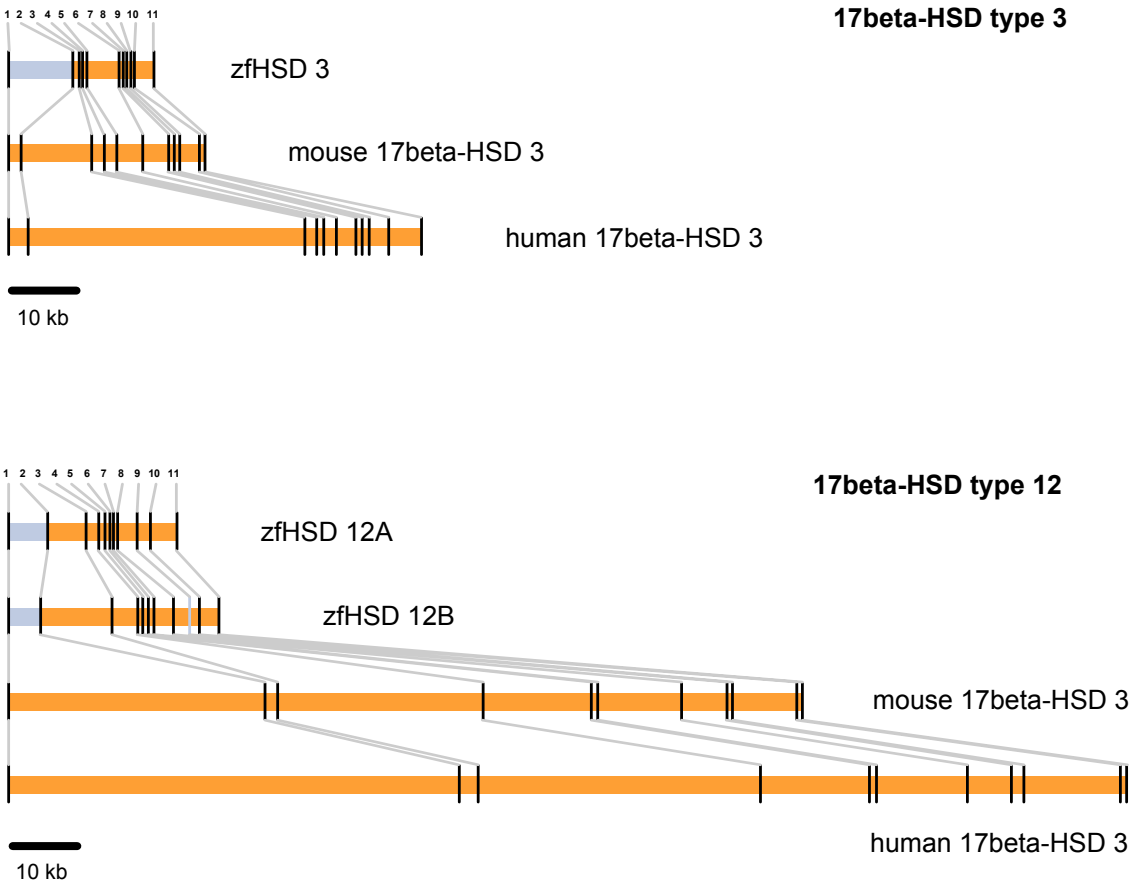


Fig. 23: Gene structures of zebrafish, mouse and human 17beta-HSD type 3 and type 12.
The coding exons are presented as black bars and introns separating them in orange. Elements in blue indicate slight uncertainties due to unfinished sequencing and mapping of the zebrafish genome. In case of the first exon of zfHSD 3, zfHSD 12A and zfHSD 12B only the minimum size is depicted according to size estimations from assemblies of WGS sequences. For zfHSD 12B the exact position of exon 9 could not be determined as no genomic sequence containing this theme could be identified.

4.2.3 Characterization on RNA level

The expression pattern of the candidate genes for the zebrafish 17beta-HSD type 3 homolog was investigated by RT-PCR (see also chapters 2.2.3 and 2.5.2 in methods). Qualitative RNA-levels were monitored in tissues of male and female adult wild type fish and during embryonic development. Data were complemented with results from *in silico* Northern.

4.2.3.1 Expression analysis by RT-PCR

In investigated stages of zebrafish embryonic development, all three zebrafish genes displayed similar expression patterns (Fig. 24). RNA from zfHSD 12A and zfHSD 3 could be detected at sphere stage already and was constantly present thereafter. Expression of zfHSD 12B could be easily monitored from 14 somite stage onwards. Although this method is not capable of giving quantitative results, signals for zfHSD 12A always were quite consistent and strong whereas other genes and especially zfHSD 12B in general produced weaker signals.

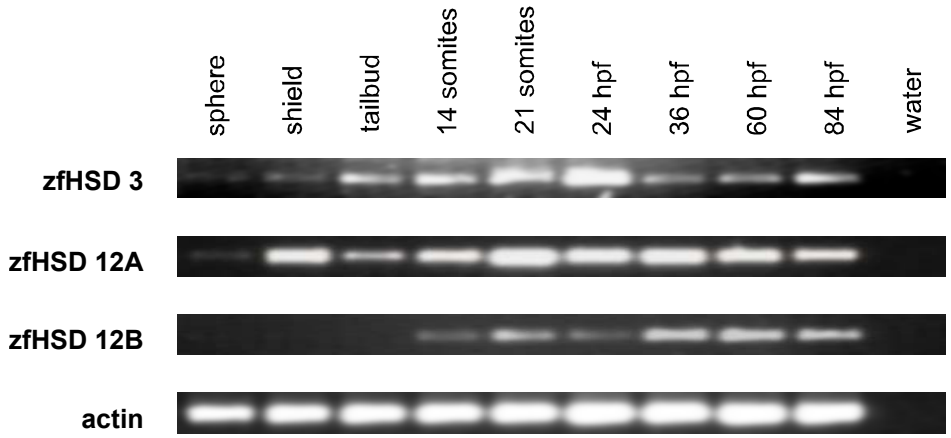


Fig. 24: Expression pattern of zebrafish 17beta-HSD type 3, type 12A and type 12B during embryonic development.
Total RNA was subjected to RT-PCR as described in chapter 2.2.3. Water (negative control) as well as actin (reference) controls were always included to enhance reliability of results and to allow for comparison of signal strength under identical conditions. hpf: hours post fertilization

Concerning organs of adult fish, all investigated zebrafish genes were found to be expressed in both sexes (Fig. 25). Sexual differences might exist especially in the expression pattern of zfHSD 3. In female fish, prominent and consistent signals were obtained from the gonads while transcript levels in liver might be significantly lower; in male fish, the reverse seems to be the case. Expression in brain of both sexes is probably present but was difficult to detect. Signals were absent from male skin as well as female spleen, kidney, heart and intestine.

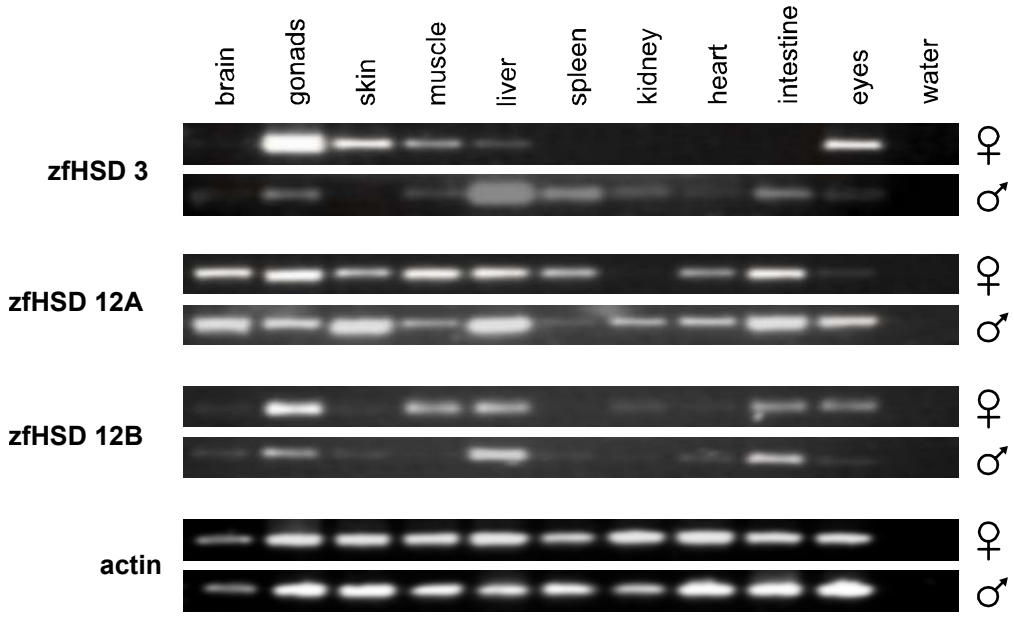


Fig. 25: Expression pattern of zebrafish 17beta-HSD type 3, type 12A and type 12B in tissues of adult male and female fish.
For sample composition and method see also chapters 2.2.3 and 2.5.2. Negative controls (water) as well as actin reference controls were always included to enhance reliability of results and to allow for comparison of signal strength under identical conditions.

Slight differences also exist between the RT-PCR expression profiles of the two zebrafish 17beta-HSDs type 12. Although in general both genes seem to be expressed in all tissues of both sexes measured, type 12A shows nearly ubiquitous moderate signal strength. Only female heart and eyes as well as male spleen produced significantly weaker signals. In contrast, transcripts of zfHSD 12B were generally difficult to detect and might even be absent in female spleen and male muscle and kidney. Moderate signals could be obtained from female gonads, muscle and liver, and male liver and intestine.

4.2.3.2 Expression analysis by *in silico* Northern blot

Not only for human and mouse but as well for zebrafish and other organisms expression information via EST databases are becoming more and more comprising. Therefore, to corroborate and expand expression information about the three zebrafish genes, *in silico* Northern (for method see chapter 2.4.2) was employed. Although the transcript sources might be quite limited by the available databases, *in silico* Northern can provide not only qualitative but also quantitative data by the general number and origin of specific ESTs.

Analysis of the *in silico* Northern data is depicted in Fig. 26. The total number of transcripts is not uniformly distributed among the three zebrafish genes. While for zfHSD 3 and zfHSD 12B eight ESTs each were identified in the databases, the double amount (16) was found for zfHSD 12A. This indicates that zfHSD 12A transcripts are more abundant and therefore the gene is most likely higher expressed than zfHSD 3 and zfHSD 12B.

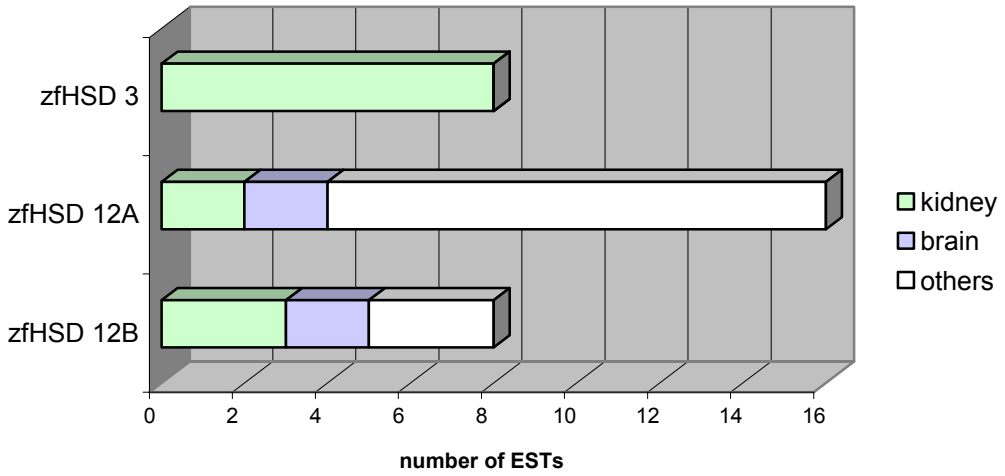


Fig. 26: Graphic analysis of *in silico* Northern data retrieved for zfHSD 3 and the two zfHSDs 12.
For each gene the total number of EST hits is depicted. The composition of library sources is given in a color code. The group “others” includes all libraries from pooled or unspecified origins.

Many transcripts originate from libraries generated from pooled tissues or even whole fish (denoted as “others” in Fig. 26). zfHSD 12A shares a high amount of ESTs originating from

these sources indicating widespread moderate levels of expression rather than a clear restriction to or prominence in a series of tissues and/or developmental stages. This effect is slightly less pronounced for zfHSD 12B where more than 60% of transcripts are from specified origin. In contrast, transcripts of zfHSD 3 were only identified in kidney libraries, indicating a moderate expression in this distinct organ.
Concerning intensity levels and overall distribution, the *in silico* Northern data reflect those obtained by RT-PCR. But it also provides additional information on the specific expression profiles: while by RT-PCR transcripts of all three zebrafish genes in kidney could be detected at best at moderate levels (zfHSD 12A, male) presence of specific mRNA in this organ is clearly sustained by *in silico* Northern.

4.2.4 Characterization on protein level

4.2.4.1 1D-structure

4.2.4.1.1 Overall sequence identity

Following complete sequencing of the zebrafish candidate coding sequences, they were translated by use of the genetic standard code. The resulting proteins were compared to the murine and human homologs for the percentage of identical residues in the complete sequence. Tab. 9 displays the identity matrix for zebrafish, mouse and human 17beta-HSDs type 3 and type 12. Interestingly, the zebrafish homolog of 17beta-HSD type 3, zfHSD 3, shows about the same degree of sequence identity to both HSD groups: ~44.5% to type 3 and ~45% to type 12. These values are comparable to those obtained by non-homologous 17beta-HSD type 3 versus type 12 sequences (mean value: ~43%). In consequence, overall sequence identity in this case does not unambiguously support affiliation of zfHSD 3 to the group of 17beta-HSDs type 3.
Both 17beta-HSD type 12 sequences show a significantly higher percentage of identical residues to this group (mean value: ~59%) than to the 17beta-HSDs type 3 (mean value: ~44%). In addition, zfHSD 12A and zfHSD 12B are more similar to each other (67%) than to the mammalian proteins, sustaining the hypothesis of a fish specific duplication event that gave rise to these two forms of 17beta-HSDs type 12 in zebrafish.
For homologous mammalian protein sequences an even higher amount of identical residues can be identified: 70% for 17beta-HSDs type 3 and 82% for 17beta-HSDs type 12. Hence, the mammalian sequences reflect a generally higher conservation in the group of 17beta-HSDs type 12 as was found for the zebrafish sequences.

Tab. 9: Identity matrix of 17beta-HSD type 3 and type 12 sequences of zebrafish, mouse and human proteins.

17beta-HSD type 3:							
zfHSD 3	100						
mouse	44	100					
human	45	70	100				
17beta-HSD type 12:							
zfHSD 12A	48	40	46	100			
zfHSD 12B	46	40	43	67	100		
mouse	46	40	43	57	57	100	
human	41	43	46	59	56	82	100
	zfHSD 3	mouse	human	zfHSD 12A	zfHSD 12B	mouse	human
	17beta-HSD type 3			**17beta-HSD type 12**			

Pairwise alignments were performed with Blast2seq; amino acid identities are listed in percent.

4.2.4.1.2 Phylogenetic analysis

The zebrafish proteins were set into correlation to a broad set of related sequences to assess their homology. Results derived from these phylogenetic analyses contributed in great part to the classification and naming of the zebrafish candidate genes.

The phylogenetic tree (Fig. 27), that was calculated by means of a Neighborjoining algorithm, displays several distinct groups. As the plant 3beta-ketoacyl reductases (3KAR) are the most distantly related sequences in this set, the tree was rooted at this bifurcation. The group closest to the plant enzymes comprises a number of enzymes from *C. elegans*. Among them LET-767 is located, whose homology to vertebrate 17beta-HSDs type 3 and type 12 has already been suggested. The fact that LET-767 rather clusters with other proteins from C. elegans instead of being located in the 17beta-HSD 12 group reflects that this C. elegans group in itself is closer related than each single sequence is related to the 17beta-HSD 12 group. This does not contradict the assumption that LET-767 is the C. elegans homolog of vertebrate 17beta-HSD type 12 and also reflects the accumulated occurrence of gene duplications typical for *C. elegans*. The same reason applies to yeast YBR159 located in the adjacent group (protozoan group I), which is the yeast homolog of 17beta-HSD type 12.
The most prominent groups and clearly sustained by high bootstrap values are the vertebrate 17beta-HSDs type 12 (blue) and type 3 (orange). The two zebrafish 17beta-HSDs type 12 appear as members of the former group whereas zebrafish 17beta-HSD type 3 matches to the latter group. This result clearly demonstrated that indeed a zebrafish homolog of 17beta-HSD type 3 and in addition two paralogous 17beta-HSDs type 12 could be identified.

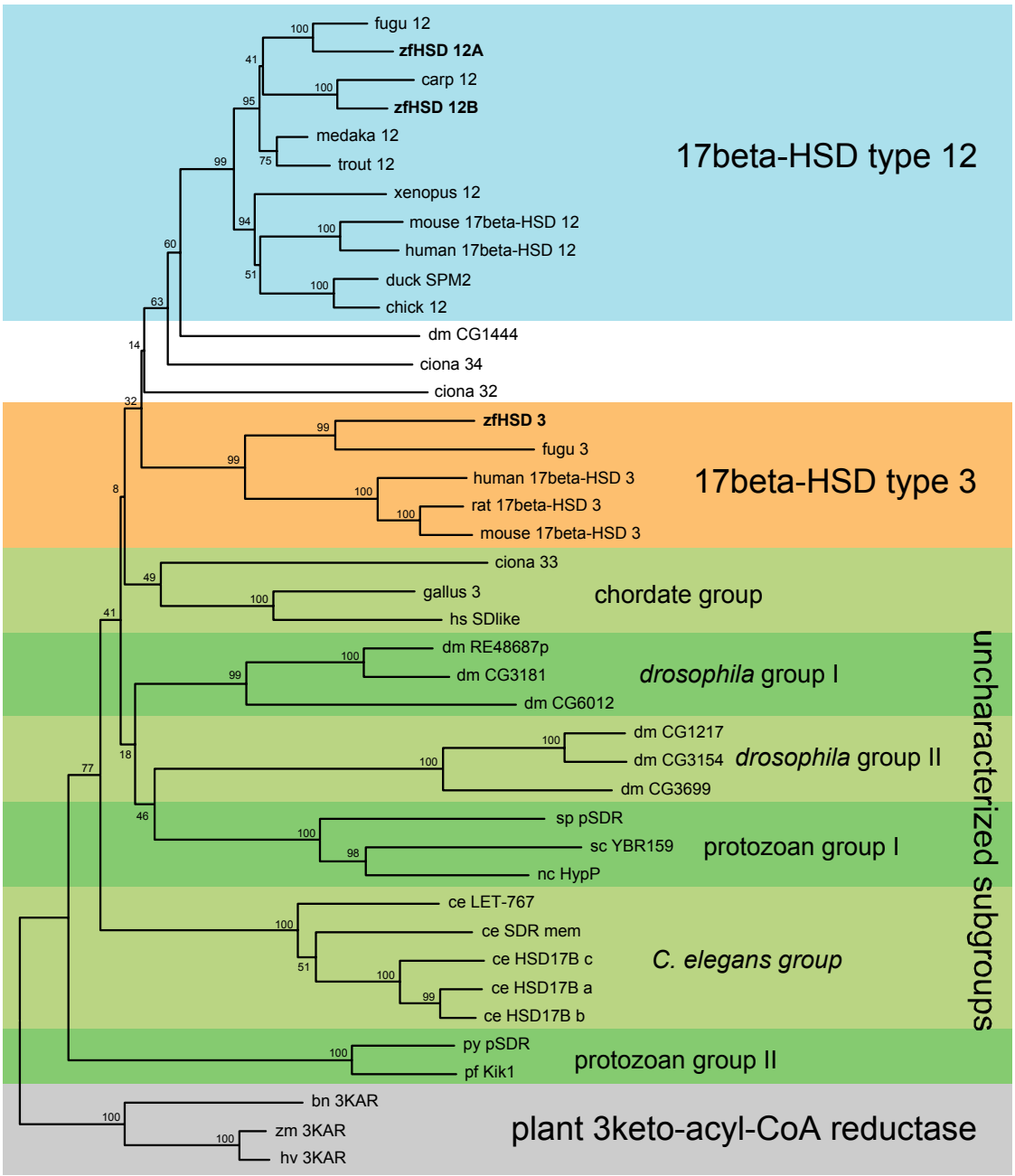


Fig. 27: Evolutionary background of vertebrate 17beta-HSD type 3 and type 12.
Distinguishable subgroups are outlined in color: 17beta-HSDs type 12 are marked in blue, 17beta-HSDs type 3 in orange and groups of yet uncharacterised sequences in green. The tree was calculated by means of Neighborjoining with bootstrap values of a 1000 pseudoreplicates, given in percent at the individual dichotomies. The tree was rooted with the plant 3keto-acyl-CoA reductases as outgroup (gray). The protein sequences employed for this calculation are listed in the appendix (chapter 6.3).

Among the vertebrate sequences available in the database, no other 17beta-HSD type 3 homologs could be identified. Only one fugu sequence (fugu 3) was identified in a direct search against the genomic database. In the 17beta-HSD type 3 group, the two fish sequences appear

in direct vicinity forming their own distinct subroup. This reflects a closer relation among these fish enzymes compared to the mammalian homologs.
For the family of 17beta-HSDs type 12, a number of sequences could be identified enabling a more detailed analysis. Again, all fish sequences appear in a separate group apart from the rest of the vertebrate sequences, forming their own cluster. Probably, the same explanation as for the C. elegans sequences applies here, too: the fish proteins (at least of the teleosts presented here) have diverged less than the vertebrate proteins. In addition, location of the two paralogous zebrafish 17beta-HSDs type 12 on separate branches indicates the underlying duplication event not to be zebrafish specific. As a consequence, duplicates may be found as well in other fish. Alternatively, due to loss of one or the other duplicate, differences between fish 17beta-HSDs type 12 may exist.

4.2.4.1.3 Conserved motifs

Like the mammalian homologs, all zebrafish proteins analyzed here belong to the family of short chain dehydrogenases/reductases (SDRs). The presence and identity of characteristic SDR motifs (see also chapter 1.3.2) is investigated and also compared to the mouse and human homologs. In addition, amino acids that might have become typical for the function of either 17beta-HSDs type 3 or type 12 are explored.

The comparison of all sequences (Fig. 28) led to the identification of residues common to both groups of 17beta-HSDs. Most of these amino acids (~90%) are located in the cofactor binding region, contributing to the motifs characteristic to SDRs. In the very N-terminus there is only one conserved glycine present, which is isolated in the 17beta-HSDs type 3 and part of a preserved cluster in the 17beta-HSDs type 12. Although the first motif starts with a highly conserved threonine, four more amino acids N-terminal of this residue are also identical to both groups of 17beta-HSDs and were therefore counted as belonging to the cofactor binding part.

SDR-motifs involved in alignment and binding of the cofactor
Motif (1), which conveys the binding of the cofactor, is highly conserved. Beside the four central amino acids (capital letters), two more residues are identical in both enzyme groups. The same applies for motif (2), that is also involved in cofactor binding. Here, the central aspartic acid is 100% conserved as is the subsequent phenylalanine. A more structural role, namely stabilization of the central β-sheet, is executed by motif (3). Although nine of 12 residues in the displayed 17beta-HSDs type 3 and 12 are identical, not all of these fit the consensus motif: the classical NNAG motif is displaced by NNVG and the length of the pattern is stretched to thirteen amino acids by insertion of one residue most likely after the first glycine. Interestingly, this additional residue is highly conserved and either resembled by the 100% preserved leucine or the subsequent acidic amino acid. The aspartic acid, another central residue of the motif, is lost and maybe replaced by the small and hydrophobic but completely conserved glycine. This non-conservative exchange of a presumably functionally important amino acid is also apparent in the next motif (motif (4)). Here, the glycine of the consensus sequence is changed to a preserved serine.

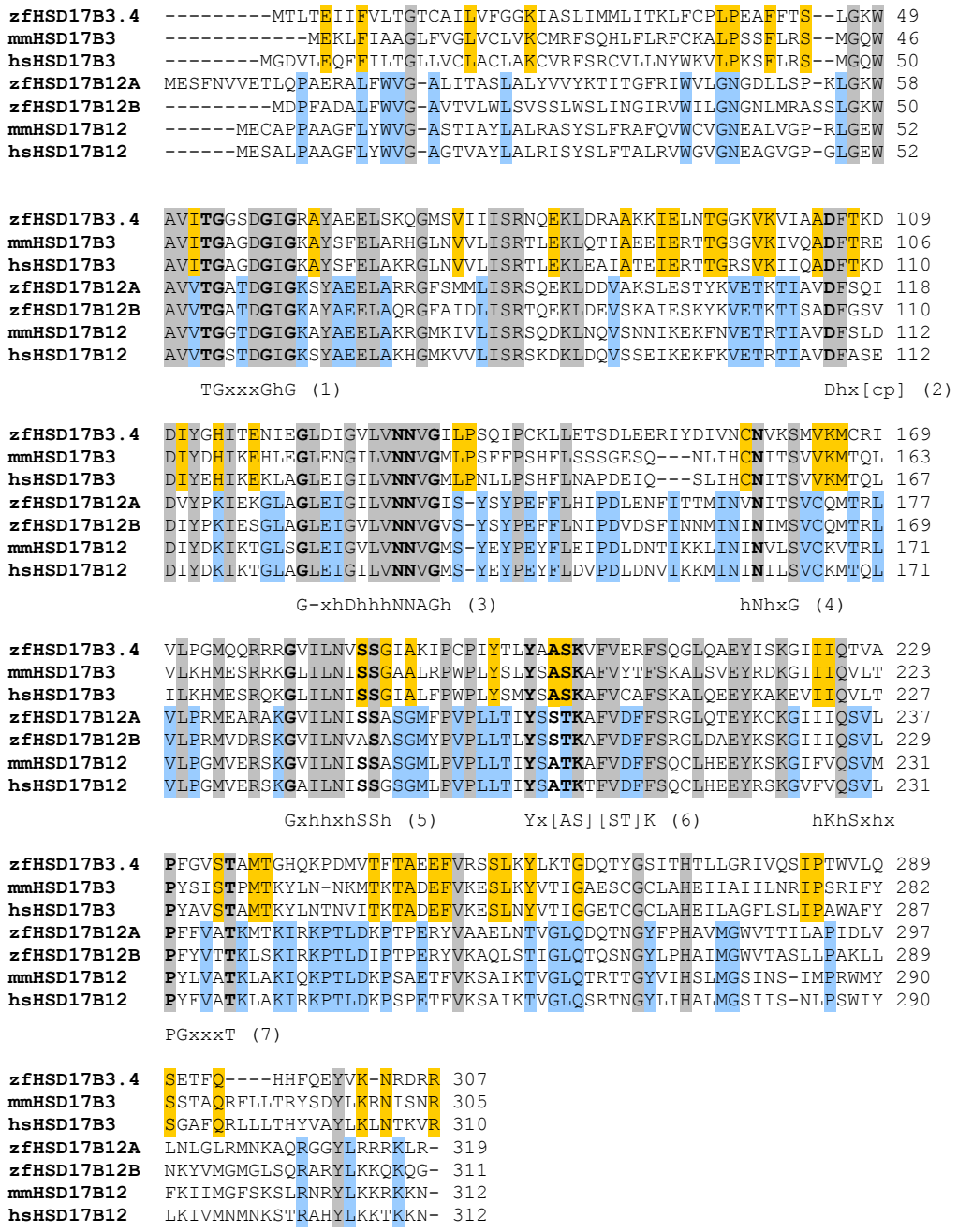


Fig. 28: Comparison of SDR-motifs in a multiple protein sequence alignment of zebrafish, mouse and human 17beta-HSDs type 3 and 12.
Residues identical in all sequences are marked in gray. Orange shading indicates amino acids preserved in the group of 17beta-HSD type 3 and blue shading those residues conserved in the group of 17beta-HSD type 12. The consensus sequence of motifs characteristic to the family of SDR proteins are given below the sequence where capital letters denote residues strictly conserved for the motif (central residues) and small letters indicate less conserved amino acids. Letters in brackets are alternative amino acids for the same position. c: charged residue; h: hydrophobic residue; p: polar residue; x: any residue.

SDR-motifs involved in modeling of the active site

Motifs (4), (5) and (6) are part of the active site, conveying orientation of the active site and the transfer of the hydride ion. Consequently, residues of both 17beta-HSD groups fit the consensus sequence of motif (5) very well. Only the first of the two serine residues is changed

to alanine in zfHSD 12B. Moreover, three amino acids inside the motif but part only of the extended pattern are completely conserved in all sequences presented here, demonstrating the close relation of the two groups of 17beta-HSDs. Tyrosine and lysine, which take part in the transfer of the hydride, are 100% preserved as are the less stringent residues of motif (6).
Finally, motif (7) is the last pattern common to all SDRs and therefore may be assumed to terminate the predominantly cofactor binding part of the enzyme. The central residues of this theme are proline and glycine as well as threonine, as they interact with the cofactor. Of these only glycine is not conserved in the 17beta-HSDs analyzed here. Instead, it is consequently displaced by aromatic residues, namely either tyrosine or phenylalanine. The N-terminal part of this motif is highly altered compared to the consensus pattern: (i) the proposed central lysine was either lost or has taken on a further upstream position by introduction of three residues in all seven sequences, (ii) the central serine was either displaced by a completely conserved glutamine or has been shifted downstream by insertion of one amino acid, where it is preserved only in the group of 17beta-HSD type 12.
These findings suggest that in the zebrafish proteins motifs of classical SDRs are present and resemble the amino acid composition of the respective mammalian homologs. Based on this, the zebrafish proteins appear to be functional HSDs. Concerning the SDR motifs, a close relation between 17beta-HSDs type 3 and type 12 is also apparent.

Conservation in 17beta-HSDs type 3 and type 12 N- and C-terminal protein parts
When the comparison is zoomed in on each of the two 17beta-HSD groups instead of an overall alignment, residues conserved and characteristic for either 17beta-HSD type 3 or type 12 emerge (see Fig. 28, colors orange and blue, respectively). Overall sequence identity rises from 21% to 39% in case of 17beta-HSDs type 3 and 46% in case of 17beta-HSDs type 12, assuming a higher conservation for the latter enzyme family. Interestingly, the arrangement of these group-specific amino acids is neither uniform along the linear sequence nor similar when comparing the two different HSD groups. In case of 17beta-HSD type 12, 27% of residues in the cofactor and 28% in the substrate binding part are group specific; about 20% are type 12 specific in the N-terminus. In contrast, 17beta-HSDs type 3 harbor less amino acids characteristic to this group: 18% in the N-terminal part, 17% in the cofactor binding region and 24% in the substrate specific area. As a result, residues specific for 17beta-HSDs type 12 are evenly distributed in cofactor and substrate binding region while in 17beta-HSDs type 3 conserved residues rather cluster in the substrate binding part.

4.2.4.1.4 Distribution of identical residues

Examination of the multiple alignment in the previous chapter has revealed intricate modes of conservation that differ in the functionally dissimilar enzyme groups and parts of the respective proteins. Identity plots were calculated as described in chapter 2.4.4 to better illustrate and investigate the clusters and distribution of conserved amino acids in both groups of 17beta-HSDs.

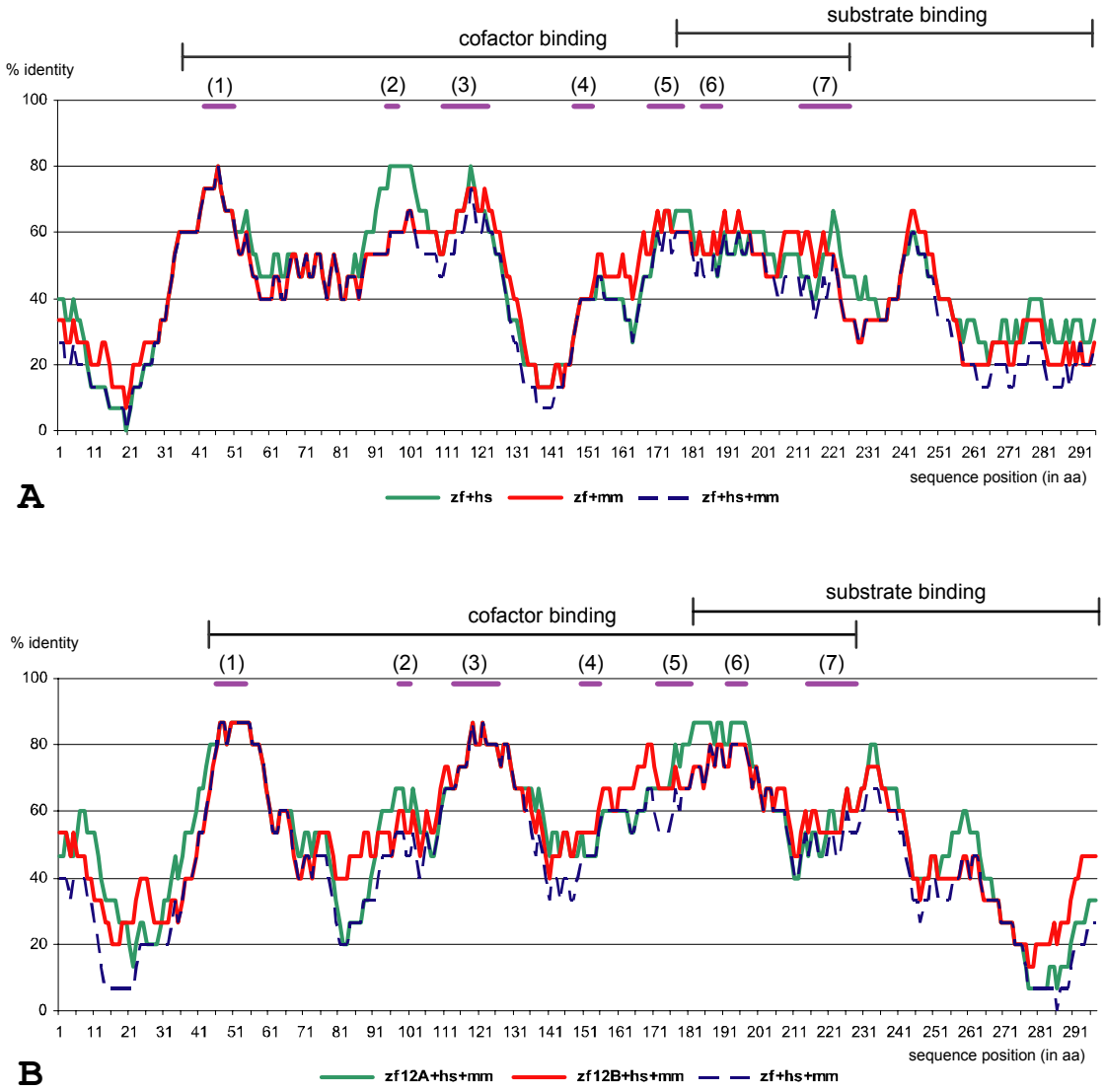


Fig. 29: Identity plots for the sequences of 17beta-HSDs type 3 (A) and 17beta-HSDs type 12 (B).
The percent identity of a 15-residue block is plotted against the first amino acid of this block; as a result, plots are cut off fifteen residues before the end of the alignment. The SDR-motifs as described in the previous chapter are inserted in violet with the respective motif number. Their position was shifted 7 aa positions to the left compared to the alignment in Fig. 28 to level out the slight distortion introduced by the 15-residue block calculation. The regions predominantly binding the cofactor and the substrate are outlined above each plot. For the cofactor binding region, borders were set in a way to mainly comprise the SDR motifs as described in the previous chapter. The substrate binding region was presumed to initiate at the end of motif (5) similar to the situation in 17beta-HSDs type 1 (see also chapter 4.1.4.2.1). Due to lack of specific information regarding substrate binding in these two enzyme groups, the end of the predominantly substrate binding region was set at the end of the sequence. The caption points out the different sequences that were taken into account for the respective comparison. hs: homo sapiens; mm: mus musculus; zf: zebrafish; zf12A: zfHSD 12A; zf12B: zfHSD 12B.

At first glance, the course of conservation for each group of 17beta-HSDs, as shown in Fig. 29, appears to be quite similar to each other. Variations are mainly present in the actual strength of conservation at a given position. In both cases about the first fifteen residues are moderately conserved and afterwards identity drops off to a minimum at position 20. Subsequent to this minimum, conservation rises steeply to a maximum of about 80% identity, including the first important SDR-motif which conveys cofactor binding (motif (1)). A similar strength of conservation can be monitored in case of motif (3). In 17beta-HSDs type 12 both peaks (motif

(1) and motif (3)) are broader than in 17beta-HSDs type 3 reflecting that in the first group more amino acids neighboring these motifs are conserved leading to the formation of extended clusters whereas in the latter group identicals of motifs are more isolated. The region connecting motifs (1) and (3) in both HSD families share about 50% identical residues. Interestingly, in 17beta-HSDs type 3 the zebrafish protein is quite conform to both mammalian enzymes except for the region comprising aa 85-105. Here, identical amino acids rise from about 60% in comparison to mouse to a maximum of 80% compared to man, in this way also including motif (2). Though this course is also present, albeit more subdued in the 17beta-HSDs type 12, the preceding region is in this case far more outstanding: while zfHSD 12A shows only low identity to the mammalian proteins, zfHSD 12B remains at a more or less constant level of 50% conservation. Since this region in SDRs is suggested to convey discrimination between the two putative cofactors NAD and NADP, dissimilar cofactor preferences for the individual enzymes in both HSD groups may be possible.
The second half of the cofactor binding area, including motifs (4)-(7) is less marked by outstanding peaks and takes a similar course in both HSD groups although in general, the amount of conservation is slightly higher for the 17beta-HSDs type 12. There are minimal variations in both cases stating zfHSD 3 to be sometimes more similar to the human (aa 220-235) or the mouse (aa 155-175) protein; similarly, to the mammalian 17beta-HSDs type 12 enzymes sometimes zfHSD 12A shows higher sequence identity (aa 180-195) and sometimes zfHSD 12B (aa 160-175). Apparently different is the connection between these two parts of the cofactor binding region. 17beta-HSDs type 3 display a deep rift stretching about fifteen residues from aa 130-145. In the 17beta-HSDs type 12 this region is altogether higher conserved. In the C-terminal region of the 17beta-HSDs type 3, subsequent to a peak in identicals common to all three sequences around position aa 245, the amount of amino acid preservation drops to roughly 30% but is slightly higher for the zebrafish-human comparison than zebrafish-mouse.
In the 17beta-HSDs type 12, two peaks appear in the corresponding area: the first slightly upstream at positions 225-235, the second slightly downstream (aa 250-260). Especially the second peak is much more pronounced for zfHSD 12B in comparison to the mammalian proteins whereas zfHSD 12A is less preserved. After a decline in conservation a third peak follows in the very C-terminal part of the protein. In case of zfHSD 12B, the amount of identical residues in the last fifteen amino acids rises to nearly 50% and about 30% in case of zfHSD 12A.

4.2.4.2 3D-structure

4.2.4.2.1 Expression of recombinant protein

Expression of the corresponding proteins of the zebrafish genes is necessary to explore their substrate specificities. For this purpose, the coding sequences were cloned into expression vectors attaching a C-terminal GST- (pGEX vector) or His_6-tag (pQE 30 or pET15b) and the constructs transformed into *E. coli* (for method see chapter 2.1.3). Expression of recombinant proteins was monitored via SDS-PAGE and Coomassie stain and these findings were substantiated by Western Blotting (Fig. 30).

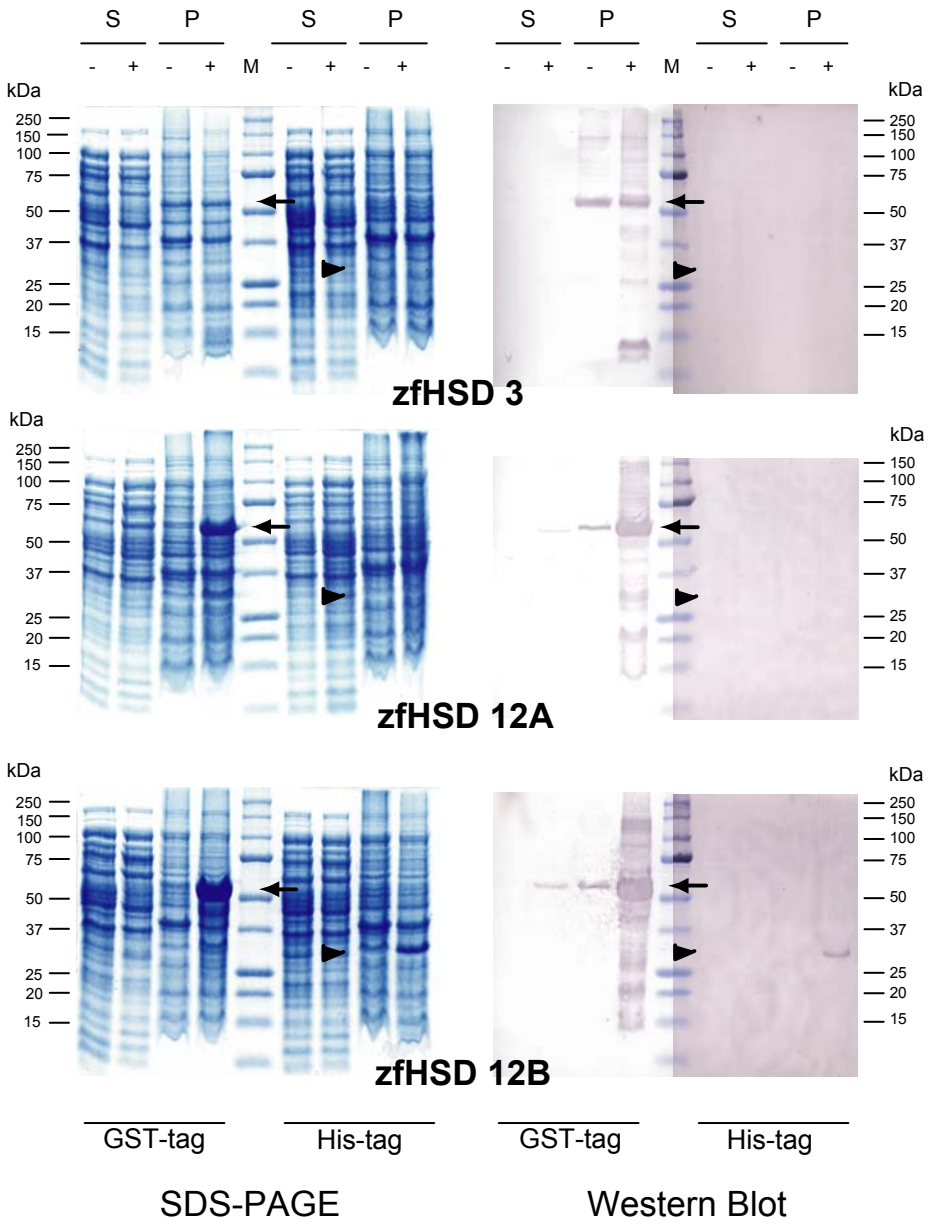


Fig. 30: Expression of recombinant zfHSD types 3, 12A and 12B.
The left side depicts Coomassie stained SDS-PAGE gels that show the various fractions after induction of expression for both kinds of protein-tags. The right side shows Western Blots of identical gels which were probed with monoclonal antibodies against the respective tag (for method see chapter 2.3.3). Arrows point at the GST-fusion proteins; arrowheads mark the position where the His_6-tagged protein would be expected. M: molecular weight marker; P: pellet fraction; S: soluble fraction; - : without induction by IPTG; + : with IPTG induction.

Following SDS-PAGE and Coomassie staining apparent production of recombinant protein is visible only in case of the GST-fusion proteins of zfHSD 12A and 12B at about 61 kDa in both cases. Arrows in Fig. 30 point at the strong band present in the fraction of pelleted proteins after lysis. In case of zfHSD 3 (~60 kDa), a signal at the expected size is also visible but present in both pelleted fractions and not enhanced by induction with IPTG. Western Blots with antibodies against GST demonstrate that the strong signals for both zfHSDs type 12 and the medium signal for zfHSD 3 are indeed due to expression of the fusion protein. In addition, both zfHSDs type 12 are shown to be present as well in the soluble fraction though to a much lower extent.

For His_6-tagged proteins Coomassie stain could scarcely reveal any formation of recombinant protein (arrowheads in Fig. 30 point at the expected size of about 34 kDa in case of zfHSD 3 and 35kDa in case of the zfHSDs type 12). Only in case of zfHSD 12B, protein might be present in the insoluble fraction after induction. These findings are corroborated by the Western Blot employing antibodies against the His_6-tag. Detection of His_6-tagged protein failed with the exception of zfHSD 12B. Here, the presence of recombinant protein already slightly visible after SDS-PAGE and Coomassie staining was confirmed.

4.2.4.2.2 Activity measurements and substrate specificity

The potential of the zebrafish proteins to convert the substances typical for the corresponding mammalian enzymes was investigated via bacterial lysate (pellet fractions, see also previous chapter) that contained the over-expressed recombinant proteins. A reaction characteristic to the known mammalian 17beta-HSD type 3 is the reduction of androstenedione to testosterone by use of NADPH as cofactor. *In vitro*, the back-reaction can also be measured by addition of NAD.

Neither zfHSD 3 nor the two types of zebrafish 17beta-HSD 12 converted androstenedione to testosterone in this assay set-up. Concordantly, the reverse reaction was also not catalyzed by the bacterial lysates containing the overexpressed zebrafish enzymes.

5 Discussion

5.1 17beta-HSD type 1: zebrafish vs. human vs. mouse

5.1.1 In every respect highly conserved

In search for a zebrafish homolog of the 17beta-HSD type 1, three candidate sequences were initially identified and selected for further analysis. The gene architecture of all three genes resembled that of known human, mouse and rat isozymes, thereby demonstrating evolution from a common progenitor. Overall sequence identity on protein level suggested one candidate, zf 1.2, to share more identical residues to mouse and human 17beta-HSD type 1 than the other two zebrafish sequences. Finally, the identity of zebrafish 17beta-HSD type 1 (zfHSD1) was consolidated by phylogenetic analysis clearly demonstrating zf 1.2 to be part of this group whereas the two other zebrafish genes were calculated to be homologous to mammalian photoreceptor-associated retinol dehydrogenases. But not only on sequence level does the zebrafish isozyme resemble a 17beta-HSD type 1. Indeed, characterization especially on RNA and protein level revealed features in common with already characterized 17beta-HSDs type 1 from other organisms.

5.1.1.1 Expression might be absent during embryogenesis

There is not much known about 17beta-HSD type 1 expression during mammalian embryogenesis. Mustonen et al. (1997) investigated occurrence of transcripts in mouse embryogenesis at days E 7, 11, 15, and 17 and detected a signal solely in stage E 7 (onset of/early gastrulation). Use of more sensitive RT-PCR also detected 17beta-HSD type 1 transcripts specifically in the testis of fetus age E 15.5 (Sha et al., 1996). In the human fetus, 17beta-HSD type 1 mRNA might be present at extremely low levels at gestational week 13 and 20 (Takeyama et al., 2000). During pig embryogenesis, 17beta-HSD type 1 might be expressed in day 12 embryos, as specific ESTs can be identified by *in silico* Northern (ESTs: AW620093, AW619829). This developmental stage is preceding implantation and gastrulation. These findings indicate that in mammals 17beta-HSD type 1 may be transiently expressed before or in early stages of gastrulation but does not seem to be expressed during the following stages of embryogenesis. Only in later, fetal development expression recurs in specific organs (e.g. mouse testis). Similar findings were obtained for 17beta-HSD type 1 expression during chicken development. There are no data available about gastrulation and preceding stages of

embryogenesis but specific transcripts were detected at 2-9 days of incubation (Nomura et al., 1999) comprising somitogenesis and the main period of organogenesis (Hamburger and Hamilton, 1992). Differentiation of gonads starts at about the 6th day of incubation (Romanoff, 1960) and in consistence, 17beta-HSD type 1 was detected by RT-PCR in developing female gonads at incubation days 7-14 (Nakabayashi et al., 1998). At this time, expression was also present in male developing gonads though at a significantly lower level (Nakabayashi et al., 1998).

These data from mammals as well as chicken are in synchrony with the absence of zebrafish 17beta-HSD type 1 mRNA during embryogenesis (up to 60 hpf) and transition to larval stages (84 hpf) observed in this study. Gonadogenesis leading to ovaries and testes starts at about 15 days post hatching (dph)(Uchida et al., 2002) and may continue up to 35 dph (Orn et al., 2003). This time-frame was not studied in this work and it may be that zfHSD 1 is expressed during this period as was found for other vertebrate homologs. Concerning possible expression during gastrulation, only mRNA from the onset (shield stage) and end (tailbud stage) of this process were investigated by RT-PCR. The sphere stadium resembles a blastula stage but it is difficult to exactly bring this developmental stage in line with the mammalian phases of embryogenesis as described above. Hence, if expression of zfHSD 1 exists during the corresponding early embryonic development the exact time point might have been missed. Alternatively, expression itself might have been extremely low and/or restricted to certain tissues and not detected due to the experimental design.

5.1.1.2 Enhanced but not exclusive expression in tissues of female adults

In all adult organisms studied for 17beta-HSD type 1 expression, transcripts were present in high amounts in organs and tissues directly associated with female sex-specific functions, i.e. ovaries, placenta, uterus and breast tissues. Especially occurrence of transcripts in ovaries seems to be common to all so far investigated species. This was demonstrated by Northern Blot in chicken (Wajima et al., 1999), mouse (Mustonen et al., 1997) and rat (Akinola et al., 1996), by RT-PCR in eel (Kazeto et al., 2000a) and zebrafish (in this work), and by RNase protection assay in human (Labrie et al., 1997). Comparison of the respective data indicates a very strong influence of the method chosen for detectability. Notably, using Northern Blot analysis 17beta-HSD type 1 was exclusively detected in ovaries of chicken and mouse reflecting high expression levels in this tissue in the respective organism. Employing more sensitive RT-PCR or RNase protection assay, presence of transcripts is also demonstrated in prostate and testis (Labrie et al., 1997) and at other sites such as brain (Stoffel-Wagner et al., 1999; Beyenburg et al., 2000) and kidney (Akinola et al., 1996). Furthermore, employing *in silico* Northern blot, presence of specific transcripts can be revealed in mouse pancreas and human skin (data not shown). Altogether, 17beta-HSD type 1 seems to be highly but not exclusively expressed in female specific tissues. Reported absence in other tissues might at least in part be due to the low sensitivity of the method selected for analysis.

The expression profile of zebrafish 17beta-HSD type 1 identified in the study on hand strongly resembled that found in other vertebrates. The highest amount of transcripts might be present in the female gonads as RT-PCR readily gave rise to a strong signal. But RT-PCR signals were also obtained from other organs, namely skin, muscle, heart and eyes. During identification and complementation of the zebrafish candidate genes a clone containing parts of zebrafish 17beta-

HSD type 1 was obtained from a liver cDNA library by screening with a specific probe (data not shown). *In silico* Northern revealed presence of 17beta-HSD type 1 mRNA in the kidney. Hence, 17beta-HSD type 1 most likely is also expressed in the zebrafish liver and kidney. Expression of 17beta-HSD type 1 was not detected in fish brain or any male tissue investigated. As aforementioned, presence of 17beta-HSD type 1 transcripts in these tissues can not be rigorously excluded due to the method chosen for investigation since low expression levels might have interfered with successful detection. Findings from my study altogether suggest that expression of zebrafish 17beta-HSD type 1 closely resembles that of other homologs from mammals, as well as birds (chicken) and other fish (eel).

5.1.1.3 Functional integrity: assessing the flexibility of SDR consensus motifs

That extensive residue identity on amino acid sequence level between 17beta-HSD type 1 from zebrafish and from other organisms exists, is clearly demonstrated by the phylogenetic analysis. A bootstrap value of 100% clearly supports membership of the zebrafish protein sequence in the family of 17beta-HSDs type 1. Nevertheless, functional integrity and substrate specificity was not predictable by these analyses. SDR motifs were thoroughly investigated not only to demonstrate affiliation of the zebrafish enzyme to the SDR family but also to assess its theoretical functionality. Investigation of the SDR motifs in an alignment additionally comprising the sequences of the characterized homologs from rat, chicken and eel (Fig. 31) demonstrates extensive concordance to the consensus sequence.

```
human       TGCSSGIG...DVRD...GRVDVLVCNAGL...VNVVG...GRVLVTGSV...YCASK...VHLSLIECGPVHT
mouse       TGCSSGIG...DVRD...GRVDVLVCNAGR...VNVLG...GRVLVTASV...YCASK...VHVSLIECGAVHT
rat         TGCSSGIG...DVRD...GRVDVLVCNAGR...VNVLG...GRVLVTASV...YCASK...VHVSLIECGAVHT
chicken     TGCSSGIG...DVTD...RHPDVLVCNAGV...VNVFG...GRILVSSSV...YCASK...IHMTLVECGPVHT
eel         TGCSSGIG...DVTD...NRVDILVCNAGV...VNLLG...GHVLVTGSI...YCASK...IHVSLIESGPVNT
zebrafish   TGCSSGIG...DVTD...GRIDILVCNAGV...VNLLG...GRILVTGSM...YCASK...IHISLIECGPVNT

consensus   TGxxxGhG...Dhxc...GxhDhhhNNAGh...hNhxG...GxhhxhSSh...YxASK...hKhSxhxPGxxxT
                          *      *   *               **             * *   *
```

Fig. 31: Classical SDR motifs in 17beta-HSDs type 1 from fish, birds and mammals.
Residues identical in all aligned sequences are shaded in blue. Deviations to the consensus sequence are marked with asterisks. Amino acids central to the consensus are in capital letters while small letters denote more flexible residues. c: charged aa; h: hydrophobic aa; x: any aa.

Furthermore, variations are neither restricted to the zebrafish sequence in comparison to human and mouse nor do they affect only the central or the more flexible residues of a given motif. Since all these 17beta-HSDs type 1 aligned in Fig. 31 have been shown to be active, the observed alterations are not likely to influence enzymatic integrity (at least in the cofactor binding part) of the zebrafish protein. Enzymatic functionality of zfHSD 1 could indeed be verified by activity measurements as described in chapter 4.1.4.2.3.

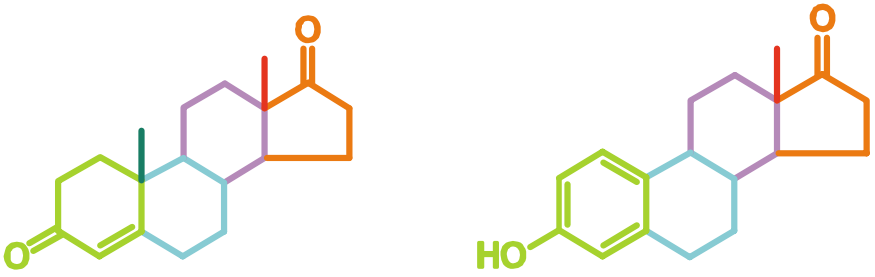
5.1.1.4 Substrate specificity: dissecting the definitely important residues

All 17beta-HSDs type 1 characterized for their substrate specificity have shown a high enzymatic activity towards estrone, catalyzing its reduction to estradiol by use of NADPH as cofactor. The rat and mouse enzyme additionally catalyze reduction of androstenedione to testosterone with a comparable efficiency (Akinola et al., 1996; Nokelainen et al., 1996). The recombinant zebrafish protein of 17beta-HSD type 1 demonstrated a high activity towards estrone and estradiol *in vitro*, readily catalyzing the redox reaction leading to nearly 100% product formation under the experimental conditions. Furthermore, androgens were not accepted as substrates, supporting the notion that this characteristic was developed exclusively in rodents.

The identity of specific residues that account for the recognition and binding of an estrogen substrate has been investigated for years and was strongly promoted by crystallization of the human 17beta-HSD type 1 complexed together with NADP and estradiol (Ghosh et al., 1995; Azzi et al., 1996). The study of Puranen et al. (1997b) made use of the extensive sequence homology between the human and rat enzyme identifying regions and residues responsible for additional androgen recognition in rat by cloning and characterization of chimeric enzymes. Extensive study of the crystal structure was performed by Azzi et al. (1996) and Lin et al. (1999), and these findings were discussed and deepened by Ghosh and Vihko, (2001). In this process a variety of amino acids possibly responsible for estrogen specificity and its orientation in the substrate binding pocket of 17beta-HSD type 1 were identified. Finally, Nahoum et al. (2003) formulated less strict rules for estrogen recognition, comparing several mammalian estrogen binding proteins in addition to human 17beta-HSD type 1.

From these sources I extracted the information about amino acids considered to be critical for estrogen recognition and summarized them in a graphic overview as depicted in Fig. 32. This should allow for comparison and expansion of the data gathered on the mammalian enzymes to that of birds and fish, including the zebrafish. In this respect, analysis of non-mammalian proteins had not been carried out before.

It becomes quite clear that the data produced in strict reference to the mammalian enzymes are corroborated but in some cases also slightly revised by integration of the zebrafish data from this work and additionally the sequences of other non-mammalian enzymes proven to specifically bind and convert estrogens. Three structural elements are different between estrogens and androgens and therefore the composition of the binding fold in these areas is especially important for successful discrimination: the A-ring, nature of the oxy group at C3 and existence of a C19 methyl-group. Binding of the planar, aromatic ring in estrogens and at the same time steric hindrance of a C19 methyl-group is thought to be generated by two hydrophobic residues on either side of the ring in a sandwich-like manner. Nahoum et al. (2003) suggest these residues in human 17beta-HSD type 1 to be L150 and V 226. Both positions were also identified in the homology model of zebrafish 17beta-HSD type 1. As L150 is also conserved in the two rodent homologs, additionally capable of androgen conversion, V226 might be the crucial residue in this case though amino acid identity in this position is obviously not necessarily strict since in the chicken 17beta-HSD type 1 it is replaced by the even larger leucine. The smaller and obviously conserved alanine in the two rodent enzymes seems to be necessary to give enough room for the non-planar A- ring of androstenedione.




```
human      TGSVGGLMGLPFNDVYCA-158..ECGPVHTAFMEK-196..YQYLAHSKQVFRE-229..ERFLP-262..TAMHREVF-285
mouse      TASVGGLMGLPFHEVYCA-158..ECGAVHTAFYEK-196..AHYLRGYEQALSE-229..NRFLP-261..AAMHQEAF-284
rat        TASVGGLMGLPFHEVYCA-158..ECGAVHTAFHEK-196..AHYQRGYEQALSE-229..NRFLP-261..EAMHREAF-284
chicken    SSSVGGLQGLPFNAVYCA-163..ECGPVHTAFMDN-201..RRYLQHCTQLFLD-236..RLPAP-269..RAMHDFVF-292
eel        TGSIGGLQGLPFNELYCA-156..ESGPVNTHFLSN-194..DTYLQHCSSVFHS-229..SALLP-262..RAMSKLIF-285
zebrafish  TGSMGGLQGLPFNEVYCA-156..ECGPVNTDFLMN-194..DQYLQHCQSVFQN-230..RALLP-263..RAMSKLIF-286
             * *    *  *  *          **                      *           *          *  *
```

Fig. 32: Identity of positions and residues supposed to convey estrogen specificity comparing the mammalian 17beta-HSDs type 1 to their homologs in chicken, eel and zebrafish.
All positions suggested to be involved in estrogen specificity are written in bold letters; in addition a color code marks the respective position where interaction is supposed to take place. To highlight positions as well as residues completely conserved to all sequences, color shading refers to the human enzyme and encloses only amino acids from other species if they are identical with these. Asterisks mark positions within 4 Å of estradiol in the zebrafish 17beta-HSD type 1 as identified by homology modeling in this work (see also chapter 4.1.4.2.1). The alignment was corrected by the 3D-homology modeling data to resemble a structure rather than a sequence based alignment.

In consistence with hydrogen bonds recognizing the hydroxy group at C3 in estrogens, H222 is conserved in all but the additionally androgen-metabolizing 17beta-HSDs type 1. It was suggested that two conserved residues are necessary to recognize and fix this hydroxy group through a bifurcated hydrogen bond. But the second residue, E283, is neither completely conserved nor exclusive for the estrogen binding isozymes and may therefore not be critical for estrogen recognition as was also suggested by Ghosh and Vihko (2001).
The correct alignment of the substrate near the active center requires stabilization of the D-ring and the C18 methyl-group. The alignment shows that for interaction with the D-ring residues S143, Y156 and F193 are indeed important as they appear to be strictly conserved. In fact, S143 and Y156 are part of the catalytic triad and essential for general functional integrity of members of the short-chain dehydrogenase/reductase (SDR) family (see also chapter 1.3.2). In contrast, the role of M194 might be less strict as formerly suggested since this position in the two fish 17beta-HSDs type 1 is occupied by leucine, another hydrophobic but slightly smaller residue. Concerning interaction with the C18 methyl-group, V143, G145 and L150 are considered to form a small hydrophobic pocket enabling the space and interaction to stabilize the substrate. V143 is displaced by isoleucine in the eel and methionine in the zebrafish enzyme. In addition, this position was not identified to be in a radius for van der Waals interaction with estradiol in the zfHSD 1 homology model. This suggests that a relatively short hydrophobic amino acid might be the only important restriction for this position whereas the character of the respective residue is not conserved. G145 and L150 are instead strictly conserved in all 17beta-HSD type 1 homologs, which is further supported by identification of their close vicinity to the substrate in the homology models.

5.1.2 Why a fish is no human is no mouse (gene): speciation effects

Analysis and characterization of zfHSD 1 revealed high similarity to the mouse and human homologs on DNA, RNA and protein level. But in addition, differences emerged reflecting individual adaptation of 17beta-HSD type 1 to the respective organism. Slight changes in the amino acid sequence that emerged in the course of evolution are best reflected in the phylogenetic analysis but are also immediately apparent when calculating an identity matrix. Aside from this linear and relatively easy to identify mode of change, other effects can be discerned. On DNA level the gene structure may alter and probably lead to different transcripts. Changes in the promoter or enhancer have influence on the expression pattern and finally substitutions of specific residues in the amino acid sequence may influence the enzymatic characteristics, the localization of protein in the cell or its binding to putative interaction partners. All these changes eventually lead to adaptation of this specific gene to the organism and, the other way round, are characteristics of the organism reflected in the genes.

5.1.2.1 Additional regulation acquired in viviparous vs. oviparous organisms? -The very C-terminus of 17beta-HSD type 1

Complementation of the coding sequence of zebrafish 17beta-HSD type 1 immediately revealed one striking difference to the mammalian homologs: it was by far shorter. Analysis of the gene structure demonstrated that this difference was due to the last coding exon which in the zebrafish comprised 168 nt, in mouse 270 nt and in human even 321 nt. Expanding this comparison to the eel, chicken and rat homolog, it appears that non-mammalian 17beta-HSDs type 1 share sizes similar to that monitored in the zebrafish whereas the mammalian proteins display an elongated C-terminus (Fig. 33).

```
                    *
human       GSNYVTAMHREVFG--DVPAKAEAGAEAGGGAGPGAEDEAGRSAVGDPELGDPPAAPQ--------------- 328
mouse       GSSYVAAMHQEAFSNLQTQENAKAGAQVPGVSDTASSALICLPECAIPRVASELGWSASDKPGQDNSCYQQKI 344
rat         GSSYVEAMHREAFSDLQVQEGAKAGAQVSGDPDTPPRALICLPECAIPRVTAELGWSASDKPGQNKSCYQQKI 344
chicken     GADYVRAMHDFVFGPGEAAEEQP------------------------------------------------- 302
eel         GSQYIRAMSKLIFSSSGPEDP--------------------------------------------------- 293
zebrafish   GSQYIRAMSKLIFSSPGTDAQK------------------------------------------------- 295
```

Fig. 33: Alignment of the very C-terminus of characterized 17beta-HSDs type1.
Several strictly conserved residues that may function in substrate recognition (see previous chapter) are present and are shaded in blue. The glutamic acid, which was suggested to form a hydrogen bond in the human 17beta-HSD type 1 to the C3 hydroxy-group of the substrate, is marked by an asterisk. The sequence parts which were investigated for putative protein interaction motifs by comparison to the ELM server database are marked in red.

Interestingly, this part seems not to be crucial for substrate binding and recognition. For example, a chimeric enzyme consisting of rat 17beta-HSD type 1 from residue 1-266 and the human enzyme in the subsequent C-terminus showed nearly identical activity as compared to

the intact rat protein (Puranen et al., 1997b). Furthermore, available structures of crystallized 17beta-HSD type 1 also miss this very C-terminal part. E283, the last amino acid in vicinity to the substrate, as identified by analysis of the crystal structure, was suggested to function in estrogen recognition by forming hydrogen bonds to the C3 hydroxy-group (Puranen et al., 1997b). But this notion was neither confirmed by independent, theoretical investigations nor by experimental data.

Apart from contribution to enzymatic function, an alternative possibility that might justify preservation of this sequence part in mammals is a putative signal or interaction function. Testing this hypothesis, the last C-terminal amino acids, which are possibly unique in mammals, (see Fig. 33) were compared against the database of the ELM server. A number of putative motifs was identified in the human, mouse and rat sequence and is listed in Tab. 10.
The results comprise two major kinds of motifs: ligand motifs (LIG) and modification motifs (MOD). The first group indicates sequences directly recognized and used by other molecules for protein-protein interaction. The second group specifies motifs which will be recognized and subsequently modified by enzymes leading to phosphorylation, glycosylation or other types of covalent modifications.
Concerning the ligand motifs, results fall into four different groups of interactors: Cyclin, FHA, SH3 and WW domain. All four motifs are present in rat where they not only fit the pattern itself but also meet secondary requirements needed to be functional: in addition to the LIG_CYCLIN_1 motif the target phosphorylation site MOD_CDK is present and phosphorylation of the threonine in LIG_FHA_1 can be performed through the MOD_PKA_2 site. Therefore, especially the rat 17beta-HSD type 1 might be involved in processes of signal transduction (SH3 and FHA) and cell cycle control (Cyclin, FHA and WW). It is interesting to note that despite an amino acid identity of nearly 80% between rat and mouse in this region, none of the motifs present in rat could be identified in mouse. It is not clear whether this indicates that the ligand patterns present in the rat sequence are random, non-functional hits or they indeed reflect diverse regulatory mechanisms between the two rodents. It is also possible that slight modifications of a motif are still recognized by the organism and therefore remain functional but have not yet been reported. In the human sequence, only the SH3 domain could be identified which is present twice in a tandem formation at the very C-terminus, a location not consistent with that in the rat protein.
Aside from the modification site for cyclin-dependant kinase (MOD_CDK), the other modification motifs fall into three different groups: phosphorylation, glycosylation and sumoylation. Human 17beta-HSD type 1 has been described to reside in the cytoplasm (Ghersevich et al., 1994) and not to be secreted or localized to the ER. In consistence with that, no signalling patterns are present in the N-terminus of the enzyme. Hence, glycosylation signals might rather be random hits and not relevant in case of 17beta-HSD type 1. Concerning phosphorylation, a putative context was described above for the rat isozyme. Indeed, the modification pattern for phosphorylation is quite simple so it can be expected to appear frequently in protein sequences without a physiological significance.

Tab. 10: ELM motifs in the very C-terminus of human, mouse and rat 17beta-HSD type 1.

Elm Name	*Match (Organism)*	*Elm Description*	*Cell Compartment*	*Pattern*
LIG_CYCLIN_1	RALI (r)	Interaction with cyclin and thereby increase of phosphorylation by cyclin/cdk complexes. Predicted proteins should have the MOD_CDK site	nucleus, cytoplasm	[RK].L.{0,1}[FYLIVMP]
LIG_FHA_1	TAEL (r)	FHA domain interaction motif 1, threonine posphorylation is required	nucleus, cytoplasm, plasma membrane	T..[ILA]
LIG_SH3_2	PDTPPR (r)	Motif recognized by class II SH3 domains	cytoplasm, plasma membrane, focal adhesion	P..P.[KR]
LIG_SH3_3	RSAVGDP (h) GDPPAAP (h) SGDPDTP (r)	Motif recognized by SH3 domains with a non-canonical class I recognition specificity	cytoplasm, plasma membrane, focal adhesion	...[PV]..P
LIG_WW_3	TPPRA (r)	WW domain of group III binding motif	cytoplasm	.PPR.
LIG_WW_4	DPDTPP (r)	Class IV WW domains interaction motif; phosphorylation-dependent interaction	nucleus, cytoplasm	...[ST]P.
MOD_CDK	DPDTPPR (r)	Substrate motif for phosphorylation by CDK	nucleus, cytoplasm, cytoplasmic cyclin-dependent protein kinase holo-enzyme complex	...([ST])P.[KR]
MOD_GlcNHglycan	RSAV (h) SSAL (m) WSAS (m) VSGD (r) WSAS (r)	Glycosaminoglycan attachment site	extracellular, Golgi apparatus	[ED]{0,3}.(S)[GA].
MOD_GSK3_1	SDTAS (m)	GSK3 phosphorylation recognition site	nucleus, cytoplasm	([ST])...S
MOD_N-GLC_1	NKS (r)	Generic motif for N-glycosylation	extracellular, Golgi apparatus, endoplasmic reticulum	(N)[^P][ST]|(N)[^P][ST][^P]
MOD_N-GLC_2	NSC (m)	Atypical motif for N-glycosylation site	extracellular, Golgi apparatus, endoplasmic reticulum	(N)[^P]C
MOD_PKA_2	RVT (r)	PKA phosphorylation site	cytoplasm	R.([ST])[VILAFP]
MOD_SUMO	AKAE (h) AKAG (m) AKAG (r)	Motif recognised for modification by SUMO-1	nucleus, PML body	(K)[EDNGP]

Abbreviations for organisms: h: human; m: mouse; r: rat.

The sumoylation signal is present in all three mammalian sequences at similar positions. Involvement of sumoylation has been demonstrated to have an inhibitory effect on the transcriptional activity of the androgen receptor (Poukka et al., 2000). In this context sumoylation was positively correlated with androgen signalling. In another case, sumoylation prevented protein degradation by competing with an ubiquitinylation site (Desterro et al., 1998). Although enzymes necessary for the transfer of the SUMO protein to targets are predominantly located in the nucleus (Rodriguez et al., 2001), RanGAP1 localization changed from cytosolic to nuclear upon SUMO-1 conjugation (Mahajan et al., 1998; Zhang et al., 2002) indicating the sumoylation process not be restricted to the nucleus. It remains unclear whether the SUMO modification pattern detected in these three mammalian sequences is indeed necessary for physiological 17beta-HSD type 1 function but the already observed connection of sumoylation with the steroid signalling pathway argues against a random hit.

Elongation of the C-terminus seems to be an interesting but yet unexplored feature of all so far available mammalian sequences in contrast to that of non-mammalian vertebrates, raising the

possibility of functions specific to viviparous organisms. These might be encoded differently in the individual species or are more flexible than suggested by the strict motif patterns since sequence identity in this very C-terminus is apparently lower compared to the enzymatically active part (e.g. in mouse vs. rat: 78% to 96%, respectively).

5.1.2.2 The difference that makes the difference: 17beta-HSD type 1 in different organisms, sexes and tissues

Upon characterization of zebrafish 17beta-HSD type 1 its extensive similarities with the homologous enzymes from other organisms were quite notable. In addition, there exist distinct differences concerning enzyme activity and expression. These two aspects reflect on the non-identical function of 17beta-HSD type 1 in the respective species.

Concerning substrate specificity, all so far characterized 17beta-HSDs type 1 catalyze the conversion of estrone to estradiol with high efficiency. For the human enzyme a variety of other reactions were observed to be catalyzed such as the reduction of dehydroepiandrosterone to 5-androstene-3,17-diol (Mendoza-Hernandez et al., 1984), oxidation of 5α-dihydrotestosterone to androstanedione (Mendoza-Hernandez et al., 1984), and reduction of 5α-dihydrotestosterone to 3β,17β-androstanediol (Gangloff et al., 2003). Most of these reactions were measured *in vitro* and their biological significance is not known. Furthermore, the efficiency of the respective catalysis was extremely low as compared to the reduction of estrone. Conversion of androstenedione to testosterone instead, is readily performed by the murine as well as the rat 17beta-HSD type 1 *in vivo* (Akinola et al., 1996; Nokelainen et al., 1996). Since a similar acceptance of androgens as a substrate could not be identified in the chicken or eel 17beta-HSD type 1, it is suggested that acquirement of this function may be rodent specific. Investigation of substrate specificity from my work support this notion as neither testosterone nor androstenedione could function as a substrate for the recombinant zebrafish 17beta-HSD type 1 whereas the conversion of estrone and estradiol *in vitro* was catalyzed with high efficiency.

The apparent enzymatic differences mirror non-identical functions of 17beta-HSD type 1 in different vertebrate species. Further discrepancies between the human and rodent enzymes consequently affect their expression pattern (Akinola et al., 1996; Labrie et al., 1997) and transcriptional regulation (Piao et al., 1995; Piao et al., 1997; Akinola et al., 1998). But even if the enzymatic properties are comparable, employment of 17beta-HSD type 1 in estradiol generation in different organisms may not. Another enzyme, 17beta-HSD type 7, has been identified to convert estrone into estradiol, though with apparently lower efficiency, and was found to be highly expressed in mouse ovaries (Nokelainen et al., 1998) and as well in the mouse uterus during pregnancy (Nokelainen et al., 2000). Further, yet uncharacterized 17beta-HSDs with a preference for estrone reduction might exist as this enzymatic activity was detected in a variety of tissues in which presence of 17beta-HSD type 1 remains to be proven (Labrie et al., 1997). Furthermore, estradiol can be generated directly from testosterone upon aromatization by P450 aromatase. This pathway may also be the main route to provide estradiol during embryogenesis since the present study on zebrafish 17beta-HSD type 1 could not detect any expression during embryogenesis. Involvement of estradiol in non-gonadal differentiation processes, which take place during the monitored stages of development, indeed seem to be rather regulated by P450 aromatase. In zebrafish, strong up-regulation of the brain specific

form, P450 aromatase type B, was detected starting at 24 hpf (Kishida and Callard, 2001). Important roles for P450 aromatase in the sexual dimorphic development of brain in fetal mice have been reported (Beyer et al., 1994; Hutchison et al., 1997; Bakker et al., 2003). Alternatively as well as in case of male zebrafish, 17beta-HSD type 1 expression levels might have been too low to be detected by the applied method. This argument is supported by identification of 17beta-HSD type 1 in gonads of male eel, chicken, mice and human upon use of highly sensitive RT-PCR and RNAse protection assay whereas Northern blot and *in situ* hybridization were not able to demonstrate the presence of specific transcripts.

5.1.2.3 Expansion of the understanding of 1D- and 3D-protein sequence features by integration of zebrafish data

The previous two chapters were focused on discrepancies concerning transcriptional regulation and possible post-translational modulation. Despite these interesting features, all 17beta-HSDs type 1 were shown not only to be homologous but also orthologous, conveying similar functions in the respective organisms. This groundwork of substrate specificity and physiological function must have been present about 450 million years ago when fish separated from the vertebrate lineage. During evolution into the present organisms the protein has undergone slight modifications that might be species-specific but do not simultaneously affect enzymatic function. Knowledge of these processes deepens the understanding of the evolvement of substrate-enzyme interaction and helps to single out the necessary requirements on structure and residue identity from an organism-specific background.

Conservation of structure-function relations

This context of structure-function relation was very well illustrated by Identity plots, employed here for the first time (Fig. 34). Despite an overall amino acid identity of only about 50% between the zebrafish and the two mammalian sequences the SDR-motifs, viable to cofactor binding and orientation, were immediately recognizable as areas of elevated conservation. In addition, it became clear that conservation of the connecting sequence parts is not uniformly stringent. Residue identity in the area between motifs (1) and (3) is primarily dictated by the assembly of the tertiary structure as amino acid identity is very low. In concurrence, homology modeling revealed a shortened helix and connecting loop in the zebrafish enzyme in this area.

In the Identity plot of 17beta-HSD type 1, percent identity rises again at position aa 105 forming a continuous area of elevated preservation that drops steeply at aa 180 subsequent to the last SDR-motif involved in cofactor binding. Comparing these findings to functional parts of the protein and homology modelling, it becomes clear that this region was formed by more than one way. Firstly, a strict succession of tertiary structures and connecting turns is necessary to shape this area which aligns cofactor and substrate via steric and hydrophobic interactions as well as hydrogen bonds and harbors the active center. Secondly, this region comprises the two long helices, αE and αF, which form the dimeric interface for a four-helix bundle in a homodimer (Benach et al., 1998). Although the formation of dimeric or tetrameric quaternary structures is characteristic for SDRs, it is not clear whether this process is a prerequisite for enzymatic activity.

Conservation in the area subsequent to motif (7) drops notably. Here, 3D-structures do not show any interaction with the substrate and in agreement with this homology modeling of the zebrafish enzyme reveals structural discrepancies in this part. The tertiary structure of this area

mainly builds a flexible loop, the substrate entry loop, which undergoes strong conformational changes upon substrate binding (Ghosh and Vihko, 2001). In conclusion, flexibility and reduced amino acid conservation are in agreement with these findings.

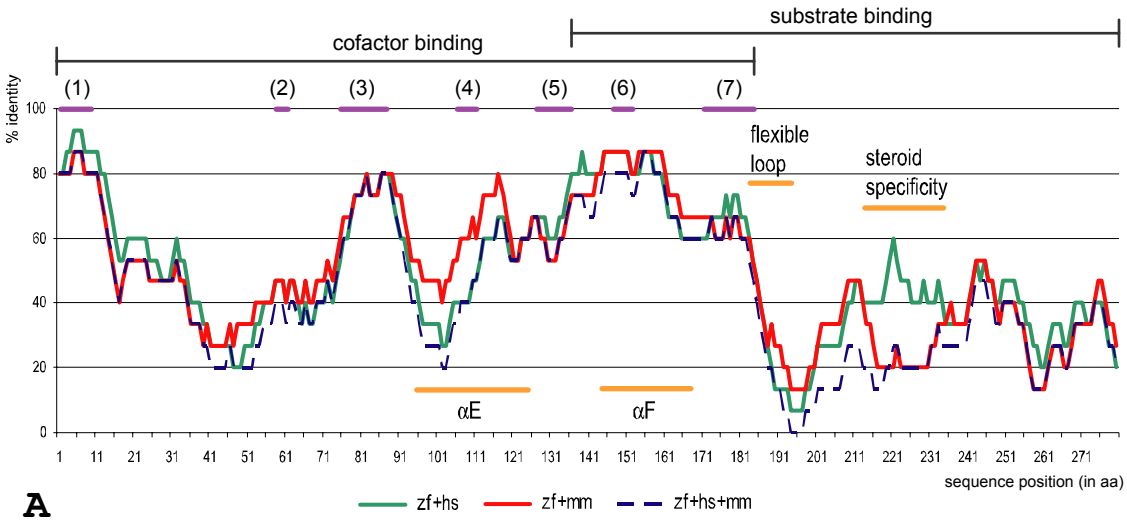
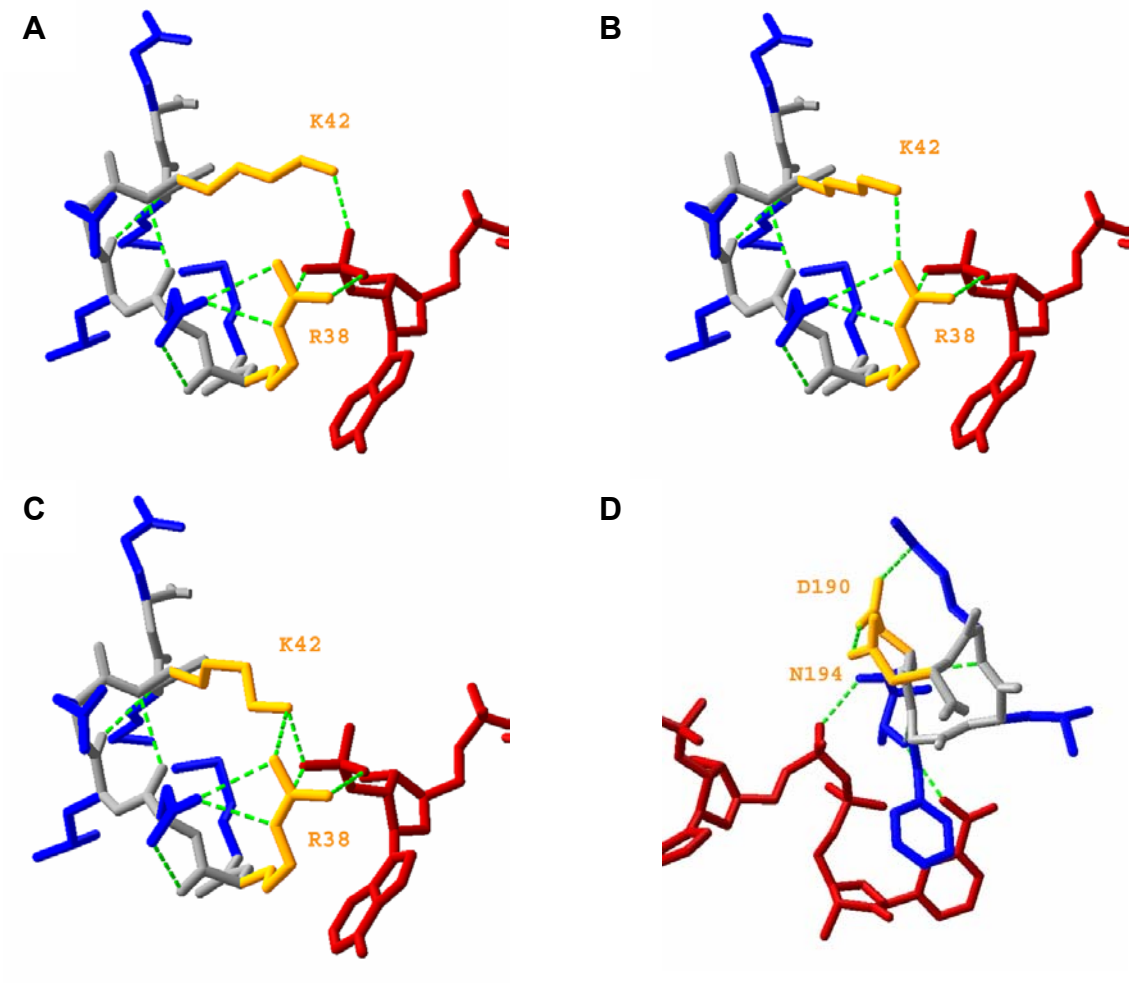


Fig. 34: Identification of conserved functional and structural aspects in Identity plots.
The ID plot resembles the one from chapter 4.1.4.1.4 comparing 17beta-HSDs type 1. Areas reflecting a context of structural conservation are marked by an orange line and descriptions as mentioned in the main text.

Enhanced cofactor-binding in the zebrafish enzymes

Homology modeling of the zebrafish 17beta-HSD type 1 revealed a high similarity to the human homolog. Especially in the cofactor binding region all residues forming hydrogen bonds to NADP in the human enzyme could as well be identified in zfHSD 1. Interestingly, the two zebrafish prRDHs were found to display additional residues capable of stabilizing NADP via hydrogen bonds at two different positions: one was R45 in zfprRDH 1 and the other was E193 in zfprRDH 1 and E194 in zfprRDH 2, respectively. At both sites, two amino acids binding to NADP were identified spaced exactly four residues apart (see also chapter 4.1.4.2.1, Fig. 16). Investigation of the corresponding positions in zfHSD 1 indicates that the respective amino acid is here either completely conserved or replaced by a conservative exchange. Hence, additional bond formation through these specific residues might also be present in zfHSD 1. This possibility was examined by testing the different rotamers of the respective amino acid side chains in zfHSD 1 and the results are presented in Fig. 35. In the first case, several rotamers were capable of hydrogen bond formation either directly with the cofactor or with the amino acid four positions before or both (Fig. 35A-C). This results in an enhanced stabilization of NADP binding and also in an increased specificity for this cofactor compared to NAD due to additional recognition of the phosphor group. In the second case (Fig. 35D), residues next to the cofactor binding and the highly conserved threonine, forming a hydrogen bond to NADP, could not produce the described synergistic mechanism by different rotamers. In all possible conformations both amino acids were too far away from NADP and instead the loop in vicinity to the cofactor was stabilized by a hydrogen bond between their respective side chains.



E

17beta-HSDs type 1:

human	L	**R**	D	L	K	T	Q	G	-	44
mouse	L	**R**	D	L	K	A	Q	G	-	44
rat	L	**R**	D	L	K	S	Q	G	-	44
chicken	M	**R**	D	L	A	K	G	E	-	51
eel	M	**R**	N	L	A	K	K	E	-	44
zebrafish	M	**R**	N	L	D	K	K	Q	-	44

prRDHs:

human	M	**R**	D	L	G	K	K	E	-	46
bovine	M	**R**	D	L	G	K	K	G	-	46
mouse	M	**R**	D	L	G	K	K	E	-	46
rat	M	**R**	D	L	G	K	K	E	-	46
zebrafish 1	M	**R**	D	L	K	R	Q	E	-	47
zebrafish 2	M	**R**	D	L	K	K	K	D	-	48

17beta-HSDs type 1:

human	H	**T**	A	F	M	E	K	V	-	197
mouse	H	**T**	A	F	Y	E	K	L	-	197
rat	H	**T**	A	F	H	E	K	L	-	197
chicken	H	**T**	A	F	M	D	N	A	-	202
eel	N	**T**	H	F	L	S	N	L	-	195
zebrafish	N	**T**	D	F	L	M	N	L	-	195

prRDHs:

human	V	**T**	E	F	E	G	K	L	-	196
bovine	V	**T**	E	F	E	G	K	L	-	197
mouse	T	**T**	D	F	E	G	K	L	-	198
rat	I	**T**	D	F	E	G	K	L	-	197
zebrafish 1	H	**T**	E	F	E	M	K	M	-	198
zebrafish 2	H	**T**	E	F	E	T	K	M	-	199

Fig. 35: Enhanced cofactor binding by synergistic hydrogen bonds in zfHSD 1 in comparison to 17beta-HSDs type 1 and prRDHs.
A-D: Hydrogen bonds of the two amino acids separated by four positions in zfHSD 1. Hydrogen bonds are shown as a green dotted line; the cofactor is depicted in red. **A-C**: Residues in the vicinity of the phosphor group of NADP. R38 is highly conserved and supposed to facilitate cofactor discrimination while K42 could have a synergistic effect in this by forming hydrogen bonds to NADP (**A**), R38 (**B**) or both (**C**). In the second position (**D**), a strictly conserved threonine functions in binding the cofactor. Here, both D190 and N194 are too far away from NADP to assist in binding. Instead they can form hydrogen bonds between their side chains thereby stabilizing this loop. **E**: Alignment of 17beta-HSDs type 1 and prRDHs in the two investigated regions. Amino acids capable of forming hydrogen bonds to NADP as suggested by homology models and alignment are shaded in orange; D190 and N194 in zfHSD 1 as depicted in (**D**) are written in orange. The highly conserved arginine in the first and threonine in the second investigated region are highlighted in bold.

Expanding the comparison to other 17beta-HSDs type 1 and prRDHs (Fig. 35E), an interesting scenario emerges: In the first described position (R45 in zfprRDH 1), the synergistic constellation is highly preserved in the prRDHs and still present in the fish enzymes as well as the chicken homolog of 17beta-HSD type 1. Concerning the second position (e.g. E193 in zfprRDH 1), prRDHs again display a similar constellation of amino acids with the exception of the two rodent proteins where the glutamic acid is replaced by the shorter aspartic acid but the lysine four positions further downstream is still present. Interaction of this lysine with the cofactor or loop stabilization via hydrogen bond formation to the amino acid four positions before is therefore theoretically possible for all these prRDHs. In the 17beta-HSDs type 1, additional hydrogen bonds to NADP at this site might be lost in the fish and chicken enzyme but the conserved lysine re-emerges in the mammalian homologs.
Altogether it seems likely that the identified additional hydrogen bonds facilitating cofactor binding have been present in the ancient progenitor of prRDHs and 17beta-HSDs type 1. While in the first group this mechanism appears to be almost completely conserved, the latter group seems to have lost it during evolution at least at one of the two identified sites.

5.1.3 The sister group: comparison to retinol dehydrogenases

At the onset of the study, three non-identical zebrafish sequences were identified as putative homologs of mammalian 17beta-HSD type 1. Extensive studies on DNA, RNA and protein level identified two of these sequences to be the zebrafish homologs of mammalian photoreceptor associated retinol dehydrogenases (prRDHs) and they were named accordingly, zfprRDH 1 and zfprRDH 2. The first identified and characterized prRDH was from *bos taurus* where it appeared to be closely associated with the visual process (Rattner et al., 2000). The authors found a putative relation of this protein to mammalian 17beta-HSD type 1 but the phylogenetic analysis was not very detailed. While these first studies on prRDH concentrated on elucidation of its vision specific function, the focus in the following chapters is rather on its evolution in comparison to 17beta-HSD type 1.

5.1.3.1 Sticking together: traces of common evolution

Among individual members of the SDR family sequence similarity on protein level is usually quite low due to the different functions exerted by these enzymes. In general, sequence identity does not rise above 30% thereby highlighting only residues and motifs basic to common SDR architecture and enzymatic integrity. The prRDHs display about 40% amino acid identity to the 17beta-HSDs type 1, indicating a much closer relation in spite of the clear difference in substrate specificity. This was also reflected in the phylogenetic analysis where a bootstrap value of 100% separated the groups of prRDHs and 17beta-HSDs type 1 from all other sequences in the analyzed data set. Interestingly, the group of retinol dehydrogenase similar proteins in this phylogenetic tree does not seem to be very closely related to the prRDHs. The retinol dehydrogenases themselves form a huge family of diverse proteins which result from a complex evolution (Duester, 2000; Dalfo et al., 2001; Zhuang et al., 2002). The majority of these RDHs share the same evolutionary origin and might have arisen from an ancestral enzyme,

which also gave rise to 3beta-hydroxybutyrate dehydrogenases (Baker, 1996), that fall into the group of oxidoreductases. Although my study indicated that microbial oxidoreductases are related to prRDHs and 17beta-HSDs type 1, no member of the RDH family was identified in the process of sequence data acquirement here. This suggests an evolution of prRDHs independent from the main family of RDHs and gain of substrate specificity maybe by convergence. Such a hypothesis is further sustained by comparison and calculation of phylogenetic trees by use of structure data of 17beta-HSDs. While 17beta-HSD type 2, which is closely related to the RDHs, is in one group together with 17beta-HSDs type 6 and 9, which clearly convert retinoids, 17beta-HSD type 1 forms a separate group together with 17beta-HSD type 7 (see Fig. 36).

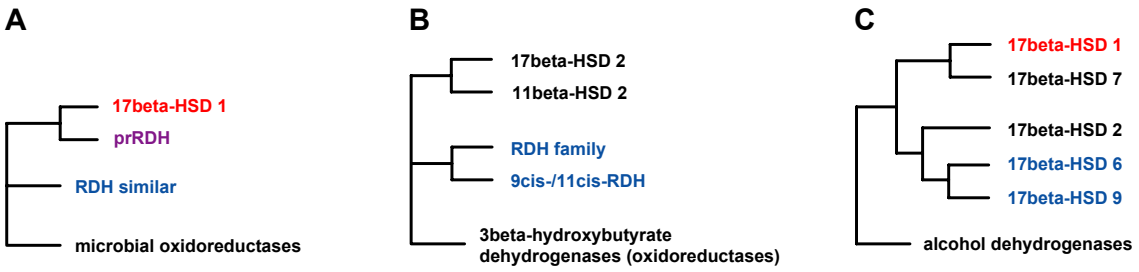


Fig. 36: Evolutionary context of photoreceptor-associated RDHs and 17beta-HSDs type 1.
(A) Simplified phylogenetic tree of prRDH and 17beta-HSD type 1 evolution. (B) Origin of the family of retinol dehydrogenases. (C) Structure based phylogenetic calculation to asses the relations of 17beta-HSDs (modified from Breitling et al. (2001a)). 17beta-HSD type 1 is shown in red; all subgroups belonging to the family of RDHs are shown in blue; prRDH are depicted in violet as they are much closer to 17beta-HSD type 1 than other RDHs and might have evolved independently from the RDH family.

The closeness of prRDH and 17beta-HSD type 1 is already apparent on DNA level where both genes display a nearly identical exon structure (see Fig. 37). While exon-intron boundaries are conserved, differences in exon size are present but mainly in exons coding substrate specific or putative signalling parts. One of the most intriguing discrepancies between zebrafish 17beta-HSD type 1 and prRDHs is present in the last exon. Alignment of the coding sequences of the three zebrafish sequences suggests that a frame shift is primarily responsible for the observed size difference (see also chapter 4.1.2.1, Fig. 7). Considering that the very C-terminal part in 17beta-HSDs as well as prRDHs probably functions in localization signalling, two putative ways for emergence of these two enzyme groups are possible. Either the shorter, 17beta-HSD type 1-like sequence is the ancestor and after duplication prRDHs were elongated, or the longer, prRDH-like sequence resembles the ancestor and fish 17beta-HSD type 1 was shortened due to the aforementioned frame-shift by insertion. In both cases, the elongated C-terminus in mammalian compared to non-mammalian 17beta-HSDs type 1 is a secondary change that occurred far later in evolution after divergence from prRDH and acquirement of enzymatic 17beta-HSD type 1 function.

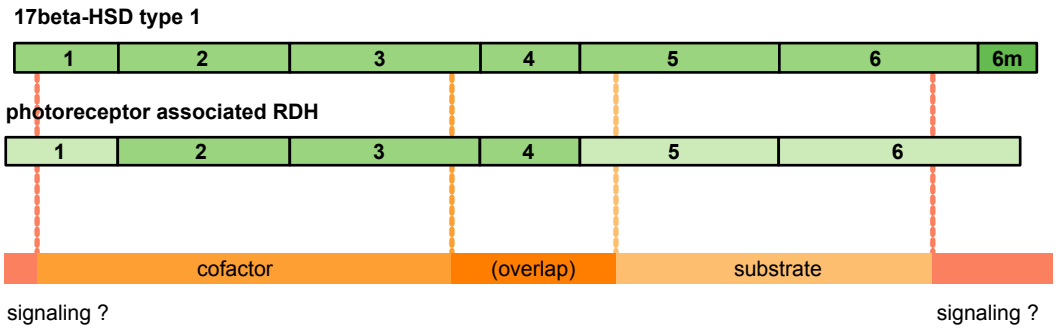


Fig. 37: Exon structure and functionally different parts of zebrafish 17beta-HSD type 1 and prRDHs. The coding sequence of both genes is shown in green while exons shaded in light green in prRDHs indicate size differences in comparison to 17beta-HSD type 1. The significantly elongated C-terminus of the mammalian 17beta-HSDs type 1 is highlighted in dark green and designated as 6m. Parts of the resulting protein which convert different functions are schematically shown underneath both sequences. The mainly cofactor and substrate binding parts were defined as in chapter 4.1.4.1.4.

5.1.3.2 Functionally different and clearly non-similar to 17beta-HSD type 1

Despite initially considering the two zebrafish prRDH sequences as putative candidates for the 17beta-HSD type 1 homolog, data soon accumulated indicating them to be functionally different genes. Characterization of the specific expression profiles and phylogenetic analyses all supported the notion that these two proteins are homologs of mammalian prRDH and not of 17beta-HSD type 1.

Substrate specificity in wild type and point-mutants

In this context, the measurements of substrate specificity were especially interesting. The evolution of the main family of retinol dehydrogenases (see also Fig. 36) has given rise to some enzymes which in addition to retinoids also accept steroids as substrates. Very impressively, this demonstrates the close evolutionary connection between RDHs and some steroid dehydrogenases e.g. 17beta-HSD 2 and 11beta-HSD 2, which have an ancient progenitor in common with the RDHs. The reason for this close relation may in part lay in the structural similarity of retinoids and steroids. Androgens, which lack the aromatic A-ring, are even slightly more similar to retinoids and in consistence with this, “cross-reactive” enzymes display a higher acceptance of androgens than estrogens.

The two prRDHs from zebrafish identified in this work did not show any substrate specificity towards the tested androgens or estrogens. From homology modeling and sequence alignment it was suggested that a prerequisite for retinoid-steroid discrimination includes a position close to the active center. In the 17beta-HSDs type 1 this position is occupied by a conserved glycine, G143 in zfHSD 1, whereas all investigated prRDH display a conserved methionine, including the two zebrafish enzymes with M146 and M147 in zfprRDH 1 and zfprRDH 2, respectively. In the homology models of zfprRDH 1 and 2, the presence of methionine was too bulky to allow enough room for the C18-methyl group of estradiol. To investigate this hypothesis, point mutants of the recombinant proteins were generated exchanging methionine to glycine in the two prRDHs and glycine to methionine in zfHSD 1. Indeed, reactivity of the 17beta-HSD type 1 was reduced though only the oxidation reaction was affected. This suggests the introduced

mutation to specifically influence the binding of the hydroxylated substrate and to be here the rate-limiting step. Although the C18-methyl group, which was in steric conflict with the G143M exchange in zfHSD1, is present in both estradiol and estrone, it seems that positioning of the keto group at C17, when binding estrone, may be less exact and less affected and therefore still allows for transfer of the hydride ion. In comparison, fitting of the stereospecific hydroxy group at C17 of estradiol may need more complex steric requirements and therefore was more strongly affected by introduction of methionine instead of glycine. Since the introduced amino acid exchange around position 145 did not affect the region mediating estrogen-androgen discrimination, rejection of androstenedione and testosterone as substrates was not altered.

Homology modeling indicates clear differences in the substrate binding area

Although the described amino acid exchange close to the catalytic site had an impact on zfHSD 1, reactivity of the zfprRDHs towards steroids was not significantly altered. The reason most likely lies in the more distal part of the enzyme predominantly responsible for substrate recognition. Homology modeling revealed three apparent deviations of the prRDHs when trying to fit them to the human 17beta-HSD type 1 template (see also chapter 4.1.4.2.1). Two of these observed differences result in the creation of loop areas which are potent to significantly change the architecture of the subsequent structure elements in comparison to 17beta-HSD type 1. Especially the third of the three deviations leads to disruption of the helix important for steroid recognition in the HSDs. Altogether, the observed changes might alter overall structure of the C-terminal part in a way to enable specific binding of the elongated retinoic substrate.

A second line of evidence comes from comparison of the Identity plots of both enzyme groups. While in 17beta-HSD type 1 the region subsequent to motif (7) shows low conservation due to a quite flexible composition of the substrate entry loop, the corresponding area in prRDHs is much higher conserved. The complete subsequent part displays a nearly uniform conservation of about 40% identical residues which most likely are involved in structural stabilization and interaction with the substrate. As a consequence, prRDHs may not undergo as pronounced conformational changes upon substrate binding as it is known for 17beta-HSD type 1.

In this context it is also interesting to note that efficient binding of the by far longer though less bulky retinoid substrate in prRDHs is performed within the same sequence stretch as in the 17beta-HSDs. The first impression that the cofactor binding part in the prRDHs is longer than in the HSDs may not hold true as, among other aspects, is discussed in the next chapter.

5.1.3.3 The second duplication: two distinct prRDHs in (zebra) fish

So far, the bovine photoreceptor associated retinol dehydrogenase remains the only isozyme of this group studied in more detail. It has been shown to be specifically expressed in the retina where it is located in the rod outer segment and to convert all-trans-retinal to all-trans-retinol with high specificity (Rattner et al., 2000). Identification of single, highly identical sequences in mouse, rat and human indicate that at least in mammals only one single homolog coding for prRDH exists. This posed the question of origin of the duplication apparent in zebrafish and its physiological consequences. Screening available databases for homologs revealed related sequences from a variety of other vertebrates. Phylogenetic analyses of these data (Fig. 38) showed that most likely duplicates exist as well in other organisms since additional and highly similar sequences were found in salmon and *Xenopus*.

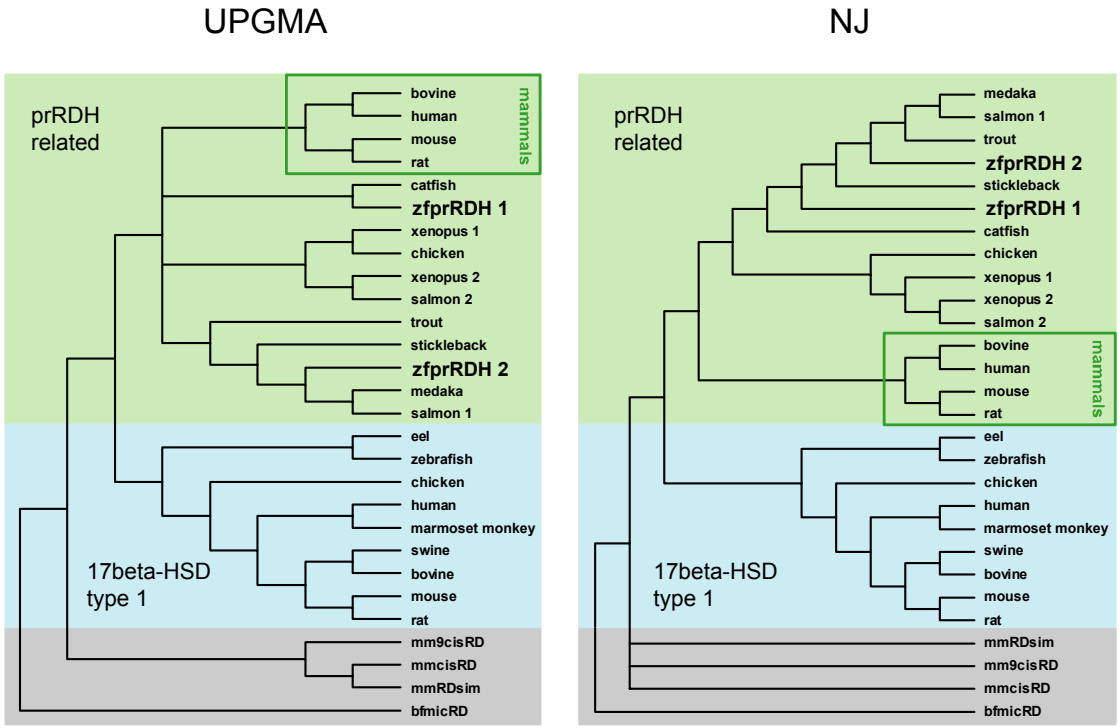


Fig. 38: Dendrograms of prRDH evolution.
Sequences similar to the two zebrafish prRDH proteins were identified by BLAST search and their relation investigated by phylogenetic analysis. As most of these sequences were not complete, different algorithms were employed to assess the evolutionary context. Distinct groups are formed by the 17beta-HSDs type 1 and the prRDHs. In the latter family, mammalian sequences cluster in a separate group but it remains unclear where the zebrafish duplication originated. Trees were calculated by unweighted pair group method with arithmetic mean (UPGMA) or Neighborjoining (NJ) with a bootstrap of 1000 pseudoreplicates. Depicted dendrograms are unrooted bootstrap consensus trees.

But the dendrograms also indicate that evolution of this gene might be rather complex. Both employed mathematical approaches of the phylogenetic analysis (Neighborjoining and unweighted pair group method with arithmetic mean) suggest the mammalian prRDHs to form a separate group with neither zfprRDH 1 nor zfprRDH 2 closer related to them. Furthermore, the duplication event giving rise to the two zebrafish paralogs might not be identical with that resulting in the two salmon paralogs since the respective paralogous proteins are not separated by the same bifurcation. The two *Xenopus* paralogs might have been generated after the divergence of fish as they are located on a separate tree and in close vicinity to each other.
Many data collected in this work suggest zfprRDH 1 to resemble the zebrafish homolog of bovine prRDH. During embryogenesis expression starts during late somitogenesis where eye and brain undergo enhanced development. In agreement with this, transcripts are most abundant in the eye and were additionally detected in brain, ovary and testis of adult fish. Overall sequence identity on protein level to human and mouse prRDH was higher for zfprRDH 1 (~51%) than for zfprRDH 2 (~47%). This was also reflected in the Identity plot where especially in the region conveying correct orientation of cofactor and substrate including the active site percent of identical residues was higher for zfprRDH 1 than for zfprRDH 2. The same was true for the majority of the specifically substrate binding part. Quite apparent was instead the drop of percent identity in the very C-terminus starting at about position aa 280. In this area zfprRDH 2 displays a higher similarity to the two mammalian sequences.

Since identity between all four compared sequences drops significantly in this part of the protein, it is most likely not involved in substrate recognition and binding. It may rather be assumed that it harbors motifs and signals necessary for localization of the protein as the N-terminus does not seem to contain such signals but the mammalian prRDH localizes to the rod outer segment membrane. Rattner et al. (2000) suggested that membrane association does not take place via hydrophobic regions in the protein itself but rather by lipid attachment to the C-terminus. Another protein specific for the vision process, rod cGMP phosphodiesterase, had been demonstrated to be post-translationally modified by methyl esterification and lipidation at a C-terminal cysteine (Ong et al., 1989). A similar mechanism including farnesylation and attachment of a palmitoyl group is essential for membrane localization of ras proteins (Willumsen et al., 1984; Hancock et al., 1989), implying indeed a signal character for the identified C-X-X-X motif at the very C-terminus of the respective proteins. Notably, the cysteines in the mammalian prRDHs as well as zebrafish prRDH 2 appear to be conserved but are not positioned at the very end of the enzyme (Fig. 39). A conserved cysteine in this region is also present in two other enzymes capable of converting all-trans retinal to all-trans retinol: retSDR1, which is expressed in retina but in addition in other tissues such as kidney, pancreas and liver (Haeseleer et al., 1998), and RalR1, which was most abundant in the prostate (Lin et al., 2001) and localized to the ER (Kedishvili et al., 2002). Therefore, it seems that the cysteine motif may indeed convey localization to the membrane but if so, it may not be vision specific. Although zfprRDH 1 lacks any conserved cysteine here, a corresponding and slightly larger region is able to form a transmembrane helix (Fig. 39) as was predicted by the SMART server. Such a structure could not be identified for any of the other prRDHs examined and might indeed explain the observed divergence in the very C-terminus as highlighted by Identity plots. Probably, zfprRDH 1 exerts similar physiological functions in the visual process in addition to others but is targeted to the membrane by a different mechanism as compared to the mammalian enzymes.
zfprRDH 2 may also function in vision but is clearly not restricted to it as expression was detected throughout embryogenesis and in all adult tissues investigated. The reduction of all-*trans*-retinal to all-*trans*-retinol is a step required in reconstituting 11-cis retinal after photoisomerization in the rod outer segment but it also yields vitamin A (all-*trans*-retinol), involved in a variety of metabolic processes. In addition, by reducing all-*trans*-retinal to its alcohol it may influence the abundance of the signal molecule retinoic acid, which can be gained by oxidation of all-*trans*-retinal (Duester, 2000). The expression pattern of zfprRDH 2 agrees well with any or several of these putative functions.

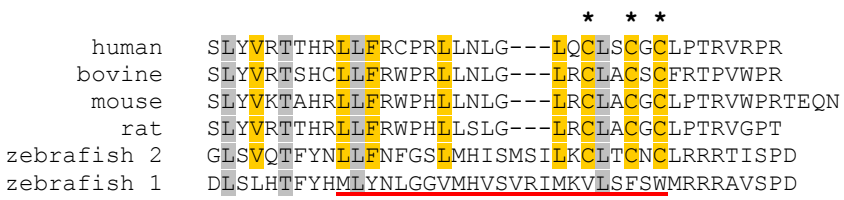


Fig. 39: Motifs associated with membrane localization in the very C-terminus of prRDHs.
The sequence alignment indicates amino acids conserved in all investigated proteins (gray) and those identical only in the mammalian enzyme and zfprRDH 2 (zebrafish 2; orange). The highly preserved cysteines which likely serve as targets for lipid attachment and thereby association of the prRDH with the membrane are marked with asterisks. A red line in zfprRDH 1 (zebrafish 1) outlines the sequence stretch capable of forming a transmembrane helix according to SMART at the ELM server.

In mammals, the family of so far identified and characterized RDHs is quite large and a variety of highly similar enzymes, non-identical concerning enzymatic properties and expression, has evolved. Although elucidation of these proteins in zebrafish is still at the beginning, it may be that diversity of RDHs in mammals is much higher than in fish. Under these circumstances, one or more enzymes recruited from another branch of RDH evolution apart from 17beta-HSD type 1 might in mammals or other vertebrates fulfill the functions that in zebrafish are assigned to zfprRDH 2.

5.2 17beta-HSD type 3: zebrafish vs. mammals

5.2.1 17beta-HSD type 3 from zebrafish: homolog, analog, ortholog?

The search for the zebrafish homolog of mammalian 17beta-HSD type 3 proved to be an even more difficult task than was the case for 17beta-HSD type 1. Two main obstacles aggravated a successful identification: on the one hand, only the mammalian genes from human, mouse and rat were known in contrast to 17beta-HSD type 1 of which the chicken and eel homolog were already characterized. On the other hand, it soon became clear that quite a set of genes closely related to 17beta-HSD type 3 does exist. As a result, several candidate sequences were identified in the *in silico* screen, that displayed very similar grades of amino acid identity towards mammalian 17beta-HSD type 3.
Three candidate genes resulted from the first screen against EST-databases but during subsequent thorough analyses none of them was identified to be the homolog of mammalian 17beta-HSD type 3. Instead, they were identified as two paralogous genes of 17beta-HSD type 12 (zfHSD 12A and zfHSD 12B) and a yet unknown putative steroid dehydrogenase. A second screen against genomic databases of *Takifugu* and zebrafish finally yielded the zebrafish 17beta-HSD type 3 (zfHSD 3), the first to be characterized non-mammalian homolog of this gene.

5.2.1.1 On common ground: shared features of zebrafish and mammalian 17beta-HSD type 3

Selection of expressed or genomic sequences as candidates in search for the homolog always involved a reverse comparison against the mammalian protein sequences. In case of 17beta-HSD type 3 though, overall sequence identity was not very conclusive as it did not yield higher values for one or more of the candidate sequences (see also chapter 4.2.4.1.1). Phylogenetic analyses proved to be very useful tools to assess relations to a set of sequences in an

evolutionary background. These calculations clearly showed homology of the zebrafish 17beta-HSD type 3 to the mammalian proteins.
Investigation of the sequence alignment (compare Fig. 28, chapter 4.2.4.1.3) revealed common features concerning SDR motifs (Fig. 40). These did not only fit the required consensus sequences but demonstrated that residues deviating from the consensus in most cases were identical to that of the mammalian homologs. This aspect is a valuable hint suggesting functional integrity of the zebrafish enzyme in addition to high similarity to features typical for the mammalian 17beta-HSDs type 3.

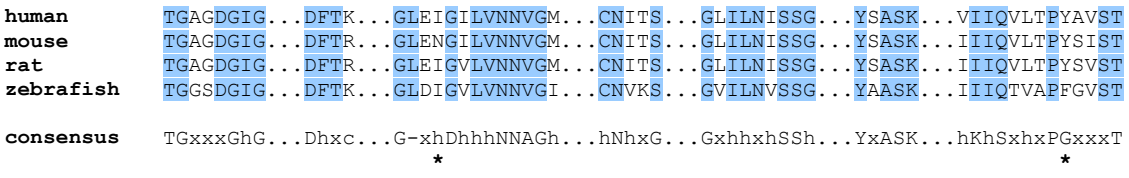


Fig. 40: Classical SDR motifs in 17beta-HSDs type 3 from zebrafish and mammals.
Residues identical in all aligned sequences are shaded in blue. Deviations to the consensus sequence which in addition are not identical in the investigated sequences are marked with asterisks. As the mammalian enzymes have been demonstrated to be active, the alterations observed in the SDR motifs are not likely to affect general functionality of the zebrafish homolog.

Characteristics of mammalian 17beta-HSD type 3 concerning gene architecture and expression were also reflected in the zebrafish homolog. This gene consisted as well of eleven exons with defined exon-intron borders similar to those observed in the mammalian genes. But in addition, size differences in the first, sixth and eleventh exon that could influence 3D structure and therefore function of the enzyme were present. An expression in testis, which is quite typical for the known mammalian 17beta-HSDs type 3 (Geissler et al., 1994), was also observed in the zebrafish homolog. Interestingly, while in mammals this is the predominant site of expression, a variety of 17beta-HSD type 3 expressing tissues aside from testis were identified in zebrafish.
Although, in terms of evolution, the zebrafish and mammalian genes are homologs, these differences may indicate that the function in zebrafish and mammals are not identical, and consequently the genes would not be orthologs. This question will be addressed in the following chapters, analyzing and discussing the observed dissimilarities.

5.2.1.2 Fish vs. mammals – different in more than one aspect. I: structure and activity

In contrast to zebrafish 17beta-HSD type 1, no enzymatic activity towards the tested androgens or estrogens could be detected in case of 17beta-HSD type 3. It can not be excluded that reasons for this lay in the laboratory set-up. All constructs, including zfHSD3 expression constructs, were sequenced prior to overexpression to exclude the possibility of mutations. Production of recombinant protein was also clearly identified by Western Blot. Nevertheless, the detected enzyme might have been inactivated due to incorrect processing caused by the *E.coli* host system. 17beta-HSD type 1 is a cytosolic protein of which at least the zebrafish homolog does not seem to contain any localization or interaction signals that might influence its activity.

In consistence with this a protein with high enzymatic activity was readily expressed in *E.coli*, which is also true for the human homolog (G. Möller, personal communication). In contrast, 17beta-HSD type 3 was reported to be microsomal and therefore would harbor terminal processing and localization signals (see also chapter 5.2.2.1). So far, 17beta-HSD type 3 activity was only characterized in transfected mammalian cells or measured in homogenates of these cells (Geissler et al., 1994; Luu-The et al., 1995; Andersson et al., 1996; Moghrabi et al., 1998; Tsai-Morris et al., 1999; Lindqvist et al., 2001; McKeever et al., 2002). In concert with these findings, attempts to express the active human or murine homolog in *E.coli* have as well failed so far (G. Möller, personal communication).

Alternatively, there is still the possibility that zfHSD 3 indeed does not catalyze the formation of testosterone from androstenedione. The observed differences in amino acid sequence compared to the mammalian homologs might have resulted in an altered substrate specificity. Effects like these were identified in other closely related if not homologous proteins. For example, a shift of substrate specificity occurred between 17beta-HSD type 1 homologs where the rodent enzyme compared to other known homologs shows a drastically improved acceptance of androstenedione. For 17beta-HSD type 6, a murine and a rat homolog exist whereas 17beta-HSD type 9 also seems to be a homolog of this gene but was only identified in mouse (Su et al., 1999). The type 9 has roughly equivalent activities as a 17beta-HSD towards estradiol and androstanediol and as a 3alpha-HSD towards androstanediol and androsterone. In contrast, the type 6 has 10-fold greater 17beta-HSD activity with androstanediol than with estradiol and has low 3alpha-HSD activity with androsterone (Su et al., 1999). A new retinol dehydrogenase-similar gene in mouse (mRDH-S) was identified by Song et al. (2003), sharing 89% identical amino acids with RDH 1. None of the major substrates of RDH 1 was recognized by wild type mRDH-S but were accepted after mutation of a few residues at three different sites present in both, the cofactor and substrate binding area (Song et al., 2003).

These findings reflect how a few exchanges or even point-mutations in the amino acid sequence might alter the protein structure sufficiently to affect its enzymatic potential. Aside from enzymatic measurements, the impact of sequence differences can be assessed by analysis of a structural context, which was successfully done in case of zebrafish 17beta-HSD type 1. For the type 3 though, no crystal structure of a homolog or close relative exists and tentative models consequently are of poor quality (see also below).

Modeling of zfHSD 1 and the two zfprRDHs also highlighted how the observed variations in exon size are reflected in the 3D-structure. Such differences were as well detected between the zebrafish and mammalian sequences (chapter 4.2.2.2) in 17beta-HSD type 3. They occur in the first and last exon as well as exon 6 (Fig. 41) and might very well influence overall structure and enzymatic activity.

The first exon is non-uniform even for the mammalian proteins, and sequence alignment suggests the differences to be due to slightly altered start sites. It is therefore unlikely that this change has an effect on enzyme function. The last exon is in zebrafish by five amino acids (15 nucleotides) shorter than in human, mouse and rat. The affected region displays a degree of conservation similar to the preceding area among the mammalian sequences and might still be involved in substrate binding, which in zebrafish could be altered due to the deletion of four residues. The gap near the very end of exon 11 should have no effect on enzyme function as this region is generally lower conserved and only a few key positions might be part of putative signaling or localization motifs (see also chapter 5.2.2.1).

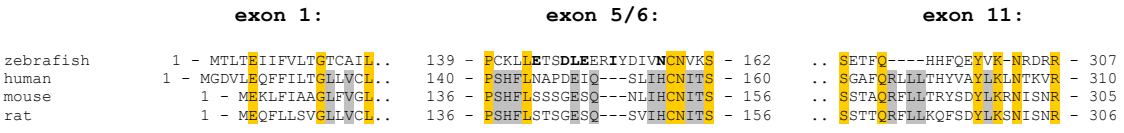


Fig. 41: Localization of exon size differences in the amino acid sequence of zfHSD 3 in comparison to the mammalian homologs.
Amino acids identical in all four sequences are shaded in orange; those identical only in the mammalian proteins are marked in gray. In the zebrafish sequence part of exon 5/6, amino acids identical to one or more 17beta-HSDs type 12 are written in bold.

Concerning the enlarged exon 6 in zebrafish, the exact position of the additional amino acids can only by hypothesized. Homology modeling in case of 17beta-HSD type 1 has demonstrated that deviations in a sequence based alignment might end up at a different position when compared to a structure based alignment (compare chapter 4.1.4.2.1). The affected region in 17beta-HSD type 3 resembles a loop area connecting the central β-strand βD with the subsequent long helix αE when general architecture of the SDR cofactor binding domain is applied. Although it remains unclear whether this change has an effect on enzyme function, it is interesting to note that comparison of zebrafish 17beta-HSD 1 with its mammalian homologs marks the same region as to be of apparently low conservation. Furthermore, the addressed region has six out of fourteen residues identical with 17beta-HSDs type 12 (Fig. 41), suggesting that the observed differences in amino acid number and identity in zebrafish might be remnants from evolution of the type 12 progenitor, that were later modified in the mammalian 17beta-HSDs type 3 (see also chapters 4.2.4.1.2 and 5.2.2.1). The origin and impact of this diversion between zebrafish and mammalian 17beta-HSD type 3 can only be investigated when yet unidentified homologs from other non-mammalian vertebrates become available.

Since impaired function of human 17beta-HSD type 3 in male individuals leads to pseudohermaphroditism (Geissler et al., 1994), a variety of mutations in this gene have been analyzed which either completely abolish or strongly reduce testosterone synthesis. Investigation of these amino acid exchanges allows for quite another way to evaluate the cause of missing enzymatic activity in zfHSD 3. Hence, I compared the fifteen so far identified missense mutations to the zebrafish as well as the two rodent protein sequences as outlined in Fig. 42.
Nine of the fifteen identified positions are completely conserved in zebrafish as well as mouse and rat and therefore can not be involved in altered enzymatic functionality of the zebrafish protein. The mutation Q176P is as well unlikely to play a role since all except the human sequence display a conserved arginine at this position and the rodent enzymes are active. Of the remaining five mutations, two display non-identical amino acids in the zebrafish sequence but may also have no effect since these residues do not resemble the dysfunctional one. For example, A56T, which in the human enzyme abolishes activity, is in zebrafish resembled by A55G but since the glycine is a conservative exchange of alanine, it is unlikely to exhibit a similar effect as threonine in the mutated human enzyme.
Three missense mutations of human 17beta-HSD type 3 remain that are reflected in the zebrafish sequence, namely S65L, N74T and A203V. The latter does exactly match the situation in zebrafish whereas the first two are represented by functionally similar amino acids. None of these amino acids resides inside one of the known SDR motifs and hence are unlikely to have direct effects on cofactor or substrate binding or primary structure scaffold. It is rather

likely that exchange into the dysfunctional amino acids might have secondary effects on protein structure which I investigated on tentative homology models of human and zebrafish 17beta-HSD type 3.

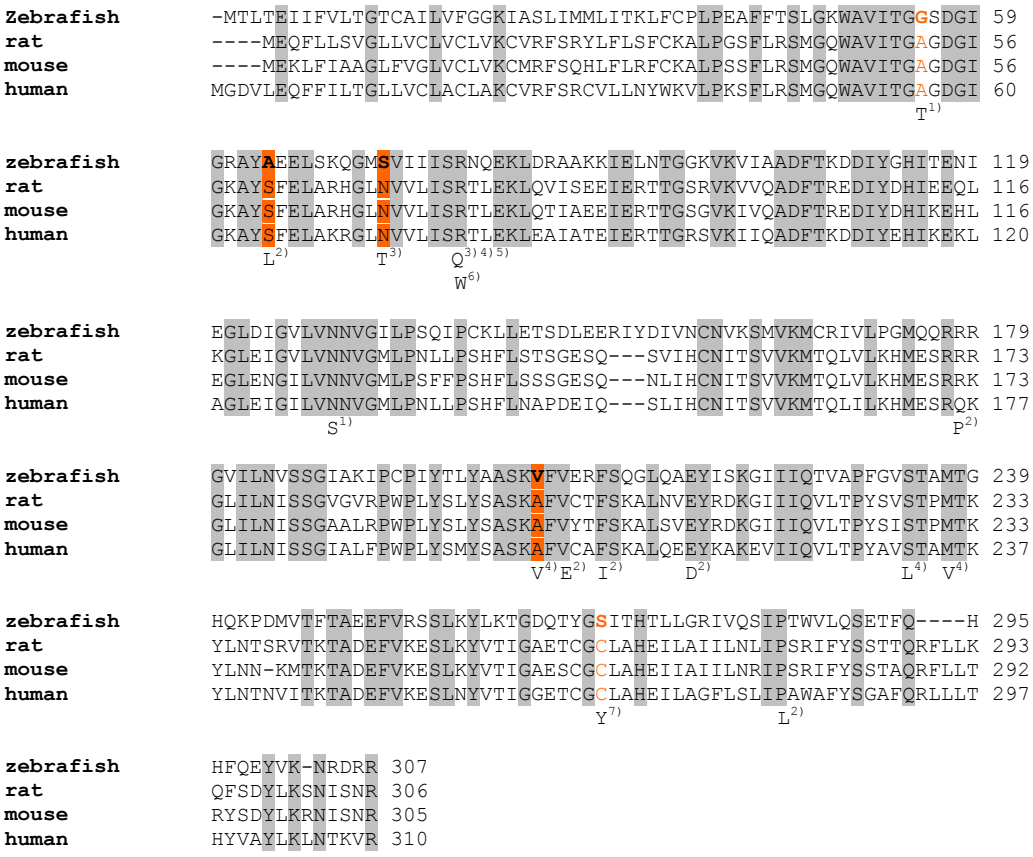


Fig. 42: Identity of inactivating point mutations of human 17beta-HSD type 3 in comparison to the homolog proteins.
Amino acid exchanges known to abolish or diminish enzymatic activity in human 17beta-HSD type 3 are given below the sequence together with footnotes indicating the respective publication. Positions where the zebrafish sequence differs from the mammalian homologs but does not resemble the inactivating residue or one of comparable function are written in orange. Positions shaded in orange mark the presence of a non-identical amino acid in zebrafish, which fits the inactivating residue or one of comparable function. Amino acids identical in all four sequences are shaded in gray. [1)] Moghrabi et al. (1998); [2)] Andersson et al. (1996); [3)] Boehmer et al. (1999); [4)] Geissler et al. (1994); [5)] McKeever et al. (2002); [6)] Bilbao et al. (1998); [7)] Lindqvist et al. (2001).

Fig. 43 summarizes the analysis of the three sites in homology models of human and zebrafish 17beta-HSD type 3. Both models were constructed using 3beta-ketoacyl reductase as a template. Albeit this enzyme is anciently related to the family of 17beta-HSDs type 3 and type 12 (see also chapter 4.2.4.1.2), both models regionally show strong deviations concerning exact length, connection and spatial orientation of secondary structure elements. Nevertheless, the effect exerted by the three mutations in the human protein is recognizable. Residues S65 and A203 in human 17beta-HSD type 3 point into a cleft where they sterically interact with residues positioned vis-à-vis. Concerning S65, these interacting amino acids are not identical in human and zebrafish but it becomes clear that at this position a bulkier amino acid side chain might not fit into the cleft.

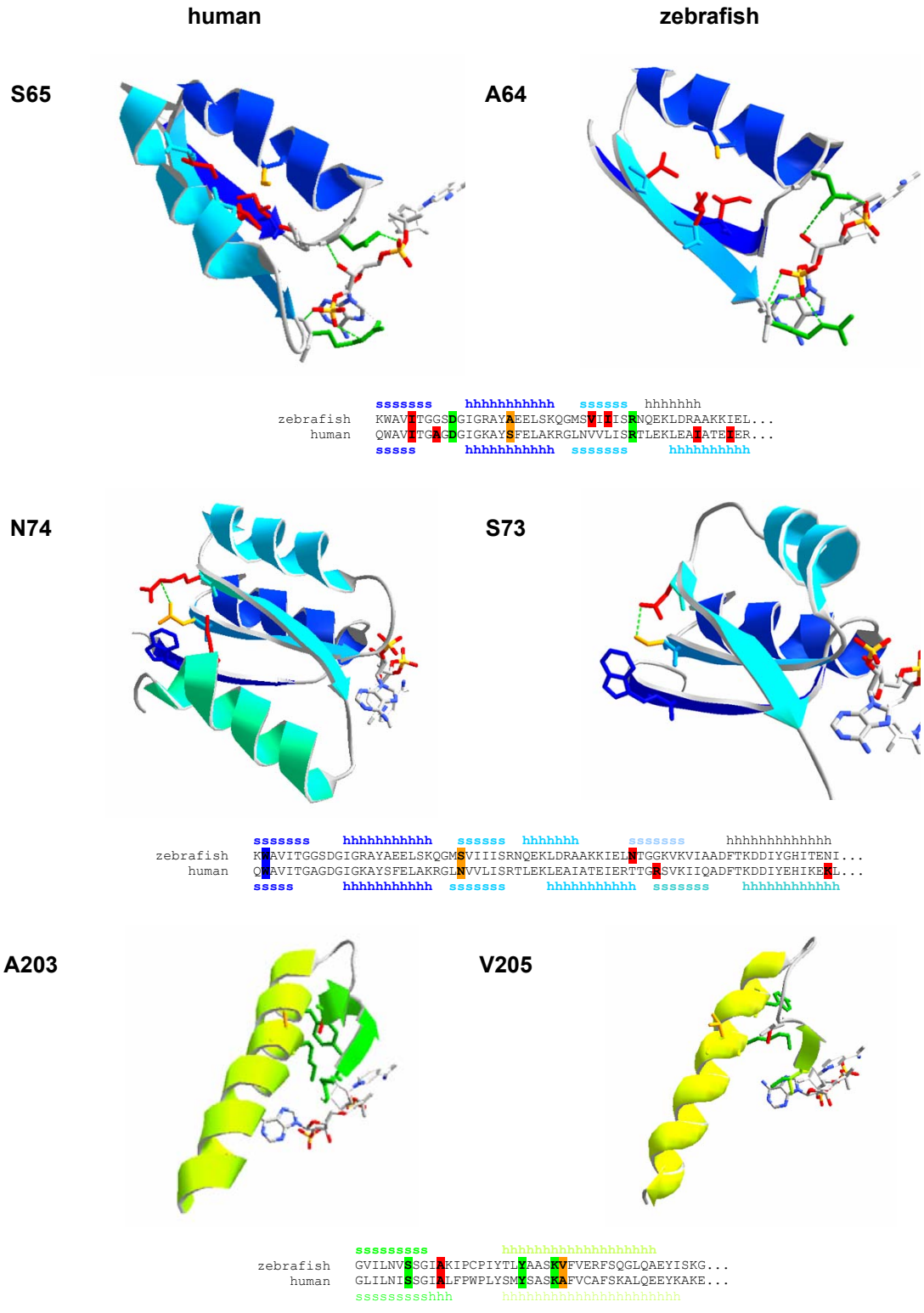


Fig. 43: Analysis of missense mutations in models of human and zebrafish 17beta-HSD type 3.
Tentative homology models of human and zebrafish 17beta-HSD type 3 were constructed using 3beta-ketoacyl reductase (swiss-pdb: 1edoA) as a template. The pictures display the modeled parts of the regarded missense mutations and their structural environment in the human and the corresponding position in zebrafish enzyme. Alignments below show the affected area and amino acids in a linear context. Secondary structures are colored by succession in the overall 3D-model where dark blue indicates N-terminal and yellow indicates C-terminal structures; mutated sites are marked in orange, positions possibly interacting with these are marked in red. Amino acids which facilitate cofactor binding or form the catalytic triad are outlined in green. Hydrogen bonds are depicted as green dashed lines. h: α-helix; s: β-strand

This would result in an expansion of the structure as the interacting residues are placed on subsequent structure elements connected by loop regions. A consequence would then be the displacement of two amino acids vital for cofactor binding (D58 and R80) and hence loss of activity. This would explain why the human mutation S65L abolishes enzyme activity. In the respective model for zfHSD 3 though, alanine (A64) is too small to and additionally has different interaction partners as compared to the human enzyme to exert a similar effect.
A comparable situation might exist in A203 where alanine and its interaction partner A188 are positioned vis-à-vis on subsequent structure elements connected by a loop. The affected α-helix and β-strand harbor three amino acids involved in the catalytic reaction mechanism, S184, Y198 and K202. Steric hindrance of A203 by a larger side chain on the opposite side might again result in a widening of the cleft and subsequent loss of correct positioning of the amino acids involved in the catalytic process. Although in zfHSD 3 the corresponding position of A203 is occupied by V205, which was demonstrated to inactivate the human enzyme, the model suggests a slightly altered positioning giving enough room for valine and its interaction partner on the opposite site of the cleft.
N74 might have an effect different from the aforementioned mutations. In the human and the zebrafish model, N74 is positioned on a β-strand on the outside of the protein but amino acids spatially close to this residue differ in the two species. In both cases hydrogen bonding to an amino acid on the neighboring β-strand and therefore additional stabilization of the central β-strand seems likely but it remains unclear why substitution of N74 by threonine in the human protein leads to an inactive enzyme. Hence, it is also unclear whether serine in the corresponding position in the zebrafish enzyme might have contributed to the observed missing activity.
Altogether, the effect of the three missense mutations in the human enzyme appear comprehensible following investigation in these tentative homology models but they are not likely to affect the zebrafish homolog. The already observed low conservation outside of regions directly interacting with the cofactor may result in a 3D-structure altered sufficiently to accommodate the exchanges as suggested by these tentative models.

5.2.1.3 Fish vs. mammals – different in more than one aspect. II: expression

In the embryo

The onset of mammalian 17beta-HSD type 3 expression is not known. It might be absent during embryogenesis (Mustonen et al., 1997) but was detected in fetal testis of mouse dpc 15.5 (Sha et al., 1996) and rat GD 19 (Barlow et al., 2003). These findings are in concert with the major period of male sexual differentiation which in mice starts at dpc 12-13 and in rat takes place at GD 12-21. In contrast, zebrafish 17beta-HSD type 3 was detected in all investigated stages of embryogenesis. But expression levels during this period might not be uniform as the amount of RT-PCR products seemed to increase slightly during gastrulation and somitogenesis with the strongest signal present in 24 hpf embryos.

In the adult organism

Mammalian 17beta-HSD type 3 with high efficiency converts androstenedione to testosterone and therefore is thought to be mainly expressed in male-specific tissues. Transcripts were demonstrated to be present in high abundance in testis of mice (Mustonen et al., 1997; Sha et

al., 1997), rat (Tsai-Morris et al., 1999) and human (Geissler et al., 1994). But lower expression levels also exist at other sites and were detected upon direct investigation of the respective tissue. In human, 17beta-HSD type 3 was identified in adipose tissue (Corbould et al., 2002) and various sites of the brain (Stoffel-Wagner et al., 1999; Beyenburg et al., 2000). In addition, the enzyme might be present in bone in human (Feix et al., 2001) and rat (Eyre et al., 1998). My *in silico* Northern analysis reflects these data and furthermore suggests the human enzyme to be also expressed in female tissues such as placenta (Fig. 44).

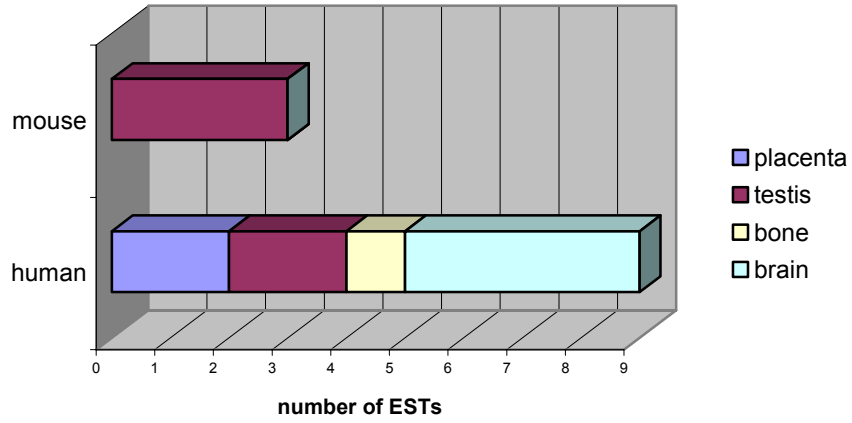


Fig. 44: *In silico* Northern blot of mouse and human 17beta-HSD type 3.
Specific transcripts were detected in a variety of tissues especially in human due to the more expansive EST-database. In case of mouse, only few ESTs were identified which exclusively originated from testis. In both cases, specific ESTs from other sources were also found but not shown here since these originated either from pooled samples or from cancerogenous tissues or cells.

The zebrafish 17beta-HSD type 3 was found to be expressed in the same tissues as in mammals, namely testis, brain and skin; zebrafish adipose tissue and bone were not investigated. In addition, zfHSD 3 was identified in nearly all analyzed sites though transcript levels might be sexually dimorphic as great differences of RT-PCR signal strength were observed comparing male and female tissues.

Comparison to other modes of testosterone synthesis

At first, the observed expression pattern in zebrafish seems to be strikingly different from the situation in mammals but appear in another light when further data on 17beta-HSDs and activity measurements are integrated.

Generation of testosterone can be performed by two main ways (Fig. 45), either having androstenedione or 3β,17β-dihydroxy-delta-5-androstene as a direct precursor. Both possibilities involve 3β-dihydroxysteroid dehydrogenase/isomerase and 17beta-HSD activity which can be exerted by one or more of three different 17beta-HSDs, namely type 1, type 3 and type 5.

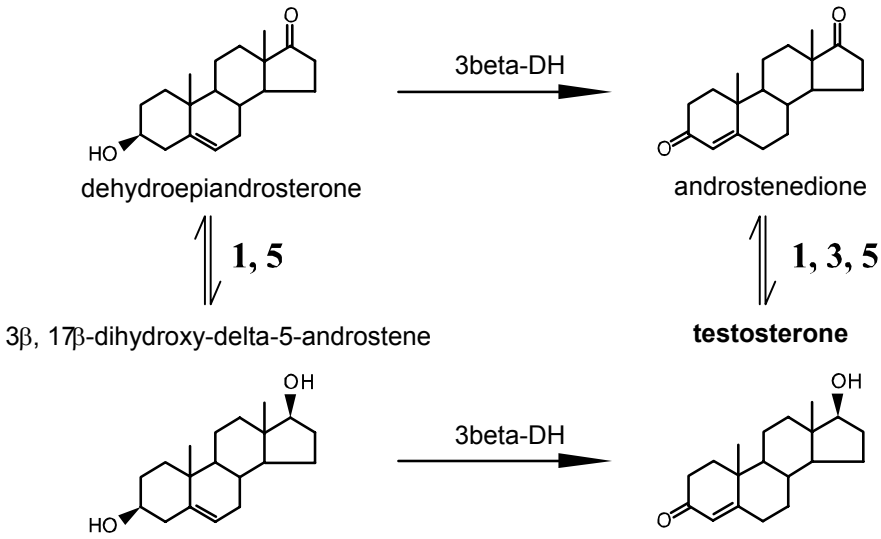


Fig. 45: Enzymatic reactions leading directly to testosterone formation.
Numbers indicate the types of 17beta-HSDs identified to catalyze the respective reactions. Though redox reactions at position C17 are reversible and also performed by 17beta-HSDs, only enzymes preferentially catalyzing the reduction are shown. 3beta-DH: 3beta-hydroxysteroid dehydrogenase /isomerase.

The conversion of androstenedione to testosterone is performed by rat 17beta-HSD type 1 with high efficiency, similar to its potential to reduce estrone. 17beta-HSD type 5 accepts a variety of steroid substrates and also exerts significant 20alpha-HSD activity in addition to the mentioned 17beta-HSD activity (Luu-The et al., 2001). Interestingly, these “additional” 17beta-HSDs beside type 3 show non-identical expression patterns in rodents and human.
17beta-HSD type 5 is expressed in a variety of human tissues (Lin et al., 1997; Dufort et al., 1999) and also in ovary (Qin and Rosenfield, 2000), where it is considered to drive formation of active androgens as 17beta-HSD type 3 is absent (Zhang et al., 1996). Transcripts of 17beta-HSD type 5 were also identified in the human Leydig cells and seminiferous tubules (Pelletier et al., 1999) while in mouse the enzyme is absent from the testis (O'Shaughnessy et al., 2000) and instead shows high expression in liver and to a much lower extent in lung and kidney (Rheault et al., 1999). In murine fetal testis 17beta-HSD type 1 was detected (O'Shaughnessy et al., 2000) and is highly expressed in ovaries (Mustonen et al., 1997). There is also evidence for the existence of another so far uncharacterized 17beta-HSD driving testosterone formation in the testis of adult mice, starting at puberty (O'Shaughnessy et al., 2000). In concordance, two biochemically different 17beta-HSDs, one located in the seminiferous tubules and the other in the interstitial tissue, were described in rats (Murono and Payne, 1976).
Further evidence on testosterone formation at sites not directly associated with sex-specific function comes from activity measurements. Such androgenic activity was identified in several tissues of rat (Martel et al., 1992) and monkey (Labrie et al., 1997). In the latter work, about 20 different male and female tissues and organs were tested (Tab. 11). These measurements identified considerable activity which rival and even exceed those in testis. Furthermore, sexually dimorphic activities were observed. In all non-gonadal tissues activities were higher in female than in male which was most apparent in intestine and liver, where ten-times more testosterone was produced from androstenedione.

Tab. 11: Activity of androstenedione to testosterone formation in monkey tissues.

Tissue	*male*	*female*
Testis	494.4 ± 14.4	
Prostate	73.8 ± 19.2	
Seminal vesicle	86.4 ± 29.4	
Ovary		79.8 ± 37.2
Oviduct		89.1 ± 7.5
Cervix		171.6 ± 70.8
Endometrium		152.4 ± 60
Myometrium		144.6 ± 60
Mammary gland		38.4 ± 6
Small intestine	34.2 ± 0.6	216.6 ± 97.2
Large intestine	293 ± 54	2740 ± 187
Kidney	15 ± 8.4	31.2 ± 6.6
Liver	329 ± 70	2296 ± 542
Lung	45 ± 27	158.4 ± 78
Heart	3 ± 0.6	14.1 ± 3.6
Spleen	12 ± 6	38.4 ± 8.8
Mesenteric fat	47.4 ± 19.2	69 ± 17.4
Skeletal muscle	8.4 ± 4.8	37.8 ± 1.2
Cerebellum	36.6 ± 6.6	112 ± 42
Cerebrum	22.2 ± 0.3	24.3 ± 13.8
Adrenal	124.2 ± 10.2	254.4 ± 160.8

Data are presented as means ± SEM (standard error of mean value) of pmol testosterone formed/mg protein/h. Modified from Labrie et al. (1997).

As compiled above, 17beta-HSD activity transforming androstenedione into testosterone is present in a variety of tissues in mammals aside from sex-specific organs. This activity shows clear sexual differences in non-gonadal tissues and is accomplished by at least two but most likely even more non-identical enzymes, which differ in their expression pattern, enzymatic efficiency and specificity. These data suggest that the observed expression pattern of zfHSD 3 is rather similar to that in mammals concerning sites of androstenedione to testosterone conversion.

5.2.1.4 Fish vs. mammals – different in more than one aspect. III: function

Androgens and estrogen, the two classes of sex steroids, and their role in sexually associated functions such as development, reproduction and behavior are present in all vertebrates. Testosterone is produced by the testis of all vertebrates but may be converted to derivatives that are more biologically active in the species in which they are formed. It is widely accepted that in mammals, testosterone is the central androgen while in teleost fish similar functions are taken over by 11-ketotestosterone or 11-hydroxytestosterone (Kime, 1993). The character of the physiologically active androgen seems to have changed in the course of evolution. In lungfish, 11-oxygenated androgens might be absent (Joss et al., 1996) whereas elasmobranchs seem to employ both classes of steroids (Garnier et al., 1999; Manire et al., 1999); in lampreys, 15alpha-hydroxylated derivatives have been identified to be the predominant gonadal steroids (Lowartz et al., 2003).

Furthermore, when it is referred to physiological active androgens generally their male specific functions in development of sexual organs and reproduction are addressed. Aside from this role, androgens also act in female specific tissues and at somatic sites either through the androgen receptor or as aromatizable precursors of estrogens through the estrogen receptor. Quite notably, application of aromatizable androgens to amphibians (Kuntz et al., 2003) or reptiles (Crews, 1994) during development can lead to both complete all-male or all-female populations depending on time-point and time span of exposure, amount, and androgen character. In mammals, 17beta-HSD type 3 seems to be predominantly employed in gonadal aspects, which is reflected in its high and predominant expression in testis, absence during embryogenesis and high substrate specificity. 17beta-HSD 3 in zebrafish is most likely involved in "somatic" aspects as is suggested by its expression throughout embryogenesis and in many tissues of male and female adults. Additionally, it might support androgen metabolism in testis where it is also present though in a lower degree than in ovaries or male liver, for instance. Although redox reactions at C17 of a variety of androgenic steroids are detectable in different fish species, the identity and character of the corresponding enzymes are still unclear as up to now, no such protein was purified or the respective gene cloned. Substrate specificity of zebrafish 17beta-HSD type 3 remains to be elucidated and will help to understand its physiological role, which does not seem to be identical to that of the mammalian homolog. A combination of different aspects of mammalian enzymes controlling androgen activity is rather likely.

5.2.2 The dawn of a new spawn: comparison to 17beta-HSD type 12 and derived groups

Aside from the zebrafish 17beta-HSD type 3, two homologs of 17beta-HSD type 12 (zfHSD 12A and zfHSD 12B) were identified. Phylogenetic analyses for the first time demonstrated the close relation of these two enzyme groups and furthermore that 17beta-HSD type 12 is an ancient gene already present in yeast. Only recently and independent from the annotation as a putative hydroxysteroid dehydrogenase it has been demonstrated that the encoded protein resembles a 3-ketoacyl-CoA-reductase involved in fatty acid synthesis (Moon and Horton, 2003). A close evolutionary connection to or even identity to fatty acid metabolizing enzymes is also known for other 17beta-HSDs such as type 2 (Baker, 1996), type 4 (Leenders et al., 1996; Breitling et al., 2001b), type 8 (Ando et al., 1996) and type 10 (He et al., 1999; He et al., 2000). The ancient predecessor of 17beta-HSD type 2 is most likely a 3beta-hydroxyacyl-CoA-reductase, an enzyme involved in mitochondrial fatty acid beta-oxidation which catalyzes the reverse reaction of 3-ketoacyl-CoA-reductase. Interestingly, in the cause of evolution 3beta-hydroxyacyl-CoA-reductase gave rise to a huge set of related but functionally different proteins including different families of retinol dehydrogenases beside 17beta- and 11beta-HSD type 2. A similar phenomenon was seen when investigating the evolution of 17beta-HSDs type 3 and type 12. Indeed, 3-ketoacyl-CoA-reductase (17beta-HSD type 12) gave rise to various groups of so far uncharacterized enzymes among which 17beta-HSDs type 3 appear as the most recently evolved group, present in vertebrates only (see chapter 4.2.4.1.2).

5.2.2.1 The legacy of features of fatty acid synthesis in 17beta-HSD type 3

In the phylogenetic analysis, 17beta-HSDs type 3 formed a distinct separate group and the functional diversity from 17beta-HSD type 12 is marked by the clearly different substrate specificity. Nevertheless are features of the fatty acid synthesizing ancestor still visible and might influence the function of the newly evolved enzyme. During the analysis of zebrafish 17beta-HSD type 3 on protein level, it was quite apparent that in contrast to the previously investigated 17beta-HSD type 1 the cofactor binding domain was positioned further downstream and the substrate binding domain was shorter (Fig. 46).
The sequence length of the substrate binding region does not hint at substrate specificity as long as the 3D-structure of this domain is not known. Changes in the architecture of this domain can facilitate specific binding of substrates of different length and scaffold as was seen in the modeling of zebrafish 17beta-HSD type 1 compared to the two prRDHs (see chapter 4.1.4.2.1). Notably, the substrate binding domain of fatty acid metabolizing enzymes is short and binds substrates of various length. Looking at human 17beta-HSDs derived from fatty acid metabolizing ancestors, the substrate binding domain in type 3, type 8 and type 10 comprises about 77, 54 and 58 amino acids, respectively, and the spectrum of accepted, structurally different substrates increases in this order.

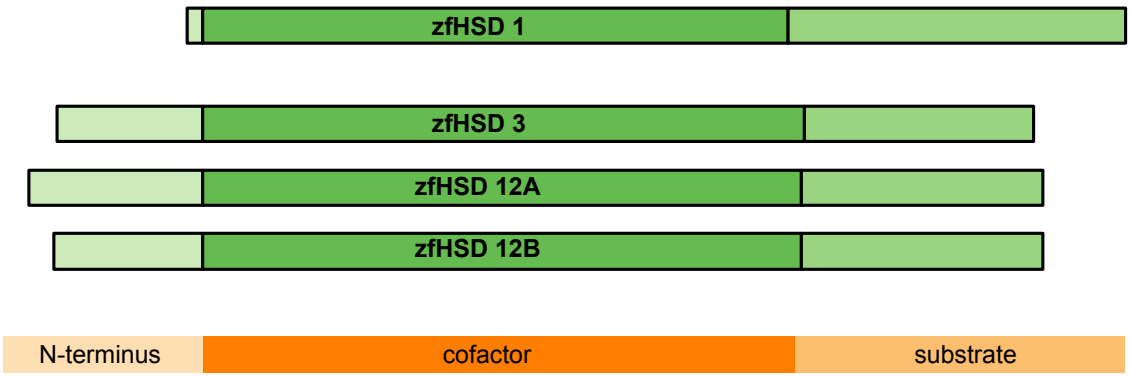


Fig. 46: Domains of zebrafish 17beta-HSD type 3 and 12 in comparison to 17beta-HSD type 1.
The coding sequences of the four zebrafish genes are subdivided into three different parts according to the cofactor binding domain, comprising the seven SDR-motifs (dark green). The preceding region is denoted as N-terminus, the subsequent region as substrate binding. The latter domain might be even shorter due to C-terminal signaling motifs. Respective sizes of the cDNAs are in scale.

The terminal regions

N-terminal as well as C-terminal regions which do not contribute to cofactor- or substrate binding are likely sites of localization, modification or interaction motifs. For all investigated 17beta-HSDs type 3 and 12, the N-terminal region preceding the cofactor binding domain is predicted by the ELM server to contain a signal peptide or transmembrane helix, conferring localization to the ER (Fig. 47). In addition, 17beta-HSDs type 12 harbor a Dilysine ER retention and retrieval signal at their very C-terminus. This motif is positioned at the end of the cytoplasmic tail of type 1 membrane proteins and may either be KKXX or KXKXX (Jackson et

al., 1990). Interplay of the transmembrane and lumenal domain with the exact composition of the Dilysine signal mediate localization to or trafficking between compartments such as the ER, the ERGIC, the cis-Golgi structure and lipid droplets. The mechanisms are not yet fully understood and experimental investigation of subcellular localization is still necessary. It has been shown that replacement of lysine by arginine in the Dilysine motif might be possible as in case of murine NSDHL (NAD(P)H steroid dehydrogenase-like) a C-terminal RKDK mediates ER retrieval and the protein is located exclusively in the Golgi upon deletion of the last four amino acids (Caldas and Herman, 2003). Interestingly, while the murine NSDHL localized to the ER as well as lipid droplets, the human homolog was almost exclusively found on the surface of lipid droplets (Ohashi et al., 2003). In human NSDHL, both lysines of the Dilysine signal are replaced by arginine thereby displaying the motif RRVK in the C-terminus.

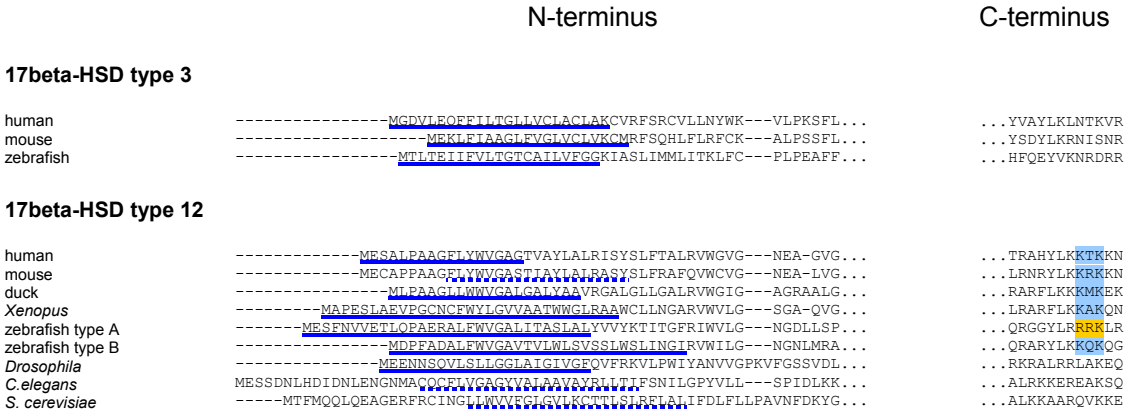


Fig. 47: N-terminal and C-terminal signal motifs in 17beta-HSD type 3 and type 12.
In the N-terminus, sequences are aligned with respect to the first SDR-motif of the cofactor binding region present further downstream (not shown). Regions predicted to facilitate localization to the ER membrane are underlined in blue: single line: putative signaling peptide; broken line: putative transmembrane helix. In the C-terminus, the last twelve amino acids of each sequence are shown. Compliance with the ER retention and retrieval signal is highlighted in blue. An orange mark indicates alteration of the motif which might still be recognized. Analysis was performed at the ELM server.

The *in vivo* function of 17beta-HSD type 12 seems to be the reduction of very long chain saturated or unsaturated fatty acids (Beaudoin et al., 2002; Han et al., 2002; Moon and Horton, 2003). This specific activity is described to be microsomal in contrast to the cytosolic fatty acid synthesis (compare also chapter 5.2.2.2). A localization to the ER could indeed be demonstrated for the human 17beta-HSD type 12 (Moon and Horton, 2003) and fits the presence of the described N- and C-terminal motifs. For the yeast homolog YBR159, localization to the ER has been demonstrated by use of GFP-tagged protein (Han et al., 2002). The C-terminal sequence does not predict a ready Dilysine motif but it might be possible that the last residues are removed upon processing and KKAA further upstream takes over the function of an ER retention/retrieval signal. A similar mechanism could also apply for the *Drosophila* and *C.elegans* enzyme.

17beta-HSDs type 3 apparently do not seem to possess the C-terminal retrieval signal although their N-terminal region suggests transmembranal localization. In concordance, characterization of the mammalian enzymes described them as microsomal, indicating association with the ER membrane or derived structures.

Presence of an N-terminal but no C-terminal motif in 17beta-HSD type 3 was also suggested by Identity plots (compare chapter 4.2.4.1.4). Here, comparison of type 3 and 12 show that conservation is relatively high at the beginning of the proteins and subsequently drops to a minimum around position 20, concordant with the end of the signal peptide/transmembrane region. Following this minimum sequence identity rises again as structures of the cofactor binding domain are formed and peaks at the first SDR-motif. In the C-terminal part of 17beta-HSDs type 12 the same trend is observed. Conservation drops to a minimum around position 280, which might resemble the end of the substrate binding region, and subsequently rises again concomitant with the emerging localization motif. Such a course is absent in the ID plot of 17beta-HSDs type 3. The region subsequent to position 250 is instead marked by a relative steady degree of 30% sequence identity (see chapter 4.2.4.1.4).
The function of a specific transmembrane or membrane-bound localization of 17beta-HSD type 3, which seems to originate from the ancestral 17beta-HSD type 12, and is most likely conserved in mammals as well as zebrafish, is not known. It may not influence performance of the enzymatic reaction as no mutant that is defective in this N-terminal region has yet been characterized. A contribution to enzyme function aside from putative regulatory actions cannot be excluded since substrate specificity has so far only been characterized for full-length proteins.

Expression
In zebrafish, 17beta-HSD type 3 demonstrated a wide ranged expression. Especially during embryogenesis specific transcripts could be found in all investigated stages of development. In chapter 5.2.1.3, this phenomenon was discussed as possibly resulting from zfHSD 3 taking over testosterone synthesis at all these sites where in mammals several different enzymes are employed. Not contradicting but rather complementing this explanation would be the notion that wide-spread expression was inherited from the ancestor 17beta-HSD type 12. Expression pattern of zfHSD 3 and likely function of 17beta-HSD type 3 in fish as well as mammals may hence still be connected to fatty acid synthesis. In this context it is interesting to note that in human 17beta-HSD type 3 was also detected in adipose tissue (Corbould et al., 1998; Corbould et al., 2002) and that at least female intra-abdominal adipose tissue may be substantially androgenic (Corbould et al., 2002). Furthermore, effects of testosterone on the homeostasis of fatty acid reservoirs have been reported. In young men, testosterone supplementation reduced total body adipose tissue (Woodhouse et al., 2004) whereas in cultured brown adipocytes testosterone-treated cells had a lower lipolytic activity (Monjo et al., 2003). These findings indicate a role for androgens and especially 17beta-HSD type 3 in the control of fatty acid homeostasis, which might be a regulatory function inherited in part from 17beta-HSD type12.

5.2.2.2 Redundancy or synergism: two types of 17beta-HSD type 12 in zebrafish

At least two different systems of fatty acid synthesis exist: One is located in the cytosol and catalyses the *de novo* formation of acyl chain lengths up to palmitat (C_{16}). The other is microsomal and elongates already existing acyl chains of medium length thereby generating long and very long chain fatty acids (LCFAs and VLCFAs, respectively). There is evidence that different enzymes are responsible for the elongation of unsaturated and saturated fatty acids and for the elongation of fatty acids of different chain lengths (Cinti et al., 1992). These

substrate specific enzymes catalyze the condensation of malonyl-CoA to the acyl-chain while the subsequent reactions of the elongating process are catalyzed by less specific enzymes. As a consequence, even exogenous substances, that are normally not present in the organism, are readily accepted as substrates and converted (Beaudoin et al., 2000; Beaudoin et al., 2002).
The identity of enzymes participating in the microsomal elongase system is not yet fully revealed. The yeast as well as the murine and human enzyme catalyzing the two reduction steps have recently been identified. In yeast, Tsc13p, the enoyl reductase, is non-redundant and essential whereas the ketoreductase function of Ybr159p, the ancient homolog of 17beta-HSD type 12 (see chapter 4.2.4.1.2), could partially be substituted by Ayr1 (Han et al., 2002). The double mutant ybr159Δayr1Δ is not viable due to complete disruption of VLCFA synthesis (Han et al., 2002).
The findings in yeast stress the importance of VLCFA synthesis which is considered to be essential for all organisms. Ybr159p seems to be the major enzyme in the described process. While Ayr1 is not related to it and predominantly functions as a 1-acyldihydroxyacetone-phosphate reductase (Athenstaedt and Daum, 2000), it can partially substitute for Ybr159p. ADR activity has also been demonstrated in mammals (Datta et al., 1990) but only in yeast have the corresponding gene and gene product so far been characterized. Therefore, it is not clear whether a similar mechanism as observed in yeast is present in vertebrates. The existence of a second 17beta-HSD type 12 in zebrafish is intriguing as it might indicate redundancy in VLCFA synthesis as well in vertebrates. In this case it could be speculated that one zfHSD 12 functions predominantly as the microsomal 3-ketoreductase whereas the other zfHSD 12 has a slightly different function but may still resemble some 3-ketoreductase activity in VLCFA synthesis. Although substrate specificity concerning fatty acids was not investigated in the work at hand, findings from expression profiles and protein sequence analysis support this notion.
Especially concerning the expression pattern, zfHSD 12A resembles a 3-ketoacyl-CoA reductase as the human and mouse homologs were found to be expressed in all tissues investigated but differed slightly in the relative expression level depending on the tissue type (Moon and Horton, 2003). A closer relation to the mammalian proteins was also suggested by overall sequence identity which was slightly higher for type A than type B, especially in comparison to the human enzyme. In the Identity Plots several areas showed higher amino acid conservation between zfHSD 12A and the mammalian enzymes. This was more pronounced in the cofactor binding region. In contrast, the C-terminal signaling motif for ER retrieval/retention is altered in zfHSD 12A and hence localization of the zebrafish protein may differ from that of the human and mouse enzyme, which localize to the ER (Moon and Horton, 2003). A better resemblance of the Dilysine motif was instead revealed in zfHSD 12B and might indicate a preservation of 3-ketoacyl-CoA reductase function in fatty acid elongation congruent with a likely microsomal localization. Although expression of the type 12B was found in many zebrafish tissues, it was distinctly different from the ubiquitous expression of type 12A. In the adult fish, expression was low and readily detectable especially in gonads, liver and intestine suggesting these tissues to have an elevated fatty acid turnover or need for specific VLCFA generated via zfHSD 12B activity. This supports the hypothesis that zfHSD 12B is involved in functions overlapping with zfHSD 12A but different concerning substrates required in greater extent only in certain tissues. This is also reflected in onset of zfHSD 12B expression subsequent to gastrulation while zfHSD 12A transcripts were present in all investigated stages indicating housekeeping-like functions.

It is not clear when the duplication event in 17beta-HSD type 12 leading to the paralogous zebrafish genes occurred. My phylogenetic analyses (chapter 4.2.4.1.2) suggest that paralogs might also be present in other teleosts and are not zebrafish specific. The high potential of 17beta-HSD type 12 for duplication and subsequent gain of function for the duplicate is mirrored in the complexity of the phylogenetic tree (see Fig. 27 in chapter 4.2.4.1.2). Although no 17beta-HSD type 12 paralogs were found in other mammals, it can not be excluded that functional orthologs have evolved from one of these earlier duplications. In addition, partial substitution could be achieved by an evolutionary non-related protein as was observed in case of Ayr1 in yeast.

6 References

Adamski, J. and Jakob, F.J. (2001) A guide to 17beta-hydroxysteroid dehydrogenases. *Mol Cell Endocrinol*, **171**, 1-4.

Akinola, L.A., Poutanen, M. and Vihko, R. (1996) Cloning of rat 17 beta-hydroxysteroid dehydrogenase type 2 and characterization of tissue distribution and catalytic activity of rat type 1 and type 2 enzymes. *Endocrinology*, **137**, 1572-1579.

Akinola, L.A., Poutanen, M., Peltoketo, H., Vihko, R. and Vihko, P. (1998) Characterization of rat 17 beta-hydroxysteroid dehydrogenase type 1 gene and mRNA transcripts. *Gene*, **208**, 229-238.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*, **25**, 3389-3402.

Andersen, L., Holbech, H., Gessbo, A., Norrgren, L. and Petersen, G.I. (2003) Effects of exposure to 17alpha-ethinylestradiol during early development on sexual differentiation and induction of vitellogenin in zebrafish (Danio rerio). *Comp Biochem Physiol C Toxicol Pharmacol*, **134**, 365-374.

Andersson, S., Geissler, W.M., Wu, L., Davis, D.L., Grumbach, M.M., New, M.I., Schwarz, H.P., Blethen, S.L., Mendonca, B.B., Bloise, W., Witchel, S.F., Cutler, G.B., Jr., Griffin, J.E., Wilson, J.D. and Russel, D.W. (1996) Molecular genetics and pathophysiology of 17 beta-hydroxysteroid dehydrogenase 3 deficiency. *J Clin Endocrinol Metab*, **81**, 130-136.

Ando, A., Kikuti, Y.Y., Shigenari, A., Kawata, H., Okamoto, N., Shiina, T., Chen, L., Ikemura, T., Abe, K., Kimura, M. and Inoko, H. (1996) cDNA cloning of the human homologues of the mouse Ke4 and Ke6 genes at the centromeric end of the human MHC region. *Genomics*, **35**, 600-602.

Aparicio, S. (2000) Vertebrate evolution: recent perspectives from fish. *Trends Genet*, **16**, 54-56.

Armen, T.A. and Gay, C.V. (2000) Simultaneous detection and functional response of testosterone and estradiol receptors in osteoblast plasma membranes. *J Cell Biochem*, **79**, 620-627.

Athenstaedt, K. and Daum, G. (2000) 1-Acyldihydroxyacetone-phosphate reductase (Ayr1p) of the yeast Saccharomyces cerevisiae encoded by the open reading frame YIL124w is a major component of lipid particles. *J Biol Chem*, **275**, 235-240.

Azzi, A., Rehse, P.H., Zhu, D.W., Campbell, R.L., Labrie, F. and Lin, S.X. (1996) Crystal structure of human estrogenic 17 beta-hydroxysteroid dehydrogenase complexed with 17 beta-estradiol. *Nat Struct Biol*, **3**, 665-668.

Baker, M.E. (1996) Unusual evolution of 11beta- and 17beta-hydroxysteroid and retinol dehydrogenases. *Bioessays*, **18**, 63-70.

Bakker, J., Honda, S., Harada, N. and Balthazart, J. (2003) The aromatase knockout (ArKO) mouse provides new evidence that estrogens are required for the development of the female brain. *Ann N Y Acad Sci*, **1007**, 251-262.

Barlow, N.J., Phillips, S.L., Wallace, D.G., Sar, M., Gaido, K.W. and Foster, P.M. (2003) Quantitative changes in gene expression in fetal rat testes following exposure to di(n-butyl) phthalate. *Toxicol Sci*, **73**, 431-441.

Beachy, P.A., Cooper, M.K., Young, K.E., von Kessler, D.P., Park, W.J., Hall, T.M., Leahy, D.J. and Porter, J.A. (1997) Multiple roles of cholesterol in hedgehog protein biogenesis and signaling. *Cold Spring Harb Symp Quant Biol*, **62**, 191-204.

Beaudoin, F., Gable, K., Sayanova, O., Dunn, T. and Napier, J.A. (2002) A Saccharomyces cerevisiae gene required for heterologous fatty acid elongase activity encodes a microsomal beta-keto-reductase. *J Biol Chem*, **277**, 11481-11488.

Beaudoin, F., Michaelson, L.V., Hey, S.J., Lewis, M.J., Shewry, P.R., Sayanova, O. and Napier, J.A. (2000) Heterologous reconstitution in yeast of the polyunsaturated fatty acid biosynthetic pathway. *Proc Natl Acad Sci U S A*, **97**, 6421-6426.

Benach, J., Atrian, S., Gonzalez-Duarte, R. and Ladenstein, R. (1998) The refined crystal structure of Drosophila lebanonensis alcohol dehydrogenase at 1.9 A resolution. *J Mol Biol*, **282**, 383-399.

Beyenburg, S., Watzka, M., Blumcke, I., Schramm, J., Bidlingmaier, F., Elger, C.E. and Stoffel-Wagner, B. (2000) Expression of mRNAs encoding for 17beta-hydroxisteroid dehydrogenase isozymes 1, 2, 3 and 4 in epileptic human hippocampus. *Epilepsy Res*, **41**, 83-91.

Beyer, C., Green, S.J. and Hutchison, J.B. (1994) Androgens influence sexual differentiation of embryonic mouse hypothalamic aromatase neurons in vitro. *Endocrinology*, **135**, 1220-1226.

Bhandari, R.K., Higa, M., Nakamura, S. and Nakamura, M. (2004) Aromatase inhibitor induces complete sex change in the protogynous honeycomb grouper (Epinephelus merra). *Mol Reprod Dev*, **67**, 303-307.

Bilbao, J.R., Loridan, L., Audi, L., Gonzalo, E. and Castano, L. (1998) A novel missense (R80W) mutation in 17-beta-hydroxysteroid dehydrogenase type 3 gene associated with male pseudohermaphroditism. *Eur J Endocrinol*, **139**, 330-333.

Biswas, M.G. and Russell, D.W. (1997) Expression cloning and characterization of oxidative 17beta- and 3alpha-hydroxysteroid dehydrogenases from rat and human prostate. *J Biol Chem*, **272**, 15959-15966.

Blomquist, C.H. (1995) Kinetic analysis of enzymic activities: prediction of multiple forms of 17 beta-hydroxysteroid dehydrogenase. *J Steroid Biochem Mol Biol*, **55**, 515-524.

Boehmer, A.L., Brinkmann, A.O., Sandkuijl, L.A., Halley, D.J., Niermeijer, M.F., Andersson, S., de Jong, F.H., Kayserili, H., de Vroede, M.A., Otten, B.J., Rouwe, C.W., Mendonca, B.B., Rodrigues, C., Bode, H.H., de Ruiter, P.E., Delemarre-van de Waal, H.A. and Drop, S.L. (1999) 17Beta-hydroxysteroid dehydrogenase-3 deficiency: diagnosis, phenotypic variability, population genetics, and worldwide distribution of ancient and de novo mutations. *J Clin Endocrinol Metab*, **84**, 4713-4721.

Borg, B., Mayer, I., Lambert, J., Granneman, J. and Schulz, R. (1992) Metabolism of androstenedione and 11-ketotestosterone in the kidney of the three-spined stickleback, Gasterosteus aculeatus. *Gen Comp Endocrinol*, **86**, 248-256.

Borg, B. (1994) Mini review-Androgens in teleost fishes. *Comp Biochem Physiol*, **109C**, 219-245.

Braat, A.K., Zandbergen, T., van de Water, S., Goos, H.J. and Zivkovic, D. (1999) Characterization of zebrafish primordial germ cells: morphology and early distribution of vasa RNA. *Dev Dyn*, **216**, 153-167.

Brantley, R.K., Wingfield, J.C. and Bass, A.H. (1993) Sex steroid levels in Porichthys notatus, a fish with alternative reproductive tactics, and a review of the hormonal bases for male dimorphism among teleost fishes. *Horm Behav*, **27**, 332-347.

Breitling, R., Laubner, D. and Adamski, J. (2001a) Structure-based phylogenetic analysis of short-chain alcohol dehydrogenases and reclassification of the 17beta-hydroxysteroid dehydrogenase family. *Mol Biol Evol*, **18**, 2154-2161.

Breitling, R., Marijanovic, Z., Perovic, D. and Adamski, J. (2001b) Evolution of 17beta-HSD type 4, a multifunctional protein of beta-oxidation. *Mol Cell Endocrinol*, **171**, 205-210.

Britt, K.L. and Findlay, J.K. (2002) Estrogen actions in the ovary revisited. *J Endocrinol*, **175**, 269-276.

Bury, N.R., Sturm, A., Le Rouzic, P., Lethimonier, C., Ducouret, B., Guiguen, Y., Robinson-Rechavi, M., Laudet, V., Rafestin-Oblin, M.E. and Prunet, P. (2003) Evidence for two distinct functional glucocorticoid receptors in teleost fish. *J Mol Endocrinol*, **31**, 141-156.

Caldas, H. and Herman, G.E. (2003) NSDHL, an enzyme involved in cholesterol biosynthesis, traffics through the Golgi and accumulates on ER membranes and on the surface of lipid droplets. *Hum Mol Genet*, **12**, 2981-2991.

Callard, G.V., Pudney, J.A., Kendall, S.L. and Reinboth, R. (1984) In vitro conversion of androgen to estrogen in amphioxus gonadal tissues. *Gen Comp Endocrinol*, **56**, 53-58.

Chiang, E.F., Yan, Y.L., Tong, S.K., Hsiao, P.H., Guiguen, Y., Postlethwait, J. and Chung, B.C. (2001) Characterization of duplicated zebrafish cyp19 genes. *J Exp Zool*, **290**, 709-714.

Cinti, D.L., Cook, L., Nagi, M.N. and Suneja, S.K. (1992) The fatty acid chain elongation system of mammalian endoplasmic reticulum. *Prog Lipid Res*, **31**, 1-51.

Corbould, A.M., Bawden, M.J., Lavranos, T.C., Rodgers, R.J. and Judd, S.J. (2002) The effect of obesity on the ratio of type 3 17beta-hydroxysteroid dehydrogenase mRNA to cytochrome P450 aromatase mRNA in subcutaneous abdominal and intra-abdominal adipose tissue of women. *Int J Obes Relat Metab Disord*, **26**, 165-175.

Corbould, A.M., Judd, S.J. and Rodgers, R.J. (1998) Expression of types 1, 2, and 3 17 beta-hydroxysteroid dehydrogenase in subcutaneous abdominal and intra-abdominal adipose tissue of women. *J Clin Endocrinol Metab*, **83**, 187-194.

Crews, D. (1994) Temperature, steroids and sex determination. *J Endocrinol*, **142**, 1-8.

Cyr, D.G., Eales, J.G. (1996) Interrelationships between thyroidal and reproductive endocrine systems in fish. *Rev Fish Biol Fish*, **6**, 165-200.

Dalfo, D., Canestro, C., Albalat, R. and Gonzalez-Duarte, R. (2001) Characterization of a microsomal retinol dehydrogenase gene from amphioxus: retinoid metabolism before vertebrates. *Chem Biol Interact*, **130-132**, 359-370.

Datta, S.C., Ghosh, M.K. and Hajra, A.K. (1990) Purification and properties of acyl/alkyl dihydroxyacetone-phosphate reductase from guinea pig liver peroxisomes. *J Biol Chem*, **265**, 8268-8274.

Desterro, J.M., Rodriguez, M.S. and Hay, R.T. (1998) SUMO-1 modification of IkappaBalpha inhibits NF-kappaB activation. *Mol Cell*, **2**, 233-239.

Di Cosmo, A., Di Cristo, C. and Paolucci, M. (2002) A estradiol-17beta receptor in the reproductive system of the female of Octopus vulgaris: characterization and immunolocalization. *Mol Reprod Dev*, **61**, 367-375.

Duester, G. (2000) Families of retinoid dehydrogenases regulating vitamin A function: production of visual pigment and retinoic acid. *Eur J Biochem*, **267**, 4315-4324.

Dufort, I., Rheault, P., Huang, X.F., Soucy, P. and Luu-The, V. (1999) Characteristics of a highly labile human type 5 17beta-hydroxysteroid dehydrogenase. *Endocrinology*, **140**, 568-574.

Dumont, M., Luu-The, V., de Launoit, Y. and Labrie, F. (1992) Expression of human 17 beta-hydroxysteroid dehydrogenase in mammalian cells. *J Steroid Biochem Mol Biol*, **41**, 605-608.

Eddy, E.M., Washburn, T.F., Bunch, D.O., Goulding, E.H., Gladen, B.C., Lubahn, D.B. and Korach, K.S. (1996) Targeted disruption of the estrogen receptor gene in male mice causes alteration of spermatogenesis and infertility. *Endocrinology*, **137**, 4796-4805.

Eyre, L.J., Bland, R., Bujalska, I.J., Sheppard, M.C., Stewart, P.M. and Hewison, M. (1998) Characterization of aromatase and 17 beta-hydroxysteroid dehydrogenase expression in rat osteoblastic cells. *J Bone Miner Res*, **13**, 996-1004.

Falkenstein, E. and Wehling, M. (2000) Nongenomically initiated steroid actions. *Eur J Clin Invest*, **30 Suppl 3**, 51-54.

Feix, M., Wolf, L. and Schweikert, H.U. (2001) Distribution of 17beta-hydroxysteroid dehydrogenases in human osteoblast-like cells. *Mol Cell Endocrinol*, **171**, 163-164.

Gangloff, A., Shi, R., Nahoum, V. and Lin, S.X. (2003) Pseudo-symmetry of C19 steroids, alternative binding orientations, and multispecificity in human estrogenic 17beta-hydroxysteroid dehydrogenase. *Faseb J*, **17**, 274-276.

Garnier, D.H., Sourdaine, P. and Jegou, B. (1999) Seasonal variations in sex steroids and male sexual characteristics in Scyliorhinus canicula. *Gen Comp Endocrinol*, **116**, 281-290.

Geissler, W.M., Davis, D.L., Wu, L., Bradshaw, K.D., Patel, S., Mendonca, B.B., Elliston, K.O., Wilson, J.D., Russell, D.W. and Andersson, S. (1994) Male pseudohermaphroditism caused by mutations of testicular 17 beta-hydroxysteroid dehydrogenase 3. *Nat Genet*, **7**, 34-39.

Ghersevich, S.A., Poutanen, M.H., Martikainen, H.K. and Vihko, R.K. (1994) Expression of 17 beta-hydroxysteroid dehydrogenase in human granulosa cells: correlation with follicular size, cytochrome P450 aromatase activity and oestradiol production. *J Endocrinol*, **143**, 139-150.

Ghosh, D., Pletnev, V.Z., Zhu, D.W., Wawrzak, Z., Duax, W.L., Pangborn, W., Labrie, F. and Lin, S.X. (1995) Structure of human estrogenic 17 beta-hydroxysteroid dehydrogenase at 2.20 A resolution. *Structure*, **3**, 503-513.

Ghosh, D. and Vihko, P. (2001) Molecular mechanisms of estrogen recognition and 17-keto reduction by human 17beta-hydroxysteroid dehydrogenase 1. *Chem Biol Interact*, **130-132**, 637-650.

Gimeno, S., Gerritsen, A., Bowmer, T. and Komen, H. (1996) Feminization of male carp. *Nature*, **384**, 221-222.

Govoroun, M., McMeel, O.M., Mecherouki, H., Smith, T.J. and Guiguen, Y. (2001) 17beta-estradiol treatment decreases steroidogenic enzyme messenger ribonucleic acid levels in the rainbow trout testis. *Endocrinology*, **142**, 1841-1848.

Grazzini, E., Guillon, G., Mouillac, B. and Zingg, H.H. (1998) Inhibition of oxytocin receptor function by direct binding of progesterone. *Nature*, **392**, 509-512.

Greenwood, A.K., Butler, P.C., White, R.B., DeMarco, U., Pearce, D. and Fernald, R.D. (2003) Multiple corticosteroid receptors in a teleost fish: distinct sequences, expression patterns, and transcriptional activities. *Endocrinology*, **144**, 4226-4236.

Guex, N. and Peitsch, M.C. (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*, **18**, 2714-2723.

Guiguen, Y., Baroiller, J.F., Ricordel, M.J., Iseki, K., McMeel, O.M., Martin, S.A. and Fostier, A. (1999) Involvement of estrogens in the process of sex differentiation in two fish species: the rainbow trout (Oncorhynchus mykiss) and a tilapia (Oreochromis niloticus). *Mol Reprod Dev*, **54**, 154-162.

Gunnarsson, C., Olsson, B.M. and Stal, O. (2001) Abnormal expression of 17beta-hydroxysteroid dehydrogenases in breast cancer predicts late recurrence. *Cancer Res*, **61**, 8448-8451.

Haeseleer, F., Huang, J., Lebioda, L., Saari, J.C. and Palczewski, K. (1998) Molecular characterization of a novel short-chain dehydrogenase/reductase that reduces all-trans-retinal. *J Biol Chem*, **273**, 21790-21799.

Hamburger, V. and Hamilton, H.L. (1992) A series of normal stages in the development of the chick embryo. 1951. *Dev Dyn*, **195**, 231-272.

Hammes, S.R. (2003) The further redefining of steroid-mediated signaling. *Proc Natl Acad Sci U S A*, **100**, 2168-2170.

Han, G., Gable, K., Kohlwein, S.D., Beaudoin, F., Napier, J.A. and Dunn, T.M. (2002) The Saccharomyces cerevisiae YBR159w gene encodes the 3-ketoreductase of the microsomal fatty acid elongase. *J Biol Chem*, **277**, 35440-35449.

Hancock, J.F., Magee, A.I., Childs, J.E. and Marshall, C.J. (1989) All ras proteins are polyisoprenylated but only some are palmitoylated. *Cell*, **57**, 1167-1177.

Hand, R.A., Jia, N., Bard, M. and Craven, R.J. (2003) Saccharomyces cerevisiae Dap1p, a novel DNA damage response protein related to the mammalian membrane-associated progesterone receptor. *Eukaryot Cell*, **2**, 306-317.

Hawkins, M.B., Thornton, J.W., Crews, D., Skipper, J.K., Dotte, A. and Thomas, P. (2000) Identification of a third distinct estrogen receptor and reclassification of estrogen receptors in teleosts. *Proc Natl Acad Sci U S A*, **97**, 10751-10756.

He, X.Y., Merz, G., Mehta, P., Schulz, H. and Yang, S.Y. (1999) Human brain short chain L-3-hydroxyacyl coenzyme A dehydrogenase is a single-domain multifunctional enzyme. Characterization of a novel 17beta-hydroxysteroid dehydrogenase. *J Biol Chem*, **274**, 15014-15019.

He, X.Y., Merz, G., Yang, Y.Z., Pullakart, R., Mehta, P., Schulz, H. and Yang, S.Y. (2000) Function of human brain short chain L-3-hydroxyacyl coenzyme A dehydrogenase in androgen metabolism. *Biochim Biophys Acta*, **1484**, 267-277.

Hines, G.A., Watts, S.A., Sower, S.A. and Walker, C.W. (1992) Sex steroid levels in the testes, ovaries, and pyloric caeca during gametogenesis in the sea star Asterias vulgaris. *Gen Comp Endocrinol*, **87**, 451-460.

Hooft, R.W., Vriend, G., Sander, C. and Abola, E.E. (1996) Errors in protein structures. *Nature*, **381**, 272.

Hsu, H.J., Hsiao, P., Kuo, M.W. and Chung, B.C. (2002) Expression of zebrafish cyp11a1 as a maternal transcript and in yolk syncytial layer. *Gene Expr Patterns*, **2**, 219-222.

Hutchison, J.B., Beyer, C., Hutchison, R.E. and Wozniak, A. (1997) Sex differences in the regulation of embryonic brain aromatase. *J Steroid Biochem Mol Biol*, **61**, 315-322.

Ijiri, S., Kazeto, Y., Lokman, P.M., Adachi, S. and Yamauchi, K. (2003) Characterization of a cDNA encoding P-450 aromatase (CYP19) from Japanese eel ovary and its expression in ovarian follicles during induced ovarian development. *Gen Comp Endocrinol*, **130**, 193-203.

Ikeuchi, T., Todo, T., Kobayashi, T. and Nagahama, Y. (1999) cDNA cloning of a novel androgen receptor subtype. *J Biol Chem*, **274**, 25205-25209.

Ikeuchi, T., Todo, T., Kobayashi, T. and Nagahama, Y. (2002) A novel progestogen receptor subtype in the Japanese eel, Anguilla japonica. *FEBS Lett*, **510**, 77-82.

Itagaki, E. and Iwaya, T. (1988) Purification and characterization of 17 beta-hydroxysteroid dehydrogenase from Cylindrocarpon radicicola. *J Biochem (Tokyo)*, **103**, 1039-1044.

Jackson, M.R., Nilsson, T. and Peterson, P.A. (1990) Identification of a consensus motif for retention of transmembrane proteins in the endoplasmic reticulum. *Embo J*, **9**, 3153-3162.

Jones, M.E., Thorburn, A.W., Britt, K.L., Hewitt, K.N., Wreford, N.G., Proietto, J., Oz, O.K., Leury, B.J., Robertson, K.M., Yao, S. and Simpson, E.R. (2000) Aromatase-deficient (ArKO) mice have a phenotype of increased adiposity. *Proc Natl Acad Sci U S A*, **97**, 12735-12740.

Joss, J.M., Edwards, A. and Kime, D.E. (1996) In vitro biosynthesis of androgens in the Australian lungfish, Neoceratodus forsteri. *Gen Comp Endocrinol*, **101**, 256-263.

Jungblut, P.W., Gaues, J., Gorlich, L., Hughes, A., Kallweit, E., Kielhorn, J., Little, M., Maschler, I., McCann, S., Parl, F., Rosenfeld, G.C., Sierralta, W., Stone, G., Szendro, P.I., Teran, C., Truitt, A.J. and Wagner, R.K. (1981) Intracellular actions of gonadal steroids. *Exp Brain Res*, **Suppl 3**, 37-60.

Kallberg, Y., Oppermann, U., Jornvall, H. and Persson, B. (2002) Short-chain dehydrogenases/reductases (SDRs). *Eur J Biochem*, **269**, 4409-4417.

Kazeto, Y., Ijiri, S., Matsubara, H., Adachi, S. and Yamauchi, K. (2000a) Cloning of 17beta-hydroxysteroid dehydrogenase-I cDNAs from Japanese eel ovary. *Biochem Biophys Res Commun*, **279**, 451-456.

Kazeto, Y., Ijiri, S., Matsubara, H., Adachi, S. and Yamauchi, K. (2003) Molecular cloning and characterization of 3beta-hydroxysteroid dehydrogenase/Delta5-Delta4 isomerase cDNAs from Japanese eel ovary. *J Steroid Biochem Mol Biol*, **85**, 49-56.

Kazeto, Y., Ijiri, S., Todo, T., Adachi, S. and Yamauchi, K. (2000b) Molecular cloning and characterization of Japanese eel ovarian P450c17 (CYP17) cDNA. *Gen Comp Endocrinol*, **118**, 123-133.

Kedishvili, N.Y., Chumakova, O.V., Chetyrkin, S.V., Belyaeva, O.V., Lapshina, E.A., Lin, D.W., Matsumura, M. and Nelson, P.S. (2002) Evidence that the human gene for prostate short-chain dehydrogenase/reductase (PSDR1) encodes a novel retinal reductase (RalR1). *J Biol Chem*, **277**, 28909-28915.

Keshan, B. and Ray, A.K. (1998) Action of estradiol-17beta on the synthetic activity of the silk gland in Bombyx mori L. *J Insect Physiol*, **44**, 491-498.

Kim, S.J., Ogasawara, K., Park, J.G., Takemura, A. and Nakamura, M. (2002) Sequence and expression of androgen receptor and estrogen receptor gene in the sex types of protogynous wrasse, Halichoeres trimaculatus. *Gen Comp Endocrinol*, **127**, 165-173.

Kime, D.E. (1993) “Classical” and “non-classical” reproductive steroids in fish. *Rev Fish Biol Fish*, **3**, 160-180.

Kimmel, C.B., Ballard, W.W., Kimmel, S.R., Ullmann, B. and Schilling, T.F. (1995) Stages of embryonic development of the zebrafish. *Dev Dyn*, **203**, 253-310.

Kishida, M. and Callard, G.V. (2001) Distinct cytochrome P450 aromatase isoforms in zebrafish (Danio rerio) brain and ovary are differentially programmed and estrogen regulated during early development. *Endocrinology*, **142**, 740-750.

Kishida, M., McLellan, M., Miranda, J.A. and Callard, G.V. (2001) Estrogen and xenoestrogens upregulate the brain aromatase isoform (P450aromB) and perturb markers of early development in zebrafish (Danio rerio). *Comp Biochem Physiol B Biochem Mol Biol*, **129**, 261-268.

Koh, E., Noda, T., Kanaya, J. and Namiki, M. (2002) Differential expression of 17beta-hydroxysteroid dehydrogenase isozyme genes in prostate cancer and noncancer tissues. *Prostate*, **53**, 154-159.

Kousteni, S., Bellido, T., Plotkin, L.I., O'Brien, C.A., Bodenner, D.L., Han, L., Han, K., DiGregorio, G.B., Katzenellenbogen, J.A., Katzenellenbogen, B.S., Roberson, P.K., Weinstein, R.S., Jilka, R.L. and Manolagas, S.C. (2001) Nongenotropic, sex-nonspecific signaling through the estrogen or androgen receptors: dissociation from transcriptional activity. *Cell*, **104**, 719-730.

Krazeisen, A., Breitling, R., Moller, G. and Adamski, J. (2001) Phytoestrogens inhibit human 17beta-hydroxysteroid dehydrogenase type 5. *Mol Cell Endocrinol*, **171**, 151-162.

Kristan, K., Rizner, T.L., Stojan, J., Gerber, J.K., Kremmer, E. and Adamski, J. (2003) Significance of individual amino acid residues for coenzyme and substrate specificity of 17beta-hydroxysteroid dehydrogenase from the fungus Cochliobolus lunatus. *Chem Biol Interact*, **143-144**, 493-501.

Krozowski, Z. (1994) The short-chain alcohol dehydrogenase superfamily: variations on a common theme. *J Steroid Biochem Mol Biol*, **51**, 125-130.

Kumar, S., Tamura, K., Jakobsen, I.B. and Nei, M. (2001) MEGA2: molecular evolutionary genetics analysis software. *Bioinformatics*, **17**, 1244-1245.

Kuntz, S., Chardard, D., Chesnel, A., Grillier-Vuissoz, I. and Flament, S. (2003) Steroids, aromatase and sex differentiation of the newt Pleurodeles waltl. *Cytogenet Genome Res*, **101**, 283-288.

Kusakabe, M., Kobayashi, T., Todo, T., Mark Lokman, P., Nagahama, Y. and Young, G. (2002) Molecular cloning and expression during spermatogenesis of a cDNA encoding testicular 11beta-hydroxylase (P45011beta) in rainbow trout (Oncorhynchus mykiss). *Mol Reprod Dev*, **62**, 456-469.

Labrie, F. (1991) Intracrinology. *Mol Cell Endocrinol*, **78**, C113-118.

Labrie, F., Luu-The, V., Lin, S.X., Labrie, C., Simard, J., Breton, R. and Belanger, A. (1997) The key role of 17 beta-hydroxysteroid dehydrogenases in sex steroid biology. *Steroids*, **62**, 148-158.

Lai, W.W., Hsiao, P.H., Guiguen, Y. and Chung, B.C. (1998) Cloning of zebrafish cDNA for 3beta-hydroxysteroid dehydrogenase and P450scc. *Endocr Res*, **24**, 927-931.

Lanisnik Rizner, T., Stojan, J. and Adamski, J. (2001a) 17beta-hydroxysteroid dehydrogenase from the fungus Cochliobolus lunatus: structural and functional aspects. *Chem Biol Interact*, **130-132**, 793-803.

Lanisnik Rizner, T., Stojan, J. and Adamski, J. (2001b) Searching for the physiological function of 17beta-hydroxysteroid dehydrogenase from the fungus Cochliobolus lunatus: studies of substrate specificity and expression analysis. *Mol Cell Endocrinol*, **171**, 193-198.

Laubner, D., Breitling, R. and Adamski, J. (2003) Embryonic expression of cholesterogenic genes is restricted to distinct domains and colocalizes with apoptotic regions in mice. *Brain Res Mol Brain Res*, **115**, 87-92.

Laudet, V. (1997) Evolution of the nuclear receptor superfamily: early diversification from an ancestral orphan receptor. *J Mol Endocrinol*, **19**, 207-226.

Le Bail, J.C., Champavier, Y., Chulia, A.J. and Habrioux, G. (2000) Effects of phytoestrogens on aromatase, 3beta and 17beta-hydroxysteroid dehydrogenase activities and human breast cancer cells. *Life Sci*, **66**, 1281-1291.

Le Curieux-Belfond, O., Moslemi, S., Mathieu, M. and Seralini, G.E. (2001) Androgen metabolism in oyster Crassostrea gigas: evidence for 17beta-HSD activities and characterization of an aromatase-like activity inhibited by pharmacological compounds and a marine pollutant. *J Steroid Biochem Mol Biol*, **78**, 359-366.

Le Guellec, D., Thiard, M.C., Remy-Martin, J.P., Deray, A., Gomot, L. and Adessi, G.L. (1987) In vitro metabolism of androstenedione and identification of endogenous steroids in Helix aspersa. *Gen Comp Endocrinol*, **66**, 425-433.

Leenders, F., Tesdorpf, J.G., Markus, M., Engel, T., Seedorf, U. and Adamski, J. (1996) Porcine 80-kDa protein reveals intrinsic 17 beta-hydroxysteroid dehydrogenase, fatty acyl-CoA-hydratase/dehydrogenase, and sterol transfer activities. *J Biol Chem*, **271**, 5438-5442.

Levin, E.R. (1999) Cellular Functions of the Plasma Membrane Estrogen Receptor. *Trends Endocrinol Metab*, **10**, 374-377.

Lieberherr, M. and Grosse, B. (1994) Androgens increase intracellular calcium concentration and inositol 1,4,5-trisphosphate and diacylglycerol formation via a pertussis toxin-sensitive G-protein. *J Biol Chem*, **269**, 7217-7223.

Lin, B., White, J.T., Ferguson, C., Wang, S., Vessella, R., Bumgarner, R., True, L.D., Hood, L. and Nelson, P.S. (2001) Prostate short-chain dehydrogenase reductase 1 (PSDR1): a new member of the short-chain steroid dehydrogenase/reductase family highly expressed in normal and neoplastic prostate epithelium. *Cancer Res*, **61**, 1611-1618.

Lin, H.K., Jez, J.M., Schlegel, B.P., Peehl, D.M., Pachter, J.A. and Penning, T.M. (1997) Expression and characterization of recombinant type 2 3 alpha-hydroxysteroid dehydrogenase (HSD) from human prostate: demonstration of bifunctional 3 alpha/17 beta-HSD activity and cellular distribution. *Mol Endocrinol*, **11**, 1971-1984.

Lin, S.X., Han, Q., Azzi, A., Zhu, D.W., Gangloff, A., Campbell, R.L. and Gongloff, A. (1999) 3D-structure of human estrogenic 17beta-HSD1: binding with various steroids. *J Steroid Biochem Mol Biol*, **69**, 425-429.

Lindqvist, A., Hughes, I.A. and Andersson, S. (2001) Substitution mutation C268Y causes 17 beta-hydroxysteroid dehydrogenase 3 deficiency. *J Clin Endocrinol Metab*, **86**, 921-923.

Loir, M. (1999) Spermatogonia of rainbow trout: II. in vitro study of the influence of pituitary hormones, growth factors and steroids on mitotic activity. *Mol Reprod Dev*, **53**, 434-442.

Lowartz, S., Petkam, R., Renaud, R., Beamish, F.W., Kime, D.E., Raeside, J. and Leatherland, J.F. (2003) Blood steroid profile and in vitro steroidogenesis by ovarian follicles and testis fragments of adult sea lamprey, Petromyzon marinus. *Comp Biochem Physiol A Mol Integr Physiol*, **134**, 365-376.

Luu The, V., Labrie, C., Zhao, H.F., Couet, J., Lachance, Y., Simard, J., Leblanc, G., Cote, J., Berube, D., Gagne, R. and et al. (1989) Characterization of cDNAs for human estradiol 17 beta-dehydrogenase and assignment of the gene to chromosome 17: evidence of two mRNA species with distinct 5'-termini in human placenta. *Mol Endocrinol*, **3**, 1301-1309.

Luu-The, V., Dufort, I., Pelletier, G. and Labrie, F. (2001) Type 5 17beta-hydroxysteroid dehydrogenase: its role in the formation of androgens in women. *Mol Cell Endocrinol*, **171**, 77-82.

Luu-The, V., Labrie, C., Simard, J., Lachance, Y., Zhao, H.F., Couet, J., Leblanc, G. and Labrie, F. (1990) Structure of two in tandem human 17 beta-hydroxysteroid dehydrogenase genes. *Mol Endocrinol*, **4**, 268-275.

Luu-The, V., Zhang, Y., Poirier, D. and Labrie, F. (1995) Characteristics of human types 1, 2 and 3 17 beta-hydroxysteroid dehydrogenase activities: oxidation/reduction and inhibition. *J Steroid Biochem Mol Biol*, **55**, 581-587.

Mahajan, R., Gerace, L. and Melchior, F. (1998) Molecular characterization of the SUMO-1 modification of RanGAP1 and its role in nuclear envelope association. *J Cell Biol*, **140**, 259-270.

Manire, C.A., Rasmussen, L.E. and Gross, T.S. (1999) Serum steroid hormones including 11-ketotestosterone, 11-ketoandrostenedione, and dihydroprogesterone in juvenile and adult bonnethead sharks, Sphyrna tiburo. *J Exp Zool*, **284**, 595-603.

Marcus, P.I. and Talalay, P. (1956) Induction and purification of alpha- and beta-hydroxysteroid dehydrogenases. *J Biol Chem*, **218**, 661-674.

Margiotti, K., Kim, E., Pearce, C.L., Spera, E., Novelli, G. and Reichardt, J.K. (2002) Association of the G289S single nucleotide polymorphism in the HSD17B3 gene with prostate cancer in Italian men. *Prostate*, **53**, 65-68.

Martel, C., Rheaume, E., Takahashi, M., Trudel, C., Couet, J., Luu-The, V., Simard, J. and Labrie, F. (1992) Distribution of 17 beta-hydroxysteroid dehydrogenase gene expression and activity in rat and human tissues. *J Steroid Biochem Mol Biol*, **41**, 597-603.

Matsumoto, T., Takeyama, K., Sato, T. and Kato, S. (2003) Androgen receptor functions from reverse genetic models. *J Steroid Biochem Mol Biol*, **85**, 95-99.

McKeever, B.M., Hawkins, B.K., Geissler, W.M., Wu, L., Sheridan, R.P., Mosley, R.T. and Andersson, S. (2002) Amino acid substitution of arginine 80 in 17beta-hydroxysteroid dehydrogenase type 3 and its effect on NADPH cofactor binding and oxidation/reduction kinetics. *Biochim Biophys Acta*, **1601**, 29-37.

Mendonca, B.B., Arnhold, I.J., Bloise, W., Andersson, S., Russell, D.W. and Wilson, J.D. (1999) 17Beta-hydroxysteroid dehydrogenase 3 deficiency in women. *J Clin Endocrinol Metab*, **84**, 802-804.

Mendoza-Hernandez, G., Calcagno, M., Sanchez-Nuncio, H.R. and Diaz-Zagoya, J.C. (1984) Dehydroepiandrosterone is a substrate for estradiol 17 beta-dehydrogenase from human placenta. *Biochem Biophys Res Commun*, **119**, 83-87.

Mindnich, R., Moller, G. and Adamski, J. (2004) The role of 17 beta-hydroxysteroid dehydrogenases. *Mol Cell Endocrinol*, **218**, 7-20.

Miura, T., Yamauchi, K., Takahashi, H. and Nagahama, Y. (1991) Hormonal induction of all stages of spermatogenesis in vitro in the male Japanese eel (Anguilla japonica). *Proc Natl Acad Sci U S A*, **88**, 5774-5778.

Miyoshi, Y., Ando, A., Shiba, E., Taguchi, T., Tamaki, Y. and Noguchi, S. (2001) Involvement of up-regulation of 17beta-hydroxysteroid dehydrogenase type 1 in maintenance of intratumoral high estradiol levels in postmenopausal breast cancers. *Int J Cancer*, **94**, 685-689.

Moghrabi, N., Hughes, I.A., Dunaif, A. and Andersson, S. (1998) Deleterious missense mutations and silent polymorphism in the human 17beta-hydroxysteroid dehydrogenase 3 gene (HSD17B3). *J Clin Endocrinol Metab*, **83**, 2855-2860.

Monjo, M., Rodriguez, A.M., Palou, A. and Roca, P. (2003) Direct effects of testosterone, 17 beta-estradiol, and progesterone on adrenergic regulation in cultured brown adipocytes: potential mechanism for gender-dependent thermogenesis. *Endocrinology*, **144**, 4923-4930.

Moon, Y.A. and Horton, J.D. (2003) Identification of two mammalian reductases involved in the two-carbon fatty acyl elongation cascade. *J Biol Chem*, **278**, 7335-7343.

Murono, E.P. and Payne, A.H. (1976) Distinct testicular 17-ketosteroid reductases, one in interstitial tissue and one in seminiferous tubules. Differential modulation by testosterone and metabolites of testosterone. *Biochim Biophys Acta*, **450**, 89-100.

Mustonen, M.V., Poutanen, M.H., Isomaa, V.V., Vihko, P.T. and Vihko, R.K. (1997) Cloning of mouse 17beta-hydroxysteroid dehydrogenase type 2, and analysing expression of the mRNAs for types 1, 2, 3, 4 and 5 in mouse embryos and adult tissues. *Biochem J*, **325 (Pt 1)**, 199-205.

Nahoum, V., Gangloff, A., Shi, R. and Lin, S.X. (2003) How estrogen-specific proteins discriminate estrogens from androgens: a common steroid binding site architecture. *Faseb J*, **17**, 1334-1336.

Nakabayashi, O., Kikuchi, H., Kikuchi, T. and Mizuno, S. (1998) Differential expression of genes for aromatase and estrogen receptor during the gonadal development in chicken embryos. *J Mol Endocrinol*, **20**, 193-202.

Nemoto, Y., Toda, K., Ono, M., Fujikawa-Adachi, K., Saibara, T., Onishi, S., Enzan, H., Okada, T. and Shizuta, Y. (2000) Altered expression of fatty acid-metabolizing enzymes in aromatase-deficient mice. *J Clin Invest*, **105**, 1819-1825.

Nieuwkoop, P.D., Sutasurya, L.A. (1979) Primordial germ cells in the chordates. London: Cambridge University Press.

Nokelainen, P., Peltoketo, H., Mustonen, M. and Vihko, P. (2000) Expression of mouse 17beta-hydroxysteroid dehydrogenase/17-ketosteroid reductase type 7 in the ovary, uterus, and placenta: localization from implantation to late pregnancy. *Endocrinology*, **141**, 772-778.

Nokelainen, P., Peltoketo, H., Vihko, R. and Vihko, P. (1998) Expression cloning of a novel estrogenic mouse 17 beta-hydroxysteroid dehydrogenase/17-ketosteroid reductase (m17HSD7), previously described as a prolactin receptor-associated protein (PRAP) in rat. *Mol Endocrinol*, **12**, 1048-1059.

Nokelainen, P., Puranen, T., Peltoketo, H., Orava, M., Vihko, P. and Vihko, R. (1996) Molecular cloning of mouse 17 beta-hydroxysteroid dehydrogenase type 1 and characterization of enzyme activity. *Eur J Biochem*, **236**, 482-490.

Nomura, O., Nakabayashi, O., Nishimori, K., Yasue, H. and Mizuno, S. (1999) Expression of five steroidogenic genes including aromatase gene at early developmental stages of chicken male and female embryos. *J Steroid Biochem Mol Biol*, **71**, 103-109.

Ohashi, M., Mizushima, N., Kabeya, Y. and Yoshimori, T. (2003) Localization of mammalian NAD(P)H steroid dehydrogenase-like protein on lipid droplets. *J Biol Chem*, **278**, 36819-36829.

Oliveira, R.F., Hirschenhauser, K., Carneiro, L.A. and Canario, A.V. (2002) Social modulation of androgen levels in male teleost fish. *Comp Biochem Physiol B Biochem Mol Biol*, **132**, 203-215.

Olsen, L.C., Aasland, R. and Fjose, A. (1997) A vasa-like gene in zebrafish identifies putative primordial germ cells. *Mech Dev*, **66**, 95-105.

Ong, O.C., Ota, I.M., Clarke, S. and Fung, B.K. (1989) The membrane binding domain of rod cGMP phosphodiesterase is posttranslationally modified by methyl esterification at a C-terminal cysteine. *Proc Natl Acad Sci U S A*, **86**, 9238-9242.

Oppermann, U., Filling, C., Hult, M., Shafqat, N., Wu, X., Lindh, M., Shafqat, J., Nordling, E., Kallberg, Y., Persson, B. and Jornvall, H. (2003) Short-chain dehydrogenases/reductases (SDR): the 2002 update. *Chem Biol Interact*, **143-144**, 247-253.

Oppermann, U., Salim, S., Hult, M., Eissner, G. and Jornvall, H. (1999) Regulatory factors and motifs in SDR enzymes. *Adv Exp Med Biol*, **463**, 365-371.

Oppermann, U.C., Belai, I. and Maser, E. (1996) Antibiotic resistance and enhanced insecticide catabolism as consequences of steroid induction in the gram-negative bacterium Comamonas testosteroni. *J Steroid Biochem Mol Biol*, **58**, 217-223.

Orn, S., Holbech, H., Madsen, T.H., Norrgren, L. and Petersen, G.I. (2003) Gonad development and vitellogenin production in zebrafish (Danio rerio) exposed to ethinylestradiol and methyltestosterone. *Aquat Toxicol*, **65**, 397-411.

Osada, M. and Nomura, T. (1990) The levels of prostaglandins associated with the reproductive cycle of the scallop, Patinopecten yessoensis. *Prostaglandins*, **40**, 229-239.

Osada, M., Takamura, T., Sato, H. and Mori, K. (2003) Vitellogenin synthesis in the ovary of scallop, Patinopecten yessoensis: Control by estradiol-17 beta and the central nervous system. *J Exp Zoolog Part A Comp Exp Biol*, **299**, 172-179.

O'Shaughnessy, P.J., Baker, P.J., Heikkila, M., Vainio, S. and McMahon, A.P. (2000) Localization of 17beta-hydroxysteroid dehydrogenase/17-ketosteroid reductase isoform expression in the developing mouse testis--androstenedione is the major androgen secreted by fetal/neonatal leydig cells. *Endocrinology*, **141**, 2631-2637.

Oz, O.K., Zerwekh, J.E., Fisher, C., Graves, K., Nanu, L., Millsaps, R. and Simpson, E.R. (2000) Bone has a sexually dimorphic response to aromatase deficiency. *J Bone Miner Res*, **15**, 507-514.

Pall, M.K., Mayer, I. and Borg, B. (2002) Androgen and behavior in the male three-spined stickleback, Gasterosteus aculeatus I.--changes in 11-ketotestosterone levels during the nesting cycle. *Horm Behav*, **41**, 377-383.

Paolucci, M., Di Cristo, C. and Di Cosmo, A. (2002) Immunological evidence for progesterone and estradiol receptors in the freshwater crayfish Austropotamobius pallipes. *Mol Reprod Dev*, **63**, 55-62.

Pearson, W.R., Wood, T., Zhang, Z. and Miller, W. (1997) Comparison of DNA sequences with protein sequences. *Genomics*, **46**, 24-36.

Pelletier, G., Luu-The, V., Tetu, B. and Labrie, F. (1999) Immunocytochemical localization of type 5 17beta-hydroxysteroid dehydrogenase in human reproductive tissues. *J Histochem Cytochem*, **47**, 731-738.

Peltoketo, H., Isomaa, V., Maentausta, O. and Vihko, R. (1988) Complete amino acid sequence of human placental 17 beta-hydroxysteroid dehydrogenase deduced from cDNA. *FEBS Lett*, **239**, 73-77.

Peltoketo, H., Luu-The, V., Simard, J. and Adamski, J. (1999) 17beta-hydroxysteroid dehydrogenase (HSD)/17-ketosteroid reductase (KSR) family; nomenclature and main characteristics of the 17HSD/KSR enzymes. *J Mol Endocrinol*, **23**, 1-11.

Penning, T.M. (1997) Molecular endocrinology of hydroxysteroid dehydrogenases. *Endocr Rev*, **18**, 281-305.

Pernet, V. and Anctil, M. (2002) Annual variations and sex-related differences of estradiol-17beta levels in the anthozoan Renilla koellikeri. *Gen Comp Endocrinol*, **129**, 63-68.

Persson, B., Kallberg, Y., Oppermann, U. and Jornvall, H. (2003) Coenzyme-based functional assignments of short-chain dehydrogenases/reductases (SDRs). *Chem Biol Interact*, **143-144**, 271-278.

Piao, Y.S., Peltoketo, H., Oikarinen, J. and Vihko, R. (1995) Coordination of transcription of the human 17 beta-hydroxysteroid dehydrogenase type 1 gene (EDH17B2) by a cell-specific enhancer and a silencer: identification of a retinoic acid response element. *Mol Endocrinol*, **9**, 1633-1644.

Piao, Y.S., Peltoketo, H., Vihko, P. and Vihko, R. (1997) The proximal promoter region of the gene encoding human 17beta-hydroxysteroid dehydrogenase type 1 contains GATA, AP-2, and Sp1 response elements: analysis of promoter function in choriocarcinoma cells. *Endocrinology*, **138**, 3417-3425.

Pietras, R.J. and Szego, C.M. (1977) Specific binding sites for oestrogen at the outer surfaces of isolated endometrial cells. *Nature*, **265**, 69-72.

Poirier, D. (2003) Inhibitors of 17 beta-hydroxysteroid dehydrogenases. *Curr Med Chem*, **10**, 453-477.

Pontius, J., Richelle, J. and Wodak, S.J. (1996) Deviations from standard atomic volumes as a quality measure for protein crystal structures. *J Mol Biol*, **264**, 121-136.

Poukka, H., Karvonen, U., Janne, O.A. and Palvimo, J.J. (2000) Covalent modification of the androgen receptor by small ubiquitin-like modifier 1 (SUMO-1). *Proc Natl Acad Sci U S A*, **97**, 14145-14150.

Poulin, R. and Labrie, F. (1986) Stimulation of cell proliferation and estrogenic response by adrenal C19-delta 5-steroids in the ZR-75-1 human breast cancer cell line. *Cancer Res*, **46**, 4933-4937.

Puntervoll, P., Linding, R., Gemund, C., Chabanis-Davidson, S., Mattingsdal, M., Cameron, S., Martin, D.M., Ausiello, G., Brannetti, B., Costantini, A., Ferre, F., Maselli, V., Via, A., Cesareni, G., Diella, F., Superti-Furga, G., Wyrwicz, L., Ramu, C., McGuigan, C., Gudavalli, R., Letunic, I., Bork, P., Rychlewski, L., Kuster, B., Helmer-Citterich, M., Hunter, W.N., Aasland, R. and Gibson, T.J. (2003) ELM server: A new resource for investigating short functional sites in modular eukaryotic proteins. *Nucleic Acids Res*, **31**, 3625-3630.

Puranen, T., Poutanen, M., Ghosh, D., Vihko, P. and Vihko, R. (1997a) Characterization of structural and functional properties of human 17 beta-hydroxysteroid dehydrogenase type 1 using recombinant enzymes and site-directed mutagenesis. *Mol Endocrinol*, **11**, 77-86.

Puranen, T., Poutanen, M., Ghosh, D., Vihko, R. and Vihko, P. (1997b) Origin of substrate specificity of human and rat 17beta-hydroxysteroid dehydrogenase type 1, using chimeric enzymes and site-directed substitutions. *Endocrinology*, **138**, 3532-3539.

Qin, K.N. and Rosenfield, R.L. (2000) Expression of 17 beta-hydroxysteroid dehydrogenase type 5 in human ovary: a pilot study. *J Soc Gynecol Investig*, **7**, 61-64.

Rattner, A., Smallwood, P.M. and Nathans, J. (2000) Identification and characterization of all-trans-retinol dehydrogenase from photoreceptor outer segments, the visual cycle enzyme that reduces all-trans-retinal to all-trans-retinol. *J Biol Chem*, **275**, 11034-11043.

Razandi, M., Oh, P., Pedram, A., Schnitzer, J. and Levin, E.R. (2002) ERs associate with and regulate the production of caveolin: implications for signaling and cellular actions. *Mol Endocrinol*, **16**, 100-115.

Razandi, M., Pedram, A., Greene, G.L. and Levin, E.R. (1999) Cell membrane and nuclear estrogen receptors (ERs) originate from a single transcript: studies of ERalpha and ERbeta expressed in Chinese hamster ovary cells. *Mol Endocrinol*, **13**, 307-319.

Redding, J.M., Patino, R. (1993) Reproductive physiology. *The Physiology of Fishes*, ed. D.H. Evans, pp. 503-534. CRC Press, Boca Raton, FL.

Rheault, P., Charbonneau, A. and Luu-The, V. (1999) Structure and activity of the murine type 5 17beta-hydroxysteroid dehydrogenase gene(1). *Biochim Biophys Acta*, **1447**, 17-24.

Rizner, T.L., Adamski, J. and Stojan, J. (2000) 17Beta-hydroxysteroid dehydrogenase from Cochliobolus lunatus: model structure and substrate specificity. *Arch Biochem Biophys*, **384**, 255-262.

Rizner, T.L., Adamski, J. and Zakelj-Mavric, M. (2001) Expression of 17beta-hydroxysteroid dehydrogenases in mesophilic and extremophilic yeast. *Steroids*, **66**, 49-54.

Rizner, T.L., Zakelj-Mavric, M., Plemenitas, A. and Zorko, M. (1996) Purification and characterization of 17beta-hydroxysteroid dehydrogenase from the filamentous fungus Cochliobolus lunatus. *J Steroid Biochem Mol Biol*, **59**, 205-214.

Robertson, K.M., O'Donnell, L., Jones, M.E., Meachem, S.J., Boon, W.C., Fisher, C.R., Graves, K.H., McLachlan, R.I. and Simpson, E.R. (1999) Impairment of spermatogenesis in mice lacking a functional aromatase (cyp 19) gene. *Proc Natl Acad Sci U S A*, **96**, 7986-7991.

Rodriguez, M.S., Dargemont, C. and Hay, R.T. (2001) SUMO-1 conjugation in vivo requires both a consensus modification motif and nuclear targeting. *J Biol Chem*, **276**, 12654-12659.

Romanoff, A.L. (1960). *The avian embryo: structrual and functional development*. pp816-853. NewYork MacMillan Ltd.

Rosler, A., Silverstein, S. and Abeliovich, D. (1996) A (R80Q) mutation in 17 beta-hydroxysteroid dehydrogenase type 3 gene among Arabs of Israel is associated with pseudohermaphroditism in males and normal asymptomatic females. *J Clin Endocrinol Metab*, **81**, 1827-1831.

Sakai, N., Tanaka, M., Adachi, S., Miller, W.L. and Nagahama, Y. (1992) Rainbow trout cytochrome P-450c17 (17 alpha-hydroxylase/17,20-lyase). cDNA cloning, enzymatic properties and temporal pattern of ovarian P-450c17 mRNA expression during oogenesis. *FEBS Lett*, **301**, 60-64.

Sakai, N., Tanaka, M., Takahashi, M., Fukada, S., Mason, J.I. and Nagahama, Y. (1994) Ovarian 3 beta-hydroxysteroid dehydrogenase/delta 5-4-isomerase of rainbow trout: its cDNA cloning and properties of the enzyme expressed in a mammalian cell. *FEBS Lett*, **350**, 309-313.

Schoenmarkers, H.J. and Voogt, P.A. (1980) In vitro biosynthesis of steroids from progesterone by the ovaries and pyloric ceca of the starfish Asterias rubens. *Gen Comp Endocrinol*, **41**, 408-416.

Schwede, T., Diemand, A., Guex, N. and Peitsch, M.C. (2000) Protein structure computing in the genomic era. *Res Microbiol*, **151**, 107-112.

Scordalakes, E.M., Imwalle, D.B. and Rissman, E.F. (2002) Oestrogen's masculine side: mediation of mating in male mice. *Reproduction*, **124**, 331-338.

Sha, J., Baker, P. and O'Shaughnessy, P.J. (1996) Both reductive forms of 17 beta-hydroxysteroid dehydrogenase (types 1 and 3) are expressed during development in the mouse testis. *Biochem Biophys Res Commun*, **222**, 90-94.

Sha, J.A., Dudley, K., Rajapaksha, W.R. and O'Shaughnessy, P.J. (1997) Sequence of mouse 17beta-hydroxysteroid dehydrogenase type 3 cDNA and tissue distribution of the type 1 and type 3 isoform mRNAs. *J Steroid Biochem Mol Biol*, **60**, 19-24.

Song, M.S., Chen, W., Zhang, M. and Napoli, J.L. (2003) Identification of a mouse short-chain dehydrogenase/reductase gene, retinol dehydrogenase-similar. Function of non-catalytic amino acid residues in enzyme activity. *J Biol Chem*, **278**, 40079-40087.

Sperry, T.S. and Thomas, P. (1999a) Characterization of two nuclear androgen receptors in Atlantic croaker: comparison of their biochemical properties and binding specificities. *Endocrinology*, **140**, 1602-1611.

Sperry, T.S. and Thomas, P. (1999b) Identification of two nuclear androgen receptors in kelp bass (Paralabrax clathratus) and their binding affinities for xenobiotics: comparison with Atlantic croaker (Micropogonias undulatus) androgen receptors. *Biol Reprod*, **61**, 1152-1161.

Stocker, U. and van Gunsteren, W.F. (2000) Molecular dynamics simulation of hen egg white lysozyme: a test of the GROMOS96 force field against nuclear magnetic resonance data. *Proteins*, **40**, 145-153.

Stoffel-Wagner, B., Watzka, M., Steckelbroeck, S., Schramm, J., Bidlingmaier, J.F. and Klingmuller, D. (1999) Expression of 17beta-hydroxysteroid dehydrogenase types 1, 2, 3 and 4 in the human temporal lobe. *J Endocrinol*, **160**, 119-126.

Su, J., Lin, M. and Napoli, J.L. (1999) Complementary deoxyribonucleic acid cloning and enzymatic characterization of a novel 17beta/3alpha-hydroxysteroid/retinoid short chain dehydrogenase/reductase. *Endocrinology*, **140**, 5275-5284.

Suzuki, T., Moriya, T., Ishida, T., Kimura, M., Ohuchi, N. and Sasano, H. (2002) In situ production of estrogens in human breast carcinoma. *Breast Cancer*, **9**, 296-302.

Takahashi, H. (1977) Juvenile hermaphroditism in the zebrafish *Brachydanio rerio*. *Bulletin Fac Fish Hokkaido University*, **28**, 57-65.

Takeo, J. and Yamashita, S. (1999) Two distinct isoforms of cDNA encoding rainbow trout androgen receptors. *J Biol Chem*, **274**, 5674-5680.

Takeyama, J., Suzuki, T., Hirasawa, G., Muramatsu, Y., Nagura, H., Iinuma, K., Nakamura, J., Kimura, K.I., Yoshihama, M., Harada, N., Andersson, S. and Sasano, H. (2000) 17beta-hydroxysteroid dehydrogenase type 1 and 2 expression in the human fetus. *J Clin Endocrinol Metab*, **85**, 410-416.

Talalay, P., Dobson, M.M. and Tapley, D.F. (1952) Oxidative degradation of testosterone by adaptive enzymes. *Nature*, **170**, 620-621.

Tarrant, A.M., Atkinson, S. and Atkinson, M.J. (1999) Estrone and estradiol-17 beta concentration in tissue of the scleractinian coral, Montipora verrucosa. *Comp Biochem Physiol A Mol Integr Physiol*, **122**, 85-92.

Temple, J.L., Scordalakes, E.M., Bodo, C., Gustafsson, J.A. and Rissman, E.F. (2003) Lack of functional estrogen receptor beta gene disrupts pubertal male sexual behavior. *Horm Behav*, **44**, 427-434.

Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*, **22**, 4673-4680.

Thornton, J.W. (2001) Evolution of vertebrate steroid receptors from an ancestral estrogen receptor by ligand exploitation and serial genome expansions. *Proc Natl Acad Sci U S A*, **98**, 5671-5676.

Tong, S.K., Chiang, E.F., Hsiao, P.H. and Chung, B. (2001) Phylogeny, expression and enzyme activity of zebrafish cyp19 (P450 aromatase) genes. *J Steroid Biochem Mol Biol*, **79**, 299-303.

Trant, J.M. (1995) Isolation and characterization of the cDNA encoding the spiny dogfish shark (Squalus acanthias) form of cytochrome P450c17. *J Exp Zool*, **272**, 25-33.

Tsai-Morris, C.H., Khanum, A., Tang, P.Z. and Dufau, M.L. (1999) The rat 17beta-hydroxysteroid dehydrogenase type III: molecular cloning and gonadotropin regulation. *Endocrinology*, **140**, 3534-3542.

Twan, W.H., Hwang, J.S. and Chang, C.F. (2003) Sex steroids in scleractinian coral, Euphyllia ancora: implication in mass spawning. *Biol Reprod*, **68**, 2255-2260.

Uchida, D., Yamashita, M., Kitano, T. and Iguchi, T. (2002) Oocyte apoptosis during the transition from ovary-like tissue to testes during sex differentiation of juvenile zebrafish. *J Exp Biol*, **205**, 711-718.

Vihko, P., Harkonen, P., Oduwole, O., Torn, S., Kurkela, R., Porvari, K., Pulkka, A. and Isomaa, V. (2002) 17 beta-hydroxysteroid dehydrogenases and cancers. *J Steroid Biochem Mol Biol*, **83**, 119-122.

Wahli, W. (1988) Evolution and expression of vitellogenin genes. *Trends Genet*, **4**, 227-232.

Wajima, Y., Furusawa, T., Kawauchi, S., Wakabayashi, N., Nakabayashi, O., Nishimori, K. and Mizuno, S. (1999) The cDNA cloning and transient expression of an ovary-specific 17beta-hydroxysteroid dehydrogenase of chickens. *Gene*, **233**, 75-82.

Warrier, S.R., Tirumalai, R. and Subramoniam, T. (2001) Occurrence of vertebrate steroids, estradiol 17beta and progesterone in the reproducing females of the mud crab Scylla serrata. *Comp Biochem Physiol A Mol Integr Physiol*, **130**, 283-294.

Wasson, K.M., Gower, B.A., Hines, G.A. and Watts, S.A. (2000) Levels of progesterone, testosterone, and estradiol, and androstenedione metabolism in the gonads of Lytechinus variegatus (Echinodermata:echinoidea). *Comp Biochem Physiol C Toxicol Pharmacol*, **126**, 153-165.

Wasson, K.M., Hines, G.A. and Watts, S.A. (1998) Synthesis of testosterone and 5alpha-androstanediols during nutritionally stimulated gonadal growth in Lytechinus variegatus lamarck (Echinodermata:Echinoidea). *Gen Comp Endocrinol*, **111**, 197-206.

Weihua, Z., Lathe, R., Warner, M. and Gustafsson, J.A. (2002) An endocrine pathway in the prostate, ERbeta, AR, 5alpha-androstane-3beta,17beta-diol, and CYP7B1, regulates prostate growth. *Proc Natl Acad Sci U S A*, **99**, 13589-13594.

Willumsen, B.M., Norris, K., Papageorge, A.G., Hubbert, N.L. and Lowy, D.R. (1984) Harvey murine sarcoma virus p21 ras protein: biological and biochemical significance of the cysteine nearest the carboxy terminus. *Embo J*, **3**, 2581-2585.

Woodhouse, L.J., Gupta, N., Bhasin, M., Singh, A.B., Ross, R., Phillips, J. and Bhasin, S. (2004) Dose-dependent effects of testosterone on regional adipose tissue distribution in healthy young men. *J Clin Endocrinol Metab*, **89**, 718-726.

Xia, Z., Patino, R., Gale, W.L., Maule, A.G. and Densmore, L.D. (1999) Cloning, in vitro expression, and novel phylogenetic classification of a channel catfish estrogen receptor. *Gen Comp Endocrinol*, **113**, 360-368.

Yoon, C., Kawakami, K. and Hopkins, N. (1997) Zebrafish vasa homologue RNA is localized to the cleavage planes of 2- and 4-cell-stage embryos and is expressed in the primordial germ cells. *Development*, **124**, 3157-3165.

Zhang, H., Saitoh, H. and Matunis, M.J. (2002) Enzymes of the SUMO modification pathway localize to filaments of the nuclear pore complex. *Mol Cell Biol*, **22**, 6498-6508.

Zhang, Y., Word, R.A., Fesmire, S., Carr, B.R. and Rainey, W.E. (1996) Human ovarian expression of 17 beta-hydroxysteroid dehydrogenase types 1, 2, and 3. *J Clin Endocrinol Metab*, **81**, 3594-3598.

Zhu, Y., Bond, J. and Thomas, P. (2003) Identification, classification, and partial characterization of genes in humans and other vertebrates homologous to a fish membrane progestin receptor. *Proc Natl Acad Sci U S A*, **100**, 2237-2242.

Zhuang, R., Lin, M. and Napoli, J.L. (2002) cis-Retinol/androgen dehydrogenase, isozyme 3 (CRAD3): a short-chain dehydrogenase active in a reconstituted path of 9-cis-retinoic acid biosynthesis in intact cells. *Biochemistry*, **41**, 3477-3483.

7 Appendix

7.1 Publications

7.1.1 Scientific papers

Mindnich, R., Deluca, D. and Adamski, J. (2004). Identification and characterization of 17 beta-hydroxysteroid dehydrogenases in the zebrafish, Danio rerio. *Mol Cell Endocrinol*. **215**,19-30.

Mindnich, R., Moller, G. and Adamski, J. (2004). The role of 17 beta-hydroxysteroid dehydrogenases. *Mol Cell Endocrinol*. **218**, 7-20.

Deluca, D., Fritz, A., **Mindnich, R.**, Möller, G. and Adamski, J. (2004). Biochemical genetics of 17beta-hydroxysteroid dehydrogenases. *Curr. Topics Steroid Res.* (in press)

7.1.2 Sequences

Proteins:

```
NP_957082
hypothetical protein MGC73320 [Danio rerio]
gi|41387186|ref|NP_957082.1|[41387186]

AAP74566
photoreceptor associated retinol dehydrogenase type 2 [Danio rerio]
gi|32250711|gb|AAP74566.1|[32250711]

AAP74565
photoreceptor associated retinol dehydrogenase type 1 [Danio rerio]
gi|32250709|gb|AAP74565.1|[32250709]

AAP74564
17-beta hydroxysteroid dehydrogenase [Danio rerio]
gi|32250707|gb|AAP74564.1|[32250707]

AAS58452
17-beta hydroxysteroid dehydrogenase type 12A, 3-ketoacyl-CoA reductase type A
[Danio rerio]
gi|45356826|gb|AAS58452.1|[45356826]

AAS58451
17-beta hydroxysteroid dehydrogenase type 3 [Danio rerio]
gi|45356824|gb|AAS58451.1|[45356824]

AAS58450
17-beta hydroxysteroid dehydrogenase type 12B, 3-ketoacyl-CoA reductase type B
[Danio rerio]
gi|45356822|gb|AAS58450.1|[45356822]
```

NP_991147
17-beta hydroxysteroid dehydrogenase [Danio rerio]
gi|45387597|ref|NP_991147.1|[45387597]

AAN62014
17-beta-hydroxysteroid dehydrogenase type 4 [Danio rerio]
gi|24461283|gb|AAN62014.1|AF439319_1[24461283]

Nucleotides:

NM_200788
Danio rerio zgc:73320 (zgc:73320), mRNA
gi|41387185|ref|NM_200788.1|[41387185]

AY306007
Danio rerio photoreceptor associated retinol dehydrogenase type 2 mRNA, complete cds
gi|32250710|gb|AY306007.1|[32250710]

AY306006
Danio rerio photoreceptor associated retinol dehydrogenase type 1 mRNA, complete cds
gi|32250708|gb|AY306006.1|[32250708]

AY306005
Danio rerio 17-beta hydroxysteroid dehydrogenase (HSD17B1) mRNA, complete cds
gi|32250706|gb|AY306005.1|[32250706]

AY551082
Danio rerio 17-beta hydroxysteroid dehydrogenase type 12A, 3-ketoacyl-CoA reductase type A (hsd17b12A) mRNA, complete cds
gi|45356825|gb|AY551082.1|[45356825]

AY551081
Danio rerio 17-beta hydroxysteroid dehydrogenase type 3 (hsd17b3) mRNA, complete cds
gi|45356823|gb|AY551081.1|[45356823]

AY551080
Danio rerio 17-beta hydroxysteroid dehydrogenase type 12B, 3-ketoacyl-CoA reductase type B (hsd17b12B) mRNA, complete cds
gi|45356821|gb|AY551080.1|[45356821]

NM_205584
Danio rerio 17-beta hydroxysteroid dehydrogenase (HSD17B1), mRNA
gi|45387596|ref|NM_205584.1|[45387596]

AF439319
Danio rerio 17-beta-hydroxysteroid dehydrogenase type 4 (hsd17b4) gene, exons 19 through 24 and complete cds
gi|24461282|gb|AF439319.1|AF439315S5[24461282]

AF439318
Danio rerio 17-beta-hydroxysteroid dehydrogenase type 4 (hsd17b4) gene, exons 15 through 18
gi|24461281|gb|AF439318.1|AF439315S4[24461281]

AF439317
Danio rerio 17-beta-hydroxysteroid dehydrogenase type 4 (hsd17b4) gene, exons 8 through 14
gi|24461280|gb|AF439317.1|AF439315S3[24461280]

AF439316
Danio rerio 17-beta-hydroxysteroid dehydrogenase type 4 (hsd17b4) gene, exons 5, 6, and 7
gi|24461279|gb|AF439316.1|AF439315S2[24461279]

AF439315
Danio rerio 17-beta-hydroxysteroid dehydrogenase type 4 (hsd17b4) gene, exons 1 through 4
gi|24461278|gb|AF439315.1|AF439315S1[24461278]

AH012110
Danio rerio 17-beta-hydroxysteroid dehydrogenase type 4 (hsd17b4) gene, complete cds
gi|24461277|gb|AH012110.1|SEG_AF439315S[24461277]

7.1.3 Scientific presentations

April 2003, Bath, UK:
Serono Foundation for the advancement of medical science – workshop on molecular steroidogenesis
Mindnich, R., Adamski, J.: "Presence of multiple 17β-HSDs in the zebrafish"

7.2 Primer

7.2.1 Sequencing

Gene	*Name*	*Number (internal)*	*Position: exon / nt (in full-lenght sequence)*	*direction*
zfHSD1	RZPD-zfHSD1.2-for2	40280	exon 2 / nt 259	forward
	RZPD-1.2for2-5	40376	exon 3 / nt 466	forward
	RZPD-1.2-for3	40339	exon 6 / nt 787	forward
zfRHD 8.1	zfHSD1.3seq-for1	43079	exon 2 / nt 285	forward
	zfHSD1.3seq-rev1	43080	3'UTR / nt 1481	reverse
zfRDH 8.2	HSD1-N2-nested-for1	39486	exon 2 / nt 213	forward
zfHSD 3	zfHSD3.4/v9rev	47599	exon 8 / nt 629	reverse
	zfHSD3.4/v11rev	47600	exon 10 / nt 756	reverse
	zfHSD3.4/v7for	47598	exon 6 / nt 487	forward
zfHSD 12A	HSD3.1-nested-for2	39642	exon 1 / nt 170	forward
	HSD3.1-nested-rev1	39643	exon 6 / nt 632	reverse
	9141-for1	41316	exon 10 / nt 854	forward
	9141-for2	41502	3'UTR / nt 1260	forward
	9141-for3	41709	3'UTR / nt 1895	forward

7.2.2 Generation of probes

Gene	*Name*	*Number (internal)*	*Position: exon / nt (in full-length sequence)*	*direction*
zfHSD 12A	zf/pHSD3.1/for1	37421	exon 1 / nt 123	forward
	zf/pHSD3.1/rev1	37422	exon 7 / nt 656	reverse

7.2.3 RT-PCR based construction of full-length cds

Gene	*Name*	*Number (internal)*	*Position: exon / nt (in full-lenght sequence)*	*direction*
zfRDH 8.2	zfHSD1.1/start	42151	5'UTR / nt 1	forward
	zfHSD1.1/stop	42152	exon 2 / nt 230	reverse
	zfHSD1.1/express/for	42153	exon 1 / nt 79 (+restriction site)	forward
	zf-HSD/N2-pGEX-express-rev	40976	exon 6 / nt 1035 (+restriction site)	reverse
	RT-PCR with primers 42151 and 42152 yielded the fragment N, fusion PCR with primers 42153 and 40976 yielded the full-length cds (5'- and 3'-UTR excluded!).			
zfHSD 12B	3.3/BamH1for	43591	exon 1 / nt 1 (+restriction site)	forward
	zfHSD3.3/RHE5rev	44837	exon 5 / nt 444	reverse
	zfHSD3.3/RHE4for	44836	exon 4 / nt 333	forward
	3.3/51/Hind3rev	43590	exon 11 / nt 936 (+restriction site)	reverse
	RT-PCR with the primer pairs 43591+44837 and 44836+43590 yielded fragment A and B, respectively. Both fragments were combined by fusion PCR using primers 43591 and 43590.			

7.2.4 Expression analysis via RT-PCR

Gene	*Name*	*Number (internal)*	*Position: exon / nt (in full-length sequence)*	*direction*
zfHSD 1	RZPD/zfHSD1.2/for2	40280	exon 2 / nt 259	forward
	ZF-HSD1-rev2	36546	exon 6 / nt 820	reverse
zfRDH 8.1	zfpHSD1.3-for1	40950	5'UTR / nt 32	forward
	zfpHSD1.3-rev1	40951	exon 3 / nt 489	reverse
zfRDH 8.2	HSD-N1-nested-for2	39484	exon 1 / nt 97	forward
	zf-pHSD1-for1/rev	37118	exon 3 / nt 524	reverse
zfHSD 3	zfHSD3.4/v5for	47601	exon 4 / nt 378	forward
	zfHSD3.4/v11rev	47600	exon 10 / nt 756	reverse
zfHSD 12A	HSD3.1/nested/for2	39642	exon 1 / nt 170	forward
	zf/pHSD3.1/rev1	37422	exon 7 / nt 655	reverse
zfHSD 12B	3.3aPCR/3Fstart	46868	exon 1 / nt 111	forward
	3.3aPCR/6Fstop	46869	exon 4 / nt 355	reverse

7.2.5 Cloning into expression vectors

Gene	*Name*	*Number (internal)*	*Restriction site at 5'-end*	*direction*
zfHSD 1	zf-HSD/1.2-pGEX-express-for	40979	BamH I	forward
	zf-HSD/C-pGEX-express-rev	40978	EcoR I	reverse
	zfHSD1.2-pQE-express-rev	41399	Kpn I	reverse
	1.2exp-H3-rev	MWG-5	Hind III	reverse
	zfHSD1.2-XhoI-for	42617	Xho I	forward
	zfHSD1.2-XhoI-rev	42620	Xho I	reverse
zfRDH 8.1	zfHSD1.3/BamHI/for2	43348	BamH I	forward
	zfHSD1.3/Hind3/rev3	43596	Hind III	reverse
zfRDH 8.2	zfHSD1.1/express/for	42153	BamH I	forward
	zf-HSD/N2-pGEX-express-rev	40976	EcoR I	reverse
	zfHSD-N2-pQE-express-rev	41400	Kpn I	reverse
	1.1exp-H3-rev	MWG-4	Hind III	reverse
	zfHSD1.1-XhoI-for	42615	Xho I	forward
	zfHSD1.1-XhoI-rev	42619	Xho I	reverse
zfHSD 3	zf3.4/BamHIfor	47896	BamH I	forward
	zf3.4/EcoRIrev	47897	EcoR I	reverse
zfHSD 12A	3.1exp-XhoI-for	MWG-9	Xho I	forward
	3.1exp-XhoI-rev	MWG-10	Xho I	reverse
	3.1exp_SacI_f	MWG-14	Sac I	forward
	3.1exp_HindIII_r	MWG-15	Hind III	reverse
zfHSD 12B	3.3/BamH1for	43591	BamH I	forward
	3.3/51/Hind3rev	43590	Hind III	reverse

7.2.6 Site-directed mutagenesis

Gene	*Name*	*Number (internal)*	*Position: exon / nt (in full-lenght sequence)*	*direction*
zfHSD 1	zfHSD1.2mutfor	47201	exon 3 / nt 479	forward
	zfHSD1.2mutrev	47202	exon 4 / nt 514	reverse
zfRDH 8.1	zfHSD1.3mutfor	47203	exon 3 / nt 561	forward
	zfHSD1.3mutrev	47204	exon 4 / nt 594	reverse
zfRDH 8.2	zfHSD1.1mutfor	47199	exon 3 / nt 501	forward
	zfHSD1.1mutrev	47200	exon 4 / nt 538	reverse

7.3 Constructs

7.3.1 cDNAs

Gene	*Name*	*vector*	*Insertion sites*	*host*	*manufacturer*
zfHSD 1	LLKMp964D226Q2	pBK-CMV	EcoR I / Xho I	GeneHogs DH10B	RZPD
zfRDH 8.1	IMAGp998K1110662Q3	pBS-SK-	EcoR I / Sal I	GeneHogs DH10B	RZPD
zfRDH 8.2 *)	zfHSD1.1-pGEX (2)	pGEXΔBamHI	BamH I / EcoR I	JM107	R. Mindnich
	zfHSD1.1-pQE30 (1)	pQE 30	BamH I / Hind III	JM107	R. Mindnich
zfHSD 3	IMAGp998N0614301Q3	pCMV-pSPORT6.1	EcoR V / Not I	DH10B TonA	RZPD
zfHSD 12A	UCDMp611M09141Q10	pZIPLOX	Not I / Sal I	DH10B	RZPD
zfHSD 12B *)	zfHSD3.3AB-pGEX (4)	pGEXΔBamHI	BamH I / Hind III	JM107	R. Mindnich
	zfHSD3.3AB-pQE (6)	pQE 30	BamH I / Hind III	JM107	R. Mindnich

*) For these genes, no constructs containing the complete sequence were prepared (compare chapters 4.1.1 and 4.2.1). Instead, fusion PCR was employed only to generate the full-length coding sequence, which was subcloned into expression vectors (see also next chapter).

7.3.2 Expression of recombinant proteins

Gene	*Name*	*vector*	*Insertion sites*	*host*
zfHSD 1	zfHSD1.2-pGEX (5)	pGEXΔBamHI	BamH I / EcoR I	JM107
	zfHSD1.2-pQE30	pQE 30	BamH I / HindIII	JM107
zfRDH 8.1	zfHSD1.3-pGEX	pGEXΔBamHI	BamH I / HindIII	JM107
	zfHSD1.3-pQE30	pQE 30	BamH I / HindIII	JM107
zfRDH 8.2	zfHSD1.1-pGEX (2)	pGEXΔBamHI	BamH I / EcoR I	JM107
	zfHSD1.1-pQE30 (1)	pQE 30	BamH I / Hind III	JM107
zfHSD 1mut	zfHSD1.2G143M-pGEX (3)	pGEXΔBamHI	BamH I / Hind III	JM107
zfRDH 8.1mut	zfHSD1.3M146G-pGEX (2)	pGEXΔBamHI	BamH I / HindIII	JM107
zfRDH 8.2mut	zfHSD1.1M147G-pGEX (7)	pGEXΔBamHI	BamH I / EcoR I	JM107
zfHSD 3	zfHSD3.4N-pGEX (3)	pGEXΔBamHI	BamH I / EcoR I	JM107
	zfHSD3.4N-pQE (6)	pQE 30	BamH I / Hind III	JM107
zfHSD 12A	zfHSD3.1-pGEX (19)	pGEXΔBamHI	Sac I / Hind III	JM107
	zfHSD3.1-pET15b	pET 15b	Xho I	JM107
zfHSD 12B	zfHSD3.3AB-pGEX (4)	pGEXΔBamHI	BamH I / Hind III	JM107
	zfHSD3.3AB-pQE (6)	pQE 30	BamH I / Hind III	JM107

7.4 Protein sequences for phylogenetic analyses

7.4.1 From databases

Manuscript name	*GeneBank Acc.*	*origin*	*organism*
ag CP5823	EAA11852.1	protein	*Anopheles gambiae str. PEST*
ag EST	XM_315499.1	nucleotide	*Anopheles gambiae str. PEST*
bfmicRD	AF283540.1	nucleotide	*Branchiostoma floridae*
bn 3KAR	AAO43448.1	protein	*Brassica napus*
bovine 1	AAF73061.1	protein	*Bos taurus*
bovine prRDH	NP_776592.1	protein	*Bos taurus*
carp 12	CA968619.1	EST	*Cyprinus carpio*
catfish prRDH	CK417333.1	EST	*Ictalurus punctatus*
ce CDNA	BJ110143.1	EST	*Caenorhabditis elegans*
ce HSD17B a	NP_506449.1	protein	*Caenorhabditis elegans*
ce HSD17B b	NP_507092.1	protein	*Caenorhabditis elegans*
ce HSD17B c	NP_506448.1	protein	*Caenorhabditis elegans*
ce HypProt	AAB69884.1	protein	*Caenorhabditis elegans*
ce LET-767	NP_498386.2	protein	*Caenorhabditis elegans*
ce SDRmem	NP_505205.1	protein	*Caenorhabditis elegans*
chick 12	BU119147.1+CD218992.1	EST+EST	*Gallus gallus*
chicken prRDH	BU236399.1+BU233503.1	EST+EST	*Gallus gallus*
chicken 17beta-HSD 1	BAA19567.1	protein	*Gallus gallus*
ciona 32	AL669220.1	EST	*Ciona intestinalis*
ciona 33	BW274894.1	EST	*Ciona intestinalis*
ciona 34	AK116587.1	nucleotide	*Ciona intestinalis*
dm BcDNA	AAD38606.1	protein	*Drosophila melanogaster*
dm CG1217	NP_649563.1	protein	*Drosophila melanogaster*
dm CG1444	NP_572420.1	protein	*Drosophila melanogaster*
dm CG3154	NP_730972.2	protein	*Drosophila melanogaster*
dm CG3181	NP_724023.1	protein	*Drosophila melanogaster*
dm CG3699	NP_569875.2	protein	*Drosophila melanogaster*
dm CG6012	NP_609817.1	protein	*Drosophila melanogaster*
dm RE48687p	AAN71421.1	protein	*Drosophila melanogaster*
duck SPM2	O57314	protein	*Anas plathyrinchos*
ec pOxRed	AAK11695.1	protein	*Enterobacter cloacae*
eel 17beta-HSD 1	JC7561	protein	*Anguilla japonica*
fugu 3	CA332508.1	EST	*Takifugu rubripes*
fugu 12	CA845898.1	EST	*Takifugu rubripes*
gallus 3	BU137375.1	EST	*Gallus gallus*
hs Prot	AAH04126.1	protein	*Homo sapiens*
hs SDlike	NP_113651.3	protein	*Homo sapiens*
hs SRD	AAC39922.1	protein	*Homo sapiens*
human 17beta-HSD 1	NP_000404.1	protein	*Homo sapiens*
human 17beta-HSD 12	NP_057226.1	protein	*Homo sapiens*
human 17beta-HSD3	NP_000188.1	protein	*Homo sapiens*

human prRDH	NP_056540.1	protein	*Homo sapiens*
hv 3KAR	AAB82766.1	protein	*Hordeum vulgare*
marmoset monkey 17beta-HSD 1	AAG01115.2	protein	*Callithrix jacchus*
medaka 12	BJ489098.1	EST	*Oryzias latipes*
medaka prRDH	BJ525178.1	EST	*Oryzias latipes*
ml OxRed	NP_103540.1	protein	*Mesorhizobium loti*
ml pOxRed	NP_104075.1	protein	*Mesorhizobium loti*
mm9cisRD	NP_598767.1	protein	*Mus musculus*
mm cisRD	AAL14860.1	protein	*Mus musculus*
mm RDsim	AAK95855.1	protein	*Mus musculus*
mouse 17beta-HSD 1	NP_034605.1	protein	*Mus musculus*
mouse 17beta-HSD 12	NP_062631.1	protein	*Mus musculus*
mouse 17beta-HSD 3	NP_032317.1	protein	*Mus musculus*
mouse prRDH	XP_134689.1	protein	*Mus musculus*
nc HypP	XP_323815.1	protein	*Neurospora crassa*
pf Kik1	NP_702865.1	protein	*Plasmodium falciparum 3D7*
porcine 1	BP144707.1+CN152937.1	EST+EST	*Sus scrofa*
py pSDR	EAA17604.1	protein	*Plasmodium yoelii yoelii*
rat 17beta-HSD 1	NP_036983.1	protein	*Rattus norvegicus*
rat 17beta-HSD 3	NP_446459.1	protein	*Rattus norvegicus*
rat prRDH	XP_233743.1	protein	*Rattus norvegicus*
salmon prRDH 1	CB513415.1	EST	*Salmo salar*
salmon prRDH 2	CA064278.1	EST	*Salmo salar*
sc pOxRed	NP_639619.1	protein	*Streptomyces coelicolor A3(2)*
sc YBR159	NP_009717.1	protein	*Saccharomyces cerevisiae*
sp SDR	CAA19583.1	protein	*Schizosaccharomyces pombe*
sp pSDR	NP_593697.1	protein	*Schizosaccharomyces pombe*
stickleback prRDH	CD505363.1	EST	*Gasterosteus aculeatus*
trout 12	CA361044.1	EST	*Oncorhynchus mykiss*
trout prRDH	BX885263.2	EST	*Oncorhynchus mykiss*
xenopus prRDH 1	CD255274.1	EST	*Xenopus laevis*
xenopus 12	AAH41194.1	protein	*Xenopus laevis*
xenopus prRDH 2	CD362901.1	EST	*Xenopus laevis*
xl EST	CA981882.1	EST	*Xenopus laevis*
xl ESTplant	BG410026.1	EST	*Xenopus laevis*
xl mRNA	CA982789.1	EST	*Xenopus laevis*
xf OxRed	NP_299361.1	protein	*Xylella fastidiosa 9a5c*
zm 3KAR	AAB82767.1	protein	*Zea mays*

7.4.2 Generated in this PhD thesis

zfHSD 1

```
MEQKVVLITG CSSGIGLSLA VHLASNPAKA YKVYATMRNL DKKQRLLESV RGLHKDTLDI LQMDVTDQQS ILDAQRNVSE
GRIDILVCNA GVGLMGPLET HSLDTIRAIM DVNLLGTIRT IQTFLPDMKK KRHGRILVTG SMGGLQGLPF NEVYCASKFA
IEGACESLAI LLQHFNIHIS LIECGPVNTD FLMNLKRTET GDKELEVEVD AHTRSLYDQY LQHCQSVFQN AAQDTEDIIQ
VYLEAMEAQT PFLRYYTNRA LLPMSSLKLT SMDGSQYIRA MSKLIFSSPG TDAQK
```

zfHSD 12A

```
MESFNVVETL QPAERALFWV GALITASLAL YVVYKTITGF RIWVLGNGDL LSPKLGKWAV VTGATDGIGK SYAEELARRG
FSMMLISRSQ EKLDDVAKSL ESTYKVETKT IAVDFSQIDV YPKIEKGLAG LEIGILVNNV GISYSYPEFF LHIPDLENFI
TTMINVNITS VCQMTRLVLP RMEARAKGVI LNISSASGMF PVPLLTIYSS TKAFVDFFSR GLQTEYKCKG IIIQSVLPFF
VATKMTKIRK PTLDKPTPER YVAAELNTVG LQDQTNGYFP HAVMGWVTTI LAPIDLVLNL GLRMNKAQRG GYLRRRKLR
```

zfHSD 12B

```
MDPFADALFW VGAVTVLWLS VSSLWSLING IRVWILGNGN LMRASSLGKW AVVTGATDGI GKAYAEELAQ RGFAIDLISR
TQEKLDEVSK AIESKYKVET KTISADFGSV DIYPKIESGL AGLEIGVLVN NVGVSYSYPE FFLNIPDVDS FINNMININI
MSVCQMTRLV LPRMVDRSKG VILNVASASG MYPVPLLTLY SSTKAFVDFF SRGLDAEYKS KGIIIQSVLP FYVTTKLSKI
RKPTLDIPTP ERYVKAQLST IGLQTQSNGY LPHAIMGWVT ASLLPAKLLN KYVMGMGLSQ RARYLKKQKQ G
```

zfHSD 3

```
MTLTEIIFVL TGTCAILVFG GKIASLIMML ITKLFCPLPE AFFTSLGKWA VITGGSDGIG RAYAEELSKQ GMSVIIISRN
QEKLDRAAKK IELNTGGKVK VIAADFTKDD IYGHITENIE GLDIGVLVNN VGILPSQIPC KLLETSDLEE RIYDIVNCNV
KSMVKMCRIV LPGMQQRRRG VILNVSSGIA KIPCPIYTLY AASKVFVERF SQGLQAEYIS KGIIIQTVAP FGVSTAMTGH
QKPDMVTFTA EEFVRSSLKY LKTGDQTYGS ITHTLLGRIV QSIPTWVLQS SETFQHHFQE YVKNRDRR
```

zfprRDH 1

```
MASAGQKVVL ITGCSSGIGL GIAVMLARDK QQRYYVIATM RDLKRQEKLV CAAGDTYGKT LTVCTLDVCS NESVRQCVDS
VKDRHIDILI NNAGVGLVGP VEGLSLDDMM KVFETNFFGA VRMIKEVMPD MKKRRSGHII VISSVMGLQG VAFNDVYAAS
KFAIEGFCES LAVQLLKFNV TMSMIEPGPV HTEFEMKMYD DVSKKEYPNT DPETMHHFRT CYLPTSVNIF QGLGQTPEDI
AKVTKKVIES PRPPFRSLTN PLYTPIVALK YADDSGDLSL HTFYHMLYNL GGVMHVSVRI MKVLSFSWMR RRAVSPD
```

zfprRDH 2

```
MASGGGQKVV LITGCSSGIG LRIAVLLARD EQKRYHVIAT MRDLKKKDRL VEAAGEVYGQ TLTLLPLDIC SDESVRQCVN
SVKDRHIDVL INNAGVGLLG PVESISMDEM KRVFETNFFG TVRMIKEVMP DMKKRQAGHI IVMSSVMGLQ GVVFNDVYTA
SKFAIEGFCE SMAVQLLKFN VKLSLIEPGP VHTEFETKMM EEVAKMEYPG ADPDTVRYFK DVYVPSSIDI FEAMGQTPDD
IAKCTKKVIE TSQPRFRNLT NSLYTPIVAM KYADETGGLS VQTFYNLLFN FGSLMHISMS ILKCLTCNCL RRRTISPD
```

Danksagung (Acknowledgements)

Dr. Jerzy Adamski, Leiter des GAC, Privatdozent an der TU München und Doktorvater dieser Arbeit, gebührt besonderer Dank. Nicht allein wegen der Überlassung dieses Themas meiner Doktorarbeit und der fachlichen Betreuung, sondern insbesondere auch wegen seiner liberalen Führung und moralischen Unterstützung, die mir geholfen haben in seiner Arbeitsgruppe wissenschaftlich und menschlich zu wachsen. Dass meine Arbeit nun endlich nach so einigen Irrungen und Wirrungen einen befriedigenden Abschluss findet, ist zum größten Teil sein Verdienst.

An dieser Stelle möchte ich auch Prof. Dr. Rudi Balling noch einmal für seine Unterstützung und entscheidenden Ratschlag zur rechten Zeit ganz herzlich danken.

Als besonders angenehm und in jeder Hinsicht förderlich habe ich das Klima in dieser Arbeitsgruppe empfunden, wofür ich allen jetzigen und auch ehemaligen Mitgliedern der AG Adamski danken möchte. Meinen besonderen Dank möchte ich gegenüber folgenden Menschen aussprechen:
Dr. Gabi Möller, für ihre moralische Unterstützung, stete Hilfsbereitschaft, sowie dem ungebrochenen Versuch Ordnung ins Chaos zu bringen sowohl im Labor als auch beim Lesen dieses Manuskriptes.
Dr. Dominga Deluca, für ihre erfrischend-lebendige Art und Unterstützung in allen chemischen Fragen.
Thomas Ohnesorg, Mitstreiter der ersten Stunde, für ein stets offenes Ohr und Diskussionbereitschaft.
Brigitte Keller, für seelisch-moralische Unterstützung.
Dr. Rainer Breitling, Vollblut-Wissenschaftler, für seine „andere Sicht der Dinge“ und dass wir es dennoch zu einer Art Freundschaft geschafft haben.
Gabi Zieglmeier, für die unendliche Nachsicht, Geduld, Hilfsbereitschaft und Freundlichkeit.

Ich möchte mich zudem bei allen Praktikanten und Praktikantinnen bedanken, die ich während meiner Zeit betreuen durfte; aus dieser Zeit habe ich viel gelernt. Besonderer Dank gilt Marc Meier, dessen unerschütterlich-positive Lebenseinstellung mir wohl immer fremd aber immer willkommen sein wird, sowie Felix Halbach, dessen Ehrgeiz mich beeindruckt und dessen Mitarbeit mir menschlich und wissenschaftlich viel bedeutet hat.

Die Zebrafische für diese Arbeit wurden von Dr. Laure Bally-Cuif freundlicherweise zur Verfügung gestellt. Hierfür, sowie für die Einführung in die Arbeit mit diesem Tiermodell als auch die Unterstützung und Unterweisung in technischen und biologischen Fragen möchte ich ihr an dieser Stelle recht herzlich danken.

Mein Dank gilt zudem Dr. Michael Lang, dessen freundschaftliche und auch wissenschaftliche Unterstützung mir eine große Hilfe waren.

Meinen Eltern habe ich zu verdanken, dass ich überhaupt soweit gekommen bin, und dass sicherlich in vielerlei Hinsicht. Danke!

Curriculum vitae

Name:	Rebekka Mindnich
Geburtsdatum:	01.12.1973
Geburtsort:	Frankfurt / Main
Seit November 2000:	Doktorand / Wissenschaftlicher Mitarbeiter in der Arbeitsgruppe „Molekulare Endokrinologie und Proteomik“ von Dr. J. Adamski, Institut für experimentelle Genetik, GSF-Forschungszentrum für Umwelt und Gesundheit, Neuherberg, Deutschland
Januar-März 2003	Forschungsaufenthalt im Department of Biology, State University of San Diego, USA
Januar 1999-Oktober 2000:	Wissenschaftlicher Mitarbeiter in der Nachwuchsgruppe „Molekulare Genetik des Verhaltens“ von Dr. B. Lutz, Max-Planck-Institut für Psychatrie, München, Deutschland
1993-1998:	Studium der Biochemie, Ruhr-Universität Bochum, Deutschland *Abschluss:* Diplom
1980-1993:	BMV-Gymnasium, Essen, Deutschland *Abschluss:* Abitur