UAB

Universitat Autònoma de Barcelona

UAB

Universitat Autònoma de Barcelona

Structural and kinetic features of three human aldehyde dehydrogenases, ALDH1A1, ALDH1A2 and ALDH1A3, active in retinoic acid biosynthesis

Memòria presentada per

RAQUEL PEQUERUL PAVÓN

per optar al Grau de Doctor en Bioquímica, Biologia Molecular i Biomedicina

Treball realitzat al Departament de Bioquímica i Biologia Molecular de la Universitat Autònoma de Barcelona, sota la direcció del Doctors

JAUME FARRÉS VICÉN, XAVIER PARÉS CASASAMPERA I SERGIO PORTÉ ORDUNA

Jaume Farrés Vicén   Xavier Parés Casasampera   Sergio Porté Orduna

Raquel Pequerul Pavón

Bellaterra, 28 de setembre de 2018

AGRADECIMIENTOS

En primer lugar, me gustaría agradecer a los Doctores Xavier Parés y Jaume Farrés la oportunidad que me dieron al aceptarme como miembro del grupo ADHs, permitiéndome iniciarme en la carrera científica. Muchas gracias por todos los buenos consejos, por el tiempo dedicado en el desarrollo de mi formación académica e investigadora y por hacer posible la obtención final de esta Tesis Doctoral. Además, quiero dar las gracias de manera especial, a todos los miembros ADHs con los que he compartido laboratorio. He tenido mucha suerte de tener compañeros como vosotros. Os adoro.

A Sergio Porté, a quién conocí en las clases de problemas de enzimología y quién finalmente aceptó ser director de mi Tesis. Durante los años que hemos compartido me has enseñado a ser exigente y rigurosa. Gracias por todos los "*podries fer…*" que se te ocurrían a cada momento y que desmoronaban mis planes experimentales diarios. He echado muchísimo de menos tu presencia en el lab en la última época. Seguro que te irá bien independientemente de lo que hagas. Eres un ejemplo a seguir. Cris, gracias a ti he tenido fuerzas para seguir adelante. Mi compañera de máster, de chocolate después de comer, de llorar simultáneamente por el mismo motivo, la que me ha enseñado tantas cosas dentro y fuera del lab. Gracias por tu paciencia y por compartir conmigo tus conocimientos tan generosamente. Cuando te fuiste a IQS casi me da algo… Sin ti, el laboratorio tiene menos alegría y menos luz (y no solo porque ya no tengo que compartir contigo el HPLC ☺). Dentro de poco, tendré que llamarte Dra. Alsina y por fin podremos gritar, que lo hemos conseguido. Te quiero mucho, amiga. I a Adri también. Joan, te echo de menos. Mucho. Has sido mi compañero del "lado oscuro" y cuando te fuiste a Japón, pensé que no llegaría este momento, que no sería capaz sin ti. Gracias por todas esas veces que has puesto en pausa tus experimentos para echarme una mano, por ser tan generoso conmigo y enseñarme tantas cosas, por hacerme reír como nunca me he reído en el lab al cambiar una bombona de nitrógeno (aún lo pienso y me muero de risa), por tus brownies de té verde y chocolate blanco, por aguantar CadenaDial día sí y día también y por aquel ensayo de tu tesis en el Carbonyl2016, después de dos (o tres) botellas de vino blanco. Aún guardo los vídeos y me invade la nostalgia. ¿Peeerooo qué me estás cooontandooo? Isidro, pusiste un pie por primera vez en el laboratorio el mismo día que yo. Tú sí que vales para esto. Eres un crack. A ti también tengo mucho que agradecerte. Gracias por acudir al rescate siempre que te he necesitado, gracias por poner a punto tantos métodos en el laboratorio, por ayudarme constantemente con los p_pockets y por poner remedio a todos los interrogantes que existen. ¿Por qué sabes tanto? Estoy segura de que llegarás a ser un gran científico y que dentro de poco serás Dr. Crespo. Julio, tú fuiste el primero que me puso a trabajar. Expresar proteína, purificar proteína, cinéticas, cinéticas, cinéticas, y más cinéticas. Gracias a ti añado el cofactor y el sustrato por la misma esquinita de la cubeta. Tú me has enseñado a ser meticulosa y disciplinada. Gracias por todas las BBQ que has organizado, por abrirnos las puertas de tu casa y por estar siempre ahí cuando te he necesitado, sobre todo cuando no he pasado por buenos momentos; siempre has sabido animarme mostrándome la mejor versión de ti mismo. No sabes lo feliz que me hace que Mayra y tú hayáis sido papás de Alicia. Os merecéis todo lo bueno que os pase en la vida. *Padawans*: Javi, Eszter, Ignacio, Adrián y Rafa. De todos he aprendido mucho. Javi, mi primer padawan. ¡Qué gran descubrimiento! A ti te debo mucho, compañero… Gracias por ser tan auténtico, por venir en "pack mariquita" con Álvaro, por ser mi única compañía de pitis en la salida de emergencia, por todas y cada una de las visitas que me has hecho después de acabar el máster y por ser el padre de mis ALDHs mutantes. Estoy segura de que harás una buena tesis en Bilbao. Eszter, guapa. Aún recuerdo la cara de Joan e Isidro la primera vez que te vieron… Boquiabiertos. Gracias por la frescura que trajiste desde Hungría a este laboratorio, por ayudarme a mejorar mi inglés después de tantas conversaciones, por tu dedicación y compromiso con el proyecto de colaboración, por todos los pasteles que nos traías cada día, y por transmitir esa dulzura que te caracteriza y que te convierte en una persona maravillosa. Ignacio, el padawan que descubrió que el precursor de la ALDH2 humana es activo… WOW. Que bien defendiste tu máster, Nacho. Espero que te

vaya bien allí donde vayas. Adrián, mi compi de las inmovilizaciones imposibles, de la fauna y la flora caribeña, de los sangoraches y los lupinos. Que currazo te pegaste, pequeñajo. Tu sonrisa me ayudó muchísimo a superar aquel curso. Contigo todo era alegría y tranquilidad. Déjenlo cinco horas… hasta las once… (con acento argentino) XD. Gracias por preocuparte siempre por mí cuando me daban brotes de "alergia", que por entonces eran bastante frecuentes; por llegar antes que yo al laboratorio por las mañanas y saber siempre lo que tenías que hacer. Hagas lo que hagas, te deseo y te mereces lo mejor. Rafa, mi menorquín favorito. Eres muy TOP. Te admiro por cómo has cogido las riendas del proyecto, por cómo te has implicado desde que llegaste al grupo, por todo lo que has evolucionado desde que te conozco y por los magníficos resultados que has obtenido durante tu máster. Rafa, tu TFM es de matrícula, no tengo la menor duda. Quien la sigue la consigue, así que ánimo, porque vales mucho.

También quiero agradecer el tiempo compartido con toda la gente del Departamento: *Llevats* (Biosca, Chari y Edu), *Kinàsics* (Fani i Jordi), *Cromatina* (Andrea), *DNAs* (Inma y Alicia), *Ribos* (Mohamed, Javi, Guillem, Lu, Pablo, Helena), *Pros* (Irantzu, Samu, Valen, Jordi y Francisca), *RMNs* (Ana Paula, Lucía, Núria A., Laura, Pilar y Helena), *PFs* (Jofre, Laia, Gisela, Gabri y Alejandro) y *SysBio* (Nuria C.). Muchas gracias por el buen ambiente que se respira en la torre. Sois geniales. De manera especial, quiero dar las gracias a Helena Carbó, no sólo por preparar todos los tampones y los medios que he necesitado, sino por todas las sesiones de terapia que hemos compartido, por abrazarme en los momentos difíciles y por disfrutar conmigo de las mejores etapas. Eres única. Gracias por todo lo que haces por nosotros. A Salva, pieza fundamental en este Departamento, un sabio viajero a quien tengo mucho cariño. A Magda, por ser una magnífica profesional y mejor persona. Gracias por facilitarnos el trabajo a los profes durante las prácticas. Eres fantástica. Santi, al final conseguiste grabar mis clases de prácticas, eh. Gracias por tener siempre buenas palabras para mí… J. Carlos, María y Mar, muchísimas gracias por todas las gestiones. Con vosotros, todo es mucho más fácil.

Por último no quiero dejar de agradecer el apoyo que me han mostrado mi familia y mis amigos durante esta etapa. Mami, sé que aunque no sabes muy bien qué es lo que hago, valoras mucho mi constancia. Eso, os lo debo a vosotros; sé que el Papa estaría orgulloso de mí. Gracias por enseñarme a volar alto. Víctor, mi hermano pequeño, mi Enano, gracias por confiar en mí y por demostrarme qué es lo verdaderamente importante en la vida. Te quiero mucho. Lidia, mi cuñi, gracias por tantos días de compañía en la facultad, por aguantar mis historias de ciencia y por darme tu apoyo. Natalia y Kevin, gracias por estar ahí, y por interesaros siempre por mi trabajo. Edu, gracias por aparecer en nuestras vidas. Eres maravilloso. Juanita y Sergio, muchas gracias por el interés que siempre habéis mostrado por lo que hago y por intentar comprender mi trabajo. Sandra, mi amiga-hermana, gracias por secarme las lágrimas y ayudarme a liberar las tensiones de los experimentos. Juan, amigo, gracias por valorarme siempre y por estar ahí incondicionalmente. Gracias a todos mis amigos, Balta, Alba S., Rocío, Jaime, Nuri, Murillo, Mari, Alcaina, Juanito, Alba O., Davinia,… por entender que tengo poco tiempo y por adaptaros a mi poca disponibilidad. Os quiero mucho. Conchi, tú me enseñaste "*El arte de no amargarse la vida*" y eso, supuso un punto de inflexión que determinó mi futuro. Esta Tesis también es gracias a ti. Gracias a todos mis compañeros de Cinesa La Maquinista, por convertir los fines de semana en un viaje alternativo a la ciencia. Juanma, gracias por todo lo que has hecho por mí, por adaptarme los horarios cada vez que no podía con mi alma y por estar ahí siempre. Eres muy chachi. Gori, Carmen y Javi, sois los mejores compañeros de oficina que podía tener. Gracias por todos los cambios de turno que siempre estáis dispuestos a hacer conmigo. Paco, mil gracias por la portada de esta Tesis. Eres un artista. Al resto de mi Maqui-Equipo, deciros que siempre os llevaré en un trocito de mi corazón. Gracias a todos.

Y por supuesto, gracias a ti, Víctor. Gracias por apoyarme cada día y por confiar en mí. Sin ti, estoy segura de que me habría costado mucho más conseguirlo. Tengo que agradecerte tantas cosas… Gracias por levantarme cada vez que me caía, por enseñarme a ver las piedras del camino como un escalón, por estar dispuesto a cualquier cosa solo por verme feliz y por quererme incondicionalmente. Soy muy afortunada…

A quien me (ad)mira desde algún
punto del universo

Structural and kinetic features of three human aldehyde dehydrogenases, ALDH1A1, ALDH1A2 and ALDH1A3, active in retinoic acid biosynthesis

Raquel Pequerul Pavón

Abstract

The aldehyde dehydrogenase (ALDH) superfamily comprises a large number of dimeric and tetrameric proteins with a subunit molecular weight of approximately 55 kDa and different subcellular localization (cytoplasm, mitochondria and endoplasmic reticulum). ALDHs include a cluster of evolutionarily related $NAD(P)^+$-dependent enzymes catalyzing the oxidation of a wide spectrum of aldehydic substrates, generated from various endogenous and exogenous precursors, to their corresponding carboxylic acids.

The Thesis dissertation is a part of the structural and functional studies performed by our group on the role of oxidoreductases in retinoid metabolism, where their catalytic constants with retinoids were determined after solubilization with bovine serum albumin and by activity analysis using an HPLC-based methodology. To complete the study of the retinoid enzymatic pathway, this Thesis aims to perform an exhaustive and robust structural and kinetic analysis on the human enzymes involved in the irreversible oxidation of retinaldehyde to retinoic acid.

The first part deals with a comparison of the substrate-binding pocket of the human ALDH1A enzymes. ALDH1A1, ALDH1A2 and ALDH1A3 exhibited similar topologies and decreasing volumes in their substrates-binding pockets. The three enzymes were subcloned, overexpressed and affinity purified in their soluble and active form. Their enzymatic activity was characterized with typical aldehyde (alkanals and alkenals) substrates. The three enzymes were active with all tested substrates. In terms of k_{cat}/K_m values, ALDH1A3 exhibits the lowest values for all substrates, suggesting a moderate role in the physiological oxidation of these aldehydes. The k_{cat}/K_m values of ALDH1A1 and ALDH1A2 indicate a potentially major role in the transformation of these substrates with slightly different substrate specificity. Among the three enzymes, ALDH1A1 showed low K_m and k_{cat} values for most of the substrates, while ALDH1A2 exhibited the highest k_{cat} values. In order to measure activity with the physiological substrate of ALDH1As, retinaldehyde, an optimization of the solvent extraction methodology was carried out. From the evaluated methods, extraction with hexane/dioxane/isopropanol was chosen, since it was the most efficient in recovering retinaldehyde and retinoic acid from the activity buffer, with a yield near 100%. The three enzymes were active with two

retinaldehyde isomers and followed Michaelis-Menten kinetics, with K_m values in the micromolar range. The three forms exhibited lower K_m values for 9-*cis*-retinaldehyde than for all-*trans*-retinaldehyde. Related to the k_{cat} values, they were higher for the all-*trans* isomer (ALDH1A2 and ALDH1A3) or similar for the two isomers (ALDH1A1). Overall, the k_{cat}/K_m values for retinaldehyde were within the upper range as compared with those for most aldehyde substrates, being ALDH1A3, the best enzyme in terms of catalytic efficiency, followed by ALDH1A2. Moreover, the activity of ALDH1A enzymes with apo-β-carotenals, derived from the eccentric cleavage of β-carotene, was described for the first time.

The second part of this work is centered on the role of specific residues in the kinetic properties of ALDH1A enzymes. We performed site-directed mutagenesis, based on structural differences of selected residues from the substrate-binding pocket. The substitution L114P in ALDH1A1 was selected to make this part of the structure more similar to that of ALDH1A2. Likewise, in ALDH1A2, four contiguous residue changes, N475G, A476V, L477V and N478S were made to mimic the structure of ALDH1A1. In the ALDH1A1 L114P mutant, the K_m values for hexanal and citral were increased by 50-100 fold related to those of the wild-type ALDH1A1. This implies that the catalytic efficiency, specifically for citral, was similar to that observed in wild-type ALDH1A2. Conversely, the mutant ALDH1A2 exhibited a 50-fold decrease in the K_m value for citral. In addition, the 5-fold decrease in the k_{cat} value made the catalytic efficiency of mutant ALDH1A2 for citral to become similar to that of wild-type ALDH1A1. In regard to kinetics with retinaldehyde isomers, the mutants did not show significant differences with the respective wild-type forms, thus the mutated residues are not critical for retinaldehyde specificity.

Finally, inhibition studies of ALDH1A enzymes were performed in order to find novel, potent and selective inhibitors against ALDH1A1, ALDH1A2 and ALDH1A3. These preliminary results appear to be very promising to develop new pharmacophores by using structure-based drug design.

CONTENTS

TABLE LIST

FIGURE LIST

ABBREVIATIONS

ADH	alcohol dehydrogenase
AKR	aldo-keto reductase
ALDH	aldehyde dehydrogenase
AOX	aldehyde oxidase
BCO1	β-carotene-15,15'-oxygenase
BCO2	β-carotene 9',10'-oxygenase
BSA	bovine serum albumin
CAT	catalase
cGMP	cyclic guanosine monophosphate
COUPTFII	chicken ovalbumin upstream promoter-transcriptional factor II
CRABP	cellular retinoic acid binding protein
CRABP I	cellular retinoic acid binding protein type-I
CRABP II	cellular retinoic acid binding protein type-II
CRBP	cellular retinol-binding protein
CRBP2	cellular retinol-binding protein type-2
CYP26	cytochrome P450 family 26
CYP450 1A1	cytochrome P450 family 1A1
DEAB	N,N-diethylaminobenzaldehyde
DMSO	dimethyl sulfoxide
DOPAL	3,4-dihydroxyphenylacetaldehyde
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
FABP5	fatty acid-binding protein
FPLC	fast protein liquid chromatography
GABA	γ-aminobutyric acid
GST	glutathione S-transferase
HNF4α	hepatocyte nuclear factor 4α
HPLC	high performance liquid chromatography
IC$_{50}$	half maximal (50%) inhibitory concentration
IF	insoluble fraction or pellet
IL-13	interleukin 13
IPTG	isopropyl β-D-1-thiogalactopyranoside
k$_{cat}$	Catalytic constant

K_m	Michaelis constant
LRAT	lecithin:retinol acyltransferase
MCS	multicloning site
MDR	medium-chain dehydrogenases/reductases
NAD^+	nicotinamide adenine dinucleotide (oxidized)
NADH	nicotinamide adenine dinucleotide (reduced)
$NADP^+$	nicotinamide adenine dinucleotide phosphate (oxidized)
NADPH	nicotinamide adenine dinucleotide phosphate (reduced)
NCBI	National Center for Biotechnology Information
O.D.	optical density
O/N	overnight
PCR	polymerase chain reaction
PDB	Protein Data Bank
PMSF	phenylmethylsulfonyl fluoride
PPAR β/δ	peroxisome proliferator-activated receptor β/δ
RA	retinoic acid
RAL	retinaldehyde
RALDH	retinaldehyde dehydrogenase
RAR	retinoic acid receptors
RARE	retinoic acid response element
RBP	retinol binding protein
RDH10	retinol dehydrogenase 10
REH	retinyl ester hydrolase
RMSD	root-mean-square deviation
ROS	reactive oxygen species
RXR	retinoid X receptors
SDR	short-chain dehydrogenase/reductase
SDS-PAGE	sodium dodecyl sulfate–polyacrylamide gel electrophoresis
SF	soluble fraction or supernatant
STRA6	stimulated by retinoic acid 6
TCEP	tris(2-carboxyethyl)phosphine
TR	thyroid hormone receptor
TTR	transthyretin
VMS	visual molecular dynamics
WIN 18,446	N,N'-bis(dichloroacetyl)-1,8-octamethylenediamine
XO	xanthine oxidase

INTRODUCTION

- 18 -

1.1   Aldehyde dehydrogenases: general perspective

Aldehydes are organic compounds that are widespread in nature. They can be formed endogenously during numerous physiological processes, including the biotransformation of amino acids, neurotransmitters, carbohydrates, alcohols, biogenic amines, vitamins and steroids, and through processes such us lipid peroxidation, in which 4-hydroxy-nonenal and malondialdehyde are synthesized [1]. Aldehydes are often generated during the metabolism of a number of drugs and environmental agents. Various aldehydes, including ethanol-derived, acetaldehyde, and the anticancer drugs cyclophosphamide and ifosfamide are important aldehyde precursors [2].

Aldehyde dehydrogenases (ALDHs) include a cluster of evolutionarily related $NAD(P)^+$-dependent enzymes catalyzing the oxidation of a wide spectrum of aldehydic substrates, generated from various endogenous and exogenous precursors, to their corresponding carboxylic acids. Aldehydes are detoxified primarily through reductive and oxidative Phase I enzyme-catalyzed reactions, including the enzyme systems for aldehyde reduction, alcohol dehydrogenase (ADH), aldo-keto reductase (AKR) and short-chain dehydrogenase/reductase (SDR), and those for aldehyde oxidation, xanthine oxidase (XO), aldehyde oxidase (AOX) and ALDH (Figure 1).
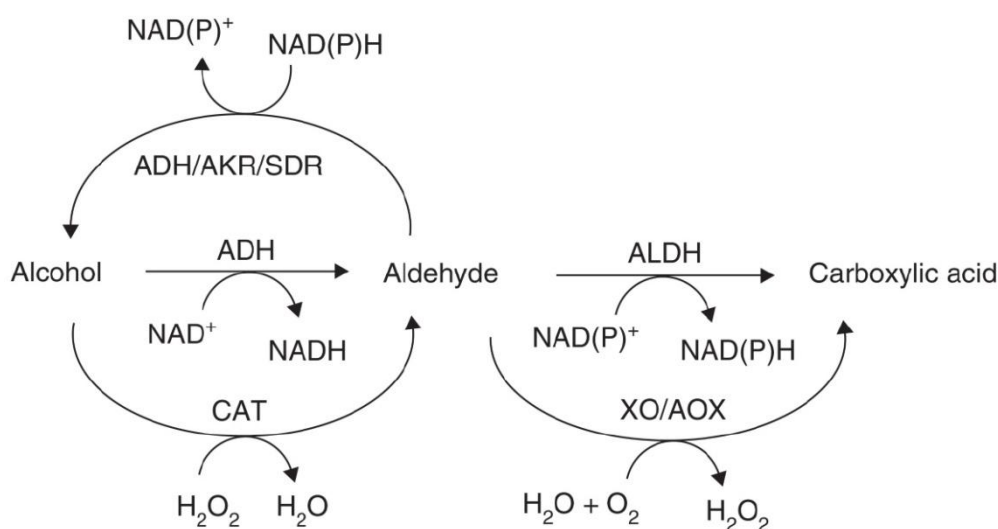
Figure 1. Enzymatic metabolism of aldehydes. ADH: Alcohol dehydrogenase; AKR: Aldo-keto reductase; ALDH: Aldehyde dehydrogenase; AO: Aldehyde oxidase; CAT: Catalase; SDR: Short chain dehydrogenase/reductase; XO: Xanthine oxidase. Reproduced from [2].

These enzymes play a particularly critical role in the cellular protection against toxic species, as evidenced by the fact that allelic variants of *ALDH* genes (leading to perturbations

in aldehyde metabolism) are the molecular basis of several disease states and metabolic anomalies [3].

1.2   Structure of *ALDH* genes

In the 70-80s, the classification of the ALDH superfamily was based on the substrate specificity, since the nomenclature was related to the catalyzed reaction. In 1998, a new standard nomenclature system based on the divergent evolution of genes was established [4]. Currently, the abbreviated symbol ALDH, followed by a number, represents the family. It is considered that ALDHs are part of the same family if they share at least 40% amino-acid identity sequence. When it is considered, a letter to assign the subfamily is added after the number. Two members of the same subfamily exhibit approximately, more than 60% amino-acid sequence identity. In addition, a number can be added to the letter that defines the subfamily to denote an individual gene within the subfamily (Figure 2).
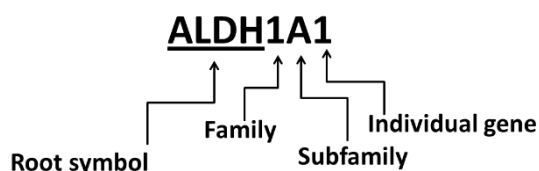
Figure 2. ALDH nomenclature system based on the divergent evolution of genes. The root symbol is followed by the number that defines the family, the letter that determines the subfamily and the number associated with the individual gene. The scheme is reproduced from the website of the *ALDH* gene superfamily resource database (www.aldh.org).

The *ALDH* gene superfamily is represented in all three taxonomic domains (*Archaea*, *Eubacteria* and *Eukarya*). In the 2002 update of the *ALDH* superfamily members [5], 555 *ALDH* genes were listed, including 32 from *Archaea*, 351 from *Eubacteria* and 172 from *Eukarya*. This update describes the *ALDH* genes of 11 representative vertebrate species (five primates, the cow, two rodents, two birds, and one fish) from which the full genome has been sequenced [6]. Recently, an update on the *ALDH* gene superfamily considers a common feature on vertebrate ALDH genes, such as the presence of duplications, which can lead to putatively protein-encoding genes or non-functional pseudogenes. The update of the gene study of these 11 species of vertebrates recently sequenced, has allowed increasing the number of records in the National Center for Biotechnology Information (NCBI) gene database. Until 2012, 51 different genes had been described for these 11 vertebrate species mentioned above, while currently, up to 207 different genes could be consulted. In this way, the latest update carried out by Vasiliou and coworkers [6] supposed an increase in the deposition of *ALDH* gene records described. In fact, in the Pfam database there are approximately 17,000 ALDH entries when pseudogenes and gene-duplication events are considered.

1.2.1 Structure of human *ALDH* genes

The *ALDH1A1*, *ALDH1A2* and *ALDH1A3* genes are conserved in most vertebrate species, although zebrafish lacks *ALDH1A1*. *ALDH1A1* and *ALDH1A3* genes consist of an average of 13 exons while *ALDH1A2* is constituted by an additional exon. Human *ALDH1A2* and *ALDH1A3* genes are located in chromosome 15, while the *ALDH1A1* gene is in chromosome 9 (Figure 3). Four ALDH1A2 isoforms may be generated from alternative splicing of the human *ALDH1A2* gene. The isoform 1 has been chosen as the canonical sequence. The isoform 2 sequence differs from the canonical sequence by missing of 228-265 residues.

(a) Human *ALDH1A1* gene (9q21.13).

(b) Human *ALDH1A2* gene (15q21.3).

(c) Human *ALDH1A3* gene (15q26.3).
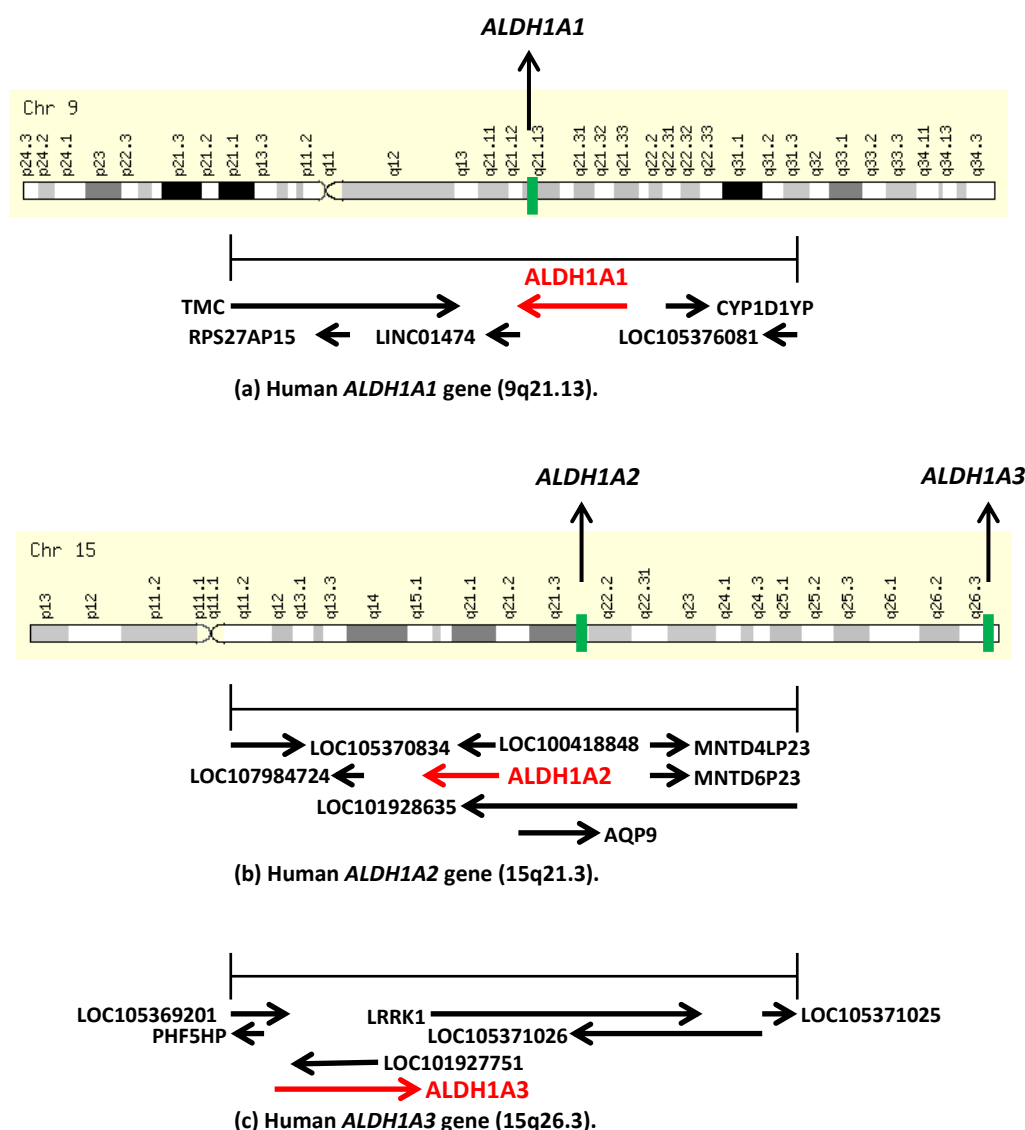
Figure 3. Position of *ALDH1A1*, *ALDH1A2* and *ALDH1A3* genes in human chromosomes. (a) *Locus* map of human *ALDH1A1* in chromosome 9. (b) *Locus* map of human *ALDH1A2* gene in chromosome 15. (c*) Locus* map of human *ALDH1A3* in chromosome 15. In the three schemes, the other *loci* present in the same chromosome localization are shown. The *loci* maps are reproduced from genomic context of NCBI resources (*ALDH1A1* Gene ID: 216, *ALDH1A2* Gene ID: 8854 and *ALDH1A3* Gene ID: 220).

1.2.2 Comparative and evolutionary studies of vertebrate *ALDH1A*-like genes

Multiple *ALDH1A*-like genes were identified in mouse, rat and marsupial genomes. Vertebrate ALDH1A1, ALDH1A2 and ALDH1A3 subunit sequences were highly conserved throughout vertebrate evolution. Comparative amino acid substitution rates showed that mammalian ALDH1A2 sequences were more highly conserved than those of ALDH1A1 and ALDH1A3. Phylogenetic studies supported the hypothesis for *ALDH1A2* as being the primordial gene originating in vertebrate genomes and undergoing sequential gene duplication to generate two additional genes, *ALDH1A1* and *ALDH1A3*, in most vertebrate genomes (Figure 4) [7].

The ALDH1 family consists of six human *ALDH* genes: *ALDH1A1*, *ALDH1A2*, *ALDH1A3*, *ALDH1B1*, *ALDH1L1* and *ALDH1L2*. The genomes of *Rattus norvegicus* (rat) and *Mus musculus* (mouse) contain an additional gene, *Aldh1a7* which is 92 % identical to mouse *Aldh1a1*. Therefore, the rodent *Aldh1a7* very likely arose as a gene duplication event after the mammalian radiation ~ 70 million years ago (mya) and then became fixed in the genome before the rat-mouse divergence ~ 17 mya [8].
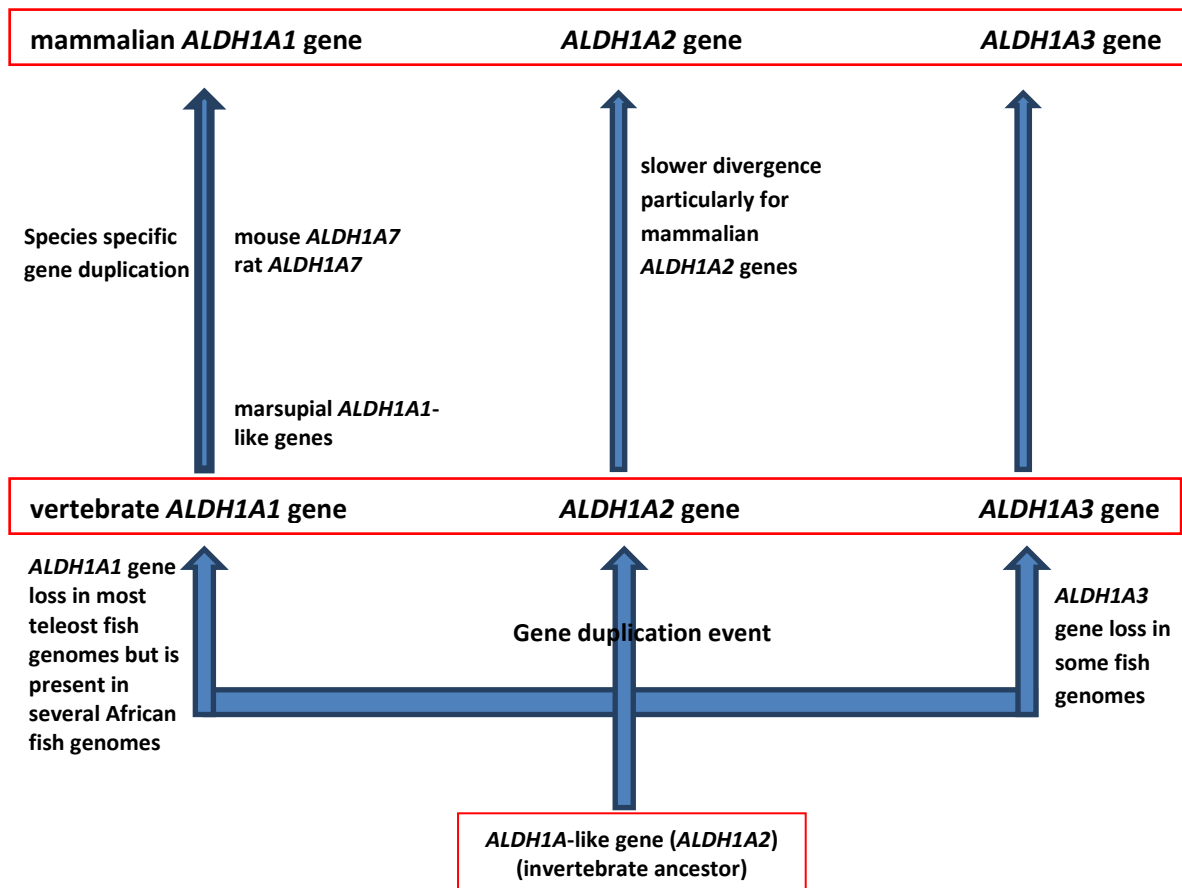
Figure 4. Hypothetical evolutionary appearance of the vertebrate *ALDH1A*-like genes. Note the presence of at least three *ALDH1A*-like genes throughout vertebrate evolution with the exception of some teleost fishes and the slower rate of sequence divergence for mammalian *ALDH1A2* genes. Reproduced from [7].

1.3   Structural and catalytic properties of ALDHs

1.3.1   Structural properties

The ALDH superfamily includes tetrameric and dimeric enzymes with subunits of approximately 55 kDa, having different subcellular localization: cytoplasm, mitochondria, endoplasmic reticulum, etc.

The crystallographic structures of sheep ALDH1A1 [9] (PDB code 1BXS) and rat ALDH1A2 [10] (PDB code 1BI9) show a subunit profile with three domains (Figure 5): an N-terminal NAD$^+$-binding domain (residues 8-135 and 159-270 in sheep ALDH1A1 and residues 1-136 and 161-270 in rat ALDH1A2), a catalytic domain (residues 271-470 and 271-484, respectively), and an oligomerization domain (residues 140-158 and 486-495 in sheep ALDH1A1, and residues 137-160 and 485-500 in rat ALDH1A2).

In sheep ALDH1A1, the N-terminal NAD$^+$-binding domain and the catalytic domain contain a five-stranded parallel β-sheets and six-stranded parallel β-sheets, respectively. Both the NAD$^+$-binding and catalytic domains are based on topologically related βαβ type polypeptide folds. In turn, the oligomerization is made up by three-stranded antiparallel β-sheets (Figure 6) [9].

It is recognized that the nucleotide-binding domain of ALDH enzymes presents a unique mode of NAD$^+$ binding that does not appear to be related to the canonical binding of this cofactor in the remaining NAD(P)-dependent dehydrogenases. The canonical binding of the cofactor in NAD(P)-dependent dehydrogenases exhibits a "fingerprint" sequence corresponding to the Gly-X-Gly-X-X-Gly (GXGXXG) residues [11]. This moiety is the characteristic pattern that reflect the turn at the end of the first β-strand in the first mononucleotide-binding unit of the Rossmann fold, which interacts with the adenine ribose of NAD$^+$ [12]. Usually, NAD(P)-binding dehydrogenases contain a Rossmann fold in their dinucleotide-binding domains and in all cases, the pyrophosphate moiety of the dinucleotide makes close contact with the first β-α-β-α-β unit, specifically in the loop between β1 and αA, of the Rossmann fold (Figure 7) [13].

An unusual feature observed in the crystallographic structure of bovine ALDH2 (PDB code 1A4Z) reveals that the cofactor molecule binds across the N terminus of the αD helix instead of the αA helix. This difference may explain why an amino acid sequence corresponding to the Gly-X-Gly-X-X-Gly motif is not found in ALDH enzymes [14]. The linear sequence that most closely resembles this motif, Gly-Ser-Thr-Glu-Val-Gly (GSTEVG; residues 245-250) is localized in

the region where the cofactor is bound. The adenine ring makes van der Waals contacts with the side chains of residues Gly225, Pro226, Ala230, Val249, Leu252 and Ile165 in bovine ALDH2, which are conserved in the sheep ALDH1A1 and rat ALDH1A2 structures (Figure 8). No hydrogen bonds are formed between the protein and the adenine ring of the cofactor molecule.
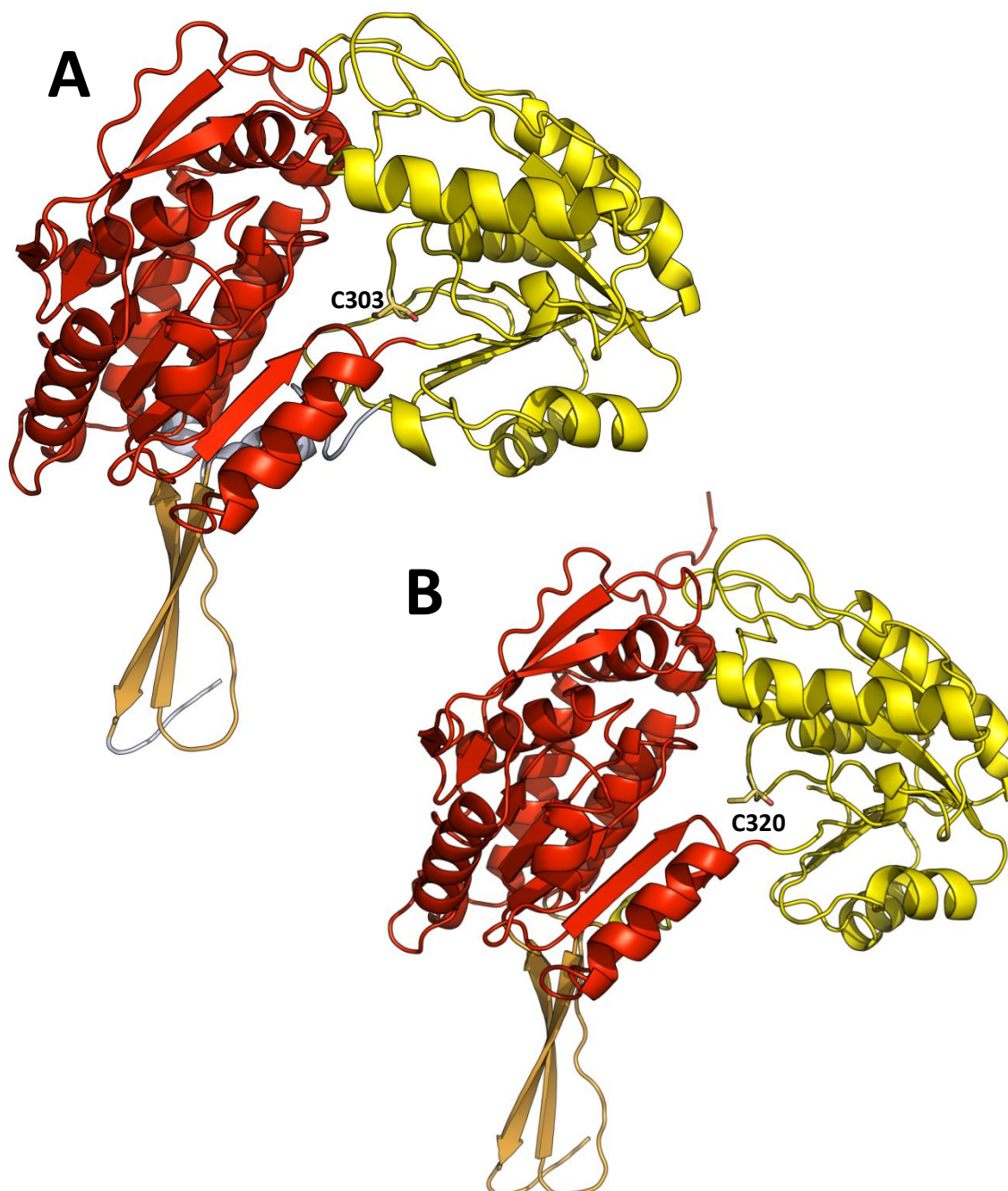
Figure 5. Stereoview showing the monomer subunit of (A) sheep ALDH1A1 (PDB code 1XBS) and (B) rat ALDH1A2 (PDB code 1BI9). In both, the NAD$^+$-binding domain is shown in red; the catalytic domain and the oligomerization domain are represented in yellow and bright orange, respectively. The nucleophile C302 (in ALDH2 nomenclature) is represented in sheep ALDH1A1 and rat ALDH1A2 monomers. Structural images were created using the program PyMOL (www.pymol.org).

Figure 6. Alignment of sheep ALDH1A1 and rat ALDH1A2 amino-acid sequence. The conserved residues are highlighted in red. Residue substitutions by amino-acid with similar properties are highlighted in yellow. Non-conservative residues are highlighted in white. Secondary structure elements, obtained from sheep ALDH1A1 crystal structure (PDB code 1BXS) are shown above the sequence: α-helices are displayed as squiggles; β-strands are rendered as arrows and strict α- and β- turns are displayed as TT [15].

Hydrogen-bonding interactions with the adenosine ribose are made by the sidechains of Lys192, Glu195 and the mainchain oxygen of Ile166, in both the ALDH1 and ALDH2 structures, whereas Glu399 makes two hydrogen bonds to the nicotinamide ribose [9,14]. Lys192 and Glu399 are two residues totally conserved in ALDH enzymes and mutations in these residues are responsible for the change on the rate-limiting step of ALDH2 from acetylation to hydride transfer. Although they are not involved in the catalytic chemistry, they are essential to ensure the correct hydride transfer [16,17].

It is known that the nicotinamide half of the NAD$^+$ adopts two major conformations in the sheep ALDH1A1 structure [9]. The nicotinamide ring of ALDH1A1 occupies nearly the same position as the nicotinamide ribose of ALDH2 (Figure 8B and C), and it corresponds with the less occupied conformation [14]. The major conformer includes a slight reorientation of the adenine phosphate being the structure displaced relative to that observed in bovine ALDH2. The reorientation of the major NAD$^+$ conformer in ALDH1A1 results in a movement of over 5 Å for the nicotinamide ribose relative to bovine ALDH2. Indeed, the interactions between the nicotinamide ribose and the protein involve the same residues in both ALDH1A1 and ALDH2,

but the hydrogen bonds and van der Waals contacts that they form are different (Figure 8A) [9].
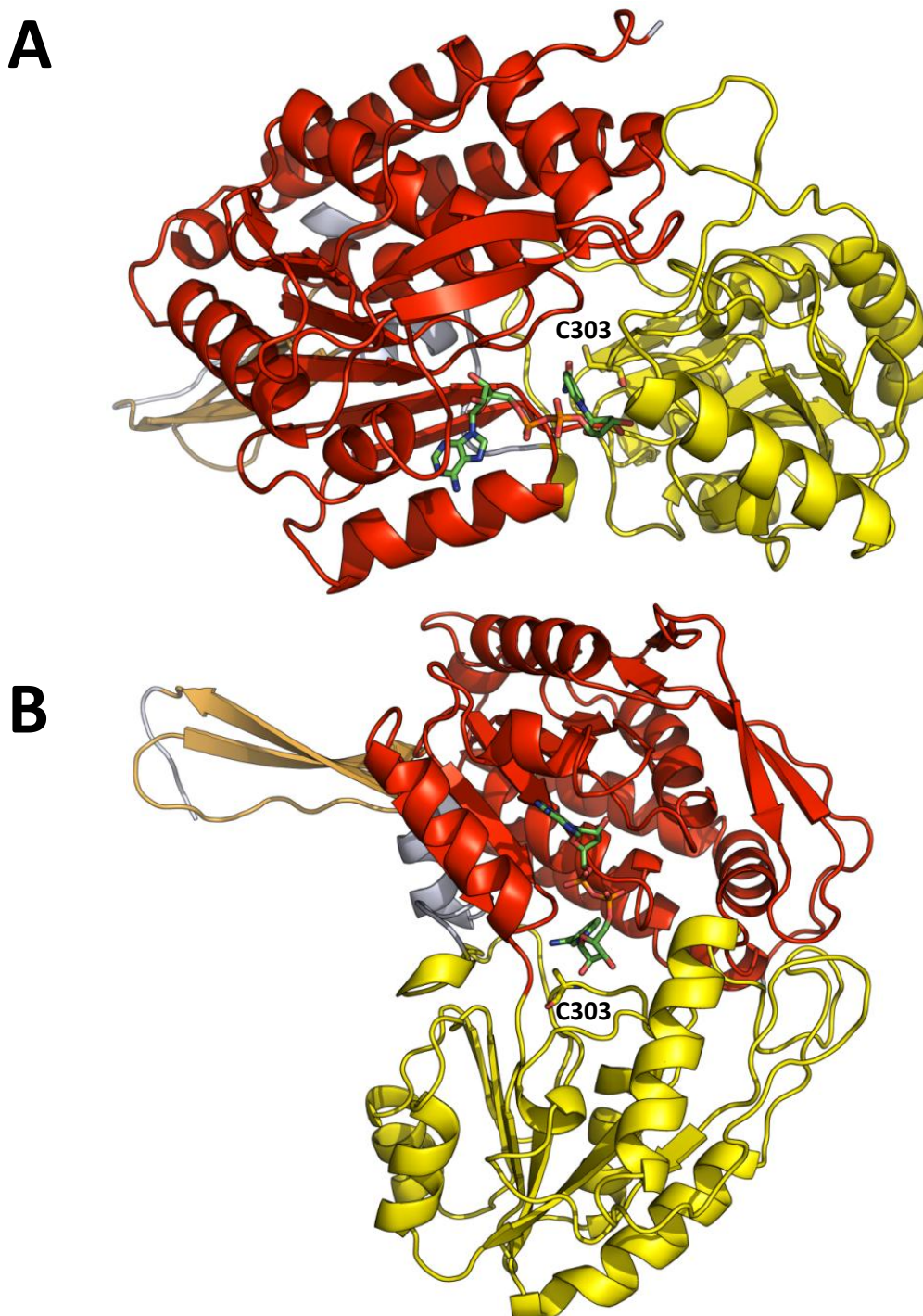
Figure 7. Crystal structure of sheep ALDH1 complexed with NAD$^+$ (PDB code 1BXS). Top view (A) and lateral view (B) of the cofactor position in the interface of the NAD$^+$-binding and catalytic domains. The nicotinamide ring is located near the nucleophile Cys 303. The NAD$^+$-binding domain and the catalytic domains are shown in red and yellow, respectively. The NAD$^+$ cofactor is represented in green.

In sheep ALDH1A1, it has been possible to describe both cofactor conformations because a different electronic density has been detected and there is evidence of a discrete disorder in the nicotinamide half of NAD$^+$. In their discussion of ALDH2 structure, the authors commented

that the conformation of the nicotinamide ring as seen is not possible for the entire catalytic cycle. It was suggested that the nicotinamide must change positions during the catalytic cycle and this suggestion would be consistent with the fact that the NAD$^+$ is in a different conformation in sheep ALDH1, or that the nicotinamide ring is mobile in rat ALDH1A2 [18].
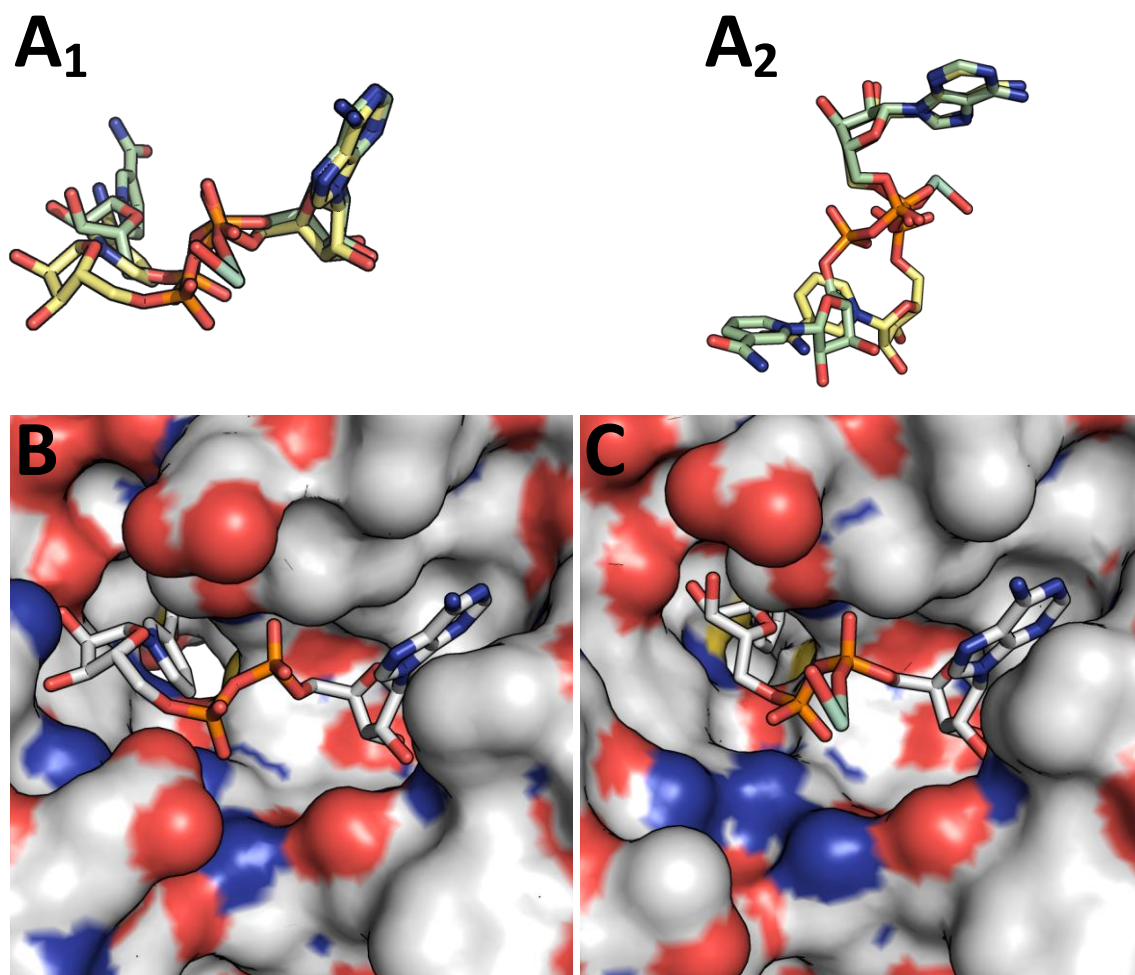
Figure 8. NAD$^+$-binding comparison between sheep ALDH1A1 and bovine ALDH2. (A) The refined coordinates of the ALDH1A1 NAD$^+$ molecule (in yellow with standard atom colors) are superimposed with the ALDH2 NAD$^+$ molecule (in green). (A$_1$) The NAD$^+$ orientations are maintained on the surfaces (B) and (C). (A$_2$) The NAD$^+$ orientations are showed from the frontal view of catalytic Cys. (B) and (C) The surface representation coloured by electrostatic potential of the NAD$^+$-binding conformations of sheep ALDH1A1 (B) and bovine ALDH2 (C). In the bovine ALDH2 structure, the A$_1$ conformation of NAD$^+$ is the only observed, while for the sheep ALDH1 structure both NAD$^+$ conformations can be detected, being the A$_2$ the most frequent.

Most of ALDHs characterized to date are specific for NAD$^+$ [9,13,14,19,20]. Also available are the structures of bacterial enzymes described for NADP$^+$-dependent ALDHs, such as the ALDH from *Vibrio harveyi* [21] and *Streptococcus mutans* [22]. The interaction between the NADP(H) phosphate group and the ALDH residues involves the formation of hydrogen bonds between the enzyme residues and the hydroxyl groups of the adenosine ribose [23]. In the case of ALDHs that bind NAD$^+$ better than NADP$^+$, there is a negatively charged amino acid residue (Glu or Asp) that would likely repeal the 2'-phosphoryl moiety of NADP$^+$. Whereas, this

residue is not negatively charged (Ala, Val, Leu, Ile, Thr or Cys) in ALDHs that preferentially bind NADP$^+$ [24].

The substrate-binding pocket contains characteristic and differential features in ALDH1A and ALDH2 enzymes. This fact implies differences in the type of substrate used by each enzyme type. The size of the substrate access channel is a crucial determinant of ALDH1A and ALDH2 function. While ALDH2 shows a narrow channel that implies the ability to process small aldehydes, ALDH1A includes a large pocket that allows the catalysis of bulky aldehydes. The residue at position 124 serves as a selective gate for the aldehyde size and would determine the substrates that can access the active site [25]. In terms of evolution, longer and larger channels are associated with the loss of ancestral bulky residues (Met124), which have been replaced by smaller ones (Ala124 or Gly124) (Figure 9). Briefly, ALDH1A1 possesses a wider opening leading to the active site, whereas ALDH2 has a much more constricted, cylindrical shaped site [26]. Therefore, the rational design of selective inhibitors for these enzymes has to take advantage of the differences found in the distinct surface topologies described above.
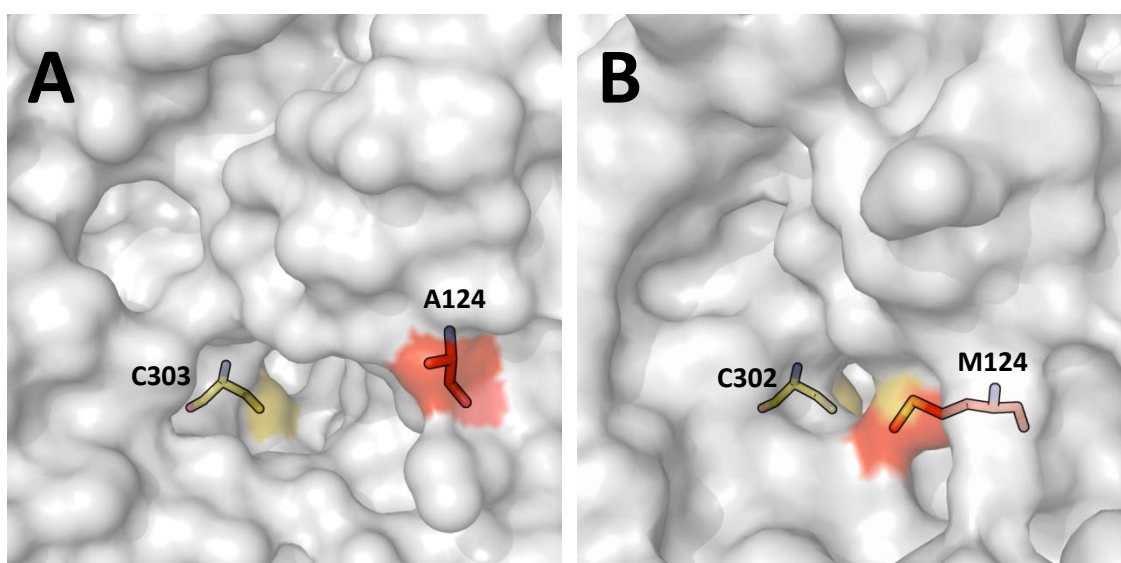
Figure 9. Top view of the substrate access channel of human ALDH1A1 (A) and human ALDH2 (B). Crystal structure of human ALDH1A1 (PDB code 4WB9) exhibits a large substrate entrance channel by presence of small residue in the position 124 (Ala). However, the human ALDH2 (PDB code 1CW3) shows a narrow substrate-binding pocket by the evolutionary maintenance of a bulky residue in position 124 (Met). The bulky residues are in consonance with small substrate access channels and determine the substrate accommodation in the active site.

Most ALDH enzymes share a high sequence identity in their cofactor-binding site and in the catalytic center. Structurally, most ALDH conserved characteristic catalytic residues, which consist of Asn169, Glu268 and Cys302 (human ALDH2 numbering). Based on the ALDH2 sequence, site-directed mutagenesis has shown that Cys302 is the essential nucleophile for

dehydrogenase reaction (Figure 10), with Glu268 acting as the proton acceptor to activate Cys302, and Asn169 having a role in the stabilization of the transition state intermediate [27].
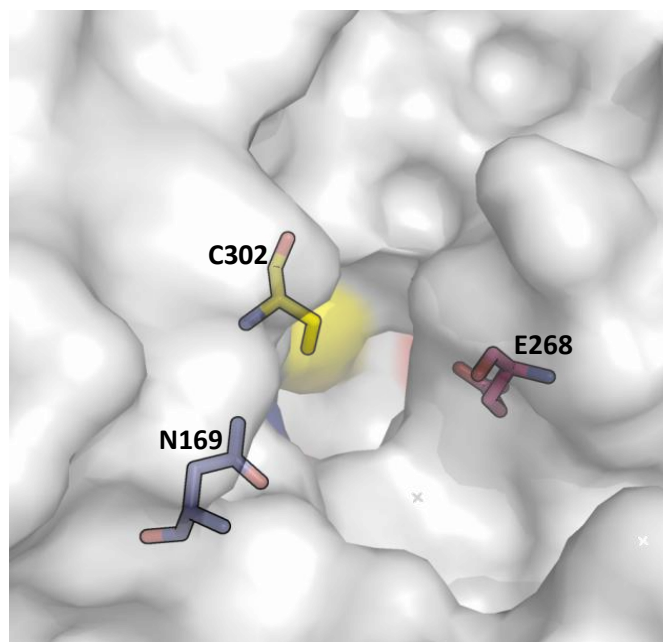
Figure 10. The active site of human ALDH2 in the vicinity of Cys302. Mutational analysis has shown that Glu268 functions as a general base necessary for both non-physiological esterase activity and the normal dehydrogenase activity of ALDH2 [28]. The sidechain amide nitrogen of Asn169 and the peptide nitrogen of Cys302 are positioned to both stabilize the developing oxyanion in the thiohemiacetal transition state and to orient the thiohemiacetal for optimal hydride transfer to NAD$^+$. Asn169 conserves specific interactions with the cofactor molecule and possibly with the transition state of the substrate [14].

1.3.2   Catalytic properties

Most ALDHs are catalytically active proteins. The crystallographic structures of ALDHs have allowed to disclose the essential feature of their catalytic mechanism. The members of the ALDH1A subfamily are tetrameric cytosolic enzymes that catalyze the oxidation of the retinol metabolite, retinaldehyde, to retinoic acid [29,30]. In contrast, ALDH2 is the primary enzyme involved in the oxidation of acetaldehyde during ethanol metabolism [31]. The kinetic mechanism seems to be similar in all ALDHs. In general, ALDHs exhibit a sequential ordered bi-bi mechanism, in which the oxidized cofactor first binds to the enzyme to allow the subsequent substrate attachment (Figure 11) [32,33].
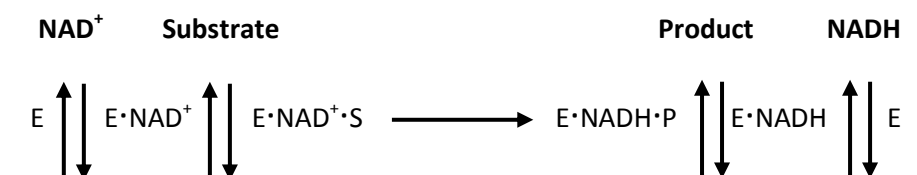
Figure 11. Sequential ordered bi-bi mechanism to the irreversible reaction catalyzed by ALDHs. The oxidized cofactor binds first to the enzyme and allows substrate binding. When the oxidation reaction occurs, the product is released prior to the reduced cofactor.

The variability of residues in the substrate-binding site clearly indicates evolutionary substrate specificities of individual ALDHs. However, farther to the interior of the active site strict conservations are compatible with a common catalytic mechanism [13]. Site-directed mutagenesis experiments have revealed that catalysis occurs in sequential steps (Figure 12): 1) activation of the catalytic residue by abstraction of a proton from a water molecule between Cys302 and Glu268, 2) consequent nucleophilic attack on the electrophilic aldehyde by the thiolate group of Cys302, 3) formation of a tetrahedral thiohemiacetal intermediate (deacylation) with concomitant hydride transfer to the pyridine ring of NAD^+, 4)hydrolysis of the resulting thioester intermediate, 5) dissociation of the reduced cofactor and 6) subsequent regeneration of the enzyme by NAD^+ binding [34,35].

It is well established that the cytoplasmic and mitochondrial ALDHs from a variety of mammalian sources also possess esterase activity toward 4-nitrophenyl esters. Cys302 is the essential nucleophile for both esterase and dehydrogenase reactions, with Glu268 acting as the general base to activate Cys302 [27,28]. Thus, the hydrolysis of *p*-nitrophenyl acetate has been shown to take place at the same catalytic center which is responsible for dehydrogenase activity (Figure 13). The hydrolysis of *p*-nitrophenyl esters by human ALDH was inhibited completely by glyceraldehyde and chloral hydrate, substrate and potent competitive inhibitor of the dehydrogenase activity, respectively [36].

Figure 12. Catalytic mechanism of aldehyde oxidation based on ALDH2. The thiolate attacks the substrate carbonyl carbon (I), to form a transition-state tetrahedral intermediate with hydride transfer (II). Thus, the thioester (III) is poised for base catalysis initiated by abstraction of a proton from a water molecule between Cys302 and Glu268 (II and III), creating a second tetrahedral intermediate (IV). Collapse of this intermediate yields the product acid which may reprotonate the thiolate (V). The enzyme can bind reactants again (VI). Reproduced from [34].

Figure 13. Hydrolysis mechanism of *para*-nitrophenyl esters. 1, Cys302 nucleophilic residue is activated from a basic residue, Glu268, by hydrogen abstraction in catalytic site. 2, the substrate *p*-nitrophenyl acetate is attacked by thiolate ion reaching the oxyanion intermediate formation. 3, the oxyanion intermediate reshuffles and nitrophenolic group leaves. 4, Glu268 residue extracts a hydrogen atom from neighboring ordered water molecule, which becomes a nucleophile attacking the thioacyl enzyme complex. 5, the tetrahedral intermediate formed is reorganized to release acetic acid and free enzyme. Reproduced from [35]**.**

1.4   Human ALDH1A subfamily

Within the superfamily of ALDHs, this Thesis focuses on the structural and kinetic features of some members of the ALDH1A subfamily:

1.4.1   ALDH1A1

The ALDH1A1 gene product, like ALDH1A2 and ALDH1A3, is a homotetramer with ~55-kDa subunits (Figure 14). ALDH1A1 is ubiquitously distributed in the adult epithelium of various organs including brain, cornea, lens, retina, liver and gastrointestinal tract [37,38]. Mouse and rat ALDH1A1 is one of the three highly conserved cytosolic enzymatic forms (together with ALDH1A2 and ALDH1A3) that catalyze the oxidation of the retinol metabolite, retinaldehyde, to retinoic acid [29,30] and is often referred to in the literature as RALDH1. The *in vivo* role of human ALDH1A1 in retinoic acid synthesis is evidenced in the human brain, where *ALDH1A1* is highly expressed in dopaminergic neurons of various areas, including medulla, striatum,

cerebellum and cortex [39], specifically *ALDH1A1* is very highly expressed in human dopaminergic cells of the substantia nigra and ventral tegmental area, which are known to require retinoic acid for their differentiation and development [40]. In addition, rodent ALDH1A1 is one of the major enzymes involved in the elimination of ethanol metabolite, acetaldehyde, with K_m values of 50-180 µM [41]. Therefore, ALDH1A1 has been implicated in several alcohol-related phenotypes including alcoholism and alcohol sensitivity. Ethanol, through its active metabolite acetaldehyde, has been shown to interfere with the metabolism of biogenic aldehydes including 3,4-dihydroxyphenylacetaldehyde (DOPAL) over the competitive inhibition of ALDH1A1, and other ALDHs, such as ALDH2 and ALDH1B1 [42].
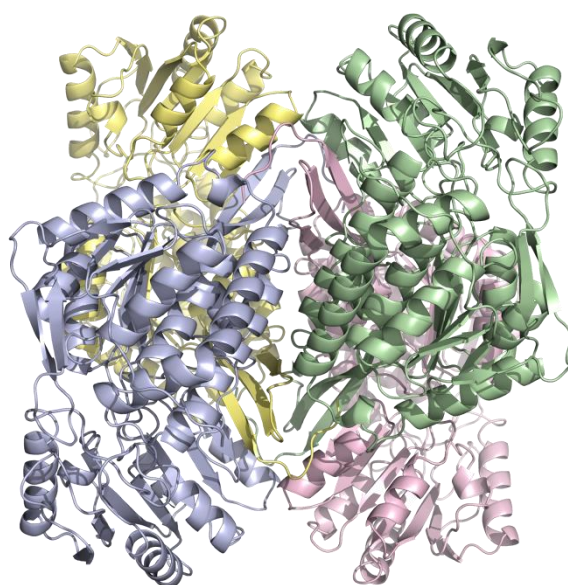
Figure 14. Human ALDH1A1 crystallographic structure (PDBe rendering based on 4WB9). The image shows the homotetramer with the individually colored subunits in green, yellow, pink and blue.

Furthermore, *ALDH1A1* is one of the 139 genes that are differentially expressed in primary human hematopoietic stem cells and, through the production of retinoic acid, ALDH1A1 has been shown to promote their differentiation [43,44]. It is widely used as a marker for normal and cancer stem cells, playing an important biological role in the metabolism of reactive oxygen species (ROS). Primitive quiescent cells exhibit low levels of ROS, while intracellular ROS increases imply proliferation and differentiation. However, an excess of ROS production or exposure leads to premature apoptosis. Particularly, ALDH1A1 plays a relevant role in tight regulation of ROS production by detoxification of strongly electrophilic and reactive compounds, which include 4-hydroxynonenal, acrolein and malonaldehyde [45]. Therefore, it has a key function in the cellular defense against oxidative stress. Similarly to other ALDHs, ALDH1A1 may also play an important role in cancer chemoresistance. Thus, it decreases the effectiveness of some oxazaphosphorine anticancer drugs, by detoxifying their major active

aldehyde metabolites [46]. In this area, it is admitted that ALDH1A1 binds some anticancer drugs and is downregulated in certain carcinomas [47]. Finally, aside from aldehyde metabolism, ALDH1A1 possesses esterase activity [36], binds thyroid hormone [48] and is induced by estrogens [49], suggesting it may be regulated by or involved in hormone signaling.

1.4.2 ALDH1A2

ALDH1A2 is a cytosolic homotetramer with 56.7-kDa subunits (Figure 15) expressed ubiquitously in various embryonic and adult tissues including brain, digestive tract, lung, heart, testis, ovary, endometrium, pancreas, liver and many others [50,51]. ALDH1A2, often referred to as RALDH2 in the literature, exhibits complex expression patterns throughout embryonic development. This suggests that ALDH1A2 plays an important role in retinoid metabolism during embryonic development where it is considered to be the major retinoic acid-synthesizing enzyme during early embryogenesis [52].

Human ALDH1A2 is one of the least characterized and understood ALDH enzymes. Only two crystal structures of human ALDH1A2 are known to date, human ALDH1A2 liganded with NAD^+ (PDB code 4X2Q) [53] and human ALDH1A2 liganded with NAD^+ and compound WIN 18,446 (PDB code 6ALJ) [54]. Previously, a rat ALDH1A2 with NAD^+ bound crystal structure was reported (PDB code 1BI9) [18]. The rat and human ALDH1A2 holoenzymes (PDB code 1BI9 and 4X2Q, respectively) exhibit a large disordered loop region including residues 475-495, which border the active site. In the ternary complex with the compound WIN 18,446 (PDB code 6ALJ), this loop displays a well-defined regular structure, may be by the intermolecular interactions with the small molecule inhibitor [54].

Until a few years ago, only the kinetic constants of rodents ALDH1A2 had been reported [55]. However, a study that reveals the kinetic parameters of human ALDH1A2 has recently been published [56]. Unlike what happens in rodent ALDH1A enzymes, where published results demonstrated that ALDH1A2 is the most efficient form in the irreversible oxidation of all-*trans*-retinaldehyde [29,30], there are no conclusive data on which is the most efficient form of human ALDH1A enzyme in the retinoic acid production. This is related with the fact that the different reports used a variety of methodologies to determine the kinetic parameters of human RALDHs with the physiological substrate all-*trans*-retinaldehyde.
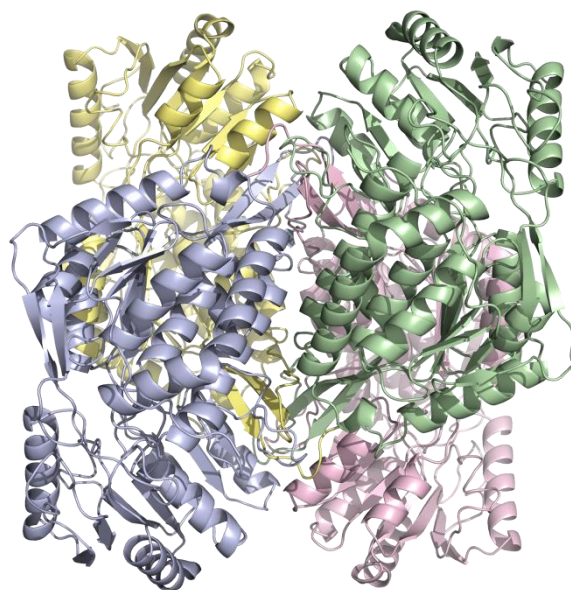
Figure 15. Human ALDH1A2 crystallographic structure (PDBe rendering based on 6ALJ). The image shows the homotetramer with the individually colored subunits in green, yellow, pink and blue.

ALDH1A2 may be a susceptible target for tumor suppression since it is expressed in epithelia from normal prostate but not in prostate cancer [57]. Furthermore, retinoic acid synthesis by ALDH1A2 promotes differentiation, cell growth arrest and apoptosis [58]. These findings implicate *Aldh1a2* as a candidate tumor suppressor gene in prostate cancer and further support a role of retinoids in the prevention or treatment of this cancer type. Besides, human ALDH1A2 is specifically involved in testicular retinoic acid biosynthesis, which is essential in spermatogenesis. Thus, ALDH1A2 is responsible for developing germ cells in men and is known that the protein levels were reduced in infertile men testes when compared with testicular tissue from men with normal spermatogenesis [59].

1.4.3 ALDH1A3

ALDH1A3 is a cytosolic homotetramer with 56-kDa subunits (Figure 16) and is expressed in an organ- and tissue-specific manner at low levels in most of them but in higher levels in the salivary gland, stomach, breast, kidney and fetal nasal mucosa [60,61]. ALDH1A3 is often referred to as the third retinaldehyde dehydrogenase, RALDH3. Similar to ALDH1A2, rodents ALDH1A3 participates in the retinoic acid biosynthesis and plays an important role in embryonic development [62]. A number of studies have demonstrated that ALDH1A3 deficiency may play a critical role in cancer. Thus, *ALDH1A3* is one of the genes that are upregulated by induction of wild-type p53 in cultured human colon cancer cells [63].

Figure 16. Human ALDH1A3 crystallographic structure (PDBe rendering based on 5FHZ). The image shows the homotetramer with the individually colored subunits in green, yellow, pink and blue.

Additionally, *ALDH1A3* is highly expressed in pancreatic [64] and ovarian [65] cancers but it is not expressed in the corresponding normal tissue samples. Moreover, *ALDH1A3* is methylation-silenced in gastric cancer cells [66] and induced by the antitumor agent interleukin 13 (IL-13) cytotoxin in glioblastoma cells [67]. In summary, *ALDH1A3* expression is regulated by many factors, and it is associated with the development, progression and prognosis of cancers. In addition, ALDH1A3 influences a diverse range of biological characteristics within cancer stem cells and acts as a marker for these cells. Thus, growing evidence indicates that ALDH1A3 has the potential to be used as a target for cancer diagnosis and therapy [68]. Besides the important role of ALDH1A3 in cancer regulation, is essential to mention the critical effect of ALDH1A3 in embryogenesis. Thus, a missense mutation in *ALDH1A3* caused anophthalmia/microphthalmia, rare developmental anomalies resulting in absent/small ocular globes, respectively [69].

1.4.4   ALDH2

The ALDH2 gene product is a mitochondrial matrix homotetramer with 56.4-kDa subunits (Figure 17). Based on the current nomenclature system, ALDH2 should be considered a member of the ALDH1A subfamily, since it shares more than 60% sequence identity. However, it was integrated in the ALDH2 family owing to the amount of literature regarding this enzyme as ALDH2. The human *ALDH2* has a broad expression pattern and is constitutively expressed in a variety of tissues including kidney, liver, heart, brain and lung [2].

Apparently, in human liver, only mitochondrial ALDH oxidizes acetaldehyde at physiological concentrations. In fact, ALDH2 is the major enzyme involved in the oxidation of acetaldehyde during ethanol metabolism [31]. It has been shown that about 50% of the Asian population is deficient in mitochondrial ALDH and these individuals possess a variant form of ALDH2, resulting from a single base pair mutation G/C → A/T, causing the E487K substitution and the consequent functional inactivation of ALDH2 under physiological conditions [70]. This variant, named ALDH2*2, induces adverse reactions in those who drink alcohol and alcohol toxicity occurs due primarily to acetaldehyde accumulation effects [71]. This fact implies a decrease in the alcoholism rate in Asian individuals in comparison with the global population [72]. However, it has been demonstrated that the *ALDH2*2* allele is robustly associated with the development of certain types of cancers, including oropharyngolaryngeal, esophageal, stomach, colon, lung, head and neck cancers [73,74].

ALDH2 is an important enzyme involved in the bioactivation of nitrates, acting as a nitrate reductase [75]. Related to the nitrates metabolism, ALDH2 is the major enzyme responsible for the activation of nitroglycerin, which is used in the treatment of angina and heart failure [76]. The variant ALDH2*2 is associated with an absence of nitroglycerin efficacy in Chinese patients [77], with myocardial infarction in Korean patients [78], and with hypertension in Japanese patients [79].

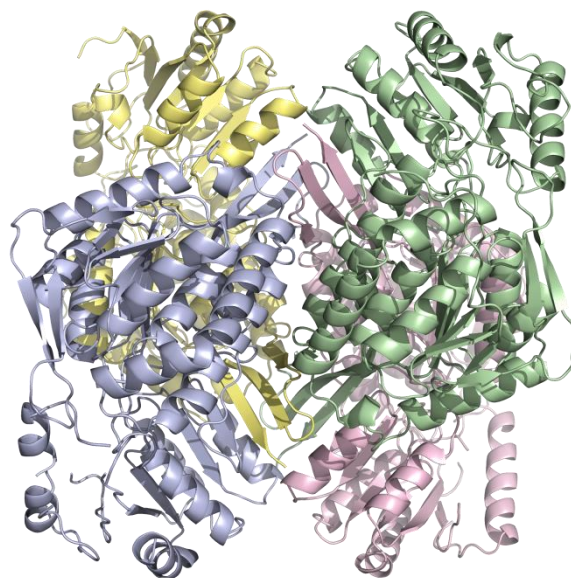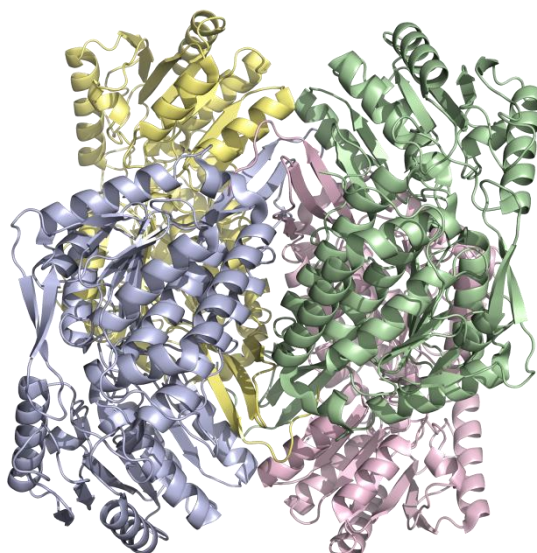**Figure 17. Human ALDH2 crystallographic structure (PDBe rendering based on 1CW3).** The image shows the homotetramer with the individually colored subunits in green, yellow, pink and blue.

1.4.5   Rate-limiting step and effect of magnesium ions

For the dehydrogenase reaction of human mitochondrial ALDH2, deacylation of the acyl intermediate is rate limiting, while the rate-determining step for the human cytosolic ALDH1A1 is the NADH release [80]. This general statement is accepted with nuances. The establishment of the rate-limiting step in the various members of the ALDH superfamily has been directly associated to the effect of divalent cations, specifically Mg^{2+}, on the enzymatic activity of ALDHs. The pioneering reports on this field were carried out with horse liver mitochondrial ALDH, in which it was possible to determine by steady-state kinetic analysis, that in the absence of Mg^{2+} ions, the enzyme exhibits half-of-the-sites reactivity functioning as a dimeric enzyme, while in presence of magnesium ions, the tetrameric form of the enzyme possesses all-of-the-sites reactivity. In addition, the metal is responsible for the dissociation of tetramer in dimers [81]. It seems that the activation mechanism of the enzyme by Mg^{2+} ions could be associated with the change in the number of functional subunits, rather than with an alteration of the catalytic properties of existing active sites. For the enzyme-catalyzed hydrolysis of *p*-nitrophenyl acetate, formation of the acyl intermediate is the rate limiting step [82].

However, in the presence of NAD^+, the deacylation step for the esterase reaction also becomes rate limiting [81]. Saturating concentrations of either NAD^+ or NADH increased both the Michaelis constant and V_{max} values for *p*-nitrophenyl acetate. NAD^+ produced the most striking increase in the Michaelis constant, more than 2-fold, while NADH produced a smaller but significant increase [36].

In other species, the effects of divalent cations on the ALDH enzymatic activity have been described. The sheep liver cytoplasmic ALDH is strongly inhibited by Mg^{2+}, Mn^{2+} and Ca^{2+} ions. In particular, concentrations in the micromolar range of these divalent cations imply a potent inhibition of the enzyme but do not produce a change in the number of functional subunits [83]. Pre-steady-state kinetic studies displayed that magnesium ions had no effect on the NAD^+ and NADH binding ratios. However, the dissociation constants for $E·NAD^+$ and $E·NADH$ were significantly decreased in the presence of $MgCl_2$ [84]. The esterase activity of sheep liver cytoplasmic enzyme with *p*-nitrophenyl acetate as a substrate, is not altered by Mg^{2+} concentrations that exhibit a total inhibitory effect on dehydrogenase activity [85]. The esterase activity of horse liver enzyme is not altered by Mg^{2+}, Ca^{2+} or other divalent ions [86]. The dehydrogenase activity of human mitochondrial ALDH2 was increased 2-fold in the presence of low Mg^{2+} ion concentrations (120-140 µM) [87]. An increase in the K_m value for

NAD$^+$ was observed simultaneously to the increase of V$_{max}$ when the concentration of Mg^{2+} ions was raised.

Thus, it is reliable to state the appearance of alterations in the catalytic properties of the enzyme. The pre-steady state burst observed for NADH was increased in the presence of Mg^{2+}, suggesting that the rate-limiting step of the dehydrogenase reaction is altered by Mg^{2+} ions. In the human mitochondrial enzyme, no alterations were found in the molecular weight properties, which indicates that the mechanism by which the human mitochondrial ALDH is activated by Mg^{2+}, is different from that of the horse mitochondrial enzyme, with the dissociation of the tetramer into dimers [87].

The human cytosolic ALDH1A1 is inhibited by magnesium ions while the human mitochondrial ALDH2 is activated. In the cytosolic form, where the NADH cofactor release is the rate-limiting step of the reaction, the Mg^{2+} ions are responsible of the binding between enzyme and cofactor in the binary complex formed. The effect mentioned above is explained by the fact that the dissociation constants for both NAD$^+$ and NADH cofactors were significantly decreased in the presence of MgCl$_2$ [80]. In turn, the human mitochondrial enzyme is activated by magnesium ions since the rate-determining step of the reaction is the deacylation, in which the carboxylic acid generated is released from the ternary complex enzyme-NADH-carboxylic acid [88].

1.5   Retinoids

The term "retinoid" is used to refer to a wide range of compounds derived from vitamin A (retinol), were they natural o synthetic, with or without biological activity. Retinoids were defined as molecules composed of four isoprenoid units joined head-to-tail. Thereby, both the precursors and the derived products of retinoids may be monocyclic compounds containing five conjugated double bonds and a functional group at the distal end of the acyclic portion of the molecules [89]. Vitamin A and its derivatives have many important functions throughout the organism including roles in visual cycle (acting as chromophores) and participate in many essential biological processes such as cell proliferation and differentiation, fetal and adult tissue development and activation of tumor suppressor genes. Some retinoids are very irritating to the skin and some are highly teratogenic. Hence, when large amounts of retinoids are handled, adequate care should be taken to avoid their undesirable toxic effects.

1.5.1 Chemical structure of retinoids

The general structure of retinoids is divided into three domains: a trimethylated cyclohexene ring named β-ionone, a conjugated tetraene side chain, and a polar group which may exhibit different oxidation states: alcohol, aldehyde and acid (resulting in retinol, retinaldehyde and retinoic acid, respectively, Figure 18). In the parent retinoids, all the double bonds are in a *trans* (or *E*) configuration. The different configuration (*cis* or *Z* configuration) of the double bonds of the aliphatic chain leads to the formation of isomers, which are described by using the lowest numbered carbon in the double bond affected. Retinoids are unstable molecules being readily oxidized and/or isomerized to altered compounds, especially in the presence of oxidants including air and excessive heat. They are labile toward strong acids and solvents containing oxygen or peroxides and are degraded by light, which yields to isomerization of double bonds. For this reason, they have to be manipulated under inert atmosphere and red or dim light [90].

$1R = CH_2OH$
$2R = CHO$
$3R = COOH$
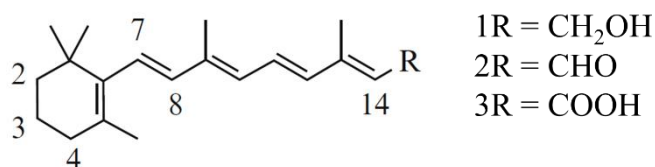
Figure 18. Chemical structure of retinoids. The carbon numbering and the carbon-oxygen functional group (1R: retinol; 2R: retinaldehyde; 3R: retinoic acid), are shown.

1.5.2 Vitamin A absorption, transport and metabolism

Vitamin A is an essential micronutrient and hence must be acquired from the diet, either as retinyl esters or retinol from animal sources, or as provitamin A carotenoids from fruits and vegetables, mainly as β-carotene [91]. The two most abundant preformed retinoids in the diet are retinol and retinyl esters. Provitamin A carotenoids are absorbed intact by the intestine. Thus, β-carotene can be directly absorbed by passive diffusion or through facilitated transport into the cells, being able to accumulate in blood and tissues. However, retinyl esters are unable to enter the intestinal mucosa and must be hydrolyzed by a luminal retinyl hydrolase to yield free retinol or can be also hydrolyzed in the intestinal lumen by non-specific pancreatic enzymes [92]. In the enterocyte, β-carotene is converted to retinaldehyde by β-carotene-15,15'-oxygenase 1 (BCO1), which cleaves β-carotene at its central double bond (15,15') to yield two symmetric molecules of all-*trans*-retinaldehyde [93]. A significant amount of β-carotene enters uncleaved in the circulation and can be metabolized by BCO1 in the peripheral tissues [91]. Inside the enterocyte, all-*trans*-retinaldehyde is then reduced by one or more retinaldehyde reductases, such as SDRs and AKRs, to retinol, which can bind to cellular retinol-binding protein type II (CRBP2). Besides, retinol is then esterified with long-chain fatty acids

through the action of lecithin:retinol acyltransferase (LRAT) for incorporation into chylomicrons and it is transported to the liver parenchymal cells, where it is captured by specific receptors and transferred to stellate cells for storage (Figure 19). CRBP2 plays a primary role in the regulation of retinol absorption in the enterocytes and its intercellular metabolism [94,95].

When retinoids are required in the peripheral tissues, retinyl esters are hydrolyzed to retinol by retinyl ester hydrolase (REH) [96]. Then, retinol is transported into circulation bound to retinol binding protein (RBP). The retinol-RBP complex is secreted from the hepatocyte into the circulation to allow for retinol delivery to retinoid-dependent peripheral tissues and is found in the blood as a complex with the thyroxine-binding protein transthyretin (TTR), which stabilizes the retinol-RBP complex, thereby reducing renal filtration of RBP and allowing the RBP recycling when retinol is taken into cells [97]. The transport of retinol across cell membranes is mediated by a multi-transmembrane domain protein that represents a new type of cell-surface receptor *stimulated by retinoic acid 6* (STRA6), which has nine transmembrane domains, five extracellular domains and four intracellular domains [98]. There are numerous hypotheses on the mechanism of STRA6-mediated cellular vitamin A uptake from RBP [99–101] but is well established that this process is not an example of primary active transport because STRA6 has no ATP-binding domain. A remaining possibility is a channel/facilitative transport, which depends on the electrochemical gradient of the free ligand. However, retinol is not a free substrate but is bound with high affinity to RBP (Figure 19). In summary, the mechanism by which retinol penetrates across cell membranes is subject to controversy, since the STRA6 mechanism is distinct from all known cellular uptake mechanisms [100].

1.5.3 Biosynthesis of all-*trans*-retinoic acid from retinol

Retinoids are transported and stored mainly in the form of retinol and retinyl esters, respectively. However, retinoic acid represents the actual form of vitamin A with the most potent biological activity. The major established pathway of retinol activation involves mobilization of retinyl esters, reversible conversion of the released retinol into retinaldehyde, and irreversible conversion of retinaldehyde into retinoic acid [102].

The biosynthesis of retinoic acid from retinol involves two oxidative reactions. Firstly, retinol is oxidized to retinaldehyde and then, a second oxidation reaction takes place, in which the irreversible oxidation of retinaldehyde to retinoic acid occurs. The first step, the oxidation of retinol to retinaldehyde, is generally considered to be rate limiting [103]. A wide variety of

enzymes, including members of several oxidoreductase superfamilies, participate in these two oxidative steps: SDRs, ADHs, AKRs and ALDHs (Figure 19). The SDR superfamily includes a group of microsomal enzymes that can oxidize retinol (NAD^+-dependent forms) and reduce retinaldehyde ($NADP^+$-dependent forms). Different families of SDRs have been implicated in the regulation of retinoic acid homeostasis. Interestingly, although the enzymes from different families share the ability to recognize retinol and retinaldehyde as substrates, their protein sequence are quite diverse (30% overall sequence identity). These significant differences in primary structures are reflected in the distinct profiles of substrates and cofactor specificities characteristic for each type of retinoid-active SDRs [104]. The retinol dehydrogenase 10 (RDH10) is the major SDR enzyme responsible for retinoic acid biosynthesis in embryo [105]. Although RDH10 is bidirectional *in vitro*, in the cells the enzyme acts unidirectionally in the oxidative direction, increasing the levels of retinaldehyde and retinoic acid, promoting cell differentiation and development [103]. The cytosolic retinoid-active ADHs are NAD^+-dependent enzymes that oxidize different retinoid isomers *in vitro*. The current consensus appears to be that the ADH enzymes may contribute to the oxidation of retinol postnatally in specific tissues during vitamin A excess (ADH1) or deficiency (ADH4). Specifically, ADH1 and ADH2 are the major medium-chain dehydrogenases/reductases (MDR) in liver retinol detoxification, while ADH3 (less active) and the most active ALDH4 participate in retinoic acid biosynthesis in tissues [106]. The spatiotemporal expression pattern of ADH1 and ADH4, and their co-localization with the onset of retinoic acid signaling in the embryonic development of mouse, suggest an active role of these enzymes in the synthesis of retinoic acid [107,108]. AKRs are NADPH-dependent enzymes that display activity with retinoids and have a clear preference for retinaldehyde reduction. The human activity was found for the first time in AKR1B1 and AKR1B10 enzymes [109–111]. It is well established that AKR1B10 is a diagnostic marker of non-small cell lung cancer, where it is found overexpressed. Moreover, an overexpression of AKR1B10 would decrease the amount of retinoic acid that could be formed, promoting metaplasia [112]. Thus, the high efficiency of AKR1B10 to reduce retinaldehyde has been proposed as a possible mechanism involved in the development of various types of carcinoma, and as a protein that could act at pre-receptor stage, regulating the amount of retinoic acid available for retinoic acid receptors (RAR) and retinoid X receptors (RXR) [113]. The irreversible NAD^+-dependent oxidation of all-*trans*-retinaldehyde to retinoic acid is carried out by members of the ALDH superfamily (Figure 19). In humans, it is unknown which ALDH displays the highest catalytic efficiency in retinaldehyde conversion to retinoic acid. Thereby, an objective of this Thesis will be to uncover this incognita. However, some studies have revealed that ALDH1A2 is the enzyme that displays the greatest catalytic efficiency in rodents

[114]. The physiological role of cytochrome P450 family 26 (CYP26) in the irreversible oxidation of retinoic acid to degradation products seems to exhibit significance. In addition, cytochrome P450 family 1A1 (CYP450 1A1) is highly effective in hydroxylating both all-*trans*- and 13-*cis* retinaldehyde, indicating that this isoform may be responsible for much of the P450-mediated retinoid metabolism [115].

Our group compared the catalytic constants with retinoids of members of the SDR, MDR and AKR superfamilies, determined with the same methodology that is solubilization of retinoids in presence of bovine serum albumin (BSA) and detection by HPLC, for the first time [109]. To complete the study of the metabolic pathway of retinoids, this Thesis aims to perform an exhaustive and robust structural and kinetic analysis on the enzymes involved in the irreversible oxidation of retinaldehyde to retinoic acid.

1.5.4   Retinoic acid

Almost 70 years ago, retinoic acid was defined as the biologically active form of vitamin A. However, its role and mechanism of action still generate an important interest in the clinical and scientific fields [116,117]. Once retinoic acid is formed within the cell, it is mainly bound to the cellular retinoic acid binding proteins (CRABPs) and transported to the nucleus (Figure 19). CRABP I and II transport only all-*trans*-retinoic acid, which has a higher affinity for CRABP I [118]. Retinoic acid regulates gene expression and exerts its pleiotropic effects mainly through the activation of two classes of nuclear ligand-dependent activators, RARs and RXRs, both with their respective subtypes α, β and γ. While RARs are activated by both all-*trans*- and 9-*cis*-retinoic acid, the RXRs are exclusively activated by the 9-*cis* isomer [119]. Whereas RAR only can form homodimers or heterodimers with another retinoic acid receptor (RAR itself or RXR), RXR can dimerize with other nuclear receptors: chicken ovalbumin upstream promoter-transcriptional factor II (COUPTFII) and thyroid hormone receptor (TR), among others. Once the receptor has bound retinoic acid, it forms a dimer and it becomes active. These dimers bind to retinoic acid response elements (RAREs) located in the promoters of retinoic acid target genes. Ligand binding induces a conformational change in the receptor homodimers and heterodimers which promotes gene transcription. RAREs can be occupied by RAR, RXR, COUPTFII, hepatocyte nuclear factor 4α (HNF4α) and peroxisome proliferator-activated receptor β/δ (PPAR β/δ) as homo or heterodimers, and their complex interaction can lead to significant differences in the expression of their target genes in various tissues [120,121].
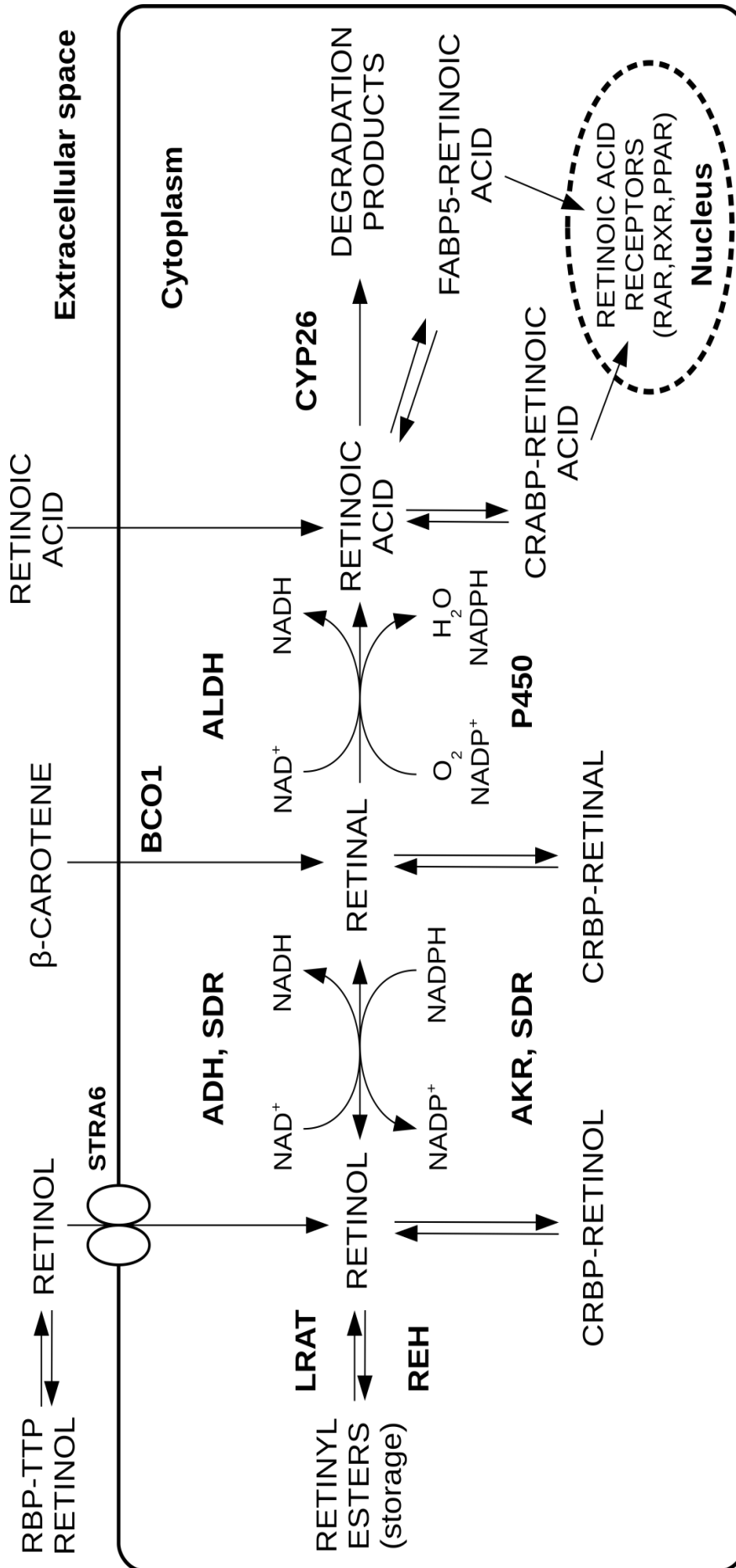
Figure 19. Proposed mechanism for retinoic acid synthesis and retinoid storage. Cellular uptake of retinol, carried by retinol-binding protein (RBP) complexed with thyroxine-binding protein transthyretin (TTR), of β-carotene and retinoic acid is indicated. In the target tissues, STRA6, a cell-surface receptor for RBP, facilitates retinol entry. Inside the cell, retinol or retinaldehyde may bind cellular retinol-binding protein (CRBP) and be esterified by lecithin:retinol acyltransferase (LRAT) for storage. Retinoic acid may bind cellular retinoic acid-binding protein (CRABP) or fatty acid-binding protein (FABP5) in keratinocytes. The reversible interconversion of retinol and retinaldehyde is catalyzed by members of ADH, SDR and AKR enzyme families. The cofactor preferences of the enzymes involved are indicated. The second and irreversible step of the pathway, the oxidation of retinaldehyde to retinoic acid, is catalyzed by ALDHs and cytochrome P450. Reproduced from [200].

Abundant evidence shows that retinoic acid, via various RAR signaling pathways, inhibits cell-cycle progression in a variety of human cancer cells. Generally, retinoic acid causes a block in the in the G1 phase of the cell cycle acting as a tumor suppressor. Moreover, retinoic acid induces apoptosis in various human cancer cells, such as melanoma and ovarian adenocarcinoma, hepatoma, breast and lung cancer. Accordingly, based on their ability to promote cell-cycle arrest and induce the process of apoptosis, retinoid acid could be useful drug for the treatment of numerous human cancers [122,123]. Moreover, retinoic acid enhances mitosis and cell proliferation in keratinocytes and oligodendrocytes, through the PPAR β/δ signaling pathways.

Thus, retinoic acid could inhibit cell growth, as in cancer cells, by binding to RAR, and stimulate, in keratinocytes, cell proliferation with antiapoptotic actions binding through PPAR β/δ [124]. Besides, retinoic acid is also involved in regulating lipid and sugar metabolism and insulin responses. Hence, the signaling pathways in adipose tissue and muscle results in upregulation of multiple protein that enhance lipid oxidation, mitochondrial respiration and sensitization of cells to insulin signaling [125]. Finally, retinoic acid could decrease the severity of tissue inflammation with therapeutic effects. It was demonstrated that retinoic acid is required for terminal differentiation of myeloid cells such as neutrophils, and it is essential in the optimal maintenance of phagocytic activities. Therefore, retinoic acid is necessary in those tissues that constantly are providing stimulatory signals for the immune system [126].

1.5.5   Carotenoids: biosynthesis, metabolism and physiological role

Carotenoids are lipophilic pigments synthesized by plants, microorganisms and some animals but not by humans. They consist of polyisoprenoid structures typically containing a series of conjugated double bonds in the central chain of the molecule, which makes them susceptible to oxidative cleavage and isomerization from *trans* to *cis* forms, with the formation of potentially bioactive metabolites. They can be broadly divided into two classes of chemical compounds: carotenes (e.g., β-carotene and lycopene) and their oxygenated derivatives termed xanthophylls (e. g., lutein, zeaxanthin and cryptoxanthin).

β-Carotene is a naturally-occurring retinol precursor and it can be metabolized in mammals via two enzymatic pathways (Figure 20). Briefly, β-carotene can be cleaved symmetrically by β-carotene 15,15'-oxygenase (BCO1) yielding two retinaldehyde molecules, that can be further oxidized to all-*trans* retinoic acid by ALDHs (as described above in section 1.7.2). For provitamin A carotenoids, such as β-carotene, α-carotene and β-cryptoxanthin, central

cleavage is a major pathway leading to vitamin A formation [127]. The second pathway of β-carotene cleavage is called the eccentric cleavage and it occurs by β-carotene 9',10'-oxygenase (BCO2). It acts on 9',10' double bond of the polyene chain of β-carotene to produce β-apo-10'-carotenal and β-ionone [128] (Figure 20). Eccentric cleavage at other double bonds, catalyzed or non-catalyzed, seems also possible (Figure 20). Within carotenoids, the term apo-carotenoid is used to designate those with a backbone of less than 40 carbon atoms [129]. β-carotene is the most effective provitamin A carotenoid in the context of BCO1 cleavage, since the enzyme catalyzed the oxidative cleavage of β-carotene with K_m = 17 µM and catalytic efficiency, k_{cat}/K_m = 6098 $M^{-1} \cdot min^{-1}$. The enzyme also catalyzed the oxidative cleavage of compounds which contain at least one non-substituted β-ionone ring, such as α-carotene, β-cryptoxanthin and 8'-apo-β-carotenal, to yield retinaldehyde.

The catalytic efficiency values of these substrates are lower than that of β-carotene. With the shorter apo-β-carotenals (10'-apo-β-carotenal, 12'-apo-β-carotenal and 14'-apo-β-carotenal), the enzyme does not show Michaelis-Menten behavior [130]. Although BCO1 functions as a major enzyme in vitamin A production, BCO2 is considered an alternative pathway for forming vitamin A. Moreover, BCO2 displays a much broader substrate specificity for carotenoid metabolism as compared to BCO1, since lutein, lycopene, β-cryptoxanthin and zeaxanthin could be cleaved by BCO2, leading to the formation of β-ionone and 10'-apo-β-carotenal. Interestingly, the production of 10'-apo-β-carotenal was significantly affected when the concentration of ferrous iron decreased. Thus, iron is a necessary cofactor for this enzymatic cleavage [127].

Different human BCO2 isoforms possess a cleavable N-terminal leader sequence critical for mitochondrial import. However, β-carotene is retained predominantly in the cytoplasm. Accordingly, there is some evidence confirming a compartmentalization of carotenoid metabolism that prevents competition between BCO1 and BCO2 for the provitamin and the noncanonical β-carotene metabolites [131].

Apo-β-carotenoids have been detected in some foods and in the blood of both human and experimentation animals. The 8'-apo-β-carotenal was detected in plasma after ingestion of β-carotene in a healthy human subject [132], while another assay found that 13'-apo-β-carotenone was present in a human plasma [133]. The biological functions of apo-β-carotenals and apo-β-carotenones in mammals are partially unknown. However, some studies have demonstrated that these compounds could interact directly with the ligand binding site of the

retinoid receptors, establishing a high affinity competition with retinoic acid for binding to the receptors.
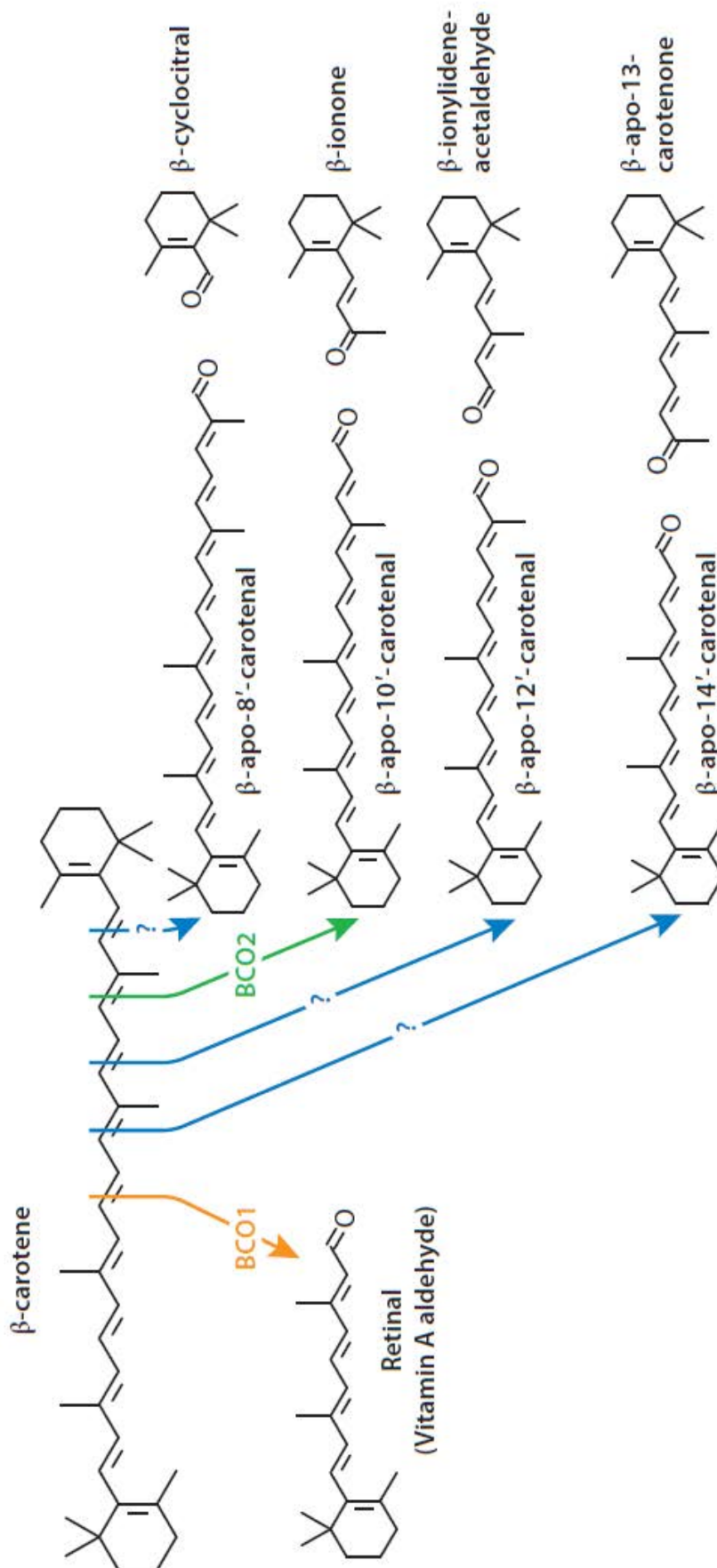
Figure 20. The cleavage products of β-carotene. A) Symmetric cleavage of β-carotene. β-carotene 15,15′-oxygenase (BCO1) catalyzes the central cleavage of β-carotene to yield two molecules of all-*trans*-retinaldehyde. B) Eccentric cleavage of β-carotene. β-carotene 9′,10′-oxygenase (BCO2) catalyzes the asymmetric cleavage of β-carotene, from which β-apo-10′-carotenal and β-ionone compounds are generated. Other compounds of different chain length are described, such as β-apo-carotenals and β-apo-carotenones, although its enzymatic production is unknown. The scheme is reproduced from [128].

- 47 -

In a recent study, the authors observed a progressive shift in the all-*trans*-retinoic acid dose-response curve with increasing concentrations of 13'-apo-β-carotenone in the nanomolar range. Higher concentrations of all-*trans*-retinoic acid were able to overcome inhibition by 13'-apo-β-carotenone, suggesting direct competition between the two compounds for binding [133].

In turn, 14'-apo-β-carotenal, but no other structurally related apo-β-carotenals, represses PPAR and RXR activation, and biologic responses induced by their respective agonists both *in vitro* and *in vivo*. Specifically, during adipocyte differentiation, 14'-apo-β-carotenal inhibited PPARγ target gene expression and adipogenesis and suppressed PPARα responses, including target gene expression and its anti-inflammatory effects [134]. Recently, it has been described that the BCO2 activity towards β-carotene would act as a defense mechanism in front of the induced damage by β-carotene, which could impair mitochondrial respiration and induce oxidative stress [135]. In addition, apo-β-carotenals are a retinoic acid source through oxidation to acid or by β-oxidation [136]. Enzymes for the oxidation to acid have not been described, but ALDH1A forms are possible candidates. In this regard, it is important to note that this Thesis described for the first time an enzymatic in *vitro* assay with the purified enzymes from the ALDH1A subfamily to determine their kinetic parameters with apo-β-carotenals.

1.6 Physiological role of ALDH

Most of the ALDHs have a wide tissue distribution and display distinct substrate specificity and different type of functions (Figure 21). The physiological role of some ALDHs, such as ALDH1A1, ALDH1A2 and ALDH1A3, involves the conversion of several metabolites that are essential in the organism. Specifically, these enzymes catalyze the irreversible oxidation of retinaldehyde to retinoic acid [114,137,138]. Considering the role of retinoic acid, it should be noted that retinoic acid isomers, all-*trans* and 9-*cis*-retinoic acid, serve as ligands for RAR and RXR, which modulate the gene expression for growth and development [139,140].

The generation of γ-aminobutyric acid (GABA) is an additional relevant function of ALDHs with regard to regulation of many critical metabolic pathways [141]. In addition to enzymes from ALDH1A subfamily, other ALDHs such as ALDH3A1 and ALDH2 catalyze ester hydrolysis [142] and the mitochondrial form also exhibits nitrate reductase activity, leading the production of cyclic guanosine monophosphate (cGMP) that results in vasorelaxation [143]. Among the non-catalytic activities of ALDHs, other functions for ALDH1A have been described

such as a binding protein to cholesterol, androgens, thyroid hormone [3], flavopyridol, daunorubicin and acetaminophen [144]. Together with other ALDHs, such as ALDH3A1, the members of ALDH1A subfamily, ALDH1A1, ALDH1A2 and ALDH1A3, generate NADH (NADPH in the ALDH3A1 instance) contributing to cellular homeostasis and maintaining redox balance [145].

Some allelic variants in *ALDH* genes are associated with distinct pathological conditions in humans and rodents [146], including Sjögren-Larsson syndrome, type II hyperprolinemia, γ-hydroxybutyric aciduria, cancer and alcohol-related disease. Finally, it was demonstrated the structural function of ALDH3A1 and ALDH1A1 in the mammalian cornea and lens crystallins [147].

Functions of ALDHs

- **Catalytic** ⟶
  - **Aldehyde oxidation**
  - **Ester hydrolysis**
  - **Nitrate reductase**
- **Binding** ⟶
  - **Endobiotics**
  - **Xenobiotics**
- **Antioxidant** ⟶
  - **NAD(P)H production**
- **Structural** ⟶
  - **Cornea and lens crystalline**
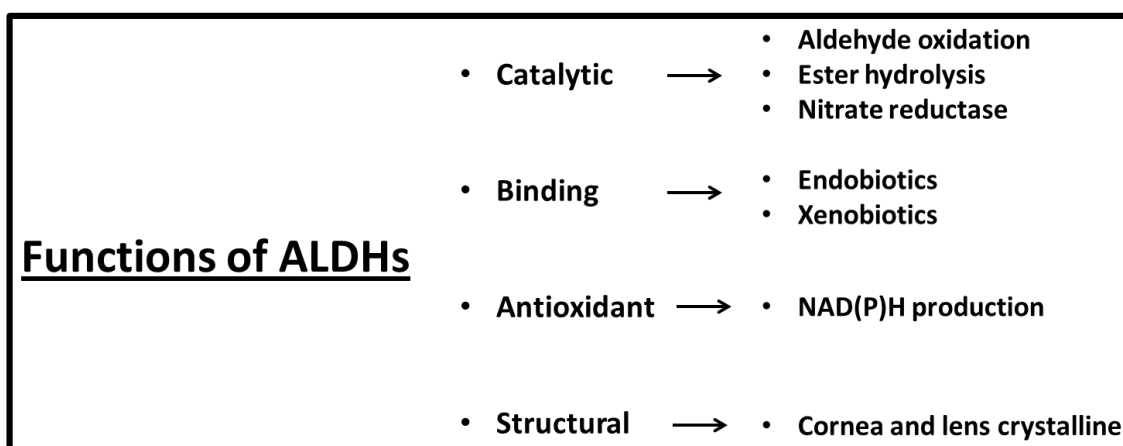
Figure 21. Multiple functions of ALDHs enzymes. The scheme shows the functional diversity of the ALDH superfamily of enzymes, which participate in catalysis and act as binding proteins of endobiotic and xenobiotic compounds. In addition, ALDHs exhibit antioxidant effects, through the production of NAD(P)H, and structural support functions in the ocular system.

1.7   Role of ALDH1A enzymes in embryogenesis

At least the three RALDHs have been identified in human, mouse and *Xenopus*, with different physiological functions in embryogenesis [148]. *ALDH1A1* is highly expressed in the dorsal retina of mouse embryo [52]. The knockout of *ALDH1A1* did not severely affect the morphology of the retina [149], indicating that other enzymes might redundantly share the function of ALDH1A1.

ALDH1A2 was identified in human, mouse, chick, zebrafish and *Xenopus* embryonic tissues [150]. Genes belonging to retinoic acid metabolic and signaling pathways are targets for investigation as direct causal agents, or modifiers of congenital heart disease. Among ALDH1As, ALDH1A2 is the major form involved in early embryonic and in cardiac development [151]. Non-conservative *ALDH1A2* mutations are associated with rare cases of human

congenital heart disease [152]. However, it is also well known that the *ALDH1A2* haploinsufficiency causes a defective retinoic acid signaling in mice, which can develop other non-cardiac defects such as vascular, laryngeal, tracheal, thymus and parathyroid alterations [153]. Interestingly, the most severe effect of retinoic acid altered expression results from the genetic deficiency of ALDH1A2, causing the mice lethality at 8.5 days post-fertilization [55]. The *ALDH1A2* knock-out mice develop severely impaired segmentation of rhombomeres with altered homeobox gene expression pattern, and neural crest cell migration [154].

ALDH1A3 has been identified in human, chick, mouse, zebrafish and *Xenopus* embryonic tissues, and is expressed in the ventral retina across various species [155,156]. ALDH1A3 is essential for the development of the central nervous system and the morphogenesis of anterior head structures. Mice lacking *ALDH1A3* have defects in nasal and ocular development and are neonatally lethal, due to respiratory tract obstruction in the nasal region, but they can be rescued by retinoic acid administration during pregnancy [157]. Recently, several groups have reported that *ALDH1A3* is associated with anophthalmia/microphthalmia in humans [69].

1.8 ALDH inhibitors

Related to the involvement of ALDHs in some pathological affections, the study of pharmacological inhibitors of these enzymes supposes a significant clinical interest. Human ALDHs show an overlapping spectrum for substrates and distinct substrate specificities, and this fact implies a limitation about the enzyme-specific effects on the course of diseases. It has to be noted, that pharmacological inhibitors have been developed for few ALDH enzymes, such as the enzymes specifically involved in the metabolism of alcohol (ALDH2), and in the chemoresistance against the anticancer oxazaphosphorine drugs (ALDH1A1 and ALDH3A1) [35].

Citral

Citral (3,7-dimethyl-2,6-octadienal, Figure 22) is a volatile α,β-unsaturated aldehyde associated with a variety of biochemical and toxicological effects. This compound acts more as a slow substrate than as a strict inhibitor of ALDH enzymes [158]. Citral is a 1:2 mixture of two isomers: the *cis* isomer, neral, and the *trans* isomer, geranial. The latter structurally resembles physiologically important retinoids and it is the preferred isomer for the inhibition of the ALDH1 enzymes. For ALDH2, neral shows Michaelis-Menten kinetics and geranial shows positive cooperativity [159]. Citral exhibits a reversible non-competitive inhibition and the Ki value for rat ALDH2 is 360 nM [160]. Furthermore, citral has been shown to possess anti-

proliferative and pro-apoptotic effects in breast cancer cells lines as a result of ALDH1A1 inhibition [161].

DEAB

N,N-diethylaminobenzaldehyde (DEAB, Figure 22) behaves as a slow substrate (like citral) rather than a classical inhibitor of ALDH enzymes. It has been used as an ALDH1-selective inhibitor for repressing retinoic acid synthesis in studies investigating the role of retinoic acid in abnormal embryonic development in vertebrates [162]. Furthermore, the ALDH inhibition by DEAB is the basis of Aldefluor™ assay, which is used to identify, evaluate, and isolate normal and cancer stem cells, that express high levels of ALDH. It is widely presumed that the assay mostly measures ALDH1A1 activity based on the early report that DEAB used in this assay is a selective inhibitor of ALDH1A1 [163]. However, a recent study about the effects of human ALDH enzymes, ALDH1A2 and ALDH2, on drug resistance and proliferation, and the selectivity of DEAB as an inhibitor concluded that DEAB is not a selective inhibitor against ALDH1A1 and that Aldefluor™ assay is not specific for the ALDH1A1 activity (IC_{50} = 60 nM), since it also inhibits ALDH1A2 (IC_{50} = 3 μM) and ALDH2 (IC_{50} = 160 nM) enzymes, which play a major role in the biology and drug resistance of various malignant cells [164]. In ALDH1A1, DEAB exhibits a competitive tight-binding inhibition, while in ALDH1A2 and ALDH2 DEAB acts as a covalent inhibitor [165]. The mechanism of ALDH inactivation by DEAB is consistent with the formation of a quinoid-like resonance state following hydride transfer that is stabilized by local structural features that exist in several of the ALDH enzymes [165].

WIN 18,446

N,N'-Bis(dichloroacetyl)-1,8-octamethylenediamine (WIN 18,446, Figure 22) reversibly inhibits spermatogenesis in many species, including humans, via inhibition of testicular retinoic acid biosynthesis catalyzed by ALDH1A2 [166]. WIN 18,446 strongly and irreversibly inhibits ALDH1A2 *in vitro* (IC_{50} = 0.19 μM) [54]. *In vivo*, WIN 18,446 treatment completely abolished spermatogenesis after 4 weeks of treatment through an effect that is tissue-dependent. This effect is similar to that of the non-selective inhibitor disulfiram. However, only disulfiram-induced inhibition was reversed by addition of antioxidants such as dithiothreitol (DTT) and tris(2-carboxyethyl)phosphine (TCEP), suggesting that the mechanism of inhibition of these two compounds is different [167]. Recently, the tridimensional structure of ALDH1A2 associated with WIN 18,446 (Table 1) allowed to confirm that the nucleophile Cys320 is critical for the interaction with reversible and covalent inhibitors and the active-site is capable to accommode chemically and structurally diverse small compounds [54].

Recently, Morgan and collaborators [26] reported the first human ALDH1A1 structure. Structural comparisons of the cofactor-binding sites in ALDH1A1 with other closely related ALDH enzymes illustrate a high degree of similarity. Currently, there are no commercially available inhibitors that target ALDH1A1. In the study cited above, Morgan and colleagues described 256 compounds that alter the esterase activity of ALDH1A1. The effects on ALDH function were also intensively analyzed. However, only two distinct chemical classes of inhibitors are selective for human ALDH1A1 compared to eight classes for other ALDH enzymes [168]. Among these compounds, CM026 and CM037 have emerged and have been used to obtain new crystallographic structures of ALDH1A1 (Table 1).

CM026

CM026 (Figure 22) is a selective inhibitor against ALDH1A1 [168]. At low concentrations (20 µM), this compound had no effect on other human ALDH enzymes, such as ALDH1A2, ALDH1A3, ALDH1B1, ALDH2, ALDH3A1 and ALDH5A1. At 100 µM, CM026 discreetly increases the ALDH1A2 activity in counterpart to what it is observed in this enzyme with the Alda-1 activator [169]. CM026 has good potency against ALDH1A1 (IC_{50} = 0.80 µM), although a complete inhibition was not observed [26]. The compound has a noncompetitive partial mode of inhibition with respect to varied substrate and an uncompetitive partial mode of inhibition with respect to varied NAD^+ cofactor [168].

CM037

CM037 (Figure 22) is a selective inhibitor against ALDH1A1 (IC_{50} = 4.6 µM) [168]. Unlike CM026, this compound displays some effect on ALDH1A3 at low concentrations. Specifically, 20% inhibition is observed in ALDH1A3 at 20 µM of CM037, while in the other enzymes tested the effect is very low. Elevated CM037 concentrations have been difficult to analyze due to solubility limitations of this compound [168].

In summary, there have been few findings regarding enzymatic inhibitors of the other members of ALDH1A subfamily, being WIN 18,446 mentioned above, the only potent inhibitory compound described for human ALDH1A2. As for ALDH1A3, no selective and potent inhibitors exist so far.
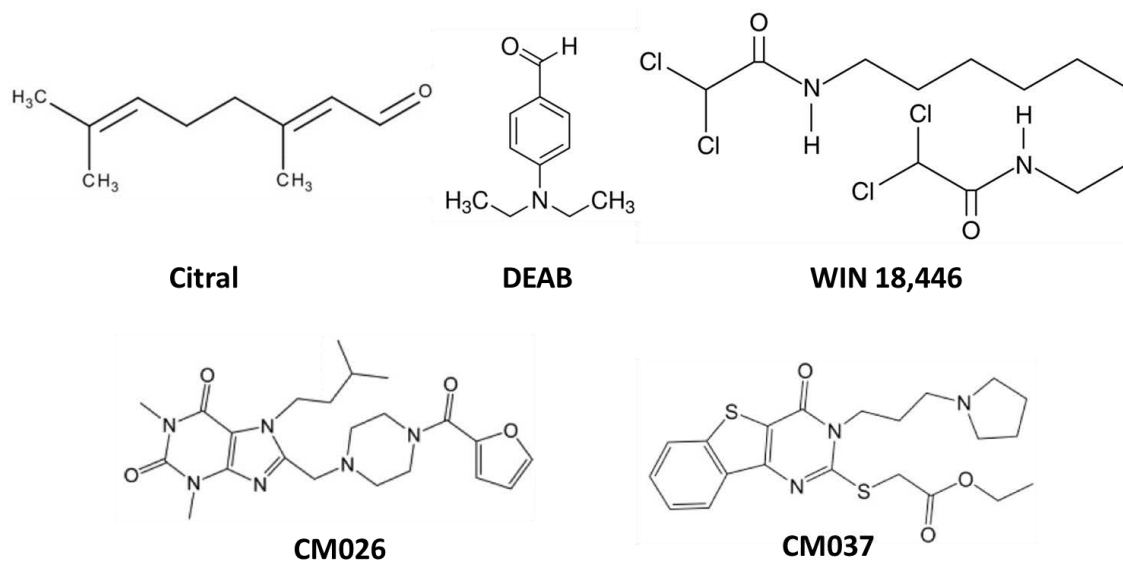
Figure 22. Molecular structure of some ALDH inhibitors.

- 54 -

OBJECTIVES

The recent availability of the three-dimensional structures of the human ALDH1A enzymes will allow performing a precise analysis of their substrate-binding pockets, in regard to their size, topology and specific residues. A first aim of this study is to identify clue amino acid residues that may help explaining the distinct substrate specificity and kinetic properties of each enzyme, supported by site-directed mutagenesis studies.

The retinoid-active ALDHs had been purified and partially characterized in several mammalian species, mainly from rodents, by independent investigators. However, not a single study had reported the kinetic properties of the three human ALDH1A enzymes with retinoids, using the same methodology and under the same experimental conditions. As a second major goal, a side-by-side kinetic characterization of the three enzymes will be performed. In addition, a well-proven method for retinoid solubilization and a sensitive HPLC-based method for activity determination will be applied to study the retinaldehyde dehydrogenase activity.

The physiological relevance of the ALDH1A enzymes stems out from their proved role in retinoic acid biosynthesis, embryonic development, cancer stem cell maintenance and several human pathologies. Thus, it becomes of prime importance to tackle the design of novel enzymatic inhibitors which might be eventually developed as potential pharmacological drugs. Despite the existence of many ALDH inhibitors, potent and selective inhibitors for each member of the 1A subfamily are still lacking. As a final goal, the structure-activity relationships gained from previous results will be used to test novel compounds as enzymatic inhibitors.

Therefore, in order to contribute to the knowledge of the three human ALDH1A enzymes, ALDH1A1, ALDH1A2 and ALDH1A3, their substrate specificity and inhibitor selectivity, the present Thesis will address the following specific objectives:

1.   To compare the topology of the substrate-binding pockets of the three enzymes and identify the clue residues that could confer specific kinetic features.

2.   To perform their heterologous expression and affinity purification.

3.   To study the effect of magnesium ions on their enzymatic activity.

4.   To determine their kinetic constants with typical ALDH substrates such as alkanals and alkenals.

5. To analyze their activity with retinaldehyde isomers and apo-β-carotenals by using a sensitive HPLC-based method.

6. To construct, express and purify ALDH1A mutant enzymes containing amino acid substitutions in clue residues of their substrate-binding pockets, by using site-directed mutagenesis.

7. To determine the kinetic features and substrate specificity of the mutant enzymes to elucidate the relevance of the substituted residues in their substrate-binding pockets.

8. To test some classical ALDH inhibitors and novel compounds as inhibitors against the three ALDH1A enzymes, and determine their potency and selectivity.

- 59 -

MATERIALS & METHODS

- 60 -

2.1   ALDH structures and volume measurement of substrate-binding pockets

Human ALDH1A1 (PDB code 4WB9), ALDH1A2 (PDB code 6ALJ) and ALDH1A3 (PDB code 5FHZ; monomer A) crystallographic structures (Table 1) were used for the study of the substrate-binding pocket. The POVME algorithm [170,171] was used to measure the volume of the substrate-binding pockets of the three ALDH1A enzymes. The first step was to define an inclusion region with focus located in the catalytic Cys residue, which entirely encompassed all the binding-pocket conformations of the trajectory. Then, a field of equidistant points was generated. After that, points that were near the receptor atoms were removed, leaving those points that were likely to be located within the binding pocket itself. Finally, the patches of points that were not contiguous with the primary binding pocket were removed. The volume of the cofactor-binding pocket was not taken into account in this study, due to the fact that it is a highly conserved region.

2.1.1   Generation of ALDH1A2 structure model and molecular dynamic method

The final study on the substrate-binding pockets of ALDH1As was performed based on the crystallographic structures mentioned above. However, the initial crystallographic data used to analyze human ALDH1A2 was that corresponding to PDB code 4X2Q, which was the only structure available when the present work started. This structure was also chosen because it did not show any molecule bound to its catalytic center other than the cofactor. Thus the holoenzyme structure was fully comparable to that of ALDH1A1 (PDB code 4WB9). However, the deposited ALDH1A2-NAD$^+$ structure lacked the sequence fragment located between Asn475 and Met495. This missing fragment of 21 residues is mostly an unstructured loop in ALDH1A1 and ALDH1A3 (corresponding to Val454-Leu478 and Ile465-Leu489, respectively).

We constructed a computer model that included the complete ALDH1A2 sequence by overlaying this unstructured loop to the corresponding ALDH1A1 sequence. In addition to the unstructured loop, a portion of the cofactor from the ALDH1A1 structure was also incorporated, as it was absent in the crystallographic structure of ALDH1A2. The modelled loop differs by four residues from ALDH1A1 (Leu477, Asn478, Ser481 and Met495 in ALDH1A2 instead of Val460, Ser461, Cys464 and Leu478 in ALDH1A1). These residues were manually mutated with the *Visual Molecular Dynamics* (VMS) program [172].

Table 1. Crystallographic structures of ALDH1A enzymes deposited in the Protein Data Base.

ALDH1A enzyme	cofactor	other ligands	resolution (Å)	PDB code	Released
1A1	–	–	1.74	4WJ9	Dec 2014
1A1	NADH	–	2.07	**4WB9**[a]	Dec 2014
1A1	–	CM026	1.80	4WP7	Feb 2015
1A1	–	CM053	1.95	4WPN	Feb 2015
1A1	NADH	CM037	1.85	4X4L	Feb 2015
1A1	–	BUC11	1.70	5L2M	Mar 2017
1A1	–	BUC22	2.05	5L2O	Mar 2017
1A1	–	BUC25	1.70	5L2N	Mar 2017
1A1	NADH	CM039	2.10	5TEI	Sept 2017
1A2	NAD^+	–	2.94	**4X2Q**[a]	Dec 2015
1A2	NAD^+	WIN18,446	1.89	**6ALJ**[a]	Jan 2018
1A2	NAD^+	6-118	2.20	6B5G	Jan 2018
1A2	NAD^+	CM121	2.30	6B5H	Jan 2018
1A2	–	CM121	2.60	6B5I	Jan 2018
1A3	NAD^+	retinoic acid	2.90	**5FHZ**[a]	Nov 2016

[a]Structures used in this work are shown in bold face.
CM026: 8-{[4-(furan-2-ylcarbonyl)piperazin-1-yl]methyl}-1,3-dimethyl-7-(3-methylbutyl)-3,7-dihydro-1H-purine-2,6-dione; CM053: 1-{[1,3-dimethyl-7-(3-methylbutyl)-2,6-dioxo- 2,3,6,7-tetrahydro-1H-purin-8-yl]methyl}piperidine-4-carboxamide; CM037: ethyl ({4-oxo-3-[3-(pyrrolidin-1-yl)propyl]-3,4-dihydro[1]benzothieno[3,2-d]pyrimidin-2-yl}sulfanyl)acetate; BUC11: 2,3,5-trimethyl-6-[3-oxo-3-(piperidin-1-yl)propyl]-7H-furo[3,2-g][1]benzopyran-7-one; BUC 22: 7-(diethylamino)-4-methyl-2H-1-benzopyran-2-one; BUC25: 3-benzyl-4-methyl-2-oxo-2H-1-benzopyran-7-yl methanesulfonate; CM039: 6-{[(3-fluorophenyl)methyl]sulfanyl}-5-(2-methylphenyl)-2,5-dihydro-4H-pyrazolo[3,4-d]pyrimidin-4-one; WIN18,446: N,N'-(octane-1,8-diyl)bis(2,2-dichloroacetamide); 6-118: (3-ethoxythiophen-2-yl){4-[4-nitro-3-(pyrrolidin-1-yl)phenyl]piperazin-1-yl}methanone; CM121: 1-(4-cyanophenyl)-N-(3-fluorophenyl)-3-[4-(methylsulfonyl)phenyl]-1H-pyrazole-4-carboxamide.

Subsequently, we prepared input files for the molecular dynamics with a molecular simulation scheme, LEaP-dynamics, which provides efficient sampling of the protein conformational space in solution [173], through the addition of water molecules and the necessary balance parameters of the electric force field of the system. NAD^+ and NADH parameters were settled as described above [174,175]. The AMBER software package [176]

was used to carry out molecular dynamics in the ALDH1A2 unstructured loop. Firstly, it was crucial to minimize the energy system to find the nearest local minimum that allowed the X-ray crystallographic structure laxity. Furthermore, it was necessary to warm the system at 300 K in progressive stages and to balance the system density to water density. Finally, it was considered the system to be ready for the study and various properties could be analyzed over time. We have focused on the root-mean-square deviation (RMSD) of atomic positions, the measure of the average distance between the atoms of superimposed proteins. The stability in the RMSD score suggested that the system was prepared for the molecular dynamics simulation. After 30-ns simulation, the system could be considered balanced. The POVME algorithm [170,171] was used to measure the volume of the substrate-binding pocket as described in section 2.1.

2.2   Cloning oh human *ALDH* cDNA

We identified the human genomic clones in GenBank (clone IDs: 2988388, 4826743 and 6208628 for ALDH1A1, ALDH1A2 and ALDH1A3, respectively). Each of human *ALDH* cDNAs were subcloned into two different expression plasmids: pGEX-4T-2 (Figure 23A) and pET-30 Xa/LIC (Figure 23B). For the pGEX-4T-2 constructs, each cDNA was PCR amplified from pOTB7/ALDH1A1, pBluescript R/ALDH1A2 or pOTB7/ALDH1A3 plasmid using the following primers: For ALDH1A1, *Eco*RI-ALDH 1A1 forward primer (5'-T<u>GAATTC</u>CC**ATGTCATCCTCAGGCA CG-3'**) and ALDH1A1-*Sal*I reverse primer (5'-<u>GTCGAC</u>**TTATGAGTTCTTCTGAGAGATTTTCA**-3'**)**, including nucleotides 1-18 and 1480-1506, respectively. For ALDH1A2, *Bam*HI-ALDH1A2 forward primer (5'CGC<u>GGATCC</u>**ATGACTTCCAGCAAGATAGAGATGC-**3') and ALDH1A2-*Eco*RI reverse primer (5'T<u>GAATTC</u>**TTAGGAGTTCTTCTGGGGG-**3'**)** including nucleotides 1-25 and 1539-1557, respectively. For ALDH1A3, *Bam*HI-ALDH1A3 forward primer (5'-CGC<u>GGATCC</u>**ATGGCCAC CGCTAACGG-**3'**)** and the ALDH1A3-*Eco*RI reverse primer (5'-T<u>GAATTC</u>**TCAGGGGGTTCTTGTCGCC-** 3'**),** including nucleotides 1-17 and 1522-1539, respectively. The new restriction sites (underlined) facilitated directional cloning in the multiple cloning site of the pGEX-4T-2 vector. The PCR steps were: 1) DNA polymerase activation (95°C for 2 min), 2) denaturing (95°C for 20 s), 3) annealing (60°C, lowest T_m −5°C, for 10 s) and 4) extension (70°C for 20 s/kb). Then, steps 2-4 were repeated for additional 34 cycles. After digestion, the PCR product was ligated into the corresponding restriction sites of pGEX-4T-2 using the Rapid DNA Dephosphorilation and Ligation kit (GE Healthcare). *ALDH* cDNAs were also subcloned into pET-30 Xa/LIC. For the generation of the pET-30 Xa/LIC-ALDH constructs, the same forward and reverse primers were designed, but without including any restriction sites. Instead, complementary nucleotides to the cohesive ends of pET-30 Xa/LIC were added (Figure 24). PCR conditions were the same as

those used for the pGEX-4T-2 insert amplification. To perform the insert ligation, the pET-30 Xa/LIC kit was used. The resulting constructs were verified by endonuclease digestion, transformed into *E. coli* DH5α cells and sequenced.
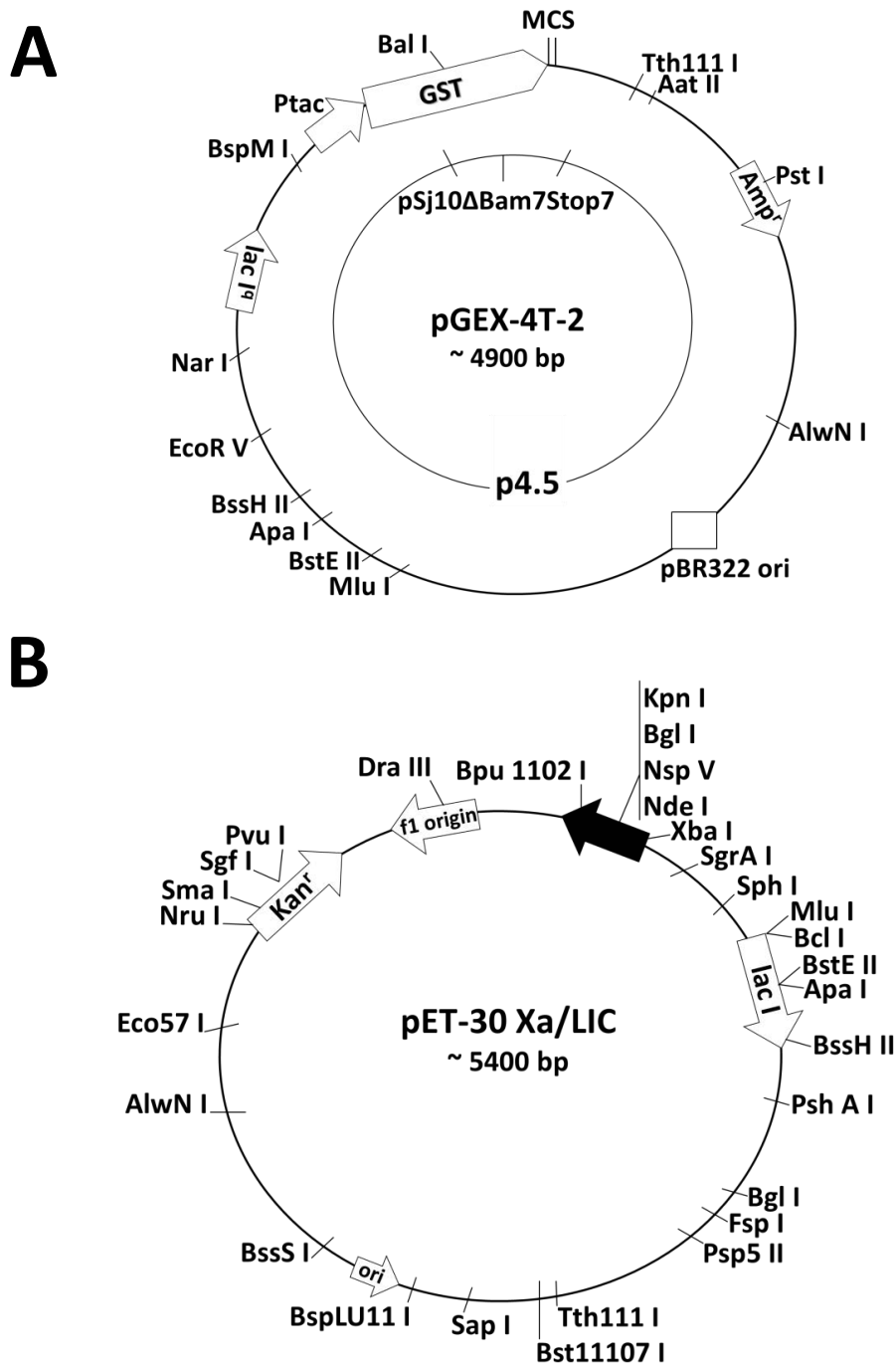
Figure 23. Heterologous expression vectors used for ALDH1A subcloning. A) pGEX-4T-2 is a bacterial expression plasmid with a *tac* promoter for chemically inducible, high-level expression of glutathione S-transferase (GST)-tagged recombinant proteins. The pGEX-4T-2 expression vector has an expanded multicloning site (MCS) that contains six restriction sites and facilitates the unidirectional cloning of cDNA inserts. It exhibits an internal *lacI*[q] gene for use in any *E. coli* host, and contains an ampicillin resistance and a thrombin recognition site. B) The pET-30 Xa/LIC is designed for cloning and high-level expression of target proteins fused with the His·tag coding sequences that are cleavable with factor Xa protease. The plasmid contains a strong T7*lac* promoter, a kanamycin resistance gene, and an optimized MCS (black arrow) that facilitates the insert transfer and includes the coding sequence for the Xa cleavage. The schemes were reproduced from product data sheets provided by the manufacturers (GE Healthcare and Merck Millipore, respectively).
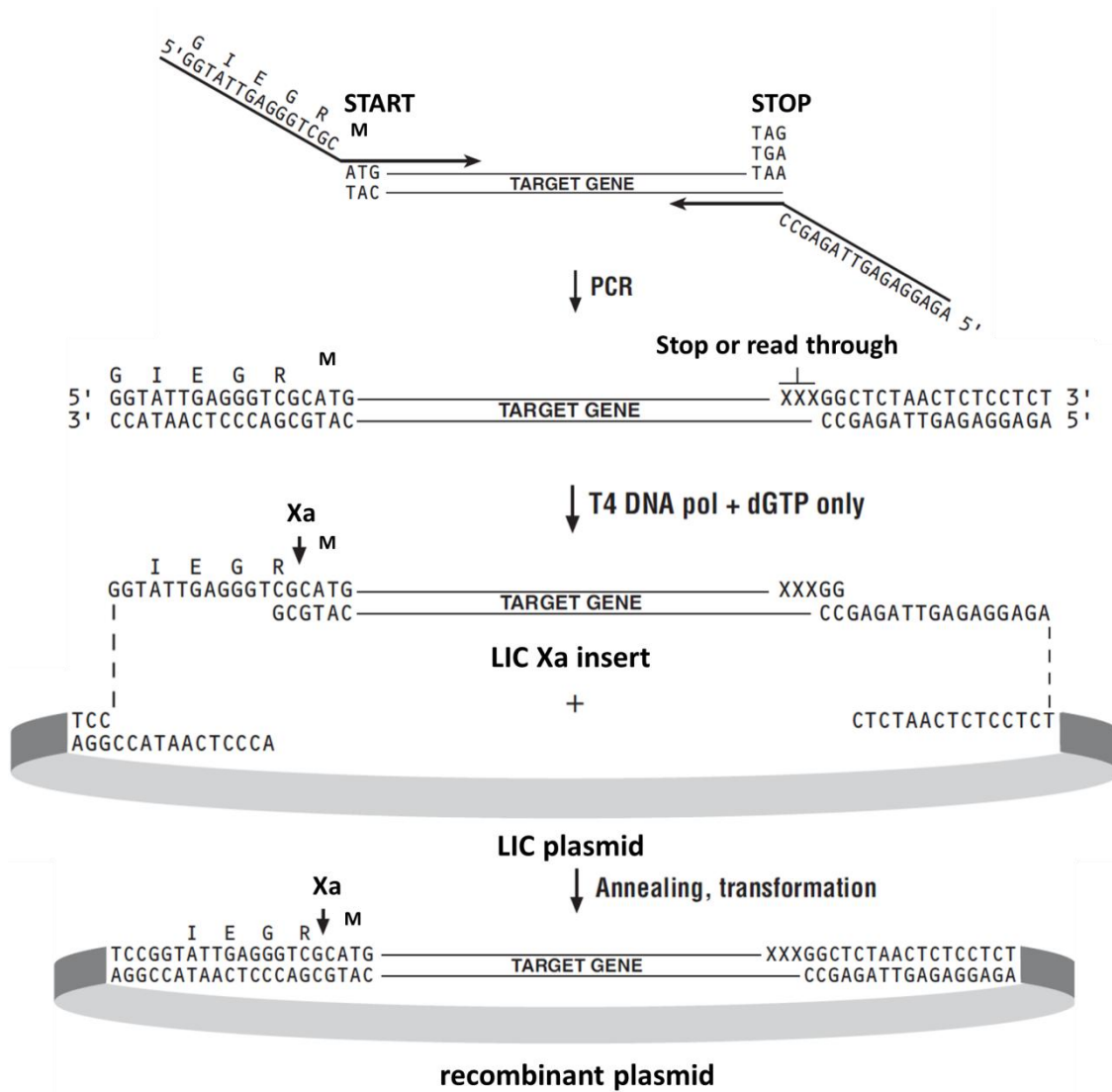
Figure 24. Diagram of the pET-30 Xa/LIC strategy. After amplification with primers that include the indicated 5'
Xa/LIC extensions, the PCR insert is treated with LIC-qualified T4 DNA polymerase (and dGTP) and annealed to the
pET-30 Xa/LIC vector. The resultant nicked, circular plasmid DNA is transformed into competent *E. coli* cells. The
scheme was reproduced from product data sheet provided by the manufacturer (Merck Millipore).

2.2.1 Construction and cloning of full-length ALDH1A2 cDNA

Four ALDH1A2 isoforms may be generated from alternative splicing of the *ALDH1A2* gene
(isoform 1 corresponds to the full-length enzyme). The difference between isoforms 1 and 2 is
the deletion of thirty-seven amino acid residues located at residues 228-265 (nucleotides 684-
795) in the isoform 2 (Figure 25). The commercial manufacturer only provided the cDNA
coding for the isoform 2 and, for this reason, a method was devised to obtain the full-length
cDNA. The complete cDNA encoding isoform 1 of human ALDH1A2 was produced by three
independent PCR amplifications. The first PCR used the forward primer referred previously in
section 2.2 (5'-GGTATTGAGGGTCGCATGACTTCCAGCAAGATAGAGATGC-3') and an internal
reverse primer (5'-CCAATGTGAGAAGCTATTGCTGCCCCAGCCGTTGGCCCATATCCTGGCAAAATAT

TGATGACCCCGGGAGGAAGCCAGCCTCCTTGATGAGGGCTCC-3') including nucleotides 725-824. The second PCR used an internal forward primer (5'-**ATGGGCCAACGGCTGGGGCAGCAATAGCT TCTCACATTGGCATAGACAAGATTGCATTCACAGGGTCTACTGAGGTTGGAAAGCTTATCCAAGAAG CAGC**-3') including nucleotides 666-764 and a reverse primer aforementioned (5'-<u>AGAGGAGA GTTAGAGCC</u>**TTAGGAGTTCTTCTGGGGG**-3'). The primer design involved an overlapping region including thirty-nine nucleotides. This fact is essential to ensure the amplification of the *ALDH1A2* complete sequence.


```
ISOFORM 1      1 MTSSKIEMPGEVKADPAALMASLHLLPSPTPNLEIKYTKIFINNEWQNSE  50
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
ISOFORM 2      1 MTSSKIEMPGEVKADPAALMASLHLLPSPTPNLEIKYTKIFINNEWQNSE  50

ISOFORM 1     51 SGRVFPVYNPATGEQVCEVQEADKADIDKAVQAARLAFSLGSVWRRMDAS 100
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
ISOFORM 2     51 SGRVFPVYNPATGEQVCEVQEADKADIDKAVQAARLAFSLGSVWRRMDAS 100

ISOFORM 1    101 ERGRLLDKLADLVERDRAVLATMESLNGGKPFLQAFYVDLQGVIKTFRYY 150
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
ISOFORM 2    101 ERGRLLDKLADLVERDRAVLATMESLNGGKPFLQAFYVDLQGVIKTFRYY 150

ISOFORM 1    151 AGWADKIHGMTIPVDGDYFTFTRHEPIGVCGQIIPWNFPLLMFAWKIAPA 200
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
ISOFORM 2    151 AGWADKIHGMTIPVDGDYFTFTRHEPIGVCGQIIPWNFPLLMFAWKIAPA 200

ISOFORM 1    201 LCCGNTVVIKPAEQTPLSALYMGALIKEAGFPPGVINILPGYGPTAGAAI 250
                 |||||||||||||||||||||||||||
ISOFORM 2    201 LCCGNTVVIKPAEQTPLSALYMGALIK---------------------- 227

ISOFORM 1    251 ASHIGIDKIAFTGSTEVGKLIQEAAGRSNLKRVTLELGGKSPNIIFADAD 300
                               ||||||||||||||||||||||||||||||||||||
ISOFORM 2    228 --------------EVGKLIQEAAGRSNLKRVTLELGGKSPNIIFADAD 262

ISOFORM 1    301 LDYAVEQAHQGVFFNQGQCCTAGSRIFVEESIYEEFVRRSVERAKRRVVG 350
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
ISOFORM 2    263 LDYAVEQAHQGVFFNQGQCCTAGSRIFVEESIYEEFVRRSVERAKRRVVG 312

ISOFORM 1    351 SPFDPTTEQGPQIDKKQYNKILELIQSGVAEGAKLECGGKGLGRKGFFIE 400
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
ISOFORM 2    313 SPFDPTTEQGPQIDKKQYNKILELIQSGVAEGAKLECGGKGLGRKGFFIE 362

ISOFORM 1    401 PTVFSNVTDDMRIAKEEIFGPVQEILRFKTMDEVIERANNSDFGLVAAVF 450
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
ISOFORM 2    363 PTVFSNVTDDMRIAKEEIFGPVQEILRFKTMDEVIERANNSDFGLVAAVF 412

ISOFORM 1    451 TNDINKALTVSSAMQAGTVWINCYNALNAQSPFGGFKMSGNGREMGEFGL 500
                 ||||||||||||||||||||||||||||||||||||||||||||||||||
ISOFORM 2    413 TNDINKALTVSSAMQAGTVWINCYNALNAQSPFGGFKMSGNGREMGEFGL 462

ISOFORM 1    501 REYSEVKTVTVKIPQKNS 518
                 ||||||||||||||||||
ISOFORM 2    463 REYSEVKTVTVKIPQKNS 480
```

Figure 25. Amino acid sequence alignment of human ALDH1A2 isoforms. The isoform 1 corresponds with the full-length protein sequence while the isoform 2 exhibits a deletion of thirty-seven amino acid residues (228-265) indicated as discontinuous lines.

Finally, a third PCR allowed obtaining the complete sequence from the two fragments amplified previously, by using outer forward and reverse primers and under the same conditions that the isoform 2 as detailed above (Figure 26). The resulting construct was transformed into *E. coli* DH5α cells and cloning was verified by sequencing.

```
601   TTGTGCTGTGGCAATACAGTAGTTATTAAGCCAGCAGAGCAAACACCACTCAGTGCACTC
601   AACACGACACCGTTATGTCATCAATAATTCGGTCGTCTCGTTTGTGGTGAGTCACGTGAG

601   TTGTGCTGTGGCAATACAGTAGTTATTAAGCCAGCAGAGCAAACACCACTCAGTGCACTC
601   AACACGACACCGTTATGTCATCAATAATTCGGTCGTCTCGTTTGTGGTGAGTCACGTGAG


661   TACATGGGAGCCCTCATCAAGGAGGCTGGCTTTCCTCCCGGGGTCATCAATATTTTGCCA
661   ATGTACCCTCGGGAGTAGTTCCTCCGACCGAAAGGAGGGCCCCAGTAGTTATAAAACGGT

661   TACATGGGAGCCCTCATCAAGGA
661   ATGTACCCTCGGGAGTAGTTCCT


721   GGATATGGGCCAACGGCTGGGGCAGCAATAGCTTCTCACATTGGCATAGACAAGATTGCA
721   CCTATACCCGGTTGCCGACCCCGTCGTTATCGAAGAGTGTAACCGTATCTGTTCTAACGT


781   TTCACAGGGTCTACTGAGGTTGGAAAGCTTATCCAAGAAGCAGCTGGAAGAAGTAATTTG
781   AAGTGTCCCAGATGACTCCAACCTTTCGAATAGGTTCTTCGTCGACCTTCTTCATTAAAC

      GAGGTTGGAAAGCTTATCCAAGAAGCAGCTGGAAGAAGTAATTTG
      AAGAGAGTAACTCTGGAACAACCTTTCGAATAGGTTCTTCGTCGACCTTCTTCATTAAAC
```

Figure 26. Scheme of the sequence involved in the design of internal primers to obtain the full-length human ALDH1A2 cDNA. The sequence of ALDH1A2 isoform 1 is shown in black and the sequence of ALDH1A2 isoform 2 is remarked in red. The absence of 111 nucleotides (nt 684-795) corresponding to the deletion of 37 amino acids comprised between residues 288-266, can be observed. The sequence corresponding to the forward primer (nt 666-764) is marked in cyan blue while the sequence corresponding to the reverse primer (nt 725-824) is marked in yellow. The last three nucleotides present in the ALDH1A2 isoform 2 and the three contiguous nucleotides in the sequence are highlighted in grey in order to ensure that the design of the internal primers anchor the complete inclusion of the absent region in ALDH1A2 full-length isoform.

2.3   Site-directed mutagenesis

Human ALDH1A1 and ALDH1A2 were mutated using the QuickChange Lightning Site-Directed Mutagenesis Kit (Agilent Technologies). For this purpose, the following primers were used: For ALDH1A1, ALDH1A1 forward primer (5'- GGTGGAAAACCCTATTCCAATGC -3') and ALDH1A1 reverse primer (5'-GCATTGGAATAGGGTTTTCCACC -3'). For ALDH1A2, ALDH1A2 forward primer (5'-GGATCAATTGTTACGGTGTCGTATCTGCCCAGAGCCCC-3') and ALDH1A2 reverse primer (5'-GGGGCTCTGGGCAGATACGACACCGTAACAATTGATCC -3'). The underlined regions correspond to the mutated nucleotides. The PCR steps were: 1) DNA polymerase activation (95°C for 2 min), 2) denaturing (95°C for 20 s), 3) annealing (50°C for ALDH1A1 and 60°C for ALDH1A2, for 10 s) and 4) extension (68°C for 30 s/kb). Then, the steps 2-4 were repeated for additional 18 cycles. Then, the PCR products were digested with *Dpn* I to

eliminate the parental DNA. The resulting constructs were transformed into *E. coli* DH5α cells, as described below, and sequenced.

2.4 Transformation of recombinant ALDHs into *E. coli* cells

The resulting constructs were cloned into *E. coli* DH5α competent cells with high transformation efficiency, according to the protocol of the manufacturer. Tubes containing plasmid and cells were incubated on an ice bath for 5 min, then heated at 42°C for 30 s and placed back on an ice bath for 2 min. Culture medium was added to the tubes and then they were incubated at 37°C while shaking for 1 h. Finally, cells were plated on selective medium with 33 μg/mL kanamycin for pET-30 Xa/LIC and 50 μg/mL ampicillin for pGEX-4T-2, and incubated overnight (O/N) at 37°C. Once cloning was confirmed by sequencing, the constructs were subcloned into *E. coli* BL21(DE3)pLys cells using the same transformation protocol as described previously for *E. coli* DH5α cells.

2.5 DNA electrophoresis and quantification

Products obtained from PCR and DNA digestion were analyzed by 0.8% agarose gel electrophoresis, using linear and supercoiled DNA ladders. DNA quantification was carried out using a NanoDrop™ spectrophotometer.

2.6 Screening of ALDH protein expression at small scale

Different *E. coli* BL21(DE3)pLys strains transformed with ALDH constructs in pGEX-4T-2 or pET-30 Xa/LIC were tested at small scale for expression levels prior to scaling-up production. For protein expression, transformed *E. coli* BL21(DE3)pLys cells were grown in 5 mL LB medium with kanamycin (pET-30 Xa/LIC) or ampicillin (pET-4T-2) O/N at 37°C. A 250-μl aliquot of this culture was used to inoculate 10 mL LB medium with the proper amount of antibiotic and was incubated at 37°C until an $O.D._{595}$ of 0.8 or 1.5 was reached (Table 2).

Protein production was then induced by the addition of 1 mM IPTG and cells were grown at 22°C. To analyze protein expression, 1-mL aliquots were collected at 4 h and O/N after induction. Each sample was centrifuged at 13200 x g and 4°C for 20 min. The resulting pellet was resuspended in Bind Buffer (containing 50 mM Tris/HCl, 0.5 M NaCl and 5 mM imidazole at pH 8.0) and then, 1 mg/mL lysozyme, 20 μg/mL DNase, 1% Triton-X 100, 1 mM protein inhibitor PMSF and 5 mM DTT were added. The samples were sonicated and then centrifuged

for 20 min at 13200 x g and 4°C. The resulting supernatant and pellet dissolved in Bind Buffer were loaded on an SDS-PAGE.

Table 2. Screening of different conditions tested for protein expression.

Conditions	Protein	Expression vector	
		pGEX-4T-2	pET-30 Xa/LIC
O.D.$_{595}$ = 0.8 4 h – O/N	ALDH1A1	IPTG	IPTG
		control	control
	ALDH1A2	IPTG	IPTG
		control	control
	ALDH1A3	IPTG	IPTG
		control	control
O.D.$_{595}$ = 1.5 4 h – O/N	ALDH1A1	IPTG	IPTG
		control	control
	ALDH1A2	IPTG	IPTG
		control	control
	ALDH1A3	IPTG	IPTG
		control	control

Protein expression with (**IPTG**) or without (control) induction by 1 mM IPTG at O.D.$_{595}$ = 0.8 or 1.5. Induction times were 4 h or O/N. After induction, temperature was set at 22°C. Each sample was sonicated, centrifuged and the resulting pellet resuspended. Two different fractions were loaded, supernatant or soluble fraction (SF) and pellet or insoluble fraction (IF), on SDS-PAGE for each sample taken under each condition.

2.7   Protein expression and purification

Human ALDHs were expressed from the pET-30 Xa/LIC constructs, which allow protein expression with an N-terminal (His)$_6$ tag under the control of T7 RNA polymerase promoter and operon *lac*. In *E. coli* BL21(DE3)pLys. The expression of T7 RNA polymerase is also under the control of the operon *lac* promoter. Consequently, the constitutive expression of the operon lac repressor (*lac I*), contained in both the *E. coli* genome and the expression vector, assures the expression of the protein only in presence of an inducer such as IPTG. For protein expression, transformed *E. coli* BL21(DE3)pLys cells were grown O/N in 25 mL LB medium in the presence of 33 μg/mL kanamycin at 37°C. This culture was used to inoculate 1 L of 2xYT medium in the presence of 33 μg/mL kanamycin and cells were incubated at 37°C until an O.D.$_{595}$ of 0.8 was reached. Protein expression was then induced by the addition of 1 mM IPTG and cells were grown O/N at 22°C. To analyze the protein expression 1-mL aliquots were collected before and after induction. After O/N incubation, cells were collected by

centrifugation at 12400 x g (8000 rpm, Beckman JLA-14 rotor) and 4°C for 10 min. The resulting pellet was resuspended in 30 mL of Bind Buffer per liter of culture. Cell lysis was performed by freezing (O/N at −20°C) and thawing at room temperature. Then, 1 mg/mL lysozyme, 20 µg/mL DNase, 1% Triton-X 100, 1 mM protein inhibitor PMSF and 5 mM DTT were added, followed by two cycles of sonication. The cellular extract was centrifuged at 15000 x g (12500 rpm, Beckman JA-25.5 rotor) and 4°C for 20 min. The soluble fraction was collected and filtered, and an aliquot of this lysate was kept for further electrophoretic analysis. Protein purification was performed by affinity chromatography on a nickel-charged chelating Sepharose™ Fast Flow 5-mL column (His Trap column), which specifically binds the protein due to the His tag, using an ÄKTA™ FPLC purification system. Protein was eluted by applying an increasing step-wise gradient (5, 60, 100 and 250 mM) of imidazole in 50 mM Tris/HCl and 0.5 M NaCl, pH 8.0. The enzyme eluted at 250 mM imidazole. The imidazole present in the eluted protein fractions was removed through a PD-10 gel filtration-desalting column. Protein including the $(His)_6$ tag was stored frozen at −80°C in 50 mM Tris/HCl, 0.5 mM NaCl, pH 8.0. To follow the purification process, 1-mL aliquots were collected from each step and analyzed by SDS-PAGE.

2.8   Fluorescence assay for the dehydrogenase activity and determination of the kinetic constants

Dehydrogenase activity with non-retinoid substrates was monitored using a fluorimeter (Cary Eclipse Varian) at 25°C. ALDH1A1 and ALDH1A2 were assayed in 50 mM HEPES, 0.5 mM EDTA, 0.5 mM DTT, pH 8.0, while ALDH1A3 assays were performed in 50 mM HEPES, 30 mM $MgCl_2$, 5 mM DTT, pH 8.0. NAD^+ concentration was 500 µM. The enzymatic reaction was initiated by the addition of the aldehyde substrate, dissolved in the corresponding reaction buffer. Fluorescence of NADH was followed at 460 nm with excitation at 340 nm and spectral bandwidth of 10 nm. The reaction mixture also contained 5 µM NADH as an internal standard to obtain absolute reaction rates, which were calculated according to the equation: $v = \frac{dF}{dt} \cdot \frac{C_{st}}{F_{st}}$, where C_{st} is the standard NADH concentration, F_{st} the standard fluorescence and *dF/dt* the slope of the time dependent fluorescence [177]. Assays without enzyme were carried out as controls. The initial velocities were measured in duplicate with at least five different substrate concentrations. The kinetic constants were calculated using the non-linear regression program Grafit 5.0 (Erithacus software), and expressed as the mean ± standard error from three independent determinations.

2.9   Effect of magnesium ions on the ALDH1A activity

Enzymatic activity was measured fluorimetrically in 50 mM HEPES, 0.5 mM DTT, pH 8.0 (for ALD1A1 and ALDH1A2) or in 50 mM HEPES, 5 mM DTT, pH 8.0 (for ALDH1A3), in the presence of 500 μM NAD$^+$. To each reaction mixture, the corresponding amount of MgCl$_2$ was added (0-400 μM final concentration, for ALDH1A1 and ALDH1A2; 0-50 mM, for ALDH1A3). Reaction was initiated with the addition of 250 μM hexanal (for ALDH1A2 and ALDH1A3) or 10 μM hexanal (for ALDH1A1) as a substrate. Reaction mixtures without enzyme served as controls. The initial rates were measured in duplicate and plotted as the percentage of remaining enzymatic activity.

2.10   HPLC-based assay for the dehydrogenase activity with retinoids

To evaluate the retinaldehyde dehydrogenase activity of human ALDHs, an HPLC-based method was used for the first time to determine the ALDH activity, as follows. Stock solutions of retinoid substrates were prepared in ethanol. Working stock solutions of retinoids were prepared by 10-min sonication in the presence of equimolar delipidated BSA [109]. The actual concentration of solubilized retinoid was determined based on the corresponding molar absorption coefficient: ε_{370} = 29500 M^{-1}·cm^{-1} for all-*trans*-retinaldehyde and ε_{364} = 26700 M^{-1}·cm^{-1} for 9-*cis*-retinaldehyde. The enzymatic reaction was carried out for 15 min at 37°C in a final volume of 500 μL, using the same buffer as that for the fluorimetric assays (see above) and a saturating concentration of cofactor (500 μM NAD$^+$). The reaction was stopped by addition of 1 mL cold hexane/dioxane/isopropanol (50:5:1, v/v/v), and retinoids were extracted by a two-step procedure with the same solvent mixture [178,179]. The aqueous phase was removed, and the organic phase was evaporated under a N$_2$ stream. Retinoids were dissolved in 200 μL of hexane and analyzed by a modification of a published method [180], as follows. Retinoids were separated by HPLC on a Novapak® Silica 4 μm, 3.9 x 150 mm in hexane/methyl-*tert*-butyl ether (96:4, v/v) mobile phase, at a flow rate of 2 mL/min, using a Waters Alliance 2695 HPLC. Elution was monitored at 370 and 364 nm for all-*trans*- and 9-*cis*-retinaldehyde isomers, respectively, and at 350 and 341 nm for the corresponding all-*trans* and 9-*cis*-retinoic acid compounds, using a Waters 2996 photodiode array detector. Quantification of retinoids was performed by interpolation of HPLC peak areas into a calibration curve of known retinoid concentrations (Annex 1.1). Kinetic constants were calculated as indicated above in section 2.8. All compound manipulations were performed under dim or red light to prevent photoisomerization.

2.11   Dehydrogenase activity with apo-β-carotenals

ALDH1A1 and ALDH1A2 activities were assayed in 50 mM HEPES, 0.5 mM EDTA, 0.5 mM DTT, pH 8.0, and ALDH1A3 was assayed in 50 mM HEPES, 30 mM $MgCl_2$, 5 mM DTT, pH 8.0. Concentration of 12'-apo- and 14'-apo-β-carotenal was determined based on the corresponding molar absorption coefficient in aqueous solutions at the appropriate wavelength (ε_{410} = 24228 $M^{-1} \cdot cm^{-1}$ and ε_{416} = 15945 $M^{-1} \cdot cm^{-1}$, for 12'-apo-β-carotenal in ALDH1A1/1A2 and ALDH1A3 reaction buffer, respectively; and ε_{397} = 9218 $M^{-1} \cdot cm^{-1}$ and ε_{340} = 9036 $M^{-1} \cdot cm^{-1}$, for 14'-apo-β-carotenal in ALDH1A1/1A2 and ALDH1A3 reaction buffer, respectively). The reaction was started by the addition of cofactor and carried out for 15 min at 37°C in a final volume of 500 μL. With the aim to measure the steady state enzymatic activity, the concentration of enzyme was kept from 50- to 100-fold lower than that of the substrate for all enzymatic assays and a saturating concentration of cofactor (500 μM NAD^+) was used. Reaction products were extracted with hexane/dioxane/isopropanol (50:5:1, v/v/v) and analyzed by the HPLC-based method described above in the section 2.10. Elution was monitored at 415 nm for 12'-apo-β-carotenal, 400 nm for 12'-apo-β-carotenoic acid and 14'-apo-β-carotenal, and at 373 nm for 14'-apo-β-carotenoic acid. Quantification of apo-β-carotenoids was performed similarly to that of retinoids, by interpolation of HPLC peak areas into a calibration curve of known apo-β-carotenoid concentrations (Annex 1.2). Kinetic constants were calculated as indicated above in section 2.8. In the same way as retinoids, all compounds were handled under dim or red light to prevent photoisomerization.

2.12   Spectrophotometric assay for esterase activity

Enzymes were assayed for esterase activity in a Cary 400 Bio (Varian) spectrophotometer. Stock solutions of 10 mM *p*-nitrophenyl acetate were prepared with ethanol as a solvent to minimize spontaneous hydrolysis, always ensuring that the final ethanol concentration in the assay did not exceed 1% (v/v) [181]. Reactions were performed at 25°C in 50 mM HEPES, 0.5 mM DTT, pH 7.5, with 500 μM NAD^+. The reactions were initiated by the addition of *p*-nitrophenyl acetate to the enzyme solutions in 1-mL cuvettes. The formation of *p*-nitrophenol was followed at 400 nm (ε_{400} = 12800 $M^{-1} \cdot cm^{-1}$) [82] to monitor esterase activity, which was corrected for the non-enzymatic hydrolysis of *p*-nitrophenyl acetate. Kinetic constants were calculated as indicated above in section 2.8.

2.13   Inhibitor screening against human ALDH1A enzymes

Single-point analyses of enzymatic activity at 10 μM of inhibitor were performed by three instrumental methods: Varian Cary 400 (UV/Vis) spectrophotometer, Varian Cary Eclipse

fluorimeter and 96-well plates in a Perkin Elmer Victor III Multilabel Plate Reader. The reaction was followed by absorbance at 340 nm or by fluorescence emission at 460 nm, both corresponding to the NADH production. Enzymatic inhibition assays were performed in a final volume of 1 mL, in an appropriate reaction buffer (50 mM HEPES, 0.5 mM EDTA, 0.5 mM DTT, pH 8.0, for ALDH1A1 and ALDH1A2, and 50 mM HEPES, 30 mM MgCl$_2$, 5 mM DTT, pH 8.0, for ALDH1A3) and with 500 μM NAD$^+$. All compounds tested as inhibitors were dissolved in DMSO and assayed in a final concentration of 0.1% (v/v) DMSO using hexanal as a substrate at 250 μM, for ALDH1A2 and ALDH1A3, and 10 μM for ALDH1A1, at 25°C. The concentration of enzyme was kept from 50- to 100-fold lower than that of the substrate for all enzymatic assays. The reaction mixture was incubated for 5 min at room temperature, before adding the substrate. The percentage of remaining activity was calculated from the ratio of activity at a given inhibitor concentration to the control activity with 1% (v/v) DMSO without inhibitor compound added. The IC$_{50}$ values were calculated by fitting the initial rates to the appropriate equation using Grafit 5.0 (Erithacus Software) and values were given as the mean ± standard error of three experiments. In all cases, standard error values were less than 30% of the mean values.

- 74 -

RESULTS

The results presented in this chapter have appeared in part in the following publications:

M. Domínguez, R. Pequerul, R. Alvarez, J. Giménez-Dejoz, E. Birta, S. Porté, R. Rühl, X. Parés, J. Farrés and A. R. de Lera. **Synthesis of apocarotenoids by acyclic cross metathesis and characterization as substrates for human retinaldehyde dehydrogenases**, Tetrahedron. 74 (2018) 2567–2574.

R. Pequerul, J. Vera, J. Giménez-Dejoz, I. Crespo, J. Coines, S. Porté, C. Rovira, X. Parés and J.

Farrés. **Structural and kinetic features of three human aldehyde dehydrogenases, ALDH1A1,**

ALDH1A2 and ALDH1A3, active in retinoic acid biosynthesis (To be submitted).

- 76 -

3.1   Comparison of the substrate-binding pocket of human ALDH1A enzymes

The recent availability of the crystal structures of the holo-ALDH1A enzymes (Table 1) has allowed us to perform for the first time a comparative topological analysis of the substrate-binding pockets of the three human ALDH1A holoenzymes. The three structures are well superimposed (pairwise RMSD values ranged from 0.429 to 0.659 Å) but they are not fully equivalent, as they correspond to different complexes with oxidized/reduced cofactor and with or without inhibitor or product molecule: ALDH1A1 with NADH, ALDH1A2 with NAD$^+$ and WIN 18,446 inhibitor, and ALDH1A3 with NAD$^+$ and retinoic acid. The active-site nucleophile Cys residue and the cofactor molecule could be well superimposed in the three structures. The contracted or open conformations of NAD$^+$ in the ALDH1A2 and ALDH1A3 ternary complexes were similar to that of NADH in the ALDH1A1 binary complex [26,54,182]. The previously deposited ALDH1A2-NAD$^+$ structure (PDB code 4X2Q) was incomplete as it included an unstructured segment between Asn475 and Met495 (corresponding to a loop region, Gly458-Leu478 in ALDH1A1 and Asn469-Leu489 in ALDH1A3), located between the β18 sheet and α13 helix, and it contained only a fragment of NAD$^+$ presumably due to incomplete ligand occupancy. Overall, ALDH1A1, ALDH1A2 and ALDH1A3 exhibited similar topologies and decreasing volumes in their substrate-binding pockets, i.e. 534, 387, 357 Å^3, respectively (Figure 27). Prior to using the most recent structure of ALDH1A2 (with PDB code 6ALJ), the former ALDH1A2 structure (with PDB code 4X2Q) was used to determine the volume of the substrate-binding pocket. After modeling the aforementioned loop (absent in the 4X2Q structure) by a molecular dynamics method, and after adding the missing region of the NAD$^+$ cofactor, a volume of 427 Å^3 was obtained. This value is close to the 387 Å^3 here obtained with the complete ALDH1A2 6ALJ structure.

Table 3 lists the residues lining the substrate-binding pocket in the ALDH1A enzymes. In the three structures, invariant Trp178 (ALDH1A1 numbering) and residue 460 (Val/Leu) define a bottle neck in the middle part of the pocket. In ALDH1A2, a new subpocket, lined by loop residues Leu477, Asn478 and Ala479, is observed near the solvent exposed area, which could be generated by inhibitor binding [54]. In an entire 20-residue ALDH1A2 segment (residues 475-495), it is possible to identify six-residue changes in comparison with ALDH1A1: Asn475, Ala476, Leu477, Asn478 Ser481 and Met495, located in ALDH1A2, are replaced by Gly458, Val459, Val460, Ser461, Cys464 and Leu478, respectively, in the ALDH1A1 enzyme. However, only three residues, Asn478, Ser481 and Met495, are replaced by Tyr472, Ala475 and Leu489, respectively, in ALDH1A3. The first four contiguous residues, 475 to 478, were chosen to construct an ALDH1A2 mutant in order to mimic the ALDH1A1 architecture.
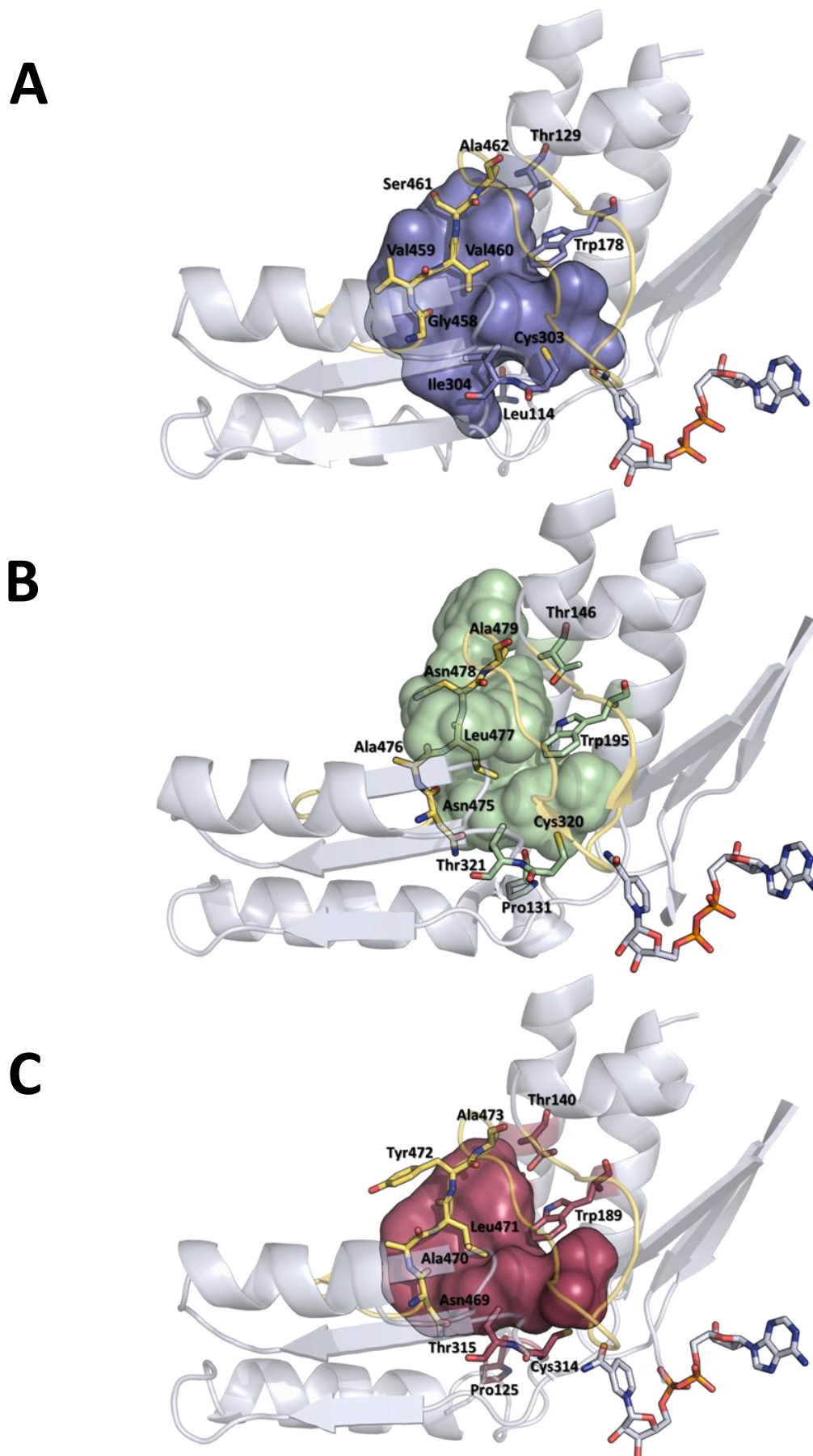
Figure 27. Crystallographic structure of the catalytic domain of the three human ALDH1A enzymes. (A), ALDH1A1; (B) ALDH1A2; (C) ALDH1A3. The contour of the substrate-binding pockets of ALDH1A1, ALDH1A2 and ALDH1A3 is colored purple, green and burgundy red, respectively. The cofactor (NAD$^+$/NADH) is displayed in bluewhite. The residues lining the pocket, including the active-site cysteine, are indicated. The secondary structure is depicted in white ribbon representation while the 475-495 loop (ALDH1A2 residue numbering) is shown in yellow.

Residue 114, located in the NAD^{+}-binding domain just before helix α3 (residues 115-136), which defines one side of the pocket, is Leu in ALDH1A1 but Pro in ALDH1A2 and ALDH1A3. Leu is typically a residue with a high α-helix-forming propensity while Pro could be helix breaker depending on its environment. Thus, the change L114P was selected for a mutant ALDH1A1 to make this part of the substrate-binding pocket more similar to that of ALDH1A2.The kinetic properties of these mutants are presented below in section 3.10.

Table 3. Comparison of homologous residues lining the substrate-binding pocket of ALDH1A enzymes.
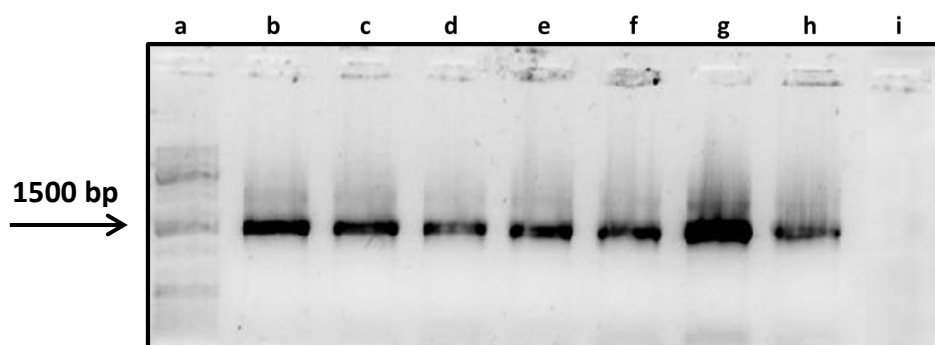
ALDH1A1	ALDH1A2	ALDH1A3
Leu114*	Pro131	Pro125
Asn121	Val138	Ile132
Gly125	**Gly142**	**Gly136**
Lys128	Lys145	Arg139
Thr129	**Thr146**	**Thr140**
Trp178	**Trp195**	**Trp189**
Ile304	Thr321	Thr315
Gly458	Asn475*	Asn469
Val459	Ala476*	Ala470
Val460	Leu477*	Leu471
Ser461	Asn478*	Tyr472
Ala462	**Ala479**	**Ala473**

*Asterisk indicates mutated residues. Bold face highlights strictly conserved residues.

3.2 Subcloning of human *ALDH1A* cDNAs into expression vectors

Proper subcloning of *ALDH* cDNAs in the expression vectors pGEX-4T-2 and pET-30 Xa/LIC was checked by PCR amplification of inserts and DNA agarose gel electrophoresis analysis (Figure 28). A single DNA band of approximately 1500 bp (1506 bp in ALDH1A1, 1557 bp in ALDH1A2 (isoform 1) and 1539 bp in ALDH1A3) was obtained in all cases. DNA amplification was verified with a positive control including a DNA which was known to be amplified. A negative control was performed by leaving out template DNA.

Figure 28. Analysis of PCR products by 0.8% agarose gel electrophoresis. (a) Gene-Ruler™ 1 kb DNA Ladder (Annex 1.3); (b) PCR positive control; (c) pGEX-4T-2/ALDH1A1; (d) pET-30 Xa/LIC/ALDH1A1; (e) pGEX-4T-2/ALDH1A2; (f) pET-30 Xa/LIC/ALDH1A2; (g) pGEX-4T-2/ALDH1A3; (h) pET-30 Xa/LIC/ALDH1A3; (i) PCR negative control.

Regarding the construction of the ALDH1A2 isoform 1, fully delineated above in section 2.2.1, the successful amplification of the full-length sequence, including the missing segment, is shown in Figure 29.
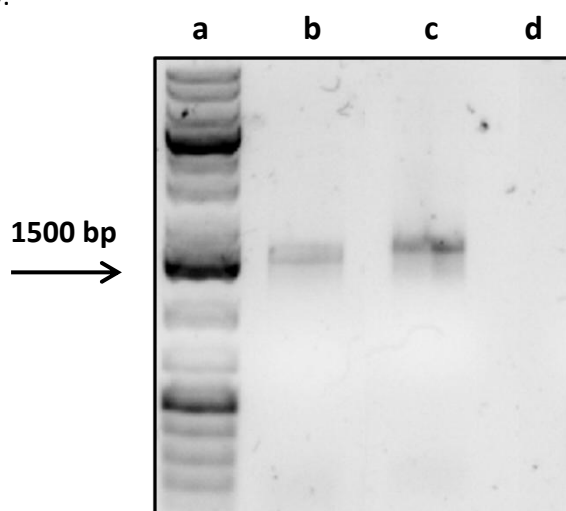

Figure 29. Analysis of PCR amplification to obtain the full-length ALDH1A2 isoform 1, using 0.8% agarose electrophoresis (a) Gene-Ruler™ 1 kb DNA Ladder (Annex 1.3); (b) PCR positive control carried out with *ALDH1A2* isoform 2 as a DNA template. (c) Full-length *ALDH1A2* cDNA. The DNA template used to obtain the band present in the agarose gel is the result of two independent PCRs using internal primers (d), PCR negative control performed in the absence of DNA template.

After inserting the amplified DNA sequences into the corresponding expression vectors, pGEX-4T-2 and pET-30 Xa/LIC, the absence of undesired mutations was verified by DNA sequencing.

## 3.3   Small-scale screening of protein expression of human ALDHs

Figures 30-33 show the SDS-PAGE analyses of the ALDH1A1 protein expression at small scale, under different conditions, according to the scheme presented in Table 2 (Materials & Methods, section 2.6).

For all the *ALDH1A1* cDNAs subcloned into pGEX-4T-2 (Figures 30-31), the expected molecular weight of the fusion proteins should be approximately 81 kDa (55 kDa for ALDH plus 26 kDa for GST). This value agrees with the observed results. In general, most of the overexpressed protein appeared in the insoluble fraction (Figures 30-31, lanes d and h), both at 4 h after induction or in samples incubated O/N, independently of the $O.D._{595}$ at which the induction took place.
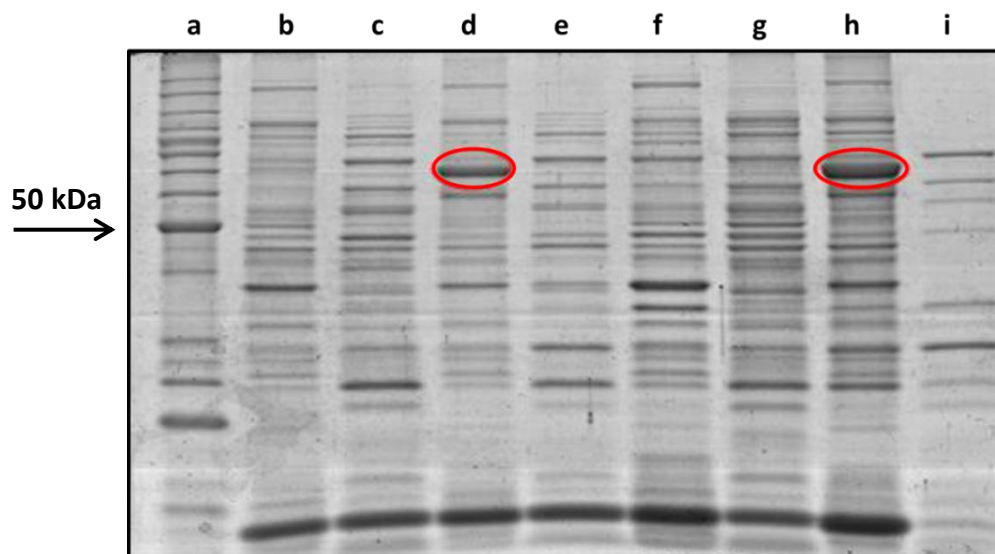
Figure 30. SDS-PAGE analysis of protein expressed (0.2 mg/ml) by pGEX-4T-2/ALDH1A1 ($O.D._{595}$ = 0.8). (a) Bench-Mark™ Protein Ladder (Annex 1.4); (b) Insoluble fraction (IF) of control after 4 h incubation without IPTG; (c) Soluble fraction (SF) of control after 4 h incubation without IPTG; (d) IF of 4 h after induction with IPTG; (e) SF of 4 h after induction with IPTG; (f) IF of control after O/N incubation without IPTG; (g) SF of control after O/N incubation without IPTG; (h) IF of O/N after induction with IPTG; (i) SF of O/N after induction of IPTG. The insoluble fraction contains the resuspended pellet from the cell fraction after lysis, and the soluble fraction is the cell lysis supernatant. ALDH1A1 bands are circled in red.

Furthermore, the amount of expressed protein is greater if the induction with IPTG is extended during O/N incubation, as compared to the 4-h IPTG sample induction. This result is not so evident in the assay of the sample ALDH1A1/pGEX-4T-2 with an $O.D._{595}$ = 1.5, since the intensity of the observed band is similar in the two induction times (Figure 31). It is remarkable that the expression of the protein is not detected in the control samples, in which IPTG is not added, neither in the electrophoretic analysis of pGEX-4T-2/ALDH1A1 with $O.D._{595}$ = 0.8 nor in that of $O.D._{595}$ = 1.5. At this point, we could state that the induction of protein expression in the heterologous system has been produced efficiently in the insoluble fraction in both trials.
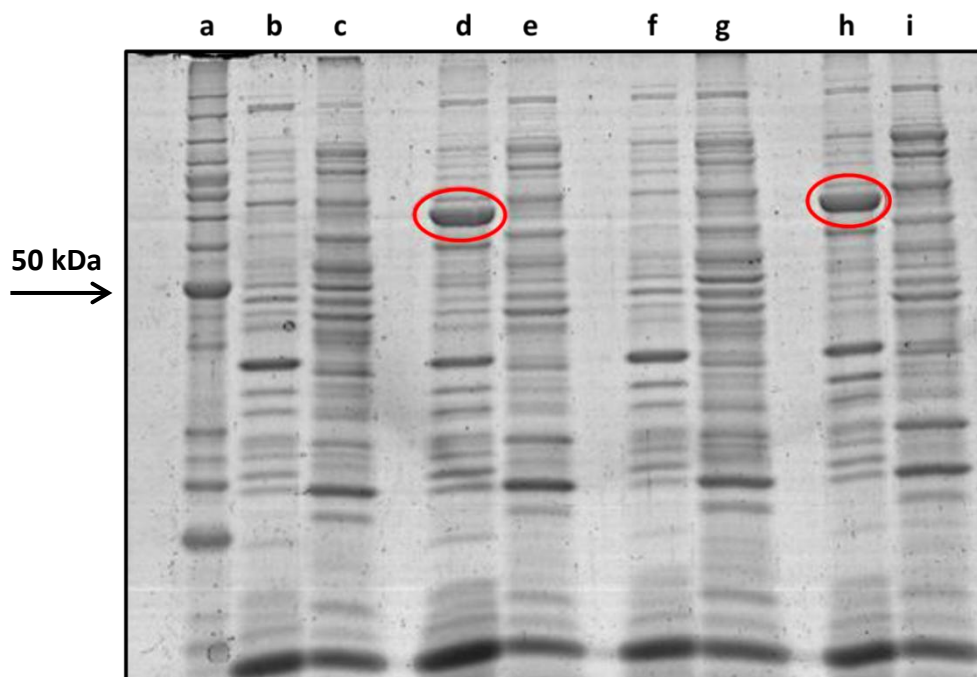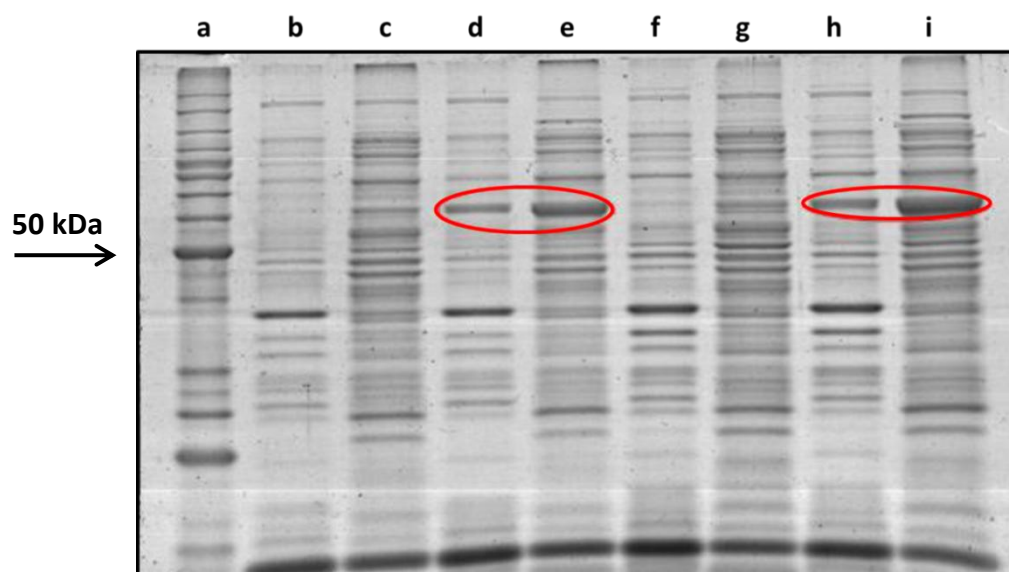
Figure 31. SDS-PAGE analysis of protein expressed (0.2 mg/ml) by pGEX-4T-2/ALDH1A1 (O.D.$_{595}$ = 1.5). (a) Bench-Mark™ Protein Ladder (Annex 1.4); (b) Insoluble fraction (IF) of control after 4 h incubation without IPTG; (c) Soluble fraction (SF) of control after 4 h incubation without IPTG; (d) IF of 4 h after induction with IPTG; (e) SF of 4 h after induction with IPTG; (f) IF of control after O/N incubation without IPTG; (g) SF of control after O/N incubation without IPTG; (h) IF of O/N after induction with IPTG; (i) SF of O/N after induction of IPTG. The insoluble fraction contains the resuspended pellet from the cell fraction after lysis, and the soluble fraction is the cell lysis supernatant. ALDH1A1 bands are circled in red.
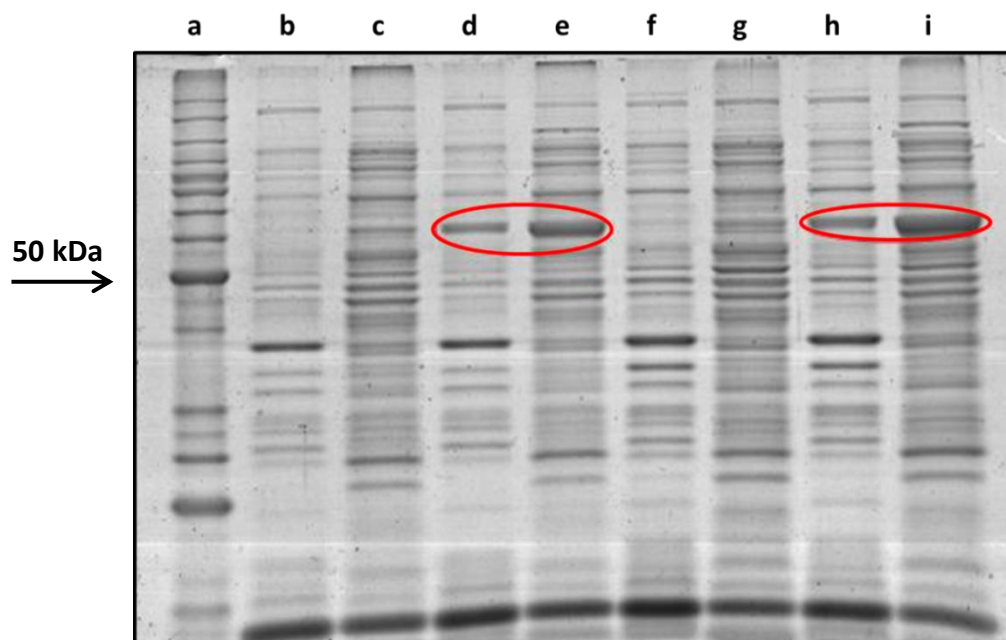
For the *ALDH* cDNA subcloned into pET-30 Xa/LIC (Figures 32-33), the expected molecular weight of the fusion proteins is approximately 60 kDa (55 kDa for ALDH plus 5 kDa for the N-terminal His tag). This value agrees with the observed results. In contrast to what was found with the pGEX-4T-2 constructs, the overexpressed protein was recovered preferentially in the supernatant as compared to the insoluble fraction (lanes e and i). This was especially true in the case of ALDH1A1, both at 4 h after induction or in samples incubated O/N, independently of the O.D.$_{595}$ at which the induction was performed. O/N induction provided a higher amount of ALDH1A1 protein (Figures 32-33). In addition, the results showed that the optimal O.D.$_{595}$ is 0.8 (Figure 32) as compared to O.D.$_{595}$ = 1.5 (Figure 33).

The results obtained for the ALDH1A2 (isoform 1) and ALDH1A3 enzymes were similar to those presented for ALDH1A1. In all cases, the pGEX-4T-2/ALDH1A constructs showed higher expression of the recombinant protein in the insoluble fraction, generally being most efficient the O/N induction as compared to the 4-h incubation samples. In addition, and also generally, the expression of the protein presented a higher yield when the sample was grown at O.D.$_{595}$ = 0.8 instead of O.D.$_{595}$ = 1.5. On the other hand, in the samples of the pET-30 Xa/LIC/ALDH1A constructs, for ALDH1A2 and ALDH1A3, a higher amount of protein was observed in the
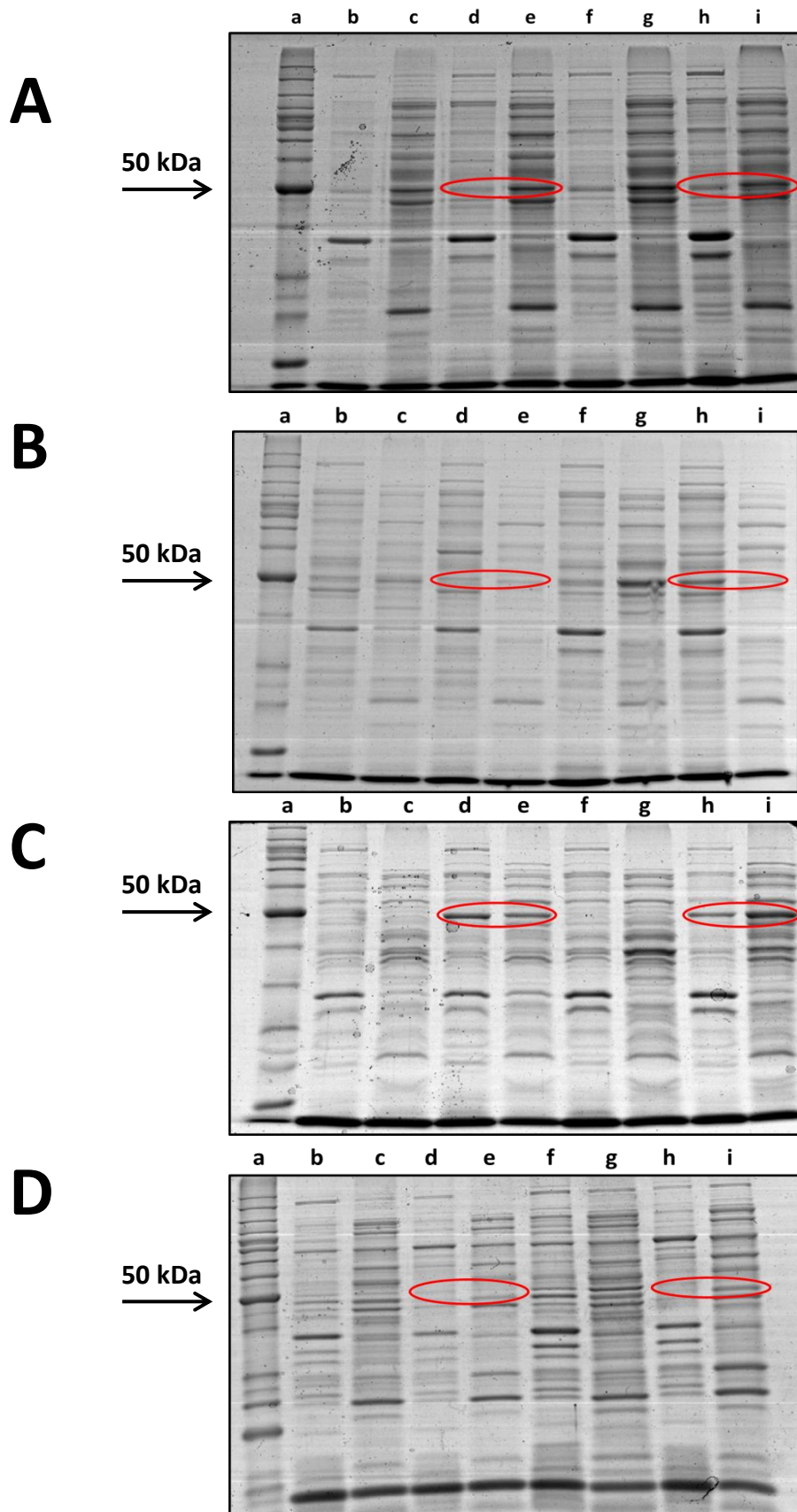
soluble fraction, being the most effective time of induction O/N as compared to 4 h, and keeping the growth before induction until an O.D.$_{595}$ = 0.8 instead of O.D.$_{595}$ = 1.5 (Figure 34).

Figure 32. SDS-PAGE analysis of protein expressed (0.2 mg/ml) by pET-30 Xa/LIC/ALDH1A1 (O.D.$_{595}$ = 0.8). (a) Bench- Mark™ Protein Ladder (Annex 1.4); (b) Insoluble fraction (IF) of control after 4 h incubation without IPTG; (c) Soluble fraction (SF) of control after 4 h incubation without IPTG; (d) IF of 4 h after induction with IPTG; (e) SF of 4 h after induction with IPTG; (f) IF of control after O/N incubation without IPTG; (g) SF of control after O/N incubation without IPTG; (h) IF of O/N after induction with IPTG; (i) SF of O/N after induction of IPTG. The insoluble fraction contains the resuspended pellet from the cell fraction after lysis and soluble fraction is the cell lysis supernatant. ALDH1A1 bands are circled in red.

Figure 33. SDS-PAGE analysis of protein expressed (0.2 mg/ml) by pET-30 Xa/LIC/ALDH1A1 (O.D.$_{595}$ = 1.5). (a) Bench- Mark™ Protein Ladder (Annex 1.4); (b) Insoluble fraction (IF) of control after 4 h incubation without IPTG; (c) Soluble fraction (SF) of control after 4 h incubation without IPTG; (d) IF of 4 h after induction with IPTG; (e) SF of 4 h after induction with IPTG; (f) IF of control after O/N incubation without IPTG; (g) SF of control after O/N incubation without IPTG; (h) IF of O/N after induction with IPTG; (i) SF of O/N after induction of IPTG. The insoluble fraction contains the resuspended pellet from the cell fraction after lysis and soluble fraction is the cell lysis supernatant. ALDH1A1 bands are circled in red.

Figure 34. SDS-PAGE analysis of protein expressed (0.2 mg/ml) by pET-30 Xa/LIC/ALDH1A2 (A) O.D.$_{595}$ = 0.8; (B) O.D.$_{595}$ =1.5 and pET-30 Xa/LIC/ALDH1A3 (C) O.D.$_{595}$ = 0.8; (D) O.D.$_{595}$ = 1.5. (a) Bench-Mark Protein Ladder (Annex 1.4); (b) Insoluble fraction (IF) of control 4 h after induction without IPTG; (c) Soluble fraction (SF) of control 4 h after induction without IPTG; (d) IF of 4 h after induction with IPTG; (e) SF of 4 h after induction with IPTG; (f) IF of control O/N induction without IPTG; (g) SF of control O/N induction without IPTG; (h) IF of O/N induction with IPTG; (i) SF of O/N induction of IPTG. The insoluble fraction is corresponding with pellet cell fraction after lysis and soluble fraction is the cell lysis supernatant. ALDH1A2-ALDH1A3 bands are circled in red.

Therefore, the conditions chosen for the expression of all the recombinant proteins of the ALDH1A subfamily were the following: pET-30 Xa/LIC/ALDH1A constructs, with an O/N incubation after IPTG induction when O.D.$_{595}$ = 0.8.

## 3.4   Purification of recombinant human ALDH1A enzymes

Once the conditions for protein overexpression had been established, large-scale protein purification from the soluble fraction was performed through affinity chromatography on a nickel-charged Sepharose column using an FPLC system. The chromatographic elution profile of ALDH1A, obtained by monitoring the column effluent at 280 nm, as well as the electrophoretic analyses of fractions, collected after various purification steps, are shown in Figure 35. Three main protein peaks were detected: the first peak corresponded to the unbound proteins (eluted with 5 mM imidazole), followed by the weakly bound proteins (eluted with 60 mM imidazole), and finally the His-tagged protein was eluted with 250 mM imidazole. The His tag was not removed from the fusion protein.

In all cases, most of the expressed protein is found in the soluble fraction. The fraction eluted with 250 mM imidazole exhibits a single band of the expected molecular weight, corresponding to the fusion protein between ALDH1A1 and the N-terminal His tag. Similar results were obtained with ALDH1A2 and ALDH1A3, although very minor contaminating bands were observed in ALDH1A1 SDS-PAGE. A final concentration of 7-11 mg/mL of pure enzyme was obtained. Enzyme aliquots, when stored at 4°C, showed protein precipitation within few days. In contrast, when aliquots were stored at −80°C, precipitates were not observed. For this reason, it was decided to store the purified enzymes in small volume aliquots at −80°C.

3.5   Effect of Mg^{2+} ions on the dehydrogenase activity of human ALDH1A enzymes

The inhibitory effect of Mg^{2+} ions and other divalent cations on ALDH1A1 (formerly referred as cytosolic ALDH or ALDH1) from human and other species had been reported, using acetaldehyde and propanal as substrates [88]. The Mg^{2+} effect using hexanal as a substrate is presented in this Thesis, since it is the substrate with the highest catalytic efficiency (Table 4). When examining the effect of Mg^{2+} ions on the dehydrogenase activity of ALDH1A enzymes, a quite distinct pattern was observed for each of the three enzymes (Figure 36). ALDH1A1 and ALDH1A2 activity decreased with increasing concentrations of Mg^{2+}, while ALDH1A3 activity increased. ALDH1A1 was extremely sensitive to divalent cations and thus it lost more than 80% of its activity in the presence of 50 μM MgCl$_2$. ALDH1A2 activity exhibited a less pronounced inhibitory pattern, losing only up to 40% of its activity when the Mg^{2+} concentration reached
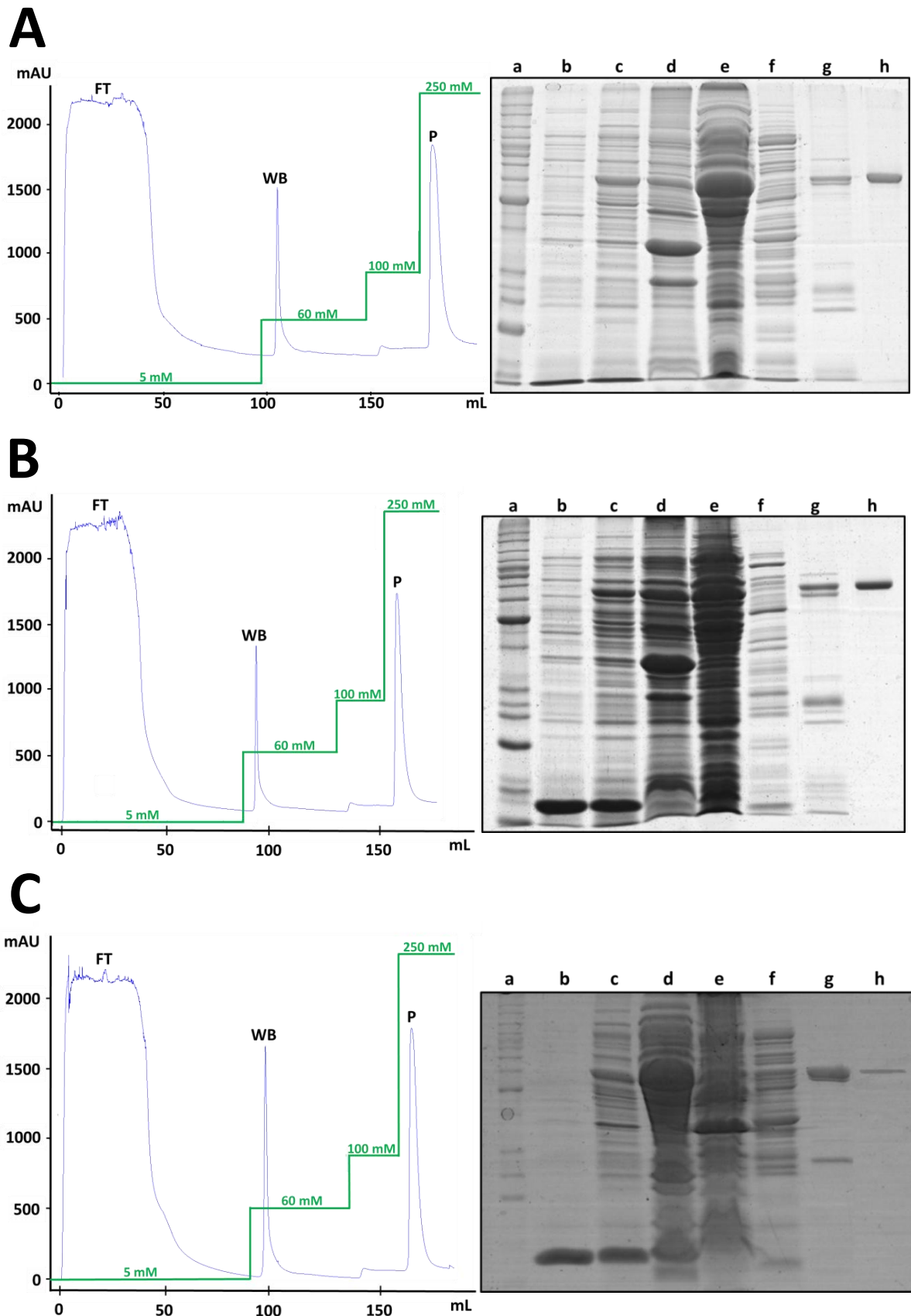
Figure 35. FPLC elution profiles of a (His)$_6$-ALDH1A purification on a nickel-charged Sepharose column (left) and SDS-PAGE analysis of ALDH1A expression and purification steps (right). (A) ALDH1A1; (B) ALDH1A2; (C) ALDH1A3; The blue line represents the amount of eluted protein measured by the absorbance at 280 nm and the green line represents the stepwise imidazole gradient. **FT**: flowthrough or unbound protein fraction. **WB**: weakly bound protein fraction. **P**: peak with protein of interest. (a) Bench-Mark[TM] Protein Ladder (reference bands are labeled in Annex 1.4); (b) Soluble protein fraction before induction; (c) Soluble protein fraction after IPTG induction; (d) Insoluble fraction after sonication; (e) Soluble fraction after sonication; (f) Flowthrough; (g) Weakly bound protein fraction; (h) peak with protein of interest (2 µg).

300 μM. In contrast, Mg^{2+} caused an opposite effect on ALDH1A3, since the enzyme activity increased as much as three-fold at 5 mM $MgCl_2$ and four-fold when the $MgCl_2$ concentration was raised up to 30 mM. Thus, thereafter, ALDH1A1 and ALDH1A2 were routinely assayed in the absence of Mg^{2+} and with 0.5 mM EDTA, while ALDH1A3 assays were performed in 30 mM $MgCl_2$ without any chelating agent.
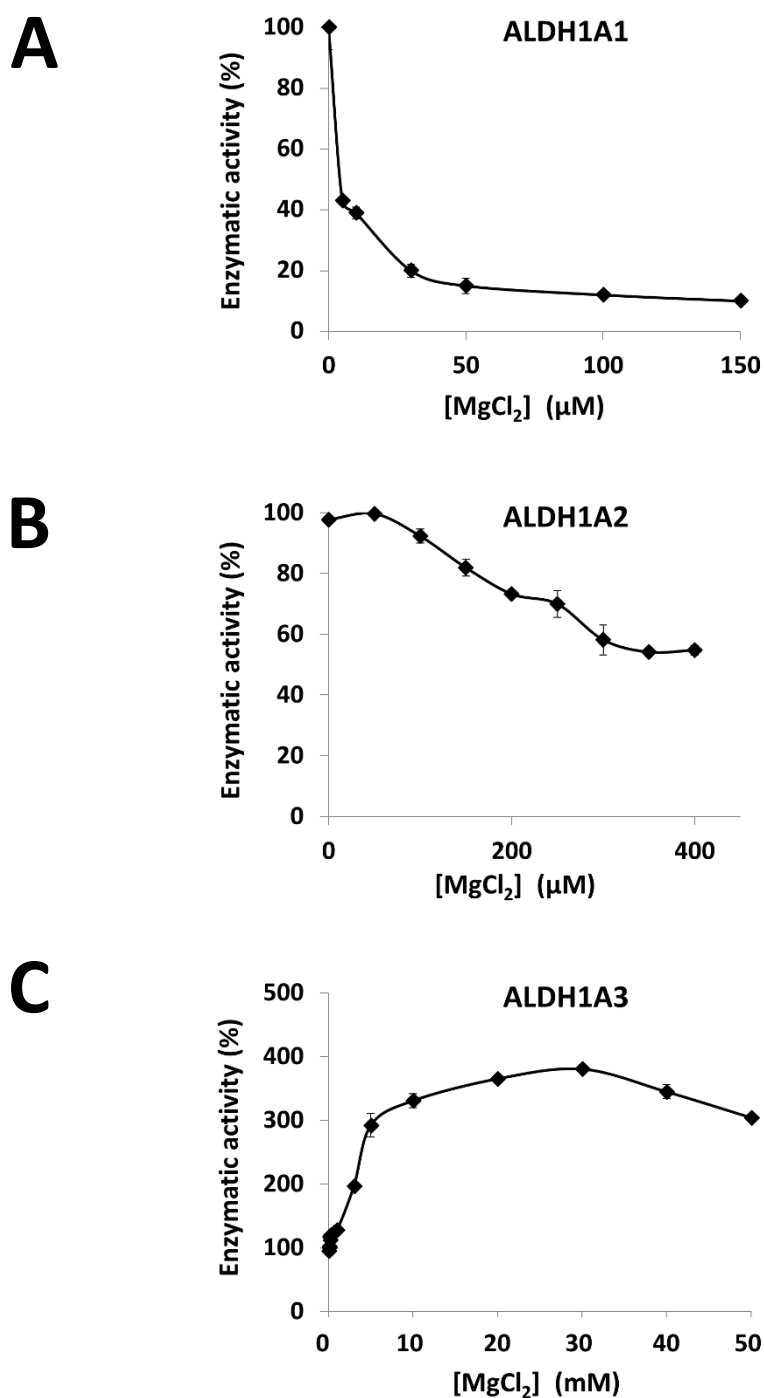
Figure 36. Effect of Mg^{2+} ions on the dehydrogenase activity of (A) ALDH1A1, (B) ALDH1A2 and (C) ALDH1A3. Enzymatic assays were carried out in 50 mM HEPES, 0.5 mM DTT, pH 8.0 (for ALDH1A1 and ALDH1A2) or 50 mM HEPES, 5 mM DTT, pH 8.0 (for ALDH1A3), with 0.5 mM NAD^+ and 250 μM hexanal, at 25°C.

3.6   Kinetic characterization of human ALDH1A enzymes with non-retinoid substrates

A wide survey of the substrate specificity of the ALDH1A forms was performed by determining their kinetic parameters with cofactor NAD^+ and a variety of aldehyde compounds, namely alkanals and alkenals, including lipid peroxidation products (Table 4). The three enzymes were active with all tested substrates. In terms of k_{cat}/K_m values, ALDH1A3 exhibits the lowest values for all substrates, suggesting a moderate role in the physiological oxidation of these aldehydes. The k_{cat}/K_m values of ALDH1A1 and ALDH1A2 indicate a potentially major role in the transformation of these substrates with slightly different substrate specificity. Among the three enzymes, ALDH1A1 showed low K_m and k_{cat} values for most of the substrates, while ALDH1A2 exhibited the highest k_{cat} values. With citral, which shares its isoprenoid nature with retinoids, ALDH1A1 showed the lowest K_m (0.085 µM) and k_{cat} (1 min^{-1}) values. In fact, citral is a slow substrate but with high affinity, and has often been used as a competitive inhibitor of ALDH1A1 [35]. Saturated aldehydes (hexanal and octanal) exhibited higher catalytic efficiency than their unsaturated counterparts (*trans*-2-hexenal and *trans*-2-octenal). Hexanal was the best substrate for the three enzymes. Regarding the aldehydes derived from lipid peroxidation, 4-hydroxy-2-hexenal and 4-hydroxy-2-nonenal, the activity of the ALDH1A forms is comparable to that other ALDHs [183], suggesting a shared role in the oxidation of these physiological substrates. ALDH1A2 is the best enzyme for 4-hydroxy-2-nonenal in terms of k_{cat}/K_m. Finally, the three enzymes showed esterase activity in the presence of NAD^+, being ALDH1A2 the form with the highest catalytic efficiency.

3.7   Optimization of the solvent extraction methodology for retinoid analysis

HPLC has become the method of choice for the analysis of retinoids. The characteristic UV absorption spectra of the various retinoids allow the use of diode-array detection for identification of resolved retinoids. However, the use of HPLC in the quantification of a broad range of biologically important retinoids in a single run, within a reasonable amount of time, presents some difficulties. It is necessary an analytically robust assay which has to combine the requirements of sensitivity, specificity and resolution to separate the different isomers and analytes of retinoids endogenously present in cells [180,184]. The HPLC/UV-based methodology, used by our group for the analysis and characterization of enzymatic reactions of different enzymes with retinoids *in vitro* [109,110,113,185–187], has been extensively tested and used, showing that the extraction efficiency for both retinaldehyde and retinol is robust and excellent. However, this methodology was not efficient enough in the extraction of all the retinoid metabolites, particularly of retinoic acid.

Table 4. Kinetic constants of ALDH1A1, ALDH1A2 and ALDH1A3 with non-retinoid substrates.

Substrate and parameter	ALDH1A1	ALDH1A2	ALDH1A3
COFACTOR			
NAD$^+$			
K_m (μM)	25 ± 5	19 ± 2	130 ± 16
k_{cat} (min^{-1})	27 ± 1	262 ± 7	166 ± 7
k_{cat}/K_m (mM^{-1}·min^{-1})	1,050 ± 220	13,800 ± 1,300	1,300 ± 170
ALKANALS			
Butanal			
K_m (μM)	3.4 ± 0.7	19 ± 2	41 ± 4
k_{cat} (min^{-1})	23 ± 1	260 ± 6	75 ± 2
k_{cat}/K_m (mM^{-1}·min^{-1})	6,700 ± 1,400	13,700 ± 1,500	1,700 ± 170
Hexanal			
K_m (μM)	0.22 ± 0.05	3.0 ± 0.2	6.0 ± 0.3
k_{cat} (min^{-1})	31 ± 2	240 ± 35	154 ± 16
k_{cat}/K_m (mM^{-1}·min^{-1})	143,500 ± 31,000	82,000 ± 13,000	25,600 ± 3,000
Octanal			
K_m (μM)	3.0 ± 0.3	3.0 ± 0.8	85 ± 15
k_{cat} (min^{-1})	26 ± 1	212 ± 13	88 ± 5
k_{cat}/K_m (mM^{-1}·min^{-1})	8,700 ± 900	70,700 ± 19,300	1,035 ± 192
ALKENALS			
***Trans*-2-hexenal**			
K_m (μM)	0.56 ± 0.12	16 ± 2	30 ± 4
k_{cat} (min^{-1})	49 ± 3	57 ± 2	31 ± 1
k_{cat}/K_m (mM^{-1}·min^{-1})	87,200 ± 19,000	3,600 ± 500	1,024 ± 148
***Trans*-2-octenal**			
K_m (μM)	4.0 ± 0.5	8.0 ± 2.0	41 ± 5
k_{cat} (min^{-1})	7.0 ± 0.2	21 ± 1	6.0 ± 0.2
k_{cat}/K_m (mM^{-1}·min^{-1})	1,900 ± 260	2,600 ± 670	150 ± 20
4-Hydroxy-2-hexenal			
K_m (μM)	13 ± 3	NS	458 ± 68
k_{cat} (min^{-1})	14 ± 1		7.4 ± 0.5
k_{cat}/K_m (mM^{-1}·min^{-1})	1,100 ± 300	134	16 ± 3
4-Hydroxy-2-nonenal			
K_m (μM)	27 ± 6	7.5 ± 0.7	40 ± 8
k_{cat} (min^{-1})	10 ± 1	31 ± 1	7.4 ± 0.7
k_{cat}/K_m (mM^{-1}·min^{-1})	380 ± 90	4,100 ± 400	185 ± 40
Citral			
K_m (μM)	0.085 ± 0.01	54 ± 10	13 ± 2
k_{cat} (min^{-1})	1.00 ± 0.02	84 ± 7	7.0 ± 0.3
k_{cat}/K_m (mM^{-1}·min^{-1})	12,400 ± 1,600	1,600 ± 320	570 ± 90
ESTERASE ACTIVITY			
4-Nitrophenyl acetate			
K_m (μM)	56 ± 17	30 ± 2	59 ± 8.4
k_{cat} (min^{-1})	40 ± 4.6	85 ± 2	26 ± 0.5
k_{cat}/K_m (mM^{-1}·min^{-1})	710 ± 240	2,800 ± 200	450 ± 60

Enzymatic activity was measured fluorimetrically at 25°C. ALDH1A1 and ALDH1A2 were assayed in 50 mM HEPES, 0.5 mM EDTA, 0.5 mM DTT, pH 8.0. ALDH1A3 assays were performed in 50 mM HEPES, 30 mM MgCl$_2$, 5 mM DTT, pH 8.0. NAD$^+$ concentration was 500 μM. To calculate k_{cat} values the molecular weight value used for ALDH1A1 was 220,000 g/mol, for ALDH1A2 was 226,800 g/mol and for ALDH1A3 was 224,000 g/mol. NAD$^+$ constants were determined using 250 μM hexanal as a substrate; NS, not saturated up to 375 μM 4-hydroxy-2-hexenal. The corresponding k_{cat}/K_m value of ALDH1A2 for 4-hydroxy-2-hexenal was determined from the slope of V/[E] *versus* [S] linear plot.

At this point, different solvent mixtures were tested to optimize retinoic acid recovery after the enzymatic reaction: hexane/dioxane/isopropanol (50:5:1) (v/v/v) [178,179], methanol/hexane (1:4) (v/v) [109], ethyl acetate [188], hexane/ethyl acetate (9:1) (v/v) [189], ethanol/hexane (1:4) (v/v) [190], and a two-step acid-base extraction [191]. From the evaluated methods, the two-step acid-base extraction was the most efficient to recover multiple retinoids, with a yield near 100% (Figure 37A, lane b). The basification step with 25 mM KOH recovers nonpolar retinoids (retinol and retinyl esters), and the acidification step with 1 M HCl (pH 1.0) facilitates the recovery of retinoic acid and other polar retinoids [180,191–193].

We applied the two-step methodology for the recovery of retinoids from the cell media and enzymatic reaction buffer (50 mM HEPES, 0.5 mM EDTA, 0.5 mM DTT, pH 8.0), with good results. Owing the extreme acidic pH to achieve an efficient retinoic acid recovery, the silica resin constituting the HPLC column was altered, and the results obtained in terms of retention time and absorbance spectra of the retinoid compounds, were not satisfactory and anomalous peaks appeared. Furthermore, it is known that retinoids are labile to strong acids, particularly under anhydrous conditions [194,195]. The UV spectra of these anomalous peaks did not correspond to any retinoid standard. Moreover, the retention time of the retinoid peaks was different from its standard value, and even we observed a decrease in the absorbance of some nonpolar retinoid peaks followed by the appearance of anomalous peaks. We concluded that the column and the retinoid compounds became damaged as a result of the methodology used. In view of these results, we tried another approach based on the acidification protocol by Kane *et al*. [191], with minor modifications. Rather than using a strong acid like HCl, we performed a single extraction step with a mild acid such as 2.5 mM ammonium acetate, pH 4.5 [189,191]. This new methodology highly improved the efficiency of retinoic acid extraction, both in culture media and enzymatic reaction buffer (Figure 37A, lane h, and Figure 37B, lanes b, d, f, h and j), without altering the retention times and absorbance spectra of the retinoid compounds. In addition, in the samples from the enzymatic reaction buffer, the solvent mixture hexane/dioxane/isopropanol (50:5:1) (v/v/v) provided a retinoic acid extraction efficiency near 100%, even without the treatment with 2.5 mM ammonium acetate. Thus, we decided to carry out the retinoic acid extraction with this solvent mixture when the enzymatic activity *in vitro* was tested. Apparently, the isopropanol present in the organic solvent mixture allowed us to recover the polar retinoids with the highest efficiency.
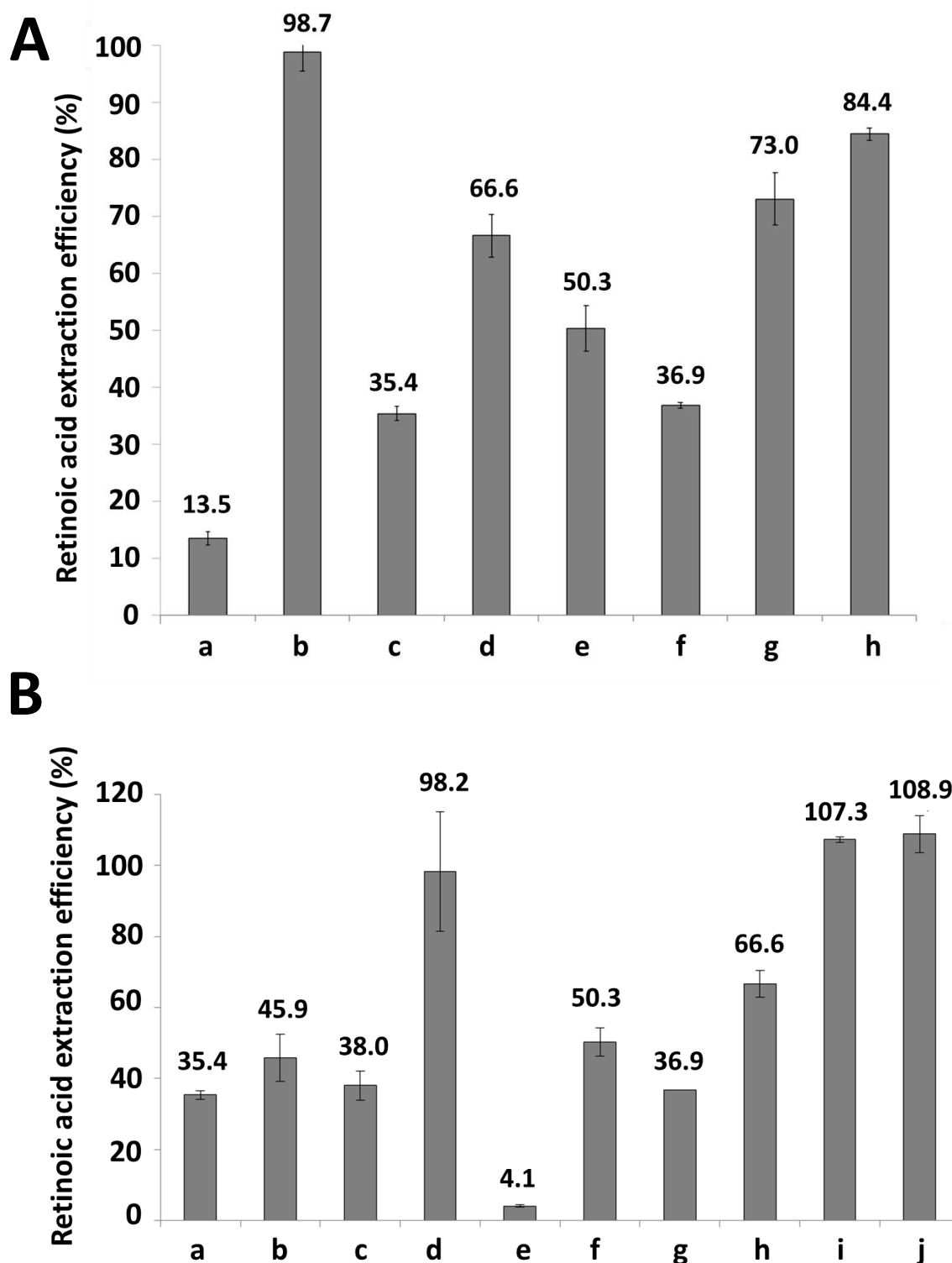
Figure 37. Extraction efficiency of retinoic acid from (A) cell media and from (B) enzymatic reaction buffer (50 mM Hepes, 0.5 mM EDTA, 0.5 mM DTT, pH 8.0), using different solvents. (A) Lanes: (a) extraction with methanol/hexane (1:4), the classical method for the extraction of retinaldehyde and retinol from *in vitro* enzymatic reactions; (b) two-step acid-base extraction; (c) hexane; (d) hexane/ethyl acetate (9:1); (e) ethanol/hexane (1:4); (f) ethyl acetate; (g) hexane/dioxane/isopropanol (50:5:1); (h) acidification of the media with 2.5 mM ammonium acetate (pH 4.5) plus methanol/hexane (1:4). **(B)** Lanes: (a) hexane; (b) acidification of the media with 2.5 mM ammonium acetate (pH 4.5) plus hexane; (c) methanol/hexane (1:4); (d) acidification of the media with 2.5 mM ammonium acetate (pH 4.5) plus methanol/hexane (1:4); (e) acetonitrile plus hexane; (f) acidification of the media with 2.5 mM ammonium acetate (pH 4.5) plus acetonitrile plus hexane; (g) ethyl acetate; (h) acidification of the media with 2.5 mM ammonium acetate (pH 4.5) plus ethyl acetate; (i) hexane/dioxane/isopropanol (50:5:1); (j) acidification of the media with 2.5 mM ammonium acetate (pH 4.5) plus hexane/dioxane/isopropanol (50:5:1).

3.8 Kinetic characterization of human ALDH1A enzymes with retinaldehyde

The activity with retinoid molecules was determined by using an HPLC-based method to analyze the substrates and reaction products (Figure 38). As already mentioned in the previous section, different solvent mixtures were tested to optimize retinoid recovery after the enzymatic reaction. From the evaluated methods, the extraction with hexane/dioxane/isopropanol (50:5:1) (v/v/v) was chosen since it was the most efficient in recovering retinaldehyde and retinoic acid from the activity buffer, with a yield near 100%. To our knowledge, this is the first time that this extraction method and the use of BSA in the activity assay have been applied to determine the enzymatic activity of ALDH1A enzymes with retinoids. In all cases, concomitant retinaldehyde consumption and retinoic acid production, under steady state conditions, could be measured. Table 5 shows the kinetic constants of the ALDH1A enzymes with all-*trans* and 9-*cis*-retinaldehyde. ALDH1A enzymes were fairly active with the two isomers and followed Michaelis-Menten kinetics, with K_m values in the micromolar range. The three forms exhibited lower K_m values for 9-*cis*-retinaldehyde than for all-*trans*-retinaldehyde. Regarding the k_{cat} values, they were higher for the all-*trans* isomer (ALDH1A2 and ALDH1A3) or similar for the two isomers (ALDH1A1). This results in, that for the most physiological isomer, all-*trans*-retinaldehyde, ALDH1A3 is the best enzyme in terms of k_{cat}/K_m, followed by ALDH1A2. Conversely, the k_{cat} value and the catalytic efficiency of ALDH1A3 are lowest for the 9-*cis* isomer. The all-*trans*-retinaldehyde specificity of ALDH1A3 within the ALDH1A subfamily is a remarkable feature given the fact that this enzyme exhibited the lowest k_{cat}/K_m values with all examined non-retinoid substrates. Overall, the k_{cat}/K_m values were within the upper range for most aldehyde substrates.

Table 5. Kinetic constants of ALDH1A1, ALDH1A2 and ALDH1A3 with retinaldehyde isomers.

Substrate and parameter	ALDH1A1	ALDH1A2	ALDH1A3
all-*trans*-retinaldehyde			
K_m (µM)	11 ± 2	4.0 ± 1.0	2.4 ± 0.5
k_{cat} (min^{-1})	18 ± 2	20 ± 2	18 ± 1
k_{cat}/K_m (mM^{-1}·min^{-1})	1,600 ± 330	5,000 ± 1,300	7,600 ± 1,600
9-*cis*-retinaldehyde			
K_m (µM)	5.0 ± 1.5	1.2 ± 0.2	1.0 ± 0.1
k_{cat} (min^{-1})	20 ± 2	3.0 ± 0.1	0.090 ± 0.002
k_{cat}/K_m (mM^{-1}·min^{-1})	4,000 ± 1,400	2,500 ± 320	90 ± 13

Enzymatic activity was measured at 37°C by using the HPLC-based method. The assay buffer for each enzyme was the same as that for the fluorimetric assays.
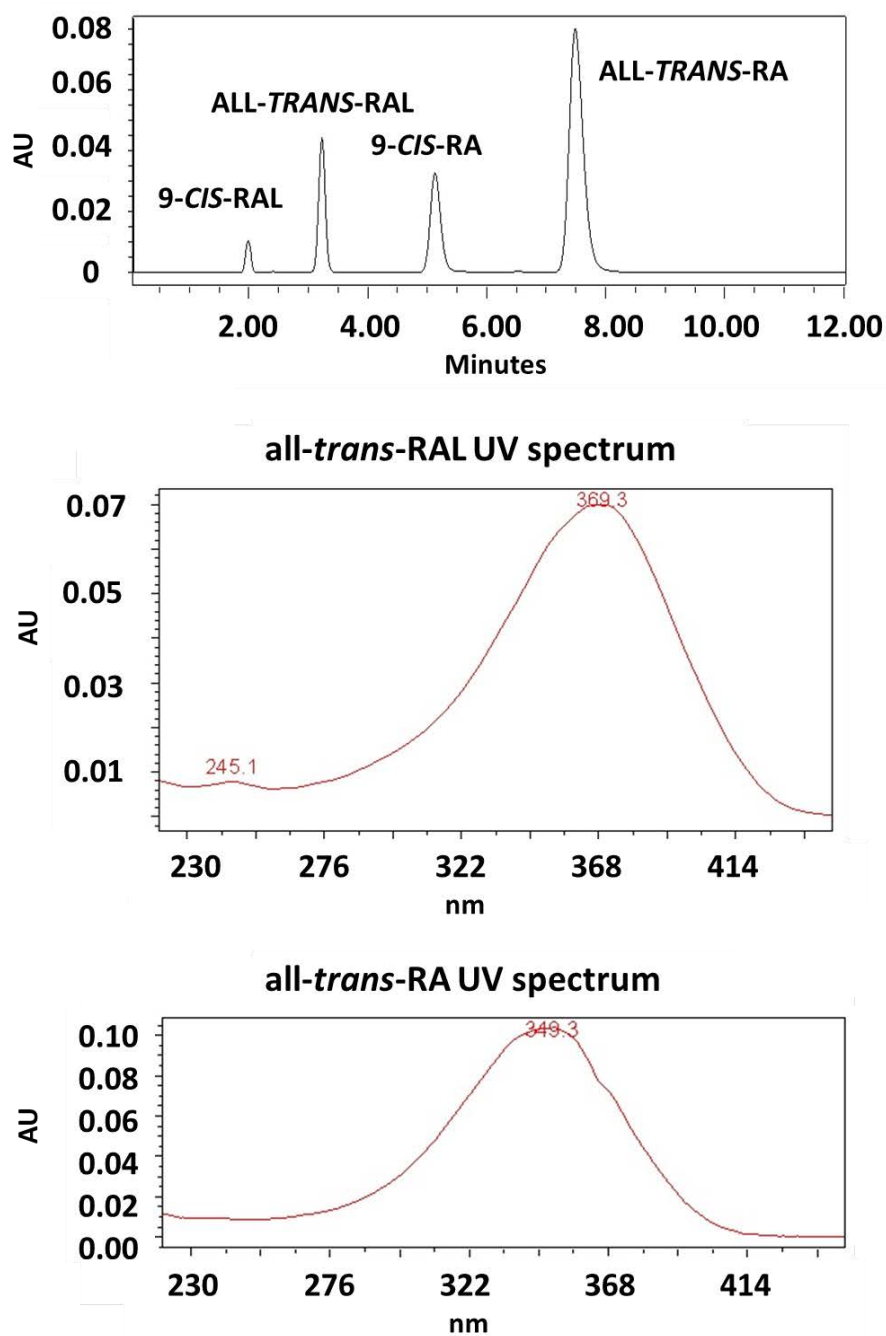
Figure 38. HPLC-based method for ALDH activity determination with retinaldehyde. (A) Elution profile of all-*trans*- and 9-*cis*-retinoids and UV-vis spectra for (B) all-*trans*-retinaldehyde and (C) all-*trans*-retinoic acid. The HPLC chromatogram shows the separation of all-*trans* and 9-*cis*-retinaldehyde substrates and the corresponding retinoic acid products after the retinaldehyde dehydrogenase reaction catalyzed by ALDH1A1. Chromatogram recorded at 350 nm.

3.9   Kinetic characterization of human ALDH1A enzymes with 14'- and 12'-apo-β-carotenals

The activity of ALDH1A enzymes with apo-β-carotenals (14'- and 12'-apo-β-carotenals) was monitored using the HPLC-based method described above in section 2.10.1. The simultaneous detection of substrates and reaction products was performed. In all cases, concomitant consumption of 14'- and 12'-apo-β-carotenals, and production of the corresponding 14'- and 12'-apo-β-carotenoic acids, respectively, were followed [196]. Thus, ALDH1A enzymes turned out to be active with 14'- and 12'-apo-β-carotenals.

Table 6 compares the kinetic constants of the ALDH1A enzymes with 12'- and 14'-apo-β-carotenals, showing some differential substrate specificity. In general, K_m values were in the micromolar range, except for ALDH1A1 and ALDH1A3, which were not saturated with 12'-apo-β-carotenal and 14'-apo-β-carotenal, respectively. While ALDH1A1 showed higher specificity for 14'-apo-β-carotenal, ALDH1A2 exhibited higher catalytic efficiency for 12'-apo-β-carotenal. This catalytic efficiency is about ten-fold higher than those of ALDH1A1 and ALDH1A3, and similar to that observed for these enzymes with all-*trans*-retinaldehyde. The k_{cat} values are in the same range in all the experiments performed, but lower than those for all-*trans*-retinaldehyde. To date, the enzymatic activity of ALDH with these compounds had only been reported in microorganisms [197]. This is the first report on the apocarotenal dehydrogenase activity of human ALDH1A enzymes [196].

Table 6. Kinetic constants of human ALDH1A enzymes with apo-β-carotenals.

Substrate and parameter	ALDH1A1	ALDH1A2	ALDH1A3
12'-apo-β-carotenal			
K_m (µM)	NS	1.8 ± 0.2	3.0 ± 0.4
k_{cat} (min^{-1})		8.5 ± 0.2	1.3 ± 0.1
k_{cat}/K_m (mM^{-1}·min^{-1})	44	$4{,}800 \pm 600$	450 ± 62
14'-apo-β-carotenal			
K_m (µM)	12.0 ± 2	5.0 ± 0.5	NS
k_{cat} (min^{-1})	3.1 ± 0.2	4.7 ± 0.2	
k_{cat}/K_m (mM^{-1}·min^{-1})	257 ± 47	930 ± 93	303

Enzymatic activity was measured in 50 mM HEPES, 0.5 mM NAD$^+$, including 0.5 mM EDTA and 0.5 mM DTT, pH 8.0, for ALDH1A1 and ALDH1A2, or 30 mM MgCl$_2$ and 5mM DTT for ALDH1A3, at 37°C. Apocarotenoic acid production was quantified by using an HPLC-based method. NS, no saturation (60 µM was the highest substrate concentration tested in the assay), where the k_{cat}/K_m value was calculated from the slope of V/[E] *vs* [S] plot.

3.10   Kinetic characterization of mutant ALDH1A enzymes

In order to test the role of specific residues in determining the kinetic properties of ALDH1A enzymes, we performed site-directed mutagenesis of selected residues based on structural differences (section 3.1). The substitution L114P, defining one site of the substrate-binding pocket, was selected for a mutant ALDH1A1 to make this part of the structure more similar to that of ALDH1A2. Concurrently, in ALDH1A2, four contiguous residue changes, N475G, A476V, L477V and N478S (corresponding to the region of the loop) were made to mimic the structure of ALDH1A1. The cofactor NAD^+, three typical aldehyde substrates and both retinaldehyde isomers were tested. Table 7 shows the comparison of the kinetic constants of the wild-type enzymes and their corresponding mutants. Cofactor kinetics did not change in the mutants, suggesting that NAD^+ binding was not affected by the mutations.

Table 7. Kinetic constants of mutant ALDH1A1 and ALDH1A2.

Substrate and parameter	ALDH1A1	ALDH1A1 L114P	ALDH1A2	ALDH1A2 mutant[a]
NAD^+				
K_m (µM)	25 ± 5	74 ± 5	19 ± 2	8 ± 0.5
k_{cat} (min^{-1})	27 ± 1	28 ± 0.6	262 ± 7	470 ± 6
k_{cat}/K_m ($mM^{-1} \cdot min^{-1}$)	1,050 ± 220	380 ± 27	13,800 ± 1,300	56,000 ± 5000
Butanal				
K_m (µM)	3.4 ± 0.7	9.0 ± 1.0	19 ± 2	38 ± 2
k_{cat} (min^{-1})	23 ± 1	58 ± 2	260 ± 6	290 ± 3
k_{cat}/K_m ($mM^{-1} \cdot min^{-1}$)	6,700 ± 1400	6,300 ± 820	13,700 ± 1,500	7,500 ± 440
Hexanal				
K_m (µM)	0.22 ± 0.50	11 ± 1	3.0 ± 0.2	13 ± 1
k_{cat} (min^{-1})	31 ± 2	32 ± 1	240 ± 35	460 ± 6
k_{cat}/K_m ($mM^{-1} \cdot min^{-1}$)	143,500 ± 31,000	2,900 ± 190	82,000 ± 13,000	36,000 ± 1,700
Citral				
K_m (µM)	0.085 ± 0.010	8.0 ± 1.0	54 ± 10	1.0 ± 0.1
k_{cat} (min^{-1})	1.0 ± 0.1	10.0 ± 0.1	84 ± 7	18 ± 0.3
k_{cat}/K_m ($mM^{-1} \cdot min^{-1}$)	12,400 ± 1,600	1,200 ± 100	1,600 ± 320	16,000 ± 1,500
All-*trans*-retinaldehyde				
K_m (µM)	11 ± 2	20 ± 4	4.0 ± 1.0	5.0 ± 2.0
k_{cat} (min^{-1})	18 ± 2	19 ± 1.5	20 ± 2	53 ± 7
k_{cat}/K_m ($mM^{-1} \cdot min^{-1}$)	1,600 ± 330	940 ± 210	5,000 ± 1,300	10,600 ± 4000
9-*Cis*-retinaldehyde				
K_m (µM)	5 ± 1.5	5 ± 1.2	1.2 ± 0.2	3 ± 0.5
k_{cat} (min^{-1})	20 ± 2	20 ± 1	3 ± 0.1	3.5 ± 0.14
k_{cat}/K_m ($mM^{-1} \cdot min^{-1}$)	4,000 ± 1,400	4,000 ± 993	2,500 ± 320	1,200 ± 200

[a]This mutant included the following substitutions: N475G, A476V, L477V and N478S. Enzymatic activity was measured by the HPLC-based method for all-*trans*-retinaldehyde and fluorimetrically for alkanals and citral. NAD^+ constants were determined using 250 µM hexanal as a substrate. To calculate k_{cat} values, the molecular weight value used, for wild-type and mutant ALDH1A1, was 220,000 g/mol and, for wild-type and mutant ALDH1A2, was 226,800 g/mol.

For the ALDH1A1 L114P mutant, there was a significant increase (50-100 fold) in the K_m values for hexanal and citral. This caused the catalytic efficiency for these substrates, especially citral, to be more similar to that of wild-type ALDH1A2. Conversely, the major effect observed with the mutant ALDH1A2 was a 50-fold decrease in the K_m value for citral. This decline, together with the 5-fold decrease in the k_{cat} value, implies that the catalytic efficiency of mutant ALDH1A2 for citral was closer to that observed in wild-type ALDH1A1. Regarding kinetics with retinaldehyde isomers, the mutants did not show significant differences with the respective wild-type forms, suggesting that the mutated residues are not critical for retinaldehyde specificity.

3.11   Determination of the selectivity of various inhibitors for ALDH1A enzymes

We first tested, with each one of the ALDH1A enzymes, some compounds already know to be good ALDH inhibitors, such as DEAB and WIN 18,446. DEAB is the compound for which a lowest IC_{50} is obtained against ALDH1A1. Furthermore, it is not a good inhibitor neither of ALDH1A2 nor of ALDH1A3. In contrast, WIN 18,446 is a potent inhibitor against ALDH1A2, with an IC_{50} value in nanomolar range, while the IC_{50} values for ALDH1A1 and ALDH1A3 are in micromolar range. Therefore, we can state that DEAB and WIN 18,446 are potent and selective inhibitors for ALDH1A1 and ALDH1A2, respectively (Table 8 and Figure 40).

**Table 8. IC_{50} values (µM) of DEAB and WIN 18,446 against ALDH1A enzymes.**

Compound	ALDH1A1	ALDH1A2	ALDH1A3
DEAB	0.18 ± 0.05	9.5 ± 3	49.5 ± 12
WIN 18,446	56.4 ± 2.2	0.07 ± 0.009	30.6 ± 8.2

IC_{50} values were obtained fluorimetrically at 25°C. ALDH1A1 and ALDH1A2 were assayed in 50 mM HEPES, 0.5 mM EDTA, 0.5 mM DTT, pH 8.0. ALDH1A3 assays were performed in 50 mM HEPES, 30 mM MgCl$_2$, 5 mM DTT, pH 8.0. NAD$^+$ concentration was 500 µM. Experimental values are shown as the mean ± standard error.

In order to find novel, potent and selective inhibitors against ALDH1As, a screening was carried out with some compounds that had been described as inhibitors of other oxidoreductases. In particular, compounds that had shown inhibitory capacity against human AKRs were tested. The reason why these compounds were chosen is because some AKR1Bs and the ALDH1As share substrate specificity for all-*trans*-retinaldehyde [109]. Many compounds had been developed as inhibitors of the AKR1B subfamily with the purpose to use them to treat the pathologies in which AKR1B1 and AKR1B10 are involved, such as diabetes

and cancer. However, very few of them showed sufficient therapeutic efficacy in clinical trials. Some inhibitors, like tolrestat, were commercialized for some time, but most of the drug candidates failed because of their poor pharmacokinetic properties and/or unacceptable side effects, mainly due to their toxicity and unspecific *in vivo* action [198–200].

Recently, a new series of compounds named 1-oxopyrimido[4,5-c]quinolone-2-acetic acid and thiazolidinedione-acetic acid derivatives, provided by Dr. Kamel Metwally, have been described as AKR inhibitors [201,202]. Some of these compounds have been selected to test their inhibitory capacity on ALDH1As (Tables 9-10). Similarly, the series of JF inhibitors, provided by Biomar Microbial Technologies S.A. and synthesized by de la Fuente and coworkers [203], were tested in order to determine the inhibitory capacity against ALDH. Marine natural products were the initial point for the synthesis of this series of polyhalogenated derivatives [204] (Figure 39).
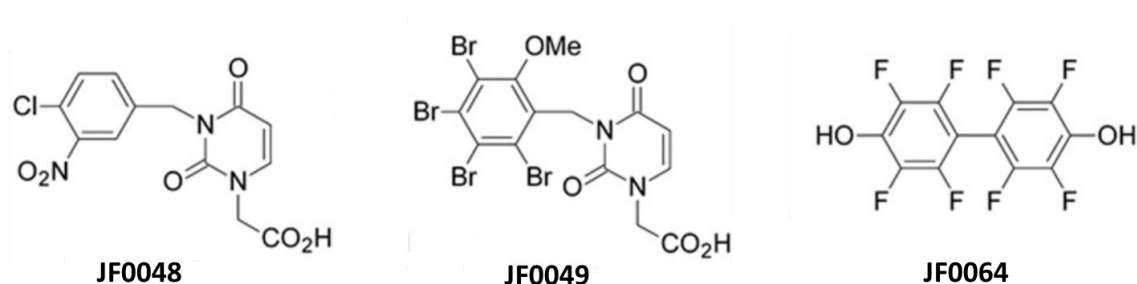
Figure 39. The series of JF compounds. The structures were reproduced from [205].

The results obtained in a quick inhibition screening with 10 μM compound are listed in the Table 11. Regarding the 1-oxopyrimido[4,5-*c*]quinoline-2-acetic acid series, compounds **64** and **65**, exhibited a high inhibitory capacity against ALDH1A2 and ALDH1A3 activities, being more potent on ALDH1A3. Compound **66** showed a high inhibitory capacity against the three ALDH1A enzymes, thus being potent but not a selective ALDH1A inhibitor. Compound **70** was a potent inhibitor of ALDH1A2, with only 33% of enzymatic activity remaining in the presence of 10 μM of this compound. Compounds **37-48** had little or null inhibitory effect on ALDH1As.

Table 9. Molecular structure of 1-oxopyrimido[4,5-c]quinoline-2-acetic acid compounds.
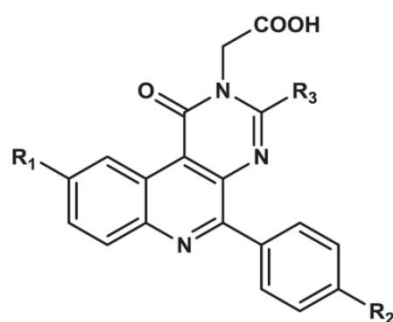
Compound	Mr	R_1	R_2	R_3
37	379.80	H	Cl	CH_3
38	424.25	H	Br	CH_3
39	397.79	F	Cl	CH_3
40	442.24	F	Br	CH_3
41	393.37	F	OCH_3	CH_3
42	414.24	Cl	Cl	CH_3
43	458.69	Cl	Br	CH_3
44	409.82	Cl	OCH_3	CH_3
45	458.69	Br	Cl	CH_3
46	503.14	Br	Br	CH_3
47	393.82	CH_3	Cl	CH_3
48	409.82	OCH_3	Cl	CH_3
64	455.89	H	Cl	ethylbenzene
65	490.34	H	Cl	4-chlorobenzyl
66	455.89	H	H	4-chlorobenzyl
67	345.35	H	H	CH_3
68	393.82	H	Cl	CH_2CH_3
69	407.85	H	Cl	$CH_2CH_2CH_3$
70	411.81	F	Cl	CH_2CH_3

The different substituent in R_1, R_2 and R_3 is detailed.

Table 10. Molecular structure of thiazolidinedione-acetic acid compounds.
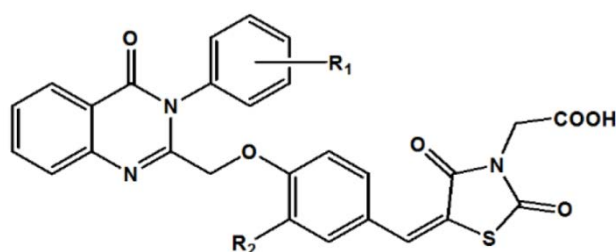
Compound	Mr	R_1	R_2
5a	513.52	H	H
5b	531.51	4-F	H
5c	547.97	3-Cl	H

The different substituent in R_1 and R_2 is detailed.

Related to the thiazolidinedione-acetic acid series, compound **5b** did not show any potent inhibitory effect on ALDH1A1, ALDH1A2 or ALDH1A3. However, compound **5c** was a good inhibitor of ALDH1A3, decreasing the enzymatic activity of this enzyme down to approximately 40%. Finally, is important to remark that compound **5a** reduced the ALDH1A1 enzymatic activity down to 10%, becoming the compound with the highest inhibitory potency against ALDH1A1. Nevertheless, this inhibitor was not selective against ALDH1A1, since it decreased the ALDH1A3 activity down to 25%.

Table 11. Inhibitory effect of various compounds at 10 μM on ALDH1A enzymes.

Compound	ALDH1A1	ALDH1A2	ALDH1A3
37	118	107	105
38	89	101	92
39	109	101	107
40	94	137	96
41	110	80	89
42	68	81	97
43	77	103	75
44	104	104	92
45	97	104	87
46	68	101	82
47	73	102	92
48	88	103	122
64	89	**38**	**11**
65	85	**34**	**10**
66	**23**	**20**	**0**
67	96	100	86
68	90	85	53
69	90	84	69
70	79	**35**	79
5a	**11**	62	**24**
5b	58	89	63
5c	94	93	**39**
JF0048	100	106	99
JF0049	97	102	125
JF0064	**40**	60	52

The inhibitory effect is expressed as percentage of remaining activity (%). In red and bold, activity values below 50%.

Those compounds resulting in remaining activity values below 50% were selected for further studies on the determination of IC_{50} values (Table 12).

Table 12. IC$_{50}$ values (µM) of selected compounds against ALDH1A enzymes.

Compound	ALDH1A1	Compound	ALDH1A2	Compound	ALDH1A3
66	16.27 ± 2.99	64	3.5 ± 0.8	64	1.17 ± 0.11
5a	5.42 ± 0.45	65	1.88 ± 0.39	65	0.82 ± 0.11
JF0064	10.5 ± 0.9	66	5.47 ± 0.71	66	2.45 ± 0.24
		70	5.49 ± 1.58	5a	22.6 ± 4.04
				5c	13.27 ± 1.32

IC$_{50}$ values were obtained fluorimetrically at 25°C. ALDH1A1 and ALDH1A2 were assayed in 50 mM HEPES, 0.5 mM EDTA, 0.5 mM DTT, pH 8.0. ALDH1A3 assays were performed in 50 mM HEPES, 30 mM MgCl$_2$, 5 mM DTT, pH 8.0. NAD$^+$ concentration was 500 µM. Experimental values are shown as the mean ± standard error.

In summary, ALDH1A1 was significantly inhibited by compounds **66** and **5a**, while ALDH1A2 was inhibited by compounds **64**, **65**, **66** and **70**, and ALDH1A3 was inhibited by compounds **64**, **65**, **66**, **5a** and **5c**.

Related to the results observed in ALDH1A3, it is remarkable that a potent or selective inhibitor for ALDH1A3 has not been described to date. In the results shown in Tables 10-11, it is noteworthy that the compounds named **64**, **65** and **66** exhibit an important effect on the ALDH1A3 enzymatic activity, with IC$_{50}$ values in the micromolar range. Although these compounds also inhibit ALDH1A2, the IC$_{50}$ values are always lower for ALDH1A3. Thus, we can state that, for the first time, very potent inhibitors for ALDH1A3 isozyme have been described (Figure 40C).

ALDH1A1
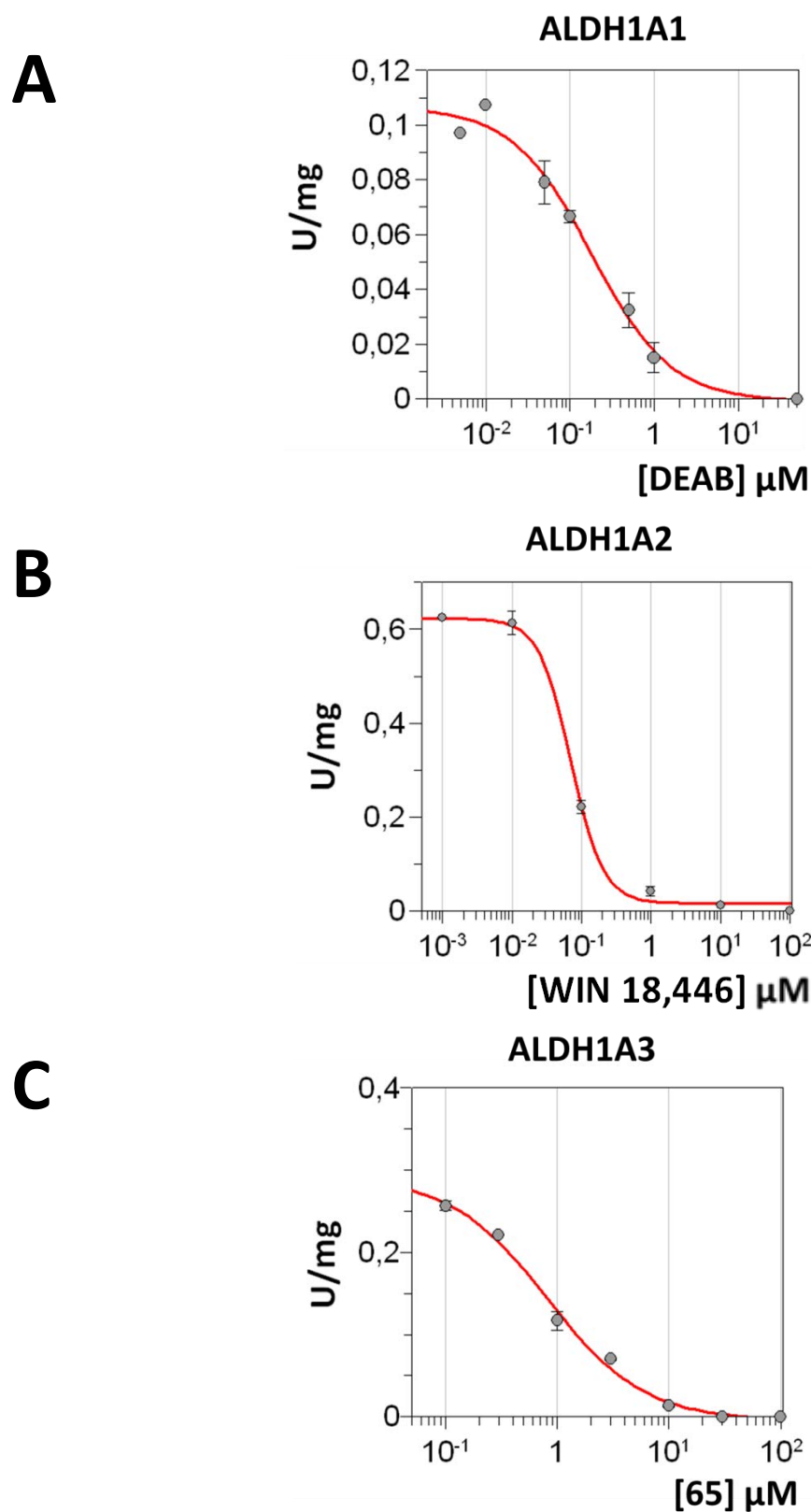
ALDH1A2

ALDH1A3

Figure 40. IC$_{50}$ plots of ALDH1As with the most potent inhibitors. (A) ALDH1A1 with DEAB; (B) ALDH1A2 with WIN 18,446; (C) ALDH1A3 with compound **65**. IC$_{50}$ values were obtained fluorimetrically at 25°C. ALDH1A1 and ALDH1A2 were assayed in 50 mM HEPES, 0.5 mM EDTA, 0.5 mM DTT, pH 8.0. ALDH1A3 assays were performed in 50 mM HEPES, 30 mM MgCl$_2$, 5 mM DTT, pH 8.0. NAD$^+$ concentration was 500 µM. The substrate of the reaction was hexanal at saturated concentrations (10 µM for ALDH1A1 and 250 µM for ALDH1A2 and ALDH1A3). Experimental values are shown as the mean ± standard error.

- 102 -

DISCUSSION

- 104 -

The enzymes from the ALDH1A subfamily are NAD^+-dependent cytosolic retinaldehyde dehydrogenases involved in the last metabolic step of retinoic acid biosynthesis. Their essential role in retinoic acid homeostasis and signaling has been widely shown by knockout mice studies [206] and, more recently, by their involvement in human genetic diseases [69,207]. In the present Thesis, the three human ALDH1A enzymes, ALDH1A1, ALDH1A2 and ALDH1A3, were recombinantly expressed in a heterologous system and purified in soluble and enzymatically active forms. For the first time, a structural and functional characterization of the three human enzymes of ALDH1A subfamily has been performed side-by-side and using a similar methodology.

The recent reports on the crystallographic structures of the human ALDH1A enzymes have enabled the close comparison of their three-dimensional features. As expected from their overall amino acid sequence identity of 70%, the three enzymes exhibit a high degree of structural similarity but demonstrate significant differences in their substrate-binding pockets. Although ALDH1A1, ALDH1A2 and ALDH1A3 show similar topologies, they have decreasing enclosed volumes in their substrate-binding pockets, with ALDH1A2 showing a distinct subpocket associated to its 471-495 loop. Most residues in the active-site channel are conserved: nine out of twelve residues, between ALDH1A2 and ALDH1A3, while only five residues are shared with ALDH1A1 (Table 3). Most residue differences in ALDH1A1 are concentrated in the loop region. Ile132, Gly136, Arg139, Thr140, Trp189, Leu471 and Ala473 establish van der Waals interactions with retinoic acid in ALDH1A3 [182]. Three residues, 125, 460 and 304 (ALDH1A1 numbering), located at the channel entrance, bottom and neck, respectively, were previously pointed as important for the substrate-binding pocket to determine the ability to harbor either small or large substrates in ALDH1 *vs* ALDH2 enzymes [25]. Importantly, residue 304 interacts with retinoic acid in ALDH1A3 [182]. Residue 125 remains an invariable Gly in the three ALDH1A enzymes and residue 460 shows a conservative substitution (a hydrophobic Val or Leu), but residue 304 involves a significant exchange (Ile/Thr) (Table 3). Residues 458 (Gly in ALDH1A1, and Asn in ALDH1A2 and ALDH1A3) and 461 (Ser in ALDH1A1, Asn in ALDH1A2 and Tyr in ALDH1A3) display even more remarkable substitutions. Such variability may be important in shaping the enclosed volume and topology of the substrate-binding site, and in modulating substrate specificity.

The three enzymes bound and used NAD^+ with high efficiency as a cofactor. Their activity was analyzed with typical aliphatic aldehydes (alkanals and alkenals) (Table 4), as well as with retinaldehyde (Table 5). Among the three enzymes, ALDH1A1 showed the most distinct kinetic

properties, with the lowest K_m and k_{cat} values for all the substrates examined. With citral (3,7-dimethyl-2,6-octadienal), which shares its isoprenoid nature with retinoids, ALDH1A1 showed the lowest K_m (0.085 μM) and k_{cat} (1 min^{-1}) values. In fact, citral was often used as a competitive inhibitor of ALDH1A1 [35] and citral specificity for ALDH1A1 could be linked in part to residue 114 (ALDH1A1 L114P mutant) and to residues 458 to 461 (ALDH1A2 mutant). Interestingly, no major differences were detected between saturated (hexanal) and unsaturated (*trans*-2-hexenal) compounds with the same number of carbons. The differences in k_{cat} values between isoprenoids and medium-chain alkanals and alkenals could be attributed to a distinct rate-limiting step in the catalysis. The low k_{cat} values are typically observed in ALDH1A enzymes with retinaldehyde and other isoprenoid compounds [208].

ALDH1A2 showed a low K_m and the highest k_{cat} value with hexanal, which made this substrate to be the most specific for this enzyme, while *trans*-2-hexenal was much less specific. In this case, K_m and k_{cat} values for citral were much higher than those of ALDH1A1, being in the same order of magnitude as for medium-chain aliphatic substrates.

ALDH1A3 displayed the highest K_m values with all the analyzed substrates (except with citral). More differences were detected between saturated and unsaturated compounds than in other enzymes. Thus, ALDH1A3 showed a five times higher K_m for *trans*-2-hexenal than for hexanal, while the k_{cat} value was five times lower. The enzyme showed the k_{cat} values in the same order of magnitude as ALDH1A1, distancing itself from the high values observed for ALDH1A2.

Given the relevance of ALDH1A enzymes to retinoic acid production, we also investigated the contribution of each enzyme to retinaldehyde oxidation. In order to determine retinaldehyde activity, we followed, for the first time in ALDH studies, a well-proven strategy for retinoid solubilization (with fatty acid-free BSA) [209] and an improved HPLC-based method to unequivocally monitor substrate disappearance and product formation using an end-point assay. This combined approach allowed us to work with a very low concentration of substrate and in the absence of any added detergent or organic solvent. This is not feasible with other methodologies, such as UV-vis spectrophotometry, which have been classically used in previous studies [56,167,182,210]. The three ALDH1A enzymes are active with the two physiological retinaldehyde isomers, meaning that their distinct active-sites allow binding of retinoid molecules, as directly observed by retinoic acid binding in the ALDH1A3 crystallographic structure [182]. Residues Asn469, Leu471 and Ala473 have been pointed out

as making hydrophobic contacts with the retinoid molecule in the ALDH1A3 structure. Only Ala473 is strictly conserved in the three ALDH1A enzymes, while Asn469 and Leu471 are conserved in ALDH1A2 (Table 3). Despite of exchanging some of these residues important for retinoid binding, between ALDH1A1 and ALDH1A2, no major effect was observed in the kinetic constants for all-*trans*-retinaldehyde, supporting the idea that some variability is allowed within the loop residues without compromising retinoid binding. This variability may be linked with the retinoid isomer specificity. In this regard, ALDH1A3 is truly specific for all-*trans*-retinaldehyde, as its catalytic efficiency with this substrate is 80-fold higher than with 9-*cis*-retinaldehyde. The unfavorable kinetics of ALDH1A3 with the 9-*cis* isomer may be explained by steric hindrance due to the smaller volume of the substrate-binding pocket, compared to those of ALDH1A1 and ALDH1A2. ALDH1A3 appears to be the most conspicuous all-*trans*-retinaldehyde dehydrogenase since it exhibits the highest k_{cat}/K_m value among the ALDH1A forms, and this substrate is the most efficiently catalyzed by ALDH1A3 other than hexanal.

We have shown that human retinaldehyde dehydrogenases are capable of using apo-β-carotenals as substrates with good catalytic efficiency. ALDH1A1 and ALDH1A2 exhibit higher substrate specificity for 12'-apo-β-carotenal and 14'-apo-β-carotenal, respectively, while ALDH1A3 shows a similar efficiency for the two compounds. Overall, the substrate specificity of ALDH1A enzymes with apo-β-carotenals is comparable to that with all-*trans*-retinaldehyde, indicating that these enzymes might play a physiological role in the metabolism of apocarotenoids.

The study of the substrate-binding pocket aims to tackle the question whether specific residue differences could be affecting enzymatic activity. Previously reported kinetic parameters were obtained using a plethora of experimental conditions, including buffer type, pH, use of Mg^{2+} ions, chelating and reducing compounds, and solubilizing agents [56,167,182,208,210,211]. Here, for the first time, the activity of the three human ALDH1A enzymes was characterized side-by-side under similar experimental conditions, making possible the most thorough comparative study reported so far. Previously, the enzymatic activity of these proteins had been described, using typical aldehyde substrates and retinoids, mostly in rodent species. Thus, the retinoid activity of rat ALDH1A1 was characterized [212] and the kinetic constants with retinoids were determined for rat [30,208] and mouse [213] ALDH1A2. The enzymatic activity of mouse ALDH1A3 was characterized with several aldehydes [62] and its K_m values were reported for retinoids [214]. Furthermore, the human ALDH1A1, ALDH1A2 and ALDH1A3 had been partially characterized as well. Thus, the enzymatic activity

of human ALDH1A1 was studied with typical aldehyde substrates [177] and their Michaelis-Menten constants were reported for retinoids [208,211]. Recently, some kinetic constants of human ALDH1A2 and ALDH1A3 for retinoids were also determined, although using a much less reliable methodology for retinoid solubilization and analysis [56,69]. Similarly to the human ortholog protein, mouse and rat ALDH1A2 catalyze the oxidation of both all-*trans*- and 9-*cis*-retinaldehyde and, like the human enzyme, mouse ALDH1A3 is highly specific for all-*trans*-retinaldehyde [214]. Despite these reports, not a single report with a common methodology and under the same experimental conditions described simultaneously the enzymatic characterization of the human ALDH1A enzymes.

Therefore, the robust methodology employed and the data obtained in this Thesis has been useful to compare the kinetic properties between the three ALDH1A enzymes and with previous reports. For human ALDH1A1, our results with typical aldehydes (*trans*-2-hexenal and citral) are quite similar to those published by Solobodowska *et al.*, [177]. For human ALDH1A2, our results show k_{cat} values much higher and K_m values 4-fold lower than those previously observed by Shabtai *et al.* [56], whose experimental methods included the use of a reaction buffer that combined the presence of $MgCl_2$ and the chelating agent EDTA, making uncertain the actual concentration of Mg^{2+} ions. For human ALDH1A3, the obtained results are comparable with those of rat ALDH1A3 [62], where the kinetic parameters were evaluated using hexanal and benzaldehyde. In all cases, the K_m values were in micromolar range.

Up to now, the effect of divalent cations only had been assessed with ALDH1A1 (formerly known as ALDH1 or cytosolic ALDH) and ALDH2 (namely mitochondrial ALDH). Interestingly, the effect of Mg^{2+} ions is quite distinct for each ALDH1A enzyme. While ALDH1A1 activity is severely inhibited, ALDH1A2 activity is moderately reduced and, in contrast, ALDH1A3 is significantly increased. The relevant Mg^{2+} concentrations fall within the range of the physiological intracellular concentrations of free Mg^{2+} ions (0.5-1 mM) [215]. The inhibitory effect of Mg^{2+} ions had been described previously for ALDH1A1 and seems to be related to coenzyme dissociation [82]. Magnesium stabilizes the cofactor binding by interacting with the pyrophosphate group and making NADH dissociation slower. The decrease in activity of ALDH1A1 with Mg^{2+} ions is therefore evidence that coenzyme dissociation is the rate-limiting step, as previously proposed [82]. The rate-limiting steps for ALDH1A2 and ALDH1A3 are currently unknown. The moderately inhibitory effect of Mg^{2+} ions on ALDH1A2 may support a limiting step similar to that of ALDH1A1, i.e. coenzyme dissociation. The activating effect on ALDH1A3 is consistent with what was previously observed for mouse ALDH1A3 [62]. This

activation is reminiscent of the reported effect on the mitochondrial ALDH2 form, where Mg^{2+} ions enhanced deacylation, which is the rate-limiting step for this enzyme [88]. The similar activation by Mg^{2+} ions observed here for ALDH1A3 gives support for deacylation being the rate-limiting step for ALDH1A3. Thus, our results are consistent with cofactor dissociation being the rate-limiting step for ALDH1A1 and ALDH1A2, and deacylation for ALDH1A3.

Regarding the effect of ALDHs on embryonic development, it is remarkable that the knockout of ALDH1A1 did not severely affect the morphology of the retina [149], indicating that other ALDH1A enzymes might redundantly share the function of ALDH1A1. In addition, ALDH1A1 is able to induce the differentiation of hematopoietic and neural stem cells and has been widely used as a marker of cancer stem cells. ALDH1A1 is also involved in chemoresistance, specifically against cyclophosphamide treatment [46]. In adults, ALDH1A2 is involved in spermatogenesis, which is the basis of inhibitor development intended as anticonceptive drugs [54,166]. ALDH1A3 is expressed in the ventral retina across various species [155,156] and is essential for the development of the central nervous system and the morphogenesis of anterior head structures. Therefore, the three ALDH1A enzymes are essential for embryonic and adult life because of their role in retinoic acid synthesis. The low micromolar range K_m values with all-*trans*-retinaldehyde, the most physiological relevant isomer, and high catalytic efficiency, common for human and murine enzymes, are suitable for this role. The higher efficiency of ALDH1A2 and ALDH1A3 correlates with the stronger effects of the respective knock-outs.

From the single-point analysis of the inhibitory effect (Table 9) and the IC_{50} values (Tables 10-11) of various synthetic compounds against ALDH1A enzymes, some points could be featured. In general, the most potent inhibitors from the 1-oxopyrimido[4,5-*c*]quinoline-2-acetic acid series are compounds **64**, **65** and **66**. Compounds **64** and **65** are excellent inhibitors against both ALDH1A2 and ALDH1A3, but not for ALDH1A1. In contrast, only **66** is a good inhibitor *versus* the three ALDH1A enzymes. According with these results, it can be deduced that the presence of a bulky benzene group in R_3 is important for binding. Regarding ALDH1A1, the presence of a chlorine atom in R_2 (as in compounds **64** and **65**) may be detrimental for their inhibitory capacity. Compound **70** shows an inhibitory effect on ALDH1A1, ALDH1A2 and ALDH1A3, being ALDH1A2 the most sensitive enzyme. This enhanced selectivity for ALDH1A2 might be due to the presence of a fluorine atom in R_1. This is supported by the fact that, in the absence of a fluorine atom, the inhibitory capacity is lost, like in compound **68**. Moreover, when the R_3 ethyl group is substituted by a methyl group, the inhibition also disappears, as in

compound **39**. As for the thiazolidinedione-acetic acid derivatives, **5a** is a much better inhibitor against ALDH1A1 and ALDH1A3 against ALDH1A2. Interestingly, compound **5c** is the most selective inhibitor, showing only inhibition for ALDH1A3, while **5a** inhibits both ALDH1A1 and ALDH1A3. The presence of a chlorine atom in R_1 (as in **5c**) seems to be detrimental for the inhibitory activity against ALDH1A1 and ALDH1A2. Finally, in regard to the JF series, only JF0064 shows inhibition against the three ALDH1A enzymes, with IC_{50} values in the micromolar range for the three enzymes. Although JF0064 contains two benzene rings and fluorine atoms, it is a smaller molecule than other members of JF series, such as JF0048 or JF0049, which do not exhibit fluorine atoms. These other compounds do not have inhibitory capacity on ALDH1A enzymes. Thus, it is likely that the fluorine atoms may play an important role in the inhibition mechanism.

In summary, it can be concluded that the three enzymes, ALDH1A1, ALDH1A2 and ALDH1A3, share similar substrate-binding pocket topologies (with some scattered point residue differences and decreasing enclosed volumes) and common substrate specificity (with some notable singularities). Some kinetic features, such as the citral specificity for ALDH1A1, have been linked in part to residue 114 and to residues 458 to 461 of the loop located between the β18 sheet and α13 helix. The specificity of ALDH1A3 for the all-*trans-* *versus* 9-*cis* isomer of retinaldehyde might be related to the smaller volume of its substrate-binding pocket. Among ALDH1As, ALDH1A3 exhibits the highest k_{cat}/K_m value for all-*trans*-retinaldehyde but the lowest values for all other aldehyde substrates, indicating a strong substrate specificity of ALDH1A3 for all-*trans*-retinaldehyde. In addition to the known role of ALDH1A enzymes in the conversion of retinadehyde to retinoic acid, they all may participate in aldehyde detoxification and, specifically, in the elimination of lipid peroxidation products (*trans*-2-hexenal, 4-hydroxy-2-hexenal and 4-hydroxy-2-nonenal). In addition, the substrate specificity of ALDH1A enzymes with apo-β-carotenals indicates a possible physiological role of these enzymes in the metabolism of these compounds and thus in the pathway derived from the eccentric cleavage of β-carotene by BCO2. Both types of activity are relevant to physiological roles of ALDH1A enzymes, such embryogenesis or stemness maintenance in normal and cancer stem cells. It is important to remark that a novel selective ALDH1A3 inhibitor has been identified, while novel, potent and selective inhibitors for ALDH1A1 and ALDH1A2 have been explored. These preliminary results appear to be very promising to develop new pharmacophores by using structure-based drug design.

CONCLUSIONS

1. ALDH1A1, ALDH1A2 and ALDH1A3 share similar topologies in their substrate-binding pockets, with some notable singularities. The three enzymes exhibited decreasing volumes in their substrate-binding pockets, 534, 387 and 357 $Å^3$, respectively.

2. In ALDH1A2, a new subpocket, lined by loop residues Leu477, Asn478 and Ala479 (ALDH1A2 numbering), is observed near the solvent exposed area, which could be generated by inhibitor binding.

3. Human ALDH1A enzymes (ALDH1A1, ALDH1A2 and ALDH1A3) were recombinantly expressed in *E. coli* in soluble and active form, and affinity purified in a single chromatographic step from the soluble fraction of bacterial lysates.

4. The effect of Mg^{2+} ions on the dehydrogenase activity of ALDH1A enzymes, using hexanal as a substrate, shows a quite distinct pattern for each of the three enzymes. While ALDH1A1 and ALDH1A2 activity decreased with increasing concentrations of Mg^{2+}, the ALDH1A3 activity was increased. This may reflect a different rate-limiting step: cofactor dissociation in ALDH1A1 and ALDH1A2, and deacylation in ALDH1A3.

5. The kinetic properties of the three human enzymes were characterized side-by-side and using a similar methodology. All analyzed ALDHs exhibited NAD^+-dependent dehydrogenase activity with typical aldehyde substrates.

6. The k_{cat}/K_m values of ALDH1A1 and ALDH1A2 for typical aldehyde substrates indicate a potentially major role in the transformation of these substrates with slightly different substrate specificity, while ALDH1A3 exhibits the lowest values in terms of k_{cat}/K_m for all substrates, suggesting a moderate role in the physiological oxidation of these aldehydes.

7. The retinaldehyde dehydrogenase activity of the ALDH1A enzymes was determined by using a novel solubilization method in the absence of detergents or organic solvents and a sensitive HPLC-based method, which allowed the detection of the substrate and the reaction product in a single chromatographic step.

8. The three ALDH1A forms exhibited lower K_m values for 9-*cis*-retinaldehyde than for all-*trans*-retinaldehyde. Regarding the k_{cat} values, they were higher for the all-*trans* isomer (ALDH1A2 and ALDH1A3) or similar for the two isomers (ALDH1A1). ALDH1A3 was highly specific for all-*trans*-retinaldehyde.

9. Human retinaldehyde dehydrogenases were capable of using apo-β-carotenals as substrates with good catalytic efficiency. ALDH1A1 and ALDH1A2 exhibited higher substrate specificity for 12'-apo-β-carotenal and 14'-apo-β-carotenal, respectively, while ALDH1A3 showed a similar specificity for the two compounds.

10. Site-directed mutagenesis was used to explore structure-function relationships in ALDH1As. For the ALDH1A1 L114P mutant, there was a significant increase in the K_m values for most substrates, including all-*trans*-retinaldehyde, being more similar to those of wild-type ALDH1A2. The ALDH1A2 mutant displayed a 50-fold decrease in the K_m value for citral, in part mimicking ALDH1A1.

11. Several compounds were tested as inhibitors with potential pharmacological interest. The most potent inhibitors from the 1-oxopyrimido[4,5-*c*]quinoline-2-acetic acid series are compounds **64, 65** and **66.** Compounds **64** and **65** are excellent inhibitors for both ALDH1A2 and ALDH1A3, but not for ALDH1A1. In contrast, only **66** is a good inhibitor for the three ALDH1A enzymes.

12. For the thiazolidinedione-acetic acid derivatives, compound **5a** is a much better inhibitor against ALDH1A1 and ALDH1A3 than against ALDH1A2. Compound **5c** is the most selective inhibitor, showing only inhibition against ALDH1A3.

13. Compound JF0064 shows inhibition against the three ALDH1A enzymes, with IC_{50} values in the micromolar range.

14. The presence of a bulky benzene group in R_3 of 1-oxopyrimido[4,5-*c*]quinoline-2-acetic acid series is important for binding, while the presence of fluorine atoms may play a role in the inhibition mechanism.

BIBLIOGRAPHY

[1]   P.J.O. Brien, A.G. Siraki, N. Shangari, P.J.O. Brien, A.G. Siraki, N. Shangari, et al., Critical Reviews in Toxicology Aldehyde Sources, Metabolism, Molecular Toxicity Mechanisms, and Possible Effects on Human Health, Crit. Rev. Toxicol. 8444 (2017) 1547–6898.

[2]   S.A. Marchitti, C. Brocker, D. Stagos, V. Vasiliou, Non-P450 aldehyde oxidizing enzymes: the aldehyde dehydrogenase superfamily, Expert Opin Drug Metab Toxicol. 4 (2008) 697–720.

[3]   V. Vasiliou, A. Pappa, T. Estey, Role of Human Aldehyde Dehydrogenases in Endobiotic and Xenobiotic Metabolism Role of Human Aldehyde Dehydrogenases in, Drug Metab. Rev. 2532 (2004) 1097–9883.

[4]   V. Vasiliou, A. Bairoch, K.F. Tipton, D.W. Nebert, Eukaryotic aldehyde dehydrogenase (ALDH) genes: human polymorphisms, and recommended nomenclature based on divergent evolution and chromosomal mapping., Pharmacogenetics. 9 (1999) 421–34.

[5]   N.A. Sophos, V. Vasiliou, Aldehyde dehydrogenase gene superfamily : the 2002 update, Chem. Biol. Interact. 144 (2003) 5–22.

[6]   B. Jackson, C. Brocker, D.C. Thompson, W. Black, D.W. Nebert, V. Vasiliou, Update on the aldehyde dehydrogenase gene (ALDH) superfamily, Hum Genomics. 5 (2012) 1–20.

[7]   R.S. Holmes, Comparative and evolutionary studies of vertebrate ALDH1A-like genes and proteins, Chem. Biol. Interact. 234 (2015) 4–11.

[8]   W.J. Black, D. Stagos, S.A. Marchitti, D.W. Nebert, F. Keith, A. Bairoch, et al., Human Aldehyde Dehydrogenase Genes: Alternatively-Spliced Transcriptional Variants and Their Suggested Nomenclature, Pharmacogenet Genomics. 19 (2012) 893–902.

[9]   S. a Moore, H.M. Baker, T.J. Blythe, K.E. Kitson, T.M. Kitson, E.N. Baker, Sheep liver cytosolic aldehyde dehydrogenase: the structure reveals the basis for the retinal specificity of class 1 aldehyde dehydrogenases., Structure. 6 (1998) 1541–51.

[10]   A.L. Lamb, X. Wang, J.L. Napoli, M.E. Newcomer, Purification, crystallization and preliminary X-ray diffraction studies of retinal dehydrogenase type II, Acta Crystallogr. Sect. D Biol. Crystallogr. 54 (1998) 639–642.

[11]   R.K. Wierenga, W.G.J. Hol, Predicted nucleotide-binding properties of p21 protein and its cancer-associated variant, Nature. 302 (1983) 842–844.

[12]   J. Hempel, H. Nicholas, R. Lindahl, Aldehyde dehydrogenases: widespread structural and functional diversity within a shared framework., Protein Sci. 2 (1993) 1890–900.

[13]   Z.J. Liu, Y.J. Sun, J. Rose, Y.J. Chung, C.D. Hsiao, W.R. Chang, et al., The first structure of an aldehyde dehydrogenase reveals novel interactions between NAD and the Rossmann fold., Nat. Struct. Biol. 4 (1997) 317–326.

[14] C.G. Steinmetz, P. Xie, H. Weiner, T.D. Hurley, Structure of mitochondrial aldehyde dehydrogenase: the genetic component of ethanol aversion., Structure. 5 (1997) 701–711.

[15] X. Robert, P. Gouet, Deciphering key features in protein structures with the new ENDscript server, Nucleic Acids Res. 42 (2014) 320–324.

[16] L. Ni, S. Sheikh, H. Weiner, Involvement of Glutamate 399 and Lysin 192 in the Mechanism of Human Liver Mitochondrial Aldehyde Dehydrogenase, J. Biol. Chem. 272 (1997) 18823–18826.

[17] S. Sheikh, L. Ni, T.D. Hurley, H. Weiner, The potential roles of the conserved amino acids in human liver mitochondrial aldehyde dehydrogenase, J Biol Chem. 272 (1997) 18817–18822.

[18] A.L. Lamb, M.E. Newcomer, The structure of retinal dehydrogenase type II at 2.7 A resolution: implications for retinal specificity., Biochemistry. 38 (1999) 6003–11.

[19] L. Di Costanzo, G. a. Gomez, D.W. Christianson, Crystal Structure of Lactaldehyde Dehydrogenase from Escherichia coli and Inferences Regarding Substrate and Cofactor Specificity, J. Mol. Biol. 366 (2007) 481–493.

[20] K. Johansson, M. El-Ahmad, S. Ramaswamy, L. Hjelmqvist, H. Jörnvall, H. Eklund, Structure of betaine aldehyde dehydrogenase at 2.1 A resolution., Protein Sci. 7 (1998) 2106–17.

[21] B. Ahvazi, C. R, M. Delarge, M. Vedadi, L. Zhang, E. Meighen, et al., Crystal structure of the NADP+-dependent aldehyde dehydrogenase from Vibrio harveyi: structural implications for cofactor specificity and affinity., Biochem. J. 349 (2000) 853–61.

[22] D. Cobessi, F. Tête-Favier, S. Marchal, S. Azza, G. Branlant, a Aubry, Apo and holo crystal structures of an NADP-dependent aldehyde dehydrogenase from Streptococcus mutans., J. Mol. Biol. 290 (1999) 161–173.

[23] J. Perozich, I. Kuo, R. Lindahl, J. Hempel, Coenzyme specificity in aldehyde dehydrogenase, Chem. Biol. Interact. 130-132 (2001) 115–124.

[24] S.O. Kotchoni, J.C. Jimenez-Lopez, D. Gao, V. Edwards, E.W. Gachomo, V.M. Margam, et al., Modeling-Dependent Protein Characterization of the Rice Aldehyde Dehydrogenase (ALDH) Superfamily Reveals Distinct Functional and Structural Features, PLoS One. 5 (2010) 1–14.

[25] T.J.P. Sobreira, F. Marlétaz, M. Simões-Costa, D. Schechtman, A.C. Pereira, F. Brunet, et al., Structural shifts of aldehyde dehydrogenase enzymes were instrumental for the early evolution of retinoid-dependent axial patterning in metazoans., Proc. Natl. Acad. Sci. U. S. A. 108 (2011) 226–31.

[26] C.A. Morgan, T.D. Hurley, Development of a high-throughput in vitro assay to identify selective inhibitors for human ALDH1A1, Chem Biol Interact. 234 (2015) 29–37.

[27]   J. Farrés, T.T. Wang, S.J. Cunningham, H. Weiner, Investigation of the active site cysteine residue of rat liver mitochondrial aldehyde dehydrogenase by site-directed mutagenesis., Biochemistry. 34 (1995) 2592–2598.

[28]   X. Wang, H. Weiner, Involvement of Glutamate 268 in the Active Site of Human Liver Mitochondrial (Class 2) Aldehyde Dehydrogenase As Probed by Site-Directed Mutagenesist, Biochemistry. (1995) 237–243.

[29]   D. Zhao, P. McCaffery, K.J. Ivins, R.L. Neve, P. Hogan, W.W. Chin, et al., Molecular identification of a major retinoic-acid-synthesizing enzyme, a retinaldehyde-specific dehydrogenase., Eur. J. Biochem. 240 (1996) 15–22.

[30]   X. Wang, P. Penzes, J.L. Napoli, J.B. Chem, X. Wang, P. Penzes, et al., Cloning of a cDNA Encoding an Aldehyde Dehydrogenase and Its Expression in Escherichia coli, J. Biol. Chem. 271 (1996) 16288–16293.

[31]   A. a Klyosov, L.G. Rashkovetsky, M.K. Tahir, W. Keung, Possible role of liver cytosolic and mitochondrial aldehyde dehydrogenases in acetaldehyde metabolism, Biochemistryiochemis. 35 (1996) 4445–4456.

[32]   S.L. Bradbury, W.B. Jakoby, Ordered binding of substrates to yeast aldehyde dehydrogenase., J. Biol. Chem. 246 (1971) 1834–1840.

[33]   E.M. Valenzuela-soto, Betaine-aldehyde Dehydrogenase, J. Biol. Chem. 268 (1993) 23818–23823.

[34]   J. Hempel, J. Perozich, T. Chapman, J. Rose, J.S. Boesch, Z.-J. Liu, et al., Aldehyde Dehydrogenase Catalytic Mechanism, Enzymol. Mol. Biol. Carbonyl Metab. 7. (1999) 53–59.

[35]   V. Koppaka, D.C. Thompson, Y. Chen, M. Ellermann, K.C. Nicolaou, R.O. Juvonen, et al., Aldehyde dehydrogenase inhibitors: a comprehensive review of the pharmacology, mechanism of action, substrate specificity, and clinical application., Pharmacol. Rev. 64 (2012) 520–39.

[36]   S. Sidhu, A. Huntley, N. Scotia, Liver aldehyde dehydrogenase, J. Biol. Chem. 250 (1975) 7894–7899.

[37]   G. King, R. Holmes, Human Corneal and Lens Aldehyde Dehydrogenases, Enzymol. Mol. Biol. Carbonyl Metab. 6. (1997) 19–27.

[38]   Y. Zhai, Z. Sperkova, J.L. Napoli, Cellular expression of retinal dehydrogenase types 1 and 2: effects of vitamin A status on testis mRNA, J Cell Physiol. 186 (2001) 220–232.

[39]   D. Galter, S. Buervenich, A. Carmine, M. Anvret, L. Olson, ALDH1 mRNA: presence in human dopamine neurons and decreases in substantia nigra in Parkinson's disease and in the ventral tegmental area in schizophrenia, Neurobiol Dis. 14 (2003) 637–647.

[40]   F.M. Jacobs, S.M. Smits, C.W. Noorlander, L. von Oerthel, A.J. van der Linden, J.P. Burbach, et al., Retinoic acid counteracts developmental defects in the substantia nigra caused by Pitx3 deficiency, Development. 134 (2007) 2673–2684.

[41]   Y. Ueshima, Y. Matsuda, M. Tsutsumi, a Takada, Role of the aldehyde dehydrogenase-1 isozyme in the metabolism of acetaldehyde., Alcohol Alcohol Suppl. 1B (1993) 15–19.

[42]   M.J. Stewart, K. Malek, Q. Xiao, K.M. Dipple, D.W. Crabb, The novel aldehyde dehydrogenase gene, ALDH5, encodes an active aldehyde dehydrogenase enzyme, Biochem. Biophys. Res. Commun. 211 (1995) 144–151.

[43]   X. He, L. González-Segura, A. Tsang, J. Thompson, T.C. Tsang, D.T. Harris, Differential gene expression profiling of CD34+ CD133+ umbilical cord blood hematopoietic stem progenitor cells, Stem Cells Dev. 837 (2006) 831–837.

[44]   J.P. Chute, G.G. Muramoto, J. Whitesides, M. Colvin, R. Safi, N.J. Chao, et al., Inhibition of aldehyde dehydrogenase and retinoid signaling induces the expansion of human hematopoietic stem cells, Proc Natl Acad Sci U S A. 103 (2006) 11707–11712.

[45]   M. Gasparetto, C. a. Smith, ALDHs in normal and malignant hematopoietic cells: Potential new avenues for treatment of AML and other blood cancers, Chem. Biol. Interact. 276 (2017) 46–51.

[46]   N.E. Sládek, Aldehyde dehydrogenase-mediated cellular relative insensitivity to the oxazaphosphorines., Curr. Pharm. Des. 5 (1999) 607–25.

[47]   D. Velázquez-Fernández, C. Laurell, J. Geli, A. Höög, J. Odeberg, M. Kjellman, et al., Expression profiling of adrenocortical neoplasms suggests a molecular signature of malignancy, Surgery. 138 (2005) 1087–1094.

[48]   J. Zhou, H. Weiner, Binding of thyroxine analogs to human liver aldehyde dehydrogenases, Eur J Biochem. 245 (1997) 123–128.

[49]   L. Deng, G.L. Shipley, D.S. Loose-Mitchell, G.M. Stancel, R. Broaddus, J.H. Pickar, et al., Coordinate regulation of the production and signaling of retinoic acid by estrogen in the human endometrium, J. Clin. Endocrinol. Metab. 88 (2003) 2157–2163.

[50]   The GTExArd Consortium, D. Welter, J. MacArthur, J. Morales, T. Burdett, P. Hall, et al., The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans., Science. 348 (2015) 648–60.

[51]   M. Nishimura, S. Naito, Tissue-specific mRNA expression profiles of human phase I metabolizing enzymes except for cytochrome P450 and phase II metabolizing enzymes., Drug Metab. Pharmacokinet. 21 (2006) 357–374.

[52]   R.J. Haselbeck, I. Hoffmann, G. Duester, Distinct functions for ALDH1 and RALDH2 in the control of ligand production for embryonic retinoid signaling pathways, Dev Genet. 25 (1999) 353–364.

[53]   A.S. Goldstein, J. Paik, J. Moreb, M. Haenisch, I. Le Trong, R.E. Stenkamp, et al., Synthesis and in vitro testing of bisdichloroacetyldiamine analogs for use as a reversible male contraceptive, To Be Published.

[54]   Y. Chen, J.-Y. Zhu, K.H. Hong, D.C. Mikles, G.I. Georg, A.S. Goldstein, et al., Structural basis of ALDH1A2 inhibition by irreversible and reversible small molecule inhibitors, ACS Chem. Biol. 13 (2018) 582–590.

[55] K. Niederreither, V. Subbarayan, P. Dolle, P. Chambon, Embryonic retinoic acid synthesis is essential for early mouse post-implantation development, Nat Genet. 21 (1999) 444–448.

[56] Y. Shabtai, H. Jubran, T. Nassar, J. Hirschberg, A. Fainsod, Kinetic characterization and regulation of the human retinaldehyde dehydrogenase 2 enzyme during production of retinoic acid, Biochem J. 473 (2016) 1423–1431.

[57] H. Kim, J. Lapointe, G. Kaygusuz, D.E. Ong, C. Li, M. van de Rijn, et al., The retinoic acid synthesis gene ALDH1a2 is a candidate tumor suppressor in prostate cancer, Cancer Res. 65 (2005) 8118–8124.

[58] L.M. De Luca, Retinoids and their receptors in differentiation, embryogenesis and neoplasia, FASEB J. 14 (1991) 2924–33.

[59] J.K. Amory, S. Arnold, M.C. Lardone, A. Piottante, M. Ebensperger, N. Isoherranen, et al., Levels of the retinoic acid synthesizing enzyme aldehyde dehydrogenase-1A2 are lower in testicular tissue from men with infertility, Fertil. Steril. 101 (2014) 960–966.

[60] X. Zhang, Q. Zhang, D. Liu, T. Su, Y. Weng, G. Ling, et al., Short Communication Expression of cytochrome P450 and other biotransformation genes in fetal and adult human nasal mucosa, Drug Metab. Dispos. 33 (2005) 1423–1428.

[61] L.C. Hsu, W.-C. Chang, L. Hiraoka, C.-L. Hsieh, Molecular cloning, genomic organization, and chromosomal localization of an additional human aldehyde dehydrogenase gene, ALDH6, Genomics. 24 (1994) 333–341.

[62] C.E. Graham, K. Brocklehurst, R.W. Pickersgill, M.J. Warren, Characterization of retinaldehyde dehydrogenase 3., Biochem. J. 394 (2006) 67–75.

[63] S. Okamura, C.C. Ng, K. Koyama, Y. Takei, H. Arakawa, M. Monden, et al., Identification of seven genes regulated by wild-type p53 in a colon cancer cell line carrying a well-controlled wild-type p53 expression system., Oncol. Res. 11 (1999) 281–5.

[64] J. Jia, H. Parikh, W. Xiao, J.W. Hoskins, H. Pflicke, X. Liu, et al., An integrated transcriptome and epigenome analysis identifies a novel candidate gene for pancreatic cancer, BMC Med. Genomics. 6 (2013) 1.

[65] Y.T. Saw, J. Yang, S.K. Ng, S. Liu, S. Singh, M. Singh, et al., Characterization of aldehyde dehydrogenase isozymes in ovarian cancer tissues and sphere cultures, BMC Cancer. 12 (2012) 1–12.

[66] S. Yamashita, Y. Tsujino, K. Moriguchi, M. Tatematsu, T. Ushijima, Chemical genomic screening for methylation-silenced genes in gastric cancer cell lines using 5-aza-2'-deoxycytidine treatment and oligonucleotide microarray, Cancer Sci. 97 (2006) 64–71.

[67] J. Han, L. Yang, R.K. Puri, Analysis of target genes induced by IL-13 cytotoxin in human glioblastoma cells, J Neurooncol. 72 (2005) 35–46.

[68] J.J. Duan, J. Cai, Y.F. Guo, X.W. Bian, S.C. Yu, ALDH1A3, a metabolic target for cancer diagnosis and therapy, Int. J. Cancer. 139 (2016) 965–975.

[69]   A. Mory, F.X. Ruiz, E. Dagan, E. a. Yakovtseva, A. Kurolap, X. Parés, et al., A missense mutation in ALDH1A3 causes isolated microphthalmia/anophthalmia in nine individuals from an inbred Muslim kindred, Eur. J. Hum. Genet. 22 (2014) 419–422.

[70]   J. Farres, X. Wang, K. Takahashi, S.J. Cunningham, T.T. Wang, H. Weiner, Effects of Changing Glutamate 487 to Lysine in Rat and Human Liver Mitochondrial Aldehyde Dehydrogenase, J. Biol. Chem. 269 (1994) 13854–60.

[71]   Y.C. Chen, G.S. Peng, T.P. Tsao, M.F. Wang, R.B. Lu, S.J. Yin, Pharmacokinetic and pharmacodynamic basis for overcoming acetaldehyde-induced adverse reaction in Asian alcoholics, heterozygous for the variant ALDH2*2 gene allele, Pharmacogenet. Genomics. 19 (2009) 588–599.

[72]   S.E. Luczak, B. Elvine-Kreis, S.H. Shea, L.G. Carr, T.L. Wall, Genetic risk for alcoholism relates to level of response to alcohol in Asian-American men and women., J. Stud. Alcohol. 63 (2002) 74–82.

[73]   A. Yokoyama, T. Muramatsu, T. Omori, T. Yokoyama, S. Matsushita, S. Higuchi, et al., Alcohol and aldehyde dehydrogenase gene polymorphisms and oropharyngolaryngeal, esophageal and stomach cancers in Japanese alcoholics., Carcinogenesis. 22 (2001) 433–9.

[74]   M. Muto, Association of aldehyde dehydrogenase 2 gene polymorphism with multiple oesophageal dysplasia in head and neck cancer patients, Gut. 47 (2000) 256–261.

[75]   Z. Chen, J. Zhang, J.S. Stamler, Identification of the enzymatic mechanism of nitroglycerin bioactivation, Proc. Natl. Acad. Sci. 99 (2002) 8306–8311.

[76]   I.S. Mackenzie, K.M. Maki-Petaja, C.M. McEniery, Y.P. Bao, S.M. Wallace, J. Cheriyan, et al., Aldehyde dehydrogenase 2 plays a role in the bioactivation of nitroglycerin in humans, Arterioscler. Thromb. Vasc. Biol. 25 (2005) 1891–1895.

[77]   Y. Li, D. Zhang, W. Jin, C. Shao, P. Yan, C. Xu, et al., Mitochondrial aldehyde dehydrogenase-2 (ALDH2) Glu504Lys polymorphism contributes to the variation in efficacy of sublingual nitroglycerin, J. Clin. Invest. 116 (2006) 506–511.

[78]   S.A. Jo, E.K. Kim, M.H. Park, C. Han, H.Y. Park, Y. Jang, et al., A Glu487Lys polymorphism in the gene for mitochondrial aldehyde dehydrogenase 2 is associated with myocardial infarction in elderly Korean men, Clin. Chim. Acta. 382 (2007) 43–47.

[79]   P. Hui, T. Nakayama, A. Morita, N. Sato, M. Hishiki, K. Saito, et al., Common single nucleotide polymorphisms in Japanese patients with essential hypertension: aldehyde dehydrogenase 2 gene as a risk factor independent of alcohol consumption., Hypertens. Res. 30 (2007) 585–592.

[80]   R.C. Vallari, R. Pietruszko, Interaction of Mg2+ with Human Liver Aldehyde Dehydrogenase. Mechanism and site of interaction, J. Biol. Chem. 259 (1984) 4927–4933.

[81]   K. Takahashi, H. Weiner, D.L. Filmer, Effects of pH on horse liver aldehyde dehydrogenase: alterations in metal ion activation, number of functioning active sites, and hydrolysis of the acyl intermediate, Biochemistry. 20 (1981) 6225–6230.

[82]   H. Weiner, H. James, H. Hu, G. Sanny, Rate-limiting steps for the esterase and dehydrogenase reaction catalyzed by horse liver aldehyde dehydrogenase, J. Biol. Chem. 251 (1976) 3853–3856.

[83]   F.M. Dickinson, G.J. Hart, Effects of Mg2+, Ca2+ and Mn2+ on sheep liver cytoplasmic aldehyde dehydrogenase, Biochem J. 205 (1982) 443–448.

[84]   A.F. Bennett, P.D. Buckley, L.F. Blackwell, Inhibition of the Dehydrogenase Activity of Sheep Liver Cytoplasmic Aldehyde Dehydrogenase by Magnesium Ions, Biochemistry. 22 (1983) 776–784.

[85]   A.K. MacGibbon, L.F. Blackwell, P.D. Buckley, Pre-steady-state kinetic studies on cytoplasmic sheep liver aldehyde dehydrogenase., Biochem. J. 167 (1977) 469–477.

[86]   K. Takahashi, H. Weiner, Magnesium stimulation of catalytic activity of horse liver aldehyde dehydrogenase, J. Biol. Chem. 2 (1980) 8206–8209.

[87]   R.C. Vallari, R. Pietruszko, Interaction of Mg2+ with human liver aldehyde dehydrogenase. II. Species difference in the mitochondrial isozyme, J. Biol. Chem. 259 (1984) 4922–4926.

[88]   K.K. Ho, A. Allali-hassani, T.D. Hurley, H. Weiner, Differential effects of Mg2+ ions on the individual kinetic steps of human cytosolic and mitochondrial aldehyde dehydrogenase, Biochemistry. 44 (2005) 8022–8029.

[89]   IUPAC-IUB Joint Commission on Biochemical Nomenclature (JCBN). Nomenclature of Retinoids. Recommendations 1981, Eur. J. Biochem. 129 (1982) 1–5.

[90]   A.B. Barua, H.C. Furr, Properties of retinoids, Mol. Biotechnol. 10 (1998) 167–182.

[91]   C.A. Redlich, J.N. Grauer, A.M. Van Bennekum, S.L. Clever, R.B. Ponn, W.S. Blaner, Characterization of carotenoid, vitamin A, and alpha-tocopheral levels in human lung tissue and pulmonary macrophages., Am. J. Respir. Crit. Care Med. 154 (1996) 1436–1443.

[92]   M.J. Haskell, The challenge to reach nutritional adequacy for vitamin A: b-carotene bioavailability and conversion—evidence in humans 1–4, Am J Clin Nutr. 96 (2012) 1193–203.

[93]   A. Eroglu, E.H. Harrison, Carotenoid metabolism in mammals, including man: formation, occurrence, and function of apocarotenoids, J. Lipid Res. 54 (2013) 1719–1730.

[94]   G. Duester, Involvement of Alcohol Dehydrogenase, Short-Chain Dehydrogenase / Reductase, Aldehyde Dehydrogenase, and Cytochrome P450 in the Control of Retinoid Signaling by Activation of Retinoic Acid Synthesis, Biochemistry. 35 (1996) 12221–12227.

[95]   B.P. Kakkad, D.E. Ong, Reduction of retinaldehyde bound to cellular retinol-binding protein (type II) by microsomes from rat small intestine, J. Biol. Chem. 263 (1988) 12916–12919.

[96] Y. Li, N. Wongsiriroj, W.S. Blaner, The multifaceted nature of retinoid transport and metabolism, Hepatobiliary Surg. Nutr. 3 (2014) 126–139.

[97] G.F. Combs, The vitamins: fundamental aspects in nutrition and health, San Diego, CA: Academic Press, 1992.

[98] R. Kawaguchi, J. Yu, P. Wiita, M. Ter-Stepanian, H. Sun, Mapping the membrane topology and extracellular ligand binding domains of the retinol binding protein receptor, Biochemistry. 47 (2008) 5387–5395.

[99] A. Isken, M. Golczak, V. Oberhauser, S. Hunzelmann, D. Wolfgang, Y. Imanishi, et al., RBP4 disrupts vitamin A uptake homeostasis in a STRA6-deficient animal model for Matthew-Wood syndrome, 7 (2008) 258–268.

[100] R. Kawaguchi, M. Zhong, M. Kassai, M. Ter-Stepanian, H. Sun, Vitamin a transport mechanism of the multitransmembrane cell-surface receptor STRA6, Membranes (Basel). 5 (2015) 425–453.

[101] R. Kawaguchi, J. Yu, J. Honda, J. Hu, J. Whitelegge, P. Ping, et al., A Membrane Receptor for Retinol, Science. 815 (2007) 820–826.

[102] J.L. Napoli, Interactions of retinoid binding proteins and enzymes in retinoid metabolism, Biochim. Biophys. Acta - Mol. Cell Biol. Lipids. 1440 (1999) 139–162.

[103] O. V. Belyaeva, M.P. Johnson, N.Y. Kedishvili, Kinetic analysis of human enzyme RDH10 defines the characteristics of a physiologically relevant retinol dehydrogenase, J. Biol. Chem. 283 (2008) 20299–20308.

[104] N.Y. Kedishvili, Enzymology of retinoic acid biosynthesis and degradation, J. Lipid Res. 54 (2013) 1744–1760.

[105] L.L. Sandell, B.W. Sanderson, G. Moiseyev, T. Johnson, A. Mushegian, K. Young, et al., RDH10 is essential for synthesis of embryonic retinoic acid and is required for limb, craniofacial, and organ development, Genes Dev. 21 (2007) 1113–1124.

[106] X. Parés, J. Farrés, N.Y. Kedishvili, G. Duester, Medium-chain and short-chain dehydrogenases/reductases in retinoid metabolism, Cell Mol Life Sci. 65 (2009) 3936–3949.

[107] H.L. Ang, L. Deltour, T.F. Hayamizu, M. Zgombic-Knight, G. Duester, Retinoic Acid Synthesis in Mouse Embryos during Gastrulation and Craniofacial Development Linked to Class IV Alcohol Dehydrogenase Gene Expression, J. Biol. Chem. 271 (1996) 9526–9534.

[108] R.J. Haselbeck, H.L. Ang, G. Duester, Class IV alcohol/retinol dehydrogenase localization in epidermal basal layer: Potential site of retinoic acid synthesis during skin development, Dev. Dyn. 208 (1997) 447–453.

[109] O. Gallego, O. V Belyaeva, S. Porté, F.X. Ruiz, A. V Stetsenko, E. V Shabrova, et al., Comparative functional analysis of human medium-chain dehydrogenases , short-chain dehydrogenases / reductases and aldo-keto reductases with retinoids, Biochem. J. 399 (2006) 101–109.

[110] O. Gallego, F.X. Ruiz, a. Ardevol, M. Dominguez, R. Alvarez, A.R. de Lera, et al., Structural basis for the high all-trans-retinaldehyde reductase activity of the tumor marker AKR1B10, Proc. Natl. Acad. Sci. 104 (2007) 20764–20769.

[111] B. Crosas, D.J. Hyndman, O. Gallego, S. Martras, X. Parés, T.G. Flynn, et al., Human aldose reductase and human small intestine aldose reductase are efficient retinal reductases: consequences for retinoid metabolism., Biochem. J. 373 (2003) 973–979.

[112] T.M. Penning, AKR1B10: a new diagnostic marker of non-small cell lung carcinoma in smokers., Clin. Cance Res. 11 (2005) 1687–1690.

[113] F.X. Ruiz, O. Gallego, A. Ardèvol, A. Moro, M. Domínguez, S. Alvarez, et al., Aldo-keto reductases from the AKR1B subfamily: Retinoid specificity and control of cellular retinoic acid levels, Chem. Biol. Interact. 178 (2009) 171–177.

[114] G. Duester, Families of retinoid dehydrogenases regulating vitamin A function. Production of visual pigment and retinoic acid, Eur. J. Biochem. 267 (2000) 4315–4324.

[115] G.M. Raner, A.D.N. Vaz, M.J. Coon, Metabolism of all-trans, 9-cis, and 13-cis isomers of retinal by purified isozymes of microsomal cytochrome P450 and mechanism-based inhibition of retinoid oxidation by citral, Mol. Pharmacol. 49 (1996) 515–522.

[116] J.F. Arens, D. a van Dorp, Synthesis of some Compounds Possessing Vitamin A Activity, Nature. 157 (1946) 190–191.

[117] D. a Van Dorp, J.F. Arens, Biological activity of vitamin A acid, Nature. 158 (1946) 60.

[118] C. Lampron, C. Rochette-Egly, P. Gorry, P. Dolle, M. Mark, T. Lufkin, et al., Mice deficient in cellular retinoic acid binding protein II (CRABPII) or in both CRABPI and CRABPII are essentially normal, Development. 121 (1995) 539–548.

[119] S. Poulain, F. Evenou, M.C. Carré, S. Corbel, J.M. Vignaud, N. Martinet, Vitamin A/retinoids signalling in the human lung, Lung Cancer. 66 (2009) 1–7.

[120] R. Zhang, Y. Wang, R. Li, G. Chen, Transcriptional factors mediating retinoic acid signals in the control of energy metabolism, Int. J. Mol. Sci. 16 (2015) 14210–14244.

[121] X.H. Tang, L.J. Gudas, Retinoids, Retinoic Acid Receptors, and Cancer, Annu. Rev. Pathol. Mech. Dis. 6 (2011) 345–364.

[122] N.P. Mongan, L.J. Gudas, Diverse actions of retinoid receptors in cancer prevention and treatment, Differentiation. 75 (2007) 853–870.

[123] S. Aebi, R. Kröning, B. Cenni, A. Sharma, D. Fink, G. Los, et al., All-trans retinoic acid enhances cisplatin-induced apoptosis in human ovarian adenocarcinoma and in squamous head and neck cancer cells, Clin. Cancer Res. 3 (1997) 2033–2038.

[124] N. Shaw, M. Elholm, N. Noy, Retinoic acid is a high affinity selective ligand for the peroxisome proliferator-activated receptor β/δ, J. Biol. Chem. 278 (2003) 41589–41592.

[125] N. Noy, Non-classical Transcriptional Activity of Retinoic Acid, Biochem. Retin. Signal. II Physiol. Vitam. A - Uptake, Transp. Metab. Signal. (2016) 179–199.

[126] C.H. Kim, Retinoic Acid, Immunity, and Inflammation, Vitam. Horm. 86 (2011) 83–101.

[127] K.Q. Hu, C. Liu, H. Ernst, N.I. Krinsky, R.M. Russell, X.D. Wang, The biochemical characterization of ferret carotene-9', 10'-monooxygenase catalyzing cleavage of carotenoids in vitro and in vivo, J. Biol. Chem. 281 (2006) 19327–19338.

[128] E.H. Harrison, L. Quadro, Apocarotenoids: Emerging Roles in Mammals, Annu. Rev. Nutr. 38 (2018) annurev–nutr–082117–051841.

[129] G. Britton, H. Pfander, H. Liaaen-Jensen, Carotenoids, biosynthesis and metabolism, Birkhäuser, Basel. 3 (1998).

[130] C. de la Seña, S. Narayanasamy, K.M. Riedl, R.W. Curley, S.J. Schwartz, E.H. Harrison, Substrate specificity of purified recombinant human β-carotene 15,15'-oxygenase (BCO1), J. Biol. Chem. 288 (2013) 37094–37103.

[131] G. Palczewski, J. Amengual, C.L. Hoppel, J. Von Lintig, Evidence for compartmentalization of mammalian carotenoid metabolism, FASEB J. 28 (2014) 4457–4469.

[132] C.C. Ho, F.F. de Moura, S.-H. Kim, A.J. Clifford, Excentral cleavage of beta-carotene in vivo in a healthy man, Am. J. Clin. Nutr. 85 (2007) 770–777.

[133] A. Eroglu, D.P. Hruszkewycz, C. dela Sena, S. Narayanasamy, K.M. Riedl, R.E. Kopec, et al., Naturally occurring eccentric cleavage products of provitamin A beta-carotene function as antagonists of retinoic acid receptors, J Biol Chem. 287 (2012) 15886–15895.

[134] O. Ziouzenkova, G. Orasanu, G. Sukhova, E. Lau, J.P. Berger, G. Tang, et al., Asymmetric cleavage of β-carotene yields a transcriptional repressor of retinoid X receptor and peroxisome proliferator-activated receptor responses, Mol. Endocrinol. 21 (2007) 77–88.

[135] J. Amengual, G.P. Lobo, M. Golczak, H.N.M. Li, T. Klimova, C.L. Hoppel, et al., A mitochondrial enzyme degrades carotenoids and protects against oxidative stress., FASEB J. 25 (2011) 948–59.

[136] C. Liu, X.-D. Wang, R.M. Russell, Biosynthesis of retinoic acid from β-apo-14'-carotenal in ferret in vivo, J. Nutr. Biochem. 8 (1997) 652–657.

[137] A. Yoshida, L.C. Hsu, Y. Yanagawa, Biological Role of Human Cytosolic Aldehyde Dehygrogenase 1: Hormonal Response, Retinal Oxidation and Implication in Testicular Feminization, Enzymol. Mol. Biol. Carbonyl Metab. 4. (1993) 37–44.

[138] A. Yoshida, A. Rzhetsky, L.C. Hsu, C. Chang, Review human aldehyde dehydrogenase gene family, Eur J Biochem. 251 (1998) 549–557.

[139] C. Mendelsohn, E. Ruberte, M. LeMeur, G. Morriss-Kay, P. Chambon, Developmental analysis of the retinoic acid-inducible RAR-beta 2 promoter in transgenic animals., Development. 113 (1991) 723–34.

[140] K. Niederreither, V. Fraulob, J.-M. Garnier, P. Chambon, P. Dollé, Differential expression of retinoic acid-synthesizing (RALDH) enzymes during fetal development and organ differentiation in the mouse., Mech. Dev. 110 (2002) 165–71.

[141] N. Seiler, B. Eichentopf, 4-Aminobutyrate in Mammalian Putrescine Catabolism., Biochem. J. 152 (1975) 201–10.

[142] N.E. Sládek, Human aldehyde dehydrogenases: potential pathological, pharmacological, and toxicological impact., J. Biochem. Mol. Toxicol. 17 (2003) 7–23.

[143] K. Sydow, A. Daiber, M. Oelze, Z. Chen, M. August, M. Wendt, et al., Central role of mitochondrial aldehyde dehydrogenase and reactive oxygen species in nitroglycerin tolerance and cross-tolerance, J. Clin. Invest. 113 (2004) 482–489.

[144] V. Vasiliou, A. Pappa, D.R. Petersen, Role of aldehyde dehydrogenases in endogenous and xenobiotic metabolism, Chem. Biol. Interact. 129 (2000) 1–19.

[145] L. Uma, J. Hariharan, Y. Sharma, D. Balasubramanian, Corneal aldehyde dehydrogenase displays antioxidant properties, Exp Eye Res. 63 (1996) 117–120.

[146] V. Vasiliou, A. Pappa, Polymorphisms of Human Aldehyde Dehydrogenases, Pharmacology. 61 (2000) 192–198.

[147] J. V Jester, T. Moller-Pedersen, J. Huang, C.M. Sax, W.T. Kays, H.D. Cavangh, et al., The cellular basis of corneal transparency: evidence for "corneal crystallins," J. Cell Sci. 112 (1999) 613–622.

[148] G. Duester, F.A. Mic, A. Molotkov, Cytosolic retinoid dehydrogenases govern ubiquitous metabolism of retinol to retinaldehyde followed by tissue-specific metabolism to retinoic acid, Chem Biol Interact. 143-144 (2003) 201–210.

[149] X. Fan, A. Molotkov, S. Manabe, C.M. Donmoyer, L. Deltour, M.H. Foglio, et al., Targeted disruption of Aldh1a1 (Raldh1) provides evidence for a complex mechanism of retinoic acid synthesis in the developing retina, Mol Cell Biol. 23 (2003) 4637–4648.

[150] A. Blentic, E. Gale, M. Maden, Retinoic acid signalling centres in the avian embryo identified by sites of expression of synthesising and catabolising enzymes, Dev Dyn. 227 (2003) 114–127.

[151] K. Niederreither, P. McCaffery, U.C. Dräger, P. Chambon, P. Dollé, Restricted expression and retinoic acid-induced downregulation of the retinaldehyde dehydrogenase type 2 (RALDH-2) gene during mouse development, Mech. Dev. 62 (1997) 67–78.

[152] M. Pavan, V.F. Ruiz, F. a. Silva, T.J. Sobreira, R.M. Cravo, M. Vasconcelos, et al., ALDH1A2 (RALDH2) genetic variation in human congenital heart disease, BMC Med. Genet. 10 (2009) 113.

[153] J. Vermot, K. Niederreither, J.M. Garnier, P. Chambon, P. Dolle, Decreased embryonic retinoic acid synthesis results in a DiGeorge syndrome phenotype in newborn mice, Proc Natl Acad Sci U S A. 100 (2003) 1763–1768.

[154] K. Niederreither, J. Vermot, B. Schuhbaur, P. Chambon, P. Dolle, Retinoic acid synthesis and hindbrain patterning in the mouse embryo, Development. 127 (2000) 75–85.

[155] R.K. Kam, Y. Deng, Y. Chen, H. Zhao, Retinoic acid synthesis and functions in early embryonic development, Cell Biosci. 2 (2012) 11.

[156] R. Suzuki, T. Shintani, H. Sakuta, A. Kato, T. Ohkawara, N. Osumi, et al., Identification of RALDH-3, a novel retinaldehyde dehydrogenase, expressed in the ventral region of the retina, Mech. Dev. 98 (2000) 37–50.

[157] N. Molotkova, A. Molotkov, G. Duester, Role of retinoic acid during forebrain development begins late when Raldh3 generates retinoic acid in the ventral subventricular zone, Dev Biol. 303 (2007) 601–610.

[158] C.S. Boyer, D.R. Petersen, The metabolism of 3,7-dimethyl-2,6-octadienal (citral) in rat hepatic mitochondrial and cytosolic fractions. Interactions with aldehyde and alcohol dehydrogenases., Drug Metab. Dispos. 19 (1991) 81–6.

[159] A. Kikonyogo, D.P. Abriola, M. Dryjanski, R. Pietruszko, Mechanism of inhibition of aldehyde dehydrogenase by citral, a retinoid antagonist, Eur. J. Biochem. 262 (1999) 704–712.

[160] C.S. Boyer, D.R. Petersen, The metabolism of 3,7-dimethyl-2,6-octadienal (citral) in rat hepatic mitochondrial and cytosolic fractions. Interactions with aldehyde and alcohol dehydrogenases., Drug Metab Dispos. 19 (1991) 81–6.

[161] S. Ebrahimi Nigjeh, S. Keong Yeap, N. Nordin, B. Kamalideghan, H. Ky, R. Rosli, Citral induced apoptosis in MDA-MB-231 spheroid cells, BMC Complement Altern Med. 18 (2018).

[162] J.E. Russo, D. Hauquitz, J. Hilton, Inhibition of mouse cytoslic aldehyde dehydrogenase by 4-(diethylamino)benzaldehyde, Biochem. Pharmacol. 37 (1988) 1639–1642.

[163] R.W. Storms, A.P. Trujillo, J.B. Springer, L. Shah, O.M. Colvin, S.M. Ludeman, et al., Isolation of primitive human hematopoietic progenitors on the basis of aldehyde dehydrogenase activity., Proc. Natl. Acad. Sci. U. S. A. 96 (1999) 9118–23.

[164] J.S. Moreb, D. Ucar, S. Han, J.K. Amory, A.S. Goldstein, B. Ostmark, et al., The enzymatic activity of human ALDH1A2 and ALDH2 is detected by Aldefluor, inhibited by DEAB and has significant effects on cell proliferation and drug resistance, Chem Biol Interact. 195 (2013) 52–60.

[165] C.A. Morgan, B. Parajuli, C.D. Buchman, K. Dria, T.D. Hurley, N,N-diethylaminobenzaldehyde (DEAB) as a substrate and mechanism-based inhibitor for human ALDH isoenzymes, Chem. Biol. Interact. 234 (2015) 18–28.

[166] J.K. Amory, C.H. Muller, J.A. Shimshoni, N. Isoherranen, J. Paik, J.S. Moreb, et al., Suppression of spermatogenesis by bisdichloroacetyldiamines is mediated by inhibition of testicular retinoic acid biosynthesis, J. Androl. 32 (2011) 111–119.

[167] J. Paik, M. Haenisch, C.H. Muller, A.S. Goldstein, S. Arnold, N. Isoherranen, et al., Inhibition of retinoic acid biosynthesis by the bisdichloroacetyldiamine WIN 18,446

markedly suppresses spermatogenesis and alters retinoid metabolism in mice, J. Biol. Chem. 289 (2014) 15104–15117.

[168] C.A. Morgan, T.D. Hurley, Characterization of two distinct structural classes of selective aldehyde dehydrogenase 1A1 inhibitors, J Med Chem. 58 (2015) 1964–1975.

[169] S. Perez-Miller, H. Younus, R. Vanam, C.-H. Chen, D. Mochly-Rosen, T.D. Hurley, Alda-1 is an agonist and chemical chaperone for the common human aldehyde dehydrogenase 2 variant., Nat. Struct. Mol. Biol. 17 (2010) 159–164.

[170] J.D. Durrant, C.A. de Oliveira, J.A. McCammon, POVME: an algorithm for measuring binding-pocket volumes, J Mol Graph Model. 29 (2011) 773–776.

[171] J.D. Durrant, L. Votapka, J. Sorensen, R.E. Amaro, POVME 2.0: An Enhanced Tool for Determining Pocket Shape and Volume Characteristics, J Chem Theory Comput. 10 (2014) 5047–5056.

[172] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics., J. Mol. Graph. 14 (1996) 33–8, 27–8.

[173] J. Kleinjung, P. Bayley, F. Fraternali, Leap-dynamics: efficient sampling of conformational space of proteins and peptides in solution, FEBS Lett. 470 (2000) 257–262.

[174] R.C. Walker, M.M. De Souza, I.P. Mercer, I.R. Gould, D.R. Klug, Large and fast relaxations inside a protein: Calculation and measurement of reorganization energies in alcohol dehydrogenase, J. Phys. Chem. B. 106 (2002) 11658–11665.

[175] J.J. Pavelites, J. Gao, P.A. Bash, A molecular mechanics force field for NAD+ NADH, and the pyrophosphate groups of nucleotides, J. Comput. Chem. 18 (1996) 221–239.

[176] D.A. Case, T. Darden, T.E. Cheatham, C. Simmerling, A. Roitberg, J. Wang, et al., Amber 14, Univ. California, San Fr. (2014).

[177] S. Sołobodowska, J. Giebułtowicz, R. Wolinowska, P. Wroczyński, Contribution of ALDH1A1 isozyme to detoxification of aldehydes present in food products., Acta Pol. Pharm. 69 (2012) 1380–3.

[178] P. McCaffery, J. Evans, O. Koul, A. Volpert, K. Reid, M.D. Ullman, Retinoid quantification by HPLC/MS [n], J. Lipid Res. 43 (2002) 1143–1149.

[179] J.E. Evans, P. McCaffery, HPLC/MS Analysis of Retinoids, in: H. Sun, G.H. Travis (Eds.), Retin. Methods Protoc. Chapter 8, Humana Press, Totowa, NJ, 2010: pp. 149–162.

[180] M.A. Kane, N. Chen, S. Sparks, J.L. Napoli, Quantification of endogenous retinoic acid in limited biological samples by LC/MS/MS, Biochem J. 388 (2005) 363–369.

[181] R.I. Feldman, H. Weiner, Horse liver aldehyde dehydrogenase. II. Kinetics and mechanistic implications of the dehydrogenase and esterase activity, J Biol Chem. 247 (1972) 267–272.

[182]   A. Moretti, J. Li, S. Donini, R.W. Sobol, M. Rizzi, S. Garavaglia, Crystal structure of human aldehyde dehydrogenase 1A3 complexed with NAD+ and retinoic acid, Sci. Rep. 6 (2016) 35710.

[183]   J. Brichac, K.H. Kwok, A. Honzatko, R. Wang, X. Lu, H. Weiner, et al., Enantioselective oxidation of trans-4-hydroxy-2-nonenal is aldehyde dehydrogenase isozyme and Mg2+dependent, Chem. Res. Toxicol. 20 (2007) 887–895.

[184]   A.B. Barua, J. a. Olson, Reversed-phase gradient high-performance liquid chromatographic procedure for simultaneous analysis of very polar to nonpolar retinoids, carotenoids and tocopherols in animal and plant samples, J. Chromatogr. B Biomed. Appl. 707 (1998) 69–79.

[185]   F.X. Ruiz, A. Moro, O. Gallego, A. Ardèvol, C. Rovira, J.M. Petrash, et al., Human and rodent aldo-keto reductases from the AKR1B subfamily and their specificity with retinaldehyde, Chem. Biol. Interact. 191 (2011) 199–205.

[186]   F.X. Ruiz, S. Porté, O. Gallego, A. Moro, A. Ardèvol, A. Del Río-Espínola, et al., Retinaldehyde is a substrate for human aldo–keto reductases of the 1C subfamily, Biochem. J. 440 (2011) 335–347.

[187]   S. Porté, F. Xavier Ruiz, J. Giménez, I. Molist, S. Alvarez, M. Domínguez, et al., Aldo-keto reductases in retinoid metabolism: Search for substrate specificity and inhibitor selectivity, Chem. Biol. Interact. 202 (2013) 186–194.

[188]   R. Ali, B. Campos, G. Dyckhoff, W.E. Haefeli, C. Herold-Mende, J. Burhenne, Quantification of retinoid concentrations in human serum and brain tumor tissues, Anal. Chim. Acta. 725 (2012) 57–66.

[189]   Y. Yamakoshi, H. Fukasawa, T. Yamauchi, H. Waki, T. Kadowaki, K. Shudo, et al., Determination of Endogenous Levels of Retinoic Acid Isomers in Type II Diabetes Mellitus Patients. Possible Correlation with HbA1c Values, Biol. Pharm. Bull. 25 (2002) 1268–1271.

[190]   Y.K. Kim, L. Quadro, HPLC/MSN Analysis of Retinoids, in: H. Sun, G.H. Travis (Eds.), Retin. Methods Protoc. Chapter 15, Humana Press, Totowa, NJ, 2010: pp. 149–162.

[191]   M.A. Kane, J.L. Napoli, HPLC/MSN Analysis of Retinoids, in: H. Sun, G.H. Travis (Eds.), Retin. Methods Protoc. Chapter 1, Humana Press, Totowa, NJ, 2010: pp. 1–54.

[192]   M. a Kane, A.E. Folias, J.L. Napoli, HPLC/UV quantification of retinal, retinol, and retinyl esters in serum and tissues, Anal. Biochem. 378 (2008) 71–79.

[193]   M. a Kane, A.E. Folias, C. Wang, J.L. Napoli, Quantitative profiling of endogenous retinoic acid in vivo and in vitro by tandem mass spectrometry, Anal. Biochem. 80 (2008) 1702–1708.

[194]   R. Wyss, F. Bucheli, Quantitative analysis of retinoids in biological fluids by high-performance liquid chromatography using column switching. I. Determination of isotretinoin and tretinoin and their 4-oxo metabolites in plasma., J. Chromatogr. 424 (1988) 303–14.

[195] K.J. Johnson, etinoids : Biology , Chemistry and Medicine, Biochem. Educ. Head. Hill Hall. 22 (1994).

[196] M. Domínguez, R. Pequerul, R. Alvarez, J. Giménez-Dejoz, E. Birta, S. Porté, et al., Synthesis of apocarotenoids by acyclic cross metathesis and characterization as substrates for human retinaldehyde dehydrogenases, Tetrahedron. 74 (2018) 2567–2574.

[197] A.F. Estrada, L. Youssar, D. Scherzinger, S. Al-Babili, J. Avalos, The ylo-1 gene encodes an aldehyde dehydrogenase responsible for the last reaction in the Neurospora carotenoid pathway, Mol. Microbiol. 69 (2008) 1207–1220.

[198] M. Foppiano, G. Lombardo, Worldwide pharmacovigilance systems and tolrestat withdrawal, Lancet. 349 (1997) 399–400.

[199] J. Liu, G. Wen, D. Cao, Aldo-keto reductase family 1 member B1 inhibitors: old drugs with new perspectives, Recent Pat. Anticancer. Drug Discov. 4 (2009) 246–253.

[200] J. Giménez Dejoz, S. Porté Orduna, J. Farrés i Vicén, X. Parés i Casasampera, Functional and structural studies of AKR1B15 and AKR1B16: Two novel additions to human and mouse aldo-keto reductase superfamily, Doctoral thesis. Universitat Autònoma de Barcelona, 2016.

[201] K. Metwally, H. Pratsinis, D. Kletsas, L. Quattrini, V. Coviello, C. La Motta, et al., Novel quinazolinone-based 2,4-thiazolidinedione-3-acetic acid derivatives as potent aldose reductase inhibitors, Future Med. Chem. 9 (2017) 2147–2166.

[202] I. Crespo, J. Giménez-Dejoz, S. Porté, A. Cousido-Siah, A. Mitschler, A. Podjarny, et al., Design, synthesis, structure-activity relationships and X-ray structural studies of novel 1-oxopyrimido[4,5-c]quinoline-2-acetic acid derivatives as selective and potent inhibitors of human aldose reductase, Eur. J. Med. Chem. 152 (2018) 160–174.

[203] J.Á. De La Fuente, S. Manzanaro, Aldose reductase inhibitors from natural sources, Nat. Prod. Rep. 20 (2003) 243–251.

[204] J.Á. De La Fuente, S. Manzanaro, M.J. Martín, T.G. De Quesada, I. Reymundo, S.M. Luengo, et al., Synthesis, Activity, and Molecular Modeling Studies of Novel Human Aldose Reductase Inhibitors Based on a Marine Natural Product, J. Med. Chem. 46 (2003) 5208–5221.

[205] A. Cousido-Siah, F.X. Ruiz, A. Mitschler, S. Porté, Á.R. de Lera, M.J. Martín, et al., Identification of a novel polyfluorinated compound as a lead to inhibit the human enzymes aldose reductase and AKR1B10: structure determination of both ternary complexes and implications for drug design, Chem Med Chem. 10 (2015) 1989–2003.

[206] S. Kumar, L.L. Sandell, P.A. Trainor, F. Koentgen, G. Duester, Alcohol and aldehyde dehydrogenases: retinoid metabolic effects in mouse knockout models, Biochim Biophys Acta. 1821 (2012) 198–205.

[207] L. Fares-Taie, S. Gerber, N. Chassaing, J. Clayton-Smith, S. Hanein, E. Silva, et al., ALDH1A3 Mutations Cause Recessive Anophthalmia and Microphthalmia, Am J Hum Genet. (2012).

[208]   R. Bchini, V. Vasiliou, G. Branlant, F. Talfournier, S. Rahuel-Clermont, Retinoic acid biosynthesis catalyzed by retinal dehydrogenases relies on a rate-limiting conformational transition associated with substrate recognition, Chem Biol Interact. 202 (2013) 78–84.

[209]   N.Y. Kedishvili, W.H. Gough, W.I. Davis, S. Parsons, T.K. Li, W.F. Bosron, Effect of cellular retinol-binding protein on retinol oxidation by human class IV retinol/alcohol dehydrogenase and inhibition by ethanol, Biochem. Biophys. Res. Commun. 249 (1998) 191–196.

[210]   A.A. Klyosov, Kinetics and specificity of human liver aldehyde dehydrogenases toward aliphatic, aromatic, and fused polycyclic aldehydes, Biochemistry. 35 (1996) 4457–4467.

[211]   P. V Bhat, H. Samaha, Kinetic properties of the human liver cytosolic aldehyde dehydrogenase for retinal isomers, Biochem Pharmacol. 57 (1999) 195–197.

[212]   J. Labrecque, F. Dumas, A. Lacroix, P. V Bhat, A novel isoenzyme of aldehyde dehydrogenase specifically involved in the biosynthesis of 9-cis and all-trans retinoic acid, Biochem J. 305 (1995) 681–684.

[213]   I. Gagnon, G. Duester, P. V Bhat, Kinetic analysis of mouse retinal dehydrogenase type-2 (RALDH2) for retinal substrates., Biochim. Biophys. Acta. 1596 (2002) 156–62.

[214]   A. Sima, M. Parisotto, S. Mader, P. V Bhat, Kinetic characterization of recombinant mouse retinal dehydrogenase types 3 and 4 for retinal substrates, Biochim Biophys Acta. 1790 (2009) 1660–1664.

[215]   A.M. Romani, Cellular magnesium homeostasis, Arch. Biochem. Biophys. 512 (2011) 1–23.

ANNEX

A
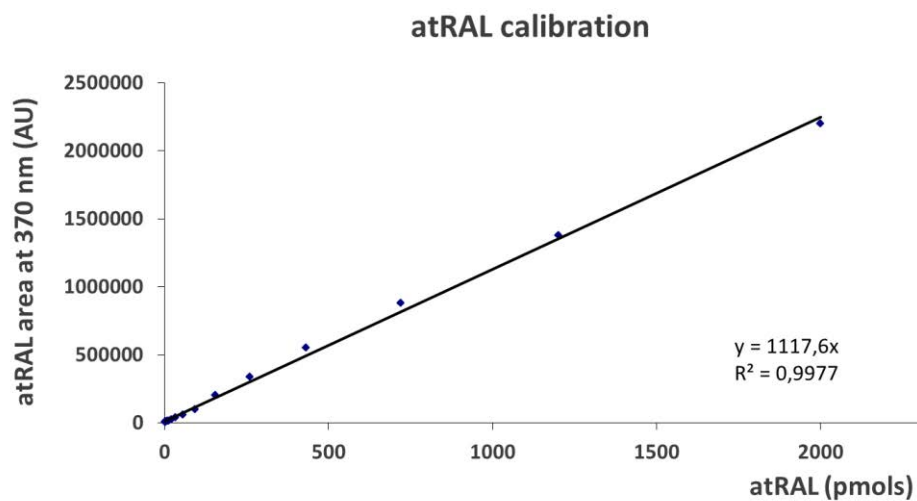
B
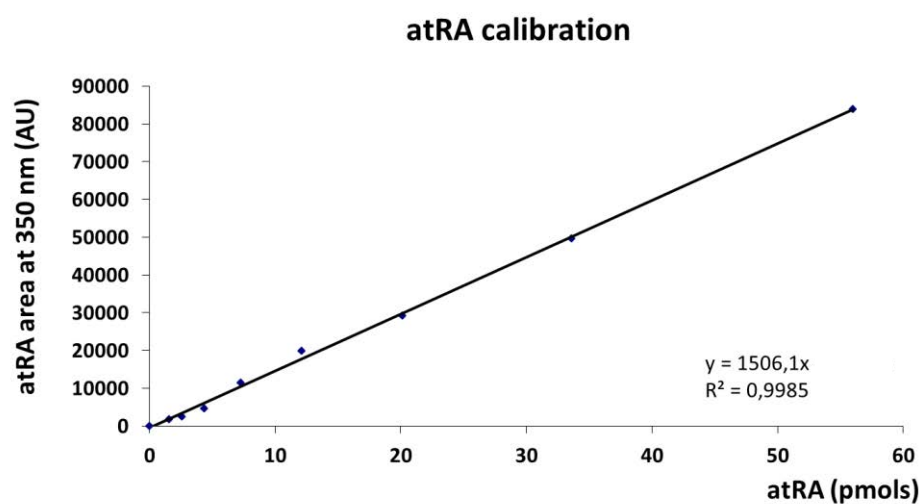
Annex 1.1. Calibration plots for (A) all-*trans*-retinaldehyde and (B) all-*trans*-retinoic acid. Duplicated standards with at least nine different concentrations (including blank) were prepared. Experimental values were fit by linear least square regression analysis. In all cases, the linearity obtained in the calibration plots showed an R^2 >99%, thus the slope of the calibration plots was used as the calibration factor to convert the peak area to pmol of compound.

12'-apo-β-carotenal calibration

A

y = 743,85x
R² = 0,9993

(x-axis: pmols 12'-apo-β-carotenal; y-axis: 12'-apo-β-carotenal area at 415 nm (AU))

12'-apo-β-carotenoic acid calibration

B

y = 609,37x
R² = 0,9991

(x-axis: pmols 12'-apo-β-carotenoic acid; y-axis: 12'-apo-β-carotenoic acid area at 400 nm (AU))

14'-apo-β-carotenal calibration

C

y = 241,29x
R² = 0,9998

(x-axis: pmols 14'-apo-β-carotenal; y-axis: 14'-apo-β-carotenal area at 400 nm (AU))

14'-apo-β-carotenoic acid calibration

D

y = 728,84x
R² = 0,9996

(x-axis: pmols 14'-apo-β-carotenoic acid; y-axis: 14'-apo-β-carotenoic acid area at 373 nm (AU))
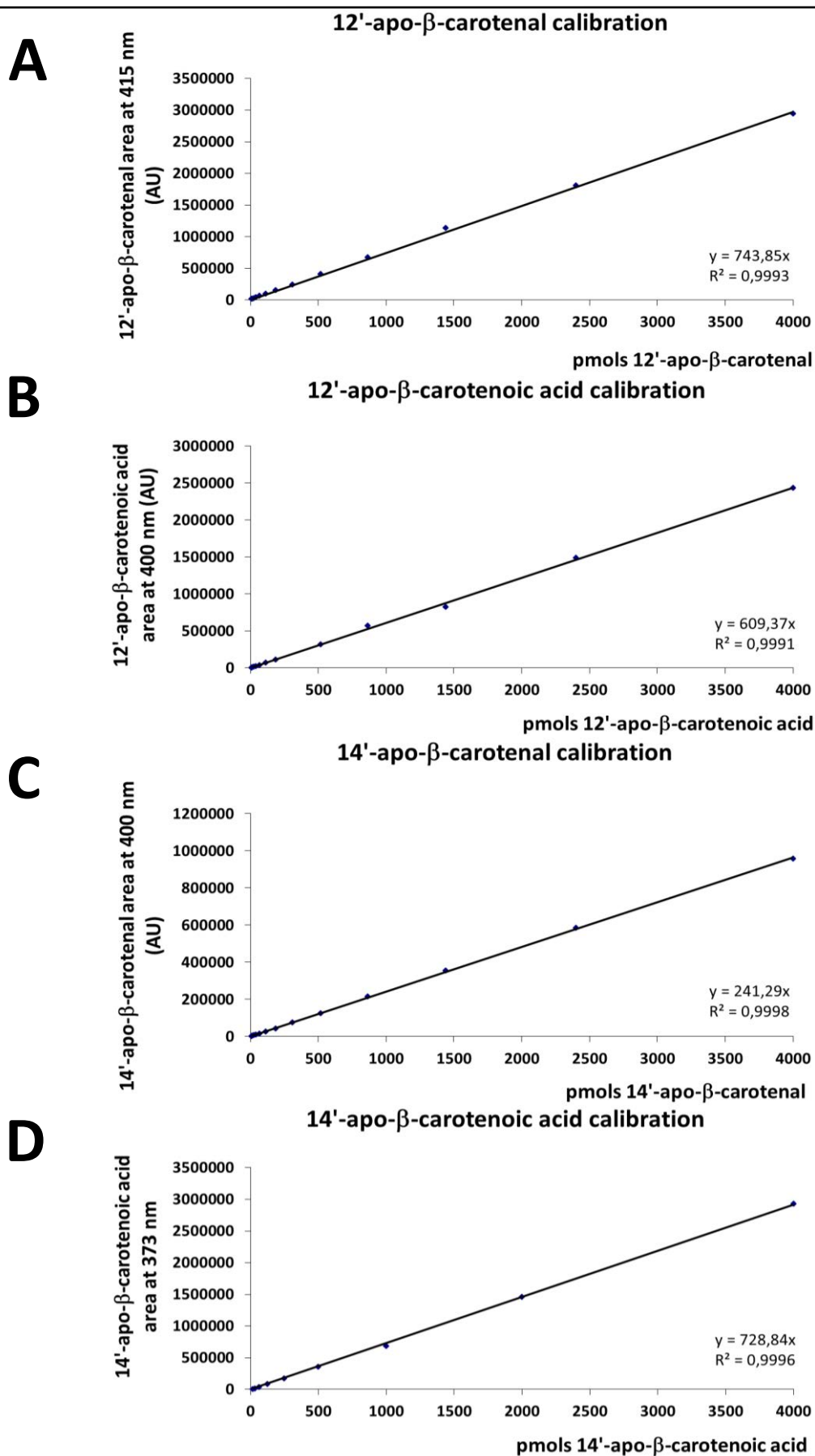
Annex 1.2. Calibration plots for (A) 12'-apo-β-carotenal, (B) 12'-apo-β-carotenoic acid, (C) 14'-apo-β-carotenal and (D) 14'-apo-β-carotenoic acid. Duplicated standards with at least twelve different concentrations (including blank) were prepared. Experimental values were fit by linear least square regression analysis. In all cases, the linearity obtained in the calibration plots showed an R^2 >99%, thus the slope of the calibration plots was used as the calibration factor to convert the peak area to pmol of compound.
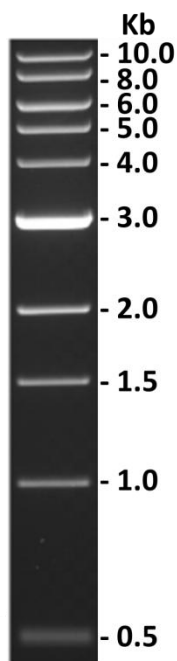
Annex 1.3. Gene-Ruler™ 1kb DNA Ladder (Thermo Scientific). The DNA marker is recommended for sizing and approximate quantification of a double-stranded DNA in the range of 500 bp to 10,000 bp on agarose gels. The DNA ladder consists of 10 DNA fragments and can be visualized after ethidium bromide or SYBR Safe staining. A uniform intensity of bands and bright reference bands allow easy sizing and orientation. A defined amount of DNA in each band enables approximate quantification of sample DNA.
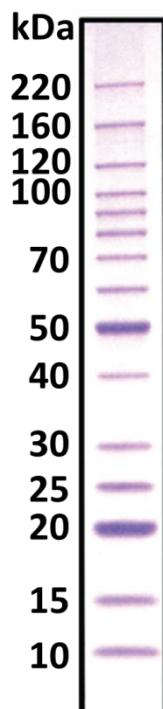
Annex 1.4. Bench-Mark^TM Protein Ladder (Invitrogen). The Bench-Mark^TM Protein Ladder is suitable for use as molecular weight standards for SDS-polyacrylamide gel electrophoresis. The ladder consists of 15 engineered proteins ranging in molecular weight from 10 to 220 kDa. The 20 and 50 kDa proteins are more prominent for easy orientation and to ensure proper identification of each protein.

- 134 -

Synthesis of apocarotenoids by acyclic cross metathesis and characterization as substrates for human retinaldehyde dehydrogenases

Marta Domínguez [a], Raquel Pequerul [b], Rosana Alvarez [a], Joan Giménez-Dejoz [b], Eszter Birta [c], Sergio Porté [b], Ralph Rühl [c, d], Xavier Parés [b], Jaume Farrés [b, *], Angel R. de Lera [a, *]

[a] Departamento de Química Orgánica, Facultade de Química, Universidade de Vigo, CINBIO and IBIV, Campus As Lagoas-Marcosende, 36310 Vigo, Spain
[b] Department of Biochemistry and Molecular Biology, Universitat Autònoma de Barcelona, Bellaterra, Barcelona E-08193, Spain
[c] Department of Biochemistry and Molecular Biology, University of Debrecen, Debrecen H-4028, Kassai út 26, PO Box 400, Hungary
[d] Paprika Bioanalytics BT, Debrecen H-4032, Hungary

ARTICLE INFO

ABSTRACT

A new synthesis of three apocarotenoids, namely 14′-apo-β-carotenal, 12′-apo-β-carotenal and 10′-apo-β-carotenal, has been achieved that is based on the acyclic cross-metathesis of the hexaene derived from retinal and the corresponding partners. These compounds can be enzymatically converted to their carboxylic acids by the human aldehyde dehydrogenases involved in retinaldehyde oxidation. Their kinetic parameters suggest that these enzymes might play a role in the physiological metabolism of apocarotenoids.

© 2018 Elsevier Ltd. All rights reserved.

1. Introduction

Carotenoids are a family of natural compounds synthesized by plants, microorganisms and some animals but not by humans.[1] Carotenoids are partly responsible for the colour in nature and play a key role both in the photosynthesis process and the photo-protection of the producing organisms. They can be broadly divided into two classes of chemical compounds: carotenes (e.g., β,β-carotene and lycopene) and their oxygenated derivatives termed xanthophylls (e.g., lutein, zeaxanthin and cryptoxanthin). Both carotenes and xanthophylls exhibit relevant physiological functions, serving as antioxidants in lipophilic environments,[2] a property that might contribute to the prevention of certain human diseases such as cardiovascular, ocular diseases and cancer.[3,4]

Within carotenoids, the term apocarotenoids is used to designate those with a backbone of less than 40 carbon atoms.[5] Apocarotenoids are formed by the oxidative degradation of one or both termini of carotenoids, a process that is catalyzed by carotenoid

cleavage enzymes. The oxidation products derived from dietary β,β-carotene **1** (Fig. 1) can be C20 all-*trans*-retinal **2** resulting from the central C15–C15′ cleavage catalyzed by BCO1[6,7] or non-symmetrical β-apocarotenoids obtained by eccentric cleavage catalyzed by BCO2.[8,9] Carotenoid metabolism appears to be cell compartmentalized,[10] as BCO1 is a cytosolic enzyme,[11] while BCO2 has been associated with mitochondria.[9] BCO2-mediated cleavage is the preferred pathway for xanthophylls.[9]

The metabolism of β,β-carotene **1** to all-*trans*-retinal **2** (Fig. 1) promoted by BCO1 is the first step in the production of the natural retinoids (including, but not restricted to, vitamin A, 11-*cis*-retinal and all-*trans*-retinoic acid **3**) required for eliciting the various functions of these compounds in the human body. Thus, enzymes responsible for isomerization and changes in the oxidation state provide the cognate ligands for receptors implicated in vision, cell proliferation, cell differentiation, immunity and development.[4] In particular, the oxidation of all-*trans*-retinal **2** promoted by aldehyde dehydrogenases (ALDHs)[12] generate all-*trans*-retinoic acid **3**, the natural hormone that binds to and activates the retinoic acid receptors (RARs),[13] members of the nuclear receptor superfamily of ligand-inducible transcription factors.[14]

Apocarotenoids (such as **4**–**11**, Fig. 1) have been detected in food and in the blood of animals,[15] and the concentration of some of

---
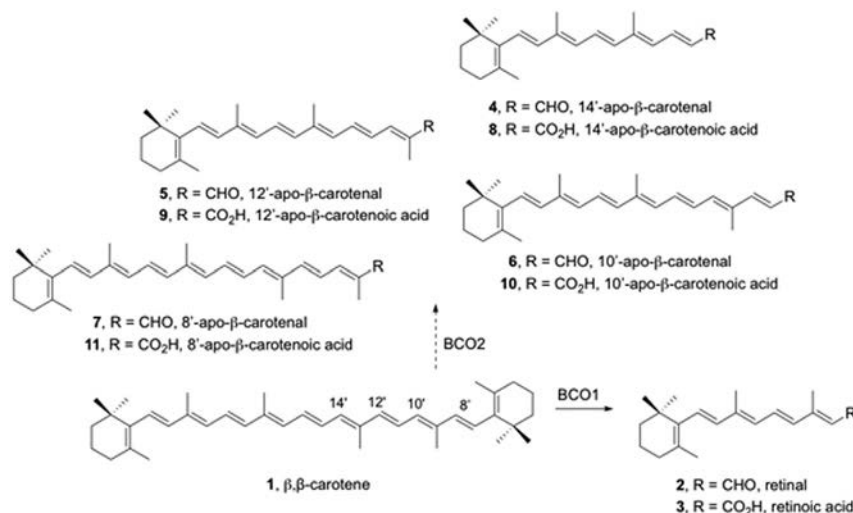
Fig. 1. β-apocarotenals **2,4**–**7** potentially formed by (enzyme-mediated) oxidative cleavage of β,β-carotene **1**, and the corresponding carboxylic acids **3**, **8**–**11**. The shorter fragments of putative BCO2 cleavage are not shown.

them are similar to that of all-*trans*-retinoic acid **3**. Their biological functions, however, remain unclear.[16,17] Recently, it has been proposed that the activity of BCO2-derived metabolites can protect against the damage induced by β,β-carotene **1**. This compound is deemed partially responsible for oxidative stress in the mitochondria, a process that can trigger signaling pathways related to cell survival and proliferation.[17] Moreover, some apocarotenoids, namely 10′-apo-β-carotenoic acid **10**, 12′-apo-β-carotenoic acid **9**,[17] 14′-apo-β-carotenoic acid **8**, 14′-apo-β-carotenal **4**, and 13′-apo-β-carotenone (not shown), have been described as low-affinity agonists potentially acting as endogenous antagonists of RARs,[16,17] and also were found to regulate other functions, such as the placental lipoprotein biosynthesis.[18]

It is believed that β-apocarotenals can be a potential source of β-apocarotenoic acids, which can be considered as vinylogues of all-*trans*-retinoic acid **3**, through the oxidation of the functional group. However the enzymes involved in these transformations have not been described. Logical candidates are aldehyde dehydrogenases (ALDHs),[12] enzymes that transform aldehydes into carboxylic acids. Within the ALDH superfamily, ALDH1A1, ALDH1A2 and ALDH1A3 are closely related enzymatic forms that catalyze the oxidation of retinaldehydes.[19] In order to shed some light on the metabolism and biological functions of β-apocarotenoids, here we present an efficient synthetic strategy for the preparation of these molecules and their biological characterization as substrates for human ALDHs.

Carotenoids have been traditionally synthesized following two different strategies.[4,20] The first comprises carbonyl condensation reactions with heteroatom-stabilized carbanions, such as Wittig, Horner–Wadsworth–Emmons and Julia reactions, which form $C_{sp2} = C_{sp2}$ bonds.[21] The second is based on the formation of $C_{sp2}-C_{sp2}$ bonds by palladium-catalyzed cross-coupling reactions[22] (primarily Negishi, Stille, and Suzuki reactions).[23] In both cases, appropriate functionalization of the intermediates for the key reaction is required. Recently, the olefin metathesis reaction[24–26] has been established as one of the most general and widely applicable synthetic methods for $C_{sp2} = C_{sp2}$ bond formation. Despite the numerous applications of olefin metathesis reactions in the synthesis of natural products,[24] its use in the preparation of conjugated polyene chains has been somehow limited because of concerns about the control of site-selectivity, stereoselectivity, and the stability of polyenes to the reaction conditions. The first application on

the synthesis of retinoids[27] and apocarotenoids starting from carotenoids was described by Wojtkielewicz and coworkers, although the products were obtained in very low yields.[28,29] We[30,31] and others[32] have also contributed to this field with the synthesis of symmetrical and non-symmetrical carotenoids by dimerization and cross-metathesis processes.[33]

We considered that β-apocarotenoids could also be prepared by cross metathesis of appropriate precursors.[33] We envisioned the synthesis of β-apocarotenoids with different lengths of the polyene chain by acyclic cross-metathesis[33] starting from a common precursor, the already known hexaene **12**[30] and the complementary component functionalized as ester.[30] More specifically, we undertook the synthesis of the three β-apocarotenoids (10′-apo-β-carotenal **4**, 12′-apo-β-carotenal **5** and 14′-apo-β-carotenal **6**; 8′-apo-β-carotenal **7** is a commercial compound) that could potentially by formed by eccentric cleavage of β,β-carotene **1** at the C9′ = C10′, C11′ = C12′ and C13′ = C14′ double bonds.[4] Only the products of cleavage at the C9′ = C10′ bond have been characterized in mice.[7] However, since β-apocarotenoids can potentially be formed from β,β-carotene **1** by non-enzymatic autoxidation processes, their availability by synthesis can provide useful tools for biochemical research. A general method for oxidation of β,β-carotene **1** with a mixture of $KMnO_4$ and H_2O_2, that provided a mixture of three apocarotenoids (8′-, 10′- and 12′-apo), was reported half a century ago.[34]

2. Results and discussion

2.1. Synthesis of 14′-apo-β-carotenal **4**

The synthesis of 14′-apo-β-carotenal **4** started with the cross metathesis[33] of previously described hexaene **12**[30] and commercial butyl (*E*)-but-2-enoate **13**. Four different ruthenium catalysts were tested with toluene as solvent based on the results described by Wojtkielewicz and coworkers[29] and special attention was paid to the reaction time in order to avoid degradation of the formed polyenes (Scheme 1). The use of Neolyst® as catalyst[35] was discouraging as the starting material was fully recovered after 48 h (entry 1, Table 1). The use of Nitro-Grela catalyst[36] allowed to obtain, after 7.5 h, the desired product in 29% yield together with 15% of β,β-carotene **1**, resulting from the competing dimerization of **12**[30] (entry 2, Table 1).
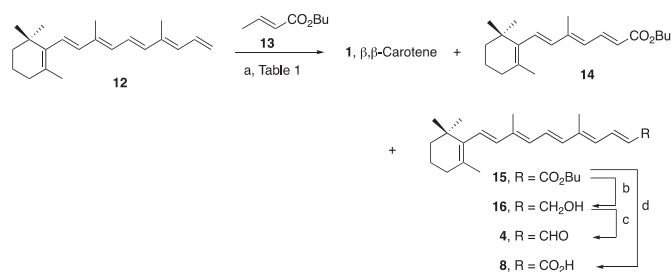
Scheme 1. Reagents and conditions. a) Table 1 b) Dibal-H, THF, −78 °C, 5 h, 71%. c) MnO₂, Na₂CO₃, CH₂Cl₂, 25 °C, 5 h, 88%. d) KOH, MeOH, 70 °C, 1 h, 62%.

Table 1
Olefin cross-metathesis of hexaene **12** and enoate **13**.

	Reagents and conditions	Yield (%)		
		1	14	15
1	**13** (6 equiv.), Neolyst® (0.15 equiv.), toluene, 25 °C, 48 h	0	0	0
2	**13** (6 equiv.), Nitro-Grela (0.15 equiv.), toluene, 25 °C, 7.5 h	15	0	29
3	**13** (6 equiv.), 2nd gen. Grubbs, toluene, 25 °C, 4.5 h	15	0	39
4	**13** (6 equiv.), 2nd gen. Hoveyda-Grubbs (0.15 equiv.), toluene, 25 °C, 4.5 h	16	1	62
5	**13** (6 equiv.), 2nd gen. Hoveyda-Grubbs (0.15 equiv.), toluene, 25 °C, 24 h	0	36	0

The 2nd generation Grubbs catalyst[26,37] proved more effective in the metathesis reaction (39% yield for **15**), although β,β-carotene **1** was also obtained as secondary product (entry 3, Table 1). Finally, the use of 2nd generation Hoveyda-Grubbs (HG) catalyst[38] afforded the best results, with a 62% yield of desired **15** and 16% of homodimer **1** obtained after 4.5 h (entry 4, Table 1). The increase of the reaction time turned out to be detrimental since it favored the cross metathesis corresponding formally to the C11 = C12 bond, with full consumption of the initially formed apo-β-carotenoid **15** (entry 5, Table 1). Compound **15** was readily transformed into the desired 14′-apo-β-carotenal **4** by Dibal-H reduction followed by oxidation of **16** with MnO₂ under basic conditions (Scheme 1). In addition, saponification of ester **15** with KOH/MeOH at 70 °C afforded 14′-apo-β-carotenoic acid **8** in 62% yield.

2.2. Synthesis of 12′-apo-β-carotenal **5**

The previously optimized reaction conditions were next applied to the synthesis of 12′-apo-β-carotenal **5**. The reaction of an excess of **17** with hexaene **12** afforded methyl 12′-apo-β-carotenoate **18** in
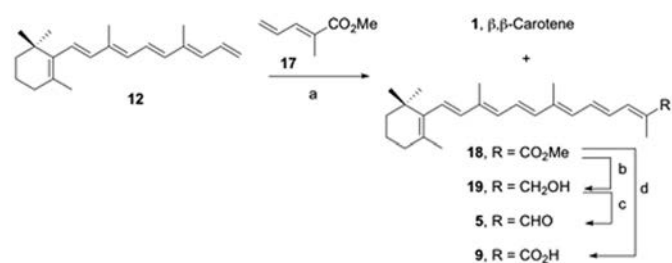
40% yield along with the product of homodimerization **1** (10%). Some reactant hexaene **12** was also recovered (11%). The geometry of the newly formed double bond was confirmed by NOE experiments. Dibal-H reduction of ester **18** and subsequent oxidation with MnO₂ under basic conditions furnished 12′-apo-β-carotenal **5** (Scheme 2). Ester hydrolysis as described led to 12′-apo-β-carotenoic acid **9** in 53% yield.

2.3. Synthesis of 10′-apo-β-carotenal **6**

The anticipated component for the synthesis of 10′-apo-β-carotenal **6**, namely methyl (2E,4E)-octan-2,4,6-enoate **24**, was prepared by a six step sequence starting from methyl (2E,4E)-3-methyl-6-oxohexa-2,4-dienoate **20**. Protection of the aldehyde of **20** as dioxolane, followed by the consecutive ester reduction with Dibal-H and MnO₂ oxidation, led to aldehyde **21** in good yields. Wittig olefination with commercially available ethyl triphenylphosphonium iodide **22** provided triene **23** in 80% yield as an inconsequential 1:1 mixture of double bond isomers. Subsequent deprotection of the acetal followed by oxidation of the trienal with MnO₂ and KCN in MeOH provided ester **24** as a mixture of isomers (Scheme 3). An excess of trienyl ester **24** was reacted with hexaene **12** under the conditions already described, which produced the desired methyl 10′-apo-β-carotenoate **25** in 24% yield along with 58% yield of β,β-carotene **1**. A two-step sequence of Dibal-H reduction and MnO₂ oxidation of ester **25** afforded 10′-apo-β-carotenal **6** in a combined 59% yield.

2.4. ALDH enzymatic assay

The activity of ALDH1A enzymes with apo-β-carotenals **4** and **5** (the study with **6** was not completed due to the low overall yield of the synthetic sequence) was monitored using an improved HPLC-based method for the simultaneous detection of substrates and reaction products. In all cases, concomitant consumption of apo-β-carotenals **4** and **5**, and production of the corresponding apo-β-carotenoic acids **8** and **9**, respectively, were observed (see Fig. 2 and Supporting Information). Thus, ALDH1A enzymes turned out to be active with 14′- and 12′-apo-β-carotenals (**4** and **5**).

Table 2 compares the kinetic constants of the ALDH1A enzymes with 12′- and 14′-apo-β-carotenals (**5** and **4**), showing some differential substrate specificity. In general, K_m values were in the micromolar range, except for ALDH1A1 and ALDH1A3, which were not saturated with 12′-apo-β-carotenal **5** and 14′-apo-β-carotenal **4**, respectively (Fig. 3). While ALDH1A1 showed higher specificity for 14′-apo-β-carotenal **4**, ALDH1A2 exhibited higher catalytic efficiency for 12′-apo-β-carotenal **5**. This catalytic efficiency is about ten-fold higher than those of ALDH1A1 and ALDH1A3, and similar to that observed for these enzymes with all-*trans*-retinaldehyde **2**.

Scheme 2. Reagents and conditions. a) 2nd generation HG, toluene, 25 °C, 7 h (10% β,β-carotene **1**, 11% starting **12**, 40% methyl 12′-apo-β-carotenoate **18**). b) Dibal-H, THF, −78 °C, 4 h, 75%. c) MnO₂, Na₂CO₃, CH₂Cl₂, 25 °C, 4 h, 38%. d) KOH, MeOH, 70 °C, 1 h, 53%.
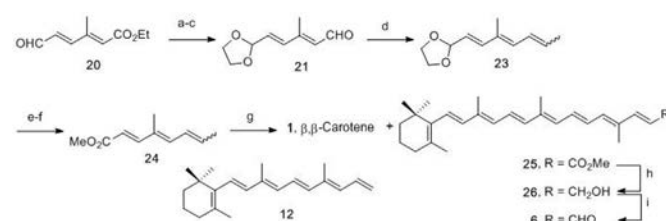
Scheme 3. Reagents and conditions: a) ethylene glycol, pTsOH, benzene, 96 °C, 16 h, 99%. b) Dibal-H, THF -78 °C, 2 h, 99%. c) MnO₂, Na₂CO₃, THF, 25 °C, 5 h, 75%. d) ethyl triphenylphosphonium iodide **22**, NaHMDS, THF, −78 °C, 2.5 h, 80%. e) pTsOH, acetone, 25 °C, 17 h, 98%. f) MnO₂, KCN, MeOH, 25 °C, 3.5 h, 79%. g) 2nd generation HG, **12**, toluene, 5 h (58% β,β-carotene **1**, 49% starting trienoate **24**, 24% methyl 10′-apo-β-carotenoate **25**). h) Dibal-H, THF, −78 °C. i) MnO₂, Na₂CO₃, CH₂Cl₂, 59% (2 steps).
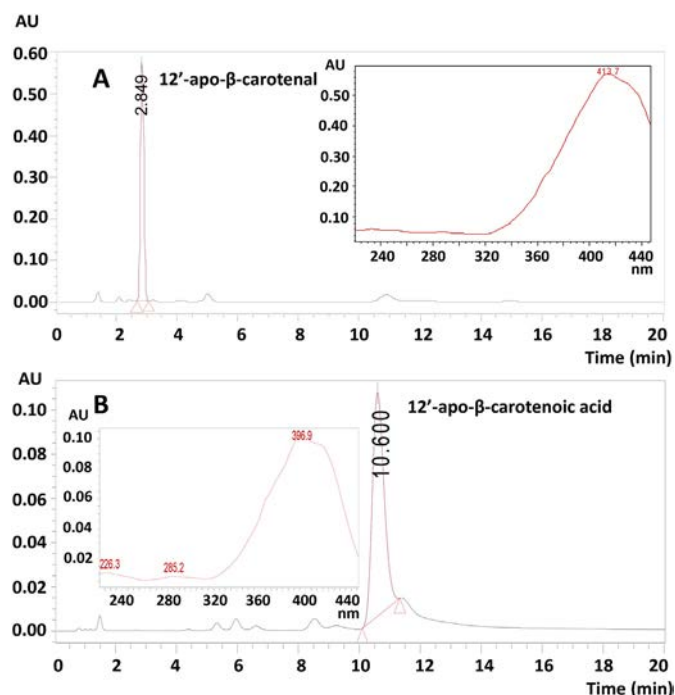
Fig. 2. HPLC elution profiles of 12′-apo-β-carotenal 5 (A) and 12′-apo-β-carotenoic acid 9 (B). Stock solutions of apocarotenoids were prepared in ethanol and 10 μM of either compound was injected. Each compound was separated by liquid chromatography on a NovaPak® silica gel column (4 μm, 3.9 × 150 mm) in hexane/tert-butyl methyl ether (96:4, v/v) mobile phase, at a flow rate of 2 mL/min using a Waters Alliance 2695 HPLC. Elution of 12′-apo-β-carotenal 5 and 12′-apo-β-carotenoic acid 9 was monitored with a Waters 2996 photodiode array detector at 415 and 400 nm, respectively. The retention times for 12′-apo-β-carotenal 5 and 12′-apo-β-carotenoic acid 9 were 2.85 and 10.60 min, respectively. Insets: UV−vis absorption spectra of the two compounds.

The k_{cat} values are in the same range in all the experiments performed, but lower than those for all-*trans*-retinaldehyde 2 (data not shown; to be reported elsewhere).

To date, the enzymatic activity of ALDH with these compounds had only been reported in microorganisms.[39] This is the first report on the apocarotenal dehydrogenase activity of human ALDH1A enzymes.

3. Conclusion

We have developed a new stereoselective synthesis of three apocarotenoids (14′-apo-β-carotenal 4, 12′-apo-β-carotenal 5 and 10′-apo-β-carotenal 6) and demonstrated that the olefin metathesis protocol is a valid synthetic method for accessing these conjugated polyenes. Whereas the acyclic cross-metathesis of the common required hexaene was efficiently performed with the shorter enoate and dienoate partners, the chemoselectivity with the longer trienoates is poor due to the presence of several *trans*-disubstituted olefins that can compete. Previous syntheses of these apo-β-

Fig. 3. Representative Michaelis-Menten kinetics of ALDH1A2 with 12′-apo-β-carotenal 5. The reaction was carried out in 50 mM HEPES, 0.5 mM EDTA, 0.5 mM DTT, pH 8.0, at 37 °C. The initial rates were measured with at least six different substrate concentrations. Experimental values were adjusted to the Michaelis-Menten equation using the non-linear regression program GraFit 5.0 (Eritacus software).

carotenoids have in general featured carbonyl condensation reactions, such as Wittig and aldol condensations (4,[40] 5[40,41] and 2-4[17]) or enol ether reactions with acetals followed by elimination and hydrolysis in the case of 4.[42]

We have shown that human retinaldehyde dehydrogenases are capable of using apo-β-carotenals as substrates with good catalytic efficiency. ALDH1A1 and ALDH1A2 exhibit higher substrate specificity for 4 and 5, respectively, while ALDH1A3 shows similar specificity for the two compounds. Overall, the substrate specificity of ALDH1 enzymes with apo-β-carotenals is comparable to that with all-*trans*-retinaldehyde 2, indicating that these enzymes might play a physiological role in the metabolism of apocarotenoids.

4. Experimental section

General experimental procedures. See E.S.I.

4.1. Butyl (2E,4E,6E,8E,10E)-5,9-dimethyl-11-(2,6,6-trimethylcyclohex-1-en-1-yl)undeca-2,4,6,8,10-pentaenoate (butyl 14′-β-apocarotenoate) (15)

General procedure for the olefin metathesis reaction. To a solution of (1E,3E,5E,7E)-2-(3,7-dimethyldeca-1,3,5,7,9-pentaenyl)-1,3,3-trimethylcyclohex-1-ene 12 (14 mg, 0.049 mmol) and butyl (E)-but-2-enoate 13 (46.4 μL, 0.294 mmol) in toluene (0.25 mL), 2nd generation Hoveyda-Grubbs catalyst (0.46 mg, 0.0073 mmol) was added and the reaction mixture was thoroughly degassed. After stirring at 25 °C for 4.5 h, the mixture was filtered through a Celite™ pad, which was washed with Et₂O and the solvent was evaporated. The residue was purified by column chromatography (silica gel, 99:1 hexane/EtOAc) to afford, in order of elution, 2.2 mg (16%) of a red solid identified as β,β-carotene 1, 11.8 mg (62%) of a

Table 2
Kinetic constants of human ALDH1A enzymes with apo-β-carotenals.

Substrate	ALDH1A1			ALDH1A2			ALDH1A3		
	K_m (μM)	k_{cat} (min⁻¹)	k_{cat}/K_m (mM⁻¹ min⁻¹)	K_m (μM)	k_{cat} (min⁻¹)	k_{cat}/K_m (mM⁻¹ min⁻¹)	K_m (μM)	k_{cat} (min⁻¹)	k_{cat}/K_m (mM⁻¹ min⁻¹)
12′-apo-β-carotenal 5	NS	–	44	1.8 ± 0.2	8.5 ± 0.2	4800 ± 600	3.0 ± 0.4	1.3 ± 0.1	450 ± 62
14′-apo-β-carotenal 4	12 ± 2	3.1 ± 0.2	257 ± 47	5.0 ± 0.5	4.7 ± 0.2	930 ± 93	NS	–	303

ᵃEnzymatic activity was measured in 50 mM HEPES, pH 8.0, 0.5 mM NAD⁺, including 0.5 mM EDTA and 0.5 mM DTT for ALDH1A1 and ALDH1A2, or 30 mM MgCl₂ and 5 mM DTT for ALDH1A3, at 37 °C. Apocarotenoic acid production was quantified by using an HPLC-based method. NS, no saturation (60 μM was the highest substrate concentration tested in the assay), where the k_{cat}/K_m value was calculated from the slope of V/[E] vs [S] plot.

red oil identified as butyl (2*E*,4*E*,6*E*,8*E*,10*E*)-5,9-dimethyl-11-(2,6,6-trimethylcyclohex-1-en-1-yl)undeca-2,4,6,8,10-pentaenoate **15** and 0.2 mg (1%) of a yellow oil identified as butyl (2*E*,4*E*,6*E*)-5-methyl-7-(2,6,6-trimethylcyclohex-1-en-1-yl) hepta-2,4,6-trienoate **14**.

Data for β,β-carotene (**1**): ^{1}H NMR (400 MHz, C_6D_6): δ 6.79 (dd, $J = 14.8$, 11.4 Hz, 2H, $2H_{11}$), 6.70–6.67 (m, 2H, H_{15}), 6.49 (d, $J = 14.8$ Hz, 2H, H_{12}), 6.42–6.31 (m, 8H, $2H_7 + 2H_8 + 2H_{10} + 2H_{14}$), 1.99 (t, $J = 5.6$ Hz, 4H, $2 \times CH_2$), 1.94 (s, 6H, $2 \times CH_3$), 1.88 (s, 6H, $2 \times CH_3$), 1.82 (s, 6H, $2 \times CH_3$), 1.64–1.56 (m, 4H, $2 \times CH_2$), 1.52–1.49 (m, 4H, $2 \times CH_2$), 1.16 (s, 12H, $4 \times CH_3$) ppm. HRMS (ESI$^+$): Calcd. for $C_{40}H_{56}$ ([M+H]$^+$), 536.4376; found, 536.4376. UV (MeOH): λ_{max} 449, 241 nm.

Data for butyl (2*E*,4*E*,6*E*,8*E*,10*E*)-5,9-dimethyl-11-(2,6,6-trimethylcyclohex-1-en-1-yl)undeca-2,4,6,8,10-pentaenoate (**15**): ^{1}H NMR (400 MHz, CO(CD$_3$)$_2$): δ 7.71 (dd, $J = 15.0$, 12.0 Hz, 1H, H_{15}), 7.02 (dd, $J = 15.0$, 11.4 Hz, 1H, H_{11}), 6.48 (d, $J = 15.0$ Hz, 1H, H_{12}), 6.34 (d, $J = 11.8$ Hz, 1H, H_{14}), 6.42–6.18 (m, 3H, $H_7 + H_8 + H_{10}$), 5.97 (d, $J = 15.0$ Hz, 1H, $H_{15'}$), 4.16 (t, $J = 6.6$ Hz, 2H, O–CH$_2$–), 2.13 (s, 3H, CH$_3$), 2.10–2.06 (m, 2H, CH$_2$), 2.06 (s, 3H, CH$_3$), 1.75 (s, 3H, CH$_3$), 1.72–1.61 (m, 4H, $2 \times CH_2$), 1.54–1.48 (m, 2H, CH$_2$), 1.47–1.39 (m, 2H, CH$_2$), 1.07 (s, 6H, $2 \times CH_3$), 0.97 (t, $J = 7.4$ Hz, 3H, CH$_2$CH$_3$) ppm. ^{13}C NMR (100 MHz, (CD$_3$)$_2$CO): 167.4 (s), 145.2 (s), 140.8 (d), 138.8 (s), 138.7 (s), 138.6 (d), 137.0 (d), 131.4 (d), 130.1 (s), 129.5 (d), 129.4 (d), 128.4 (d), 121.4 (d), 64.4 (t), 40.4 (t), 34.9 (s), 33.6 (t), 31.6 (t), 29.3 (q, 2×), 21.9 (q), 19.9 (t), 19.8 (t), 14.0 (q), 13.1 (q), 12.9 (q) ppm. HRMS (ESI$^+$): Calcd. for $C_{26}H_{39}O_2$ ([M+H]$^+$), 383.2945; found, 383.2946. IR (NaCl): ν 2957 (m, C–H), 2928 (m, C–H), 2866 (w, C–H), 1708 (s, C=O), 1615 (m), 1563 (m), 1135 (s), 970 (m) cm^{-1}. UV (MeOH): λ_{max} 394 nm.

Data for butyl (2*E*,4*E*,6*E*)-5-methyl-7-(2,6,6-trimethylcyclohex-1-en-1-yl)hepta-2,4,6-trienoate (**14**): ^{1}H NMR (400 MHz, CDCl$_3$): δ 7.71 (dd, $J = 15.0$, 11.8 Hz, 1H, H_3), 6.38 (d, $J = 16.1$ Hz, 1H, H_6 or H_7), 6.18–6.10 (m, 2H, H_6 or $H_7 + H_4$), 5.87 (d, $J = 15.5$ Hz, 1H, H_2), 4.16 (t, $J = 6.7$ Hz, 2H, O–CH$_2$), 2.04 (s, 3H, CH$_3$), 2.05–1.95 (m, 2H, CH$_2$), 1.71 (s, 3H, CH$_3$), 1.68–1.58 (m, 4H, $2 \times CH_2$), 1.50–1.37 (m, 4H, $2 \times CH_2$), 1.03 (s, 6H, $2 \times CH_3$), 0.95 (t, $J = 7.4$ Hz, 3H, CH$_2$CH$_3$) ppm. ^{13}C NMR (100 MHz, CDCl$_3$): δ 167.8 (s), 144.4 (s), 140.7 (d), 137.6 (s), 136.9 (d), 130.9 (s), 130.8 (d), 127.4 (d), 120.2 (d), 64.3 (t), 39.7 (t), 34.4 (s), 33.3 (t), 30.9 (t), 29.1 (q, 2×), 21.9 (q), 19.4 (t), 19.3 (t), 13.9 (q), 13.2 (q) ppm. HRMS (ESI$^+$): Calcd. for $C_{21}H_{33}O_2$ ([M+H]$^+$), 317.2475; found, 317.2472. IR (NaCl): ν 2958 (m, C–H), 2929 (m, C–H), 1711 (s, C=O), 1598 (m), 1144 (m), 977 (m) cm^{-1}. UV (MeOH): λ_{max} 329 nm.

4.2. (2E,4E,6E,8E,10E)-5,9-dimethyl-11-(2,6,6-trimethylcyclohex-1-en-1-yl) undeca-2,4,6,8,10-pentaenal (14′-apo-β-carotenal) (**4**)

General procedure for Dibal-H reduction of esters. To a cooled (−78 °C) solution of (2*E*,4*E*,6*E*,8*E*,10*E*)-5,9-dimethyl-11-(2,6,6-trimethylcyclohex-1-en-1-yl)undeca-2,4,6,8,10-pentaenoate **15** (13.2 mg, 34.5 mmol) in THF (0.175 mL) was added Dibal-H (0.103 mL, 1 M in hexanes, 0.103 mmol). After stirring for 5 h at −78 °C, H$_2$O was added and the mixture was extracted with EtOAc (3×). The combined organic layers were washed with brine (3×), dried (Na$_2$SO$_4$) and the solvent was evaporated. The residue was purified by column chromatography (silica gel, 93:4:3 hexane/EtOAc/Et$_3$N) to afford 7.7 mg (71%) of a yellow oil identified as (2*E*,4*E*,6*E*,8*E*,10*E*)-5,9-dimethyl-11-(2,6,6-trimethylcyclohex-1-en-1-yl)undeca-2,4,6,8,10-pentaene-1-ol **16**, which was used without further purification.

General procedure for MnO$_2$ oxidation of alcohols. To a cooled (0 °C) solution of (2*E*,4*E*,6*E*,8*E*,10*E*)-5,9-dimethyl-11-(2,6,6-trimethylcyclohex-1-en-1-yl)undeca-2,4,6,8,10-pentaene-1-ol **16** (28.5 mg, 0.0912 mmol) in CH$_2$Cl$_2$ (1.7 mL) MnO$_2$ (0.143 g,

1.64 mmol) and Na$_2$CO$_3$ (0.174 g, 1.64 mmol) were added. After stirring for 5 h at 25 °C, the reaction mixture was filtered through a Celite™ pad and the solvent was removed. The residue was purified by column chromatography (silica gel, 95:2:3 hexane/EtOAc/Et$_3$N) to afford 24.8 mg (88%) of a yellow oil identified as (2*E*,4*E*,6*E*,8*E*,10*E*)-5,9-dimethyl-11-(2,6,6-trimethylcyclohex-1-en-1-yl)undeca-2,4,6,8,10-pentaenal **4**. ^{1}H NMR (400 MHz, (CD$_3$)$_2$CO): δ 9.65 (d, $J = 8.0$ Hz, 1H, $H_{14'}$), 7.76 (dd, $J = 14.9$, 12.0 Hz, 1H, H_{15}), 7.07 (dd, $J = 15.0$, 11.5 Hz, 1H, H_{11}), 6.50 (d, $J = 15.0$ Hz, 1H, H_{12}), 6.48 (d, $J = 11.8$ Hz, 1H, H_{14}), 6.38–6.18 (m, 3H, $H_7 + H_8 + H_{10}$), 6.16 (dd, $J = 14.9$, 8.0 Hz, 1H, $H_{15'}$), 2.17 (s, 3H, CH$_3$), 2.08–2.04 (m, 2H, CH$_2$), 2.05 (s, 3H, CH$_3$), 1.73 (s, 3H, CH$_3$), 1.68–1.60 (m, 2H, CH$_2$), 1.52–1.47 (m, 2H, CH$_2$), 1.05 (s, 6H, $2 \times CH_3$) ppm. ^{13}C NMR (100 MHz, (CD$_3$)$_2$CO): δ 193.7 (d), 148.4 (d), 147.2 (s), 139.5 (s), 138.5 (s), 138.4 (d), 136.8 (d), 131.9 (d), 131.4 (d), 130.5 (d), 130.2 (s), 129.7 (d), 128.8 (d), 40.3 (t), 34.8 (s), 33.6 (t), 29.3 (q, 2×), 22.0 (q), 19.9 (t), 13.2 (q), 12.9 (q) ppm. HRMS (ESI$^+$): Calcd. for $C_{22}H_{30}O$ ([M+H]$^+$), 311.2369; found, 311.2376. IR (NaCl): ν 2926 (m, C–H), 2863 (w, C–H), 1670 (s, C=O), 1559 (m), 1124 (m), 967 (m) cm^{-1}. UV (MeOH): λ_{max} 402 nm ($\varepsilon = 45\,950$ mol^{-1} L cm^{-1}).

4.3. (2E,4E,6E,8E,10E)-5,9-dimethyl-11-(2,6,6-trimethylcyclohex-1-en-1-yl)undeca-2,4,6,8,10-pentaenoic acid (14′-β-apocarotenoic acid) (**8**)

General procedure for hydrolysis of esters with KOH. To a solution of (2*E*,4*E*,6*E*,8*E*,10*E*)-5,9-dimethyl-11-(2,6,6-trimethylcyclohex-1-en-1-yl)undeca-2,4,6,8,10-pentaenoate **15** (19.6 mg, 51.2 mmol) in MeOH (3.53 mL) was added KOH (0.82 mL, 2 M in H$_2$O, 1.64 mmol). After stirring at 70 °C for 1 h, the reaction mixture was cooled down to 25 °C, CH$_2$Cl$_2$ and brine were added and the layers were separated. The organic layer was washed with H$_2$O (3×). The combined aqueous layers were acidified until acidic pH and extracted with CH$_2$Cl$_2$ (3×). The combined organic layers were dried (Na$_2$SO$_4$) and the solvent was evaporated. The residue was purified by column chromatography (silica gel, 70:30 hexane/EtOAc) to afford 10.4 mg (62%) of an orange solid identified as (2*E*,4*E*,6*E*,8*E*,10*E*)-5,9-dimethyl-11-(2,6,6-trimethylcyclohex-1-en-1-yl)undeca-2,4,6,8,10-pentaenoic acid **8**. ^{1}H NMR (400 MHz, (CD$_3$)$_2$CO): δ 7.68 (dd, $J = 14.9$, 12.0 Hz, 1H, H_{15}), 6.98 (dd, $J = 15.0$, 11.5 Hz, 1H, H_{11}), 6.45 (d, $J = 15.0$ Hz, 1H, H_{12}), 6.35 (d, $J = 12.1$ Hz, 1H, H_{14}), 6.38–6.16 (m, 3H, $H_7 + H_8 + H_{10}$), 5.92 (d, $J = 15.0$ Hz, 1H, $H_{15'}$), 2.09 (s, 3H, CH$_3$), 2.07–1.97 (m, 2H, CH$_2$), 2.02 (s, 3H, CH$_3$), 1.71 (s, 3H, CH$_3$), 1.68–1.57 (m, 2H, CH$_2$), 1.51–1.46 (m, 2H, CH$_2$), 1.04 (s, 6H, $2 \times CH_3$) ppm. ^{13}C NMR (100 MHz, (CD$_3$)$_2$CO): 168.0 (s), 145.0 (s), 141.2 (d), 138.8 (s), 138.6 (s), 138.5 (d), 137.1 (d), 131.4 (d), 130.1 (s), 129.4 (d), 129.3 (d), 128.4 (d), 121.6 (d), 40.4 (t), 34.9 (s), 33.6 (t), 29.3 (q, 2×), 21.9 (q), 19.9 (t), 13.1 (q), 12.9 (q) ppm. HRMS (ESI$^+$): Calcd. for $C_{22}H_{31}O_2$ ([M+H]$^+$), 327.2319; found, 327.2316. IR (NaCl): ν 2924 (m, C–H), 2856 (w, C–H), 1680 (s, C=O), 1560 (m), 1162 (s), 969 (m) cm^{-1}. UV (MeOH): λ_{max} 375 ($\varepsilon = 46\,100$ mol^{-1} L cm^{-1}).

4.4. Methyl (2E,4E,6E,8E,10E,12E)-2,7,11-trimethyl-13-(2,6,6-trimethylcyclohex-1-en-1-yl)trideca-2,4,6,8,10,12-hexaenoate (methyl 12′-apo-β-carotenoate) (**18**)

Following the general procedure for the olefin metathesis, the reaction of (1,3,3-trimethyl-2-((1*E*,3*E*,5*E*,7*E*)-3,7-dimethyldeca-1,3,5,7,9-pentaenyl)cyclohex-1-ene **12** (21 mg, 0.074 mmol) with methyl (*E*)-2-methylpenta-2,4-dienoate **17** (37.5 mg, 0.298 mmol) and 2nd generation Hoveyda-Grubbs catalyst (7 mg, 0.011 mmol) in toluene (0.8 mL) at 25 °C for 7 h, afforded, after purification by column chromatography (silica gel, gradient from 100:0 to 99:1 hexane/EtOAc), in order of elution, 2.4 mg (11%) of a yellow oil

identified as 1,3,3-trimethyl-2-((1E,3E,5E,7E)-3,7-dimethyldeca-1,3,5,7,9-pentaenyl)cyclohex-1-ene **12**, 1.9 mg (10%) of a red solid identified as β-β-carotene **1** and 10.3 mg (40%) of a red solid identified as methyl (2E,4E,6E,8E,10E,12E)-2,7,11-trimethyl-13-(2,6,6-trimethylcyclohex-1-en-1-yl)trideca-2,4,6,8,10,12-hexaenoate **18**. ^{1}H NMR (400 MHz, (CD$_3$)$_2$CO): δ 7.34 (d, J = 11.8 Hz, 1H, H$_{14'}$), 7.10 (dd, J = 14.3, 12.0 Hz, 1H, H$_{15}$), 6.89 (dd, J = 14.9, 11.5 Hz, 1H, H$_{11}$), 6.69 (dd, J = 14.2, 12.1 Hz, 1H, H$_{15'}$), 6.46 (d, J = 15.0 Hz, 1H, H$_{12}$), 6.39 (d, J = 11.8 Hz, 1H, H$_{14}$), 6.34–6.16 (m, 3H, H$_7$ + H$_8$ + H$_{10}$), 3.74 (s, 3H, OCH$_3$), 2.09 (s, 3H, CH$_3$), 2.10–2.02 (m, 2H, CH$_2$), 2.03 (s, 3H, CH$_3$), 1.99 (s, 3H, CH$_3$), 1.75 (s, 3H, CH$_3$), 1.70–1.62 (m, 2H, CH$_2$), 1.54–1.48 (m, 2H, CH$_2$), 1.07 (s, 6H, 2 × CH$_3$) ppm. ^{13}C NMR (100 MHz, (CD$_3$)$_2$CO): δ 168.8 (s), 140.8 (s), 139.4 (d), 138.8 (d), 138.7 (s), 137.8 (d), 137.6 (s), 137.0 (d), 132.3 (d), 131.8 (d), 129.9 (s), 128.9 (d), 127.8 (d), 127.7 (d), 126.9 (s), 51.9 (q), 40.4 (t), 34.9 (s), 33.6 (t), 29.3 (q, 2×), 22.0 (q), 19.9 (t), 13.0 (q), 12.9 (q), 12.8 (q) ppm. HRMS (ESI$^+$): Calcd. for C$_{26}$H$_{37}$O$_2$ ([M+H]$^+$), 381.2788; found, 381.2786. IR (NaCl): ν 2926 (m, C–H), 2862 (w, C–H), 1703 (m, C=O), 1235 (m) cm^{-1}. UV (MeOH): λ$_{max}$ 410 nm.

4.5. (2E,4E,6E,8E,10E,12E)-2,7,11-trimethyl-13-(2,6,6-trimethylcyclohex-1-en-1-yl)trideca-2,4,6,8,10,12-hexaenal (12′-apo-β-carotenal) (5)

Following the general procedure for Dibal-H reduction of esters, the reaction of methyl (2E,4E,6E,8E,10E,12E)-2,7,11-trimethyl-13-(2,6,6-trimethylcyclohex-1-en-1-yl)trideca-2,4,6,8,10,12-hexaenoate **18** (15.7 mg, 0.041 mmol) with Dibal-H (0.124 mL, 0.124 mmol, 1 M in hexanes) in THF (0.209 mL) at −78 °C for 4 h, afforded, after purification by column chromatography (silica gel, 90:7:3 hexane/EtOAc/Et$_3$N) 10.9 mg (75%) of a reddish oil identified as (2E,4E,6E,8E,10E,12E)-2,7,11-trimethyl-13-(2,6,6-trimethylcyclohex-1-en-1-yl)trideca-2,4,6,8,10,12-hexaene-1-ol **19**, which was used without further purification.

Following the general procedure for MnO$_2$ oxidation of alcohols, the reaction of (2E,4E,6E,8E,10E,12E)-2,7,11-trimethyl-13-(2,6,6-trimethylcyclohex-1-en-1-yl)trideca-2,4,6,8,10,12-hexaene-1-ol **19** (8.4 mg, 0.0024 mmol) with MnO$_2$ (37.3 mg, 0.429 mmol) and Na$_2$CO$_3$ (45.4 mg, 0.429 mmol) in CH$_2$Cl$_2$ (0.443 mL) at 25 °C for 4 h, afforded, after purification by column chromatography (silica gel, 95:2:3 hexane/EtOAc/Et$_3$N) 3.2 mg (38%) of an orange-reddish oil identified as (2E,4E,6E,8E,10E,12E)-2,7,11-trimethyl-13-(2,6,6-trimethylcyclohex-1-en-1-yl)trideca-2,4,6,8,10,12-hexaenal **5**. ^{1}H NMR (400 MHz, (CD$_3$)$_2$CO): δ 9.46 (s, 1H, H$_{12'}$), 7.24 (dd, J = 14.1, 12.1 Hz, 1H, H$_{15}$), 7.13 (d, J = 11.8 Hz, 1H, H$_{14'}$), 6.92 (dd, J = 14.9, 11.5 Hz, 1H, H$_{11}$), 6.84 (dd, J = 14.1, 11.9 Hz, 1H, H$_{15'}$), 6.46 (d, J = 15.2 Hz, 1H, H$_{12}$), 6.42 (d, J = 12.2 Hz, 1H, H$_{14}$), 6.32–6.15 (m, 3H, H$_7$ + H$_8$ + H$_{10}$), 2.08 (s, 3H, CH$_3$), 2.09–2.01 (m, 2H, CH$_2$), 2.01 (s, 3H, CH$_3$), 1.83 (s, 3H, CH$_3$), 1.72 (s, 3H, CH$_3$), 1.71–1.55 (m, 2H, CH$_2$), 1.56–1.44 (m, 2H, CH$_2$), 1.04 (s, 6H, 2 × CH$_3$) ppm. ^{13}C NMR (100 MHz, (CD$_3$)$_2$CO): δ 194.4 (d), 149.3 (d), 142.3 (s), 138.7 (d), 138.6 (s), 138.5 (d), 138.2 (s), 137.6 (d), 137.5 (s), 132.1 (d), 131.7 (d), 130.0 (s), 128.6 (d), 128.4 (d), 128.1 (d), 40.3 (t), 34.9 (s), 33.6 (t), 29.3 (q, 2×), 22.0 (q), 19.9 (t), 13.0 (q), 12.8 (q), 9.5 (q) ppm. HRMS (ESI$^+$): Calcd. for C$_{25}$H$_{35}$O 351.2682; found, 351.2695. IR (NaCl): ν 2924 (m, C–H), 2861 (w, C–H), 1667 (s, C=O), 1544 (m), 1185 (m) cm^{-1}. UV (MeOH): λ$_{max}$ 424 nm (ε = 69 600 mol^{-1} L cm^{-1}).

4.6. (2E,4E,6E,8E,10E,12E)-2,7,11-trimethyl-13-(2,6,6-trimethylcyclohex-1-en-1-yl)trideca-2,4,6,8,10,12-hexaenoic acid (12′-apo-β-carotenoic acid) (9)

Following the general procedure for hydrolysis of esters, the reaction of (2E,4E,6E,8E,10E,12E)-2,7,11-trimethyl-13-(2,6,6-trimethylcyclohex-1-en-1-yl)trideca-2,4,6,8,10,12-hexaenoate **18** (10.1 mg, 0.026 mmol) with KOH (0.425 mL, 2 M in H$_2$O, 0.85 mmol) in MeOH (1.83 mL) at 70 °C for 1 h, afforded, after purification by column chromatography (silica gel, 70:30 hexane/EtOAc) 5.1 mg (53%) of a red solid identified as (2E,4E,6E,8E,10E, 12E)-2,7,11-trimethyl-13-(2,6,6-trimethylcyclohex-1-en-1-yl)trideca-2,4,6,8,10,12-hexaenoic acid **9**. ^{1}H NMR (400 MHz, CDCl$_3$): δ 7.42 (d, J = 11.8 Hz, 1H, H$_{14'}$), 6.95 (dd, J = 14.0, 12.1 Hz, 1H, H$_{15}$), 6.76 (dd, J = 14.9, 11.5 Hz, 1H, H$_{11}$), 6.54 (dd, J = 13.8, 12.1 Hz, 1H, H$_{15'}$), 6.36 (d, J = 15.0 Hz, 1H, H$_{12}$), 6.27 (d, J = 12.4 Hz, 1H, H$_{14}$), 6.23–6.10 (m, 3H, H$_7$ + H$_8$ + H$_{10}$), 2.02 (s, 3H, CH$_3$), 1.99 (s, 3H, CH$_3$), 2.10–1.99 (m, 2H, CH$_2$), 1.72 (s, 3H, CH$_3$), 1.66–1.57 (m, 2H, CH$_2$), 1.51–1.47 (m, 2H, CH$_2$), 1.03 (s, 6H, 2 × CH$_3$) ppm. ^{13}C NMR (100 MHz, CDCl$_3$): δ 173.1 (s), 141.0 (d), 140.8 (s), 138.0 (s), 137.7 (d), 137.6 (s), 137.1 (d), 136.6 (d), 131.1 (d), 130.6 (d), 129.8 (s), 127.8 (d), 127.7 (d), 127.2 (d), 125.2 (s), 39.8 (t), 34.4 (s), 33.3 (t), 29.2 (q, 2×), 21.9 (q), 19.4 (t), 13.1 (q), 13.0 (q), 12.6 (q) ppm. HRMS (ESI$^+$): Calcd. for C$_{25}$H$_{35}$O$_2$ ([M+H]$^+$), 367.2632; found, 367.2629. IR (NaCl): ν 2924 (m, C–H), 2858 (w, C–H), 1670 (s, C=O), 1419 (m), 1248 (m), 968 (m) cm^{-1}. UV (MeOH): λ$_{max}$ 402 (ε = 61 450 mol^{-1} L cm^{-1}).

4.7. Ethyl (2E,4E)-5-(1,3-dioxolan-2-yl)-3methylpenta-2,4-dienoate (20a)

To a 500 mL round-bottomed flask armed with a Dean-Stark trap, were added ethyl (2E,4E) 3-methyl-6-oxohexa-2,4-dienoate **20** (2.5 g, 14.8 mmol), benzene (256.3 mL), p-TsOH (170 mg, 0.89 mmol) and ethyleneglycol (8.7 mL, 156 mmol), and the resulting mixture was heated under reflux for 16 h. The cooled reaction mixture was diluted with H$_2$O, an aqueous solution of NaHCO$_3$ was added and the mixture was extracted with Et$_2$O (3×). The combined organic layers were dried (Na$_2$SO$_4$) and the solvent was removed to obtain 3.12 g (99%) of a colorless oil identified as ethyl (2E,4E)-5-(1,3-dioxolan-2-yl)-3-methylpenta-2,4-dienoate **20a**. ^{1}H NMR (400 MHz, C$_6$D$_6$): δ 6.29 (d, J = 15.8 Hz, 1H, H$_4$), 5.96 (dd, J = 15.8, 5.4 Hz, 1H, H$_5$), 5.82 (s, 1H, H$_2$), 5.23 (d, J = 5.4 Hz, 1H, H$_6$), 3.99 (q, J = 7.1 Hz, 2H, OCH$_2$CH$_3$), 3.53–3.45 (m, 2H, CH$_2$), 3.43–3.36 (m, 2H, CH$_2$), 2.23 (s, 3H, CH$_3$), 0.96 (t, J = 7.1 Hz, 3H, OCH$_2$CH$_3$) ppm. ^{13}C NMR (101 MHz, C$_6$D$_6$): δ 166.3 (s), 150.8 (s), 137.2 (d), 131.8 (d), 121.7 (d), 103.4 (d), 65.0 (t, 2×), 59.8 (t), 14.3 (q), 13.7 (q) ppm. HRMS (ESI$^+$): Calcd. for C$_{11}$H$_{17}$O$_4$ ([M+H]$^+$), 213.1121; found, 213.1120. IR (NaCl): ν 2979 (m, C–H), 2887 (m, C–H), 1713 (s, C=O), 1616 (m), 1231 (m), 1154 (s), 772 (m) cm^{-1}. UV (MeOH): λ$_{max}$ 253 nm.

4.8. (2E,4E)-5-(1,3-dioxolan-2-yl)-3-methylpenta-2,4-dien-1-ol (20b)

Following the general procedure for Dibal-H reduction of esters, the reaction of ethyl (2E,4E)-5-(1,3-dioxolan-2-yl)-3-methylpenta-2,4-dienoate 20a (1.39 g, 6.59 mmol) with Dibal-H (16.5 mL, 1 M in hexanes, 16.5 mmol) in THF (33.4 mL) at −78 °C for 2 h, afforded 1.11 g (99%) of a colorless oil identified as (2E,4E)-5-(1,3-dioxolan-2-yl)-3-methylpenta-2,4-dien-1-ol **20b**. ^{1}H NMR (400 MHz, C$_6$D$_6$): δ 6.43 (d, J = 15.8 Hz, 1H, H$_4$), 5.77 (dd, J = 15.8, 5.9 Hz, 1H, H$_5$), 5.54 (t, J = 5.4 Hz, 1H, H$_2$), 5.32 (d, J = 5.9 Hz, 1H, H$_6$), 3.90 (d, J = 5.4 Hz, 2H, 2H$_1$), 3.65–3.54 (m, 2H, CH$_2$), 3.47–3.41 (m, 2H, CH$_2$), 1.47 (s, 3H, CH$_3$) ppm. ^{13}C NMR (101 MHz, C$_6$D$_6$): δ 138.8 (d), 134.1 (s), 134.0 (d), 125.3 (d), 104.4 (d), 65.0 (t, 2×), 59.2 (t), 12.4 (q) ppm. HRMS (ESI$^+$): Calcd. for C$_9$H$_{15}$O$_3$ ([M+H]$^+$), 171.1016; found, 171.1020. IR (NaCl): ν 3500-3100 (br, O–H), 2950 (m, C–H), 2883 (m, C–H), 1386 (m), 1156 (m), 1085 (m), 1014 (m), 950 (s) cm^{-1}. UV (MeOH): λ$_{max}$ 233 nm.

4.9. Methyl (2E,4E,6E,8E,10E,12E,14E)-4,9,13-trimethyl-15-(2,6,6-trimethylcyclohex-1-en-1-yl)pentadeca-2,4,6,8,10,12,14-heptaenoate (methyl 10′-β-apo-carotenoate) (25)

Following the general procedure for olefin metathesis, the reaction of (1,3,3-trimethyl-2-((1E,3E,5E,7E)-3,7-dimethyldeca-1,3,5,7,9-pentaenyl)cyclohex-1-ene **12** (10 mg, 0.035 mmol), methyl (2E,4E,6E) and (2E,4E,6Z)-4-methylocta-2,4,6-trienoate **24** (35.3 mg, 0.212 mmol) and 2nd generation Hoveyda-Grubbs catalyst (4.4 mg, 0.007 mmol) in toluene (0.18 mL) at 25 °C for 5 h, afforded, after purification by column chromatography (silica gel, 98:2 hexane/EtOAc) in order of elution, 5.5 mg (58%) of a red solid identified as β,β-carotene **1**, 17.2 mg (49%) of a yellowish oil identified as starting methyl (2E,4E,6E) and (2E,4E,6Z)-4-methylocta-2,4,6-trienoate **24** and 3.4 mg (24%) of a red oil identified as methyl (2E,4E,6E,8E,10E,12E,14E)-4,9,13-trimethyl-15-(2,6,6-trimethylcyclohex-1-en-1-yl)pentadeca-2,4,6,8,10,12,14-heptaenoate **25**. ^{1}H NMR (400.16 MHz, (CD$_3$)$_2$CO): δ 7.35 (d, $J = 15.5$ Hz, 1H, H$_{12'}$), 6.96 (dd, $J = 13.9$, 11.9 Hz, 1H, H$_{15}$), 6.84 (dd, $J = 14.9$, 11.5 Hz, 1H, H$_{11}$), 6.80—6.72 (m, 1H, H$_{15'}$), 6.67 (d, $J = 11.8$ Hz, 1H, H$_{14'}$), 6.43 (d, $J = 15.0$ Hz, 1H, H$_{12}$), 6.37 (d, $J = 11.8$ Hz, 1H, H$_{14}$), 6.26 (d, $J = 15.8$ Hz, 1H, H$_7$), 6.23 (d, $J = 11.0$ Hz, 1H, H$_{10}$), 6.17 (d, $J = 16.1$ Hz, 1H, H$_8$), 5.90 (d, $J = 15.5$ Hz, 1H, H$_{11'}$), 3.69 (s, 3H, OCH$_3$), 2.07—2.05 (m, 2H, CH$_2$), 2.03 (s, 3H, CH$_3$), 2.00 (s, 3H, CH$_3$), 1.96 (s, 3H, CH$_3$), 1.71 (s, 3H, CH$_3$), 1.68—1.57 (m, 2H, CH$_2$), 1.51—1.45 (m, 2H, CH$_2$), 1.03 (s, 6H, 2 × CH$_3$) ppm. ^{13}C NMR (100.62 MHz, (CD$_3$)$_2$CO): δ 167.8 (s), 149.3 (d), 140.3 (d), 139.8 (s), 138.8 (d), 138.7 (s), 137.9 (d), 137.4 (s), 135.0 (d), 134.2 (s), 132.8 (d), 131.9 (d), 129.9 (d), 129.9 (s), 127.6 (d), 127.3 (d), 116.8 (d), 51.5 (q), 40.3 (t), 34.9 (s), 33.6 (t), 29.3 (q, 2×), 22.0 (q), 19.9 (t), 12.9 (q), 12.8 (q), 12.6 (q) ppm. HRMS (ESI$^+$): Calcd. for C$_{28}$H$_{39}$O$_2$ ([M+H]$^+$), 407.2945; found, 407.2954. IR (NaCl): ν 2921 (m, C—H), 2856 (m, C—H), 1714 (m, C=O), 1701 (m), 1539 (m) cm^{-1}. UV (MeOH): 429 nm.

4.10. (2E,4E,6E,8E,10E,12E,14E)-4,9,13-trimethyl-15-(2,6,6-trimethylcyclohex-1-en-1-yl)pentadeca-2,4,6,8,10,12,14-heptaenal (10′-apo-β-carotenal) (6)

Following the general procedure for Dibal-H reduction of esters, the reaction of methyl (2E,4E,6E,8E,10E,12E,14E)-4,9,13-trimethyl-15-(2,6,6-trimethylcyclohex-1-en-1-yl)pentadeca-2,4,6,8,10,12,14-heptaenoate **25** (11.2 mg, 0.028 mmol) with Dibal-H (6.9 μL, 0.069 mmol, 1 M in hexanes) in THF (0.138 mL) at −78 °C for 2 h, afforded, after purification by column chromatography (silica gel, gradient from 85:12:3 to 80:20:0 hexane/EtOAc/Et$_3$N) 10 mg of a red solid identified as (2E,4E,6E,8E,10E,12E,14E)-4,9,13-trimethyl-15-(2,6,6-trimethylcyclohex-1-en-1-yl)pentadeca-2,4,6,8,10,12,14-heptaen-1-ol **26** which was used without further purification.

Following the general procedure for MnO$_2$ oxidation of alcohols, the reaction of (2E,4E,6E,8E,10E,12E,14E)-4,9,13-trimethyl-15-(2,6,6-trimethylcyclohex-1-en-1-yl)pentadeca-2,4,6,8,10,12,14-heptaen-1-ol **26** (10.7 mg, 0.028 mmol) with MnO$_2$ (52.2 mg, 0.51 mmol) and Na$_2$CO$_3$ (54.1 mg, 0.51 mmol) in THF (0.423 mL) at 25 °C for 2 h afforded, after purification by column chromatography (silica gel, gradient from 98:0:2 to 97:3:0 hexane/EtOAc/Et$_3$N) 6.1 mg (58% combined yield) of a red solid identified as (2E,4E,6E,8E,10E,12E,14E)-4,9,13-trimethyl-15-(2,6,6-trimethylcyclohex-1-en-1-yl)pentadeca-2,4,6,8,10,12,14-heptaenal **6** which was purified by HPLC (Waters Spherisorb™ 10 μm CN, 10 × 250 mm Semipreparative column with NovaPak CN 4 μm precolumn, 95:5 hexane/acetone, flow rate: 3 mL/min, t$_R$ = 12 min). ^{1}H NMR (400.16 MHz, (CD$_3$)$_2$CO): δ 9.59 (d, $J = 7.6$ Hz, 1H, H$_{10'}$), 7.33 (d, $J = 15.3$ Hz, 1H, H$_{12'}$), 7.08—6.96 (m, 1H, H$_{15}$ or H$_{15'}$), 6.86 (dd, $J = 15.0$, 11.4 Hz, 1H, H$_{11}$), 6.81—6.74 (m, 2H, H$_{14'}$ or H$_{14'}$ + H$_{15}$ or

H$_{15'}$), 6.44 (d, $J = 15.0$ Hz, 1H, H$_{12}$), 6.39 (d, $J = 12.0$ Hz, 1H, H$_{14}$ or H$_{14'}$), 6.30—6.11 (m, 4H, H$_7$ + H$_8$ + H$_{10}$ + H$_{11'}$), 2.07—2.04 (m, 2H, CH$_2$), 2.04 (s, 6H, 2 × CH$_3$), 2.00 (s, 3H, CH$_3$), 1.71 (s, 3H, CH$_3$), 1.66—1.59 (m, 2H, CH$_2$), 1.50—1.46 (m, 2H, CH$_2$), 1.03 (s, 6H, 2× CH$_3$) ppm. ^{13}C NMR (100.62 MHz, (CD$_3$)$_2$CO): δ 193.7 (d), 156.9 (d), 141.7 (d), 140.7 (s), 138.7 (d), 138.7 (s), 137.9 (d), 137.6 (s), 136.1 (d), 134.7 (s), 132.6 (d), 131.9 (d), 129.9 (s), 129.9 (d), 128.1 (d), 127.8 (d), 127.6 (d), 40.3 (t), 34.9 (s), 33.6 (t), 29.3 (q, 2×), 22.0 (q), 19.9 (t), 13.0 (q), 12.8 (q), 12.70 (q) ppm. HRMS (ESI$^+$): Calcd. for C$_{27}$H$_{37}$O ([M+H]$^+$), 377.2839; found, 377.2842. IR (NaCl): ν 2920 (m, C—H), 2852 (m, C—H), 1671 (s, C=O), 1123 (s), 968 (m) cm^{-1}. UV (MeOH): 450 nm (ε 66 400 mol^{-1} L cm^{-1}).

4.11. Purification of recombinant human ALDHs and enzymatic assay

Human ALDH1A1, ALDH1A2 and ALDH1A3 were recombinantly expressed from the pET-30 Xa/LIC vector and affinity purified onto a Ni^{2+}-NTA Chelating Sepharose™ Fast Flow column (GE Heathcare). Activity assays with apo-β-carotenals were carried out using detergent-free solubilization and end-point reaction, followed by HPLC, originally devised for retinoid analysis.[43] ALDH1A1 and ALDH1A2 were assayed in 50 mM HEPES, 0.5 mM EDTA, 0.5 mM DTT, pH 8.0, while ALDH1A3 was assayed in 50 mM HEPES, 30 mM MgCl$_2$, 5 mM DTT, pH 8.0. Concentration of 12′- and 14′-apo-β-carotenal was determined based on the corresponding molar absorption coefficient in aqueous solutions at the appropriate wavelength ($\varepsilon_{410} = 24\,228$ M^{-1} cm^{-1} and $\varepsilon_{416} = 15\,945$ M^{-1} cm^{-1}, for 12′-apo-β-carotenal **5** in ALDH1A1/1A2 and ALDH1A3 reaction buffer, respectively; and $\varepsilon_{397} = 9218$ M^{-1} cm^{-1} and $\varepsilon_{340} = 9036$ M^{-1} cm^{-1} for 14′-apo-β-carotenal **4** in ALDH1A1/1A2 and ALDH1A3 reaction buffer, respectively). The reaction was started by the addition of cofactor and carried out for 15 min at 37 °C in a final volume of 0.5 mL. With the aim to measure the steady state enzymatic activity, the concentration of enzyme was kept from 50- to 100-fold lower than that of the substrate for all enzymatic assays and a saturating concentration of cofactor (0.5 mM NAD$^+$) was used. Reaction products were extracted with hexane/dioxane/iso-propanol (50:5:1, v/v) and analyzed by an HPLC-based method.[44] The organic layer was evaporated, apo-β-carotenoids were dissolved in hexane and injected onto a NovaPak® silica gel column (4 μm, 3.9 × 150 mm, Waters) in hexane:*tert*-butyl methyl ether (96:4, v/v) mobile phase, at a flow rate of 2 mL/min using a Waters Alliance 2695 HPLC instrument. Elution was monitored at 415 nm for 12′-apo-β-carotenal **5**, 400 nm for 12′-apo-β-carotenoic acid **9** and 14′-apo-β-carotenal **4**, and 373 nm for 14′-apo-β-carotenoic acid **8**, using a Waters 2996 photodiode array detector. All compound manipulations were performed under dim or red light to prevent photoisomerization.

Operational Program (TAMOP) of Hungary.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.tet.2018.03.050.

References

1. (a) . In: Britton G, Liaaen-Jensen S, Pfander HE, eds. *Carotenoids. Part 1A. Isolation and Analysis*. Basel: Birkhäuser; 1995;
   (b) . In: Britton G, Liaaen-Jensen S, Pfander H, eds. *Carotenoids. Part 1B. Spectroscopy*. Basel: Birkhäuser; 1995;
   (c) . In: Landrum T, ed. *Carotenoids. Physical, Chemical and Biological Functions and Properties*. vol. 1. Boca Raton, FL: CRC Press; 2010.
2. Britton G, Liaaen-Jensen S, Pfander H, eds. *Carotenoids*. Basel: Birkhäuser; 2008: . Natural Functions; vol. 4. Basel: Birkhäuser; 2008.
3. Britton G, Liaaen-Jensen S, Pfander H, eds. *Carotenoids*. Basel: Birkhäuser; 2009: . Nutrition and Health; vol. 5. Basel: Birkhäuser; 2009.
4. Alvarez R, Vaz B, Gronemeyer H, de Lera AR. *Chem Rev*. 2014;114:1–125.
5. Britton G, Liaaen-Jensen S, Pfander H, eds. *Carotenoids*. Basel: Birkhäuser; 1998: . Biosynthesis and Metabolism; vol. 3. Basel: Birkhäuser; 1998.
6. (a) von Lintig J, Vogt K. *J Biol Chem*. 2000;275:11915–11920;
   (b) Wyss A, Wirtz G, Woggon WD, et al. *Biochem Biophys Res Commun*. 2000;271:334–336;
   (c) Kiefer C, Hessel S, Lampert JM, et al. *J Biol Chem*. 2001;276:14110–14116;
   (d) dela Seña C, Riedl KM, Narayanasamy S, Curley RW, Schwartz SJ, Harrison EH. *J Biol Chem*. 2014;289:13661–13666.
7. Lobo GP, Amengual J, Palczewski G, Babino D, von Lintig J. *Biochim Biophys Acta Mol Cell Biol Lipids*. 2012;1821:78–87.
8. (a) Hu KQ, Liu C, Ernst H, Krinsky NI, Russell RM, Wang XD. *J Biol Chem*. 2006;281:19327–19338. Amengual J, Widjaja-Adhi MAK, Rodriguez-Santiago S, Hessel S, Golczak M, Palczewski K, von Lintig J. J. Biol. Chem. 2013; 288: 34081-34096;
   (b) Sui X, Kiser PD, von Lintig J, Palczewski K. *Arch Biochem Biophys*. 2013;539:203–213;
   (c) Harrison PJ, Bugg TDH. *Arch Biochem Biophys*. 2014;544:105–111.
9. Amengual J, Lobo GP, Golczak M, et al. *Faseb J*. 2011;25:948–959.
10. Palczewski G, Amengual J, Hoppel CL, von Lintig J. *Faseb J*. 2014;28:4457–4469.
11. (a) Lindqvist A, Andersson S. *J Biol Chem*. 2002;277:23942–23948;
    (b) Kowatz T, Babino D, Kiser P, Palczewski K, von Lintig J. *Arch Biochem Biophys*. 2013;539:214–222.
12. (a) Parés X, Farrés J, Kedishvili N, Duester G. *Cell Mol Life Sci*. 2008;65:3936–3949;
    (b) Ruiz FX, Porté S, Pares X, Farrés J. *Front Pharmacol*. 2012;3:58.
13. Germain P, Chambon P, Eichele G, et al. *Pharmacol Rev*. 2006;58:712–725.
14. (a) Mangelsdorf DJ, Thummel C, Beato M, et al. *Cell*. 1995;83:835–839;
    (b) Renaud JP, Moras D. *Cell Mol Life Sci*. 2000;57:1748–1769;
    (c) Gronemeyer H, Gustafsson JA, Laudet V. *Nat Rev Drug Discov*. 2004;3:950–964.
15. Eroglu A, Harrison EH. *J Lipid Res*. 2013;54:1719–1730.
16. Eroglu A, Hruszkewycz DP, Curley Jr RW, Harrison EH. *Arch Biochem Biophys*. 2010;504:11–16.
17. Eroglu A, Hruszkewycz DP, dela Sena C, et al. *J Biol Chem*. 2012;287:15886–15895.
18. Costabile BK, Kim YK, Iqbal J, et al. *J Biol Chem*. 2016;291:18525–18535.
19. (a) Wang X, Penzes P, Napoli JL. *J Biol Chem*. 1996;271:16288–16293;
    (b) Zhao D, McCaffery P, Ivins KJ, et al. *Eur J Biochem*. 1996;240:15–22;
    (c) Yoshida A, Rzhetsky A, Hsu LC, Chang C. *Eur J Biochem*. 1998;251:549–557;
    (d) Niederreither K, Subbarayan V, Dolle P, Chambon P. *Nat Genet*. 1999;21:444–448;
    (e) Grün F, Hirose Y, Kawauchi S, Ogura T, Umesono K. *J Biol Chem*. 2000;275:41210–41218.
20. Britton G, Liaaen-Jensen S, Pfander H, eds. *Carotenoids. Part 2. Synthesis*. Basel: Birkhäuser; 1996.
21. (a) Bernhard K, Mayer H. *Pure Appl Chem*. 1991;63:35–44;
    (b) Nicolaou KC, Härter MW, Gunzner JL, Nadin A. *Liebigs Ann Chem*. 1997:1283–1301;
    (c) Yamano Y, Ito M. *J. Chem. Soc. Perkin Trans*. 1993;1:1599–1610;
    (d) Yamano Y, Tode C, Ito M. *J. Chem. Soc., Perkin Trans*. 1995;1:1895–1904;
    (e) Tode C, Yamano Y, Ito M. *J. Chem. Soc., Perkin Trans*. 1999;1:1625–1626;
    (f) Yamano Y, Yoshizawa M, Ito M. *J Nutr Sci Vitaminol*. 1999;45:49–62;
    (g) Furuichi N, Hara H, Osaki T, Mori H, Katsumura S. *Angew Chem Int Ed*. 2002;41:1023–1026;
    (h) Kajikawa T, Okumura S, Iwashita T, Kosumi D, Hashimoto H, Katsumura S. *Org Lett*. 2012;14:808–811.
22. Nicolaou KC, Bulger PG, Sarlah D. *Angew Chem Int Ed*. 2005;44:4442–4489.
23. (a) Olpp T, Brückner R. *Angew Chem Int Ed*. 2006;45:4023–4027;
    (b) Vaz B, Domínguez M, Álvarez R, de Lera AR. *Chem Eur J*. 2007;13:1273–1290;
    (c) Burghart J, Brückner R. *Angew Chem Int Ed*. 2008;47:7664–7668;
    (d) Woerly EM, Cherney AH, Davis EK, Burke MD. *J Am Chem Soc*. 2010;132:6941–6943;
    (e) Fujii S, Chang SY, Burke MD. *Angew Chem Int Ed*. 2011;50:7862–7864;
    (f) Fontan N, Vaz B, Alvarez R, de Lera AR. *Chem Commun*. 2013;49:2694–2696;
    (g) Otero L, Vaz B, Alvarez R, de Lera AR. *Chem Commun*. 2013;49:5043–5045;
    (h) Vaz B, Otero L, Alvarez R, de Lera AR. *Chem Eur J*. 2013;19:13065–13074;
    (i) Woerly EM, Roy J, Burke MD. *Nat Chem*. 2014;6:484–491.
24. (a) Nicolaou KC, Bulger PG, Sarlah D. *Angew Chem Int Ed*. 2005;44:4490–4527;
    (b) Hoveyda AH, Zhugralin AR. *Nature*. 2007;450:243–251;
    (c) Fürstner A. *Chem Commun*. 2011;47:6505–6511.
25. Fürstner A. *Science*. 2013;341:1229713.
26. Vougioukalakis GC, Grubbs RH. *Chem Rev*. 2010;110:1746–1787.
27. Maj J, Morzycki JW, Rárová L, Wasilewski G, Wojtkielewicz A. *Tetrahedron Lett*. 2012;53:5430–5433.
28. Jermacz I, Maj J, Morzycki JW, Wojtkielewicz A. *Toxicol Mech Meth*. 2008;18:469–471.
29. (a) Wojtkielewicz A, Maj J, Morzycki JW. *Tetrahedron Lett*. 2009;50:4734–4737;
    (b) Wojtkielewicz A, Maj J, Dzieszkowska A, Morzycki JW. *Tetrahedron*. 2011;67:6868–6875.
30. Fontán N, Domínguez M, Álvarez R, de Lera AR. *Eur J Org Chem*. 2011:6704–6712.
31. Fontán N, Alvarez R, de Lera AR. *J Nat Prod*. 2012;75:975–979.
32. Kajikawa T, Iguchi N, Katsumura S. *Org Biomol Chem*. 2009;7:4586–4589.
33. Chatterjee AK, Choi TL, Sanders DP, Grubbs RH. *J Am Chem Soc*. 2003;125:11360–11370.
34. Al-Hasani SMA, Parrish DB. *J Agric Food Chem*. 1967;15:943–944.
35. Fürstner A, Guth O, Düffels A, et al. *Chem Eur J*. 2001;7:4811–4820.
36. (a) Grela K, Harutyunyan S, Michrowska A. *Angew Chem Int Ed*. 2002;41:4038–4040;
    (b) Michrowska A, Bujok R, Harutyunyan S, Sashuk V, Dolgonos G, Grela K. *J Am Chem Soc*. 2004;126:9318–9325;
    (c) Samojtowicz C, Bieniek M, Grela K. *Chem Rev*. 2009;109:3708–3742.
37. (a) Scholl M, Ding S, Lee CW, Grubbs RH. *Org Lett*. 1999;1:953–956;
    (b) Nguyen ST, Grubbs RH, Ziller JW. *J Am Chem Soc*. 1993;115:9858–9859.
38. Garber SB, Kingsbury JS, Gray BL, Hoveyda AH. *J Am Chem Soc*. 2000;122:8168–8179.
39. (a) Estrada AF, Youssar L, Scherzinger D, Al-Babili S, Avalos J. *Mol Microbiol*. 2008;69:1207–1220;
    (b) Díaz-Sánchez V, Estrada AF, Trautmann D, Al-Babili S, Avalos J. *FEBS J*. 2011;278:3164–3176;
    (c) Trautmann D, Beyer P, Al-Babili S. *FEBS J*. 2013;280:3685–3696.
40. Veera Reddy P, Rabago-Smith M, Borhan B. *J. Labelled Cpd Radiopharm*. 2002;45:79–89.
41. (a) Bernhard K, Englert G, Mayer H, et al. *Helv Chim Acta*. 1981;64:2469–2484;
    (b) Haugan JA, Liaaen-Jensen S. *Acta Chem Scand*. 1994;48:899–904.
42. Rüegg R, Lindlar H, Montavon M, et al. *Helv Chim Acta*. 1959;42:847–853.
43. Gallego O, Belyaeva OV, Porté S, et al. *Biochem J*. 2006;399:101–109.
44. Kane MA, Chen N, Sparks S, Napoli JL. *Biochem J*. 2005;388:363–369.