Transworld Research Network
37/661 (2), Fort P.O.
Trivandrum-695 023
Kerala, India

Recent Advances in Pharmaceutical Sciences, 2011: 247-273 ISBN: 978-81-7895-528-5
Editor: Diego Muñoz-Torrero

11. Crystallization of the lipoxygenase of *Pseudomonas aeruginosa* 42A2, evolution and phylogenetic study of the subfamilies of the lipoxygenases

Albert Garreta[1,4], Xavier Carpena[2], Montse Busquets[3]
Carmen Fusté[1], Ignacio Fita[4] and Àngels Manresa[1]

[1]Departament de Microbiologia i Parasitologia Sanitàries, Universitat de Barcelona, Barcelona, Spain [2]Institute for Research in Biomedicine (IRB-Barcelona) and Institut Químic de Sarrià, IQS-URL Barcelona, Spain; [3]Departament de Bioquimica i Biologia Molecular. Universitat de Barcelona Barcelona. Spain; [4]Institute of Research in Biomedicine (IRB-Barcelona) and Institut de Biologia Molecular (IBMB-CSIC), Parc Científic, Barcelona, Spain

Abstract. Lipoxygenases are non-heme iron enzymes essential in eukaryotes, where they catalyze the formation of the fatty acid hydroperoxides that are required by a large diversity of biological and pathological processes. In prokaryotes, most of them totally lacking in polyunsaturated fatty acids, the possible biological roles of lipoxygenases have remained obscure. In this study, it is reported the crystallization of a lipoxygenase of *Pseudomonas aeruginosa* (Pa_LOX), the first from a prokaryote. High resolution data has been acquired which is expected to yield structural clues to the questions adressed. Besides, a preliminar phylogenetic analysis

Correspondence/Reprint request: Dr. Albert Garreta, Institute of Research in Biomedicine (IRB-Barcelona) and Institut de Biologia Molecular (IBMB-CSIC), Parc Científic, Barcelona, Spain
E-mail: albertgarreta@hotmail.com

using 14 sequences has confirmed the existence of this subfamily of bacterial lipoxygenases, on one side, and a greater diversity than in the corresponding eukaryotic ones, on the other. Finally, an evolutionary study of bacterial lipoxygenases on the same set of lipoxygenases, show a selection pressure of a basically purifying or neutral character except for a single aminoacid, which would have been selected after a positive selection event.

Introduction

The oxygenation of polyunsaturated fatty acids mediated by lipoxygenase (LOX) enzymes is the initial step in the formation of a diverse array of lipid mediators [1–4]. LOX products in animals commonly act by signaling through G-protein-coupled receptors. Different eicosanoids appear to be involved in cell homeostasis, proliferation and differentiation, as well as in pathophysiological processes (inflammation and cancer). Generally, plant LOX pathway products are involved in the resistance to environmental stress and defense against microbe and herbivore attack. Some of the volatile products, like jasmonic acid and short-chain aldehydes, have a function in plant-plant communication. Arachidonic acid (20.4ω6) is the prototypical substrate in animals; and linoleic or linolenic acids (18.2ω6, 18.3ω3), in plants. The initial products are fatty acid peroxides that give rise to local hormones classified collectively as ‘‘eicosanoids’’ (20-carbon lipids) or by the more general term ‘‘oxylipins’’ [5].

LOX described in plants and animals are non-heme iron enzymes consisting of a single polypeptide chain that is folded into two domains, the mostly α-helical catalytic domain and the N-terminal β barrel domain that is involved in membrane binding [6–12]. In humans there are five LOX enzymes that act on arachidonic acid and catalyze four distinct oxygenations (5S,12R, 12S or 15S), the products being individual fatty acid hydroperoxides [1]. There are another five distinct reactions (5R, 8R, 8S, 9S and 11R) in other animal species [13–15]. A large number of plant LOX enzymes use linoleic or linolenic acid as substrate, many of which are quite specific to these fatty acids [5,16].

The primary structure of LOX genes is highly diverse, with only three of the five amino acids that serve as iron ligands totally conserved within the LOX gene family. There is considerable variation in the polypeptide chain, perhaps not surprisingly in view of the wide spectrum of specific oxygenations that can be catalyzed. Even so, as each specific oxygenation mediated by a LOX involves a 1,4-Z,Z- pentadiene group (a pair of conjugated double bonds) on the fatty acid substrate and its coupling with O_2, there may be a common mechanism underlying the different LOX reaction specificities.

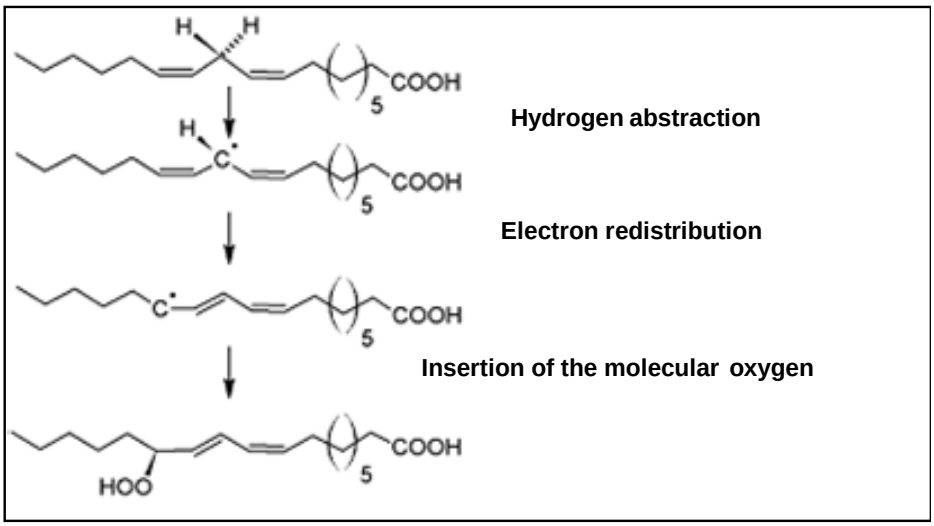

Figure 1. The three consecutive steps of the mechanism of action of lipoxygenases.

The mechanism of LOX (Fig. 1) consists of three consecutive steps. The first is the stereoselective abstraction of one hydrogen of the group 1,4-*cis*,*cis*-pentadiene to form a pentadienyl radical. The second step is the electron redistribution modifying the *cis*-*cis* pentadiene group to a *cis*-*trans* pentadiene group. The third and last step consists of the insertion of the molecular oxygen and the consequent reduction of the hydroperoxide radical.

One of the most interesting aspects is that molecular oxygen is not covalently bound in the enzyme prior to reaction with the substrate. Since polyunsaturated fatty acids, which are the substrate of the reaction, are achiral, a question arises: how do the LOX oxygenate the substrate with such regio- and stereo-selectivity? When trying to describe possible mechanisms for the control of oxygenation in the enzymatic reaction, there are four possible explanations relevant to the LOX [18].

Steric shielding

The access of the oxygen to the reactive carbon centers of the pentadienyl radical is not free. The structure of the protein restricts the access to one of these reactive carbons.

Oxygen channeling

The molecular oxygen is directed to the desired site by a channel that starts at the surface of the protein and finishes close to the activated fatty acid substrate, thus controlling position and stereospecificity.

Selective peroxyl radical trapping

This hypothesis entails that the oxygenation specificity of the LOX depends on the trapping of the peroxyl radical resulting in the final product. The molecular oxygen reacts with many positions of the pentadiene of the activated fatty acid, but a strategically placed hydrogen donor traps only the desired product by reduction to hydroperoxide.

Radical location

During oxidation of fatty acids, the pentadienyl radical with the unpaired electron delocalized over all 5 carbons adopts a planar structure. This permits the oxygen to react in multiple positions and in R or S configuration. The unpaired electron could be localized if a force is applied and the radical is twisted, with this location of the electron giving way to selective reaction with O_2 at this position.

1. Molecular structure of lipoxygenases

1.1. General structure

Despite more than 60 years of study of lipoxygenases, it was not until 1993 that the first lipoxygenase was crystallized and diffracted [6], soy bean "arachidonic acid 15-lipoxygenase". Later, in 1996, the structure was refined up to 1.4 Å [4]. More recently, other soy bean lipoxygenases have been crystallized, such as 3-lipoxygenase and various complexes. The first crystal structure of an animal lipoxygenase was published in 1997, the rabbit reticulocyte lipoxygenase, a 15-lipoxygenase [8]. The most recent lipoxygenase crystallized was the lipoxygenase of *Plexaura homomolla*, a coral, in 2005 [12].

Lipoxygenases have a single polypeptide chain (monomeric proteins) with a molecular mass of approximately 75-80 kDa for animal lipoxygenases between 660 and 680 residues [22], and 94-104 kDa for vegetable lipoxygenases [1] between 839 and 923 residues [23]. Plant and animal LOXs are composed of two domains: a larger α-helical catalytic domain and a smaller N-terminal β-barrel domain (Figs. 2 and 3). The fold of the N-terminal β-barrel domain resembles that of a “C2” module, a kind of calcium-dependent membrane-targeting domain found in many protein kinases and phospholipases. Similarly, the β-barrel domains of LOXs mediate membrane binding, such as the C2-like membrane-binding domain of gangrene α-toxin, a domain with a similar topology to the N-terminal domain of plant and animal LOXs [19].

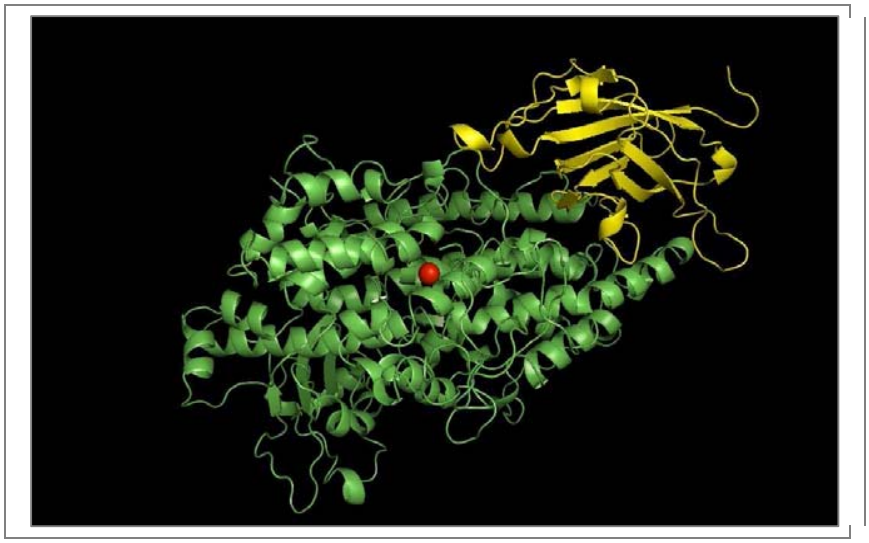

Figure 2. Structure of the soybean lipoxigenaseL-1(Glycine max) from 1.4 Å resolution (PDB code 1YGE). Yellow: N-terminal domain Green: Carboxy-terminal catalytic domain. The red sphere represents an iron atom coordinated in the active site of the protein.

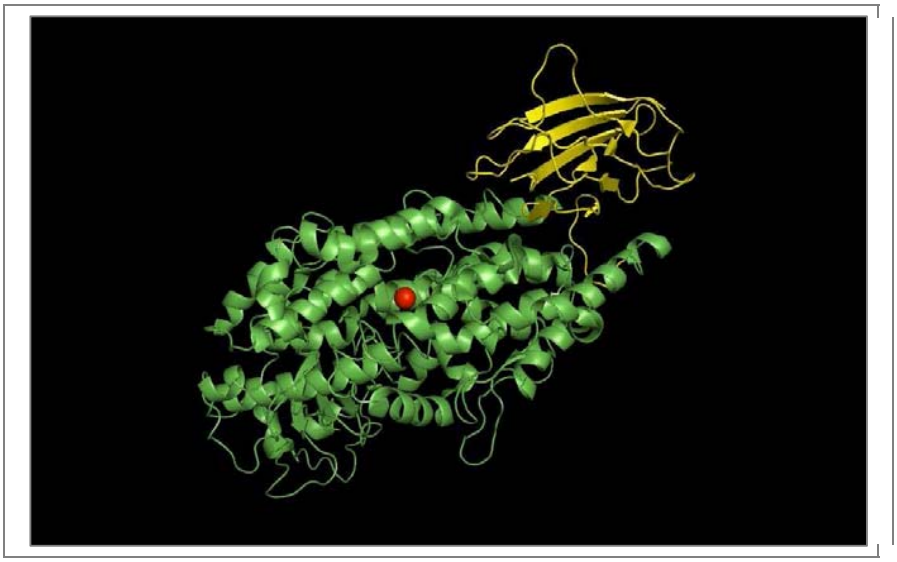

Figure 3. Structure of 8R-lipoxigenase (*Plexaura homomolla*) from 3.2 Å resolution (PDB code 2FNQ). Yellow: N-terminal domain Green: Carboxy-terminal catalytic domain. The red sphere represents an atom of iron coordinated in the active site of the protein.

The larger α-helical catalytic domain contains the non-heme iron positioned in a cavity that accepts the hydrocarbon substrate [8]. This domain consists of 23 α-helix and 2 anti-parallel β-sheets for the lipoxygenase of *Glycine max* [20], and 18 α-helix and one anti-parallel β-sheet in the case of animal lipoxygenases [21].

Iron is bound to protein by five amino acids and, in some cases, one hydroxyl group. For vegetable lipoxygenases, these five amino acids are three histidines, one asparagine and the carboxyl group of the terminal isoleucine. In the case of soy bean (*Glycine max*) lipoxygenes, these five amino acids are histidines 499 and 504 of the helix α9, histidine 690 and asparagine 694 of the helix α18 and isoleucine 839 of the extreme carboxy terminal (Fig. 4) [24].

In the case of mammals, lipoxygenase has its iron coordinated by four histidines and, again, the terminal isoleucine [25]. These amino acids, together with a hydroxyl group, form an octahedral arrangement called a slanted bipyramid [1].

To examine the second coordination sphere of iron, soy lipoxygenase was studied. The second coordination sphere consists of two amino acids, glycine 495 and glycine 697, which form a network of hydrogen bonds between the cavity of the substrate and the first coordination sphere (asparagine 694). Some studies reflect how mutants of soy bean lipoxygenase (Fig. 5), like Q495E, Q495A, Q697E or Q697N, affect enzyme activity because the network of hydrogen bonds is modified, affecting the cleavage of the C-H bond (Fig. 5 b, c, d, and e). An efficient break of this bond is achieved through the correct position of the substrate with steric interactions with the side chain of Gln495 (Fig. 5a). The network of hydrogen bonds defined by Gln495, Gln697 and Asn694 (Fig. 5a) will be responsible for the reorganization of hydrogen bonds after the substrate binding [26].

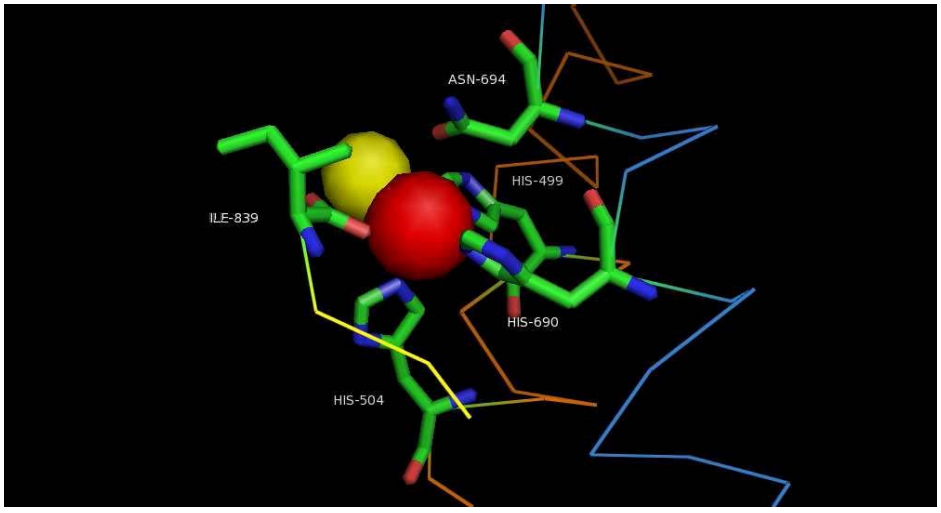

Figure 4. Coordination sphere of the iron atom for lipoxygenase L-1 of soybeans. (*Glycine max*) from 1.4 Å resolution (PDB code 1YGE). In yellow the hydroxyl group, in red the iron atom. In orange, helix 9, in blue helix 21 and in yellow, part of the chain of residues that located appropriately the terminal isoleucine.

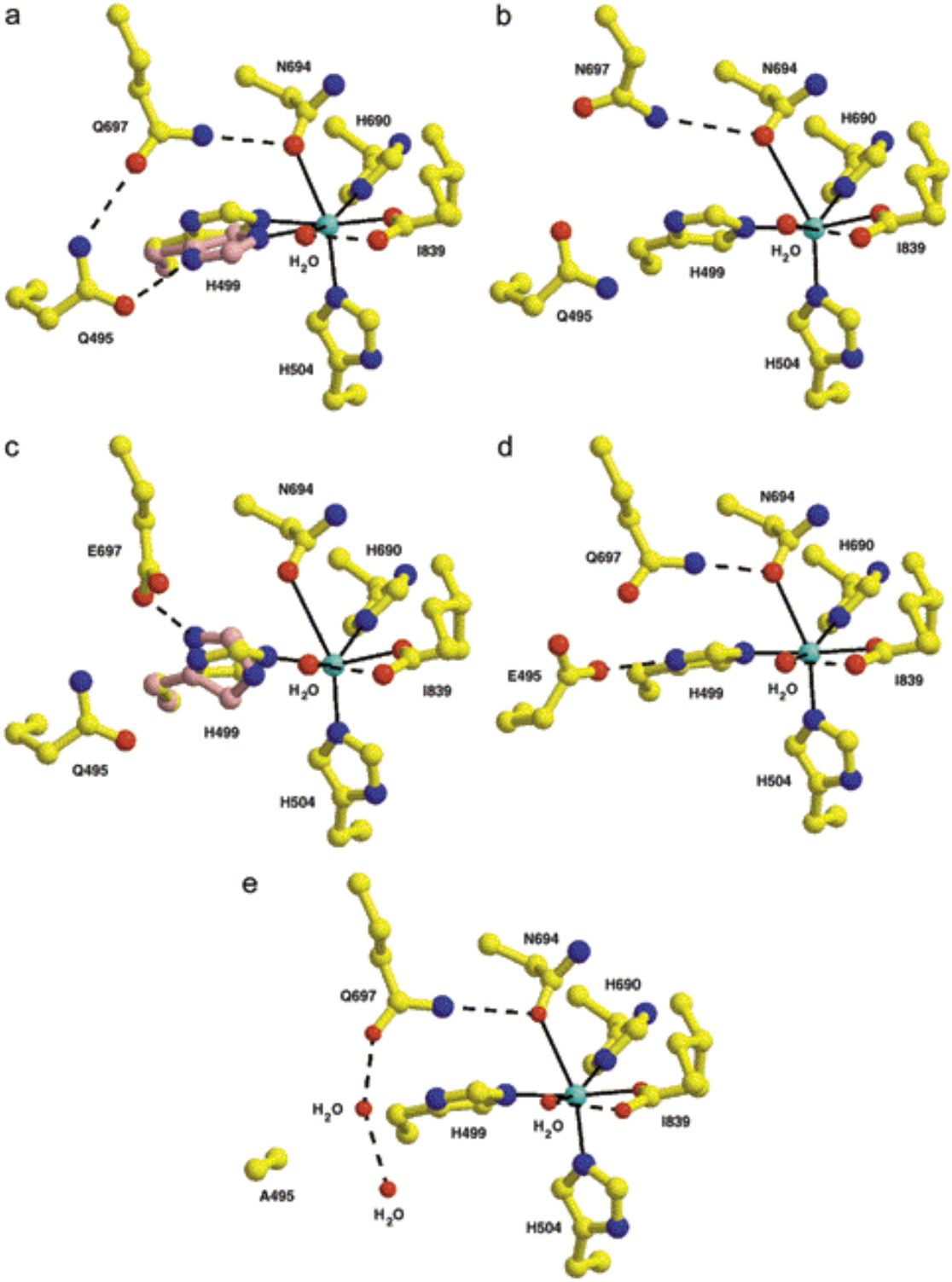

Figure 5. Geometry of the coordination of the Fe^{+2} of the soybean lipoxygenase-1 (R. Leung and D. Sloane) Native lipoxygenase (a) Q697N (b), Q697E (c), Q495E (d) and Q495A (e). The links between iron and the first coordination sphere are drawn as continuous lines and hydrogen interactions are drawn as dashed lines. The oxygen atoms are drawn in red, nitrogen in blue and carbon in yellow. The iron atom is colored in cyan. The alternative conformations of His 499 are shown in pink [26].

The core of the catalytic domain contains two long central helices (Figs. 2 and 3), adopting a conformation of helix π for various residues. These sections contain four of the five helical iron lipoxygenase ligands (His 499, His 504, His 690, His 694).

1.2. Cavities

As soy bean (*Glycine max*) lipoxygenase is the most commonly studied of all lipoxygenases, the best option, in order to study this area of the protein, is to take as a reference the cavities of soy bean lipoxygenase. This protein has two cavities (I and II), which were identified in 1993. Both are in the carboxy terminal domain, the catalytic domain of the protein [6].

The residues that define cavity I are: Cys357, Val358, Ile359, Arg360, Tyr409, Ile412, Tyr493, Met497, Ser498, Leu501, Asn502, Thr503, Val570, Asn573, Trp574, Val575, Asp578, Gln579, Leu581, Asp584, Lys587, Arg588, Tyr610, Trp684, Leu689, and Val693.

Cavity I is conical with an opening of about 10Å at the surface of the protein (Fig. 6). It is defined by the carbonyl group of serine 498 and the side chains of threonine 503, leucine 689 and valine 693: at this point, the presence of a molecule of water is allowed. This area is located at 8Å of the iron atom. The final part of the cavity was in a line that ends at the iron atom and cut the line between histidines 499 and 690, near the coordination sphere of the iron atom [7].

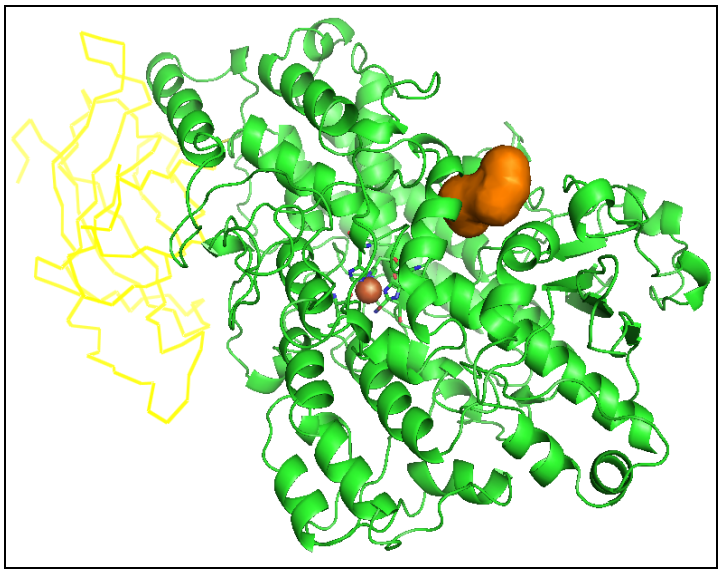

Figure 6. Structure of the soybean lipoxygenase (*Glycine max*). In orange, the cavity, in yellow, the N-terminal domain, and in green the C-terminal domain. In red, the iron atom with five residues coordinating it.

Although Boyington [6] held that most of the residues that define the cavity are hydrophobic, it is a mistake to believe that this cavity is hydrophobic, because water molecules form hydrogen bonds with the protein [7]. According to Boyington, the characteristics of this cavity make it ideal to allow the passage of molecular oxygen from outside the protein to one of two vacant sites of coordination of iron [6]. Contrary to the findings of Boyington and according to Minor, the passage of molecular oxygen required a substantial reorganization of the structure of the protein, since a reorientation of the side chain of residue Asn 694 is not enough to facilitate the access of the oxygen to iron [7].

Thus, cavity I may not be not ideal to allow the passage of molecular oxygen into the active site. The most plausible alternative is cavity II, which allows direct access to iron [7].

Cavity II, 40Å long, stretches in two directions (a and b) from the area near iron (Fig. 7). The farthest end of the cavity is defined by the side chain of residue 341 (methionine) and 480 (leucine), which are 30Å from the iron atom.

Cavity II has two subunits separated by a narrow passage of 3.5 Å, determined by arginine 707 and valine 354. Thus, speaking of cavities IIa and IIb, the first is closest to the iron and is probably the one that links with the fatty acid substrate. Both cavities are interspersed between two layers of helix, helix 9 and 11 on one side and helix 2, 6, 18 and 21 on the other.

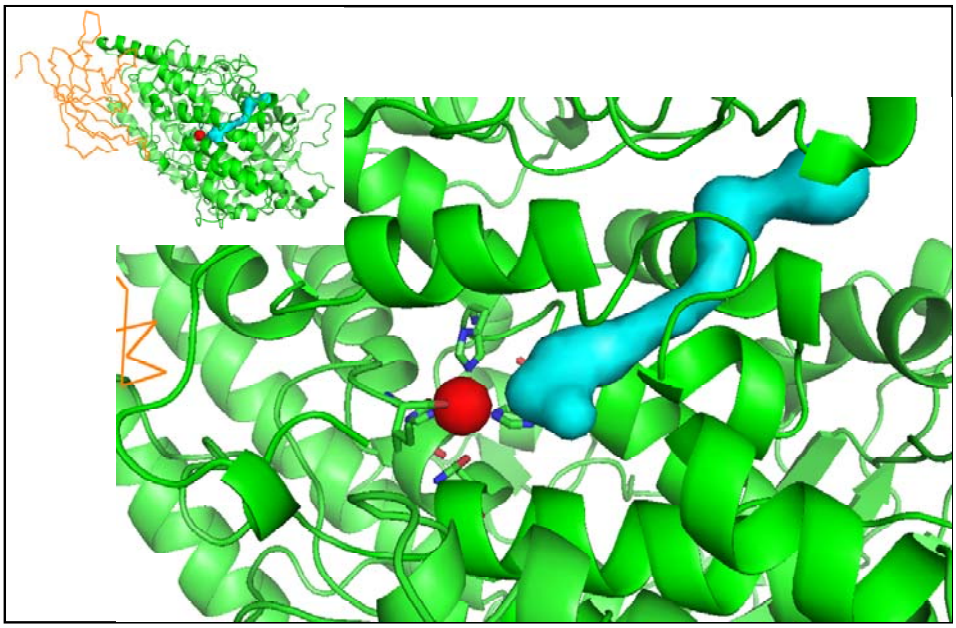

Figure 7. Structure of the soybean lipoxygenase (*Glycine max*). Top image: orange, N-terminal domain, C-terminal domain in green. In cyan, cavity IIa. In red, the iron atom with five residues that coordinate it.

The residues that define cavity IIa are Tyr214, Leu255, Glu256, Gly258, Thr259, Leu496, His499, Trp500, His504, Ile538, Leu541, Ala542, Ile547, Ile553, Val564, Ser567, and Ile839.

The residues that define cavity IIb are: Trp340, Met341, Glu345, Phe346, Glu349, Met350, Gly353, Val354, Asn355, Val358, Ile359, Leu407, Leu480, Lys483, Ala484, Val486, Ile487, Asn490, Tyr700, Met705, Asn706, Arg707, Pro708, and Tyr734 [7].

2. Philogeny of lipoxygenases

The function of a protein is closely related to its structure. Thus, determining the amino acid sequences of many proteins should enable the evolutionary relationships between them to be described. Ideally, alignments for evolutionary studies should be done with the sequence of the protein and respecting the provision in the same space as its tertiary structure. However, as we do not know the structure of most proteins, this is impossible and, therefore, phylogenetic analysis has to be based on the alignments of the nucleotide or amino acid sequences of the various proteins described.

To perform phylogenetic and evolutionary analysis of a protein family, the mutations produced over time on a sequence are usually analyzed. In phylogenetic studies performed with the proteins of the lipoxygenase family, we observe various subgroups depending on their origin, location and specificity because these are determined by the sequence residues located at the active site. Fig. 8 shows a phylogenetic tree for mammal and plant lipoxygenases, which clearly shows the separation of these two groups and the formation of subgroups. It is remarkable that the formation of a particular product for the catalysis of lipoxygenase is not necessarily associated with similar sequences. In contrast, proteins with functional homology of different species located in different subgroups have sequence identities of between 70 and 95 % [1].

Although the databases of nucleotide sequences contain various sequences described as a lipoxygenase or as potential lipoxygenases for microorganisms (bacteria and fungi), there are very few publications on this field. The most recent shows the presence of lipoxygenases in prokaryotic organisms and suggests that lipoxygenases found in cyanobacteria (*Nostoc punctiforme*) could be the evolutionary precursors of 13-LOX, the lipoxygenases of the plastids of plants [27] and that the phenomena of horizontal transfer between Cyanobacteria and other bacteria were responsible for the dispersal of this gene between bacteria [27].

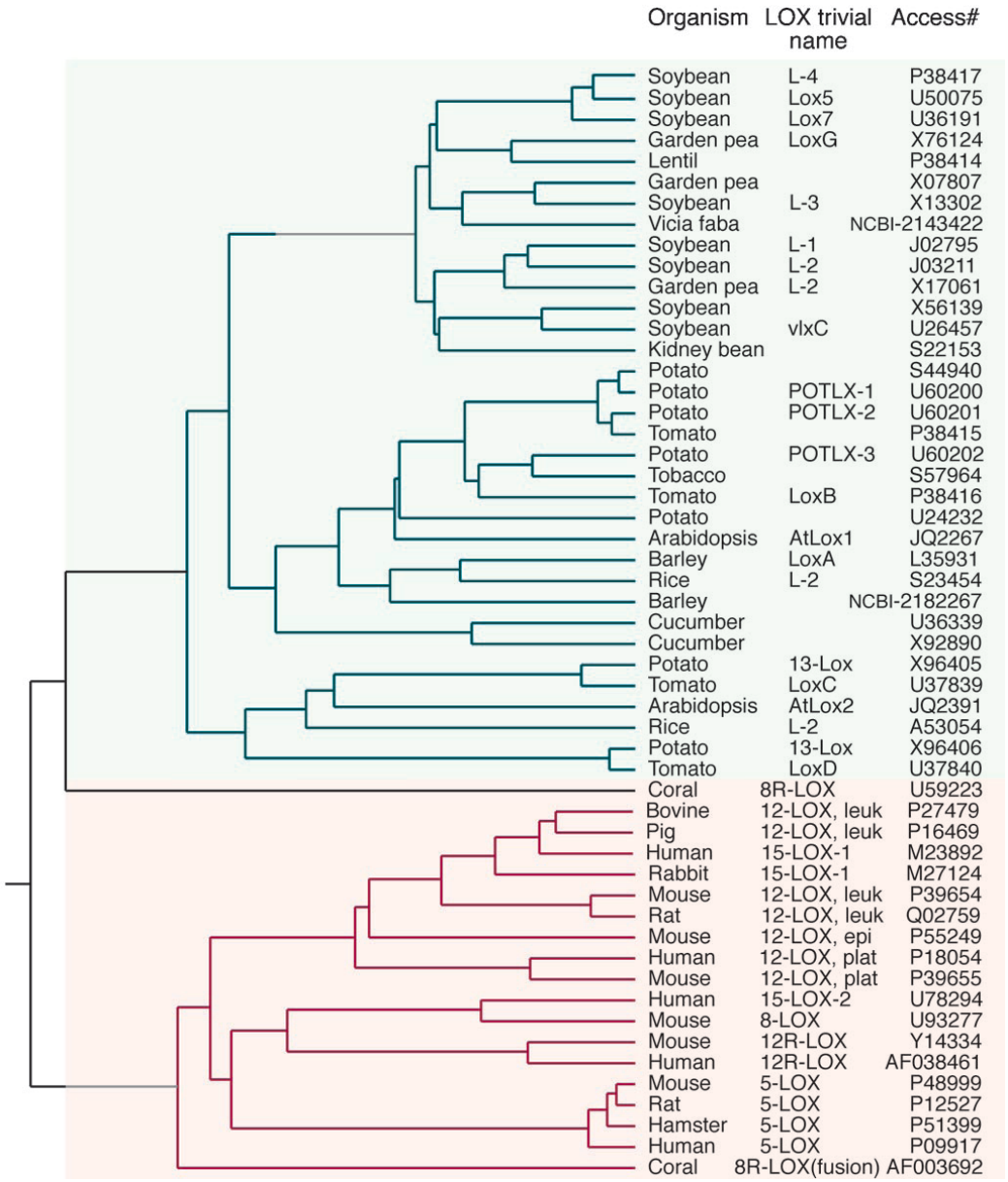

Figure 8. Phylogenetic tree of plant (blue background) and mammals lipoxygenases (pink background) [1].

The phylogenetic tree of Fig. 9 shows the phylogenetic grouping of plant lipoxygenases (13-LOX type 2 groups, 9-LOX 13-LOX type 1 and type 1), animal lipoxygenases, represented only by mammal ones, and the grouping of prokaryotic lipoxygenases. This last group is not very compact and is between the animal and plant lipoxygenases [27]. These results are consistent with those obtained by Porta and Rocha-Sosa in a less detailed phylogenetic study [28].

This phylogenetic tree (Fig. 9) shows the relation between the sequence and the functional properties of each lipoxygenase.

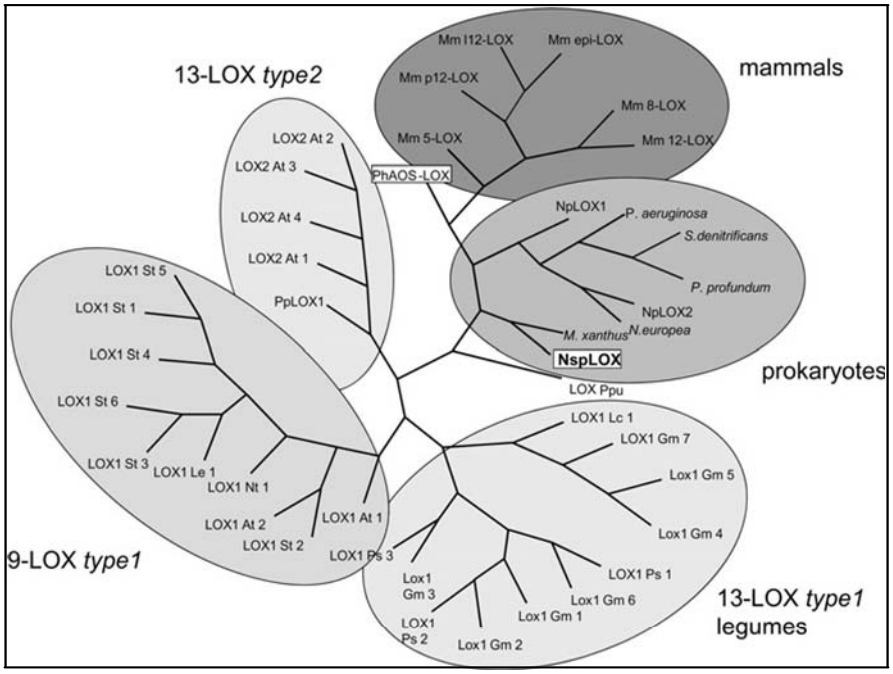

Figure 9. Phylogenetic analysis of sequences of mammals, plants, corals, algae and bacteria lipoxigenases, using the program PHYLIP 3.5 [27]. Access numbers: (i) *Mus musculus* (Moore E.R.B.): 5-LOX AAC37673; 8-LOX CAA75003; 12R LOX CAA74714; l12-LOX AAA20658; p12LOX AAA20659; e12LOX NP663717; (ii) *P. homomalla*: PhAOS–LOX O16025; (iii) *Arabidopsis thaliana* type 2 LOX: LOX2 At1 Q06327; LOX2 At2 CAB56692; LOX2 At3 CAC19364; LOX2 At4 CAG38328; (iv) *Physcomitrella patens* (Pp):CAE47464; (v) *Solanum tuberosum*: St1 CAA5572; St2 AAD09202; St3 AAB31252; St4 CAA64766; St5 CAA64765; St6 AAB67860; (vi) *A. thaliana* type 1 9-LOX: LOX1 At1 NP175900; LOX1 At2 NP188879; (vii) *Lycopersicon esculentum*: Le1 P38415; (viii) *Nicotiana tabacum*: Nt1 CAA58859; (ix) *Glycine max*: Gm1 CAA47717; Gm2 P09439; Gm3 CAA31664; Gm4 P38417; Gm5 AAB67732; Gm6 AAA96817; Gm7 AAC49159; (x) *Pisum sativum*: Ps1 AAB71759; Ps2 CAA55318; (xi) *Lens culinaris*: Lc1 CAA50483; (xii) *Porphyra purpureum*: Ppu AAA61791; (xiii) *Pseudomonas aeruginosa*: AF479686; (xiv) *Nitrosomonas europaea* : BX321860; (xv) *Shewanella denitrificans* OS-217: Q3P217; (xvi) *Photobacterium profundum* 3TCK: ZP 01218321; (xvii) *M. xanthus*: DK 1622 hypothetical protein MXAN 1745 YP 629995; (xviii) *N. punctiforme* (PCC73102): NpLOX1 ZP 00106490; NPLOX2 ZP 00107030; (xvix) *Nostoc sp*. SAG 25.82: NspLOX NP 478445.

3. Results

3.1. *Pseudomonas aeruginosa* 42A2 protein purification

Obtention of the crystal structure of *Pseudomonas aeruginosa* 42A2 lipoxygenase would allow the study of its cavities, the active site and the residues that act as iron ligands.

It is with this goal that a construct of gene AF479686.2 was inserted into pET28a vector (pET28a-LOX). *Escherichia coli* BL21 (DE3) was transformed with the plasmid pET 28a-LOX. The transformants were grown in 200 mL of Luria-Bertani medium containing 100 µg/ml of Kanamycin at 37ºC for 2 hours in an orbital shaker at 150 rpm. When the absorbance reached a value of 0.3 at 580 nm, the culture was supplemented with IPTG (isopropyl-β-D-thiogalactopyranoside) at a final concentration of 0.1 mM and grown for 18 additional hours at 18ºC at 150 rpm. Cells of 1 L of culture were harvested by centrifugation (8000 rpm, 10 min at 4ºC) and resuspended in 10 mL of 20 mM sodium phosphate buffer (pH 7.5) containing 100 mM NaCl and one tablet of antiprotease cocktail "Complete mini EDTA free" (Roche diagnostics, Mannheim, DE).

The bacterial suspension was freezed at -30ºC overnight. After incubation at room temperature for 15 min, cells were disrupted by sonication (3 cycles of 30 seg at 60 V) at 4ºC with an ultrasonic processor (Branson Ultrasonics, Co, USA). After centrifugation at 14,000 rpm at 4°C, the soluble fraction was filtered with 0.2 µm Millipore syringe filters and loaded onto a 1 ml HiTrap Ni-Chelating column (Amersham Pharmacia

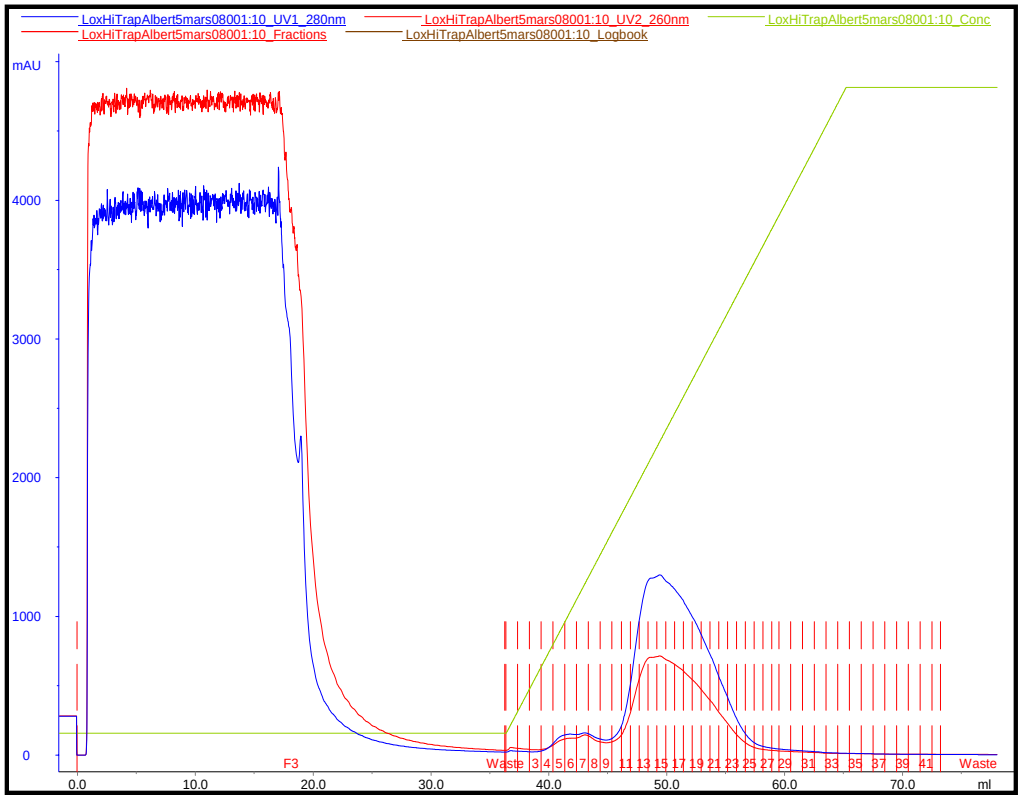

Figure 10. HiTrap Ni-Chelating UV280nm (in blue) profile of the first step of the purification of *Pseudomonas aeruginosa* 42A2 lipoxygenase.

Biotech) equilibrated with 20 mM KH_2PO_4 and 500 mM NaCl at pH 7.4 (buffer A). The column was washed with 10 volums of buffer A supplemented with 10 mM imidazol (Fig. 10). After a 10 to 500mM imidazol gradient elution at 1ml/min flow, 1ml fractions were collected and analysed by SDS-PAGE.

The corresponding fractions were pooled and concentrated using 10,000 MWCO HY Vivaspin concentrators to a final volume of 500 µl. This volume was loaded onto a Superdex 200 10/300 GL column (Amersham Pharmacia Biotech) washed with 50 mM potassium phosphate and 150 mM NaCl at pH 7.0. Fractions of 750 µl were collected and analysed by SDS and Native-PAGE (Fig. 11).

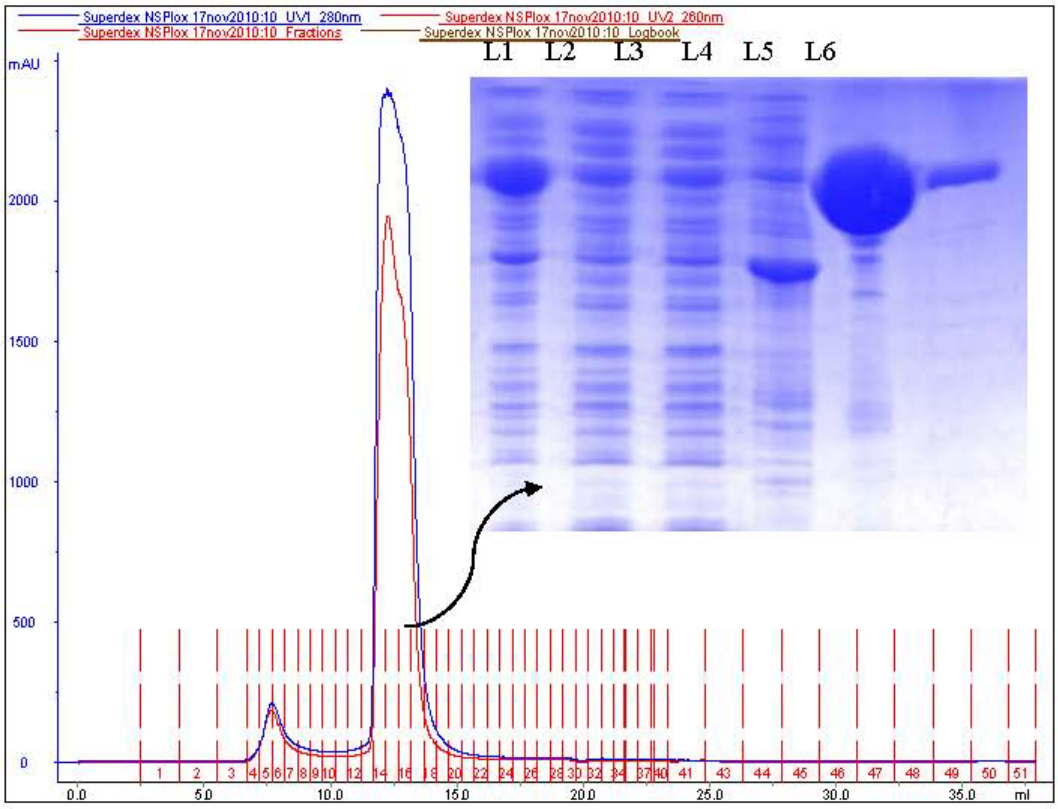

Figure 11. Superdex 200 10/300 UV280nm (in blue) profile of the second step of the purification of *Pseudomonas aeruginosa* 42A2 lipoxygenase. In the inset, the fraction of the peak (lane L5) shows the quality and purity of the purified sample.

3.2. Crystallization of Pseudomonas aeruginosa 42A2 lipoxygenase

The protein sample was crystallized at 20ºC by the vapor diffusion hanging drop method at a protein concentration of 15mg/ml over a reservoir

containing 10% PEG 3350, 50mM $MgCl_2$ and 0.1M HEPES pH 7.5. Crystals were orthorombic and belonged to space group $P2_12_12$ with unit cell parameters of **a**= 132.7 Å, **b**= 116.0 Å, **c**= 42.6 Å and $\alpha=\beta=\gamma=90^o$. A diffraction data set from a cryocooled crystal (Fig. 12) using an extra 20% of glycerol applied to the original mother liquor was collected at a fixed wavelength of 0.9334 Å at beamline ID14eh1 from the ESRF in Grenoble. The data set was integrated and scaled using the program package DENZO/SCALEPACK. The structural resolution is currently underway.

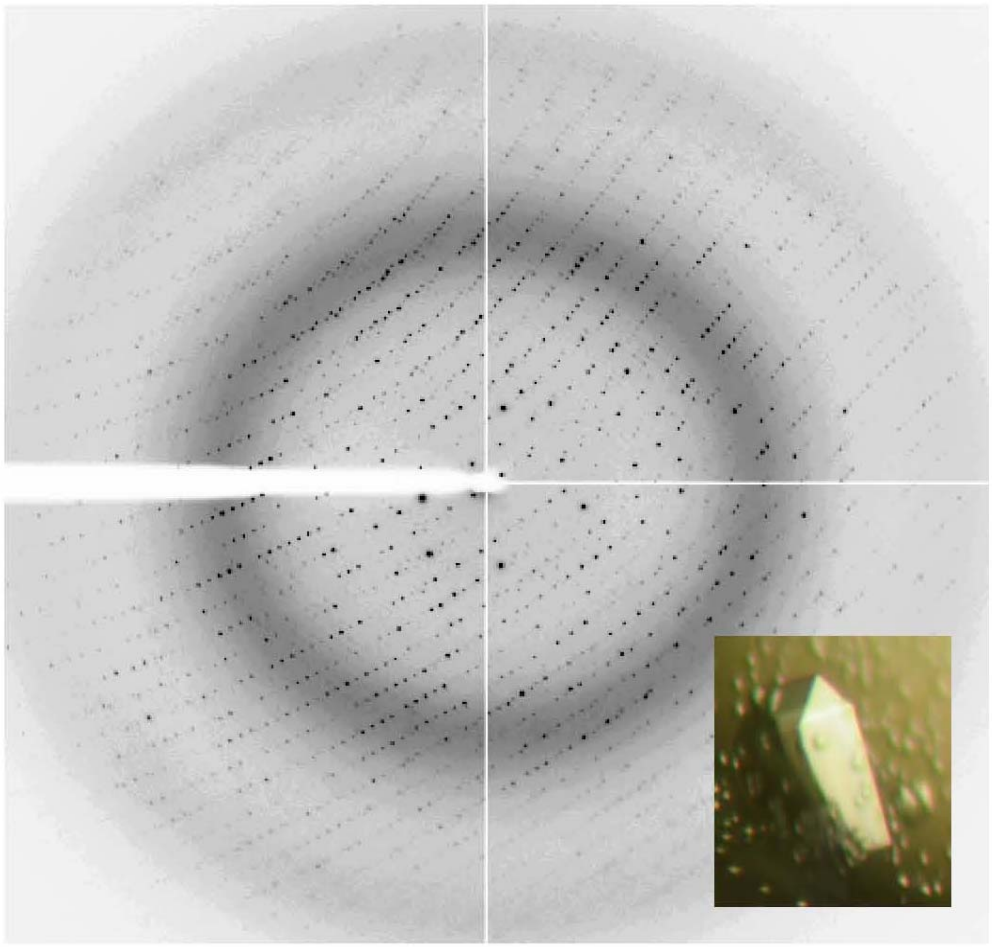

Figure 12. Diffraction pattern of a crystal of *Pseudomonas aeruginosa* 42A2 lipoxygenase (shown in the inset) obtained at beamline ID14eh1 from the ESRF in Grenoble.

3.3. Study of the phylogeny of lipoxygenases

For this study we used 14 sequences of bacterial lipoxygenases, 9 animal lipoxygenases, 10 vegetable lipoxygenases and 4 fungi ones. In the case of

bacterial and fungal sequences, the nature of the lipoxygenases was confirmed by alignments of the highly conserved amino acids that coordinate the iron atom [33] and an amino acid sequence characteristic of lipoxygenases, WXXAK [24]. The two tables below (Tables 1 and 2) show the positions of the iron ligands in relation to the sequences of bacterial and fungal lipoxygenases and how well they are aligned within the groups concerned. In the case of bacterial sequences, some of the singularities present in the sequences of ATCC 51908 *Shewanella woodyi* (ZP_01543473.1) should be mentioned. In this case, the fourth ligand that normally varies between asparagine (N) or histidine (H) is replaced by an arginine (R).

The fifth ligand, the C-terminal amino acid, is also altered in two of the three sequences of lipoxygenases present in the genome of ATCC 51908 *Shewanella woodyi*. This position usually has an isoleucine (I) or a valine (V) while, in two of the three lipoxygenase sequences, there is an asparagine (N) (Table 1).

Table 1. Alignment and position of amino acids that bind the iron atom in the sequence typical of lipoxygenases (WXX AK) of all bacterial sequences used in this study.

Species	GenBank acces	pb gen	aa gen	W - - A K	Ligands of Fe
Burkholderia thailandensis E264	CP000086.1	2088	695	368 WQMAK 372	388 H, 393 H, 565 H, 569 N i 695 I
Myxococcus xanthus DK 1622	NC_008095.1	2028	675	346 WLAAK 350	366 H, 371 H, 549 H, 553 N i 675 V
Nitrosomonas europaea ATCC 19718	NP_841292.1	1698	565	244 WLAAK 248	264 H, 269 H, 445 H, 449 N i 565 V
Nitrosospira multiformis ATCC 25196	NC_007614.1	1662	553	232 WLTAK 236	352 H, 357 H, 433 H, 437 H i 553 I
Photobacterium profundum 3TCK	ZP_01218321.1	1845	614	289 WQAAM 293	309 H, 314 H, 487 H, 491 N i 614 I
Pseudomonas aeruginosa 42A2	AAL85880.2	2058	685	357 WQMAK 361	377 H, 382 H, 555 H, 559 N i 658 I
Pseudomonas aeruginosa PAO1	NC_002516.2	2058	685	357 WQMAK 361	377 H, 382 H, 555 H, 559 N i 658 I
Pseudomonas aeruginosa UCBPP-PA14	CP000438.1	2058	685	357 WQMAK 361	377 H, 382 H, 555 H, 559 N i 658 I
Shewanella denitrificans OS217	NC_007954.1	1857	618	293 WQAAM 297	313 H, 318 H, 491 H, 495 N i 618 I
Shewanella woodyi ATCC 51908	ZP_01543473.1	2145	714	384 WQMAK 388	403 H, 407 H, 593 H, 597 R i 714 N
Shewanella woodyi ATCC 51908	NZ_AAUO01000023.1	2271	756	422 WKMAK 426	441 H, 445 H, 635 H, 639 N i 756 N
Shewanella woodyi ATCC 51908	NZ_AAUO01000010.1	2178	725	383 WQIAK 387	403 H, 408 H, 600 H, 604 N i 725 I

Table 2. Alignment and position of amino acids that bind the iron atom in the sequence typical of lipoxygenases (WXXAK) of all sequences of fungi used in this study.

Species	GenBank acces	pb gen	aa gen	W - - A K	Ligands of Fe
Gibberella zeae	XP_382392.1	2238	745	387 WRYAK 391	407 H, 412 H, 596 H, 600 N i 745 I
Chaetomium globosum	XP_001225066.1	2256	751	409 WRYAK 413	429 H, 434 H, 611 H, 615 N i 751 I
Neurospora crassa	CAD37061.1	2268	755	411 WRYAK 415	431 H, 436 H, 615 H, 619 N i 755 I
Aspergillus fumigatus	XP_746844.1	2235	744	398 WRYAK 402	418 H, 423 H, 603 H, 607 N i 744 I

Regarding the characteristic sequence of lipoxygenases (WXXAK), the high conservation of tryptophan (W) and alanine (A) is clear, as is, to a lesser extent, the lysine (K), which can be altered by methionine (M) in the case of sequences of bacterial origin (Tables 1 and 2).

There was no need to check the animal and plant sequences, as these lipoxygenases have been studied thoroughly and, in some cases, such as the lipoxygenase of *Plexaura homomolla*, the crystal structure is published [34].

As can be seen in the phylogenetic tree (Fig. 18), the lipoxygenases form a heterogeneous family of enzymes with clearly different subfamilies, corresponding to animal organisms, plant organisms, fungal organisms and bacterial organisms. This kind of grouping (animal, plant and bacterial lipoxygenases) was described in 2007 [35] and is confirmed by the addition of a new lipoxygenase subfamily, the bacterial lipoxygenases consisting of lipoxygenases of *Nitrospira*, *Nitrosomonas*, *Photobacterium*, and *Shewanella woodyi*.

There are several interesting points about this phylogenetic tree (Fig. 13). First, the case of *Myxococcus xanthus*, which is outside the group of bacterial lipoxygenases, closer to animal lipoxygenases than the vegetable ones, a phenomenon published in 2008 [27], might suggest, as was proposed in 2001, a case of horizontal transfer [28]. The same occurs for the Plexaura homomolla lipoxygenase that appears outside the group of animal lipoxygenases in a phylogenetic tree published in 2007 [35].

The case of the two *Shewanella woodyi* sequences (a and b), which are clearly outside the group of bacterial lipoxygenases, has never been published. It seems quite reasonable to attribute this to a failure in the annotation of these genes, because although they retain the typical sequence (WXXAK) of lipoxygenases, the typical valine or isoleucine terminal is not present.

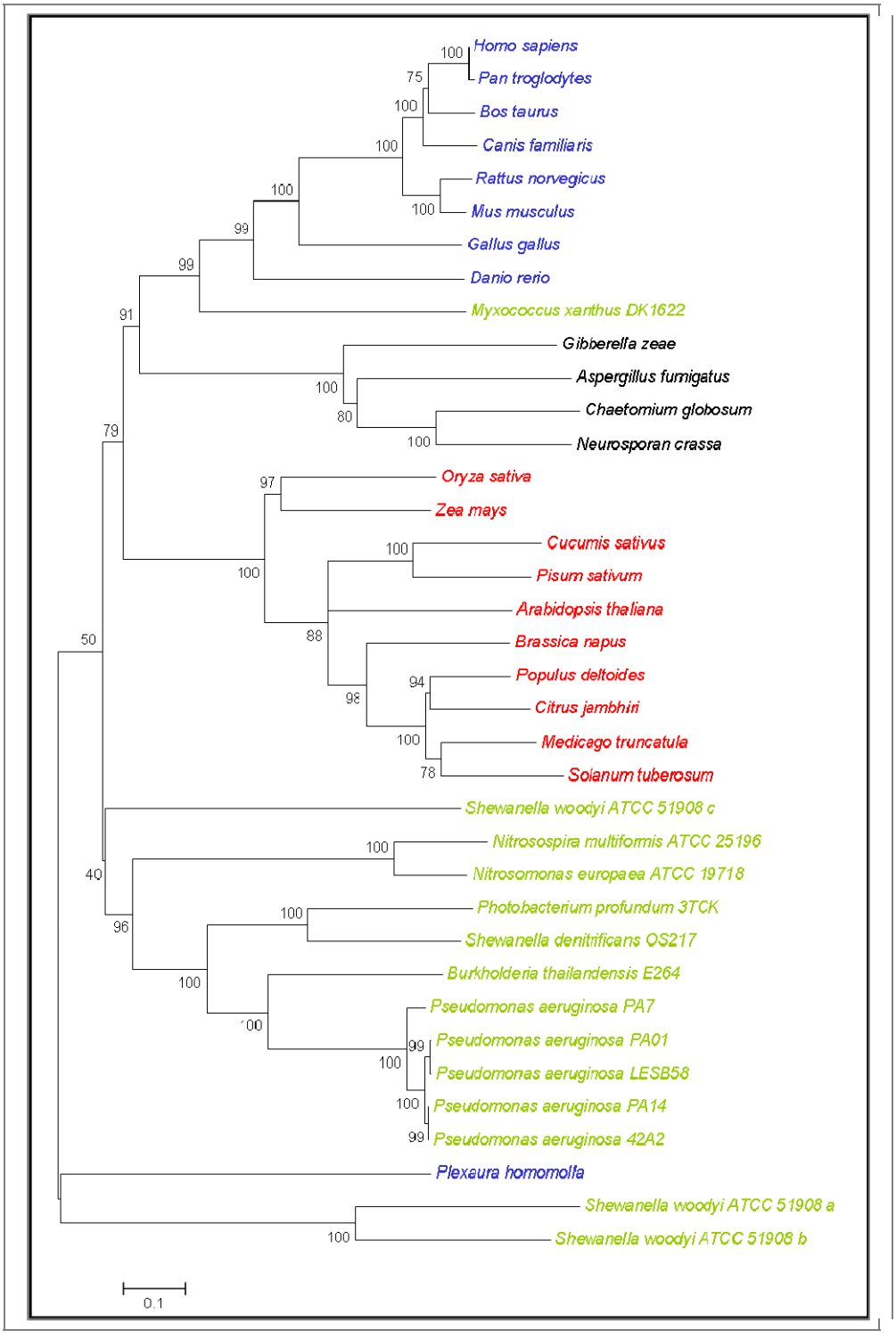

Figure 13. Consensus phylogenetic tree constructed using MEGA4 program under the Neighbor-Joining method with 1000 bootstrap replicates and the Juke-Cantor model (calculation of distance) from an alignment of nucleotide sequences. All positions containing gaps were removed from the dataset (Complete deletion option). In red, vegetal lipoxygenases, in blue, animal lipoxigenases, in black fungal lipoxygenases and in green, bacterial lipoxygenases.

The subfamily of bacterial lipoxygenases does not show degrees of similarity to the animal and plant subfamilies, but the subgroup formed by the lipoxygenases of *Pseudomonas*, *Burkholderia*, *Photobacterium*, and *Shewanella denitrificans* (all these organisms are proteobacteria) do have a degree of similarity with animal and plant lipoxygenase subfamilies.

It is necessary to take into account two key aspects in this kind of analysis. First, the quality of the sequences that are to be worked with must be checked and, second, a high number of sequences is recommended. In the case of this study, this second aspect was improved, due to the few sequences classified as lipoxygenase ones present in databases. As regards the first aspect, it is revealed that, although some sequences have features that are typical of sequences of lipoxygenases, they may not be such sequences.

3.4. Evolution of lipoxygenases

For this In order to study the evolution of lipoxygenases of prokaryotic origin, an alignment of the 14 sequences has been performed yielding 2670 sites. Taking into account that the longest sequence is that of Shewanella woodyi ATCC 51908 (2271 nucloeotides), this so different sequence length by itself is already indicating a great sequence diversity. The alignment has been analyzed with CODEML program (*), by previous removal of the stop codon

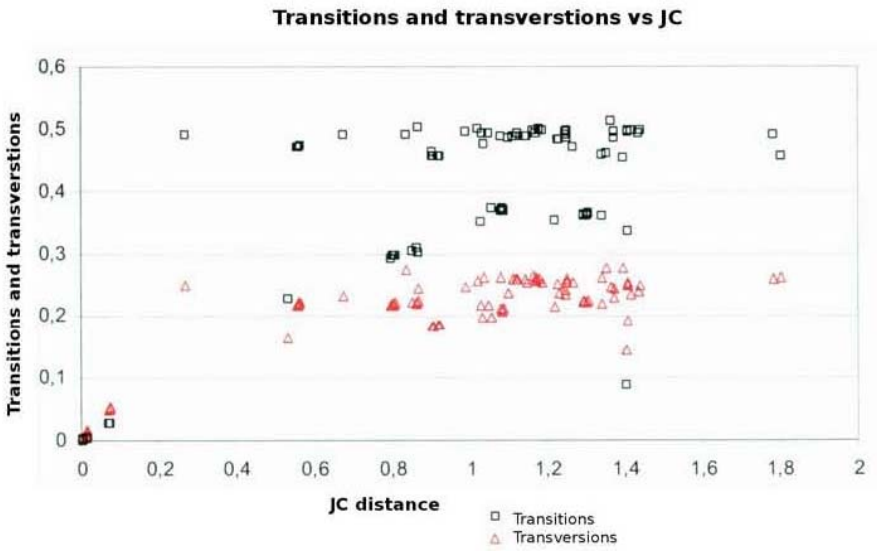

Figure 14. Graph showing the substitutions saturation (transition –squares- and transversion –triangles) with respect to the phylogenetic distance (JC) for the bacterial lipoxygenases set.

and sites of less than three sequences which represented nucleotide triplets that codified for any aminoacid. An alignment of 2457 sites is thus obtained, which only suffers a loss of an 8% of information with respect to the total alignment length.This modified alignment presents 47 invariable sites and a G+C percentatge which ranges from 68.2% (Myxococcus Xanthus) since 42.1% (Shewanella woody, GI:118072427), with an average of 55.9%. Regarding the lipoxygenases from the five Pseudomonas aeruginosa strains, the G+C percentage is of 66.5% with a standard deviation of 0.38 whereas for the 14 sequences set is of 5.45. This data indicates, once more, a great sequence diversity among bacterial lipoxygenases group. This procedure has also been applied to other groups (animals, plants and fungi).

The average rate of transitions / transversions for the modified alignment is of 0.76, i.e. there are more transversions (change of a purine base, adenine or guanine, by a pirimidinic one, cytosine or timine) than transition (change of a purine base to another purine base or change of a pirimidinic base by another one). Fig. 14 shows the substitutions rate (transversion and transition) with respect to the calculated distance using the Jukes-Cantor method (JC) for the bacterial lipoxygenases group.

This graph shows a high saturation of substitutions from a divergence of 0.3, showing multiple substitutions occurrence and even the existence of a homoplasia case (similarities between species with a different ancestor as a result of a convergent evolution).

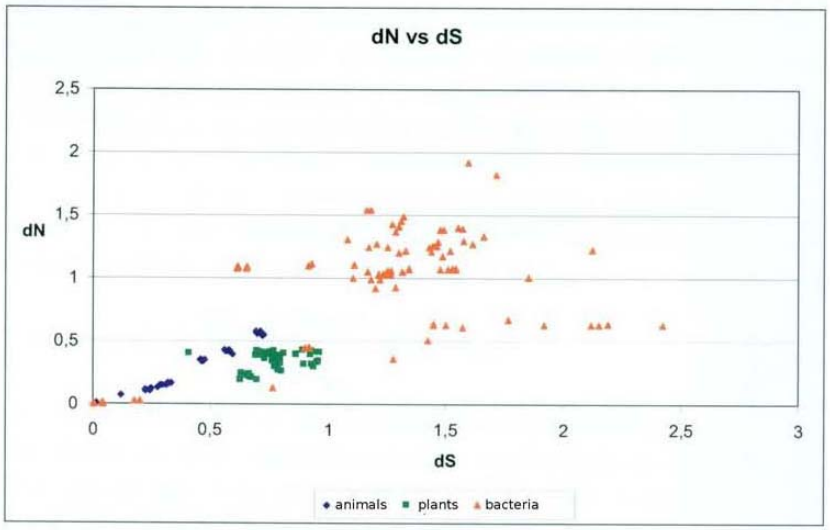

Figure 15. Graph showing synonym (dS) vs. non-synonym (dN) substitution rates from an alignment performed using the maximum likelihood method. Orange squares account for bacterial lipoxygenases, green ones for plants and blue rombohSSS, for animals.

Using Jukes-Cantor method and comparing pairs of sequences, synonym substitions rates (dS) and non-synonym rates (dN) have been estimated for bacterial lipoxygenases and for those of animal and plant origins. Those results are shown in Fig. 15 where the difference between prokaryotic and eukaryotic lipoxygenases are clearly evident.

Those results match perfectly with the phylogenetic trees from chapter 3.3, since they confirm that prokaryotic lipoxygenases present remarkable differences, at a nucleotide sequence level, with eukaryotic ones. ω

From the synonym and non-synonym substitution rates, the ω rate is defined as the dN/dS ratio. In Fig. 16 the relationship between ω and the phylogenetic distance (obtained by JC method from the alignment of the set of lipoxygenases of this study) is defined. This figure allows us to observe three differentiated sets of data corresponding to animal, plant and bacterial lipoxygenases.

The behavior of these data sets present significative differencies. For bacterial lipoxygenases, its regression graph (data not shown) has an 0.1424 of PENDENT, whereas 0.2789 and 0.2976 is applied in plant and animal types, respectively. This latter result shows how ω is already saturated in bacterial lipoxygenases but not in the others. We can clearly observe in the graph how JC distance values greater than 1.3 result in little variation of ω values. Hence, from this point the relationship between synonym and non-synonym changes do not vary regardless of an increasing phylogenetic distance, thus the variation accumulation is now saturated. Besides, the greater value of PENDENT of animal and plant lipoxygenases shows that saturation has not been reached in those subfamilies. To say it in other words,

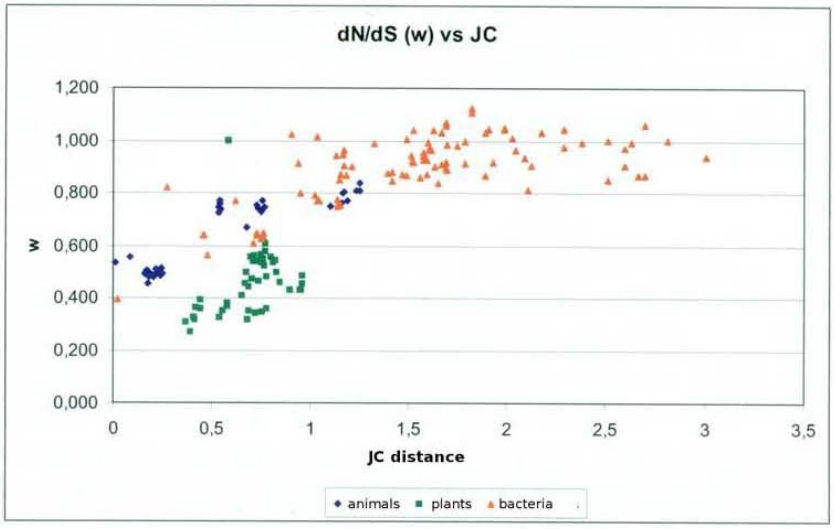

Figure 16. Graph of ω (dN/dS) vs. phylogenetic distance (calculated by JC method). Orange triangles account for bacterial lipoxygenases, green squares for those from plants and blue ROMB for those of animal origin.

while animal and plant lipoxygenases are still accumulating mutations that could alter its evolution, those from bacterial origin, despite mutations are in principle still possible, those will not affect its future evolution.

This saturation argument matches perfectly with the high dS values and its wide range (Fig. 15) and with the clear transitions and transversions saturation of Fig. 14 for divergence values higher than 0.5.

It is worthy to remark that this 14 lipoxygenases set study (Figs. 15, 16 are an example of this) is not representative of the great diversity of the lipoxygenase family.

In order to unveil the nature of the selective pressure which is acting on the alignment with which we have studied, several models developed by Nielsen and Yang have been used. These models are integrated in CODEML program of the PAML package. Models M0, M1, M2, M7 and M8, which assume ω variations among sites but not between lines, have been applied to the alignment. The model chosen, the parameters estimation and the likelihood factor are shown on Tab. 3.

Table 3. Results obtained by CODEML depending on the model used. ***p***, number of free parameters for each model.**ω**, average of **ω** values of all the sites of the alignment of the bacterial lipoxygenases of the study. ***l***, log-likelihood for each model. ***k***, transition / transversion rate.

cleandata 1						
Model	***p***	**ω**	***l***	***k***	**Parameters estimation**	**Positively selected sites *p* (w>1)**
M0	1	0,08039	-13369,71177	1,65319	w = 0.08039	-
M1	2	1,0000	-13751,52324	2,19726	*p*0 = 0.00001 w0 = 0 *p*1 = 0.99999 w1 = 1	-
M2	4	0,92342	-13751,51963	2,19726	*p*0 = 0; *p*1 = 1; *p*2 = 0 w0 = 0; w1 = 1; w2 = 1	**127 R 0.988**
M7	2	0,04706	-13307,37354	1,70077	*p* = 1.82107; *q* = 15.60717	-

M8	4	0,93942	-13751,52324	2,19726	*p*0 = 0.00001; p = 0.00500;
					q = 1.88239;
					*p*1 = 0.99999; w = 1
						127 R 0.992

cleandata 0						
Model	***p***	**ω**	***l***	***k***	**Parameters estimation**	**Positively selected sites *p* (w>1)**
M 0	1	0,10265	-25039,04654	1,6681	w = 0,10265	-
M 1	2	0,23539	-24968,70447	2,19726	*p*0 = 0.00001 w0 = 0 *p*1 = 0.99999 w1 = 1	-
M 2	4	0,326123	-24968,70447	1,86405	*p*0 = 0,83; *p*1 = 0,0536; *p*2 = 0.1163 w0 = 0,0788; w1 = 1; w2 = 1	-
M 7	2	0,130452	-24914,88305	1,77087	*p* = 1,48857; *q* = 9,62348	-
M 8	4	0,200475	-24910,60091	1,79986	*p*0 = 0,95910; p = 1,70908; q = 13.34477 *p*1 = 0,04090; w = 1	373 R 0,602* 554 L 0.715 576 A 0.632 768 R 0,691

Two sets of results have been calculated depending if the cleandata option was set to 0 or to 1. In the former case, the alignment employed suffers no alteration whereas in the latter, all sites showing any gap are removed altering the alignment. Despite a lot of information is lost (we reduce the number of sites from 2457 to 1083) this option is recommended by the program. Besides, as it can be seen on Tab. 3, only with this option can positively selected sites be obtained (i.e. sites with a $\omega>1$ with a probability higher than 0.95) since the program treats the detected gaps as undefined data and this reduces a lot ω values. It is thus for this reason that, when working with cleandata 0, ω values are always lower than with cleandata 1. The removal of sites with some gap translates into a dS increase and hence, dN/dS (ω) ratio diminishes.

The results of the Likelihood Ratio Test (LRT), restricted to nested models where the simplest model is a special case of the most complex case, are shown in Tab. 4. With this test, it can be evaluated which one of the two models used better adjusts to the data applied to the program.

Using the simplest model (M0) which assumes an uniform ω for all the alignment codons, the ***l*** value (maximum likelihood logarithm) is of -3369,71 with an estimated ω of 0.08039. M1 model assumes a positive pressure that is variable among sites but not a positive selection. It is model M2 which assumes a positive selection. ***l*** values for M1 and M2 models are -13751.51963 and -13751.51963 with ω of 1.0 and 0.92342, respectively.

The transition / transversion ratio values (***k***) for the five applied models is constant, regardless of cleandata value since this parameter remains generally constant when the gen has reached saturation (Fig 14).

In order to determine which one of the three models (M0, M1 or M2) adjusts better to the data, the LRT ($2\Delta\boldsymbol{l}=2(\boldsymbol{l}_1-\boldsymbol{l}_0)$) is performed. The *p*-value obtained of a LRT ($2\Delta\boldsymbol{l}$) between models M1 (close to neutral selection) and M2 (positive selection) is of 9.861800652 (Tab. 5), thus accepting the alternative hipothesis making model M2 (positive selection) the one that adjusts to the data better. Using the same procedure but in this case between model M0 and M2, the *p*-value obtained is of $3.3643\text{x}10^{-165}$, far lower indicative of a null hipothesis showing once more that model M2 is the one that better adjusts to the data (Tab. 4).

In models M0 and M1, no one of the studied aminoacids have been selected with a higher probability of 95%. This was expectable since the model that better adjusts to the data is positive selection model M2. In this model, arginine 127 shows a 0.988 probability for having evolved under a positive selection (Tab. 3). This aminoacid is also selected by model M8 which assumes a beta distribution with a ω higher than 1. Hence, it can be concluded that only aminoacid 127 follows an evolution with a positive selection whereas the resting studied aminoacids follow an evolution with a neutral or purifying selection pattern.

Table 4. Likelihood Ratio Test (LRT) results for different evolutionary models available in CODEML program. $2\Delta l=2(l_1-l_0)$ (l. *p*-value obtained from a χ^2 measured with the degrees of freedom shown in the table.

cleandata1

Alternative hipothesis	Null hipothesis	$2\Delta l$	Freedom degrees	*p*-value
M0	M2	763,615714	3	3.3643×10^{-165}
M2	M1	0,00722	2	9,861800652
M7	M8	762,299396	2	2.943×10^{-166}

Since LRT can only be applied over nested models, $2\Delta l$ value need to be determined comparing model M8 against model M7 which solely presupposes a beta distribution (continuous probability distribution with parameters *a* and *b* with a density function values ranging between 0 and 1). ***l*** values for models M8 and M7 are -13751.5234 and -13307.37354, respectively, yielding a p-value for a LRT ($2\Delta l$) between models M7 and M8 of 2.943×10^{-166} (Tab. 4), this value agrees with a null hipothesis situation showing that model M8 adjusts better to the data introduced into the program.

Since models M2 and M8 are not nested, a χ^2 can't be applied and thus can not be compared to know which one of the two adjusts better to the data. In such situation, those models with a higher ***l*** value will be the ones that adjust better. Despite not very different from that of model M2 (0.988), model M8 presents the highest probability (0.992) for the selected aminoacid.

According to model M8, the average ω for the 361 studied codons is of 0.9394 with a 0.1915 variance. Fig. 17 shows ω distribution through all the 361 codons following a beta distribution with a ω in the majority of sites that ranges from 0.5 and 1.0, suggesting that either the purifying (ω<1) or the neutral (ω=1) are the predominant evolution forces. From all of those sites with a ω value higher than 1, there's only one (Fig. 17, in red) presents a 0.992 probability which exceeds a 99% of confidence and that points out that this site (arginine 127) follows a beta distribution with a ω higher than 1, meaning that it has suffered a positive selection evolution.

This statement is further reinforced by the fact that using model M2, the same results are obtained despite its associated probability (0.998) lies below the 99% degree of confidence (but above the 95% degree of confidence).

All in all, the above exposed results confirm a neutral or purifying selection pressure for the vast majority of the gene and a positive selection of arginine 127. This arginine corresponds to number 325 in *Pseudomonas aeruginosa* 42A2 aminoacidic sequence.

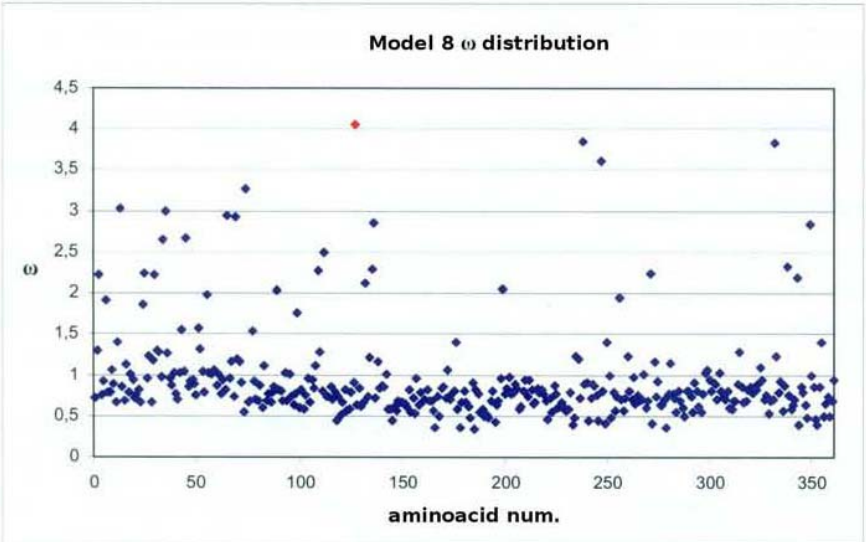

Figure 17. Distribution of ω obtained for each of the sites of the three bacterial lipoxygenases alignment using model M8 and a beta distribution (CODEML program). In red, arginine 127 W.

The differences observed applying cleandata 1 or 0 show that, despite having certain implications in the estimation of the model parameters, seem not to alter the identification of the sites under positive selection. However, the insertions and deletions might have a role in the divergence of the different lipoxygenase groups defined in the present study (animal, plant and bacterial lipoxygenases).

Aminoacid 325 detection as the unique aminoacid which suffered a positive selection pressure opens new questions about Pseudomonas aeruginosa 42A2 structure. To solve these questions, site directed mutagenesis studies should be performed so as to analyse how mutating this aminoacid has a role in the enzyme activity or in its structure.

4. Conclusions

1. We have obtained the first crystal of a bacterial lipoxygenase using 10% PEG 3350, 50mM $MgCl_2$ and 0.1M HEPES pH 7.5. The crystal structural resolution, which is currently underway, will presumably become a structural model for prokaryotic lipoxygenases.
2. The existence of a subfamily of bacterial lipoxygenases has been confirmed. This subfamily has a greater diversity than in the corresponding eukaryotic lipoxygenases.
3. The evolution study of bacterial lipoxygenases show a selection pressure of a basically purifying or neutral character. This is valid for the whole

length of the gene except for aminoacid 325 which would have been selected after a positive selection event.

References

1. Brash, A. R. 1999, *J. Biol. Chem.* 274, 23679.
2. Funk, C. D. 2001, *Science* 294, 1871.
3. Serhan, C. N., Arita, M., Hong, S., Gotlinger, K. 2004, *Lipids* 39, 1125.
4. Smith, W. L., DeWitt, D. L., Garavito, R. M. 2000, *Annu. Rev. Biochem.* 69, 145.
5. Liavonchanka, A., and Feussner, I. 2006, *Plant Physiol.* 163, 348.
6. Boyington, J. C., Gaffney, B. J., Amzel, L. M. 1993, *Science* 260, 1482.
7. Minor, W., Steczko, J., Stec, B., Otwinowski, Z., Bolin, J. T., Walter, R., Axelrod, B. 1996, *Biochemistry* 35, 10687.
8. Gillmor, S. A., Villaseñor, A., Fletterick, R., Sigal, E., Browner, M. F. 1997, *Nat. Struct. Biol.* 4, 1003.
9. Skrzypczak-Jankun, E., Bross, R. A., Carroll, R. T., Dunham, W. R., Funk, M.O., Jr. 2001, *J. Am. Chem. Soc.* 123, 10814.
10. Dainese, E., Sabatucci, A., van Zadelhoff, G., Angelucci, C. B., Vachette, P., Veldink, G. A., Agro, A. F., Maccarrone, M. 2005, *J. Mol. Biol.* 349, 143.
11. Hammel, M., Walther, M., Prassl, R., Kuhn, H. 2004, *J. Mol. Biol.* 343, 917.
12. Oldham, M. L., Brash, A. R., Newcomer, M. E. 2005, *J. Biol. Chem.* 39, 39545.
13. Krieg, P., Marks, F., Fürstenberger, G. 2001, *Genomics* 73, 323.
14. Funk, C. D., Chen, X. S., Johnson, E. N., Zhao, L. 2002, *Prostaglandins Other Lipid Mediat.* 68–69, 303.
15. Schneider, C., Brash, A. R. 2002, *Prostaglandins Other Lipid Mediat.* 68–69, 291.
16. Feussner, I., Wasternack, C. 2002, *Annu. Rev. Plant Biol.* 53, 275.
17. Yin, H., Porter, N. A. 2005, *Antioxid. Redox Signal.* 7, 170.
18. Schneider, C., Pratt, D. A., Porter, N. A., Brash R. A. 2007, *Chemistry & Biology* 14, 473.
19. Rizo, J., Sudhof, T. C. 1998, *J. Biol. Chem.* 273, 15879.
20. Youn, B., Sellhorn, G. E., Mirchel, R. J., Gaffney, B. J., Grimes, H. D., Kang, C. 2006, *Proteins*, 65, 1008.
21. Kuhn, H., Saam, J., Eibach, S., Holzhütter, H. G., Ivanov, I., Walther, M. 2005, *Biochem. Biophys. Res. Commun.*, 338, 93
22. Funk, C. D. 1996, *Biochim. Biophys. Acta Lipids Lipid Metabol.* 1304, 65.
23. Shibata, D., Axelrod, B. 1995, *J. Lipid Mediators Cell Signalling* 12, 213.
24. Gaffney, B. J. 1996, *Annu. Rev. Biophys. Biomol. Struct.* 25, 431.
25. Andreou, A., Feussner, I. 2009, *Phytochemistry* 70, 1504.
26. Tomchick, D. R., Phan, P., Cymborowski, M., Minor, W., Holman, T. R. 2001, *Biochemistry* 40, 7509.
27. Lang, I., Göbel, C., Porzel, A., Heilmann, I.; Feussner, I., 2008, *Biochem. J.* 410, 347.
28. Porta, H., Rocha-Sosa, M. 2001, *Microbiology* 147, 3199.

29. Neau, D. B., Gilbert, N. C., Bartlett, S. G., Boeglin, W., Brash, A. R., Newcomer, M. E., 2009, *Biochemistry* 48, 7906.
30. Choi, J., Chon, J. K., Kim, S., Shin, W. 2008, *Proteins* 70, 1023.
31. Ratledge, C., Wilkinson, S. G. 1988, "Microbial lipids Volume 1, Chapter 7: Gram-negative bacteria."
32. Coffa, G., Brash, A. R., 2004, *Proc. Natl. Acad. Sci. USA* 101, 15579.
33. Solomon, E. I., Zhou, J., Neese, F., Pavel, E. G., 1997, *Chem. Biol* 4, 795.
34. Brash, A. R., Boeglin, W. E., Chang, M. S., Shieh, B. H. 1996, *J. Biol. Chem.* 271, 20949.
35. Koeduka, T., Kajiwara, T., Matsui, K. 2007, *Curr. Microbiol.* 54, 315.