



OPEN

Three-dimensional structure of human cyclooxygenase (*hCOX*)-1

Morena Miciaccia^{1,5}, Benny Danilo Belviso^{2,5}, Mariacarla Iaselli¹, Gino Cingolani³, Savina Ferorelli¹, Marianna Cappellari⁴, Paola Loguercio Polosa⁴, Maria Grazia Perrone¹, Rocco Caliandro²✉ & Antonio Scilimati¹✉

The beneficial effects of Cyclooxygenases (COX) inhibitors on human health have been known for thousands of years. Nevertheless, COXs, particularly COX-1, have been linked to a plethora of human diseases such as cancer, heart failure, neurological and neurodegenerative diseases only recently. COXs catalyze the first step in the biosynthesis of prostaglandins (PGs) and are among the most important mediators of inflammation. All published structural work on COX-1 deals with the ovine isoenzyme, which is easier to produce in milligram-quantities than the human enzyme and crystallizes readily. Here, we report the long-sought structure of the human cyclooxygenase-1 (*hCOX*-1) that we refined to an R/R_{free} of 20.82/26.37, at 3.36 Å resolution. *hCOX*-1 structure provides a detailed picture of the enzyme active site and the residues crucial for inhibitor/substrate binding and catalytic activity. We compared *hCOX*-1 crystal structure with the ovine COX-1 and human COX-2 structures by using metrics based on Cartesian coordinates, backbone dihedral angles, and solvent accessibility coupled with multivariate methods. Differences and similarities among structures are discussed, with emphasis on the motifs responsible for the diversification of the various enzymes (primary structure, stability, catalytic activity, and specificity). The structure of *hCOX*-1 represents an essential step towards the development of new and more selective COX-1 inhibitors of enhanced therapeutic potential.

Prostaglandin endoperoxide H synthases (PGHSs), also known as cyclooxygenases (COXs), catalyze the conversion of free arachidonic acid (AA), released from cell membrane upon mitogen or mechanical stimuli, into Prostaglandin H₂ (PGH₂), that represents the committed step in the biosynthesis of prostaglandins, prostacyclin, and thromboxane¹. The known COX isoforms (COX-1 and COX-2) are the product of two different genes^{2,3}. The gene that encodes for COX-1 is considered a “housekeeping gene” due to its low levels of expression in most cell types. However, high levels of constitutive COX-1 expression have been detected in the gastric mucosa and platelets. In contrast to COX-1, COX-2 expression is typically inducible in response to pro-inflammatory events and growth factors². COX-1 and COX-2 are of particular interest because they are the major target of non-steroidal anti-inflammatory drugs (NSAIDs)³, which bind to the COX active site preventing that AA reaches the catalytic pocket¹ and, hence, prostaglandins biosynthesis. Prostaglandins, in particular PGE₂, are crucial mediators implicated in inflammation, angiogenesis and support the growth of several solid tumors^{4,5}. Therefore, inhibition of COXs is relevant to reduce inflammation, pain, and fever, as well as cancer progression.

Compared to COX-2, COX-1 involvement in tumorigenesis has been less studied⁵. Recent work demonstrates a direct involvement of COX-1 in breast, ovarian⁶, head and neck cancer, renal cell carcinoma, and hematological tumors^{5,7,8}. For all these cancers, overexpression of COX-1 could be exploited as a diagnostic biomarker^{9–12}.

X-ray crystallographic analysis has proven a powerful tool to investigate COX active site and in unraveling the binding mode of COX inhibitors. The ovine isoform of COX-1 (*oCOX*-1) has been crystallographically studied over a dozen times in complex with inhibitors and substrates¹³. More difficult has been working with the human isoform (*hCOX*-1). Herein, we describe the expression and purification of *hCOX*-1 from baculovirus-infected insect cells that yielded milligram-quantities of the enzyme. Ectopically expressed *hCOX*-1 yielded crystals suitable for structure determination by using X-ray crystallographic analysis.

¹Department of Pharmacy - Pharmaceutical Sciences, University of Bari “Aldo Moro”, Via E. Orabona 4, 70125 Bari, Italy. ²Istituto di Cristallografia, Consiglio Nazionale delle Ricerche, Via Amendola 122/o, 70126 Bari, Italy. ³Department of Biochemistry and Molecular Biology, Thomas Jefferson University, 1020 Locust Street, Philadelphia, PA 19107, USA. ⁴Department of Biosciences, Biotechnologies, and Biopharmaceutics, University of Bari “Aldo Moro”, Via E. Orabona 4, 70125 Bari, Italy. ⁵These authors contributed equally: Morena Miciaccia and Benny Danilo Belviso. ✉email: rocco.caliandro@ic.cnr.it; antonio.scilimati@uniba.it

Results

***hCOX-1* expression and purification.** The production of milligram quantities of correctly folded and active *hCOX-1* is vital for the structural, biochemical and pharmacological analysis of human COX-1, a key target of a plethora of pharmaceuticals, clinically used to treat several human diseases. In this study, we used the BacPAK™ Baculovirus Expression System to express large quantities of folded and post-translationally modified *hCOX-1* suitable for structural analysis.

hCOX-1 is 599 amino acids long and includes at the *N*-terminus a 24-amino acid signal peptide that targets the protein to the endoplasmic reticulum for post-translational modification. We aimed to construct a recombinant *hCOX-1* baculovirus that expresses the protein suitable for nickel affinity chromatography purification by using the BacPAK Baculovirus Expression System. The *hCOX-1* DNA was PCR amplified from a previously generated pFastBac-*hCOX-1* plasmid that also contains an 8xHistidine tag (8XHis-tag) placed just after the signal peptide, followed by a 7-amino acid TEV protease cleavage site. The amplification was carried out using the proofreading Phusion HF DNA polymerase and specific primers containing at their 5' ends BglII and NotI restriction sites. The amplified DNA was evaluated by agarose gel electrophoresis, showing, as expected, a single product of about 1800 bp (Supplementary Fig. 1). The purified products were digested by BglII and NotI restriction enzymes and ligated with the shuttle plasmid pBacPAK9. Then, the ligation products were amplified in bacteria and, once extracted, they were restriction digested to assess the presence of the target gene; the correct sequence of the insert was verified by sequencing of both strands. The cloning efficiency was about 100% (Supplementary Fig. 2). To generate the recombinant expression virus, the transfer vector pBacPAK9 containing 8XHis-tag *hCOX-1* cDNA was transfected into *Spodoptera frugiperda* cells along with Bsu36 I-digested BacPAK6 Viral DNA; an in vivo homologous recombination between the plasmid DNA and viral DNA occurred, producing recombinant baculovirus¹⁴.

To isolate a pure clone of a recombinant virus, the supernatants from co-transfection containing progeny viruses were collected and subjected to plaque assay to produce individual plaques (plaque-pick). All plaque-picks were first amplified to produce Virus Stock I (VSI), then viral DNA was extracted and analyzed by PCR, confirming that all clones were positive¹⁵ (Supplementary Fig. 3).

Hereafter, VSI was amplified twice, according to the manufacturer's instructions, to maximize the yield of recombinant protein production. A scale-up production and purification protocol for the *hCOX-1* isoenzyme was substantially the same as previously described for *oCOX-1*¹⁶.

SDS-PAGE analysis of metal affinity chromatography fractions revealed *hCOX-1* eluted around 0.25 M imidazole; the eluted protein was highly pure as no low molecular weight species were detectable after Blue Coomassie staining of the gel (Supplementary Fig. 4).

hCOX-1 protein fractions were pooled and further purified by performing in-solution cleavage, with direct addition of TEV_{6His} protease to the eluted fraction. Following cleavage, the protein solution was loaded on nickel beads, where his-tag and TEV_{6His} were recaptured on the resin. Untagged *hCOX-1* was collected in the flow-through, buffer exchanged and concentrated to 10.4 mg ml⁻¹.

Stability, catalytic activity, and specificity of *oCOX-1* and *hCOX-1*. The availability of both ovine and human COX-1 let us to verify if their stability, catalytic activity, and specificity are similar or different. In particular, their thermal and chemical stability was determined at +4, -20, and -80 °C. Within three months, both enzymes stored at the first two temperatures exhibited an impaired catalytic activity, that in turn remains unchanged at -80 °C for the same period of time. Their specific activity, instead, was determined in the presence of arachidonic acid and one of the different well-known NSAIDs-COX-1 inhibitors (Table 1), such as mofezolac (a diarylisoxazole sold in Japan as Disopain® to treat rheumatoid arthritis and that could be worldwide repurposed for neurodegenerative diseases with a marked neuroinflammatory component)^{17,18}; the ibuprofen, one of the arylpropanoic acids (profen-series) representative of the most commonly prescribed NSAIDs, and indomethacin belonging to the chemical class of indole-acetic acids actually used for treating mild to moderate acute pain in adult. The specific activities found were a little higher for *hCOX-1* than *oCOX-1*, even if the two values are very similar in all three cases. The similarity could be ascribed to the high sequence identity (92%) among the two proteins (Fig. 1).

Mofezolac inhibited *oCOX-1* to a greater extent than the corresponding human enzyme (*oCOX-1* IC₅₀ = 9.2 μM; *hCOX-1* IC₅₀ = 23 μM), ibuprofen *oCOX-1* IC₅₀ and *hCOX-1* IC₅₀ are comparable being 31 and 44 μM, respectively. The indomethacin inhibitory potency towards *hCOX-1* was found to be of two orders of magnitude higher than the ovine enzyme (*oCOX-1* IC₅₀ = 158 μM and *hCOX-1* IC₅₀ = 2.9 μM) and with an enzyme preference index (EPI) for *hCOX-1* equal to 54 (Table 1).

***hCOX-1* crystallization experiments.** *hCOX-1* crystallization experiments were carried out by following the crystallization procedure reported for *oCOX-1*¹⁹. The best diffracting crystals (Supplementary Fig. 5) were obtained in the presence of 0.5–0.6 M lithium chloride, 10–11 mg/ml of protein concentration, 0.7 M sodium citrate, and 1:1 molar ratio between protein and heme (Fe³⁺-protoporphyrin IX).

***hCOX-1* crystal structure.** *hCOX-1* crystals were subjected to X-ray diffraction experiments. Data collection and refinement statistics are shown in Table 2. The data resolution cutoff of 3.36 Å allowed meeting requirements of $\langle I/\sigma(I) \rangle > 1.5$ and $CC_{1/2} > 50\%$ for the high-resolution shell. The data resolution limit and the high value of R_{sym} in the high-resolution shell could not be improved by varying crystallization conditions and trying different crystals. The *hCOX-1* crystal shows P_{6₅22} symmetry, and its asymmetric unit contains a single protein molecule (Fig. 2a). According to the PISA server²⁰, the protein in the asymmetric unit interacts with a symmetry-related molecule (symmetry operation $y, x, -z-1/3$) to form a dimeric biological unit (Fig. 2b), whose

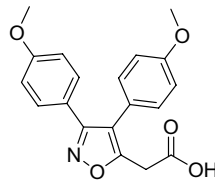
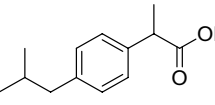
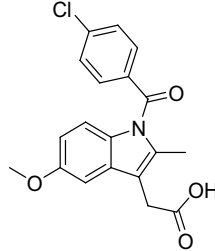
Substrate	Inhibitor	IC ₅₀ (μM) ^a		EPI ^b	Specific activity (μmol min ⁻¹ mg ⁻¹)	
		<i>o</i> COX-1	<i>h</i> COX-1		<i>o</i> COX-1	<i>h</i> COX-1
Arachidonic acid (100 μM)	 mofezolac	9.2	23	0.4	3.5	4.5
	 ibuprofen	31	44	0.7	5.9	8.2
	 indomethacin	158	2.9	54	3.9	4.4

Table 1. *o*COX-1 IC₅₀ and *h*COX-1 IC₅₀ and the specific activity of the two enzymes. ^aDetermined by Cyclooxygenase O₂ Electrode Assay. ^bEPI (enzyme preference index) = *o*COX-1 IC₅₀/*h*COX-1 IC₅₀.

	1	10	20	30	40	50	60	70	80																																																																													
hCOX1	MSR	SLLL	WFL	LL	LLLL	PPL	PVLL	AAD	PGAP	T	VN	PC	CC	YPC	QH	QC	IC	VR	FG	LD	RY	QC	DC	TR	TG	YS	SG	PNC	TIP	G	L	WT	WL	RN	SL	RP	SP	SFT																																																
<i>o</i> COX1	MSR	OS	IS	LR	WFL	LL	LLLL	SP	SP	VFS	.	AD	PG	AP	T	VN	PC	CC	YPC	QH	QC	IC	VR	FG	LD	RY	QC	DC	TR	TG	YS	SG	PNC	TIP	G	L	WT	WL	R	TL	RP	SP	SFI																																											
	90	100	110	120	130	140	150	160	170																																																																													
hCOX1	HFL	LTH	GR	W	F	V	N	A	T	F	I	R	E	M	L	M	L	V	L	T	V	R	S	N	L	I	P	S	P	P	T	Y	N	S	A	H	D	Y	I	S	W	E	S	F	S	N	V	S	Y	T	R	I	L	P	S	V	P	R	D	C	P	T	M	G	T	K	G	K	Q	L	P	D	A	Q	L	A	R									
<i>o</i> COX1	HFL	LTH	GR	W	F	V	N	A	T	F	I	R	D	T	L	M	L	V	L	T	V	R	S	N	L	I	P	S	P	P	T	Y	N	I	A	H	D	Y	I	S	W	E	S	F	S	N	V	S	Y	T	R	I	L	P	S	V	P	R	D	C	P	T	M	D	T	K	G	K	Q	L	P	D	A	E	L	S	R									
	180	190	200	210	220	230	240	250	260																																																																													
hCOX1	R	F	L	L	R	R	K	F	I	P	D	P	G	T	N	L	M	F	A	F	A	Q	H	F	T	H	Q	F	F	K	T	S	G	K	M	G	P	G	F	T	K	A	L	H	G	V	D	L	G	H	I	Y	G	D	N	L	R	Q	Y	L	R	L	F	K	D	G	K	K	Y	Q	V	L	D	G	E	M	Y	P	P	S	V	E	E	A		
<i>o</i> COX1	R	F	L	L	R	R	K	F	I	P	D	P	G	S	T	N	L	M	F	A	F	A	Q	H	F	T	H	Q	F	F	K	T	S	G	K	M	G	P	G	F	T	K	A	L	H	G	V	D	L	G	H	I	Y	G	D	N	L	R	Q	Y	L	R	L	F	K	D	G	K	K	Y	Q	V	L	D	G	E	M	Y	P	P	S	V	E	E	A	
	270	280	290	300	310	320	330	340	350																																																																													
hCOX1	P	V	L	M	H	Y	P	R	G	I	P	P	Q	S	M	A	V	G	Q	E	V	F	G	L	L	P	G	L	M	L	Y	A	T	I	W	L	R	E	H	N	R	V	C	D	L	L	K	A	E	H	P	T	W	G	D	E	Q	L	F	Q	T	A	R	L	I	L	I	G	E	T	I	K	I	V	E	E	Y	V	Q	L	S	G	F	L	Q	L
<i>o</i> COX1	P	V	L	M	H	Y	P	R	G	I	P	P	Q	S	M	A	V	G	Q	E	V	F	G	L	L	P	G	L	M	L	Y	A	T	I	W	L	R	E	H	N	R	V	C	D	L	L	K	A	E	H	P	T	W	G	D	E	Q	L	F	Q	T	A	R	L	I	L	I	G	E	T	I	K	I	V	E	E	Y	V	Q	L	S	G	F	L	Q	L
	360	370	380	390	400	410	420	430	440																																																																													
hCOX1	K	F	D	P	E	L	L	F	G	A	Q	F	Q	Y	R	N	R	I	A	E	F	N	H	L	Y	H	W	H	P	L	M	P	D	S	F	R	V	G	S	O	E	Y	S	Y	E	Q	L	F	N	T	S	M	L	V	D	Y	G	V	E	A	L	V	D	A	F	S	R	Q	A	G	R	I	G	G	R	N	D	H	H	I	L	H	V	A		
<i>o</i> COX1	K	F	D	P	E	L	L	F	G	A	Q	F	Q	Y	R	N	R	I	A	E	F	N	O	L	Y	H	W	H	P	L	M	P	D	S	F	R	V	G	S	O	E	Y	S	Y	E	Q	L	F	N	T	S	M	L	V	D	Y	G	V	E	A	L	V	D	A	F	S	R	Q	A	G	R	I	G	G	R	N	D	H	H	I	L	H	V	A		
	450	460	470	480	490	500	510	520	530																																																																													
hCOX1	D	V	I	R	E	S	R	V	L	R	L	Q	P	F	N	E	Y	R	K	R	F	G	M	K	P	Y	T	S	F	Q	E	L	G	E	K	E	M	A	E	L	E	Y	G	D	I	D	A	L	E	F	Y	P	G	L	L	E	K	C	H	P	N	S	I	F	G	E	S	M	I	E	H	G	A	P	F	S	L	K	G	L	L	G	N	P	I	
<i>o</i> COX1	D	V	I	R	E	S	R	V	L	R	L	Q	P	F	N	E	Y	R	K	R	F	G	M	K	P	Y	T	S	F	Q	E	L	G	E	K	E	M	A	E	L	E	Y	G	D	I	D	A	L	E	F	Y	P	G	L	L	E	K	C	H	P	N	S	I	F	G	E	S	M	I	E	H	G	A	P	F	S	L	K	G	L	L	G	N	P	I	
	540	550	560	570	580	590	600																																																																															
hCOX1	C	S	P	E	Y	W	K	P	S	T	F	G	G	E	V	G	F	N	I	V	K	T	A	T	L	K	K	V	C	L	N	T	K	T	C	P	Y	V	S	F	R	V	P	D	S	Q	D	D	G	P	A	V	E	R	P	S	T	E	L																											
<i>o</i> COX1	C	S	P	E	Y	W	K	P	S	T	F	G	G	E	V	G	F	N	I	V	K	T	A	T	L	K	K	V	C	L	N	T	K	T	C	P	Y	V	S	F	H	V	P	D	S	R	Q	E	D	R	P	G	V	E	R	P	S	T	E	L																										

Figure 1. Alignment of human and ovine COX-1. Amino acid sequences of COX-1 from human, *h*, (NM_000962.4) and ovine, *o*, (NM_001009476.1) were aligned with the ClustalW program. The alignment was performed at NPS@ Web server of the PBIL using GONNET weight matrix and formatted with ESPrpt 3.0²¹. The red boxes include identical amino acids in both sequences. Unconserved residues are included in the white boxes. Among these amino acids, the proline residues present only in *o*COX-1 are shown in the green boxes, while the glycine residues present only in *h*COX-1 are displayed in the yellow boxes.

Data collection		
Wavelength (Å)	0.9795	
Space group	P 6 ₂ 22	
Cell constants a, c (Å)	182.11 103.24	
	Before STARANISO	After STARANISO
Resolution range (Å)	91.05–3.36 (3.45–3.36)	91.05–3.36 (3.48–3.36)
Completeness (%)	99.9 (99.9)	81.1 (25.16)
Redundancy	39.1 (42.2)	39.1 (42.2)
Total number of reflections	582,091 (45,649)	581,081 (61,045)
Unique reflections	14,886 (1081)	14,838 (363)
R sym (%)	0.467 (4.244)	0.493 (3.356)
R meas (%)	0.473 (4.295)	0.499 (3.578)
<I/σ(I)>	14.3 (1.4)	14.3 (1.5)
CC _{1/2} (%)	99.9 (63.6)	99.9 (69.0)
Wilson B-factor (Å ²)	94.32	94.47
Refinement		
R/Rfree (%)	20.82/26.37 (29.17 /36.30)	
MolProbity score	1.59	
Clashscore	7	
Average B-factor (Å ²) (protein atoms)	100.54	
Average B-factor (Å ²) (water atoms)	61.12	
Average B-factor (Å ²) (heterogen atoms)	111.55	
Protein atoms	4518	
Water molecules	17	
Heterogen atoms	155	
RMSD bonds length (Å)	0.002	
RMSD bond angles (°)	0.53	
Ramachandran favored (%)	96.76	
Ramachandran allowed (%)	3.24	
Ramachandran outlier (%)	0.00	

Table 2. Data collection and refinement statistics of the *h*COX-1 crystal structure (PDB code 6Y3C). Outer shell statistics are in parenthesis. For the sake of completeness, statistics before and after anisotropy correction by the STARANISO server are shown.

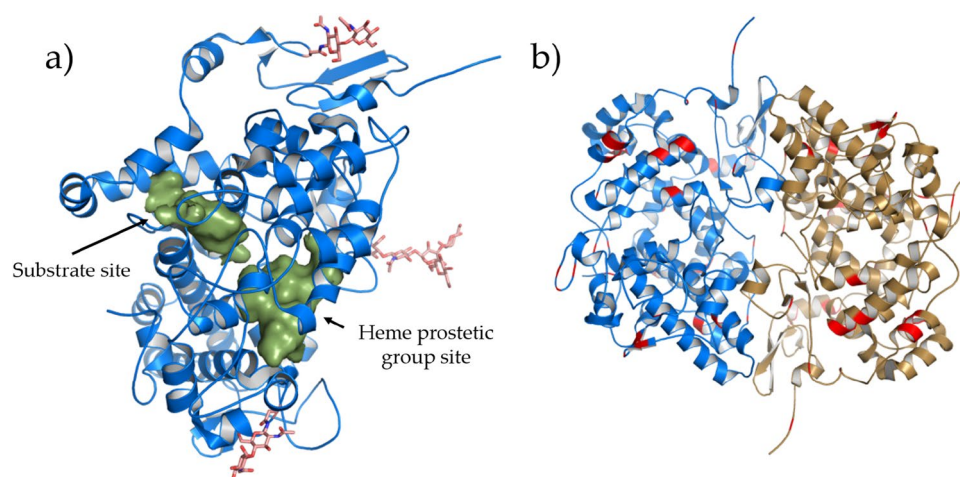


Figure 2. *h*COX-1 crystal structure (PDB code 6Y3C). (a) Asymmetric unit of *h*COX-1: protein is shown in blue cartoon, and glycosylation along with glycosylation-linked asparagine residues are in pink stick representation. The solvent excluded surface (determined at solvent radius 1.4 Å) of heme and substrate site is shown in green. (b) Biological unit of *h*COX-1: the two protein molecules forming a functional dimer are shown in cartoon representation in blue and sand color. Different residues between *h*COX-1 and *o*COX-1 are in red color.

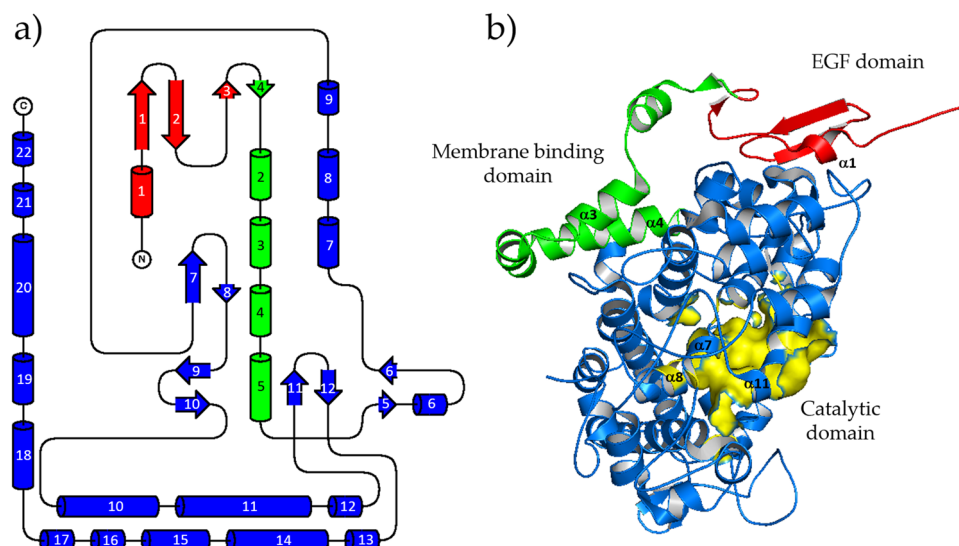


Figure 3. *hCOX-1* folding. (a) *hCOX-1* topology diagram: secondary structure elements are shown in red, green, and blue for EGF, membrane-binding domain, and catalytic domain, respectively. β -strands and α -helices are shown as arrows and cylinders, respectively; (b) *hCOX-1* domains are shown in cartoon representation in the same color code of (a). Solvent excluded surface of the catalytic domain is in yellow color.

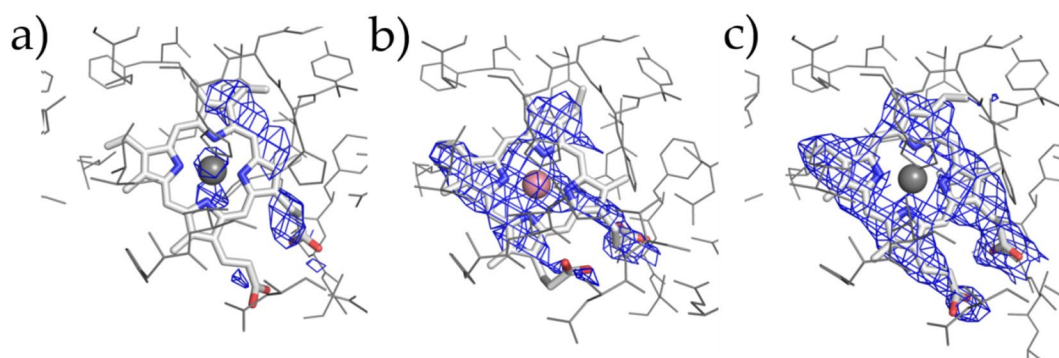


Figure 4. Polder map of heme molecule. Protein residues are shown in gray lines and heme molecules are as gray sticks. (a–c) show the polder map contoured at 2.5σ above background for *hCOX-1* (PDB code 6Y3C), *oCOX-1* (PDB code 1U67), and *oCOX-1* (PDB code 5U6X).

interaction interface is about 26% of the solvent-exposed area of the monomer. The high free energy calculated for dimer dissociation (17.6 kcal/M) suggests that such a biological unit is thermodynamically stable. Diffraction data quality is enough to model the expected *N*-acetyl glucosamine (NAG) chain on N68, N410, N144 that are shown in stick representation in Fig. 2a. The validation procedure of the Protein Data Bank server shows quite a constant electron density coverage along the entire protein chain. Poorly covered residues having $RSRZ > 2$, where $RSRZ$ is the difference between the experimental electron density and that obtained by the model, normalized by residue type and resolution, are only 4% of the all protein residues and are located at the *N*-terminus and in the regions that span 130–150 and 200–213. Such a density coverage points out the good agreement between data and the modeled protein chain. The final model has good geometry relatively to the resolution, with 18% of residues having one geometry outliers, and only 1% showing two outliers.

Secondary structure elements and their mutual orientation are shown in Fig. 3a. *hCOX-1* protein fold reveals the peculiar arrangement of an epidermal growth factor (EGF) and a membrane-binding domain placed on the surface of a catalytic domain, quite far from the heme site Fig. 3b. With respect to the EGF domain, the membrane-binding domain has more interactions with the catalytic domain and, particularly, with the region forming the substrate pocket of the protein. Such protein folding follows that of the *oCOX-1*, the protein having the highest sequence identity with *hCOX-1*, and that of the *hCOX-2*.

The heme molecule, which is important for COX activity, was added at 1:1 molar ratio with *hCOX-1* monomer before crystallization. Such a stoichiometric quantity of heme promoted crystallization resulted in poor electron density in the region where heme molecule is expected. Refining the structure with the heme molecule resulted in an increase of R/R_{free} from 20.82/26.37 to 21.50/26.58. In Fig. 4, the polder map calculated by omitting the

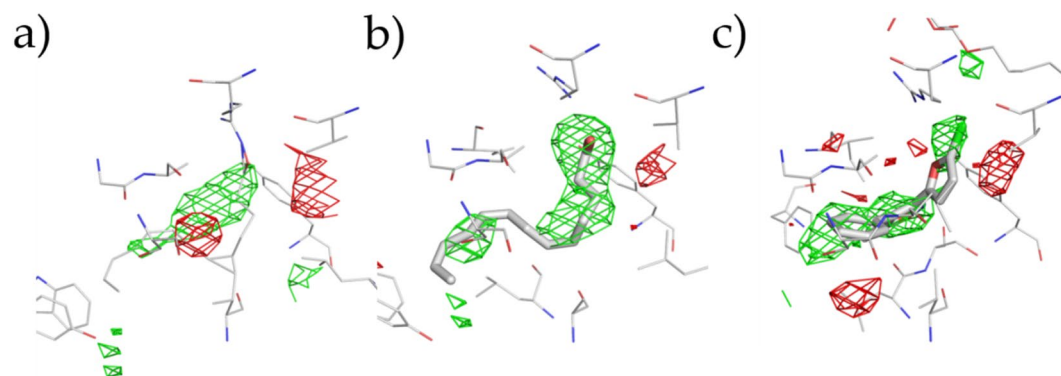


Figure 5. Fourier difference (Fo-Fc) electron density map contoured at $>|3\sigma|$ (positive and negative peaks are in green and red mesh, respectively) in the substrate site for (a), 6Y3C (*hCOX-1*), (b) 1U67 (*oCOX-1*:AA complex crystal), and (c) 5U6X (*oCOX-1*:P6 complex crystal). Enzymatic protein residues are in gray lines and AA or P6 in the substrate site in gray sticks. Fourier difference electron density map was obtained by removing ligand molecules from the structures.

heme group in our structure is compared with those calculated in the same way for two previously determined *oCOX-1* structures [1U67 (P 6_5 22 symmetry) and 5U6X (P 6_5 symmetry)]. The density for heme is much weaker in the *hCOX-1* than the other two structures, confirming the poor occupancy of this prosthetic group in our structure. Interestingly, the polder electron density covers only partially the heme molecule, mainly, the portion that interacts with protein residues. Such a feature is compatible with a flexible molecule rather than a stable molecule at low crystallographic occupancy. To make sure that such electron density features related to the heme molecule are not an artifact due to an incorrect choice of space group, data were reprocessed in the alternative space group P 6_5 , suggested by the indexing process, where the two protein chains can have different conformations, but no difference was observed. Unable to confirm the presence of the heme in our crystal structure, we decided not to include such a prosthetic group in the final model.

Another region of the enzyme that deserves attention is the substrate site. Even though our *hCOX-1* crystals were obtained in the absence of molecules that bind the substrate site, a peak of difference (Fo-Fc) electron density can be observed in this region (Fig. 5a). Fo-Fc density map was compared with that obtained by removing arachidonic acid (COXs endogenous substrate) and P6 (highly selective COX-1 inhibitor)^{22–24} ligand from 1U67 (Fig. 5b)¹³ and 5Y6X (Fig. 5c)¹³, respectively. Although less defined, the peak in our structure shows similar shape and orientation to those in the other two structures. All attempts to fit and refine arachidonic acid, or molecules present in the crystallization cocktail, were unsuccessful. This prompted us not to model this electron density peak in the final structural model.

Comparative analysis of *hCOX-1*, *oCOX-1* and *hCOX-2* crystal structures. There are twenty-nine crystal structures of *oCOX-1* deposited in the Protein Data Bank (thirteen of which crystallized in the space group I 222, nine in the space group P 6_5 , and seven in P 6_5 22) and seven crystal structures for *hCOX-2* (all crystallized in the space group I 222). Differences in protein conformation among these crystal structures could be due to different crystallization conditions, crystal packing, or protein sequence. While the first two are properties of the crystal state, the latter one is related to the intrinsic conformation of the protein. In the case of COX proteins, deciphering the intrinsic conformation of the human enzyme is particularly important for drug design and development. COX-1 is a functional dimer, and dimerization is strictly required for structural integrity and activity²⁵. Nonetheless, it remains unclear if the COX-1 dimer is built by conformationally equivalent or *quasi*-equivalent protomers of COX. With this rationale in mind, we performed a comparative analysis of our *hCOX-1* with the crystal structures of *oCOX-1* and *hCOX-2* deposited in the PDB using different metrics and considering the chains of the biological unit of each structure separately.

As a first step, we compared the experimental coordinate errors of the structural models, as estimated by the diffraction-data precision indicator (DPI) (Supplementary Fig. 6). Our *hCOX-1* crystal structure holds a coordinate error of 0.59 Å, mainly due to the medium data resolution (3.36 Å). Comparable errors are shown by *oCOX-1* crystal structures with P 6_5 22 (average error of 0.41 Å) and I 222 symmetry (average error of 0.46 Å). In this latter case, the spread of DPI values is very large, ranging from 0.10 Å for 2AYL and 1Q4G to 1.24 Å for 1PGF and 1PGG. Crystal structures of *hCOX-2* and *oCOX-1* with P 6_5 symmetry are systematically more precise, having, in both cases, an average error of 0.28 Å. The systematic trend of experimental errors among COXs reflects the quality of diffraction data, and suggests the following trend: first, *hCOX-1* is the most challenging isoform to crystallize; second, COX-2 is intrinsically more stable than COX-1 and hence more likely to crystallize; third, among *oCOX-1* structures deposited in the database, there is an empirical correlation between model accuracy and crystal symmetry with the COX-1 structure solved in space P 6_5 being the most accurate, followed by P 6_5 22 and finally I 222.

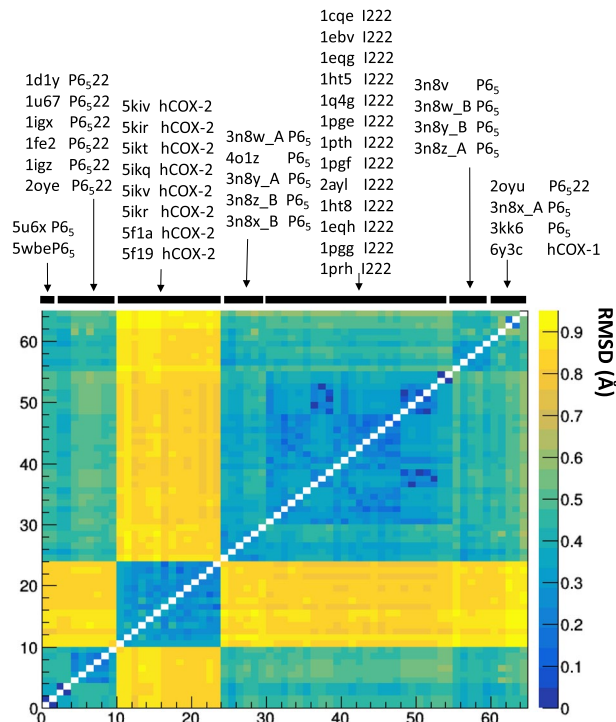


Figure 6. Matrix of the Ca pairwise RMSD values calculated between individual chains of the 37 structures of *h*COX-1, *o*COX-1, and *h*COX-2 present in the PDB. Chains are ordered based on a hierarchical clustering algorithm having RMSD as distance metrics. They are labelled on the top by using the PDB code, the type (*h*COX-1 or *h*COX-2), the space group (only for *o*COX-1), and the chain letter (A or B are reported only if they do not fall in the same cluster).

Comparison based on Cartesian coordinates. We superimposed each independent chain of *h*COX-1, *o*COX-1, and *h*COX-2 with each other to generate a matrix (65 chains \times 65 chains) of overall Ca RMSD, which was subjected to hierarchic clustering. Despite the large heterogeneity of diffraction data quality of COXs crystals (Supplementary Fig. 6), the crystal structures that are homogeneous by protein, source, and crystallographic symmetry were grouped consistently in the resulting matrix (Fig. 6), as they have similar RMSD values. It can be noted that the two chains of the same crystal structure are grouped in most of the cases, indicating that the functional dimer is preferentially made of two conformationally-equivalent units. Moreover, Fig. 6 clearly shows the cluster of *h*COX-2 chains, which are very different from the *o*COX-1 chains, and the clusters of *o*COX-1 chains with P₆₅22 and I 222 symmetry, pointing out the high structural similarity (inter-cluster RMSD < 0.4 Å) of these three sets of structures.

*o*COX-1 structures with P₆₅ symmetry are spread into three clusters, indicating the higher structural heterogeneity related to this crystal packing, where two symmetry-independent chains form the biological unit. A further cluster has slightly higher inter-cluster RMSD values (between 0.4 and 0.6 Å) and includes the following outliers: our *h*COX-1 crystal structure, 2OYU, a *o*COX-1 structure with P₆₅22 symmetry, the chain A of the structure 3N8X and the structure 3KK6, both *o*COX-1 crystallized in P₆₅. The papers related to these two latter outliers show that data merging by using P₆₅ and P₆₅22 provides comparable R_{sym} values^{26,27}, a feature that could explain why these structures group with *o*COX-1 structures crystallized in P₆₅22 rather than with those crystallized in P₆₅. Interestingly, the sum of all RMSD values obtained in the case of the human isoform is the highest compared to those obtained for the other COX-1 structures (Supplementary Fig. 7). In the case of pairwise comparison between *o*COX-1 structures and our *h*COX-1, the average of the RMSD values is \sim 0.5 Å, and it is \sim 0.8 Å in the case of comparison with *h*COX-2 crystal structures, despite the low sequence identity (60%).

An in-depth view of the above structural deviations is provided in Fig. 7, where the *h*COX-1 crystal structure is used as a reference for pairwise superposition, and the residue-by-residue averaged RMSD values for each group of structures are plotted against the residue number. It can be noted that the RMSD values of A (red) and B (black) chains do not show significant differences within the same group. A second observation regards specific regions of the protein affected by high RMSD values. By focusing on the left column (superposition between *h*COX-1 and *o*COX-1 structures), it is possible to observe that the region between 270 and 290 in the catalytic domain is systematically affected by high RMSD values regardless the crystal symmetry. We verified that this very flexible region is mainly responsible for the clustering shown in Fig. 6, as it holds the same conformation in crystal structures within the same cluster, and slightly different conformations for inter-cluster structures (Supplementary Fig. 8).

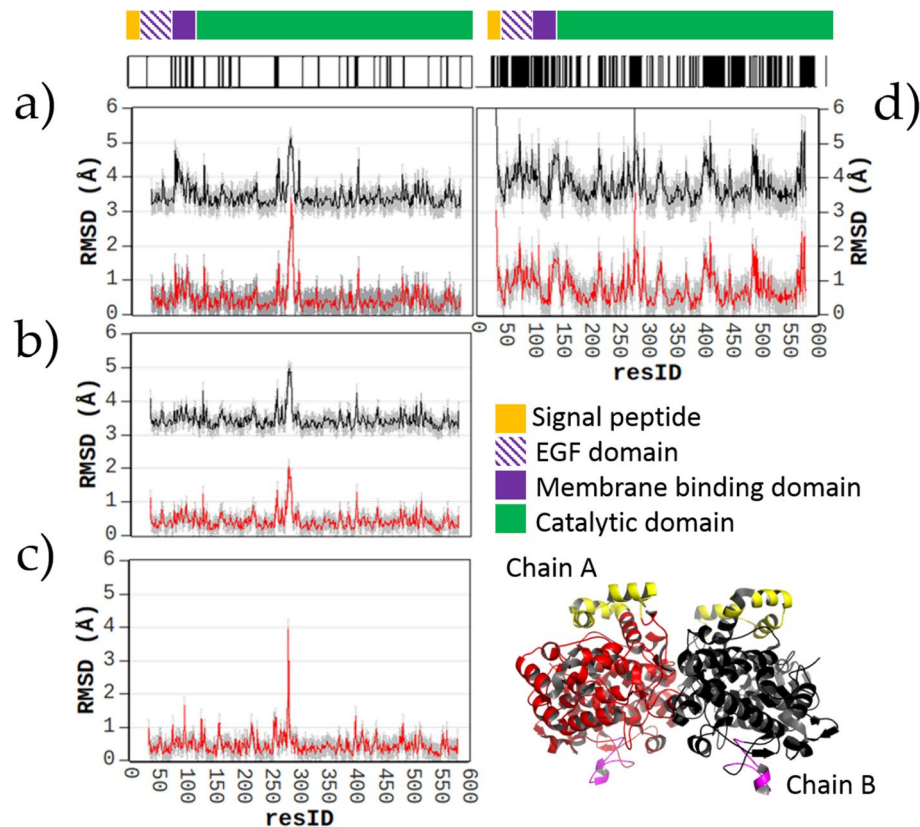


Figure 7. Average residue-by-residue RMSD values calculated after superposition of *hCOX-1* (6Y3C) with (a) *oCOX-1* crystallized in the I 222 space group (PDB codes 1Q4G, 2AYL, 1PTH, 1PRH, 1PGG, 1PGE, 1PGE, 1HT8, 1EQH, 1HT5, 1CQE, 1EBV, 1EQG), (b) *oCOX-1* crystallized in the P 6₅ space group (PDB codes 3KK6, 3N8Z, 3N8X, 5WBE, 3N8V, 4O1Z, 3N8Y, 3N8W, 5U6X), (c) *oCOX-1* crystallized in the P 6_{5,22} space group (PDB codes 1U67, 2OYE, 2OYU, 1IGZ, 1DIY, 1FE2, 1IGX), and (d) *hCOX-2* crystallized in the I 222 space group (PDB codes 5F19, 5IKQ, 5IKT, 5KIR, 5F1A, 5IKR, 5IKV, 5KIR) to the *hCOX-1* (6Y3C) crystal structure. Standard deviation for each residue is shown in gray color. RMSD values related to chain A and chain B (nomenclature refers to that adopted for 5U6X shown at bottom right in the figure) are shown in red and black, respectively. Here, regions 70–110 and 270–290 are shown in yellow and magenta color. In the case of *oCOX-1* crystallized in P 6_{5,22}, the biological unit contains a homo-conformational dimer, therefore only chain A is shown. The numbering of residues refers to *hCOX-1* structure. On the top of the figure, variations between *hCOX-1* and *oCOX-1* (left) and *hCOX-2* (right) are shown, along with *hCOX-1* domains.

Interestingly, such a region is poor of residue variations among *hCOX-1* and *oCOX-1*; thus, conformational changes are unlikely related to differences in the protein sequence. Differently from the region 270–290, residues between 70 and 110, located in the first two α -helices of the membrane-binding domain, show high RMSD values only compared to *oCOX-1* protein crystallized in the space group I 222. We verified that the same feature arises when comparing *oCOX-1* crystal structures with P 6_{5,22} (the same symmetry of our *hCOX-1* crystal structure) or P 6₅ symmetry with those with I 222 symmetry (Supplementary Fig. 8). Thus, this conformational variation was ascribed to the different effect that crystal packing exerts on the protein conformation in an orthorhombic versus hexagonal lattice. It is worth noting that the same effect does not exist if our *hCOX-1*, or *oCOX-1* with P 6_{5,22} symmetry if compared with *oCOX-1* crystallized in the P 6₅ space group. Such a result can be explained by considering that these structures have similar unit cells and crystal packing and differ only in the extension of the asymmetric unit. In addition to crystal packing, also differences in ligand interactions could contribute to the high value of RMSD found in the region 70–110. Indeed, ligands found for *oCOX-1* in I 222 are localized in different way with respect to the citrate ions found in our structure (Supplementary Fig. 9, Supplementary Discussion 1). Similar differences can be observed also by comparing our structure with *hCOX-2*. It is worth noting that, since differences in crystal packing and ligand interaction are not directly related to the protein sequence, the high values of RMSD observed in the region 70–110 are unlikely related to intrinsic protein conformation, thus they should not affect inhibitor binding.

When comparing *hCOX-1* and *hCOX-2* (Fig. 7d), it can be noted that in addition to the regions 270–290 and 70–110, highlighted in the comparison between *hCOX-1* and *oCOX-1*, also regions 125–175 (N-terminus of the catalytic domain), 400–420 and 470–520 both included in the catalytic domain show large deviations.

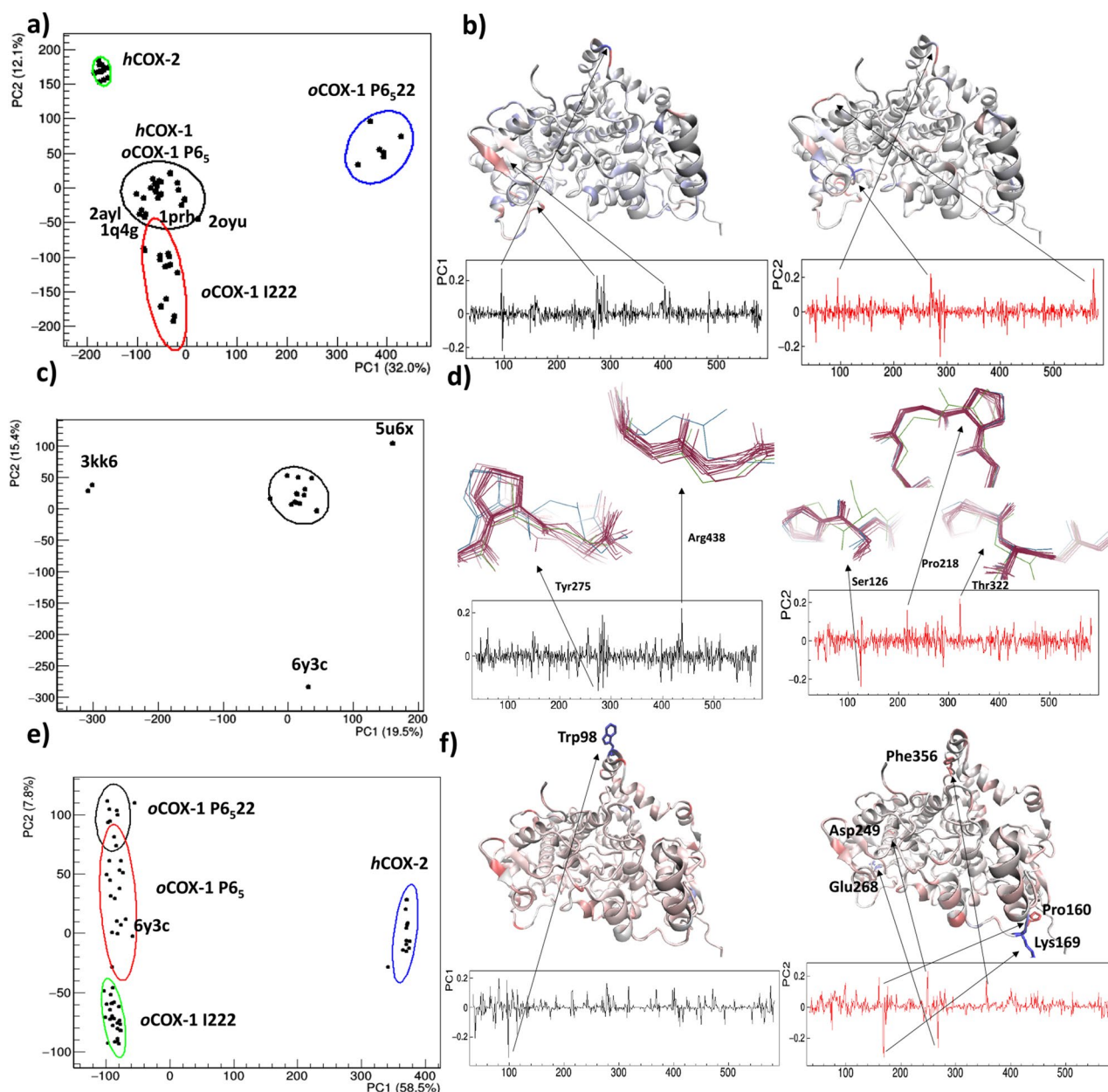


Figure 8. Comparison of *hCOX-1*, *oCOX-1*, and *hCOX-2* structural models by principal component analysis applied on PAV (a–d) and SASA (e,f) profiles. A restricted number of COX-1 structural models is considered in (c,d). [On the left is shown the score plot of the first two principal components, where representative points are grouped according to a hierarchic clustering algorithm. 85% confidence level ellipses are shown, and PDB codes associated with representative points are reported. The percentage of the total data variance explained by each principal component is shown on the axes. On the right are shown the loadings of the first (black) and second (red) principal components, the structural model 6Y3C colored according to the loading values, from red for higher positive to blue for higher negative (b,f), and the superposed structural models with 6Y3C and 3KK6 colored respectively in green and blue (d). Arrows point to protein residues having the highest and lowest loadings values. Residues falling in these regions are drawn in “bonds” representation in (f)].

Comparison based on backbone dihedral angles. In addition to the analysis based on Cartesian coordinates, we performed a comparative analysis based on backbone dihedral angles more focused on differences due to hinge motion. Profiles describing the residue-by-residue hinge flexibility have been calculated in terms of the Protein Angular Value (PAV). The 74×550 (chains \times residues) matrix formed by the PAV profiles was processed by principal component analysis (PCA) to reduce the dimensionality of the problem and assess the separation among structural models in the 2D space defined by the first two principal components. Data points, representing structural models, can be grouped according to their position in the PC2 vs PC1 scores plot (Fig. 8a). We found that the chains belonging to the same structure have very close PC1/PC2 coordinates

Residue number	Structural element	Structural region	Principal components interested	Crystal structures mainly separated
95, 96	α_3 - α_4	Substrate site	PC1, PC2	<i>o</i> COX-1 P _{6₅} 22
269, 270	α_6 - α_7 loop	Close to heme site	PC2	<i>h</i> COX-2
274, 281, 287	α_6 - α_7 loop	Close to Asn410 Glycosylation site	PC1	<i>o</i> COX-1 P _{6₅} 22, <i>o</i> COX-1 I222
400, 401, 402	β_{12}	NAG site	PC1	<i>o</i> COX-1 P _{6₅} 22
574, 575	C-term loop	C-term	PC2	<i>h</i> COX-2

Table 3. Key residues determined by backbone dihedral angles comparative analysis applied on *h*COX-1, *o*COX-1, and *h*COX-2 crystal structures. In the table are reported: the residue number, the structural elements and the structural region where the residues are located, the principal components having high loadings values in the specific residue range (fourth column) and the crystal structures mainly separated by applying PCA to the specific residue range (last column).

in the score plot to such an extent that the two chains cannot be distinguished. This result points out that each analyzed structure is made by two chains having the same hinge points, thus enforcing the hypothesis that their biological unit contains a homo-conformational dimer. Moreover, Fig. 8a shows that *h*COX-2 structures are consistently separated from COX-1 structures and are grouped in a very narrow cluster, confirming their higher structural homogeneity already observed by RMSD analysis. Instead, COX-1 structures exhibit a larger spread, with a huge separation of *o*COX-1 structures with P_{6₅}22 symmetry, located at high positive PC1 values, and of *o*COX-1 structures with I 222 symmetry, located at high negative PC2 values. A cluster, including *o*COX-1 structures having symmetry P_{6₅}, our *h*COX-1 structure and the outlier structures 2OYU, 1PRH, 2AYL, and 1Q4G is located at PC1 and PC2 values. Residues that are responsible for the above separation, as summarized in Table 3, can be identified among the highest values (positive or negative) of the PCA loadings (Fig. 8b). The effect of each individual range of hinge residues on structural separation has been determined by applying PCA on the PAV matrix restricted to that range. With this method, we were able to identify the residues 269–270 and 574–575 as main responsible for the separation of *h*COX-2 crystal structures, the residues 95–96 and 400–402 as responsible for the separation of *o*COX-1 P_{6₅}22 structures and the region 274–287, which is able to separate *o*COX-1 structures with P_{6₅}22 and I 222 symmetry. This latter region is interested in inter-molecular hydrogen bond interactions with the glycosylation site Asn410 in *o*COX-1 crystal structures (see Supplementary Table 1).

A comparative structural analysis more focused on the cluster containing our *h*COX-1 structure has been carried out by removing from the selection the structures having higher deviations, i.e., *h*COX-2 and *o*COX-1 with P_{6₅}22 and I 222 symmetry (Fig. 8c). By using this strategy, the *h*COX-1 (6Y3C) structure is separated from remaining *o*COX-1 crystal structures by PC2, while PC1 discriminates the 3KK6 chains, confirming that this latter structure is peculiar respect to the other *o*COX-1 with P_{6₅} symmetry, as found by RMSD analysis, and 5U6X, which was indeed used to drive the phasing of 6Y3C by molecular replacement. Loading analysis (Fig. 8d) allows identifying residues mainly responsible for the discrimination among selected structures: 438 and 275, which separates 3KK6 along PC1, and 322, 218 and 126, which separate 6Y3C along PC2. It should be noted that the conformational changes shown by the above residues in the discriminated crystal structures could reflect uncertainties due to limited data resolution. Thus, at this level of details, the discrimination is not significant, as confirmed by the limited data variability explained by the first two principal components (35%) in Fig. 8c. PCA applied to PAV appears able to distinguish crystal structures according to conformation changes, even if small, and, therefore, it represents a valuable diagnostic tool to check crystal structures against an ensemble of homologous crystal structures. For *h*COX-1, we verified that aligning the orientation of residues around Thr322, Pro218 and Ser126 to that shown by the P_{6₅} *o*COX-1 chains does not lead to a better structural refinement against diffraction data.

Due to the nature of the crystal structure used as input for PAV analysis, such results rely on static information only. In an attempt to analyze flexibility from more dynamic information, we performed a coarse-grained modelling on our *h*COX-1 (6Y3C) and on the *o*COX-1 crystal structure having the highest resolution among those ones with the same space group of 6Y3C (2OYU) (Supplementary Discussion 2). The residue-by-residue Root Mean Squared Fluctuations (RMSF) values obtained by this modelling appear slightly higher in the case of *h*COX-1 (Supplementary Fig. 10a). By plotting RMSF values against Solvent Accessible Surface Area (SASA) values (Supplementary Fig. 10b), it is possible to observe that residues showing largest changes of flexibility between the two isoforms also show very large changes of SASA due to crystal contacts, but no differences for SASA due to dimer formation. This result suggests that residues characterized by largest differences in flexibility between the two isoforms are involved in crystal contacts, but do not interfere with dimer formation. In addition, it would provide a rationale to the fact that the two proteins have similar activity (related to dimer formation) and different crystallizability (related to crystal packing).

Comparison based on COXs solvent accessibility. Conformational differences in protein structures usually result in significant differences in protein properties, also in the case of proteins with high homology or similar functions. Among these properties, solvent accessibility is of paramount importance because it determines the possibility of exogenous molecules to target protein pockets. RMSD and flexibility analysis suggested that our *h*COX-1 structure is comparable to *o*COX-1 structures and well separated with respect to *h*COX-2 structures. To assess how solvent accessibility discriminates COXs, we calculated the residue-by-residue SASA

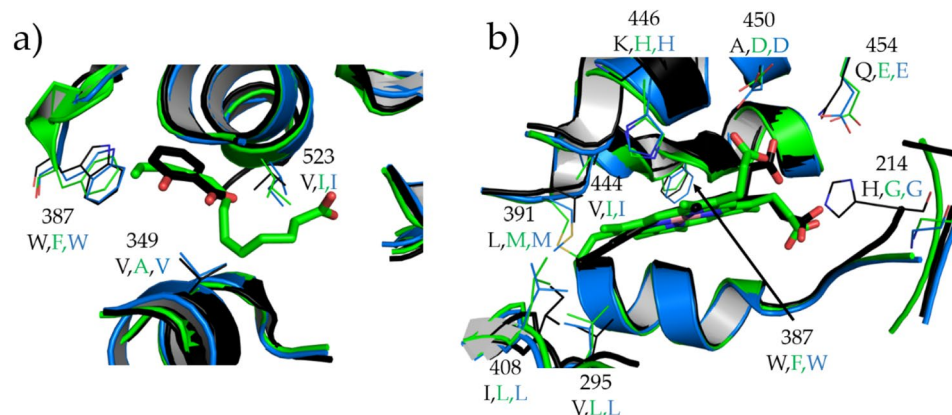


Figure 9. Comparison of the substrate (a) and heme (b) sites of COX proteins. *h*COX-1 (6Y3C) is in cyan color, *o*COX-1 (1U67) in green color, and *h*COX-2 (5F19) in black color. Residues not conserved among the proteins are drawn in line representation, whereas ligand molecules related to *o*COX-1 and *h*COX-2 are drawn in stick representation.

by scanning the protein surface with a probe of 1.4 Å of radius (water radius), and we performed PCA on SASA profiles. The resulting score plot (Fig. 8e) shows a clear separation between *h*COX-2 and *h*COX-1 structures along PC1, which explains 58.5% of the total data variance. According to the loading plots (Fig. 8f), the residue 98, which is a highly exposed tryptophan in COX-1 structures (SASA ≈ 200 Å²) and a glycine in COX-2 structures (SASA ≈ 50 Å²), is mainly responsible for this separation. Instead, COX-1 structures are separated along PC2, which explains only a small percentage of the total data variance (7.8%). *o*COX-1 crystal structures are grouped according to their crystal symmetry, with I 222 and P 6₃22 structures located respectively at large negative and positive values of PC2 scores, and P 6₅ structures positioned at small values of PC2 scores. As in the case of PAV analysis, our *h*COX-1 crystal structure is located in a peripheral region of the I 222 cluster of *o*COX-1 structures, pointing out slightly different solvent accessibility of this structure with respect to most of the other *o*COX-1 structures with same packing. Loadings analysis (Fig. 8f) clarifies that the solvent accessibility separation among COX-1 structures is due to the leading contribution of few differently exposed residues: Lys169 and Glu268, having the highest negative PC2 loadings values, hence responsible for the separation of the I 222 crystal structures, and Asp249, Pro160, and Phe356, having the largest positive PC2 loadings values, hence responsible for the separation of the P 6₃22 crystal structures. These residues are not mutated and do not assume particular conformations in *h*COX-1, so that it is not discriminated along PC2 (the 6y3c PC2 score value is nearly zero in Fig. 8e).

Highly exposed residues dominate structural comparison based on solvent accessibility, and in fact, all the above residues are found on the protein surface. However, it is interesting to perform a solvent accessibility analysis dedicated to two pockets that are important for COX activity, i.e., the substrate and heme sites. To this aim, SASA was calculated by increasing the probe radius from 1.4 to 6.0 Å for residues located in these sites and for three representative crystal structures: 6Y3C (*h*COX-1), 1U67 (*o*COX-1), and 5F19 (*h*COX-2). Such range of probe radius allowed to get information both from small (at low radius) and large (at high radius) solvent channels, thus providing a complete view of the accessibility of the two sites. SASA has been calculated after the removal of ligand molecules from protein structures. The ligand pocket volume can be characterized by the curves describing the SASA values averaged over the pocket residues as a function of the probe radius (Supplementary Fig. 12). For averaging, we consider only the contributes from conserved residues among proteins. In this way, SASA takes into account the steric hindrance of all residues but is not affected by the difference in the molecular surface of the residues exploited for averaging. The result of this analysis shows similar accessibility for *h*COX-1 and *o*COX-1. In the case of the substrate site, we found that the solvent accessibility of *h*COX-2 is higher than COX-1 proteins and decreases faster by increasing the probe radius. For the heme site, the accessibility of the three proteins is the same at 1.4 Å probe radius and, once again, it decreases faster for *h*COX-2. Therefore, also by restricting the analysis to the two pockets that are important for COX activity, we found similarity between *h*COX-1 and *o*COX-1. Interestingly, SASA decreases significantly between 1.4 and 2.5 Å probe radius regardless of protein sequence, and such a decrease is more pronounced in the case of *h*COX-2.

In the case of the heme site, SASA can be ascribed to a different steric hindrance due to residue variations rather than to different conformations. Indeed, the SASA analysis performed on protein molecules converted in polyaniline chains, so that the effect of side chains is excluded from calculations (Supplementary Fig. 13), provides the same accessibility for each of the three protein molecules. Instead, in the case of the substrate site, accessibility appears to be related to different conformations, because the SASA plot obtained for polyaniline-converted protein molecules has the same features.

A detailed view of the two pockets (Fig. 9) can be used to interpret the above analysis result. In the case of the heme site (Fig. 9b), *h*COX-1 and *o*COX-1 have the same residues, which explains the similar trend of their solvent accessibility. The faster drop of SASA curves of *h*COX-2 can be explained by the presence of residues that are bulkier in the *h*COX-2 than COX-1 proteins (particularly for 387, 446, and 214). In the case of the substrate

site (Fig. 9a), *h*COX-1 and *o*COX-1 have different residues, but these variations do not affect their SASA trends, which are still similar.

Comparison of glycosylation sites. Our crystal structure clearly shows the same glycosylation site of *o*COX-1 and *h*COX-2 (asparagine residues 68, 144, and 410). To compare intra- and inter-molecular hydrogen bond interactions occurring in these glycosylation sites, the crystal structures 2OYU, 1Q4G, 4O1Z, and 5F19 have been considered, as they have the highest data resolution among the *o*COX-1 crystallized in the P 6522, I222, P 65 space group, and the *h*COX-2 crystallized in the I 222 space group, respectively. Inter- and intra-molecular interactions are shown in Supplementary Tables 1 and 2, respectively. Five intra-molecular hydrogen bonds are conserved among the proteins under investigation: Asn144-Ser146, Asn410-Gln406, Asn68-Gln42, Asn68-Tyr55, and Asn410-Leu408. The last interaction is conserved in *h*COX-2, despite Leu408 is mutated in Ile. In addition, *h*COX-1 and most of the *o*COX-1 crystal structures share the interactions Asn144-Glu140, Asn410-Ser412, Asn410-Met413, and Asn68-Pro67. In the case of *h*COX-2, Asn410-Ser412 occurs at a distance slightly higher than that considered for our comparison (3.2 Å), while the other two interactions are likely lost due to the mutations of Met413 and Pro67 in Ile and Glu residues, respectively. Instead, Asn410-Asp416 appears conserved in *o*COX-1 crystallized in P 6₅ space group (4O1Z) and in *h*COX-2, where Asp416 is mutated in Glu. Importantly, we do not found interactions neither present nor absent only in *h*COX-1.

On the contrary, inter-molecular hydrogen bonds do not involve glycosylation sites in *h*COX-1, differently from *o*COX-1 and *h*COX-2. For these two proteins, we found that the glycosylation on Asn144 plays an important role in dimer formation (Supplementary Fig. 11a). Asn144 interacts with Glu239, which is mutated in alanine in *h*COX-2. Here, the interaction is formed with Leu238 that, interestingly, appears to shift the α -helix 237–245 much more than the Asn144-Glu239 interaction does in *o*COX-1. The glycosylation site on Asn410 is involved in inter-molecular hydrogen bonds only in *o*COX-1 crystal structures. Particularly, Asn410-Gln282 characterizes *o*COX-1 crystallized in the I222 space group (1Q4G), and Asn410-Gly278 characterizes those crystallized in the P 6₅22 and P 6₅ space groups (2OYU and 4O1Z) (Supplementary Fig. 11b–d). It is noteworthy that in the case of 2OYU and 4O1Z, the position of the amidic carbon atom of the *N*-acetylglucosamine bound to Asn410 should be inverted with that of the oxygen atom of the same group to allow the interaction between this glycosylation and Gly278. Finally, glycosylation on Asn68 interacts with Asp584 and Leu176 in the case of *o*COX-1 crystallized in P₆₅ (4O1Z) and *h*COX-2 (5F19), respectively (Supplementary Fig. 11e–f).

Discussion

For decades, *o*COX-1 has been exploited as a surrogate for the human enzyme, which is more challenging to produce in milligram-quantity and recalcitrant to crystallization¹³. In this paper, we report the first three-dimensional crystal structure of *h*COX-1 along with an extensive structural comparison with known crystal structures of the constitutive (COX-1) and inducible (COX-2) isoforms of the prostaglandin-endoperoxide H synthase. *h*COX-1 has a very high sequence identity with *o*COX-1 (92%) and, like the ovine enzyme, also crystallizes in a hexagonal space group in the presence of lithium chloride and sodium citrate^{19,28}. The remaining crystallization cocktail exploits PEG 4 K as a precipitant and leads to an orthorhombic unit cell, in which protein molecules are more tightly packed than those in the hexagonal cell (the solvent content is \approx 72% and \approx 65% for orthorhombic and hexagonal unit cell, respectively), likely due to the lower ionic strength of the crystallization cocktail. A comprehensive analysis of the diffraction-data precision indicator (DPI) among known COX crystal structures (Supplementary Fig. 6) pointed out the lower crystalline order of *h*COX-1 with respect to *o*COX-1 s and *h*COX-2 s.

In the comparative analysis of the *h*COX-1 crystal structure with those of the ovine enzyme and *h*COX-2, we considered different types of variables connected to crystal structures: (1) Cartesian coordinates, aiming at identifying specific regions showing conformational changes, (2) backbone dihedral angles, aiming at locating hinge points responsible for variations in protein flexibility, (3) solvent accessibility, aiming at finding differently exposed residues. All the analyses agree in confirming that conformationally equivalent protein units make the biological unit of *h*COX-1, *o*COX-1 and *h*COX-2 crystal structures. We delineated two regions characterized by positional deviations between *h*COX-1 and all the other structures investigated: the α_6 - α_7 loop, comprising residues from 270 to 290, and the region 70–110, covering the first two α -helices of the membrane-binding domain. The first region is a key region for determining separations among crystal structures due to its intrinsic flexibility, which is influenced by changes in sequence or symmetry. Here, no significant differences have been observed between *o*COX-1 and *h*COX-2 with respect to our *h*COX-1 by observing ligand interactions and interactions involving glycosylation sites. The second region is mainly affected by crystal packing variations (orthorhombic *versus* hexagonal). Such region shows differences in ligand interaction between *h*COX-1 and *o*COX-2 crystallized in I 222 and difference in interactions involving the glycosylation sites Asn68 of *h*COX-1 and *h*COX-2. Other regions characterized by large deviations emerged from the comparison of our *h*COX-1 structure with *h*COX-2 ones and could be explained by the differences between the interaction involving the glycosylation sites Asn144 and Asn410 of these two proteins.

The analysis based on backbone dihedral angles confirmed the key role played by the above regions to determine separations among COX-1 and COX-2 structures, as specific hinge residues have been systematically found within them (Supplementary Table 3). In other words, we have determined the residues that, changing their orientation, induce large variations of the position of nearby residues, which means that detected differences in the relative position of protein regions in different structures can be attributed to hinge motion. For example, within the region 270–290 we were able to identify residues 274, 281 and 287 as primarily responsible for the separation of *o*COX-1 structures with P 6₅22 symmetry, and residues 269, 170 and 287 as responsible for the separation of *h*COX-2 structures and *o*COX-1 structures with I 222 symmetry (Table 3). Residues 95 and 96 are,

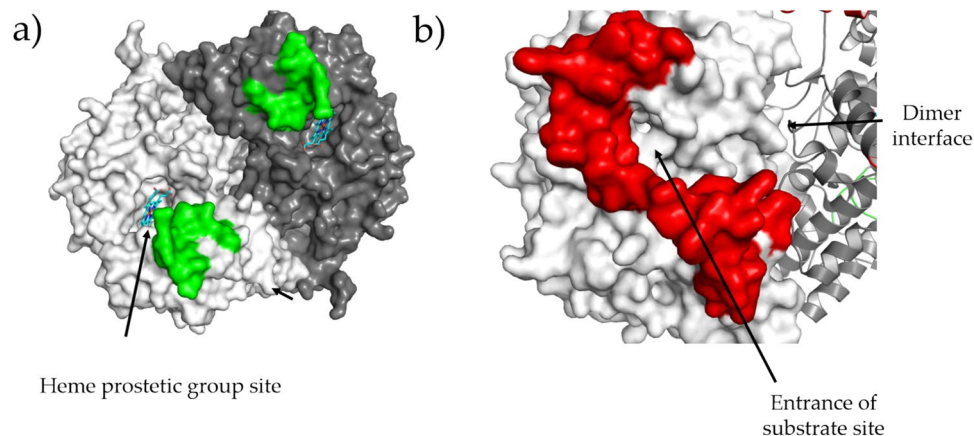


Figure 10. Regions whose conformation is peculiar for *hCOX-1*. Protein units forming the dimer are shown in white and gray surface. (a) The heme molecule is shown in stick representation and the region 270–290 in green color. (b) The region 60–100 is shown in red color.

instead, hinge points that determine unique conformations of the 70–110 region in P 6₅22 *oCOX-1* structures and in 1PRH (Table 3). Additional separation of *hCOX-2* structures is also characterized by unique conformations of their C-term loop, determined by changes in orientations of the hinge residues 574 and 575 (Table 3).

A flexibility analysis performed on similar COX-1 structures highlighted residues showing conformational changes, which are responsible for the discrimination of few outlier crystal structures. In particular, peculiar conformations of Thr322, Pro218, and Ser126 were found in our *hCOX-1* crystal structure when compared with those of the P6₅ *oCOX-1* crystal structures. It is not clear if such differences are originated by variations in the primary structure or are an effect of the limited data resolution. In this respect, the combined PAV-PCA analysis has shown its potentiality in pinpointing divergences in the orientation of residues within an ensemble of homologous crystal structures.

The comparative analysis based on solvent accessibility highlighted the role of residue 98 in discriminating *hCOX-2* structures due to the huge change in its side chain going from *hCOX-2* to *hCOX-1/oCOX-1* sequence (Gly to Trp). This residue still falls in the region 70–110 emerged from the RMSD and dihedral angle analyses. A study dedicated to characterizing the solvent accessibility of the substrate and heme sites has revealed quite a similar shape of these cavities in *hCOX-1* and *oCOX-1* crystal structures.

The two regions whose conformation is peculiar for our *hCOX-1* are shown in Fig. 10. The region between residue 270–290 (Fig. 10a) is part of the loop surrounding the heme site and does not appear directly involved in the electron-transfer mechanism occurring at this site between prostaglandin G₂ (PGG₂) and prosthetic group; however, by considering its position on the protein surface, a role of this region in the binding mechanism of PGG₂ to COX-1 can not be excluded. The region 70–110 (Fig. 10b, red color) links EGF to the membrane-binding domain and, therefore, is close to the substrate pocket. Picot et al. divide such a pocket into two regions named lobby and active site, both important for inhibitor binding^{1,5,28}. The lobby site is very close to the region between residues 70 and 110; therefore, differences in the conformation of this region are expected to affect the lobby region, thus the inhibitor binding mechanism.

In summary, the crystal structure of *hCOX-1* described in this paper represents a significant achievement in structural biology as well as medicinal chemistry, which paves the way for the development of new highly selective inhibitors^{29–32}. Co-crystallization studies with substrate molecules and small-molecule inhibitors will be instrumental in clarifying the unique biochemical properties of the human enzyme. Future studies that build upon the structure presented here will help in the development of novel therapeutics and imaging-probe that target *hCOX-1* for unmet clinical diagnosis of diseases in which *hCOX-1* is a biomarker^{9,10}.

Materials and methods

Cell culture reagents were purchased from EuroClone (Milan, Italy). Protoporphyrin IX (Hemin), trizma base, potassium chloride, imidazole, sodium citrate tribasic dihydrate and lithium chloride, Amicon Ultra-4 10 kDa MWCO centrifugal filter device was purchased from Merck. Nickel-NTA agarose beads (low density), *n*-octyl β-D-glucopyranoside (β-OG) and phenylmethylsulfonyl fluoride (PMSF) were from Gold Biotechnology. Hexaethylene Glycol Monodecyl Ether (C₁₀E₆) detergent was purchased from Anatrache (Maumee, OH). Phusion High Fidelity (HF) DNA was from Biolabs. Sf-900 II SFM medium, HALT protease inhibitor single-use cocktail EDTA-free, Pierce BCA protein assay kit, Snakeskin dialysis tubing (10 kDa MWCO, 35 mm), and glycerol were purchased from Thermo Fisher Scientific Italia (Monza, Italy). TGX 10% precast polyacrylamide gels and all reagents for SDS-PAGE were purchased from BIO-RAD Laboratories Srl (Milan, Italy). Tobacco Etch Virus (TEV) protease was produced as described³³. Oxygraph buffer: 200 mM Tris-HCl (pH 8.3), 2 M KCl, 1 mM phenol.

Construction of the baculovirus recombinant expression vector, protein production and purification. Cloning into baculovirus DNA of *hCOX-1* cDNA bearing an import signal peptide followed by an N-terminal 8X-His tag and a TEV cleavage was carried out as follows. Briefly, *hCOX-1* cDNA was obtained

by PCR using Phusion High Fidelity (HF) DNA polymerase and two specific primers For and Rev bearing at their 5'-end a BglII and a NotI site, respectively. Primer sequences (5' to 3') are: PBac.*hCOX-1*.BglII.For: CAC AAGATCTGATATGAGCCGGAGTCTCTT and PBac.*hCOX-1*.NotI.Rev: CACAGCGGCCGCTCAGAGCTC TGTGGATG. DNA template was a recombinant pFastBac1/*hCOX-1* cDNA gently provided by Prof. William Smith (Michigan University). The PCR reaction was carried out in a 50 μ l volume, using 3 ng of DNA, 0.2 mM dNTPs, 0.5 μ M for each primer, 1X Phusion HF buffer containing MgCl₂ and 1 U of Phusion High-Fidelity DNA Polymerase (BioLabs). After heating at 98 °C for 40 s, the reaction was subjected to 30 cycles consisting of 98 °C for 30 s, 68 °C for 1 min and 72 °C for 40 s; the final elongation was performed at 72 °C for 7 min. The amplification product was purified using Eurogold Cycle Pure kit (Euroclone) according to the manual instructions, and quantified by measuring the absorbance at 260 nm with the NanoDrop 1000 (Euroclone); it was then restriction-digested with BglII and NotI, purified, quantified as before, and finally inserted into the BglII/NotI sites of the pBacPAK9 baculovirus shuttle vector (Clontech), using standard DNA manipulation techniques. For DNA ligation the Rapid DNA Ligation Kit (Roche) was used according to the manufacturer's protocol, to yield the pBacPAK9/*hCOX-1* recombinant vector. Ligation mixture was transformed into XL1-blue competent *E. coli* cells according to the standard protocol employing CaCl₂. Positive clones were identified by restriction digestion of the miniprep DNA; the correct sequence of the insert was verified by automatic sequencing of both strands (Eurofins Genomics Italy). pBacPAK9/*hCOX-1* plasmid and linearized baculovirus DNA (BacPAK6 Viral DNA Bsu36 I digest) were co-transfected in *Spodoptera frugiperda* (Sf9) cells according to the manufacturer's protocol³⁴. The co-transfection supernatant was subjected to plaque assay to obtain individual viral plaques; four single and well separated agarose plaques corresponding to pure viruses were picked and immersed into 500 μ l of Sf-900II SFM (Gibco) with 1% penicillin/streptomycin (P/S) (Gibco) and 5% fetal bovine serum (FBS) (Gibco) for 16 h at +4 °C in the dark, to allow viruses to diffuse out of the agarose. Putative recombinant viruses from each plaque were tested for the presence of *hCOX-1* insert; for that, each plaque supernatant was subjected to a one-step amplification, yielding Virus Stock I (VSI), from which the DNA viral was extracted according to McCarthy CB et al. with some modifications¹⁵. A 200 μ l volume of four VSI was lysed by gentle mixing with an equal volume of Lysis Buffer (100 mM Tris-HCl pH 7.6, 10 mM EDTA, 0.25% SDS) followed by a 10-min incubation at room temperature. Samples were extracted once with phenol/chloroform/isoamyl alcohol and once with chloroform/isoamyl alcohol; the nucleic acids were ethanol precipitated, dissolved in water and quantified spectrophotometrically. A 100 ng amount of viral DNA was analyzed by PCR in a final volume of 50 μ l, using Taq Polymerase TaKara (Clontech) (2.5 U), dNTPs 0.2 mM, 1 μ M of each specific primer For and Rev used for cloning (see above). After heating at 98 °C for 15 min, the reaction was subjected to 30 cycles of 98 °C for 30 s, 52 °C for 1 min and 72 °C for 2 min, followed by an extension at 72 °C for 5 min. All of the four analyzed plaques resulted positive. One VSI was chosen, that was amplified twice yielding VSII and VSIII; the latter was titred by plaque assay ($\sim 5 \times 10^8$ pfu/ml) and used as inoculum stock for recombinant protein expression.

Sf9 insect cell culture suspension at a cell density of 2×10^6 cells/ml in 4 L of the Sf-900 II SFM medium (5% FBS, 1% P/S) was infected with the recombinant virus stock VSIII at a multiplicity of infection (pfu/cell) of 10. After 72 h, infected Sf9 cells were harvested by centrifugation at 900xg at +4 °C for 20 min. The cell pellet was rapidly cooled in liquid nitrogen and stored at -80 °C until use. The following steps were performed at +4 °C. For 8xHis-tagged *hCOX-1* purification, cell pellet was resuspended in 30 ml of cold lysis buffer [20 mM Tris-HCl pH 8.0, 100 mM KCl, containing Complete EDTA-free protease inhibitor (Roche) and 1 mM PMSF] and disrupted by sonication (7 microtip of potency, 1" pulse on and 4" pulse off for 5 min, sonicator Branson) on ice (*ovine* or *human* COX-1/Sf9 sonicated pellet). C₁₀E₆ detergent was added (1.12% w/v) and solubilization was performed with gentle shaking at +4 °C overnight (o/n). The lysate was clarified by centrifugation at +4 °C for 75 min at 62,000xg, and the resulting supernatant was incubated with Ni²⁺-NTA agarose beads (low density, 50% slurry) pre-equilibrated twice with 3 column volumes of Ni-Buffer [20 mM Tris pH 8.0, 100 mM KCl, 5 mM imidazole, 0.1% (w/v) C₁₀E₆, 5% glycerol] and rocked for 3 h at +4 °C. One milliliter of Ni²⁺-NTA agarose resin binds 5 to 10 mg of 8X-His fusion protein. The protein-resin slurry was poured into the chromatography column, and the flowthrough (FT) was collected slowly. The beads were washed sequentially with ten-bed volumes of pre-chilled wash buffer I [20 mM Tris-HCl, pH 8.0, 500 mM KCl, 10 mM imidazole, 0.1% (w/v) C₁₀E₆, 5% glycerol] and five-bed volumes of wash buffer II [20 mM Tris-HCl pH 8.0, 40 mM KCl, 20 mM imidazole, 0.1% (w/v) C₁₀E₆, 5% glycerol]. Finally, the His-tagged *hCOX-1* protein was eluted with 1-3 bed-volumes of elution buffer [20 mM Tris-HCl pH 8.0, 40 mM KCl, 250 mM Imidazole, 0.1% (w/v) C₁₀E₆, 5% glycerol]. The FT fraction was subjected to a further chromatography step by re-incubating it with the resin in batch with gentle rotation, o/n. The *hCOX-1* eluted fractions were combined and dialyzed o/n against 500 ml of no-imidazole buffer [20 mM Tris-HCl pH 8.0, 150 mM KCl, 0.1% (w/v) C₁₀E₆, 5% glycerol]. Histidine tag was cleaved off with TEV_{6His} protease at 1:40 (w/w) protease to protein ratio. After digestion, the reaction mixture was passed on fresh Ni-beads to recapture the protease¹⁶. *hCOX-1* was recovered in the FT and beads were washed with buffer without imidazole [20 mM Tris-HCl pH 8.0, 100 mM KCl, 0.1% (w/v) C₁₀E₆, 5% glycerol]. Protein purity was assessed by SDS-PAGE analysis followed by staining with Coomassie Brilliant Blue G-250. Protein concentration was assessed spectrophotometrically by NanoDrop 2000 (Thermo Scientific) measuring the absorbance at 280 nm using the theoretical molar extinction coefficient ($96,720 \text{ M}^{-1} \text{ cm}^{-1}$) determined from the amino acid sequence and molecular weight to 70 kDa.

Cyclooxygenase activity evaluation by O₂ measurement-based assay. Functional analysis of *ovine* and *human* COX-1 catalytic activity was determined by monitoring O₂ consumption by using an Oxytherm electrode unit (Hansatech), in the presence of arachidonic acid and different inhibitors. The instrument is equipped with a Clark-type oxygen electrode to monitor the dissolved oxygen concentration in a sealed measurement chamber over the time. Oxygen consumption was measured at 37 °C directly in a 1000 μ l reaction

vessel. Briefly, it follows the description of a typical assay: Oxygraph buffer (the opportune volume is adjusted to 1000 μL of final volume according to the volume of the protein), 5 μM hemin (1 μL), 100 μM arachidonic acid (20 μL) and inhibitor (10 μL) at the opportune concentrations. By this feasible approach, COX-inhibitory activity (IC_{50}) of known NSAIDs such as mofezolac, ibuprofen and indomethacin was evaluated. Mofezolac was chosen as an example of highly selective COX-1 inhibitor, while indomethacin and ibuprofen as non-selective COXs inhibitors.

Oxygen consumption determination ($\mu\text{mol}/\text{ml O}_2$) started after the addition of the ovine or human COX-1/Sf9 sonicated pellet ($C = 2.6 \text{ mg}/\text{ml}$) to the mixture containing Oxygraph buffer, arachidonic acid and test compounds at the opportune concentrations. Assay mixture was allowed to equilibrate in the reaction vessel at 37 $^\circ\text{C}$ for several minutes until steady baseline was recorded. Uninfected Sf9 sonicated cells were used as a negative control. Stock solutions were prepared in a minimum volume of vehicle (DMSO).

Data were collected by a PC and analyzed by an Oxygraph Plus Software. The change in O_2 concentration, expressed as 1 micromole of oxygen per minute at 37 $^\circ\text{C}$ (O_2 consumed/min) were obtained from the first derivative of $[\text{O}_2]$ versus time.

Protein crystallization, data collection, and structure determination. *hCOX-1* was buffer exchanged against 20 mM HEPES at pH 7.3, 40 mM NaCl and 0.4% w/v β -octyl glucoside using an Amicon Ultra 4 ml concentrator (with membrane Ultracel-PL PLGC, MWCO 10 kDa) and concentrated to 15 mg/ml. As for COX-2, also COX-1 requires heme (Fe^{3+} -protoporphyrin IX) as a cofactor. Therefore, 1 mM protoporphyrin in DMSO was added to the protein solution in a 1:1 molar ratio. The solution was incubated at 4 $^\circ\text{C}$ for 10 min and used for crystallization. Protein crystals were obtained in 2 weeks by using the sitting-drop vapor diffusion method at 293 $^\circ\text{K}$ by mixing 1.5 μL of protein solution with 1.5 μL of reservoir solution consisting of sodium citrate 0.7 and 0.58 M LiCl. The pH value of the reservoir solution was not adjusted. Crystallization experiments were performed at 10.4 mg/ml of protein concentration based on the previous optimization done using the Hampton Research Pre-Crystallization Test. Crystals were flash-frozen in liquid nitrogen after a quick dipping in a solution containing reservoir solution supplemented with 20% glycerol as cryoprotectant. X-ray diffraction data collection was performed at the I04 beamline of Diamond Light Source Ltd Synchrotron (Didcot, Oxfordshire, UK) by using 0.9795 \AA as wavelength. Data reduction was performed by XIA2³⁵ in 3dii mode, thus using XDS³⁶ for indexing, scaling and merging. TRUNCATE³⁷ included in CCP4 crystallographic suite³⁸ was used to convert reflection intensities to amplitudes and to select reflections for R_{free} calculation. Data have been processed by STARANISO³⁹ to remove anisotropy. The REMO program⁴⁰ included in the package SIR2014⁴¹ was used to solve the structure by Molecular Replacement. Here, the crystallographic structure of ovine COX-1 (PDB code 5U6X)¹³, which has 92% of sequence identity with respect to the *hCOX-1*, was used as a starting model. Structural refinement was performed by using Phenix.refine⁴² included in the crystallographic suite PHENIX⁴³. Once refined, the *hCOX-1* crystal structure has been validated using the PDB validation server⁴⁴ and deposited in the Protein DataBank with accession code 6Y3C. PISA server²⁰ was used to define *hCOX-1* biological unit and PRO-ORIGAMI⁴⁵ along with TOPDRAW⁴⁶ to determine and draw the protein topology.

Comparative structural analysis. The *hCOX-1* crystal structure was compared with known COX-1 from *ovis aries* (found by setting a search query in the PDB based on the sequence of *oCOX-1* having UniProt code P05979, identity E-value = 10^6 and cutoff identity = 95%) and *hCOX-2* structures (found by setting a search query in the PDB based on the sequence of *hCOX-2* having UniProt code P35354, identity E-value = 10^6 , and cutoff identity = 95%) by using residue-by-residue, Solvent Accessibility Surface Area (SASA) values, Cartesian coordinates of the C α atoms and backbone dihedral angles. The first ones were calculated by the program AREAIMOL⁴⁷ included in the crystallographic suit CCP4, the second ones were used to calculate the root mean square deviation (RMSD) between pair of structures, though the program SUPERPOSE⁴⁸, the third ones were used to calculate the Protein Angular Value (PAV) associated to each residue⁴⁹. PAV is defined as:

$$PAV_i = \frac{180}{\pi} \cos^{-1}(\cos(\psi_i + \phi_i)) \quad (1)$$

where ψ_i and ϕ_i are the backbone dihedral angles of the i -th residue. PAV values range between 0 $^\circ$ and 180 $^\circ$ and represent the $\psi_i + \phi_i$ values expressed in degrees, avoiding the problem of range definition connected with the circular nature of the angular variables. PAV profiles of each structure were calculated through the script TPAID⁵⁰ run on VMD⁵¹ SASA, RMSD, and PAV profiles from different structures were separately analyzed by using Principal Component Analysis (PCA) and hierarchic clustering implemented in the program RootProf⁵². Crystal structures have been compared by considering individual chains of the biological unit. Solvent molecules, metal ions and ligands were removed from the PDB files, and protein residues between 34 and 583 were considered, as terminal residues were missing in some structural models. *hCOX-2* and *oCOX-1* sequences were aligned with the *hCOX-1* one, so that to consider 550 residues regardless of the type of protein. PCA was performed by considering two chains for each crystal structure to ensure a uniform weighting. In the case of SASA, residues having missing side chain in the final model of at least one crystal structure are excluded from PCA. The coordinate errors of the structural models used for comparative analysis were estimated by the diffraction-data precision indicator⁵³, defined by:

$$DPI = \sqrt{\frac{N_{atoms}}{N_{obs}} C^{-1} / 3 d_{min} R_{free}} \quad (2)$$

where N_{atoms} is the number of atoms in the model, N_{obs} is the number of independent reflections, C is the fractional data completeness, R_{free} is the crystallographic R-factor and d_{min} is the data resolution. DPI has been calculated by using the program SFCHECK⁵⁴ when diffraction data are available, or by using Eq. (2) with information given in the PDB^{55–65}.

Received: 10 June 2020; Accepted: 2 February 2021

Published online: 22 February 2021

References

- Smith, W. L., Garavito, R. M. & DeWitt, D. L. Prostaglandin endoperoxide H synthases (cyclooxygenases) -1 and -2. *J. Biol. Chem.* **271**, 33157–33160 (1996).
- Smith, W. L., Dewitt, D. L. & Garavito, R. M. Cyclooxygenases: Structural, cellular, and molecular biology. *Annu. Re. Biochem.* **69**, 145–182 (2000).
- Rouzer, C. A. & Marnett, L. J. Cyclooxygenases: Structural and functional insights. *J. Lipid Res.* **50**, S29–S34 (2009).
- Zappavigna, S. *et al.* Anti-inflammatory drugs as anticancer agents. *Int. J. Mol. Sci.* **21**(2605), 1–29 (2020).
- Perrone, M. G., Scilimati, A., Simone, L. & Vitale, P. Selective COX-1 inhibition: A therapeutic target to be reconsidered. *Curr. Med. Chem.* **17**, 3769–3805 (2010).
- Perrone, M. G. *et al.* Translational theragnostic of ovarian cancer: Where do we stand?. *Curr. Med. Chem.* **27**, 5675–5715 (2020).
- Pati, M. L. *et al.* Translational impact of novel widely pharmacological characterized mofezolac-derived COX-1 inhibitors combined with bortezomib on human multiple myeloma cell lines viability. *Eur. J. Med. Chem.* **164**, 59–76 (2019).
- Casalino, G. *et al.* Intelligent microarray data analysis through non-negative matrix factorization to study human multiple myeloma cell lines. *Appl. Sci.* **9**(5552), 1–18 (2019).
- Perrone, M. G. *et al.* PET radiotracer [¹⁸F]-P6 selectively targeting COX-1 as a novel biomarker in ovarian cancer: Preliminary investigation. *Eur. J. Med. Chem.* **80**, 562–568 (2014).
- Scilimati, A. *et al.* Targeting COX-1 by mofezolac-based fluorescent probes for ovarian cancer detection. *Eur. J. Med. Chem.* **179**, 16–25 (2019).
- Vitale, P., Scilimati, A. & Perrone, M. G. Update on SAR studies toward new COX-1 selective inhibitors. *Curr. Med. Chem.* **22**, 4271–4292 (2015).
- Vitale, P., Panella, A., Scilimati, A. & Perrone, M. G. COX-1 inhibitors: Beyond structure toward therapy med. *Res. Rev.* **36**, 641–671 (2016).
- Cingolani, G. *et al.* Structural basis for selective inhibition of Cyclooxygenase-1 (COX-1) by diarylisoxazoles mofezolac and 3-(5-chlorofuran-2-yl)-5-methyl-4-phenylisoxazole (P6). *Eur. J. Med. Chem.* **138**, 661–668 (2017).
- Kitts, P. A. & Possee, R. D. A method for producing recombinant baculovirus expression vectors at high frequency. *Biotechniques* **14**, 810–817 (1993).
- McCarthy, C. B. & Romanowski, V. A simplified method for the extraction of baculoviral DNA for PCR analysis: A practical application. *J. Virol. Methods* **148**, 286–290 (2008).
- Miciaccia, M., Iaselli, M., Ferorelli, S., Loguercio Polosa, P., Perrone, M. G., & Scilimati, A. In-depth ovine COX-1 isoenzyme bio-production. *submitted* (2020).
- Calvello, R. *et al.* Novel selective COX-1 inhibitors suppress neuroinflammatory mediators in LPS-stimulated N13 microglial cells. *Pharmacol. Res.* **65**, 137–148 (2012).
- Calvello, R. *et al.* Highly selective cyclooxygenase-1 inhibitors P6 and mofezolac counteract inflammatory state both *in vitro* and *in vivo* models of neuroinflammation. *Front. Neurol.* **8**(251), 1–10 (2017).
- Malkowski, M. G., Ginell, S. L., Smith, W. L. & Garavito, R. M. The Productive conformation of arachidonic acid bound to prostaglandin synthase. *Science* **289**, 1933–1937 (2000).
- Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).
- Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**, W320–W324 (2014).
- Di Nunno, L., Vitale, P., Scilimati, A., Tacconelli, S. & Patrignani, P. Novel synthesis of 3,4-diarylisoxazole analogues of valdecoxib: reversal cyclooxygenase-2 selectivity by sulfonamide group removal. *J. Med. Chem.* **47**, 4881–4890 (2004).
- Vitale, P. *et al.* Synthesis, pharmacological characterization, and docking analysis of a novel family of diarylisoxazoles as highly selective cyclooxygenase-1 (COX-1) inhibitors. *J. Med. Chem.* **56**, 4277–4299 (2013).
- Vitale, P., Perrone, M. G., Malerba, P., Lavecchia, A. & Scilimati, A. Selective COX-1 inhibition as a target of theranostic novel diarylisoxazoles. *Eur. J. Med. Chem.* **74**, 1–13 (2014).
- Xiao, G., Chen, W. & Kulmacz, R. J. Comparison of structural stabilities of prostaglandin H synthase-1 and -2. *J. Biol. Chem.* **273**, 6801–6811 (1998).
- Rimon, G. *et al.* Coxibs interfere with the action of aspirin by binding tightly to one monomer of cyclooxygenase-1. *Proc. Natl. Acad. Sci. USA* **107**, 28–33 (2010).
- Sidhu, R. S. *et al.* Comparison of cyclooxygenase-1 crystal structures: cross-talk between monomers comprising cyclooxygenase-1 homodimers. *Biochemistry* **49**, 7069–7079 (2010).
- Picot, D., Loll, P. J. & Garavito, R. M. The X-ray crystal structure of the membrane protein prostaglandin H₂ synthase-1. *Nature* **367**, 243–249 (1994).
- Perrone, M. G. *et al.* Selective cyclooxygenase-1 inhibition by P6 and gastrotoxicity: Preliminary investigation. *Pharmacology* **95**, 22–28 (2015).
- Perrone, M. G. *et al.* Effect of mofezolac-galactose distance in conjugates targeting cyclooxygenase (COX)-1 and CNS GLUT-1 carrier. *Eur. J. Med. Chem.* **141**, 404–416 (2017).
- Perrone, M. G., Miciaccia, M. & Scilimati, A. *et al.* An attempt to chemically state the cross-talk between monomers of COX homodimers by double/hybrid inhibitors mofezolac-spacer-mofezolac and mofezolac-spacer- arachidonic acid. *Eur. J. Med. Chem.* **209**, 112919–112937 (2021).
- Perrone, M.G., Vitale, P., Panella, P., Fortuna, C.G., Scilimati, A. General role of the amino and methylsulfamoyl groups in selective cyclooxygenase(COX)-1 inhibition by 1,4-diaryl-1,2,3-triazoles and validation of a predictive pharmacometric PLS model. *Eur. J. Med. Chem.* **94**, 252–264 (2015).
- Koksal, A. C., Nadozzi, J. D. & Cingolani, G. Dimeric quaternary structure of the prototypical dual specificity phosphatase VHI1. *J. Biol. Chem.* **284**, 10129–10137 (2009).

34. BacPAK™ Baculovirus Expression System User Manual, Clontech.
35. Winter, G. Xia2: An expert system for macromolecular crystallography data reduction. *J. Appl. Cryst.* **43**, 186–190 (2010).
36. Kabsch, W. *et al.* Overview of the CCP4 suite and current developments. *Acta Cryst. Sect. D Biol. Cryst.* **66**, 125–132 (2010).
37. French, S. & Wilson, K. On the treatment of negative intensity observations. *Acta Cryst.* **A34**, 517–525 (1978).
38. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Cryst. Sect. D Biol. Cryst.* **67**, 235–242 (2011).
39. <http://staraniso.globalphasing.org/cgi-bin/staraniso.cgi>
40. Caliendo, R. *et al.* Molecular replacement: The probabilistic approach of the program REMO9 and its applications. *Acta Cryst. Sect. A Found. Crystal.* **65**, 512–527 (2009).
41. Burla, M. C. *et al.* Crystal structure determination and refinement via SIR2014. *J. Appl. Cryst.* **48**, 306–309 (2015).
42. Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta Cryst. Sect. D Biol. Cryst.* **68**, 352–367 (2012).
43. Liebschner, D. *et al.* Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in Phenix. *Acta Cryst. Sect. D Struct. Biol.* **75**, 861–877 (2019).
44. <https://validate-rcsb-2.wwpdb.org/>
45. Stivala, A., Wybrow, M., Wirth, A., Whisstock, J. C. & Stuckey, P. J. Automatic generation of protein structure cartoons with procrigami. *Bioinformatics* **27**, 3315–3316 (2011).
46. Bond, C. S. TopDraw: A sketchpad for protein structure topology cartoons. *Bioinformatics* **19**, 311–312 (2003).
47. Lee, B. & Richards, F. M. The interpretation of protein structures: Estimation of static accessibility. *J. Mol. Biol.* **55**, 379–400 (1971).
48. Krissinel, E. & Henrick, K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Cryst. Sect. D Biol. Cryst.* **60**, 2256–2268 (2004).
49. Liuzzi, V. C. *et al.* Enniatin and beauvericin biosynthesis in *Fusarium* species: Production profiles and structural determinant prediction. *Toxins* **9**, 45 (2017).
50. Caliendo, R., Rossetti, G. & Carloni, P. Local fluctuations and conformational transitions in proteins. *J. Chem. Theory Comput.* **8**, 4775–4785 (2012).
51. Humphrey, W., Dalke, A. & Schultze, K. V. M. D. Visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).
52. Caliendo, R. & Belviso, D. B. RootProf: Software for multivariate analysis of unidimensional profiles. *J. Appl. Cryst.* **47**, 1087–1096 (2014).
53. Cruickshank, D. W. J. In the Refinement of Macromolecular structures Proceedings of CCP4 Study weekend. pp 11–22 (1996).
54. Vaguine, A. A., Richelle, J. & Wodak, S. J. SFCHECK: A unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. *Acta Cryst. Sect. D Biol. Cryst.* **55**, 191–205 (1999).
55. Perrone, M. G. *et al.* Baker's yeast-mediated reduction of ethyl 2-(4-chlorophenoxy)-3-oxoalkanoates intermediates for potential PPARα ligands. *Tetrahedron Asymmetry*. **15**, 3501–3510 (2004).
56. Vitale, P., D'Introno, C., Perna, F., Perrone, M. G. & Scilimati, A. *Kluyveromyces marxianus* CBS 6556 growing cells as a new biocatalyst in the asymmetric reduction of substituted acetophenones. *Tetrahedron Asymmetry*. **24**, 389–394 (2013).
57. Catalano, A. *et al.* Constrained analogues of tocinide as potent skeletal muscle sodium channel blockers towards the development of antimyotonic agents. *Eur. J. Med. Chem.* **43**, 2535–2540 (2008).
58. Di Nunno, L., Vitale, P. & Scilimati, A. *et al.* Stereoselective dimerization of 3-arylisoxazoles to caged bis-β-lactams syn 2,6-diaryl-3,7-diazatricyclo[4.2.0.0_{2,5}] octan-4,8-diones induced by hindered lithium amides. *Tetrahedron*. **63**, 12388–12395 (2007).
59. De Luca, A., Talon, S., De Bellis, M., Desaphy, J. F. & Lentini, G. *et al.* Optimal requirements for high affinity and use-dependent block of skeletal muscle sodium channel by N-benzyl analogs of tocinide-like compounds. *Mol. Pharmacol.* **64**, 932–945 (2003).
60. Di Nunno, L. & Scilimati, A. Decomposition of arylazides by THF/n-butyllithium-II-isolation of 1-aryl-4,5-dihydro-5-hydroxy-1h-1,2,3-triazoles. *Tetrahedron*. **42**, 3913–3939 (1986).
61. Colabufo, N. A. *et al.* Naphthalenyl derivatives for hitting P-gp/MRP1/BCRP transporters. *Bioorg. Med. Chem.* **21**, 1324–1332 (2013).
62. Perrone, M. G. *et al.* Synthesis and biological evaluation of new clofibrate analogues as potential PPARα agonists. *Eur. J. Med. Chem.* **40**, 143–154 (2005).
63. Perrone, M. G. *et al.* Diastereo- and enantioselective bioreduction of ethyl 2-(4-chlorophenoxy)-3-oxobutanoate clofibrate analogues by *Kluyveromyces marxianus* and other whole cell biocatalysts. *Tetrahedron Asymmetry*. **15**, 3511–3517 (2004).
64. Di Nunno, L. & Scilimati, A. Decomposition of arylazides by thf/-butyllithium-II-isolation of 1-aryl-4,5-dihydro-5-hydroxy-1h-1,2,3-triazoles. *Tetrahedron*. **42**, 3913–3920 (1986).
65. Scilimati, A. *et al.* Patient Bone Marrow Aspiration to Explore the Cyclooxygenases (COXs) Involvement in Multiple Myeloma. *J. Cancer Res. Therap. Oncol.* **9**, 1–19 (2021).

Acknowledgements

Authors wish to thank the Diamond Light Source for access to beamline I04 (Proposal No. MX15832-1), that contributed to the results presented here.

Author contributions

M.G.P., S.F. and A.S. conceived and supervised the project; P.L.P., M.C. and M.M. performed *hCOX-1* cloning in the baculovirus vector; M.M. performed protein expression, purification and crystallization experiments; M.I. and G.C. participated to protein expression and purification; B.D.B. and R.C. performed X-ray diffraction experiments and determined the crystal structure; B.D.B. refined and validated the crystal structure; B.D.B. and R.C. performed the comparative structural analysis; all authors wrote, reviewed the manuscript and agreed to its published version.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-83438-z>.

Correspondence and requests for materials should be addressed to R.C. or A.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021