The FEBS Journal

FEBS PRESS science publishing by scientists

Resurrection and characterization of ancestral CYP11A1 enzymes

Philip Hartz[1], Silja J. Strohmaier[2], Basma M. EL-Gayar[1], Ammar Abdulmughni[1], Michael C. Hutter[3], Frank Hannemann[1], Elizabeth M. J. Gillam[2] and Rita Bernhardt[1]

1 Department of Biochemistry, Saarland University, Saarbrücken, Germany
2 School of Chemistry and Molecular Biosciences, The University of Queensland, St. Lucia, Australia
3 Center for Bioinformatics, Saarland University, Saarbrücken, Germany

Keywords
ancestral sequence reconstruction; cholesterol; cytochromes P450; pregnenolone formation; steroid hormone biosynthesis

Correspondence
E. M. J. Gillam, School of Chemistry and Molecular Biosciences, The University of Queensland, St. Lucia 4072, Australia
Tel: +61 7 336 51410
E-mail: e.gillam@uq.edu.au
R. Bernhardt, Department of Biochemistry, Saarland University, Campus Building B2.2, 66123 Saarbrücken, Germany
Tel: +49 681 302 4241
E-mail: ritabern@mx.uni-saarland.de

Philip Hartz and Silja J. Strohmaier contributed equally to this article

(Received 29 January 2021, revised 22 May 2021, accepted 4 June 2021)

doi:10.1111/febs.16054

Mitochondrial cytochromes P450 presumably originated from a common microsomal P450 ancestor. However, it is still unknown how ancient mitochondrial P450s were able to retain their oxygenase function following relocation to the mitochondrial matrix and later emerged as enzymes specialized for steroid hormone biosynthesis in vertebrates. Here, we used the approach of ancestral sequence reconstruction (ASR) to resurrect ancient CYP11A1 enzymes and characterize their unique biochemical properties. Two ancestral CYP11A1 variants, CYP11A_Mammal_N101 and CYP11A_N1, as well as an extant bovine form were recombinantly expressed and purified to homogeneity. All enzymes showed characteristic P450 spectral properties and were able to convert cholesterol as well as other sterol substrates to pregnenolone, yet with different specificities. The vertebrate CYP11A_N1 ancestor preferred the cholesterol precursor, desmosterol, as substrate suggesting a convergent evolution of early cholesterol metabolism and CYP11A1 enzymes. Both ancestors were able to withstand increased levels of hydrogen peroxide but only the ancestor CYP11A_N1 showed increased thermostability (~ 25 °C increase in T_{50}) compared with the extant CYP11A1. The extraordinary robustness of ancient mitochondrial P450s, as demonstrated for CYP11A_N1, may have allowed them to stay active when presented with poorly compatible electron transfer proteins and resulting harmful ROS in the new environment of the mitochondrial matrix. To the best of our knowledge, this work represents the first study that describes the resurrection of ancient mitochondrial P450 enzymes. The results will help to understand and gain fundamental functional insights into the evolutionary origins of steroid hormone biosynthesis in animals.

Introduction

Cytochromes P450 (P450s) constitute a large group of heme-thiolate oxygenases that catalyze the incorporation of one atom of molecular oxygen into nonactivated C-H bonds of organic substrates [1]. Since P450s are distributed among all domains of life [2], they are believed to originate from a common

Abbreviations
AdR, adrenodoxin reductase; Adx, adrenodoxin; ASR, ancestral sequence reconstruction; CPR, cytochrome P450 reductase; DOC, 21-hydroxyprogesterone (11-deoxycorticosterone); ETS, electron transfer system; FdR, ferredoxin reductase; Fdx, ferredoxin; IEC, ion exchange chromatography; IMAC, immobilized metal ion affinity chromatography; MSA, multiple sequence alignment; MTS, mitochondrial targeting sequence; ROS, reactive oxygen species; SCC, side-chain cleavage; T_m, melting temperature; β-CD, 2-hydroxypropyl-β-cyclodextrin.

ancestor that underwent sequential diversification processes during the evolution of life, leading to a variety of P450 species with different activities and specialized functions. P450s were first found in liver microsomes of animals [3] and were subsequently also discovered in adrenocortical tissues [4]. Further studies on distribution and function of different P450 species in vertebrates revealed two classes of P450s based on subcellular localization. Mitochondrial P450s are localized on the matrix side of the inner mitochondrial membrane and are primarily involved in biosynthesis of steroid hormones, which act as potent regulators of fundamental physiological processes including development, reproduction as well as homeostasis in animals. By contrast, microsomal P450s are found on the cytoplasmic side of the endoplasmic reticulum (ER) membrane and are, with some exceptions, responsible for metabolic clearance of endogenous and exogenous substrates. Numerous microsomal P450s with diverse functions have been identified and characterized in other eukaryotes including plants [5], fungi [6], and insects [7], but mitochondrial P450s seem to be unique to animal cells. Hence, there have been considerable efforts to investigate and understand the development and evolution of mitochondrial P450s, which presumably started shortly after the beginning of the animal lineage [8].

A common hypothesis assumes that mitochondrial P450s might have evolved early from a common microsomal ancestor. Ongoing phylogenetic studies, indeed, suggest that sequential mutations in the signal-anchor sequence of microsomal P450s, which targets them to the cytoplasmic site of the ER membrane, promoted the evolution of a mitochondrial targeting sequence (MTS) and eventually allowed the posttranslational transport of microsomal P450s into the mitochondrial matrix [9]. However, once relocated to the mitochondrial matrix, the formerly microsomal ancestor was spatially separated from its auxiliary electron transfer system, an ER membrane-bound NADPH-dependent cytochrome P450 reductase (CPR) [10]. In order to retain their oxygenase function in the mitochondrial matrix, the relocated microsomal P450s presumably adapted to an alternative electron transfer system consisting of the already existing ferredoxin (Fdx) and cognate ferredoxin reductase (FdR) [11–14]. The whole process seems to have coincided with gradual changes in the substrate spectrum of microsomal P450s at the end of which mitochondrial P450s emerged as enzymes with highly specialized functions in steroid hormone metabolism in animals.

In vertebrates, steroid hormone biosynthesis is mediated by several P450s of which only members of the CYP11 family belong to the mitochondrial P450 class. Among them, CYP11A1 catalyzes the initial and rate-limiting step of steroid hormone biosynthesis in vertebrates, which is the side-chain cleavage (SCC) of cholesterol. The sequential SCC process involves three monooxygenase reactions and starts with the hydroxylation of the cholesterol side-chain at position C22. The second step involves the hydroxylation of 22*R*-hydroxycholesterol at position 20α and proceeds with the cleavage of the C20–C22 bond of 20α,22*R*-dihydroxycholesterol to form pregnenolone, the universal precursor of all steroid hormones [15]. For this reason, the emergence of the mitochondrial CYP11 family in basal vertebrates is considered as one of the most pivotal steps during evolution of steroid hormone biosynthesis.

Conventional phylogenetic analyses may allow estimation of the time when steroidogenic activities were acquired by mitochondrial P450s or identification of conserved amino acid sequences as well as structural motifs within the evolving steroidogenic P450 enzymes. However, they do not provide insights into the effects of substituted amino acids and associated functional changes that were necessary for the evolution of steroid metabolism in animals. These investigations exclusively rely on biochemical analyses of isolated or recombinantly produced P450 enzymes. However, despite the ever-increasing number of P450 sequences in animals, biochemical characterization of steroidogenic P450s has been mostly limited to a few extant members of mammalian P450 families including CYP11 [15–21], CYP17 [22–24], CYP19 [25], and CYP21 [26]. The problem is that the scarce knowledge, derived from functional analyses of P450 enzymes of closely related species, does not provide sufficient information to draw reliable conclusions about the underlying mechanisms that promoted early development of steroid hormone biosynthesis. Furthermore, biochemical analyses of extant P450 enzymes disregard ancient enzyme properties that may have emerged transiently but actually were essential for the evolution of steroid metabolism.

Just recently, the promising bioinformatic approach of ancestral sequence reconstruction (ASR) was successfully applied for the resurrection of protein ancestors of selected microsomal P450s, including the drug-metabolizing mammalian enzymes CYP3A4 [27], CYP2D6 [28], and CYP1B1 [29]. The ASR process can in general be described in three steps: firstly known amino acid sequences of the respective phylogenetic group are collated and aligned to produce a multiple sequence alignment (MSA); secondly, the MSA is used to build a phylogenetic tree, which provides

a visual and computational representation of the evolutionary relationships; and thirdly, both the phylogenetic tree and the MSA are used in tandem with an amino acid substitution rate matrix to infer the most likely ancestral state at each position in the alignment [30]. The ancestral state is predicted at each position independently and the results are concatenated to generate an entire ancestral sequence. The prediction is thereby determined by three components, the residues present in extant sequences and how they are aligned, the evolutionary relationship between the respective sequences provided in the MSA, and the most likely substitution that occurred according to the substitution rate matrix.

The ASR of microsomal P450 enzymes resurrected to date has provided intriguing insights into changes that translated to exceptional enzyme features such as increased thermostability, altered product selectivity, or improved catalytic activity [28,29]. Despite the recent success in resurrection of microsomal P450s, ASR has not been applied yet to mitochondrial cytochromes P450. For this reason, the aim of the present study was to use ASR for the first resurrection and biochemical characterization of mitochondrial P450 ancestors. The CYP11A1 subfamily was selected as target due to its key role in steroid hormone biosynthesis in vertebrates.

Results and discussion

Ancestral sequence reconstruction of CYP11A1

For the inference of the ancestor of CYP11A1 enzymes, a maximum-likelihood, joint reconstruction approach was employed. The MSA used for the reconstruction contained a total of 339 extant sequences, of which 185 sequences were CYP11A1 sequences and 154 were CYP11B/C sequences. The phylogenetic tree obtained conformed to the expected tree as proposed for CYP11s previously [8] and was in agreement with the tree of life to the level of vertebrate classes (Fig. 1). Accordingly, the CYP11A clade and the CYP11B/C clade branched separately and the land vertebrate CYP11Bs branched from fish CYP11Cs. Given that the two clades are predicted to be ohnologs that duplicated in the 2R whole-genome duplication event (the proposed two rounds of genome duplication occurring in early vertebrate evolution [31]), the estimated age of the ancestor of the vertebrate CYP11As (CYP11A_N1) is ~ 450 million years [32]. In order to gain insights into the evolution of CYP11A enzymes, we selected two nodes for synthesis and subsequent biochemical characterization of the respective ancestral CYP11A enzymes. CYP11A_N1 was of special interest since it showed only 55.0% sequence identity to the extant CYP11A1 form, suggesting potentially altered product or substrate selectivity which could provide novel insights into ancient steroid hormone biosynthesis. Furthermore, recent studies have revealed that microsomal ancestral P450s of similar age display a significant increase (up to 30 °C) in thermostability [27,28]. Thus, CYP11A_N1 was of additional interest to test whether the stabilization of the P450 fold via ASR also applies to mitochondrial P450 enzymes. The ancestor of mammalian CYP11A enzymes (CYP11A_Mammal_N101) shared 71.5% sequence identity with extant bovine CYP11A1 and was selected as an intermediate ancestor for characterization to gain insights into gradual evolutionary changes with respect to activity and stability compared with extant bovine CYP11A1 and CYP11A_N1. A multiple sequence alignment covering the whole amino acid sequences of all CYP11A1 variants is presented in Fig. 2.

Recombinant production of selected CYP11A1 variants

The ancestral CYP11A1 enzymes, CYP11A_Mammal_N101 and CYP11A_N1, as well as the bovine CYP11A1 variant were successfully co-expressed with the molecular chaperones GroES and GroEL in recombinant *E. coli* C43(DE3) cells. Initial measurements of the CYP11A1 content in the cell homogenates suggested very high expression levels for all CYP11A1 variants ranging from 1200 to 1600 nmol·L^{-1} cell culture, which represent the highest reported values for CYP11A1 yields in recombinant *E. coli* cells so far [33]. The nearly twofold improvement of CYP11A1 expression levels could be attributed to the utilization of the T7 promoter which is considerably stronger than the previously used trc promoter [34]. However, it is notable that other ancestral P450s have also shown high recombinant expression yield, so a beneficial effect of protein stability on expression yield cannot be excluded. The his-tagged CYP11A1 proteins were purified from cell homogenates via a two-step purification process that involved immobilized metal ion affinity chromatography (IMAC) and subsequent ion exchange chromatography (IEC). Homogeneity of the purified CYP11A1 enzymes (54 kDa) was confirmed via SDS/PAGE analysis for both ancestors, CYP11A_Mammal_N101 and CYP11A_N1, whereas the bovine CYP11A1 showed minor contamination with the molecular chaperone, GroEL (60 kDa) (Fig. 3). UV-Vis spectroscopy of the

purified enzymes revealed the characteristic low spin state spectrum of oxidized cytochromes P450 including the major Soret band (γ) at 418 nm as well as the minor α- and β-bands at in the Q-band region, respectively. Upon reduction with sodium dithionite and binding of carbon monoxide (CO), all recombinant CYP11A1 proteins exhibited the typical bathochromic shift of the γ-band toward 450 nm, suggesting proper protein folding and orientation of the prosthetic heme group in the active center of the enzymes [35].

Both ancestral CYP11A1 enzymes are active

CYP11A1 is particularly known for its role in initiating *de novo* steroidogenesis in vertebrates by converting cholesterol to pregnenolone [19]. For this reason, we first aimed to test the purified CYP11A1 enzymes for side-chain cleavage (SCC) activity toward the natural substrate, cholesterol, in reconstituted *in vitro* assays. All CYP11A1 enzymes, including notably both ancestral CYP11A1 enzymes, were found to be active and showed identical product patterns with pregnenolone being the main reaction product (data not shown). As demonstrated by time-dependent conversions (Fig. 4A), the extant CYP11A1 variant showed a nearly twofold higher cholesterol conversion rate than the ancestral enzymes under the tested reaction conditions. After 30 min of incubation, bovine CYP11A1 had converted a total of 83.2 ± 4.2 μM cholesterol while the ancestral enzymes CYP11A_Mammal_N101 and CYP11A_N1 showed significantly lower cholesterol conversion values of 41.4 ± 7.7 μM and 36.8 ± 1.4 μM, respectively. Since the CYP11A1 ancestors CYP11A_Mammal_N101 and CYP11A_N1 share just 71.6% and 55.0% identity with the protein sequence of the extant CYP11A1 enzyme, the

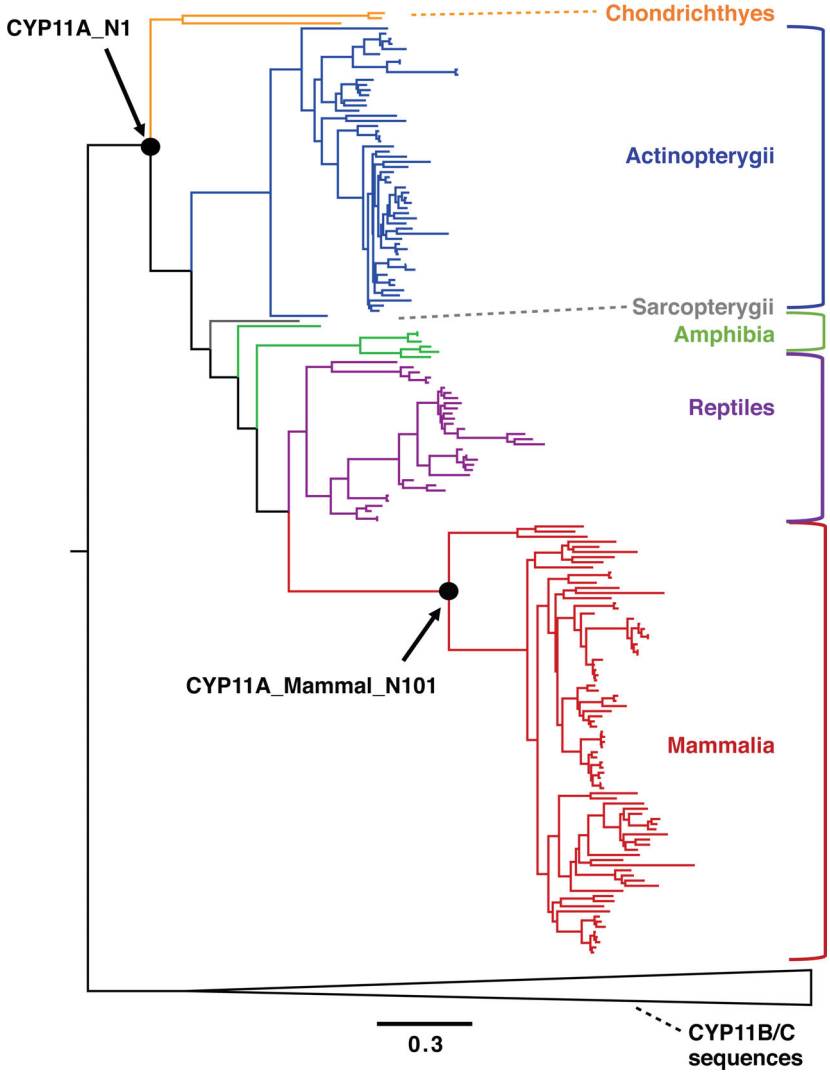

Fig. 1. Maximum-likelihood phylogenetic tree derived from the set of CYP11 sequences used here. The branches representing the CYP11B and CYP11C subfamilies are collapsed to better show the topology of the CYP11A branch. The bar indicates branch length in substitutions per site, a measure of evolutionary age.

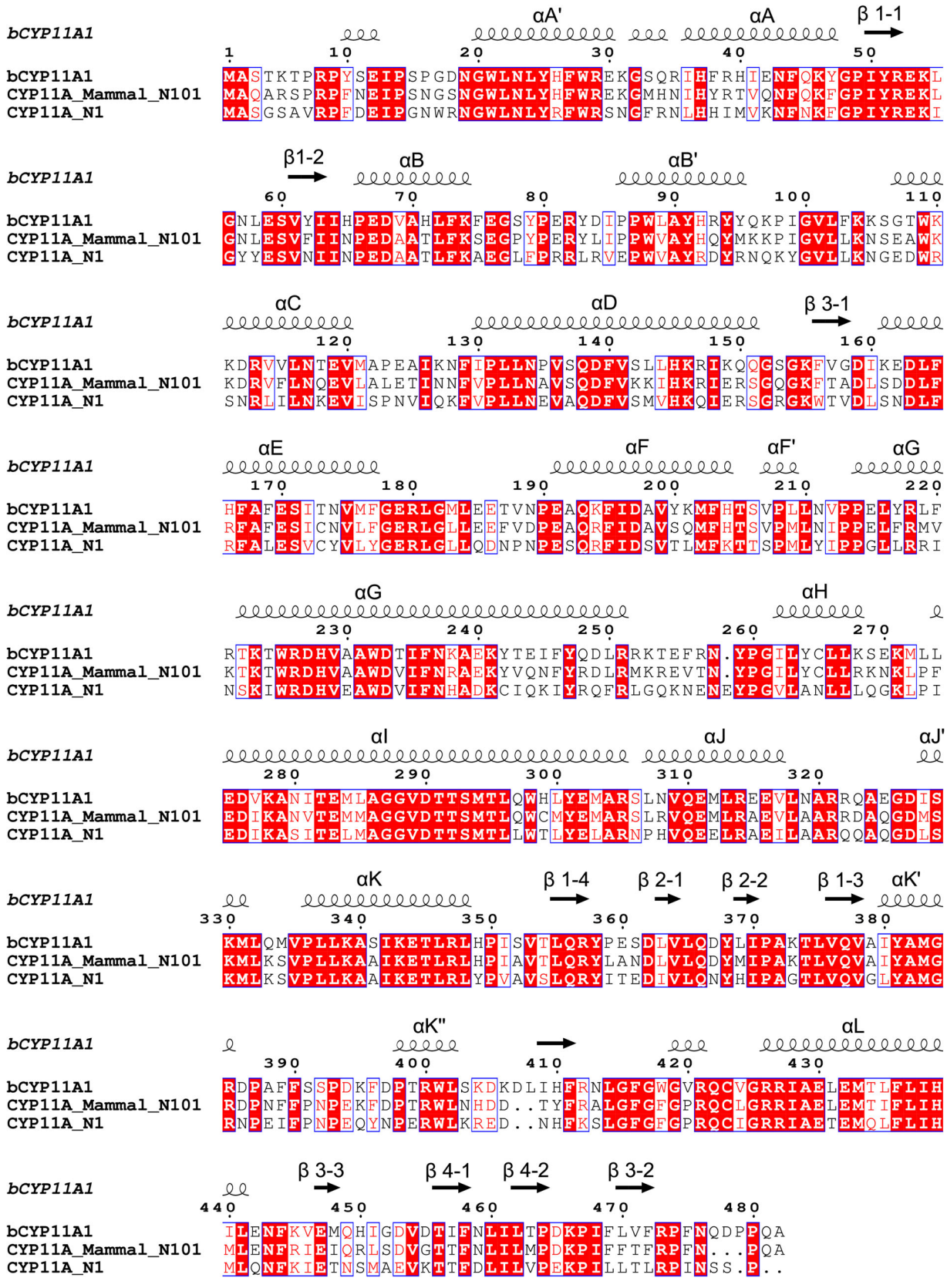

Fig. 2. MSA of ancestral CYP11A enzymes and bovine CYP11A1. Numbers on top are shown for the N-terminal modified form of bovine CYP11A1 according to the N-terminal modification by Wada et al. [62]. Red boxes indicate conservation with respect to amino acid identity, red characters show amino acid similarity, and blue frames show homologous regions. The MSA was created using MAFFT version 7 [81], and secondary structure elements of bovine CYP11A1 were mapped with the ESPript 3.0 server [82] using bovine CYP11A1 (PDB entry: 3MZS) as the structural template [16]. The sequence of bCYP11A1 shows the bCYP11A1 mutant used in this study with asparagine to aspartic acid mutation at position corresponding to position 290 [65]. Annotation of the secondary structure elements was adopted from elsewhere [83].

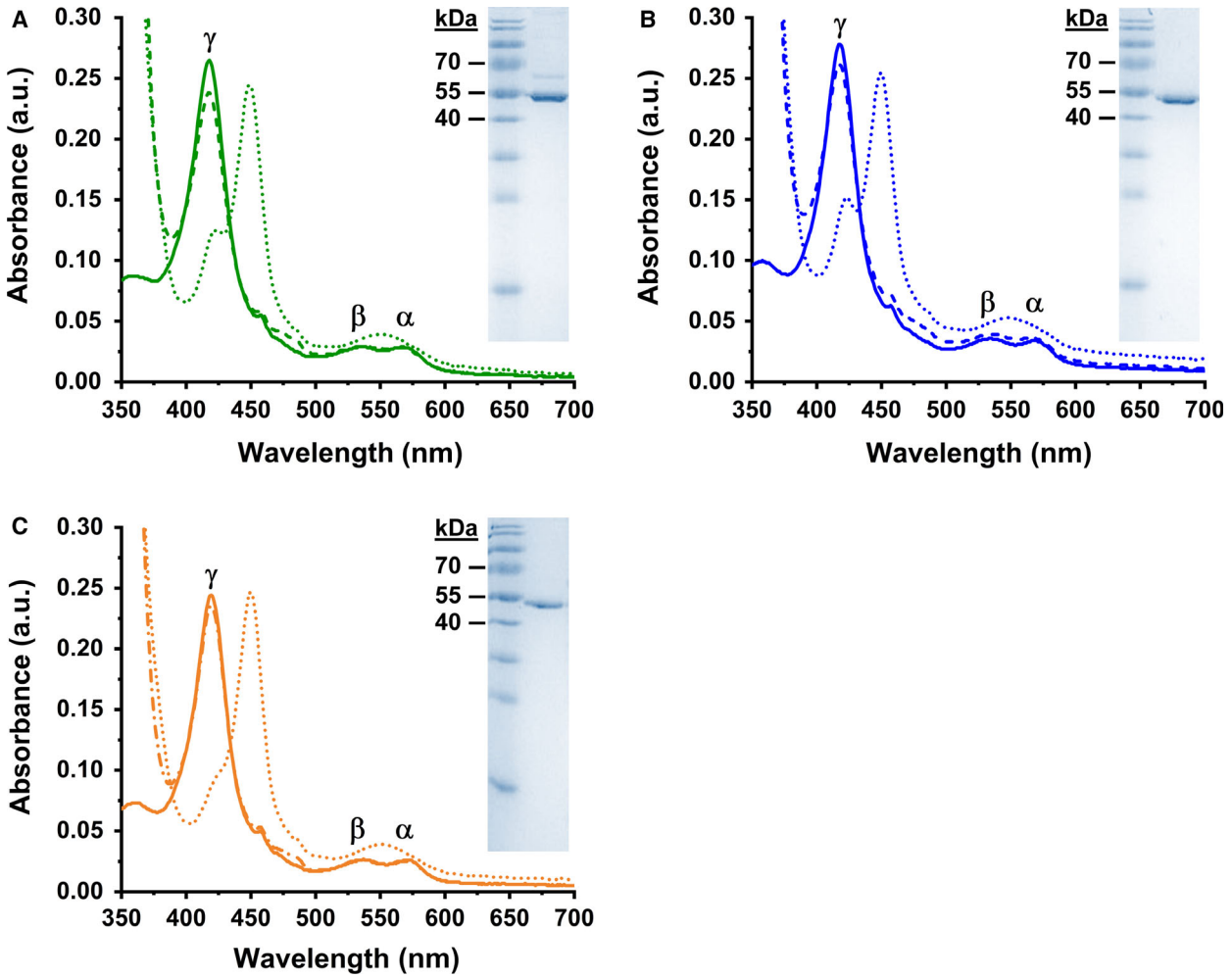

Fig. 3. UV-Vis spectral features of the purified P450 enzymes. A, Bovine CYP11A1, B, ancestor CYP11A_Mammal_N101, and C, ancestor CYP11A_N1. Characteristic absorption spectra are shown for the enzymes in the oxidized state (solid lines), dithionite-reduced state (dashed lines), and dithionite-reduced CO-bound state (dotted lines). Insets display the SDS/PAGE analysis of the corresponding CYP11A1 enzyme following IMAC and IEC purification. The peqGOLD prestained protein marker IV (VWR) was used as standard.

reduced cholesterol SCC activity could be due to several factors. These include compromised electron transfer, limited access of the sterol substrate to the active site, or spatial constraints within the substrate-binding pocket of the enzyme. For this reason, the spectral dissociation constant (K_d value) of cholesterol was determined for each CYP11A1 variant (Table 1). The binding affinity of cholesterol to both ancestral enzymes, CYP11A_N1 and CYP11A_Mammal_N101, was 2–2.4 times lower than that of cholesterol to bovine CYP11A1 and approximately reflected the reduced SCC activities observed for the time-dependent cholesterol conversion experiment. Similarly, the NADPH consumption rates for CYP11A_Mammal_N101 (15.8 ± 0.8 nmol·min^{-1}·nmol^{-1} P450) and for CYP11A_N1 (18.4 ± 0.9 nmol·min^{-1}·nmol^{-1} P450) were considerably low, whereas the high

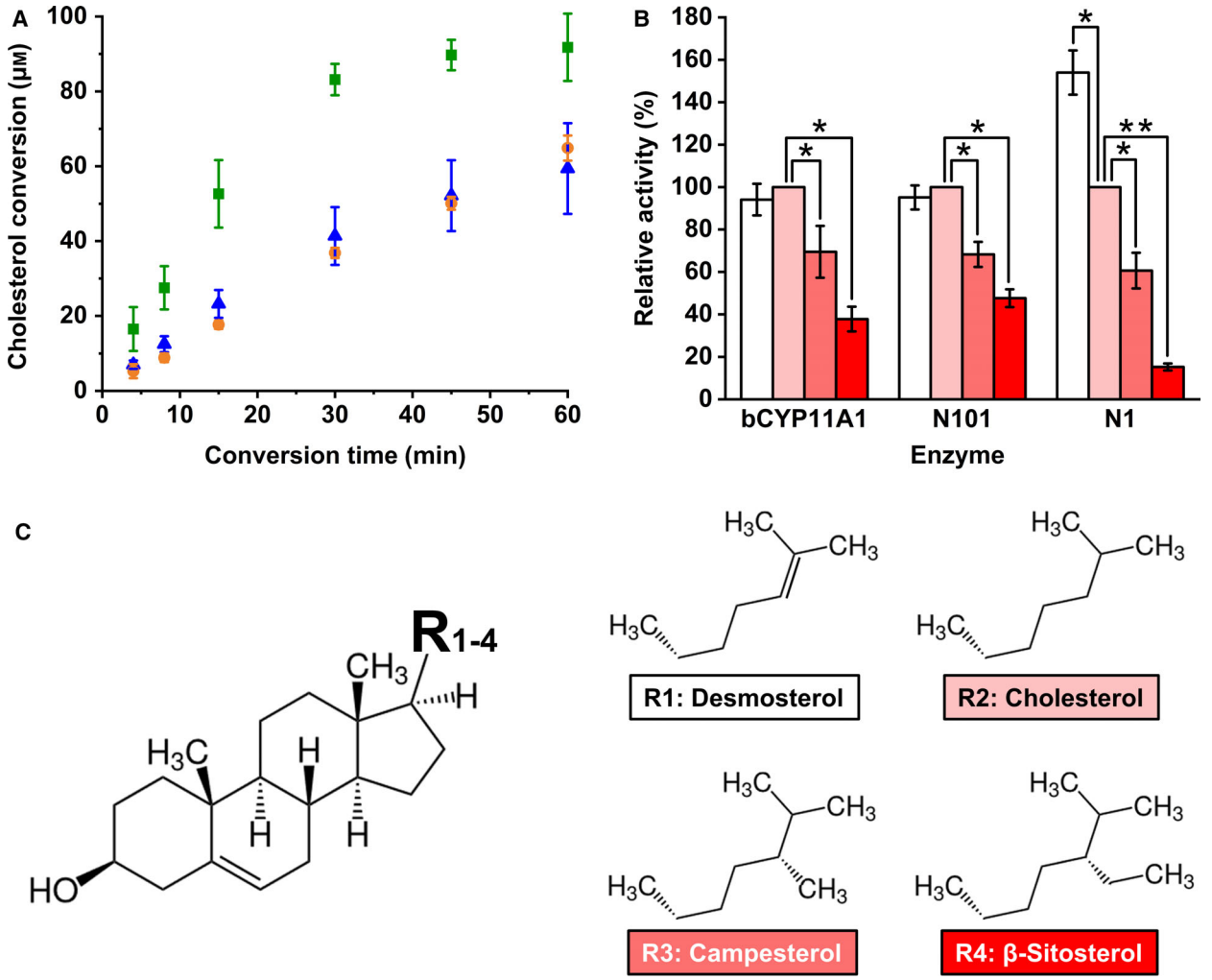

Fig. 4. Activity and substrate plasticity of the purified CYP11A1 enzymes. A, Time-dependent conversion of cholesterol with bovine CYP11A1 (green symbols), CYP11A_Mammal_N101 (blue symbols), and CYP11A_N1 (orange symbols). Data are presented as mean values and standard deviations of three independent experiments. B, Relative activities of the CYP11A1 enzymes with different sterol derivatives desmosterol (white bars), cholesterol (pink bars), campesterol (light red bars), and β-sitosterol (red bars). Data were collected from three independent experiments and scaled to the corresponding enzyme activity with cholesterol. Error bars represent the relative standard deviation (RSD) of the scaled enzyme activities. Statistical analysis was performed using two-sample t-tests, $*P \leq 0.05$, $**P \leq 0.001$. C, Molecular structures of the tested sterol derivatives desmosterol (R1), cholesterol (R2), campesterol (R3), and β-sitosterol (R4). Note that the complexity of the sterol side chain increases from R1–R4.

Table 1. Substrate affinity of bovine CYP11A1, CYP11A_Mammal_N101, and CYP11A_N1 toward sterol substrates with varying side chains. Dissociation constants (K_d values) were determined by difference spectroscopy (A_{390nm}–A_{420nm}) in the presence of bovine adrenodoxin ($Adx_{1–128}$). Data were fitted using hyperbolic regression analysis in OriginPro 2020 and represent the mean values and standard deviation of three individual experiments. n. q. Not quantifiable. The substrate-induced spectral change was too low for proper K_d value determination. n. d. Not detectable. Titration of the substrate did not induce spectral changes.

Variant	Cholesterol [μM]	Desmosterol [μM]	Campesterol [μM]	β-Sitosterol [μM]
Bovine CYP11A1	3.3 ± 0.4	4.0 ± 0.2	5.9 ± 0.3	8.8 ± 1.1
CYP11A_Mammal_N101	7.8 ± 0.4	8.3 ± 1.1	8.8 ± 1.2	12.4 ± 0.8
CYP11A_N1	6.8 ± 0.7	4.7 ± 0.5	n. q.	n. d.

NADPH consumption rate of bovine CYP11A1 (25.8 ± 3.8 nmol·min^{-1}·nmol·$^{-1}$ P450) indicated more efficient interactions with bovine Adx than both CYP11A1 ancestors. Previous studies indeed demonstrated that binding of cholesterol to CYP11A1 increases binding of Adx and *vice versa*, thereby promoting cholesterol conversion in a cooperative way [36]. Substrate-induced structural changes to the electrostatic landscape at the P450 surface seem to play a pivotal role in this heterotrophic modulation process [37]. Four positively charged residues in the bovine CYP11A1 protein sequence, corresponding to positions K268, K404, K406, and R427, were identified to be important for electrostatic interactions of CYP11A1 with Adx [38]. While the key residue R427 was well conserved within the protein sequences of extant and ancestral CYP11A1 enzymes, the other residues differed significantly from their counterparts in the extant enzyme. Most notably K406, which is involved in electrostatic interactions with D72 of Adx, was substituted in both CYP11A1 ancestors by the negatively charged residue aspartate leading to strong repulsive electrostatic forces between Adx and CYP11A_Mammal_N101 or CYP11A_N1 [39]. Since substitution of K406 with the neutral amino acid glutamine has already been reported to impair the binding of Adx to bovine CYP11A1 significantly [38], the K406D mutation in both ancestral CYP11A1 enzymes might have comparable or even worse effects on Adx binding and would explain the observed decrease in cholesterol SCC activity of CYP11A_Mammal_N101 and CYP11A_N1. Amino acid substitutions of K404 had comparably low effects, while the contribution of K268 to Adx binding has not been determined yet. Assuming that mitochondrial P450s evolved from a common microsomal ancestor, favorable changes of redox partner interaction sites toward components of an alternative ETS in the mitochondrial matrix may not only have been crucial for microsomal P450s to retain their oxygenase function but also to increase affinity to novel substrates of the emerging steroidogenic pathway. The distinct differences of redox partner interaction sites and associated catalytic activities between the ancestral CYP11A1 enzymes and the extant CYP11A1, indeed, support this hypothesis and might provide unique functional insights into sequential ETS adaptation processes.

The CYP11A1 ancestor, CYP11A_N1, prefers the cholesterol precursor, desmosterol

The CYP11A1 enzymes were also tested for substrate specificity with diverse cholesterol analogs that mainly differed in size and complexity of the sterol side chain. These included the cholesterol precursor, desmosterol (5,24-cholestadien-3β-ol), as well as the plant sterols, campesterol (24α-methylcholesterol), and β-sitosterol (24β-ethylcholesterol). The activity of all CYP11A1 variants appeared to be largely affected by the structural complexity of the sterol side chain, since bulkier chains evidently decreased the efficiency of sterol conversions with β-sitosterol being the less preferred sterol substrate (Fig. 4B). Determination of the binding affinities for each cholesterol analog further solidified this observation since substrates with bulkier sterol side chains showed increasingly worse spectral K_d values compared with that of cholesterol (Table 1). This trend was particularly apparent for the K_d values of both mammalian CYP11A1 variants and is in good agreement with previous studies that reported desmosterol as an equivalent substrate to cholesterol [40,41] while phytosterols were converted less efficiently to pregnenolone by mammalian CYP11A1 enzymes [42,43]. Despite supplementation of bovine Adx to improve the spectral response, binding of phytosterols to the active site of CYP11A_N1 was not observed spectrophotometrically under the tested conditions. Although substrate binding to cytochromes P450 does not necessarily induce the characteristic type I spectral shift [44,45], the fact a spectral shift was exclusively lacking when these bulky sterols were titrated to the whole subfamily ancestor CYP11A_N1 suggests impaired access of the substrate to the active site of CYP11A_N1 or decreased ability to replace water as the sixth ligand. Intriguingly, the earliest ancestor CYP11A_N1 was the only CYP11A1 variant that showed a lower K_d value, and consequently, higher conversion rates for the cholesterol precursor desmosterol than for cholesterol itself, suggesting sequential adaptation processes of the CYP11A1 active site that allowed the conversion of novel intermediates of the evolving cholesterol metabolism. Since desmosterol also represented the sterol with the least bulky side-chain (Fig. 4C), spatial constraints within the active site of CYP11A_N1 could be the determining factors in this adaption process.

The CYP11A1 ancestor, CYP11A_N1, is a robust biocatalyst

As mentioned previously, there is evidence that mitochondrial P450s originated from an early microsomal P450 ancestor [8,9]. Once relocated to the mitochondrial matrix, this ancestor would have been faced with substantial changes in the redox environment, most notably, with the lack of a compatible ETS. While residing on the cytoplasmic side of the ER membrane, microsomal P450s received electrons from a NADPH-dependent reductase (CPR). In the mitochondrial matrix, however, a cognate CPR enzyme was not present and consequently, ancient mitochondrial P450s would have had to receive electrons from an

alternative ETS to retain their monooxygenase functions. Initial interactions of ancient mitochondrial P450s with incompatible components of an already existing mitochondrial ETS [46] might have been inefficient and, thus, might have resulted in enhanced formation rates of reactive oxygen species (ROS) during the P450 catalytic cycle. Extensive ROS formation due to poor electron coupling has been reported to lead to rapid inactivation of extant P450 enzymes, particularly in steroidogenic tissues of the adrenal cortex [47]. For this reason, we hypothesized that an increased tolerance against high ROS levels was crucial for ancient P450s to stay active and evolve specialized steroidogenic functions in the mitochondria. In order to test this hypothesis, the extant CYP11A1 enzyme as well as its ancestors CYP11A_Mammal_N101 and CYP11A_N1 were exposed for 10 min to increasing levels of hydrogen peroxide (H_2O_2) to get a broad picture on the stability of the ancestors against ROS (Fig. 5A), although exact H_2O_2 concentrations in adrenal mitochondria are not known yet. The residual cholesterol SCC activity was tested as described in the Materials and methods section. Intriguingly, half of the extant CYP11A1 enzymes were already inactivated at H_2O_2 levels of 11 ± 1 nmol·pmol^{-1} P450 whereas its closely related ancestor CYP11A_Mammal_N101 was able to withstand approximately sixfold higher H_2O_2 levels of up to 69 ± 7 nmol·pmol^{-1} P450. However, the most ancient CYP11A1 enzyme CYP11A_N1 showed a striking increase in ROS stability, tolerating 20-fold higher H_2O_2 levels of up to 190 ± 35 nmol·pmol^{-1} P450.

These results might indicate that ancient mitochondrial P450s were not as susceptible to inactivation by ROS as their extant counterparts. Furthermore, this ability seems to have been lost as evolution of steroidogenic P450 enzymes proceeded to a point at which enhanced stability against ROS no longer provided any evolutionary advantage for ancient mitochondrial P450s. A gradual loss of ROS stability might hint toward convergent adaption processes of the redox partner interaction sites that ultimately resulted in an optimized electron transfer and, in turn, to lower ROS formation. The extraordinary resistance against ROS raised the question as to whether ancestor CYP11A_N1 might have acted as a peroxygenase-like P450, which is able to utilize H_2O_2 to drive catalysis. However, incubation of CYP11A_N1 with cholesterol and H_2O_2 levels of up to 100 nmol·pmol^{-1} P450 did not lead to any detectable product formation (data not shown). To the best of our knowledge, enhanced stability toward ROS has not been reported for the resurrected microsomal P450 ancestors and could, in fact, be a unique enzyme feature of resurrected mitochondrial P450s.

Microsomal P450 ancestors were shown in previous studies to possess substantially increased thermostability over their descendants [27–29]. For this reason, we tested the thermostability properties of the purified ancestral CYP11A1 enzymes via thermal unfolding studies. Initial CD spectroscopy in the far UV demonstrated characteristic signatures of helix-rich proteins with two signal minima at wavelengths of approximately 210 nm and 220 nm, suggesting similar secondary structure element compositions for all tested CYP11A1 enzymes (Fig. 5B). The enzymes were exposed to a linear temperature gradient, and signal changes at a wavelength of 210 nm were used to calculate the fraction of unfolded CYP11A1 at the indicated temperature. In general, all tested CYP11A1 variants showed sigmoidal melting curves suggesting highly cooperative protein unfolding kinetics. The transition, at which half of the protein fraction still existed in a folded state, represented the individual melting temperature (T_m) of the CYP11A1 enzymes (Fig. 5C). Apparently, the T_m of the ancestral CYP11A_Mammal_N101 enzyme did not differ notably from that of the extant CYP11A1 enzyme. Both showed mesophilic T_m values of approximately 49.2 ± 0.3 °C. It is noteworthy, that a second transition was evident in the melting curve of the purified bovine enzyme, which might derive from the previously described contamination with a small fraction of the molecular chaperone GroEL. However, the T_m value of the vertebrate ancestor CYP11A_N1 was strikingly increased by 25 °C to 74.2 ± 0.4 °C and even exceeded the reported melting temperatures of some naturally thermostable P450s including CYP154H1 (67 °C) from *Thermobifida fusca* [48], CYP231A2 (65 °C) from *Picrophilus torridus* [49] or CYP450-T2 (56.8 °C), a not-yet-classified P450 derived from metagenomic data of the Binh Chau hot spring [50]. In contrast to these P450s, CYP11A_N1 was not likely to have existed in a hot environment where thermal robustness was required, given that it is expected to have evolved around the time vertebrates emerged on earth and where temperatures were similar to those presently [9]. However, such high thermal robustness is in line with other ancestral vertebrate P450s previously reconstructed and characterized [27–29]. We propose that the observed high level of thermostability seen in these ancestors is a carryover from older P450 enzymes from primordial organisms which may have existed in hot, marine environments on the early Earth. This thermal stability may have facilitated the diversification of the mitochondrial P450 enzymes toward new

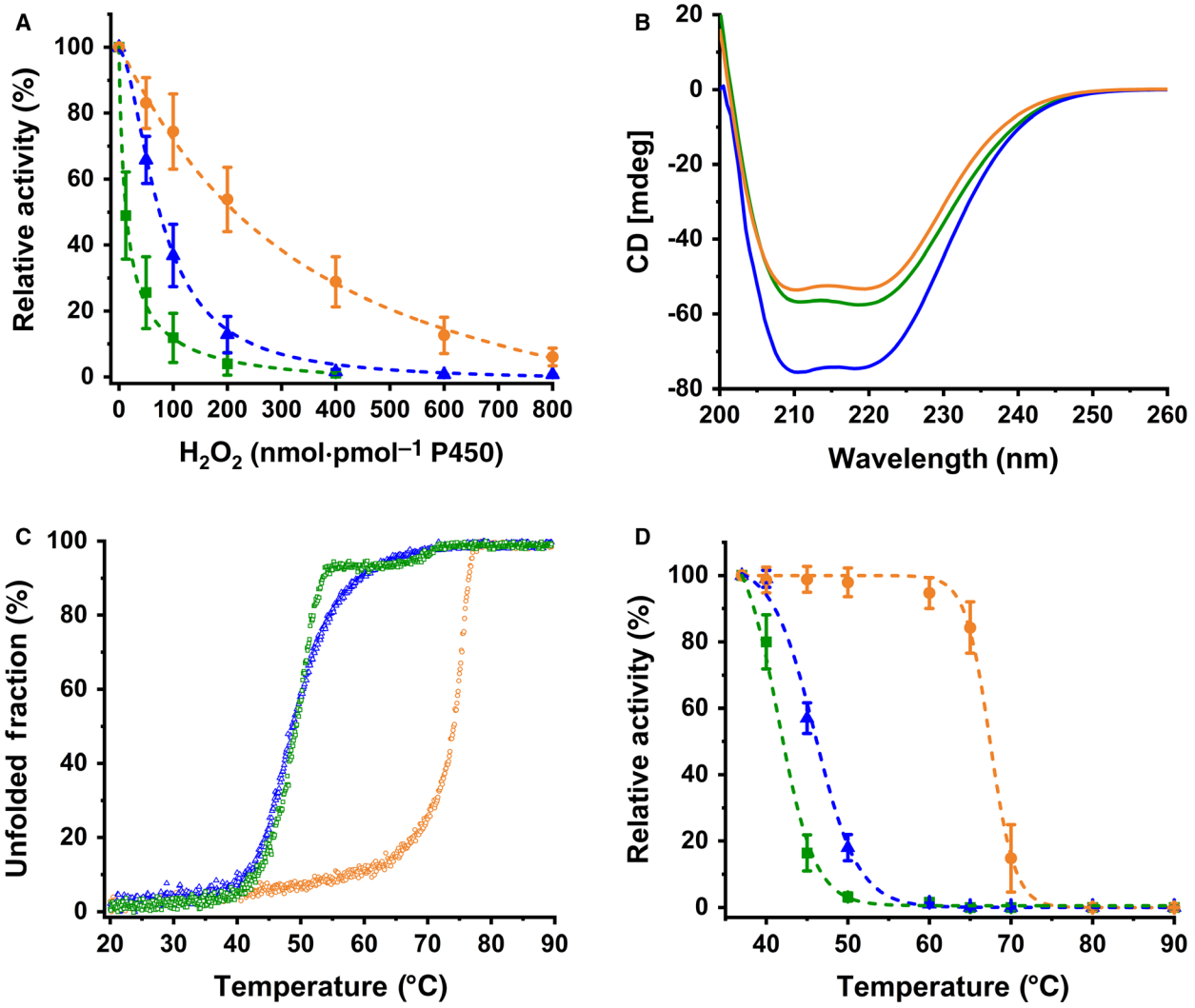

Fig. 5. Comparison of the ROS resistance and thermostability of the bovine CYP11A1 enzyme (green lines and symbols) and the ancestral CYP11A1 enzymes CYP11A_Mammal_N101 (blue lines and symbols) and CYP11A_N1 (orange lines and symbols). A, Residual activity of the CYP11A1 enzymes following incubation with increasing concentrations of H_2O_2 for 10 min. Data were collected from three independent experiments and scaled to the corresponding cholesterol SCC activity of the untreated enzymes. Error bars represent the RSD of the scaled enzyme activities. B, Representative CD spectra for each enzyme were recorded at 20 °C with enzyme concentrations of 5 μM in 50 mM potassium phosphate buffer, pH 7.4. C, Thermal melting curves of ligand-free CYP11A1 enzymes. Changes in the molar ellipticity at 210 nm were used to calculate the fraction of unfolded proteins at the indicated temperature. Data represent the mean values of two technical replicates. D, Residual activity of the CYP11A1 enzymes following exposure to elevated temperatures for 10 min. Data were collected from three independent experiments and scaled to the corresponding cholesterol SCC activity at 37 °C. Error bars represent the RSD of the scaled enzyme activities.

functions in steroid metabolism by buffering the potentially destabilizing effect of mutations. However, in the absence of an explicit selection pressure for thermal robustness, such diversification may have led to a loss of thermal robustness over time, finally resulting in the rather thermolabile extant P450 forms observed today. Alternatively, we cannot exclude the possibility that the increase in thermal robustness is an artifact of the maximum-likelihood method used for inferring the ancestral state [51]. However, the degree of stabilization expected from such an artifact would be expected to lead to an increase in T_m of a few degrees [52] rather than the ~ 25 °C between N1 and the extant bovine enzyme seen here.

Determination of the $^{10}T_{50}$ values, the temperature at which the initial cholesterol SCC activity is reduced

to 50% after 10 min of heating, confirmed the previously described results of the thermal unfolding studies (Fig. 5D). While the extant CYP11A1 enzyme showed a moderate $^{10}T_{50}$ value of 42.0 ± 0.9 °C, the corresponding value for the ancestor CYP11A_N1 was significantly increased to 67.5 ± 0.8 °C. Again, the difference in thermostability between bovine CYP11A1 and CYP11A_N1 was calculated to be 25 °C, which is consistent with that of the CD measurements. Although the T_m values for the ancestral CYP11A_Mammal_N101 and the extant CYP11A1 enzymes were determined to be identical in CD measurements, the $^{10}T_{50}$ of CYP11A_Mammal_N101 seemed to be slightly increased by 3 °C to 45.4 ± 0.5 °C, suggesting a different segmental unfolding process for CYP11A_Mammal_N101 that might favor the integrity of its catalytic center.

Determinants of thermostability for the CYP11A1 ancestor, CYP11A_N1

The thermal stability of proteins, particularly in comparison with their mesophilic counterparts within the same family, has been attributed to a number of factors such as additional disulfide bonds, expanded aromatic and electrostatic networks, differences in hydrogen-bonding, shorter loop structures, and a higher degree of buried hydrophobic surface area [53,54]. We computed and analyzed these factors on the basis of homology models for the extant and ancestral CYP11A1 enzymes. First of all, we found a unique tyrosine to cysteine substitution at position 242 in the G-helix of the thermostable vertebrate ancestor CYP11A_N1, which was found in a promising position to form a putative disulfide bond with a corresponding cysteine at position 173 in the E-helix. While the cysteine at position 173 was also present in the sequence of the ancestral CYP11A1 enzyme CYP11A_Mammal_N101, the extant bovine CYP11A1 variant did not contain either of the cysteine residues that possibly contributed to the enhanced thermostability of the CYP11A1 ancestor CYP11A_N1. Based on the atomic coordinates in the obtained homology model of CYP11A_N1, the distance between the thiol groups of C173 and C242 is 6.4 Å and, thus, beyond the contact distance for disulfide bond formation (Fig. 6A) [55]. However, we hypothesized that a certain degree of structural flexibility, particularly known for the F/G-loop region and adjacent helices of cytochromes P450, could allow both cysteine residues to come close enough for disulfide bond formation. In order to test this hypothesis, the cysteine at position 242 was replaced via site-directed mutagenesis by a tyrosine, the corresponding amino acid in the protein sequences of the mesophilic bovine CYP11A1 and CYP11A_Mammal_N101. Thermal unfolding studies with the purified C242Y mutant of CYP11A_N1, however, did not show any difference in melting temperature compared with that of the parental CYP11A_N1 (Fig. 6B), thus excluding disulfide bond formation as a potential factor for stabilization of ancestor CYP11A_N1. Detailed investigations of the CYP11A1 structures showed not only that all three homology models exhibited the same arrangement of secondary structural elements in space but also identical lengths of loop structures. Thus, this factor can also be ruled out as explanation for the thermostability of the CYP11A1 ancestor, CYP11A_N1. On the other hand, this is not surprising, as the same structural template (human CYP11A1, PDB entry 3N9Y) was used [19]. Furthermore, there was no indication that differences in amino acid composition (hydrophobic vs. hydrophilic) would be decisive, since CYP11A_Mammal_N101 exhibited the highest fraction of hydrophobic amino acids (Table S1) but was not the most thermostable enzyme. Consideration of intramolecular contacts (hydrogen bonds, salt bridges, and Cβ beta contacts) showed likewise no conclusive picture (Table S2). Although the three CYP variants do not substantially differ in their amino acid composition, the accessibility of their hydrophobic surface area, which is reflected by calculated solvation energies, could still be different. Surprisingly, the ancestral CYP11A_N1 showed the most unfavorable value. Thus, an optimized polar surface could not be the driving force for thermostability in this case. Still, the degree of hydrophobic internal packing may vary in between these P450 variants and, thus, contribute to increased thermostability as recently shown for ancestral enzyme variants of the human drug-metabolizing cytochrome P450 CYP3A4 [27]. The solvation score of the SWISS-Model is indicative of that and points toward a preferred situation in the ancestral CYP11A_N1 (Table S3). A better hydrophobic packing would also be reflected by better nonbonding (van der Waals) interactions. Moreover, a more favorable arrangement of hydrogen bonds and salt bridges will give rise to preferable electrostatics. These factors were computed as force field energies (Table S4) and indicate that CYP11A_N1 is optimal regarding electrostatic interactions (hydrogen bonds and salt bridges), but not in terms of hydrophobic packing, which is also apparent by the worse solvation energy. Since CYP11A_Mammal_N101 contains more hydrogen bonds, but the same number of salt bridges, it becomes apparent that those in the ancestral CYP11A_N1 must be more optimally arranged.

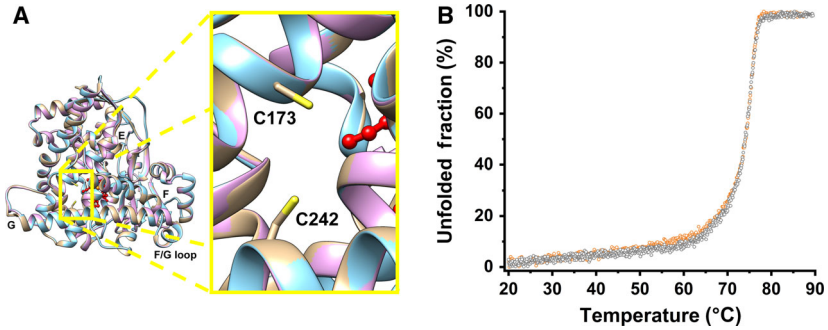

Fig. 6. Evaluation of a putative disulfide bond in the structure of the thermostable CYP11A_N1. A, Superposition of the homology models of the bovine CYP11A1 enzyme (cyan) and the ancestral variants, CYP11A_Mammal_N101 (magenta), and CYP11A_N1 (apricot). The crystal structure of human CYP11A1 (PDB entry: 3N9Y) was used as template for the protein structure modeling via the SWISS-Model web server [74]. The yellow rectangle displays a detailed view of the protein region where the thiol groups of two cysteine residues (at positions 173 and 242) might be involved in the formation of a putative disulfide bond in CYP11A_N1. B, Thermal melting curves of the ligand-free parental CYP11A_N1 (orange open circles) and the C242Y mutant (light gray open circles). Changes in the molar ellipticity at 210 nm were used to calculate the fraction of unfolded proteins at the indicated temperature. Data represent the mean values of two technical replicates.

Therefore, we conclude that the thermostability of the ancestral CYP11A_N1 is caused by a more favorable network of salt bridges that, in summary, stabilize the interactions between the secondary structural elements better than in the other CYP11 enzymes.

Conclusion

The emergence of mitochondrial P450 seems to coincide with the development of steroid hormone biosynthesis in chordates. Several studies on the phylogenetic relationships between cytochromes P450 from basal vertebrates suggested a plausible model for the development and diversification process of mitochondrial P450s from a common microsomal ancestor [8,9]. However, these studies did not provide any insights into the structural changes that were necessary for mitochondrial P450s to adapt to the novel environment in the mitochondrial matrix or to develop novel steroidogenic functions. Here, we used the strategy of ancestral sequence reconstruction to resurrect the first ancestral enzymes of mitochondrial P450s. Our findings derived from diverse *in vitro* studies revealed unique properties of the ancient CYP11A1 enzymes. For the first time, cytochrome P450 ancestors were shown to exhibit increased stability toward elevated levels of ROS resulting from poor electron coupling due to inefficient redox partners' interactions, as demonstrated for both ancestral CYP11A1 enzymes using an extant mitochondrial electron transfer system. Since ROS formation leads to rapid inactivation of P450 biocatalysts, the tolerance of ancient CYP11A1 enzymes toward high levels of ROS actually might have helped ancestral P450s to stay active and undergo further crucial diversification processes in the mitochondrial matrix. Furthermore, the ancestral enzyme CYP11A_N1 seemed to prefer the cholesterol precursor, desmosterol, as sterol substrate, a novel finding that suggests convergent evolution processes for steroidogenic P450s and enzymes involved in cholesterol biosynthesis. Finally, the ancestor N1 exhibited a striking increase in thermostability compared with the extant CYP11A1 enzyme, a desirable feature that will significantly contribute to the development and application of CYP11A_N1 as promising P450 biocatalyst for the biotechnological production of steroid hormones.

Materials and methods

Ancestral sequence reconstruction

ASR was performed using extant sequences obtained by searching the UniProt and NCBI databases using bovine CYP11A1 (UniProt accession no. P00189) and bovine CYP11B1 (UniProt accession no. P15150) sequences and Basic Local Alignment Search Tool (BLAST) [56] to search for homologue sequences. The collected sequences were clustered using CD-HIT [57] with a cutoff of 99.99% to remove duplicate sequences. Subsequently, sequences were aligned iteratively using MAFFT [58], followed by iterative rounds of manual curation using Jalview [59] to remove anomalous sequences (containing apparent frameshifts, internal gaps and insertions, and short sequence fragments (< 400 ~ amino acids). A total of 339 sequences were included in the final alignment, which was then used to

generate the phylogenetic tree with the RaxML program [60]. The GRASP web server was used for the ancestral inference [61]. The N termini of the inferred ancestors were modified according to Wada et al. [62] and the N-terminal truncation of inferred sequences was done based on the MSA and the alignment with the bovine CYP11A1 extant form. The resulting novel amino acid sequences were reverse translated to cDNA avoiding the *Nde*I, *Hin*dIII, *Sal*I, and *Xba*I restriction sites.

Reagents and enzymes

Dehydrated culture media were acquired from Becton, Dickinson and company (Franklin Lakes, NJ, USA). PCRs were performed with Phusion™ High-Fidelity DNA-Polymerase from Thermo Fisher Scientific GmbH (Karlsruhe, Germany). Oligonucleotides were synthesized by Eurofins Genomics GmbH (Ebersberg, Germany). Restriction enzymes as well as the Instant Sticky-end Ligase were purchased from New England Biolabs GmbH (Frankfurt am Main, Germany). Glucose-6-phosphate dehydrogenase was purchased from Roche (Basel, Switzerland). Cholesterol oxidase and catalase were acquired from Merck KGaA (Darmstadt, Germany). Sterols were purchased from TCI Deutschland GmbH (Eschborn, Germany). Isopropyl-β-D-thiogalactopyranoside, 5-aminolevulinate, and the reduced form of NADPH were acquired from Carbolution Chemicals GmbH (Saarbrücken, Germany). HPLC grade solvents and water were obtained from Fisher Scientific GmbH (Kehl, Germany). All other chemicals were acquired from Sigma-Aldrich (Schnelldorf, Germany).

Bacterial strains, expression vectors, and culture conditions

All bacterial strains and vectors used in this study are listed in Table S5. *Escherichia coli* TOP10 cells were used for the assembly and propagation of all vector constructs. Recombinant production of the CYP11A1 variants was carried out in *E. coli* C43 (DE3) cells. The bovine redox partners' adrenodoxin ($Adx_{1–128}$) and adrenodoxin reductase (AdR) were expressed in *E. coli* BL21 (DE3) cells. All *E. coli* strains were transformed using a standard calcium chloride ($CaCl_2$) transformation process [63]. Growth media for recombinant *E. coli* cells were supplemented with 50 $\mu g \cdot mL^{-1}$ kanamycin for the chaperone GroES/EL encoding vector pGro12 and 100 $\mu g \cdot mL^{-1}$ ampicillin for pET17b-derived expression vectors. Seed cultures for all bacterial cells were prepared in LB-Miller medium [64] supplemented with appropriate antibiotics. Expression cultures for recombinant protein production were prepared accordingly in 200 mL of TB medium [64] supplemented with 1 mM 5-aminolevulinate and inoculated with 2 mL of the corresponding seed culture.

DNA manipulation

The gene strings of all CYP11A1 variants were synthesized by Eurofins Genomics GmbH. The extant CYP11A1 form used in this study is a bovine CYP11A1 mutant with an asparagine to aspartic acid substitution at position 290 [65]. The 3′ ends of the cDNAs were extended with a hexa-histidine coding sequence (5′-CATCATCACCATCAT CAT-3′) to facilitate enzyme purification via affinity chromatography. Standard restriction digestion and ligation were used to insert the genes between the *Nde*I and *Hind*III sites of the pET17b expression vector (Novagen, Madison, WI, USA). In order to evaluate the contribution of a putative disulfide bond to the enzyme stability of the younger CYP11A1 ancestor CYP11A_Mammal_N101, site-directed mutagenesis was carried out according to standard protocols [66]. The mutagenic primer pair (5′-CATGCCGATAAATACATCCAGAAAATCTATCGTC AG-3′ and 5′-TCTGGATGTATTTATCGGCATGGTT 0AAAAATCAC-3′) contained a single codon substitution (underlined) to replace one of the involved cysteine residues at position 242 of the CYP11A_N1 sequence by a tyrosine residue. The codon exchange (C242Y) was confirmed by sequencing at Eurofins Genomics GmbH.

Protein expression and purification

Recombinant production of all CYP11A1 variants was carried out as described previously [33] with slight changes to the purification process. Following protein expression, cell pellets were suspended in 50 mL lysis buffer (pH 7.4) containing 50 mM potassium phosphate, 150 mM sodium chloride, 150 mM sodium acetate, 0.1 mM PMSF, 1% (v/v) Tween-20, 1.5% (w/v) sodium cholate, and 20% (v/v) glycerol. The bacterial cells were disrupted by sonication on ice for 20 min (15 s pulse on, 15 s pulse off, 13% amplitude). Following removal of cell debris by centrifugation at 30 000 ***g*** and 4 °C for 30 min, the supernatants of the lysates were applied to individual Protino™ Ni-NTA agarose columns (Macherey-Nagel, Düren, Germany) equilibrated with lysis buffer. The columns were washed three times with 50 mL of lysis buffer containing 10 mM, 20 mM, and 40 mM imidazole, respectively. The column bound protein was eluted with lysis buffer supplemented with 250 mM imidazole and subsequently dialyzed extensively overnight against dilution buffer (pH 6.8) containing 20 mM potassium phosphate, 0.1 mM PMSF, 1% (v/v) Tween-20, 1.5% (w/v) sodium cholate, and 20% (v/v) glycerol. The dialyzed protein solutions were applied to individual SP Sepharose™ Fast Flow columns (Cytiva Europe GmbH, Freiburg, Germany) equilibrated with dilution buffer. The columns were washed 3 times with 25 mL of dilution buffer containing 25 mM, 50 mM, and 100 mM sodium chloride, respectively. The bound protein was eluted with dilution buffer

supplemented with 300 mM sodium chloride. Finally, the obtained protein solutions were washed several times with dilution buffer using Amicon™ Ultra Centrifugal Filter Units (Merck KGaA) with MWCOs of 30 kDa, concentrated to 250 μM stock solutions and stored at −80 °C until use. Carbon monoxide (CO) difference spectroscopy was used to determine the concentration of the purified CYP11A1 enzymes based on the molar extinction coefficient of 91 $mM^{-1}cm^{-1}$ [67]. Homogeneity of the purified CYP11A1 enzymes was assessed following separation on 15% SDS polyacrylamide gels and subsequent Coomassie staining.

Recombinant bovine adrenodoxin ($Adx_{1–128}$) and bovine adrenodoxin reductase (AdR) were produced and purified as described elsewhere [68,69]. The final protein concentrations were calculated based on the molar extinction coefficients of $\varepsilon_{414} = 9.8\ mM^{-1}\cdot cm^{-1}$ for Adx [70] and $\varepsilon_{450} = 10.9\ mM^{-1}\cdot cm^{-1}$ for AdR, [71]. The Q values for Adx (A_{414}/A_{276}) and AdR (A_{450}/A_{276}) were in good agreement with those from previous reports.

CD spectroscopy

CD spectroscopy was performed on a JASCO 715 spectropolarimeter equipped with a Jasco PTC-348WI temperature controller (JASCO Deutschland GmbH, Pfungstadt, Germany). CD spectra were recorded at 20 °C in the far-UV region between 200 nm and 260 nm with data intervals of 0.5 nm and a spectral bandwidth of 1 nm. CYP11A1 samples were diluted to concentrations of 5 μM using 50 mM potassium phosphate buffer (pH 7.4) and transferred to QS-100 macrocuvettes (Hellma Analytics GmbH & Co. KG, Müllheim, Germany) with a path length of 0.1 cm. During measurements, the cuvette chamber was constantly purged with nitrogen gas (5 $L\cdot min^{-1}$). At least 10 scans were accumulated to obtain the mean values for ellipticity. For thermal unfolding studies, the CYP11A1 samples were exposed to a temperature gradient from 20 to 90 °C with an increment of 0.1 °C at a heating rate of 1 °C per min. Changes in the molar ellipticity were monitored at 210 nm. The fractions of unfolded proteins were calculated as described elsewhere [72]. The resulting melting curves were fitted to the Boltzmann function in OriginPro 2020, and individual T_m values for the CYP11A variants were derived from the calculated inflection point of the sigmoidal fits.

Ligand binding assay

Dissociation constants (K_d values) for all sterol substrates were determined on a CLARIOstar® Plus microplate reader (BMG Labtech, Ortenberg, Germany) using difference spectroscopy as described elsewhere with slight modifications [73]. Samples containing 1 μM CYP11A1 and 20 μM $Adx_{1–128}$ were prepared in 200 μL conversion buffer (50 mM HEPES, 0.05% (v/v) Tween-20, pH 7.4) and incubated with increasing concentrations of sterol substrates dissolved in 2-hydroxypropyl-β-cyclodextrin (β-CD). The final concentration of β-CD in all mixtures was 0.45% (w/v) to avoid any effect of β-CD supplementation on sterol binding. A sample supplemented with pure β-CD was used as reference. Mean values for the peak-to-through differences ($\Delta A_{390nm–420nm}$) of three individual experiments were plotted using OriginPro 2020, and K_d values were calculated from hyperbolic regression analysis.

In vitro conversion of sterols by CYP11A1 variants

CYP11A1-mediated conversion of sterols was carried out in 1.5-mL reaction tubes. The reconstituted reaction mixtures (250 μL) consisted of 0.5 μM CYP11A1, 10 μM Adx, 1.5 μM AdR, and 100 μM of the corresponding sterol substrate (prepared as 10 mM stock solutions in 45% (w/v) 2-hydroxypropyl-β-cyclodextrin in water) in 50 mM HEPES buffer (pH 7.4) supplemented with 0.05% (v/v) Tween-20. In addition, a NADPH regeneration system, comprised of 1 mM $MgCl_2$, 5 mM glucose-6-phosphate, and 1 U glucose-6-phosphate dehydrogenase per reaction, was applied. The conversion was started by adding 500 μM NADPH to the reaction tubes. Following conversion at 37 °C and 1000 r.p.m. in a ThermoMixer® (Eppendorf, Wesseling, Germany) for the indicated times, the reactions were stopped by boiling for 2 min in a water bath. To enable UV detection of the reaction product pregnenolone, the reaction mixtures were subsequently treated with 0.2 U of cholesterol oxidase. Prior to extraction, the samples were supplemented with 100 μM of the internal standard 21-hydroxyprogesterone (DOC). Thermal inactivation studies of the purified CYP11A1 variants were carried out following exposure to elevated temperatures for 10 min. In order to avoid negative effects on other enzyme constituents of the *in vitro* system, the CYP11A1 solutions were heated separately and chilled to room temperature prior to addition to the reaction mixtures. Similarly, H_2O_2 tolerance assays were carried out following preincubation of the CYP11A1 enzymes with increasing concentrations of H_2O_2 for 10 min and subsequent inactivation of residual H_2O_2 with 10 U catalase per reaction. The remaining side-chain cleavage activity for cholesterol was determined after 15 min of conversion as described above.

Steroid extraction and analysis

In vitro reactions were extracted twice with double the volume of ethyl acetate under vigorous shaking. The organic supernatants were combined and evaporated to dryness using a UNIVAPO 100 H vacuum concentrator centrifuge

UniEquip, Planegg, Germany). Residual steroid extracts were suspended in 150 μL of pure acetonitrile and subjected to RP-HPLC analysis on a JASCO 2000 series HPLC system using a Nucleodur 100-5 C18 EC column (Macherey-Nagel). The mobile phase consisted of 10% acetonitrile in water on channel A and pure acetonitrile on channel B. The following gradient was used for separation of the steroids: 0–10 min, 0% B to 100% B; 10–17.5 min, 100% B; 17.5–17.6 min, 100% B to 0% B; and 17.6–20 min, 0% B. The flow rate was changed as follows: 0–10 min, 1.0 mL·min^{-1} to 2.0 mL·min^{-1}; 10–17.5, 2.0 mL·min^{-1}; 17.5–17.6 min, 2.0 mL·min^{-1} to 1.0 mL·min^{-1}; and 17.6–20 min, 1.0 mL·min^{-1}. The column oven temperature was set to 40 °C, and the injection volume was 40 μL per sample. Detection of the internal standard, DOC, and the reaction product, progesterone, was carried out at a wavelength of 240 nm. The corresponding progesterone yields were calculated based on the peak area ratios of progesterone to DOC, which were shown to be linear ($y = 0.00846x - 0.00216$; $R^2 = 0.999$) over a broad range of concentrations.

Computational methods

Homology models for the three CYP11A1 variants were generated using the SWISS-Model web server [74,75]. The crystal structure of human CYP11A1 (PDB entry 3N9Y, chain A) was selected as template for all three P450 variants due to its higher resolution (2.10 Å) compared with that of bovine CYP11A1 (PDB entry 3MZS, 2.50 Å). The corresponding evaluation scores are given in Table S1. The heme (HEM) moiety was added manually to each model. The WHAT IF web interface was used to obtain the number of (potential) hydrogen bonds, salt bridges, and Cβ contacts, respectively [76]. Solvation energies were computed using the PDB2PQR and APBS web servers [77,78]. Hydrogen atoms were added to the homology models employing the ProToss functionality of the ProteinsPlus web service [79] prior to calculation of the force field energies using the SWISS-PDB Viewer [80].

Acknowledgments

This work was supported in part by Australian Research Council Discovery Project DP160100865 (to EMJG), by an Australian Government Research Training Program Scholarship (to SJS), and by a travel grant of the Universitätsgesellschaft des Saarlandes e.V. (to PH).

Conflict of interest

The authors declare no conflict of interest.

Author contributions

PH and SJS contributed equally to the study. SJS performed the ancestral sequence reconstruction of the CYP11A1 family with the help of EMJG. PH carried out recombinant protein production and purification, ligand binding studies, as well as circular dichroism spectroscopy. RB and EMJG conceived the project and organized its realization. PH and SJS drafted and wrote the manuscript with input from RB and EMJG. *In vitro* substrate conversion experiments were carried out by BME. AA designed, prepared, and provided expression vectors. MCH performed the computational methods and interpreted the structural data of the CYP11A1 enzymes. FH, EMJG, and RB supervised the experiments and participated in the interpretation and discussion of the data as well as the writing of the manuscript.

Peer Review

The peer review history for this article is available at https://publons.com/publon/10.1111/febs.16054.

References

1 Werck-Reichhart D & Feyereisen R (2000) Cytochromes P450: a success story. *Genome Biol* **1**, 1–9.
2 Nelson DR (2009) The Cytochrome P450 homepage. *Hum Genomics* **4**, 59–65.
3 Omura T & Sato R (1962) A new cytochrome in liver microsomes. *J Biol Chem* **237**, 1375–1376.
4 Harding BW, Wong SH & Nelson DH (1964) Carbon monoxide-combining substances in rat adrenal. *Biochim Biophys Acta BBA* **92**, 415–417.
5 Schuler MA (1996) Plant cytochrome P450 monooxygenases. *Crit Rev Plant Sci* **15**, 235–284.
6 Crešnar B & Petrič S (2011) Cytochrome P450 enzymes in the fungal kingdom. *Biochim Biophys Acta* **1814**, 29–35.
7 Feyereisen R (1999) Insect P450 enzymes. *Annu Rev Entomol* **44**, 507–533.
8 Nelson DR, Goldstone JV & Stegeman JJ (2013) The cytochrome P450 genesis locus: the origin and evolution of animal cytochrome P450s. *Philos Trans R Soc B Biol Sci* **368**, 20120474.
9 Omura T & Gotoh O (2017) Evolutionary origin of mitochondrial cytochrome P450. *J Biochem (Tokyo)* **161**, 399–407.
10 Hannemann F, Bichet A, Ewen KM & Bernhardt R (2007) Cytochrome P450 systems-biological variations of electron transport chains. *Biochim Biophys Acta BBA* **1770**, 330–344.

11 Lambeth JD, Seybert DW, Lancaster JR, Salerno JC & Kamin H (1982) Steroidogenic electron transport in adrenal cortex mitochondria. *Mol Cell Biochem* **45**, 13–31.

12 Porter TD & Kasper CB (1986) NADPH-cytochrome P-450 oxidoreductase: flavin mononucleotide and flavin adenine dinucleotide domains evolved from different flavoproteins. *Biochemistry* **25**, 1682–1687.

13 Ewen KM, Kleser M & Bernhardt R (2011) Adrenodoxin: The archetype of vertebrate-type [2Fe–2S] cluster ferredoxins. *Biochim Biophys Acta BBA* **1814**, 111–125.

14 Ewen KM, Ringle M & Bernhardt R (2012) Adrenodoxin-A versatile ferredoxin. *IUBMB Life* **64**, 506–512.

15 Mast N, Annalora AJ, Lodowski DT, Palczewski K, Stout CD & Pikuleva IA (2011) Structural basis for three-step sequential catalysis by the cholesterol side chain cleavage enzyme CYP11A1. *J Biol Chem* **286**, 5607–5613.

16 Pikuleva IA, Mackman RL, Kagawa N, Waterman MR & Demontellano PRO (1995) Active-site topology of bovine cholesterol side-chain cleavage cytochrome P450 (P450scc) and evidence for interaction of tyrosine 94 with the side chain of cholesterol. *Arch Biochem Biophys* **322**, 189–197.

17 Nonaka Y, Fujii T, Kagawa N, Waterman MR, Takemori H & Okamoto M (1998) Structure/function relationship of CYP11B1 associated with Dahl's salt-resistant rats–expression of rat CYP11B1 and CYP11B2 in Escherichia coli. *Eur J Biochem* **258**, 869–878.

18 Zöllner A, Kagawa N, Waterman MR, Nonaka Y, Takio K, Shiro Y, Hannemann F & Bernhardt R (2008) Purification and functional characterization of human 11β hydroxylase expressed in *Escherichia coli*: Functional characterization of hCYP11B1. *FEBS J* **275**, 799–810.

19 Strushkevich N, MacKenzie F, Cherkesova T, Grabovec I, Usanov S & Park HW (2011) Structural basis for pregnenolone biosynthesis by the mitochondrial monooxygenase system. *Proc Natl Acad Sci USA* **108**, 10139–10143.

20 Hobler A, Kagawa N, Hutter MC, Hartmann MF, Wudy SA, Hannemann F & Bernhardt R (2012) Human aldosterone synthase: recombinant expression in *E. coli* and purification enables a detailed biochemical analysis of the protein on the molecular level. *J Steroid Biochem Mol Biol* **132**, 57–65.

21 Schiffer L, Anderko S, Hannemann F, Eiden-Plach A & Bernhardt R (2015) The CYP11B subfamily. *J Steroid Biochem Mol Biol* **151**, 38–51.

22 Bhasker CR, Adler BS, Dee A, John ME, Kagimoto M, Zuber MX, Ahlgren R, Wang XD, Simpson ER & Waterman MR (1989) Structural characterization of the bovine CYP17 (17 alpha-hydroxylase) gene. *Arch Biochem Biophys* **271**, 479–487.

23 Imai T, Globerman H, Gertner JM, Kagawa N & Waterman MR (1993) Expression and purification of functional human 17 alpha-hydroxylase/17,20-lyase (P450c17) in *Escherichia coli*. Use of this system for study of a novel form of combined 17 alpha-hydroxylase/17,20-lyase deficiency. *J Biol Chem* **268**, 19681–19689.

24 Pechurskaya TA, Lukashevich OP, Gilep AA & Usanov SA (2008) Engineering, expression, and purification of "soluble" human cytochrome P45017alpha and its functional characterization. *Biochem Biokhimiia* **73**, 806–811.

25 Kagawa N (2011) Efficient expression of human aromatase (CYP19) in *E. coli*. *Methods Mol Biol* **705**, 109–122.

26 Arase M, Waterman MR & Kagawa N (2006) Purification and characterization of bovine steroid 21-hydroxylase (P450c21) efficiently expressed in *Escherichia coli*. *Biochem Biophys Res Commun* **344**, 400–405.

27 Gumulya Y, Baek J-M, Wun S-J, Thomson RES, Harris KL, Hunter DJB, Behrendorff JBYH, Kulig J, Zheng S, Wu X *et al.* (2018) Engineering highly functional thermostable proteins using ancestral sequence reconstruction. *Nat Catal* **1**, 878–888.

28 Gumulya Y, Huang W, D'Cunha SA, Richards KE, Thomson RES, Hunter DJB, Baek J-M, Harris KL, Boden M, De Voss JJ *et al.* (2018) Engineering thermostable CYP2D enzymes for biocatalysis using combinatorial libraries of ancestors for directed evolution (CLADE). *ChemCatChem* **12**, 1750–1761.

29 Bart AG, Harris KL, Gillam EMJ & Scott EE (2020) Structure of an ancestral mammalian family 1B1 cytochrome P450 with increased thermostability. *J Biol Chem* **295**, 5640–5653.

30 Gumulya Y & Gillam EMJ (2017) Exploring the past and the future of protein evolution with ancestral sequence reconstruction: the 'retro' approach to protein engineering. *Biochem J* **474**, 1–19.

31 Ohno S (1970) Evolution by gene duplication. Springer-Verlag, Berlin Heidelberg.

32 Nelson DR, Koymans L, Kamataki T, Stegeman JJ, Feyereisen R, Waxman DJ, Waterman MR, Gotoh O, Coon MJ, Estabrook RW *et al.* (1996) P450 superfamily: update on new sequences, gene mapping, accession numbers and nomenclature. *Pharmacogenetics* **6**, 1–42.

33 Mosa A, Neunzig J, Gerber A, Zapp J, Hannemann F, Pilak P & Bernhardt R (2015) 2β- and 16β-hydroxylase activity of CYP11A1 and direct stimulatory effect of estrogens on pregnenolone formation. *J Steroid Biochem Mol Biol* **150**, 1–10.

34 Tegel H, Ottosson J & Hober S (2011) Enhancing the protein production levels in *Escherichia coli* with a strong promoter: Protein production levels in *E. coli*. *FEBS J* **278**, 729–739.
35 Luthra A, Denisov IG & Sligar SG (2011) Spectroscopic features of cytochrome P450 reaction intermediates. *Arch Biochem Biophys* **507**, 26–35.
36 Lambeth JD, Seybert DW & Kamin H (1980) Phospholipid vesicle-reconstituted cytochrome P-450SCC. Mutually facilitated binding of cholesterol and adrenodoxin. *J Biol Chem* **255**, 138–143.
37 Yablokov EO, Sushko TA, Kaluzhskiy LA, Kavaleuski AA, Mezentsev YV, Ershov PV, Gilep AA, Ivanov AS & Strushkevich NV (2020) Substrate-induced modulation of protein-protein interactions within human mitochondrial cytochrome P450-dependent system. *J Steroid Biochem Mol Biol* **208**, 105793.
38 Usanov SA, Graham SE, Lepesheva GI, Azeva TN, Strushkevich NV, Gilep AA, Estabrook RW & Peterson JA (2002) Probing the interaction of bovine cytochrome P450scc (CYP11A1) with adrenodoxin: Evaluating site-directed mutations by molecular modeling. *Biochemistry* **41**, 8310–8320.
39 Grinberg AV, Hannemann F, Schiffler B, Müller J, Heinemann U & Bernhardt R (2000) Adrenodoxin: structure, stability, and electron transfer properties. *Proteins* **40**, 590–612.
40 Arthur JR, Blair HA, Boyd GS, Mason JI & Suckling KE (1976) Oxidation of cholesterol and cholesterol analogues by mitochondrial preparations of steroid-hormone-producing tissue. *Biochem J* **158**, 47–51.
41 Mason JI, Arthur JR & Boyd GS (1978) Control of sterol metabolism in rat adrenal mitochondria. *Biochem J* **174**, 1045–1051.
42 Morisaki M, Duque C, Ikekawa N & Shikita M (1980) Substrate specificity of adrenocortical cytochrome P-450scc–I. Effect of structural modification of cholesterol side-chain on pregnenolone production. *J Steroid Biochem* **13**, 545–550.
43 Tuckey RC & Cameron KJ (1993) Side-chain specificities of human and bovine cytochromes P-450scc. *Eur J Biochem* **217**, 209–215.
44 Simgen B, Contzen J, Schwarzer R, Bernhardt R & Jung C (2000) Substrate binding to 15β-Hydroxylase (CYP106A2) probed by FT infrared spectroscopic studies of the iron ligand CO stretch vibration. *Biochem Biophys Res Commun* **269**, 737–742.
45 Conner KP, Woods CM & Atkins WM (2011) Interactions of cytochrome P450s with their ligands. *Arch Biochem Biophys* **507**, 56–65.
46 Ewen KM, Hannemann F, Iametti S, Morleo A & Bernhardt R (2011) Functional characterization of Fdx1: evidence for an evolutionary relationship between P450-type and ISC-type ferredoxins. *J Mol Biol* **413**, 940–951.
47 Hanukoglu I, Rapoport R, Weiner L & Sklan D (1993) Electron leakage from the mitochondrial NADPH-adrenodoxin reductase-adrenodoxin-P450scc (Cholesterol Side Chain Cleavage) system. *Arch Biochem Biophys* **305**, 489–498.
48 Schallmey A, den Besten G, Teune IGP, Kembaren RF & Janssen DB (2011) Characterization of cytochrome P450 monooxygenase CYP154H1 from the thermophilic soil bacterium *Thermobifida fusca*. *Appl Microbiol Biotechnol* **89**, 1475–1485.
49 Ho WW, Li H, Nishida CR, de Montellano PRO & Poulos TL (2008) Crystal structure and properties of CYP231A2 from the thermoacidophilic archaeon *Picrophilus torridus*. *Biochemistry* **47**, 2071–2079.
50 Nguyen K-T, Nguyen N-L, Milhim M, Nguyen V-T, Lai T-H-N, Nguyen H-H, Le T-T-X, Phan T-T-M & Bernhardt R (2020) Characterization of a thermophilic cytochrome P450 of the CYP203A subfamily from Binh Chau hot spring in Vietnam. *FEBS Open Bio* **11**, 124–132.
51 Williams PD, Pollock DD, Blackburne BP & Goldstein RA (2006) Assessing the accuracy of ancestral protein reconstruction methods. *PLoS Comput Biol* **2**, e69.
52 Jaenicke R (1999) Stability and folding of domain proteins. *Prog Biophys Mol Biol* **71**, 155–241.
53 Vogt G, Woell S & Argos P (1997) Protein thermal stability, hydrogen bonds, and ion pairs. *J Mol Biol* **269**, 631–643.
54 Harris KL, Thomson RES, Strohmaier SJ, Gumulya Y & Gillam EMJ (2018) Determinants of thermostability in the cytochrome P450 fold. *Biochim Biophys Acta Proteins Proteomics* **1866**, 97–115.
55 Sun M, Wang Y, Zhang Q, Xia Y, Ge W & Guo D (2017) Prediction of reversible disulfide based on features from local structural signatures. *BMC Genom* **18**, 279.
56 Altschul SF, Gish W, Miller W, Myers EW & Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* **215**, 403–410.
57 Li W & Godzik A (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**, 1658–1659.
58 Katoh K, Rozewicki J & Yamada KD (2019) MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform* **20**, 1160–1166.
59 Waterhouse AM, Procter JB, Martin DMA, Clamp M & Barton GJ (2009) Jalview Version 2–a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191.
60 Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313.
61 Foley G, Mora A, Ross CM, Bottoms S, Sützl L, Lamprecht ML, Zaugg J, Essebier A, Balderson B,

Newell R *et al.* (2019) Identifying and engineering ancient variants of enzymes using Graphical Representation of Ancestral Sequence Predictions (GRASP). *Bioinformatics* **12**, 30.
62 Wada A & Waterman MR (1992) Identification by site-directed mutagenesis of two lysine residues in in cholesterol side chain cleavage cytochrome P450 that are essential for adrenodoxin binding. *J Biol Chem* **267**, 22877–22882.
63 Mandel M & Higa A (1970) Calcium-dependent bacteriophage DNA infection. *J Mol Biol* **53**, 159–162.
64 Sambrook J & Russell DW (2001) Molecular cloning: a laboratory manual, 3rd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
65 Bertin M & Dumas B (2017) Cytochrome P450 polypeptide with increased enzyme activity. No. US9765307B2.
66 Laible M & Boonrod K (2009) Homemade site directed mutagenesis of whole plasmids. *J Vis Exp* **11**, 1135.
67 Omura T & Sato R (1964) The carbon monoxide-binding pigment of liver microsomes. I. Evidence for its hemoprotein nature. *J Biol Chem* **239**, 2370–2378.
68 Uhlmann H, Beckert V, Schwarz D & Bernhardt R (1992) Expression of bovine adrenodoxin in *E. coli* and site-directed mutagenesis of /2 FE-2S/ cluster ligands. *Biochem Biophys Res Commun* **188**, 1131–1138.
69 Sagara Y, Wada A, Takata Y, Waterman MR, Sekimizu K & Horiuchi T (1993) Direct expression of adrenodoxin reductase in *Escherichia coli* and the functional characterization. *Biol Pharm Bull* **16**, 627–630.
70 Huang JJ & Kimura T (1973) Adrenal steroid hydroxylases. Oxidation- reduction properties of adrenal iron-sulfur protein (adrenodoxin). *Biochemistry* **12**, 406–409.
71 Chu JW & Kimura T (1973) Studies on adrenal steroid hydroxylases. Molecular and catalytic properties of adrenodoxin reductase (a flavoprotein). *J Biol Chem* **248**, 2089–2094.
72 Liu Z, Lemmonds S, Huang J, Tyagi M, Hong L & Jain N (2018) Entropic contribution to enhanced thermal stability in the thermostable P450 CYP119. *Proc Natl Acad Sci USA* **115**, E10049–E10058.
73 Jefcoate CR (1978) Measurement of substrate and inhibitor binding to microsomal cytochrome P-450 by optical-difference spectroscopy. *Methods Enzymol* **52**, 258–279.
74 Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L *et al.* (2018) SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* **46**, W296–W303.
75 Benkert P, Biasini M & Schwede T (2011) Toward the estimation of the absolute quality of individual protein structure models. *Bioinforma Oxf Engl* **27**, 343–350.
76 Vriend G (1990) WHAT IF: a molecular modeling and drug design program. *J Mol Graph* **8**, 52–56.
77 Baker NA, Sept D, Joseph S, Holst MJ & McCammon JA (2001) Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci* **98**, 10037–10041.
78 Dolinsky TJ, Nielsen JE, McCammon JA & Baker NA (2004) PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res* **32**, W665–667.
79 Lippert T & Rarey M (2009) Fast automated placement of polar hydrogen atoms in protein-ligand complexes. *J Cheminformatics* **1**, 13.
80 Guex N, Peitsch MC & Schwede T (2009) Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective. *Electrophoresis* **30**, S162–173.
81 Katoh K & Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* **30**, 772–780.
82 Robert X & Gouet P (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* **42**, W320–W324.
83 Strushkevich N, MacKenzie F, Cherkesova T, Grabovec I, Usanov S & Park H-W (2011) Structural basis for pregnenolone biosynthesis by the mitochondrial monooxygenase system. *Proc Natl Acad Sci USA* **108**, 10139–10143.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. Differences in amino acid composition of the three CYP11A1 variants.

Table S2. Number of hydrogen-bonds, salt bridges, and Cβ contacts of the three CYP11A1 variants.

Table S3. Scores of the generated homology models of the three CYP11A1 variants as reported by the SWISS-Model server.

Table S4. Force field energies and solvation energies of the three CYP11A1 variants.

Table S5. Bacterial strains and vectors used in this study.