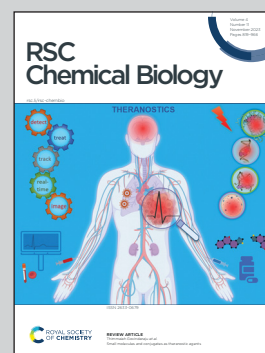**Showcasing research from Professor Distefano's laboratory, Department of Chemistry, University of Minnesota, Minneapolis, MN, USA.**

Thinking outside the CaaX-box: an unusual reversible prenylation on ALDH9A1

ALDH9A1 catalyzes the oxidation of certain biogenic aldehydes producing key metabolites including carnitine and GABA, which in turn play central roles in cellular energy production and neurotransmission. Here we report the discovery of ALDH9A1 prenoylation, a new post-translational modification involving thioesterification of a cysteine residue with a farnesoyl or geranylgeranoyl group. Since the site of ALDH9A1 acylation is an active site nucleophile, covalent modification of that residue results in enzyme inhibition. By modulating global ALDH9A1 activity through changes in its acylation state, a range of cellular processes could be regulated.

ROYAL SOCIETY OF CHEMISTRY

rsc.li/rsc-chembio

Check for updates

Thinking outside the CaaX-box: an unusual reversible prenylation on ALDH9A1†

Kiall F. Suazo, [ID] [a] Jakub Bělíček, [ID] [b] Garrett L. Schey, [ID] [c] Shelby A. Auger, [ID] [a] Alexandru M. Petre, [ID] [a] Ling Li, [ID] [d] Katarzyna M. Błażewska, [ID] [e] David Kopečný [ID] [b] and Mark D. Distefano [ID] *[a]

Protein lipidation is a post-translational modification that confers hydrophobicity on protein substrates to control their cellular localization, mediate protein trafficking, and regulate protein function. In particular, protein prenylation is a C-terminal modification on proteins bearing canonical motifs catalyzed by prenyltransferases. Prenylated proteins have been of interest due to their numerous associations with various diseases. Chemical proteomic approaches have been pursued over the last decade to define prenylated proteomes (prenylome) and probe their responses to perturbations in various cellular systems. Here, we describe the discovery of prenylation of a non-canonical prenylated protein, ALDH9A1, which lacks any apparent prenylation motif. This enzyme was initially identified through chemical proteomic profiling of prenylomes in various cell lines. Metabolic labeling with an isoprenoid probe using overexpressed ALDH9A1 revealed that this enzyme can be prenylated inside cells but does not respond to inhibition by prenyltransferase inhibitors. Site-directed mutagenesis of the key residues involved in ALDH9A1 activity indicates that the catalytic C288 bears the isoprenoid modification likely through an NAD⁺-dependent mechanism. Furthermore, the isoprenoid modification is also susceptible to hydrolysis, indicating a reversible modification. We hypothesize that this modification originates from endogenous farnesal or geranygeranial, the established degradation products of prenylated proteins and results in a thioester form that accumulates. This novel reversible prenoyl modification on ALDH9A1 expands the current paradigm of protein prenylation by illustrating a potentially new type of protein–lipid modification that may also serve as a novel mechanism for controlling enzyme function.

Introduction

Protein prenylation is an essential post-translational modification (PTM) required for certain proteins to properly localize in cellular membranes and for mediating protein–protein interactions and protein trafficking.[1] In this type of PTM, a farnesyl or geranylgeranyl group(s) is irreversibly appended through a thioether linkage on a cysteine residue near the C-terminus of a protein. There are currently four characterized prenyltransferases that catalyze such modifications. Farnesyltransferase (FTase) attaches a 15-carbon farnesyl moiety from farnesyl diphosphate while geranylgeranyltransferase type 1 (GGTase-I) links a 20-carbon geranylgeranyl group from geranylgeranyl diphosphate onto their substrates.[2] These substrate proteins generally bear a canonical CaaX-box prenylation sequence where C is the prenylated cysteine and a and X can be various amino acids. However, recent studies have shown that proteins terminating in CXXXX or CXX may also potentially be substrates of prenylation.[3–5] The third type of prenylation involves geranylgeranyltransferase type II (GGTase-II or RabGGTase) that usually appends two geranylgeranyl units on dual cysteine motifs on proteins, particularly those that belong to the Rab family. These substrates are typically recognized first by the Rab escort protein 1 or 2 (REP-1/2) and modified by GGTase-II at their C-terminal cysteines in their –CCXX, –CXC, or –XXCC canonical sequences.[6] Finally, the recently discovered GGTase-III appears to install a second isoprenoid geranylgeranyl moiety, onto a pre-farnesylated protein substrate.[7,8]

[a] Department of Chemistry, University of Minnesota, Minneapolis, MN 55455, USA. E-mail: suazo003@umn.edu, auger054@umn.edu, petre096@umn.edu, diste001@umn.edu

[b] Department of Experimental Biology, Faculty of Science, Palacký University, CZ-78371, Czech Republic. E-mail: jakub.belicek01@upol.cz, david.kopecny@upol.cz

[c] Department of Medicinal Chemistry, University of Minnesota, Minneapolis, MN 55455, USA. E-mail: schey013@umn.edu

[d] Department of Experimental and Clinical Pharmacology, University of Minnesota, Minneapolis, MN, 55455, USA. E-mail: lil@umn.edu

[e] Institute of Organic Chemistry, Faculty of Chemistry, Lodz University of Technology, Łódź, Poland. E-mail: katarzyna.blazewska@p.lodz.pl

Identifying prenylated proteins and the lipid-modified proteome, in general, has been an important research goal for more than a decade.[9,10] Recent methods for high-throughput analysis of these lipid-modified proteins take advantage of chemical proteomics that allows for the selective identification of protein substrates *via* an enrichment strategy.[11–13] In this approach, a bio-orthogonal probe (usually containing an alkyne moiety) that mimics the native form of a small molecule substrate used for enzymatic protein modification is added to cells and is incorporated by the host machinery into the target proteins. Alternatively, labeling of proteins can be achieved *in vitro* through the addition of the lipid probe and enzymes that catalyze the PTM.[14] These labeled proteins present in lysates are then subjected to a click reaction with a biotin-containing reagent for selective enrichment and subsequent proteomic identification of the modified proteins. Chemical proteomic approaches to identify prenylated proteins have been developed for more than a decade, wherein a bio-orthogonal isoprenoid analogue is synthesized and used to tag and identify the prenylated proteins in a given cell line.[10] The goal of defining the prenylome has been a challenge, reflected by the lesser number of proteins identified in pre-nylomic studies compared to the total number of prenylation substrates derived from predictions and annotations. Recently, two independent studies reported 80 prenylated proteins identified in a given cell line, which consists of known and novel prenylation substrates.[11,12] One study employed a dual labeling strategy where a farnesyl and a geranylgeranyl probe analogue were used, while the other used C15AlkOPP as a single probe that labels all classes of prenylation substrates. Despite the differences in these approaches, a comparable number of proteins were identified with slightly varying identities, particularly in the novel prenylated proteins discovered. In many cell lines that have been studied, one peculiar protein, ALDH9A1, has been consistently enriched. Here, we describe our discovery of ALDH9A1 as a potentially prenylated protein that has possible biological consequences in controlling this enzyme's function. Through the use of tools of chemical and molecular biology, we characterized this potentially novel lipid modification and propose a mechanism *via* which this may occur in nature.

Results and discussion

Chemical proteomic analysis of prenylated proteins identified ALDH9A1

Recent advances in proteomic technologies have enabled high throughput profiling of post-translationally modified (PTM) proteins of interest. In particular, chemical proteomic approaches involve an enrichment strategy to identify a set of modified proteins such as the prenylated proteome. Along with other groups, we have conducted prenylomic profiling across multiple cell lines and showed that the scope of prenylated proteins varies to some extent in these various cell types (Fig. 1).[12] COS-7 (Fig. 1(A)) displayed the largest number of

proteins profiled including novel prenylated proteins bearing the canonical CaaX-box motifs that were not previously identified in other prenylomic studies reported. We have found that the differential labeling of the prenylomes in these various cell lines may be attributed to the levels of expression of their prenyltransferases and cognate substrates or their ability for probe uptake.[12]

In our prenylomic studies, ALDH9A1, an aldehyde dehydrogenase, has been consistently enriched in a number of cell lines studied including COS-7, HeLa, MDA-MB-231, and OPM2 (Fig. 1(A)), as well as in studies from other groups where it was enriched in RAW264.7 cells.[12,15] This enzyme does not bear a canonical prenylation motif but rather contains two adjacent cysteines (Cys288/Cys289) within its protein sequence, with Cys288 ascribed as a catalytic residue (Fig. 1(B)).[16,17] We initially suspected that this protein may undergo proteolytic processing to reveal a cryptic C-terminal Cys–Cys motif (Cys288 and Cys289) that could serve as a potential substrate for dual prenylation by RabGGTase. However, examining the peptides identified in the proteomic analysis revealed that residues downstream of Cys288/Cys289 are enriched (Fig. 1(B)), suggesting that the intact full-length protein was isolated in the enrichment procedure. Therefore, labeling of ALDH9A1 with C15AlkOPP appears to deviate from the well-established rules for substrate recognition manifested by prenyltransferases.

C15AlkOPP labels ALDH9A1 but is unresponsive to prenyltransferase inhibitors

To determine whether the C15AlkOPP probe indeed labels ALDH9A1 in the metabolic labeling experiments, the enzyme was overexpressed as a GFP fusion in COS-7 cells followed by subsequent metabolic labeling. A plasmid expressing ALDH9A1 with a C-terminal GFP tag (ALDH9A1-GFP) was initially transfected in COS-7 cells followed by incubation with C15AlkOPP post-transfection. By placing the GFP tag in a C-terminal position downstream of ALDH9A1, we hoped it would have been possible to evaluate whether endoproteolytic cleavage to expose a C-terminal prenylation motif was occurring. However, this approach did not result in successful expression of the protein (data not shown), perhaps owing to the fact that the C-terminal region of ALDH9A1 is important for homotetramerization that is essential for its stability.[16] Instead, an N-terminal-tagged version (GFP-ALDH9A1) was used followed by metabolic labeling with C15AlkOPP and the lysates were subjected to click reaction with TAMRA-N$_3$. Importantly, that construct displayed an intense fluorescent band near 75 kDa (Fig. 2(A), lane 4), indicating successful labeling of the full-length intact protein. Since C15AlkOPP has the ability to label all three classes of prenylation substrates, it is not clear whether the ALDH9A1 is potentially farnesylated or geranylgeranylated.

To further establish the identity of the prenyl group modifying ALDH9A1, the C15AlkOPP probe was competed with the native isoprenoids. Co-treatment with 1.5-fold excess of FPP or GGPP with C15AlkOPP substantially but not completely abolished the labeling on ALDH9A1 (Fig. 2(A), lanes 5 and 7). Under these conditions, virtually complete inhibition of the labeling
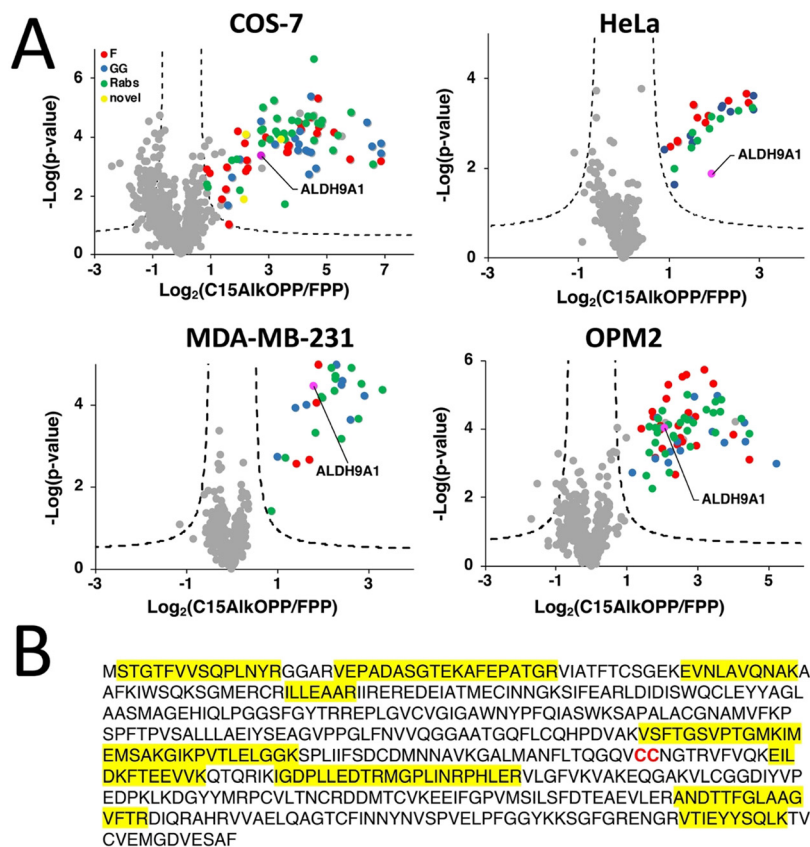
914 | *RSC Chem. Biol.*, 2023, **4**, 913–925

© 2023 The Author(s). Published by the Royal Society of Chemistry

Fig. 1 ALDH9A1 is consistently enriched in prenylomic analyses in various cell lines. (A) Volcano plots (FDR = 0.01, s0 = 0.5) of prenylomic analysis on various cell lines show that ALDH9A1 is consistently enriched. (B) The complete protein sequence of the human ALDH9A1. Highlighted in yellow are the tryptic peptides identified in the prenylomic analyses.

of known prenylated proteins is readily achieved as indicated by the abolished signal in the 25 kDa region. Treatment with excess FPP inhibits both farnesylation and geranylgeranylation since FPP is elongated to GGPP by the GGPP synthase.[18] As both FPP and GGPP competition reduce its labeling, ALDH9A1 may indeed be labeled with either isoprenoid. Next, the prenyltransferases were inhibited *in cellulo* by treating the transfected cells with known prenyltransferase inhibitors (Fig. 2(B)). Treatment with the FTase inhibitor L-744,832 (Fig. 2(A), FTI, lane 6) or GGTase-I inhibitor GGTI-286 (GGTI, lane 8) did not significantly impact the labeling of ALDH9A1 under concentrations previously reported to induce observable inhibition of intracellular prenylation.[19,20] In contrast, the reduction in the labeling of many of the normally prenylated proteins in the presence of FTI and GGTI is readily seen in lanes 6 and 8 in the 25 and 50 kDa region of the gel.

The absence of a CaaX-box motif in the full length ALDH9A1 suggests that it should not be a substrate of FTase and GGTase-I. However, this enzyme does contain two adjacent cysteines Cys288/Cys289 with Cys288 shown to be the catalytic residue for its aldehyde dehydrogenase function. As noted above, dual prenylation on adjacent cysteine residues is common among the Rab family of proteins that are substrates of RabGGTase, albeit only when they are near the C-terminus. Interestingly, a dual cysteine motif of Rab5b in *Plasmodium falciparum* was

suggested to be prenylated although these putative prenylatable cysteines are 65 amino acids upstream from the C-terminus of that protein.[21] Therefore the ability of known RabGGTase inhibitors (RRGTI-1-3) in diminishing the C15AlkOPP-labeling of GFP-ALDH9A1 was evaluated. RGGTI-1 and RGGTI-2 were previously shown to be effective in inhibiting Rab geranylgeranylation in HeLa cells (IC$_{50}$ = 358 µM and 546 µM, respectively)[22,23] while RGGTI-3 was less effective with IC$_{50}$ = 850 µM.[24] Treatment with these RGGT inhibitors in GFP-ALDH9A1-transfected COS-7 cells did not reduce the observed C15AlkOPP-labeling of GFP- ALDH9A1 (Fig. 2(C)). However, using this RGGTI concentration (250 µM), the prenylome labeling in the 25 kDa region was significantly diminished where most Rab prenylation substrates migrate (lanes 3 4, 5 and 6). If indeed RabGGTase prenylates ALDH9A1, a diminished labeling should have been observed in the presence of these RGGTIs. Therefore, ALDH9A1 does not appear to be a substrate of RabGGTase or any of the known prenyltransferases.

Key residues in ALDH9A1 function are involved in isoprenoid labelling

The aldehyde dehydrogenase (ALDH) superfamily is a family of NAD$^+$-dependent enzymes that catalyze the conversion of aldehydes to carboxylic acids with varying chain lengths and

© 2023 The Author(s). Published by the Royal Society of Chemistry

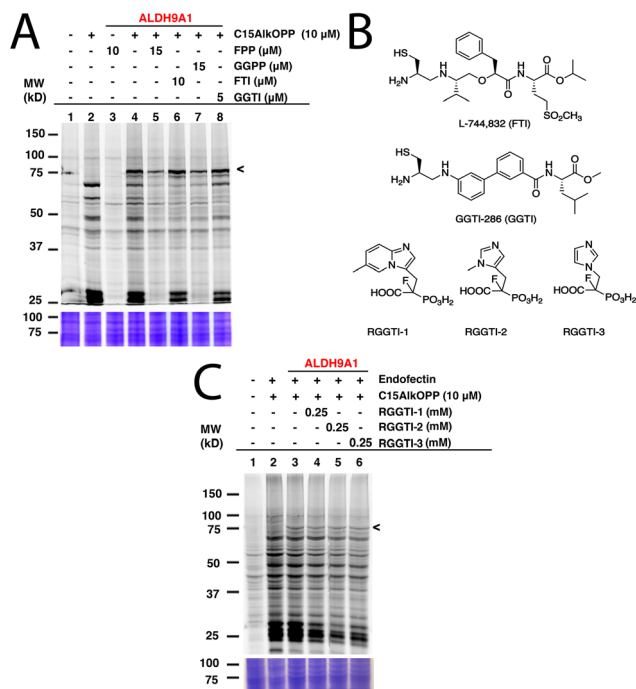*RSC Chem. Biol.*, 2023, **4**, 913–925 | **915**

Fig. 2 The C15AlkOPP probe labels ALDH9A1 in cells and cannot be inhibited with prenyltransferase inhibitors. (A) Labeling with C15AlkOPP, inhibition experiments with prenyltransferase inhibitors (FTI and GGTI), and competition with the native isoprenoids (FPP and GGPP). The C15AlkOPP labelling of ALD9A1 appears to be unaffected by the inhibitors although somewhat diminished by the competitors. That contrasts with the substantial inhibition of other prenylated proteins. (B) Structures of the prenyltransferase inhibitors used to potentially inhibit the labeling of ALDH9A1. (C) Inhibition studies of ALDH9A1 labeling by C15AlkOPP using RabGGTase inhibitors (RGGTIs). No apparent decrease in labeling of GFP-ALDH9A1 was observed although there was a significant decrease of labeling in the 25 kDa region where most Rab proteins migrate.
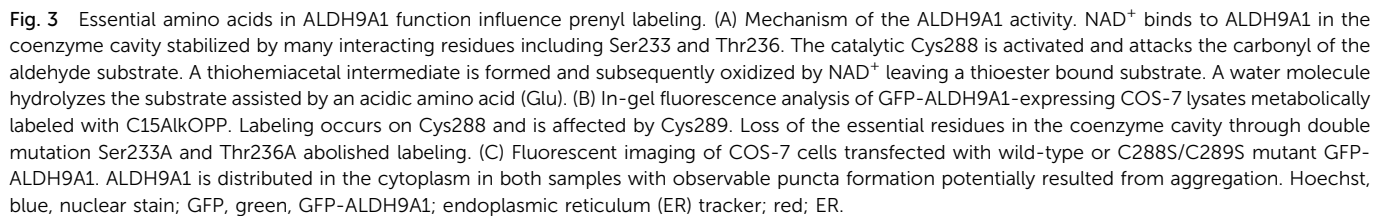
structures.[25–27] ALDH9A1 in higher eukaryotes including humans is a homotetrameric enzyme that mediates the NAD$^+$-dependent oxidation of several aldehydes including betaine aldehyde, the carnitine precursor 4-trimethylaminobuteraldehyde (TMABAL), and the GABA precursor aminobutyraldehyde (ABAL).[28–30] The general mechanism of catalysis for ALDH9A1 and for most of these ALDHs initially involves NAD$^+$ cofactor binding that induces a conformational change resulting in activation of the thiol of a catalytic cysteine as a thiolate (Fig. 3(A)).[26] In ALDH9A1, NAD$^+$ binding in the coenzyme cavity is stabilized by many electrostatic interactions, which include interactions between the pyrophosphate moiety of NAD$^+$ and residues Trp156, Ser233, and Thr236.[17] The activated thiol then attacks the carbonyl of the aldehyde substrate, generating a thiohemiacetal intermediate stabilized by NH groups from the ALDH9A1 peptide main chain. A hydride transfer to NAD$^+$ then follows, resulting in reduction of NAD$^+$ to NADH and concomitant oxidation of the substrate to a thioester. This covalent intermediate then undergoes hydrolysis with the aid of a charged amino acid (usually Glu) in the active site to release the carboxylic acid product. The residue C288 is known to be the catalytic cysteine in ALDH9A1 while an adjacent Cys289 has no currently known function in this enzyme. A related yeast enzyme

ALDH4 also contains this dual cysteine motif Cys324/Cys325 with Cys324 being the catalytic residue. A recent study has shown that Cys325 regulates the yeast enzyme's function by forming a disulfide bond with Cys324, in order to promote an oxidative stress response.[31] Similarly, the human ALDH1A1 exhibits the same disulfide bond formation between its active nucleophile Cys303 and the adjacent Cys302.[31]

Based on the mechanism described above, we explored the role of several of the residues essential for ALDH9A1-mediated catalysis on labeling by C15AlkOPP. Therefore, site-directed mutagenesis was performed on the aforementioned protein construct to introduce mutations in some of these key residues and the resulting mutants transfected into COS-7 cells, followed by metabolic labeling with C15AlkOPP and click chemistry. Double mutation of the Cys288,Cys289 pair to serine residues (Cys288S,Cys289S, abbreviated SS) completely abolished labeling (Fig. 3(B)). Mutating the cysteine adjacent to the catalytic residue (C289S, CS) diminished the probe incorporation, whereas mutating the catalytic cysteine only (C288S, SC) completely abrogated the signal. This suggests that probe labeling occurs on the catalytic Cys288 with Cys289 contributing in an indirect manner. As a control experiment, metabolic labeling was also performed in cells transfected with GFP-ALDH1A1 bearing this dual cysteine motif (Cys302/Cys303) but no labeling was observed (data not shown). We also examined the potential participation of NAD$^+$ in the course of probe incorporation. Mutating residues Ser233 and Thr236, key amino acids essential for NAD$^+$ binding, into alanine also completely abolished the observed labeling. Thus, these latter experiments suggest that NAD$^+$ may be involved in this process of isoprenoid labeling of the catalytic Cys288 of ALDH9A1.

Protein prenylation is often associated with membrane targeting of the modified substrates, with the prenyl moiety serving as an anchor.[2] Among the many protein prenylation substrates, mouse ALDH3B2 and ALDH3B3 are examples of aldehyde dehydrogenase enzymes known to be prenylated, both possessing the C-terminal sequence CTLL.[32] Although both are prenylated, they differ in their cellular localization—ALDH3B3 localizes in the plasma membrane while ALDH3B2 resides in lipid droplets—influenced by some key residues in their corresponding C-termini.[32] We therefore assessed the localization of the transfected GFP-ALDH9A1 fusion protein employed above and compared it with the corresponding double C288S/C289S mutant using fluorescence microscopy (Fig. 3(C)). There were no observable differences in the localization of the wild-type *versus* mutant. The protein appeared to be distributed across the cytoplasm, although discrete puncta were present in both samples. This protein species also did not localize in the endoplasmic reticulum (ER) where clusters of proteins reside during protein synthesis. Therefore, these puncta may be protein aggregates, which is a common phenomenon in protein overexpression in cells, particularly with proteins bearing a considerable number of cysteine residues (16 Cys in ALDH9A1).[33] It is also possible that these puncta are lipid droplets, although the potential isoprenoid modification on

Fig. 3 Essential amino acids in ALDH9A1 function influence prenyl labeling. (A) Mechanism of the ALDH9A1 activity. NAD$^+$ binds to ALDH9A1 in the coenzyme cavity stabilized by many interacting residues including Ser233 and Thr236. The catalytic Cys288 is activated and attacks the carbonyl of the aldehyde substrate. A thiohemiacetal intermediate is formed and subsequently oxidized by NAD$^+$ leaving a thioester bound substrate. A water molecule hydrolyzes the substrate assisted by an acidic amino acid (Glu). (B) In-gel fluorescence analysis of GFP-ALDH9A1-expressing COS-7 lysates metabolically labeled with C15AlkOPP. Labeling occurs on Cys288 and is affected by Cys289. Loss of the essential residues in the coenzyme cavity through double mutation Ser233A and Thr236A abolished labeling. (C) Fluorescent imaging of COS-7 cells transfected with wild-type or C288S/C289S mutant GFP-ALDH9A1. ALDH9A1 is distributed in the cytoplasm in both samples with observable puncta formation potentially resulted from aggregation. Hoechst, blue, nuclear stain; GFP, green, GFP-ALDH9A1; endoplasmic reticulum (ER) tracker; red; ER.

ALDH9A1 may not necessarily influence this mechanism of localization.

The isoprenoid modification is hydrolyzable

To gain insight into the origin of the isoprenoid incorporated into ALDH9A1, metabolic labeling experiments in COS-7 cells transfected with ALDH9A1 were performed with different isoprenoid analogues. The C15AlkOH probe (the alcohol form lacking the diphosphate group) produced appreciable labeling of ALDH9A1 although it was less compared to that observed with C15AlkOPP (Fig. 4(A) lane 3). It should be noted that isoprenyl alcohols readily become phosphorylated and incorporated into prenylated proteins.[34] A previous report of using this analogue in profiling prenylated proteins identified ALDH9A1 as one of the highly enriched prenylated proteins.[15] Thus, the isoprenoid modification on ALDH9A1 is probably derived from the diphosphate analogue produced from the phosphorylation of C15AlkOH by host cell kinases to C15AlkOPP.[34] We also evaluated the ability of the alkyne-modified analogue of

farnesal (C15AlkCHO) and the methyl ester C15AlkCOOMe in labeling ALDH9A1 since the aldehyde and acid are both known isoprenoid metabolites in cells; of particular note, farnesal (FCHO) and geranylgeranial (GGCHO) are known products of the deprenylating enzyme prenylcysteine oxidase (Pcyox 1);[35] presumably, they are also the products of the recently discovered homologue, Pcyox1L.[36] We rationalized that the neutral species C15AlkCOOMe could efficiently penetrate the plasma membrane and be hydrolyzed by cellular esterases to C15AlkCOOH, an analogue of FCOOH or GGCOOH. Neither of these analogues efficiently modified ALDH9A1 to the extent observed with C15AlkOPP (Fig. 4(A) lanes 4 and 5) although some labeling was observed with C15AlkCHO. This reduced labeling was only present in wild-type ALDH9A1 and was abolished when mutations were introduced into the key residues (Fig. S1, ESI†). It is possible that these analogues were bound to or reacted with serum proteins in the media and were not able to enter the cells since they both contain α,β-unsaturated carbonyl moieties that are known to readily react with various nucleophiles.
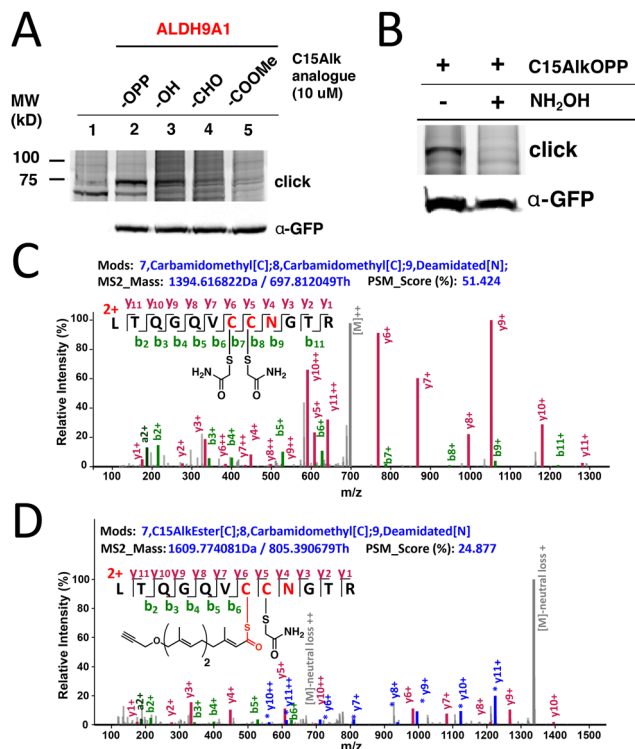
Fig. 4 Hydrolyzable modification on ALDH9A1 (A) In-gel fluorescence analysis on lysates from COS-7 cells transfected with GFP-ALDH9A1 and treated with various alkyne-modified isoprenoid analogues. (B) Hydrolysis of the observed labeling on C15AlkOPP-labeled GFP-ALDH9A1 using NH_2OH. (C) Mass spectrum of reduced/alkylated active site peptide of ALDH9A1. (D) Mass spectrum of the C15Alk-thioester-modified Cys288 of the active site peptide of ALDH9A1. Shown in blue and gray correspond to neutral loss of Alk-farnesoyl ($C_{18}H_{25}O_2$). The colors of peaks are indicated for the parent ions (gray), y ions (red), y ions from parent peptide with the neutral loss (blue with *), and b ions (green).

While the aforementioned data with prenyltransferase inhibitors suggest that ALDH9A1 prenylation is not a process mediated directly by prenyltransferases, it does not rule out the conversion of C15AlkOPP into a variety of metabolites that could potentially result in subsequent ALDH9A1 modification. Accordingly, we investigated whether this observed modification could be hydrolyzed by hydroxylamine. Basic hydrolysis with potent α-nucleophiles is commonly employed to distinguish reversible modifications such as *S*-palmitoylation on proteins since the thioester bond on palmitoylated cysteines is labile to hydroxylamine.[13] Treatment of C15AlkOPP-labeled GFP-ALDH9A1-expressing COS-7 cells with hydroxylamine completely abolished the isoprenoid labeling (Fig. 4(B)). Thus, C15AlkOPP-mediated labeling appears to involve a hydrolyzable modification on ALDH9A1 instead of a thioether linkage that is resistant to hydrolysis that typically occurs in prenyltransferase-catalyzed protein modification.

The observed hydrolyzable modification on ALDH9A1 suggests that the isoprenoid modification may involve a thioesterified cysteine residue. Hence, we attempted to obtain a mass spectrum of the peptide modified with the alkyne-functionalized probe. COS-7 cells expressing the GFP-ALDH9A1 were treated

with C15AlkOPP and the enzyme was immunoprecipitated using an anti-GFP resin. The immobilized protein was reduced, alkylated, and digested on-bead, generating tryptic peptides that were processed for LC–MS analysis. The reduced and alkylated form of the peptide containing Cys288 and Cys289 with deamidated N290 was efficiently detected with a retention time of 45.03 min (Fig. 4(C) and Fig. S2A, ESI†), indicating that some ALDH9A1 exists with free Cys288 and Cys289 thiols; however, some or all of this species could have arisen due to the hydrolytic instability of the thioester. Importantly, we also detected the C15Alk thioester-modified ALDH9A1 with deamidated N290 eluting at a longer retention time (relative to the bis-carboxyamidomethylated form) (Fig. S2B, ESI†), consistent with lipid-modified peptides; however, that peptide gave less intense signals (Fig. 4(D)). Deamidation is a commonly observed modification in MS analysis of peptides digested with trypsin.[37] The detection of the fragment ions $y6+$, $y7+$, $y9+$ and $y10+$ (Table S1, ESI†) are all consistent with the active site C288 thiol being modified with a thioester-linked isoprenoid. Numerous fragments originating from a neutral loss of the thioprenyl group were also observed (*e.g.* $*y6+$, $*y7+$, $*y8+$, $*y9+$, $*y10+$ and $*y11+$). Such neutral loss is commonly observed in the MS/MS fragmentation of thioacylated peptides such as those that are palmitoylated, as well as the thioether linkage in *S*-prenylated peptides.[38,39] Efforts to repeat this experiment with FPP or GGPP were not successful. This may be related to the difficulty of detecting thioester-linked species in proteomic workflows due to their lability in the commonly used reduction and alkylation steps. With 16 cysteine residues, ALDH9A1 is particularly difficult to work with. Moreover, the presence of hydrophobic modifications from the isoprenoids on peptides commonly impacts their ionization efficiency in LC–MS analysis, thereby lowering the signal-to-noise ratio and complicating the detection of the lipid-modified peptide.[40] It is also possible that the farnesoylated or geranylgeranoylated forms of the enzyme undergo ALDH9A1-mediated hydrolysis at a faster rate relative to the probe-modified form resulting in less accumulation or more rapid decay during isolation, rendering them harder to detect.

In vitro analysis of ALDH9A1 with isoprenoids

To gain additional insight into the interactions of isoprenoids with ALDH9A1, *in vitro* experiments were undertaken using the purified enzyme. Since ALDH9A1 is an aldehyde dehydrogenase, initial efforts focused on studying isoprenyl aldehydes as possible substrates. Initially, citral (a C_{10} isoprenyl aldehyde) was studied since the lesser number of carbons increases its aqueous solubility. Kinetic analysis showed that ALDH9A1 manifested a catalytic efficiency (V_{max}/K_M) with citral that was 0.2% of that obtained with the natural substrate, TMABAL (4-*N*-trimethylaminobutryaldehyde) (Table S2 and Fig. S3, ESI†). Given the low solubility of the isoprenoids farnesal and geranylgeranial, it was not possible to work above 50 μM and hence their enzymatic activity was measured at 25 μM aldehyde. Under those conditions, the rate of farnesal oxidation was found to be approximately 5% of that of TMABAL while the
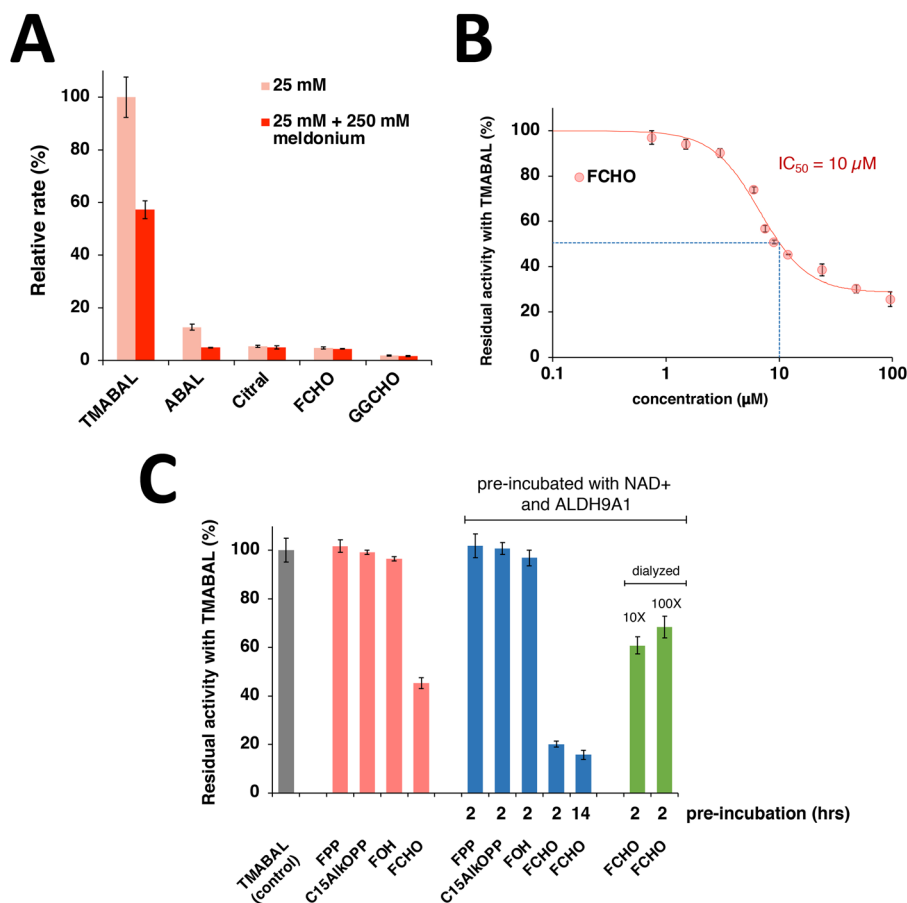
Fig. 5 (A) Relative rates of ALDH9A1 with selected aldehydes. These rates were measured at 25 μM aldehyde in 150 mM sodium pyrophosphate, pH 7.5 containing 1.0 mM NAD$^+$ at 30 °C. Specific activity values with TMABAL were arbitrarily taken as 100%. Error bars indicate S.D. from 4 measurements. (B) Inhibition of ALDH9A1 with farnesal (FCHO) when using TMABAL as a substrate. The data were measured with 6 μM TMABAL, 1.0 mM NAD$^+$ in the same buffer. (C) Effect of FCHO on the TMABAL activity of ALDH9A1. Activities were measured in the presence of 6 μM TMABAL alone (control), and with TMABAL after the addition of 10 μM farnesol, FPP, C15AlkOPP or FCHO (no incubation) or preincubated for 2 and 14 hours with 1.0 NAD$^+$ and enzyme and eventually dialyzed 10× or 100× after 2 hour-long preincubation with FCHO and NAD$^+$. Measured at 30 °C, error bars indicate S.D. from 4 measurements.

rate for geranylgeranial oxidation was closer for to 2.5% (Fig. 5(A)). Those experiments indicate that both farnesal and geranylgeranial are substrates for ALDH9A1 and that they are processed at rates that could have physiological relevance. Assays performed in the presence of meldonium, a product analogue, showed that while TMABAL and ABAL oxidation were significantly inhibited using 250 μM meldonium (3-(2,2,2-trimethylhydrazinium) propionate), isoprenoid oxidation was not inhibited (Fig. 5(A)) suggesting that the modes of TMABAL and long isoprenoid aldehyde binding may be somewhat different. Nevertheless, our results show that meldonium can inhibit the oxidation of the major substrate of ALDH9A1. Interestingly, meldonium is a known inhibitor of γ-butyrobetaine hydroxylase, the enzyme that follows the ALDH9A1 reaction leading to carnitine synthesis.[41] Direct measurements of the dissociation constant obtained via MST confirmed that the K_D for meldonium was 1.2 mM while that of TMABAL was 11 μM (Fig. S4, ESI†).

Next, the ability of farnesal to inhibit TMABAL was studied. Since farnesal is a slower substrate, it was expected that its inclusion would attenuate the rate of TMABAL oxidation through a competitive mechanism. In initial assays where farnesal, TMABAL, and NAD$^+$ were combined at essentially the same time, it was found that farnesal inhibited TMABAL oxidation with an IC$_{50}$ value of approximately 10 μM (Fig. 5(B) and Fig. S5, ESI†); however, it should be noted that the IC$_{50}$ analysis is complicated by the fact that farnesal is a slow alternative substrate. Importantly, no significant inhibition was observed using either farnesol, FPP or C15AlkOPP under those conditions emphasizing the importance of the aldehyde functionality in this inhibitory process (Fig. 5(C)). Next, the preincubation of ALDH9A1 with farnesal in the absence of TMABAL was studied. In that case, a 2 h pre-incubation in the presence of NAD$^+$ led to further inhibition of the TMABAL oxidation reaction, decreasing it from 55% inhibition (without pre-incubation) to 80% inhibition (with pre-incubation) (Fig. 5(C)). Substantial restoration of activity was obtained by dialysis of the inactivated sample, returning to approximately 60% of the initial activity. Slow restoration is consistent with the formation of a covalently modified enzyme species such as

a thioester intermediate that undergoes hydrolysis at a low rate to regenerate the active form of the enzyme. The accumulation of a thioester intermediate and its slow hydrolysis provide a compelling rationale for the labeling of ALDH9A1 by C15AlkOPP observed in the metabolic labeling experiments described above.

Mechanistic interpretation and biological implications

Overall, a compelling case for the prenylation of ALDH9A1 is presented here. Initial work using chemical proteomics demonstrated that ALDH9A1 can be modified by the probe C15AlkOPP in at least 4 different cell types. Mutational analysis of the protein suggests that it is modified on Cys288 however experiments with various prenyltransferase inhibitors indicate that the modification is not catalyzed by a canonical prenyltransferase. The existence of a base-labile thioester and the requirement for an intact NAD$^+$ binding site (inferred by mutagenesis) suggest a simple mechanism for this modification process where ALDH9A1 catalyzes the NAD$^+$-dependent oxidation of a prenyl aldehyde to the corresponding thioester (see Fig. 3(A)) that is limited by a slow rate of hydrolysis resulting in accumulation of the covalently modified enzyme. Such a thioester species is a well-established intermediate in the mechanism catalyzed by aldehyde dehydrogenases[42] and the slow hydrolysis of such an acylated enzyme species is known to be responsible for the ability of certain aldehydes to serve as covalent ALDH inhibitors.[43] In this proposed mechanism, farnesal and/or geranylgeranial would originate from the prenylcysteine oxidase-mediated deprenylation of farnesyl- or geranylgeranylcysteine.[35] Previous work with ALDH9A1 demonstrated that the related isoprenoid, geranial was both an inhibitor and substrate that is slowly converted to geranic acid.[27,44] Here we demonstrated that the longer isoprenoids farnesal and geranylgeranial are also substrates and inhibitors for that enzyme. It is interesting to note that a substantial conformational change in the enzyme occurs upon nucleotide binding. In the NAD$^+$ free form, the enzyme active site is open and the catalytic residue Cys288 is solvent exposed[16] while in the presence of NAD$^+$ that cysteine residue is sequestered (Fig. S6, ESI†).[17] In the later form, the hydrolysis of the putative thioester intermediate could be retarded, resulting in the accumulation of the prenylated species. To gain more understanding concerning the ALDH9A1-substrate complexes, models were created by docking either TMABAL or farnesal into the active site (Fig. 6, top). Using the structure of the related enzyme ALDH10 covalently linked to an aldehyde substrate *via* a thiohemiacetal, the docked conformation of farnesal was manually adjusted to create a similar species that would serve as the progenitor for the subsequent thioester intermediate in ALDH9A1-catalyzed oxidation of farnesal. Inspection of that latter structure shows how farnesal could form the requisite catalytically competent thiohemiacetal intermediate (Fig. 6, bottom). Interestingly, in that model, Cys289 (the noncatalytic cysteine) is 3.7 Å from C-3 of farnesal. It is possible that residue could undergo Michael addition to C-3 to form a cyclic adduct that would be resistant to hydrolytic cleavage (see Fig. S7, ESI†
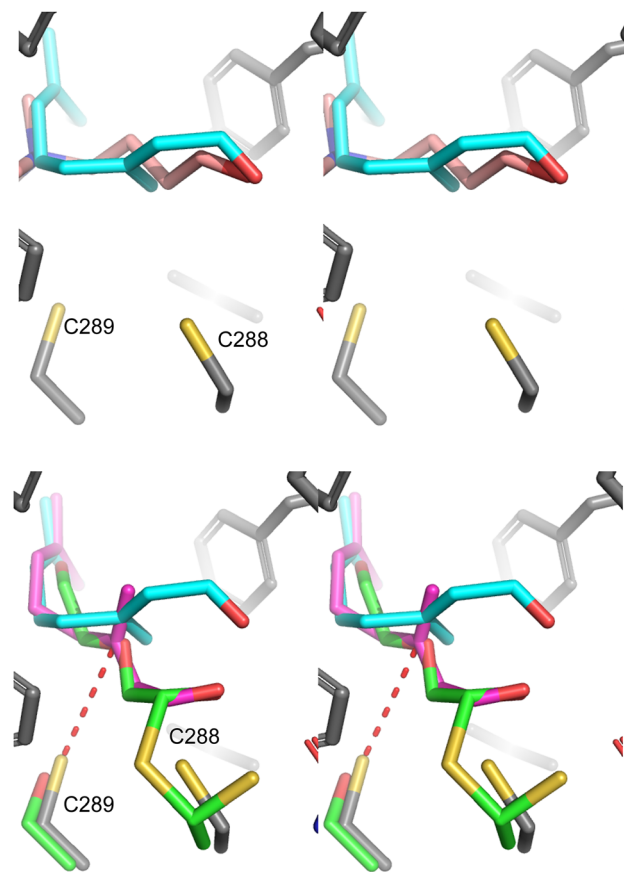

Fig. 6 Structural analysis of ALDH9A1 complexed with farnesal. (Top) Stereo view of the structure of ALDH9A1 (pdb code 6VR6) complexed with FCHO (cyan C) or TMABAL (pink C) obtained *via* docking. (Bottom) Superposition of ALDH9A1 complexed with FCHO (magenta C, obtained by docking) superimposed on the structure of ALDH10 linked to an aldehyde substrate, 2-(2-hydroxyethoxy)acetaldehyde, (green C) *via* a thiohemiacetal (pdb code 4I9B). The structure of the FCHO moiety was manually rotated (first 5 C−C bonds) to position the aldehyde on top of the aldehydic carbon of 2-(2-hydroxyethoxy)acetaldehyde. The unrotated docked structure of FCHO (cyan C) is also shown for comparison. In both structures, carbons from ALDH9A1 are shown in grey and carbons from ALDH10 are shown in green. Oxygen (red), nitrogen (blue) and sulfur (yellow) atoms are also highlighted. Cys289 of ALDH9A1 (or a serine at the homologous position in ALDH10) is shown at the bottom left and Cys288 of ALDH9A1 is shown at the bottom right. The distance between the S atom of Cys289 and C-3 of farnesal (red dashed line) is 3.7 Å.

for chemical structures). While no evidence of that species was detected *via* mass spectrometry, poor ionization could have precluded its detection. The presence of that adduct could explain why an accumulation of ALDH9A1, covalently labeled by C15AlkOPP, occurs in metabolic labelling experiments and why extensive dialysis of the enzyme after preincubation with farnesal does not result in complete recovery of activity. Such an adduct is not consistent with the reversibility of labelling obtained upon hydroxylamine treatment but that data cannot be used to rule out the formation of limited amounts of that species.

An important question concerning these observations is what is their biological significance? The ALDH superfamily

920 | *RSC Chem. Biol.*, 2023, **4**, 913–925

© 2023 The Author(s). Published by the Royal Society of Chemistry

constitutes an important class of enzymes that play key roles in modulating oxidative/electrophilic stress inside cells.[45] In fact, dysregulation in their expression and function has been implicated in a variety of cancers and inhibitors specific to these enzymes have been developed for cancer therapeutics.[46] Similar to sorbitol and alcohol dehydrogenases, ALDHs have been known to possess reactive nucleophilic cysteines in their active sites that drive the conversion of toxic aldehyde species to innocuous carboxylic acid derivatives, or in deactivating reactive aldehyde and oxygen species.[45,47] In the case of ALDH9A1, it may play an important physiological role in converting the highly reactive α,β-unsaturated aldehydes farnesal and geranylgeranial, the natural degradation products of prenylated proteins, to their less reactive α,β-unsaturated acids.

Recent developments in chemical proteomic strategies have allowed for profiling of proteins bearing reactive cysteines in which several ALDHs have been identified.[48–50] In particular, ALDH9A1 has been consistently identified in these chemical proteomic studies, underscoring the high reactivity of its active Cys288 residue. Another salient characteristic of ALDH9A1 is the presence of Cys289 adjacent to the nucleophilic Cys288. Other ALDHs also contain this dual (even triple) cysteine motif: ALDH2 (Cys318/Cys319*/Cys320), ALDH1A1 (Cys302/Cys303*), ALDH1A2 (Cys319/Cys320*), ALDH1A3 (Cys313/Cys314*), and ALDH1B1 (Cys318/Cys319*/Cys320), where * denotes the catalytic cysteine. Among these ALDHs, only ALDH9A1 was consistently enriched in our prenylomic analyses. Previous studies in ALDH1A1 have shown that Cys302 adjacent to the nucleophilic Cys303 may control its activity through disulfide bond formation and is potentially important for cellular response to oxidative stress.[31] Prenylation of Cys288 in ALDH9A1 may be a related mechanism for controlling the activity of that enzyme. Inhibition of ALDH9A1 by modification of Cys288 could reduce levels of carnitine and GABA, which are both produced by biosynthetic pathways involving that enzyme.[51,52] Changes in carnitine levels could in turn impact fatty acid metabolism while perturbations in GABA levels could have neurological effects. There is growing evidence that Alzheimer's disease (AD) is associated with dysregulation of fatty acid metabolism[53] and reductions in GABA levels have been described in severe cases of AD.[54] Interestingly, we recently identified ALDH9A1 as a protein with increased levels of prenylation in AD.[55] Since prenylation of ALDH9A1 on the catalytic residue Cys288 must inhibit its activity, this observation may provide a link between oxidative stress, neuroinflammation and Alzheimer's disease.[56]

As a final point, it is worth noting that protein prenylation typically results in membrane localization. However, in the case of ALDH9A1, that does not appear to occur as noted in the aforementioned experiments with GFP fusions (Fig. 3(C)). One reason for lack of membrane localization is probably related to the solvent exposure of the lipid. Cys288 is located in a deep pocket within the enzyme. Modelling of the farnesoylated form of the enzyme (Fig. S8, ESI†) reveals that the prenyl group is largely sequestered by the enzyme, preventing it from inserting into a lipid membrane. Additionally, the electrostatic potential surface (Fig. S8, ESI†) is predominantly negative in the region

surrounding the active site pocket, rendering it less likely to interact with the anionic phospholipids that make up cellular membranes. Those features are in contrast to many prenylated proteins such as K-Ras4B where the prenyl group is present on the C-terminus of the protein, making it is freely accessible for membrane insertion and proximal to a polybasic region directly upstream of the prenylation site that can enhance binding to the anionic membrane.[57] Finally, it is important to note that while prenylation is essentially irreversible, acylation modifications such as palmitoylation are not. Thus, the biological activity of palmitoylated proteins is often regulated by deacylation catalyzed by thioesterases.[58] Given the limited accessibility of the prenoylated form of ALDH9A1, it is difficult to envision how a thioesterase could catalyze hydrolysis of the lipidated enzyme. However, it is quite possible that regulatory proteins could bind to ALDH9A1 and either stimulate or retard hydrolysis of the thioester modified protein. By modulating global ALDH9A1 activity through changes in its acylation state, such proteins could regulate a range of cellular processes.

Conclusion

In this work, we described the potential isoprenoid modification on ALDH9A1 discovered through chemical proteomics. Through the use of tools from chemical and molecular biology, we demonstrated that this isoprenoid labeling on ALDH9A1 is robust but reversible. This potentially new type of isoprenoid modification opens new doors into the chemistry of lipid modifications on proteins, particularly in modifying the catalytic residue of an enzyme that may have functional roles in regulating its activity. The experiments described here set the stage for future work aimed at understanding this unusual post-translational modification. More studies of this novel protein modification are therefore anticipated.

Experimental

General reagents and cell lines

COS-7 and HeLa cells were generously provided by Dr Elizabeth Wattenberg at the University of Minnesota, USA. The FTI L-744,832 was purchased from Merck and GGTI-286 from CalBiochem. The RabGGTase inhibitors 1, 2, and 3 were prepared as previously described.[22,23]

Plasmid preparation

A plasmid expressing human ALDH9A1 was purchased from Genecopoeia. The ALDH9A1 gene was inserted in a vector (EX-Z3075-M29) to contain an N-terminal GFP tag controlled by the CMV promoter. Plasmids were transformed in 100 μL of competent DH5α E. coli (New England Biolabs) following the manufacturer's protocol. Transformed bacteria were plated in Luria broth (LB) agar plates containing ampicillin and colonies were isolated for bacterial cultures in 50 mL LB media. Cultures were grown overnight at 37 °C and bacteria were pelleted by centrifugation at 6000 × g for 20 min at 4 °C. The bacterial

pellets were suspended in sterile water and plasmids were extracted using a PureYield™ Plasmid Miniprep System (Promega) following the manufacturer's protocol. Plasmid concentrations were determined by taking the absorbance at 260 and 280 nm using a NanoPhotometer P330 (IMPLEN) and using the equation: concentration $(\mu g\ mL^{-1}) = (A_{260} - A_{320}) \times$ dilution factor $\times 50\ \mu g\ mL^{-1}$. Only good-quality plasmids ($A_{260}/A_{280} \sim$ 1.7 to 1.8) were used for subsequent transfections.

Transfection, in-gel fluorescence and cellular imaging

COS-7 cells were grown in 100 mm culture plates under DMEM media (Gibco) supplemented with 10% fetal bovine serum (Gibco) and 1% penicillin–streptomycin (Gibco). Cells were passaged before reaching 80% confluency and repeated at least three times prior to transfection. A day before transfection, cells (1.6×10^6 per well) were seeded in a 6-well plate and grown overnight until 90% confluency was achieved. The media was removed and replaced with 10 mL of Opti-MEM™ Reduced Serum Media (Thermo Fisher Scientific). Plasmid DNA (15 μg) and 50 μL of Endofectin Max (Genecopoeia) were each diluted to 750 μL with Opti-MEM™ in separate tubes. The solutions were combined and added to the cells dropwise. After 8 hours post-transfection, 10 μM of C15AlkOPP, isoprenoid analogues, and inhibitors (when indicated) were added and allowed to incubate for 16 more hours. Cells were washed and collected by scraping and subjected to lysis in 1× PBS + 1% SDS as described in a previously reported protocol.[59] Protein concentrations were determined using BCA assay (Thermo Scientific).

For in-gel fluorescence analysis, 100 μg of protein in lysate were used for the click reaction along with 25 μM TAMRA-N$_3$, 1 mM TCEP, 0.1 mM TBTA, and 1 mM CuSO$_4$. Labeled proteins were resolved in a 12% SDS-PAGE gel. TAMRA fluorescence (542/568 nm excitation/emission) was detected using a Typhoon FLA 9500 instrument (GE Healthcare). Gels were then stained with Coomassie blue and destained. For western blot analysis, proteins in gels were transferred to a PVDF membrane using a Mini Trans-Blot® Electrophoretic Transfer Cell (Bio-Rad) and blocked with 5% milk in 1× TBST. Membranes were incubated with mouse anti-GFP at 1:5000 dilution in 5% milk in 1× TBST (ABclonal, #AE012) overnight followed by incubation with a secondary goat HRP-conjugated anti-mouse at 1:10 000 dilution in milk (ABclonal, #AS003) and treated with Clarity™ ECL (Bio-Rad). Chemiluminescence was detected using iBright (Thermo Fisher). Gel images were processed using ImageJ.

For fluorescence cellular imaging, cells (500 000) were seeded in 35 mm glass-bottomed dish (Ibidi USA) and grown overnight until 90% confluency. The DMEM media was replaced with 2 mL of Opti-MEM™. DNA plasmids (2.5 ng in 125 μL Opti-MEM™) and Endofectin (7.5 μL diluted to 125 μL with Opti-MEM™) were mixed and added dropwise to the cells. After 24 h, the media was removed and cells were washed with cold 1× PBS (2 mL) twice, followed by staining with Hoechst 33 342 (Thermo Scientific) and ER-Tracker™ Red (Invitrogen) following the manufacturers' protocols. Cells were washed twice with PBS and visualized using a FluoView FV1000IX2

Inverted Confocal Microscope (Olympus) with a 60× objective. The .oib files were imported in Fiji and formatted.

Site-directed mutagenesis

Human ALDH9A1 mutants were generated from the template plasmid (EX-Z3075-M29) described above using a QuickChange XL II Site-Directed Mutagenesis Kit (Agilent). Primers were designed with the corresponding mutations as prescribed by the QuickChange Primer Design tool with minor alterations at the 3′ end and the oligos were purchased from Integrated DNA Technologies. In brief, the template plasmid (50 ng) was incubated with the forward (125 ng) and reverse (125 ng) primers, dNTP mix, and PfuTurbo DNA Polymerase (2.5 U). PCR mutagenesis was carried out in an Arktik Thermal Cycler (Thermo Scientific) under the following thermal program: initial denaturation (95 °C, 1 min); 18 cycles of denaturation (95 °C, 50 s), annealing (60 °C, 50 s), and extension (68 °C, 8 min); final extension (68 °C, 7 min). The resulting mixture was digested with DpnI and 5 μL of the mutagenesis reaction was transformed in XL10-Gold® Ultracompetent Cells (Agilent) and selected on ampicillin-containing LB-agar plates. Colonies were isolated, cultured and extracted with plasmids using PureYield™ Plasmid Miniprep System. Successful mutations were verified via Sanger-type DNA sequencing performed by the University of Minnesota Genomics Center. Purified plasmids were transformed in DH5α E. coli, cloned, purified, and assayed for concentration using NanoPhotometer P330 (IMPLEN).

Mapping of the modification site

Expressed GFP-ALDH9A1 was enriched using anti-GFP beads following the manufacturer's protocol for GFP-Trap (Chromotek). Briefly, COS-7 cells ($\sim 1 \times 10^6$) transfected with GFP-ALDH9A1 and treated with C15AlkOPP, FPP or GGPP were lysed in lysis buffer (100 μL, 10 mM tris–HCl pH 7.5, 10 mM NaCl, 0.5 mM EDTA, 0.5% Nonidet™ P40 Substitute (Sigma-Aldrich)) by pipetting in and out in a 1 mL syringe with 16 gauge needle on ice. The soluble proteins were separated from the cellular debris via centrifugation (20 000 × g at 4 °C for 15 min). Pre-washed GFP-Trap beads (25 μL) were added to the lysates and incubated for 1 h at 4 °C with head-to-tail mixing. The beads were then washed with wash buffer (10 mM tris–HCl pH 7.5, 10 mM NaCl, 0.5 mM EDTA) and the remaining immobilized protein was reduced and digested in 50 mM tris–HCl pH 7.5 containing 2 M urea, 10 μg mL^{-1} sequencing grade trypsin (Promega), 20 μg mL^{-1} chymotrypsin (Promega), 1 mM CaCl$_2$, and 0.5 mM TCEP for 30 min at 32 °C. The supernatant was collected using a spin trap (Pierce) and the beads were suspended in 50 mM tris–HCl pH 7.5 containing 2 M urea and 5 mM iodoacetamide. The eluate was collected and pooled with the first eluate. Samples were dried via lyophilization and redissolved in 2% CH$_3$CN in H$_2$O with 0.1% HCO$_2$H. The peptides were desalted using STAGE tips following standard protocols. The recovered peptides were dried and redissolved in 0.1% HCO$_2$H in H$_2$O.

The samples were resolved using a reversed-phase column and analyzed with an Orbitrap Lumos instrument. Separation was performed in 5–30% of 0.1% HCO_2H in CH_3CN for 140 min. A targeted database determined using Skyline was used corresponding to the m/z of the active site peptide resulting from tryptic and chymotryptic digestion with a maximum of 3 missed cleavages. MS1 scans were collected at 60 000 resolution over 320–2000 m/z range with an AGC target of 500 000 and max IT of 50 ms. Dynamic exclusion was allowed for 90 s. HCD fragmentation was performed at NCE of 42% with 1.5 m/z isolation window, AGC of 50 000, and max IT of 200 ms. The.raw files were analyzed in pFind to search for the following modifications on cysteine residue: carbamidomethylation (C_2H_3NO, 57.0215), C15Alk ($C_{18}H_{26}O$, 258.1984 Da), C15Alk-thioester ($C_{18}H_{24}O_2$, 272.1776 Da), C15Alk-thiohemiacetal ($C_{18}H_{26}O_2$, 274.1933 Da), farnesyl ($C_{15}H_{24}$, 204.1878 Da), farnesyl ester ($C_{15}H_{22}O$, 218.1671 Da), farnesyl thiohemiacetal ($C_{15}H_{24}O$, 220.1827 Da), geranylgeranyl ($C_{20}H_{32}$, 272.2504 Da), geranylgeranyl ester ($C_{20}H_{30}O$, 286.2297 Da), and geranylgeranyl thiohemiacetal ($C_{20}H_{32}O$, 288.2453 Da). The neutral losses were also included. Other dynamic modifications include methionine oxidation (16.0000 Da) and arginine and glutamine deamidation (0.9840 Da). Matching spectra were exported and formatted for visualization.

In vitro experiments with recombinant ALDH9A1

Enzyme activity was measured by monitoring the NADH formation (ε_{340} = 6.22 mM^{-1} cm^{-1}) using a UV-Vis spectrophotometer 8453 (Agilent 8453) at 30 °C in 150 mM sodium pyrophosphate, pH 7.5 containing 1.0 mM NAD^+ as previously described.[16] Relative activities with several aldehydes were compared at 25 μM due to the low solubility of the long aliphatic aldehydes. Inhibition experiments were measured using a TMABAL concentration close to its K_m value (6 μM or 10 μM concentrations were used). An MST method was employed to determine binding affinities of ALDH9A1 for meldonium and TMABAL. The enzyme was labeled with the 2nd generation RED-tris-NTA dye and adjusted to 90 nM concentration with 40 mM sodium phosphate buffer, pH 7.4, supplemented with 0.05% Tween 20. A series of sixteen 1 : 1 ligand dilutions was prepared using the identical buffer. Measurements were obtained using a Monolith NT.115 instrument (NanoTemper Technologies) at 30 °C using premium capillaries.

Docking of TMABAL and farnesal into the active site of ALDH9A1

In silico docking was performed to gain insight into TMABAL and farnesal binding modes.[60–62] Both molecules were docked into the active site of the ALDH9A1 (PDB 6VR6)[17] using PyRx software 1.0 (**https://pyrx.sourceforge.io/**)[62] and AutoDock Vina 1.05 (**https://vina.scripps.edu/**).[61] PDBQT files were generated using AutoDockTools package (**https://mgltools.scripps.edu/**). Energy grids for docking were 35 × 35 × 35 Å in dimension and they were centered on the catalytic cysteine. Docking calculations were carried out using the Lamarckian genetic

algorithm.[60] The population was 150, the maximum number of generations was 27 000, and the maximum number of energy evaluations was 3.5 × 10^6. The resulting ligand orientations and conformations were scored based on their binding free energies. To model the structure of the thiohemiacetal intermediate, the structure of farnesal docked into ALDH9A1 was superimposed on the structure of ALDH10 (PDB 4I9B) containing a thiohemiacetal with 2-(2-hydroxyethoxy)acetaldehyde.[63] C–C bonds 1–5 of farnesal were manually rotated to superimpose C-1 of farnesal onto the aldehydic carbon from 2-(2-hydroxyethoxy)acetaldehyde.

Data availability

The raw mass spectrometry files for the proteomics data have been deposited to the ProteomeXchange Consortium *via* the PRIDE partner repository with the dataset identifier PXD036365.

Author contributions

KF performed prenylomic analysis, cellular imaging, transfection and mutagenesis experiments. JB, GS and SA assisted with some gel experiments, enzyme assays and bioinformatic analysis. AP synthesized enzyme substrates. KB provided enzyme inhibitors. KS, LL, KB, DK and MD conceptualized, advised, and guided project development. KS, DK and MD wrote the paper with input and guidance from all authors. KS, DK, and MD interpreted the data. All authors critically read and reviewed the manuscript.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

References

1 C. J. Marshall, *Science*, 1993, **259**, 1865–1866.

2 F. L. Zhang and P. J. Casey, *Annu. Rev. Biochem.*, 1996, **65**, 241–269.

3 M. J. Blanden, K. F. Suazo, E. R. Hildebrandt, D. S. Hardgrove, M. Patel, W. P. Saunders, M. D. Distefano, W. K. Schmidt and J. L. Hougland, *J. Biol. Chem.*, 2017, **293**, 2770–2785.

4 S. Ashok, E. R. Hildebrandt, C. S. Ruiz, D. S. Hardgrove, D. W. Coreno, W. K. Schmidt and J. L. Hougland, *Biochemistry*, 2020, **59**, 1149–1162.

5 G. L. Schey, P. H. Buttery, E. R. Hildebrandt, S. X. Novak, W. K. Schmidt, J. L. Hougland and M. D. Distefano, *Int. J. Mol. Sci.*, 2021, **22**, 12042.

6 B. T. Kinsella and W. A. Maltese, *J. Biol. Chem.*, 1992, **267**, 3940–3945.

7 S. Kuchay, H. Wang, A. Marzio, K. Jain, H. Homer, N. Fehrenbacher, M. R. Philips, N. Zheng and M. Pagano, *Nat. Struct. Mol. Biol.*, 2019, **26**, 628–636.

8 R. Shirakawa, S. Goto-Ito, K. Goto, S. Wakayama, H. Kubo, N. Sakata, D. A. Trinh, A. Yamagata, Y. Sato, H. Masumoto, J. Cheng, T. Fujimoto, S. Fukai and H. Horiuchi, *EMBO J.*, 2020, **39**, e104120.

9 E. W. Tate, K. A. Kalesh, T. Lanyon-Hogg, E. M. Storck and E. Thinon, *Curr. Opin. Chem. Biol.*, 2015, **24**, 48–57.

10 K. F. Suazo, G. Schey, C. Schaber, A. R. Odom John and M. D. Distefano, *Mass Spectrom. Chem. Proteomics*, 2019, 317–347.

11 E. M. Storck, J. Morales-Sanfrutos, R. A. Serwa, N. Panyain, T. Lanyon-Hogg, T. Tolmachova, L. N. Ventimiglia, J. Martin-Serrano, M. C. Seabra, B. Wojciak-Stothard and E. W. Tate, *Nat. Chem.*, 2019, **11**, 552–561.

12 K. F. Suazo, A. Jeong, M. Ahmadi, C. Brown, W. Qu, L. Li and M. D. Distefano, *Sci. Rep.*, 2021, **11**, 4367.

13 K. F. Suazo, K.-Y. Park and M. D. Distefano, *Chem. Rev.*, 2021, **121**, 7178–7248.

14 W. Qu, K. F. Suazo, W. Liu, S. Cheng, A. Jeong, D. Hottman, L.-L. Yuan, M. D. Distefano and L. Li, *Mol. Neurobiol.*, 2021, **58**, 1128–1144.

15 G. Charron, M. M. H. Li, M. R. MacDonald and H. C. Hang, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 11085–11090.

16 R. Končitíková, A. Vigouroux, M. Kopečná, M. Šebela, S. Moréra and D. Kopečný, *Biosci. Rep.*, 2019, **39**, BSR20190558.

17 J. W. Wyatt, D. A. Korasick, I. A. Qureshi, A. C. Campbell, K. S. Gates and J. J. Tanner, *Arch. Biochem. Biophys.*, 2020, **691**, 108477.

18 T. Kuzuguchi, Y. Morita, I. Sagami, H. Sagami and K. Ogura, *J. Biol. Chem.*, 1999, **274**, 5888–5894.

19 C. C. Palsuledesai, J. D. Ochocki, M. M. Kuhns, Y.-C. Wang, J. K. Warmka, D. S. Chernick, E. V. Wattenberg, L. Li, E. A. Arriaga and M. D. Distefano, *ACS Chem. Biol.*, 2016, **11**, 2820–2828.

20 E. C. Lerner, Y. Qian, A. D. Hamilton and S. M. Sebti, *J. Biol. Chem.*, 1995, **270**, 26770–26773.

21 J. E. Gisselberg, L. Zhang, J. E. Elias and E. Yeh, *Mol. Cell. Proteomics*, 2017, **16**, S54–S64.

22 A. Kaźmierczak, D. Kusy, S. P. Niinivehmas, J. Gmach, Ł. Joachimiak, O. T. Pentikäinen, E. Gendaszewska-Darmach and K. M. Błażewska, *J. Med. Chem.*, 2017, **60**, 8781–8800.

23 Ł. Joachimiak, A. Marchwicka, E. Gendaszewska-Darmach and K. M. Błażewska, *ChemMedChem*, 2018, **13**, 842–851.

24 F. P. Coxon, Ł. Joachimiak, A. K. Najumudeen, G. Breen, J. Gmach, C. Oetken-Lindholm, R. Way, J. E. Dunford, D. Abankwa and K. M. Błażewska, *Eur. J. Med. Chem.*, 2014, **84**, 77–89.

25 N. A. Sophos and V. Vasiliou, *Chem. – Biol. Interact.*, 2003, **143–144**, 5–22.

26 S. A. Marchitti, C. Brocker, D. Stagos and V. Vasiliou, *Expert Opin. Drug Metab. Toxicol.*, 2008, **4**, 697–720.

27 V. Koppaka, D. C. Thompson, Y. Chen, M. Ellermann, K. C. Nicolaou, R. O. Juvonen, D. Petersen, R. A. Deitrich, T. D. Hurley and V. Vasiliou, *Pharmacol. Rev.*, 2012, **64**, 520–539.

28 M. K. Chern and R. Pietruszko, *Biochem. Biophys. Res. Commun.*, 1995, **213**, 561–568.

29 F. M. Vaz, S. W. Fouchier, R. Ofman, M. Sommer and R. J. A. Wanders, *J. Biol. Chem.*, 2000, **275**, 7390–7394.

30 G. Kurys, W. Ambroziak and R. Pietruszko, *J. Biol. Chem.*, 1989, **264**, 4715–4721.

31 Y. Zhang, M. Wang and H. Lin, *ACS Chem. Biol.*, 2020, **15**, 28–32.

32 T. Kitamura, S. Takagi, T. Naganuma and A. Kihara, *Biochem. J.*, 2014, **465**, 79–87.

33 A. L. Fink, *Folding Des.*, 1998, **3**, R9–R23.

34 M. Bentinger, J. Grünler, E. Peterson, E. Swiezewska and G. Dallner, *Arch. Biochem. Biophys.*, 1998, **353**, 191–198.

35 W. R. Tschantz, J. A. Digits, H.-J. Pyun, R. M. Coates and P. J. Casey, *J. Biol. Chem.*, 2001, **276**, 2321–2324.

36 A. Petenkova, S. A. Auger, J. Lamb, D. Quellier, C. Carter, O. T. To, J. Milosevic, R. Barghout, A. Kugadas, X. Lu, J. Geddes-McAlister, R. Fichorova, D. B. Sykes, M. D. Distefano and M. Gadjeva, *Nat. Commun.*, 2023, **14**, 2761.

37 P. Hao, Y. Ren, A. J. Alpert and S. K. Sze, *Mol. Cell. Proteomics*, 2011, **10**, O111.009381.

38 Y. Ji, N. Leymarie, D. J. Haeussler, M. M. Bachschmid, C. E. Costello and C. Lin, *Anal. Chem.*, 2013, **85**, 11952–11959.

39 D. G. Mullen, K. Kyro, M. Hauser, M. Gustavsson, G. Veglia, J. M. Becker, F. Naider and M. D. Distefano, *Bioorg. Med. Chem.*, 2011, **19**, 490–497.

40 E. W. Tate, *J. Chem. Biol.*, 2008, **1**, 17–26.

41 K. Tars, J. Rumnieks, A. Zeltins, A. Kazaks, S. Kotelovica, A. Leonciks, J. Sharipo, A. Viksna, J. Kuka, E. Liepinsh and M. Dambrova, *Biochem. Biophys. Res. Commun.*, 2010, **398**, 634–639.

42 H. Morrison in *Enzyme Active Sites and their Reaction Mechanisms*, ed. H. Morrison, Academic Press, New York, 2021, pp.21–26.

43 C. A. Morgan, B. Parajuli, C. D. Buchman, K. Dria and T. D. Hurley, *Chem. – Biol. Interact.*, 2015, **234**, 18–28.

44 A. Kikonyogo, D. P. Abriola, M. Dryjanski and R. Pietruszko, *Eur. J. Biochem.*, 1999, **262**, 704–712.

45 S. Singh, C. Brocker, V. Koppaka, Y. Chen, B. C. Jackson, A. Matsumoto, D. C. Thompson and V. Vasiliou, *Free Radicals Biol. Med.*, 2013, **56**, 89–101.

46 S. S. Dinavahi, C. G. Bazewicz, R. Gowda and G. P. Robertson, *Trends Pharmacol. Sci.*, 2019, **40**, 774–789.

47 J. Johansson, L. Fleetwood and H. Jörnvall, *FEBS Lett.*, 1992, **303**, 1–3.

48 B. Ford, L. A. Bateman, L. Gutierrez-Palominos, R. Park and D. K. Nomura, *Cell Chem. Biol.*, 2017, **24**, 133–140.

49 L. R. Whitby, R. S. Obach, G. M. Simon, M. M. Hayward and B. F. Cravatt, *ACS Chem. Biol.*, 2017, **12**, 2040–2050.

50 Z. Huang, D. Ogasawara, U. I. Seneviratne, A. B. Cognetta 3rd, C. W. Am Ende, D. M. Nason, K. Lapham, J. Litchfield, D. S. Johnson and B. F. Cravatt, *ACS Chem. Biol.*, 2019, **14**, 192–197.

51 M. Almannai, M. Alfadhel and A. W. El-Hattab, *Molecules*, 2019, **24**, 3251.

52 S. W. Lin, J. C. Chen, L. C. Hsu, C.-L. Hsieh and A. Yoshida, *Genomics*, 1996, **34**, 376–380.

53 H. Chew, V. A. Solomon and A. N. Fonteh, *Front. Physiol.*, 2020, **11**, 598.

54 M. Solas, E. Puerta and J. M. Ramirez, *Curr. Pharm. Des.*, 2015, **21**, 4960–4971.

55 A. Jeong, S. A. Auger, S. Maity, K. Fredriksen, R. Zhong, L. Li and M. D. Distefano, *ACS Chem. Biol.*, 2022, **17**, 2863–2876.

56 A. Nunomura, G. Perry, G. Aliev, K. Hirai, A. Takeda, E. K. Balraj, P. K. Jones, H. Ghanbari, T. Wataya, S. Shimohama, S. Chiba, C. S. Atwood, R. B. Petersen and M. A. Smith, *J. Neuropathol. Exp. Neurol.*, 2001, **60**, 759–767.

57 Y. Zhou, P. Prakash, H. Liang, K.-J. Cho, A. A. Gorfe and J. F. Hancock, *Cell*, 2017, **168**, 239–251.e16.

58 A. Main and W. Fuller, *FEBS J.*, 2022, **289**, 861–882.

59 K. F. Suazo, A. K. Hurben, K. Liu, F. Xu, P. Thao, C. Sudheer, L. Li and M. Distefano, *Curr. Protoc. Chem. Biol.*, 2018, **10**, e46.

60 G. M. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew and A. J. Olson, *J. Comput. Chem.*, 1998, **19**, 1639–1662.

61 O. Trott and A. J. Olson, *J. Comput. Chem.*, 2010, **31**, 455–461.

62 S. Dallakyan and A. J. Olson, in *Chemical Biology: Methods and Protocols*, ed. J. E. Hempel, C. H. Williams and C. C. Hong, Springer, New York, 2015, pp.243–250.

63 D. Kopečný, R. Končitíková, M. Tylichová, A. Vigouroux, H. Moskalíková, M. Soural, M. Šebela and S. Moréra, *J. Biol. Chem.*, 2013, **288**, 9491–9507.