Published in final edited form as:

Chem Biol Interact. 2017 October 01; 276: 23–30. doi:10.1016/j.cbi.2017.04.011.

Modeling of interactions between functional domains of ALDH1L1

David A. Horita[1] and **Sergey A. Krupenko**[1,2,*]

[1]Nutrition Research Institute, University of North Carolina at Chapel Hill, Kannapolis, NC 28081

[2]Department of Nutrition, University of North Carolina at Chapel Hill, Kannapolis, NC 28081

Abstract

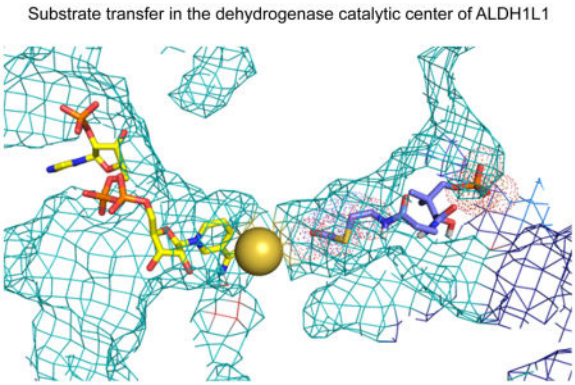
ALDH1L1, a member of the aldehyde dehydrogenase superfamily of enzymes, catalyzes the conversion of 10-formyltetrahydrofolate to tetrahydrofolate and CO_2. The enzyme is a tetramer of identical subunits, with each subunit consisting of three functional domains that originated from unrelated genes. The N- and C-terminal domains are catalytic, while the intermediate domain transfers the reaction intermediate from the N- to the C-terminal domain. The intermediate domain is an acyl carrier protein, possessing the covalently attached 4′-phosphopantetheine (4-PP) prosthetic group. This prosthetic group is known to function as a swinging arm transferring intermediates between enzymes in complex biosynthetic reactions. Here we have applied computer modeling using available structures of the three functional domains of ALDH1L1 to evaluate the extent of flexibility within the full-length protein. This approach allowed us to define positions of the 4-PP arm within the two catalytic domains and to predict N-terminal:intermediate and intermediate:C-terminal domain interfaces. Our models further suggested high degree of flexibility within the full-length enzyme.

Graphical Abstract

*Corresponding author. Tel. + 1 704 250 5053, sergey_krupenko@unc.edu.

Conflict of interest

None.



Keywords

Aldehyde dehydrogenase; ALDH1L1; formyltransferase; computational docking; acyl carrier protein; 4′-phosphopantetheine prosthetic group

1. Introduction

ALDH1L1 (aldehyde dehydrogenase 1 family member L1; 10-formyltetrahydrofolate dehydrogenase) is a member of the aldehyde dehydrogenase (ALDH) superfamily of enzymes (1,2). This enzyme plays a key role in the regulation of folate-mediated one-carbon metabolism and downstream cellular processes (3–9). ALDH1L1 catalyzes oxidative decarboxylation of 10-formyltetrahydrofolate (10-fTHF) to tetrahydrofolate with concomitant reduction of $NADP^+$ to NADPH (Fig. 1) (10). The two-step reaction is carried out by structurally and functionally distinct catalytic domains linked by a physical transfer of a formyl group from the first reaction center to the second (11–13). The need to couple two discrete reaction steps led to the fusion of three unrelated primordial genes, yielding a complex three-domain protein (1). Full-length ALDH1L1 is a homotetramer of 902 amino acid subunits, with each monomeric subunit having a modular organization (Fig. 1). The 500-residue C-terminal core domain includes the tetramer interface (14). This domain is a structural and functional homolog of aldehyde dehydrogenases and is responsible for the original assignment of the protein to the ALDH superfamily. The 310-residue N-terminal domain resembles methionine-tRNA formyltransferase, binds the folate substrate, and can function (at least *in vitro*) as a 10-fTHF hydrolase (12,15). The two catalytic domains are connected by and communicate *via* the 90-residue intermediate domain, which is a structural and functional homolog of acyl carrier proteins (ACP) (13). Another example of such a carrier protein in humans is the fatty acid synthase complex, where it functions in transferring of growing acyl chain (16). The distinctive feature of acyl-carrier proteins is the presence of a covalently linked 4′-phosphopantetheine (4-PP) prosthetic group which functions as a long (~20 Å), flexible arm that can reach into buried catalytic centers and carrying the reaction intermediate from one center to another (16).

The two-step conversion of 10-fTHF to tetrahydrofolate and CO_2 catalyzed by ALDH1L1 requires a concerted action of three enzyme domains (Fig. 1) (1). In the first catalytic step, which takes place in the N-terminal formyltransferase domain, the formyl group is

transferred to the 4-PP arm of the intermediate ACP domain (Fig. 1). In the second step, catalyzed by the C-terminal dehydrogenase domain, the formyl group delivered by 4-PP is oxidized to CO_2 through an $NADP^+$-dependent aldehyde dehydrogenase-like mechanism (Fig. 1). *In vitro*, both the full-length enzyme as well as the exogenously expressed formyltransferase domain are capable of the hydrolytic cleavage of 10-fTHF, releasing formate (10,12). The requirement of millimolar concentrations of 2-mercaptoethanol (2-ME) for this hydrolysis (10,12) suggests a mechanism where the formyl group is likely transferred first to the sulfhydryl of 2-ME, which in this case plays the role of the 4-PP sulfhydryl, before being released.

Although the mechanisms of the two catalytic steps were proposed (14,17,18), it remains unclear how the formyl product from the first reaction in the formyltransferase domain is physically transported to the catalytic center of the dehydrogenase domain. An NMR solution structure of the ALDH1L1 ACP domain has been solved (PDB 2CQ8), and several crystal structures of the dehydrogenase domain of the rat enzyme and the formyltransferase domain from several species have been reported (14,18–22). The structure of the full-length protein, however, is still not available. Our attempts to crystallize the full-length ALDH1L1 yielded crystals, which did not diffract beyond 7.8 Å at a synchrotron source. While an interpretable electron density was seen for the dehydrogenase domain, and it was possible to model this domain using its high-resolution crystal structure, we were unable to assign the ACP or formyltransferase domains. The lack of defined electron density for either the formyltransferase or ACP domains is likely the consequence of these domains adopting multiple orientations within the crystal lattice relative to the C-terminal domain. The existence of multiple domain-domain arrangements is consistent with the movement of the ACP domain between the catalytic domains as a mechanism of the formyl group transport. Furthermore, ALDH1L1 is a tetramer and there is no evidence that catalysis or domain rearrangements are synchronized between subunits. In the absence of additional experimental data, we have here applied a molecular modeling approach, using existing structures of the individual ALDH1L1 domains, to predict interactions between the ACP and the two catalytic domains of the enzyme consistent with known catalytic mechanisms.

2. Materials and methods

PDB structures used for modeling are listed in Table 1. We used AutoDock Vina (23) and SwissDock (24) to perform blind ligand docking of 4-PP to the formyltransferase domain (1S3I). We also used FTSite (25,26) to predict likely binding sites on the formyltransferase domain, also with 1S3I as the target coordinate set. Restrained docking calculations were carried out with HADDOCK 2.2 (27,28), using the position of the 2-ME sulfur in 1S3I as the target for the 4-PP sulfur. We used ACP domain coordinates 2CQ8 (modified with a 4-PP attached to S354) and formyltransferase domain coordinates. Initial restraints served to evenly distribute the ACP domain around the binding cleft, while subsequent iterations brought the 4-PP sulfur to the active site.

Docking of 4-PP to the dehydrogenase domain used AutoDock Vina (23) with 2O2Q as the target and an S-formylated 4-PP as the ligand. The grid included Cys 707 and was iteratively refined to 28 × 22 × 32 Å.

To dock the ACP domain to the dehydrogenase domain, we used Haddock 2.2 (27) with the coordinates of 2CQ8 (Ser 354 modified to phosphoserine) and 2O2Q, respectively. The conformationally disordered ACP:dehydrogenase linker (residues 399–406) was built by hand as an extension of the ACP domain. A first round of rigid body docking loosely restrained residue 354 to near the position of the crystallographic sulfate. A second round of flexible docking tightly restrained the phosphorus atom to the position of the sulfur atom and introduced a 3 Å restraint between residue 406 (ACP domain) and residue 407 (dehydrogenase domain) to simulate a peptide bond. In this flexible docking round, the linker residues 398–406 were unconstrained. Calculations were run assuming (1) intramolecular formyl transfer (residue 406 precedes subunit *A* residue 407, where the active site Cys 707 is also in subunit *A*) or (2) intermolecular transfer (residue 406 precedes subunit *C* residue 407, where the active site Cys 707 remains in subunit *A*).

3. Results and discussion

3.1. Defining the position of 4-PP prosthetic group in the formyltransferase catalytic center

The crystal structure of the ALDH1L1 formyltransferase domain (4TT8), solved in the complex with the substrate analog 10-formyl-5,8-dideazafolate (10-fDDF), shows that the folate molecule sits at the end of a long cleft between the N- and C-terminal lobes of this domain (22). The formyl leaving group of the coenzyme is adjacent to catalytic residues H106 (17) and D142 (29) on one side while exposed to nucleophilic attack on the other. Such an arrangement allows easy access to 4-PP, the sulfhydryl group of which abstracts a formyl group from 10-fTHF, converting it to tetrahydrofolate. We first sought to model possible 4-PP conformations using untargeted docking of the 4-PP moiety to the hydrolase domain. However, because the 4-PP binding cleft is comparatively large (Fig. 2), untargeted docking using AutoDock 4.2.6 (30) AutoDock Vina (23), or SwissDock (24) failed to produce convergence.

As a complement to calculating explicit docked conformations, we searched for likely ligand binding sites on the formyltransferase domain using FTSite (25,26). FTSite uses a series of small molecule probes to identify binding pockets based on the clustering of multiple probes to a particular region of the target protein. Applied to the formyltransferase domain, FTSite identified three contiguous ligand-binding sites (Fig. 3). Notably, the first site contains the 10-fTHF position (which was not included in the calculation) and fills the bottom of the binding cleft.

The actual formyltransferase reaction imposes several physical constraints on the positions of the SH and PO_4 moieties of the 4-PP. Thus, (i) the sulfur atom must be proximal to the 10-formyl group, and (ii) because the phosphate is covalently attached to S354 of the ACP domain, it cannot be buried in the ligand cleft. Intriguingly, the *apo* structure of the formyltransferase domain (PDB 1S3I) has a 2-ME molecule crystalized in the active site (18). Overlay of this *apo* structure with the 10-fDDF-loaded structure shows that the 2-ME sulfur atom would be ideally positioned to accept a formyl group (Fig. 2B). Filtering the 4-PP conformations calculated by SwissDock with these spatial restrictions yielded two general clusters: one in which the 4-PP lies entirely along the bottom of the cleft with the phosphate exiting along the floor of pocket defined by S9, Y164, and K229 (Fig. 3A), and

one in which the 4-PP initially lies along the bottom of the cleft but bends upwards and exits near I203, Q204, and K205 (Fig. 3B). In this second cluster, the 4-PP occupies much of the three principle binding sites identified by FTSite. Restrained docking simulations using HADDOCK 2.2 showed that either conformation is compatible with the physical requirements of the sulfur and phosphate positions described above and the restrictions on the relative domain orientations imposed by the 10-residue flexible linker connecting the formyltransferase and ACP domains.

3.2. Comparison of the ALDH1L1 N-terminal domain with other formyltransferases

There is no structure of the ALDH1L1 formyltransferase domain bound to a pantetheinyl group, but there are structures available of several other 10-fTHF N-formyltransferases with bound formyl group acceptors. Remarkably, while the nature of the formyl receiving group varies substantially, the proteins themselves have a high degree of structural conservation with a similar orientation of substrates subjected to formylation.

Campylobacter protein C8J_1081, also known as WlaRD, transfers a formyl group from 10-fTHF to dTDP-3,6-dideoxy-3-amino-D-glucose (dTDP-Qui3N). Structures of WlaRD (31) show an overall formyltransferase fold matching that of the ALDH1L1 formyltransferase domain, but with an altered positioning of the loop connecting the second β-strand to the second α-helix (residues 35–46) in the N-terminal lobe, and of the loop connecting the α-helix to the first β-strand in the C-terminal lobe (residues 217–225). Together, these structural changes block the bottom exit of the binding cleft and direct the dTDP-Qui3N ligand upwards towards R192–L197, residues which are structurally analogous to I203–K205 in ALDH1L1 (Fig. 4A).

The N-terminal domain of the bifunctional *E. coli* protein ArnA formylates UDP-sugars, also using 10-fTHF. The recently solved structure of this domain complexed with UDP-Ara4N also shows the formyltransferase fold (32). As with WlaRD, the β2-α2 loop (H32–G43) in the N-terminal lobe, and the α–β loop in the C-terminal lobe (V224–A231) block the bottom of the binding cleft. Surface access to the binding cleft passes by R200, R201, and T202, which are structurally analogous to I203–K205 in ALDH1L1 (Fig. 4B).

E. coli methionyl-tRNAfMet transformylase transfers a formyl group from 10-fTHF to the N-terminus of the Met residue bound to initiator tRNA (33). While the substrate is in this case a peptido-nucleic acid, rather than a sugar-nucleotide, the formyltransferase fold remains the same, as does the position of the formyl-receiving group as shown in the structure of the enzyme-tRNA complex (34). In comparison to ALDH1L1, the fMet transformylase has a 6-residue, R/K-rich insertion in the β2-α2 loop that makes numerous contacts with the tRNA and which dictates substrate specificity (34). It is notable that the fMet and the terminal tRNA base occupy space equivalent to the binding sites predicted by FTSite for ALDH1L1 (Fig. 4C).

Overall, the substrate-bound structures of the bacterial formyltransferases are highly consistent with the binding pocket delineated by FTSite for ALDH1L1, and particularly, with the second cluster of SwissDock predicted poses for 4-PP, in which the phosphate exits

(and thus the ACP domain docks) through the top of the cleft (Fig. 4D), rather than across helix α6. Verification of the bound conformation of 4-PP awaits experimental determination.

3.3. Modeling 4-PP prosthetic group within the aldehyde dehydrogenase catalytic center

Cys 707 is critical for electron transfer to $NADP^+$, the final step of the ALDH1L1-catalyzed reaction (35). This residue is buried in the core of each C-terminal subunit (~15 Å from the protein surface), and accessible only *via* two channels that approach the sulfhydryl from opposite sides (14,19,20). One channel holds the electron acceptor, $NADP^+$, while the other presumably provides access to the S-formylated 4-PP electron donor. We sought to gain insight into the second part of the ALDH1L1-catalyzed reaction through the application of untargeted docking of 4-PP to the dehydrogenase domain (2O2Q).

Visual inspection of the ALDH1L1 crystal structures 2O2Q ($NADP^+$-bound), 2O2R (NADPH-bound), and 2O2P (*apo*) revealed a crystallographic glycerol molecule at the base of the putative 4-PP channel in each structure. We also identified a bound sulfate ion in 2O2Q and 2O2P stabilized by K520, Q528, N548, and K865. Notably, in many of the low energy ligand conformations calculated by AutoDock Vina, the 4-PP phosphate group occupied the same position as the sulfate ion and the 4-PP S-formyl group occupied the same space as the glycerol (Fig. 5). Collectively, these findings are strongly suggestive that our modeling identified the physiologically relevant conformation.

3.4. Modeling of inter- versus intra-subunit communication

While there is as yet no structure of the ACP-dehydrogenase complex, our hypothesis that the sulfate ion defines the pSer 354 phosphate position substantially restricts conformational flexibility of the ACP domain. We therefore sought to identify possible complexed structures through restrained docking protocols.

We performed semi-rigid restrained molecular docking using HADDOCK 2.2 to dock the ACP domain (2CQ8, Ser 354 modified to phosphoserine) to the dehydrogenase domain (2O2Q, full tetramer). Restraints included the position of the pSer phosphate and the domain orientation restraints imposed by the flexible 9-residue ACP-dehydrogenase linker (residues 398–406, see Methods). Surprisingly, initial geometric calculations suggested that intra- or inter-subunit communication was possible. Computational docking was in agreement with this conclusion, yielding two distinct conformational arrangements depending on whether we consider the ACP-dehydrogenase interaction to be intra- (i.e., the ACP domain of subunit A brings the formyl group to C707 of subunit A) or inter-subunit (the ACP domain of subunit C brings the formyl group to C707 of subunit A, where subunits are labeled as per 2O2Q). In the intramolecular arrangement (Fig. 6A), ACP helix α2 lies across the surface of dehydrogenase subunit A, while ACP helix α3 interacts with the dehydrogenase domains of subunits A and C and the loop connecting helices α1 and α2 makes contacts with dehydrogenase subunits B and C. In the intermolecular arrangement (Fig. 6B), ACP helices α2 and α3 and the α2-α3 connecting loop all interact with dehydrogenase subunit C, while the α1-α2 loop dehydrogenase subunit A. In either structural arrangement, the ACP domain interaction surface is largely the same.

Based solely on these docking calculations, we cannot say whether the ACP- dehydrogenase interaction is intra- or inter-subunit. In the intra-subunit case, the ACP- dehydrogenase linker is maximally extended, but this condition could be altered depending on the rigidity of residues immediately preceding the first β-strand of the dehydrogenase domain (i.e., residues 407–408). Additionally, it is possible that the N-terminal formyltransferase domain associates with the C-terminal dehydrogenase domain in a way that precludes one ACP-dehydrogenase arrangement. Explicit determination of intra- *vs.* inter-subunit catalysis awaits further structure or mutagenesis studies.

3.5. Overall flexibility within the ALDH1L1 subunit

ALDH1L1 catalysis includes multiple distinct steps that require participation of different domains: (i) binding of the folate substrate and the hydrolytic cleavage of the formyl group in the formyltransferase domain; (ii) the transfer of the formyl group to the 4-PP moiety of the ACP domain; (iii) the transfer of the formyl group from the formyltransferase to the dehydrogenase domain; (iv) $NADP^+$-dependent oxidation of the formyl group in the dehydrogenase domain. During the catalysis, the formyltransferase and ACP domains can adopt a continuum of conformations relative to the dehydrogenase domain, depending on corresponding step. We suggest that two most distinct conformations in this spectrum of conformations are closed (where the ACP domain is associated with the dehydrogenase domain and the 4-PP group inserted into the dehydrogenase catalytic center) and fully open (the formyltransferase, ACP, and dehydrogenase domains are dissociated from each other). While the closed conformation is suggested by the necessity to bring the 4-PP S-formyl moiety in proximity with the dehydrogenase catalytic cysteine (14,19,20,35), the nature of the extended conformation is less obvious. Such an extended conformation, however, is suggested not only by the multicenter ALDH1L1 catalytic mechanisms, but also by the necessity for the ACP domain to interact with 4-PP transferase, the only enzyme in humans that adds the 4-PP to acyl carrier proteins (36,37). For example, the structure of the complex between the human ACP with its phosphopantetheinyl transferase (Fig. 7) shows that the interaction between the two proteins takes place on a large interface (38). Such an arrangement suggests that the ALDH1L1 ACP domain likewise requires a large degree of orientational freedom to allow the initial modification of Ser 354 by its corresponding phosphopantetheinyl transferase. In fact, the importance of the intermediate domain flexibility within the protein has been demonstrated experimentally in our previous study: the introduction of a rigid α-helix between the ACP and dehydrogenase domain completely abolished 10-fTHF dehydrogenase activity (39).

4. Conclusion

The nature of ALDH1L1 catalysis suggests the existence of several, likely very different, orientations of the formyltransferase/ACP domains relative to the tetrameric core dehydrogenase domain. Two factors could be major contributors to such conformational diversity. (i) During catalysis, ALDH1L1 must undergo substantial conformational re-arrangements to engage and disengage functional domains (39). (ii) There is no evidence of synchronization of catalytic steps across the tetramer, so positions of the formyltransferase and ACP domains relative to the core dehydrogenase domain are unlikely to be identical

across the four subunits. Such conformational heterogeneity will cause diffuse electron density in crystallographic studies and hinder particle averaging in single-molecule analyses. Previously reported crystal structures of the formyltransferase and dehydrogenase domains of ALDH1L1 provided clues for possible conformations of the 4-PP arm in the two catalytic centers (14,18,19). Based on these data, we modeled the binding of the 4-PP to the formyltransferase and dehydrogenase domains using ligand-protein docking approaches. We have further employed protein-protein docking that was based on the modeled conformations of bound 4-PP. This approach allowed the prediction of interfaces between the ACP domain and the formyltransferase and dehydrogenase domains of the protein. Our models further suggested a high degree of flexibility within the full-length enzyme. Unexpectedly, our modeling also indicated the structural arrangement consistent with the inter-subunit catalysis. Further structural and functional studies should allow the assignment of the catalytic mechanism to either inter- or intra-subunit communication.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This study was supported by the National Institutes of Health grant DK054388 (SAK)

Abbreviations

ALDH	aldehyde dehydrogenase
ACP	acyl carrier protein
4-PP	4′-phosphopantetheine
10-fTHF	10-formyltetrahydrofolate
10-fDDF	10-formyl-5,8-dideazafolate
2-ME	2-mercaptoethanol

References

1. Krupenko SA. FDH: an aldehyde dehydrogenase fusion enzyme in folate metabolism. Chem Biol Interact. 2009; 178:84–93. [PubMed: 18848533]
2. Jackson B, Brocker C, Thompson DC, Black W, Vasiliou K, Nebert DW, Vasiliou V. Update on the aldehyde dehydrogenase gene (ALDH) superfamily. Hum Genomics. 2011; 5:283–303. [PubMed: 21712190]
3. Krupenko SA, Oleinik NV. 10-formyltetrahydrofolate dehydrogenase, one of the major folate enzymes, is down-regulated in tumor tissues and possesses suppressor effects on cancer cells. Cell Growth Differ. 2002; 13:227–236. [PubMed: 12065246]
4. Oleinik NV, Krupenko NI, Priest DG, Krupenko SA. Cancer cells activate p53 in response to 10-formyltetrahydrofolate dehydrogenase expression. Biochem J. 2005; 391:503–511. [PubMed: 16014005]

5. Anguera MC, Field MS, Perry C, Ghandour H, Chiang EP, Selhub J, Shane B, Stover PJ. Regulation of Folate-mediated One-carbon Metabolism by 10-Formyltetrahydrofolate Dehydrogenase. J Biol Chem. 2006; 281:18335–18342. [PubMed: 16627483]
6. Oleinik NV, Krupenko NI, Krupenko SA. Cooperation between JNK1 and JNK2 in activation of p53 apoptotic pathway. Oncogene. 2007; 26:7222–7230. [PubMed: 17525747]
7. Oleinik NV, Krupenko NI, Krupenko SA. ALDH1L1 inhibits cell motility via dephosphorylation of cofilin by PP1 and PP2A. Oncogene. 2010; 29:6233–6244. [PubMed: 20729910]
8. Hoeferlin LA, Oleinik NV, Krupenko NI, Krupenko SA. Activation of p21-Dependent G1/G2 Arrest in the Absence of DNA Damage as an Antiapoptotic Response to Metabolic Stress. Genes Cancer. 2011; 2:889–899. [PubMed: 22593801]
9. Hoeferlin LA, Fekry B, Ogretmen B, Krupenko SA, Krupenko NI. Folate stress induces apoptosis via p53-dependent de novo ceramide synthesis and up-regulation of ceramide synthase 6. J Biol Chem. 2013; 288:12880–12890. [PubMed: 23519469]
10. Krupenko SA, Wagner C, Cook RJ. Recombinant 10-formyltetrahydrofolate dehydrogenase catalyses both dehydrogenase and hydrolase reactions utilizing the synthetic substrate 10-formyl-5,8-dideazafolate. Biochem J. 1995; 306(Pt 3):651–655. [PubMed: 7702556]
11. Krupenko SA, Wagner C, Cook RJ. Expression, purification, and properties of the aldehyde dehydrogenase homologous carboxyl-terminal domain of rat 10-formyltetrahydrofolate dehydrogenase. J Biol Chem. 1997; 272:10266–10272. [PubMed: 9092577]
12. Krupenko SA, Wagner C, Cook RJ. Domain structure of rat 10-formyltetrahydrofolate dehydrogenase. Resolution of the amino-terminal domain as 10-formyltetrahydrofolate hydrolase. J Biol Chem. 1997; 272:10273–10278. [PubMed: 9092578]
13. Donato H, Krupenko NI, Tsybovsky Y, Krupenko SA. 10-formyltetrahydrofolate dehydrogenase requires a 4′-phosphopantetheine prosthetic group for catalysis. J Biol Chem. 2007; 282:34159–34166. [PubMed: 17884809]
14. Tsybovsky Y, Donato H, Krupenko NI, Davies C, Krupenko SA. Crystal structures of the carboxyl terminal domain of rat 10-formyltetrahydrofolate dehydrogenase: implications for the catalytic mechanism of aldehyde dehydrogenases. Biochemistry. 2007; 46:2917–2929. [PubMed: 17302434]
15. Reuland SN, Vlasov AP, Krupenko SA. Modular organization of FDH: Exploring the basis of hydrolase catalysis. Protein Sci. 2006; 15:1076–1084. [PubMed: 16597835]
16. Finzel K, Lee DJ, Burkart MD. Using modern tools to probe the structure-function relationship of fatty acid synthases. Chembiochem. 2015; 16:528–547. [PubMed: 25676190]
17. Krupenko SA, Vlasov AP, Wagner C. On the role of conserved histidine 106 in 10-formyltetrahydrofolate dehydrogenase catalysis: connection between hydrolase and dehydrogenase mechanisms. J Biol Chem. 2001; 276:24030–24037. [PubMed: 11320079]
18. Chumanevich AA, Krupenko SA, Davies C. The crystal structure of the hydrolase domain of 10-formyltetrahydrofolate dehydrogenase: mechanism of hydrolysis and its interplay with the dehydrogenase domain. J Biol Chem. 2004; 279:14355–14364. [PubMed: 14729668]
19. Tsybovsky Y, Krupenko SA. Conserved Catalytic Residues of the ALDH1L1 Aldehyde Dehydrogenase Domain Control Binding and Discharging of the Coenzyme. J Biol Chem. 2011; 286:23357–23367. [PubMed: 21540484]
20. Tsybovsky Y, Malakhau Y, Strickland KC, Krupenko SA. The mechanism of discrimination between oxidized and reduced coenzyme in the aldehyde dehydrogenase domain of Aldh1l1. Chem Biol Interact. 2013; 202:62–69. [PubMed: 23295222]
21. Kursula P, Schuler H, Flodin S, Nilsson-Ehle P, Ogg DJ, Savitsky P, Nordlund P, Stenmark P. Structures of the hydrolase domain of human 10-formyltetrahydrofolate dehydrogenase and its complex with a substrate analogue. Acta Crystallogr D Biol Crystallogr. 2006; 62:1294–1299. [PubMed: 17057331]
22. Lin CC, Chuankhayan P, Chang WN, Kao TT, Guan HH, Fun HK, Nakagawa A, Fu TF, Chen CJ. Structures of the hydrolase domain of zebrafish 10-formyltetrahydrofolate dehydrogenase and its complexes reveal a complete set of key residues for hydrolysis and product inhibition. Acta Crystallogr D Biol Crystallogr. 2015; 71:1006–1021. [PubMed: 25849409]

23. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem. 2010; 31:455–461. [PubMed: 19499576]
24. Grosdidier A, Zoete V, Michielin O. SwissDock, a protein-small molecule docking web service based on EADock DSS. Nucleic Acids Res. 2011; 39:W270–277. [PubMed: 21624888]
25. Brenke R, Kozakov D, Chuang GY, Beglov D, Hall D, Landon MR, Mattos C, Vajda S. Fragment-based identification of druggable 'hot spots' of proteins using Fourier domain correlation techniques. Bioinformatics. 2009; 25:621–627. [PubMed: 19176554]
26. Ngan CH, Hall DR, Zerbe B, Grove LE, Kozakov D, Vajda S. FTSite: high accuracy detection of ligand binding sites on unbound protein structures. Bioinformatics. 2012; 28:286–287. [PubMed: 22113084]
27. Dominguez C, Boelens R, Bonvin AM. HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. J Am Chem Soc. 2003; 125:1731–1737. [PubMed: 12580598]
28. van Zundert GC, Rodrigues JP, Trellet M, Schmitz C, Kastritis PL, Karaca E, Melquiond AS, van Dijk M, de Vries SJ, Bonvin AM. The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. J Mol Biol. 2016; 428:720–725. [PubMed: 26410586]
29. Krupenko SA, Wagner C. Aspartate 142 is involved in both hydrolase and dehydrogenase catalytic centers of 10-formyltetrahydrofolate dehydrogenase. J Biol Chem. 1999; 274:35777–35784. [PubMed: 10585460]
30. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. J Comput Chem. 2009; 30:2785–2791. [PubMed: 19399780]
31. Thoden JB, Goneau MF, Gilbert M, Holden HM. Structure of a sugar N-formyltransferase from Campylobacter jejuni. Biochemistry. 2013; 52:6114–6126. [PubMed: 23898784]
32. Genthe NA, Thoden JB, Holden HM. Structure of the Escherichia coli ArnA N-formyltransferase domain in complex with N(5) -formyltetrahydrofolate and UDP-Ara4N. Protein Sci. 2016; 25:1555–1562. [PubMed: 27171345]
33. Schmitt E, Blanquet S, Mechulam Y. Structure of crystalline Escherichia coli methionyl-tRNA(f)Met formyltransferase: comparison with glycinamide ribonucleotide formyltransferase. Embo J. 1996; 15:4749–4758. [PubMed: 8887566]
34. Schmitt E, Panvert M, Blanquet S, Mechulam Y. Crystal structure of methionyl-tRNAfMet transformylase complexed with the initiator formyl-methionyl-tRNAfMet. Embo J. 1998; 17:6819–6826. [PubMed: 9843487]
35. Krupenko SA, Wagner C, Cook RJ. Cysteine 707 is involved in the dehydrogenase activity site of rat 10-formyltetrahydrofolate dehydrogenase. J Biol Chem. 1995; 270:519–522. [PubMed: 7822273]
36. Strickland KC, Hoeferlin LA, Oleinik NV, Krupenko NI, Krupenko SA. Acyl carrier protein-specific 4′-phosphopantetheinyl transferase activates 10-formyltetrahydrofolate dehydrogenase. J Biol Chem. 2010; 285:1627–1633. [PubMed: 19933275]
37. Strickland KC, Krupenko NI, Dubard ME, Hu CJ, Tsybovsky Y, Krupenko SA. Enzymatic properties of ALDH1L2, a mitochondrial 10-formyltetrahydrofolate dehydrogenase. Chem Biol Interact. 2011; 191:129–136. [PubMed: 21238436]
38. Bunkoczi G, Pasta S, Joshi A, Wu X, Kavanagh KL, Smith S, Oppermann U. Mechanism and substrate recognition of human holo ACP synthase. Chem Biol. 2007; 14:1243–1253. [PubMed: 18022563]
39. Reuland SN, Vlasov AP, Krupenko SA. Disruption of a calmodulin central helix-like region of 10-formyltetrahydrofolate dehydrogenase impairs its dehydrogenase activity by uncoupling the functional domains. J Biol Chem. 2003; 278:22894–22900. [PubMed: 12684508]

Highlights

- We docked the 4′-phosphopantetheine arm in the catalytic centers of ALDH1L1
- Physiologically relevant conformations of the arm were identified
- Modeling indicates high degree of flexibility of ALDH1L1 domains

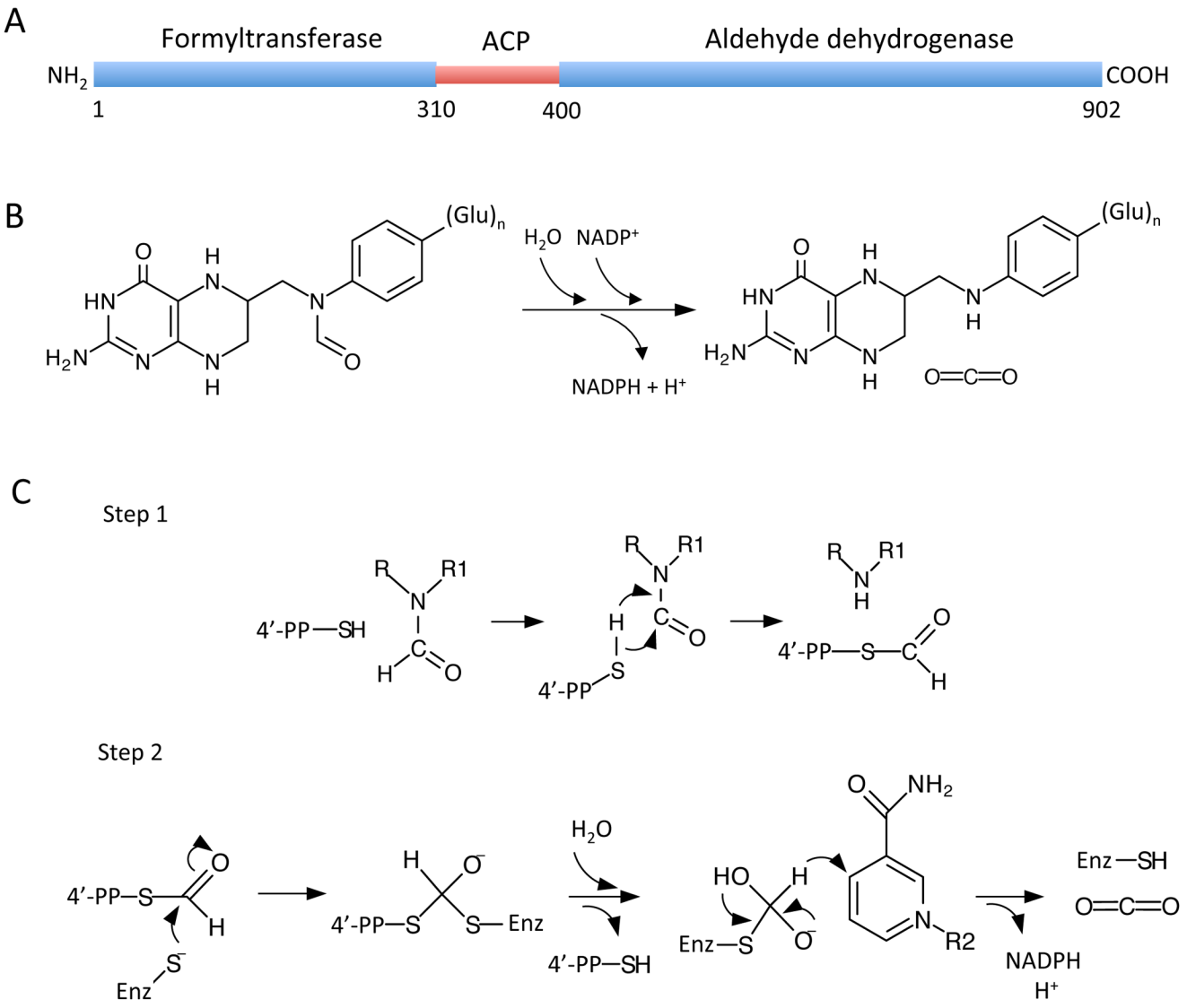


Fig. 1.

Diagram depicting domain organization of ALDH1L1 and the catalyzed reaction. (A) The ALDH1L1 subunit consists of three domains: N-terminal fornyltransferase; intermediate ACP; and C-terminal aldehyde dehydrogenase. Numbers indicate amino acid residues at domain boundaries. (B) The overall 10-fTHF dehydrogenase reaction. (C) Proposed steps of the ALDH1L1 catalysis (step 1, formyltransferase catalysis taking place in the N-terminal domain; step 2, dehydrogenase catalysis taking place in the C-terminal domain; the transfer of formyl group covalently attached to the 4-PP moiety connects the two catalytic steps).

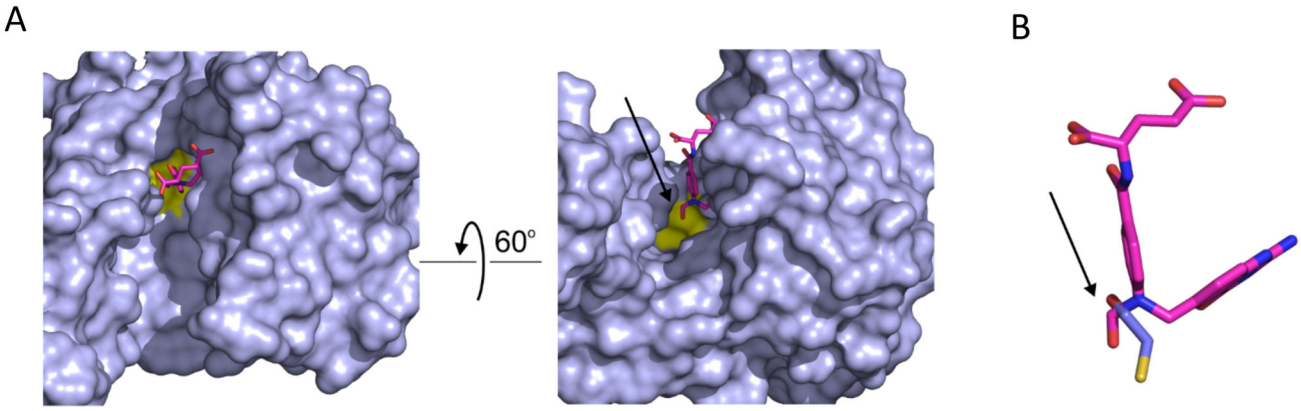


Fig. 2.

(A) The 4-PP binding cleft of the ALDH1L1 formyltransferase domain. Protein is shown in surface representation and substrate analogue 10-fDDF in *stick* representation. Catalytic residues H106 and D142 are highlighted in *yellow*, the transferred formyl group is shown by the *arrow*. In these views, the N-terminal lobe is to the left and the C-terminal lobe to the right. B. Relative positions of 10-fDDF (*magenta*) and 2-ME (*blue*) when corresponding formyltransferase coordinates (4TT8, 1S3I, respectively) are aligned. Notably, a 4-PP sulfur situated at the location of the 2-ME sulfur would be ideally placed to receive the substrate formyl group. Orientation is the same as in panel A, *right*.

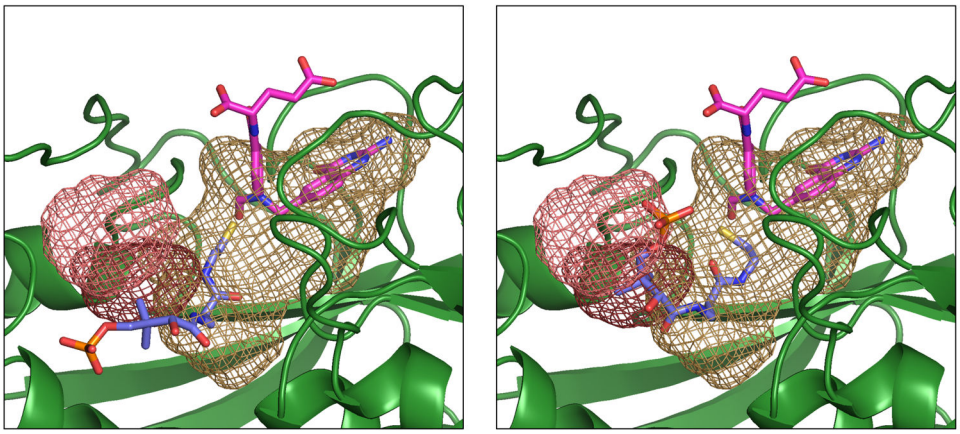


Fig. 3.

Comparison of SwissDock (*stick* representation) and FTSite (*mesh*) substrate binding calculations for 4-PP and the ALDH1L1 formyltransferase domain. (A) 4-PP (*blue sticks*) lies entirely along the floor of the cleft (FTSite binding site 1, *brown mesh*). (B) The N-(2-sufanylethyl)-β-alaninamide lies along the floor of the cleft, but the PO_4 exits through FTSite binding sites 2 (*dark red*) and 3 (*light red*). 10-fDDF is shown as magenta sticks.

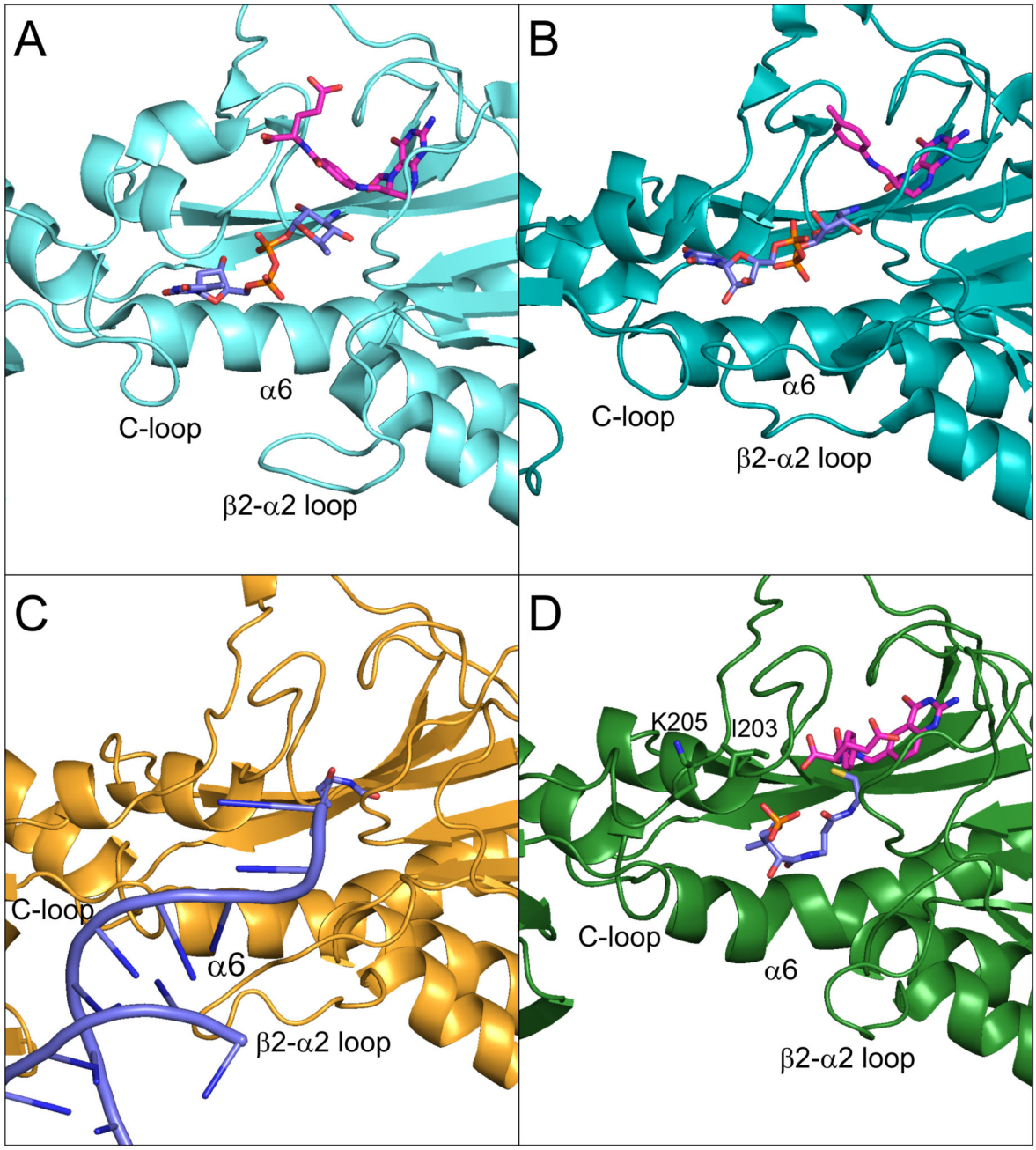


Fig. 4.

Comparison with bacterial formyltransferases suggests a 4-PP binding pose. Substrate (formyl receiver) is shown in *blue stick* representation, formyl donor analog is shown in *magenta stick* representation. In WlaRD (A) and AraN (B), the β2-α2 loop and the C-lobe loop block a potential substrate exit across helix α6, forcing the sugar nucleotide upwards, out of the binding cleft. In the fMet transformylase (C), the β2-α2 loop defines specificity for substrate tRNA, with the tRNA exiting the cleft at the same site as the nucleotides substrates in WlaRD and AraN. Substrate position in all three structures is matched by the

combined FTSite/SwissDock prediction for 4-PP binding to ALDH1L1 (D), with predicted cleft exit adjacent to I203–K205 (*green sticks*).

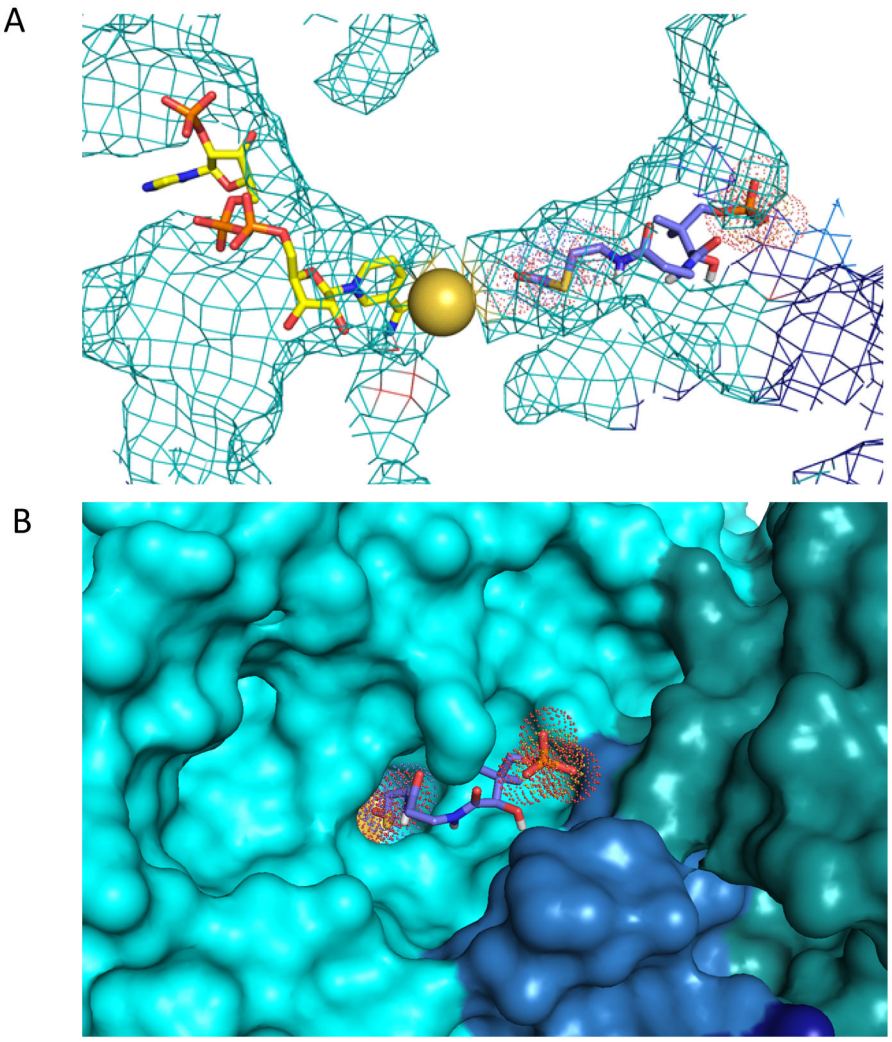


Fig. 5.

Predicted binding conformation of 4-PP to the dehydrogenase domain. (A) The Cys 707 sulfur (*yellow sphere*) is accessed through opposing ligand channels (*wireframe* protein surface). Docking S-formyl-4-PP (*blue sticks*) to 2O2Q yielded a cluster of conformations in which the S-formyl group resided at the base of the ligand channel, proximal to C707 and coincident with a crystallographic glycerol (*dots*), and with the phosphate moiety coincident with a crystallographic sulfate ion (*dots* at right). The electron-receiving $NADP^+$ is shown as *yellow sticks*. (B) Docked pose of 4-PP in the prosthetic group binding channel with the S-formyl group adjacent to the C707 sulfur (*yellow*). Dots show positions of crystallographic glycerol and sulfate. Individual subunits of the tetrameric dehydrogenase domain are shown in different colors.

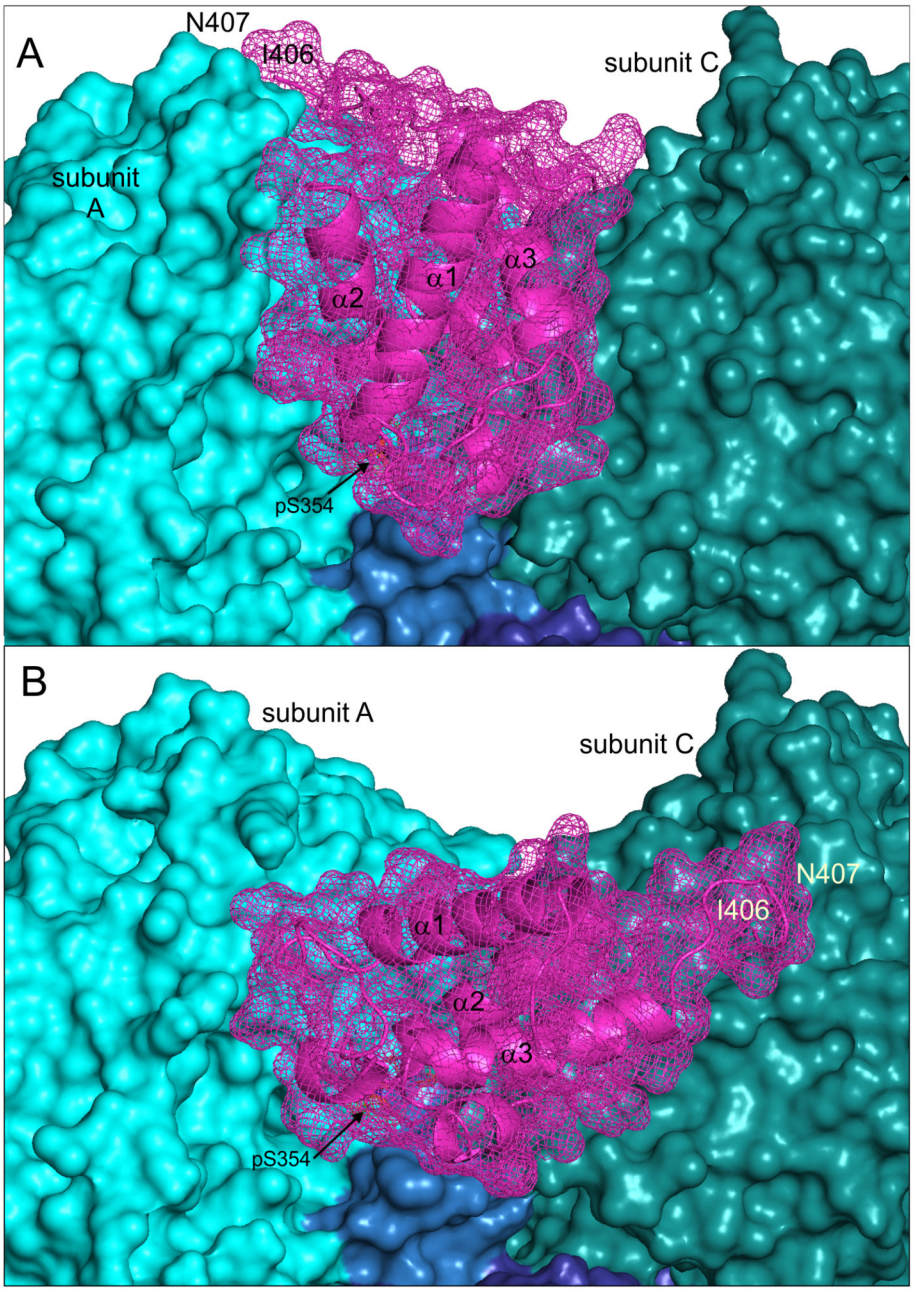


Fig. 6.

Restrained docking is consistent with both intra- (A) or inter- (B) subunit association. The ACP domain is colored *magenta*, subunit A of the dehydrogenase domain is *cyan*, and subunit C of the dehydrogenase domain is *teal*. pSer354 (ACP domain) is labeled, as are the positions of the connecting C- and N- termini of the ACP domain (I406) and dehydrogenase domain (N407), respectively. In either arrangement, the ACP domain rests in a broad pocket on the dehydrogenase tetramer surface with the principal difference being rotation around pS354.

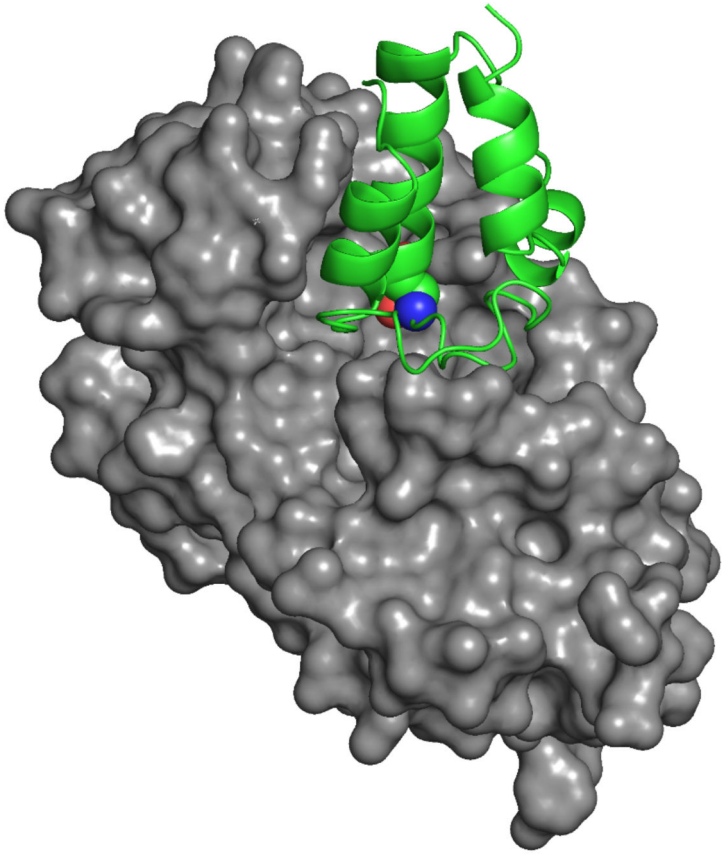


Fig. 7.

The structure (2CG5) of a phosphopantetheinyl transferase (*gray surface*) in complex with an ACP domain (*green ribbon*) shows that initial modification of the ACP domain serine (*spheres*) requires substantial access to the ACP surface. ACP helices 1 and 2 and the connecting loop lie on the surface of the transferase.

Table 1

PDB structures used for modeling.

Structure	PDB	Reference
Formyltransferase domain	1S3I	(18)
Formyltransferase domain with bound substrate	4TT8	(22)
ACP domain	2CQ8	(13)
Dehydrogenase domain with bound $NADP^+$	2O2Q	(14)
Phosphopantetheinyl transferase in complex with ACP domain	2CG5	(38)