Published in final edited form as:

Curr Opin Struct Biol. 2016 December ; 41: 1–9. doi:10.1016/j.sbi.2016.04.003.

Enzymes Required for the Biosynthesis of *N*-formylated Sugars

Hazel M. Holden^{1,*}, James B. Thoden¹, and Michel Gilbert²

¹Department of Biochemistry, University of Wisconsin, Madison, WI 53706

²Human Health Therapeutics, National Research Council Canada, Ottawa, Ontario K1A 0R6, Canada

Abstract

The *N*-formyltransferases, also known as transformylases, play key roles in *de novo* purine biosynthesis where they catalyze the transfer of formyl groups to primary amine acceptors. These enzymes require *N*¹⁰-formyltetrahydrofolate for activity. Due to their biological importance they have been extensively investigated for many years, and they are still serving as targets for antifolate drug design. Most of our understanding of the *N*-formyltransferases has been derived from these previous studies. It is now becoming increasingly apparent, however, that *N*-formylation also occurs on some amino sugars found on the O-antigens of pathogenic bacteria. This review focuses on recent developments in the biochemical and structural characterization of the sugar *N*-formyltransferases.

Graphical Abstract

*To whom correspondence should be addressed: hazel_holden@biochem.wisc.edu, FAX: 608-262-1319, PHONE: 608-262-4988.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Introduction

The lipopolysaccharide or LPS is the major structural component of the outer membrane of Gram-negative bacteria where it has been estimated to occupy ~75% of the total surface area [1]. It is a complex glycoconjugate, which varies from species to species (and within species) in specific content, but in all cases, is thought to provide a permeability barrier to hydrophobic or negatively charged molecules. Conceptually, the LPS can be thought of in terms of three specific regions: the lipid A component, the core oligosaccharide, and the O-specific polysaccharide as highlighted in Figure 1a [2]. It is the O-specific polysaccharide region, which extends farthest away from the bacterium, that displays the most variation from species to species (and between serotypes of the same species), and it is highly immunogenic [3]. This region, also referred to as the O-antigen, consists of repeating units, which typically contain three to five sugars. The O-antigens are thought to play a role in the virulence of a bacterium and also in its ability to evade antibacterial agents [3].

For more than 30 years it has been known that some O-antigens contain quite unusual deoxysugars. Due to the increased sensitivities of such techniques as NMR, however, it is becoming apparent that the O-antigens are far more complicated than originally thought. Recent research has demonstrated, for example, that the O-antigens of some Gram-negative bacteria contain quite remarkable formylated dideoxysugars including 3-formamido-3,6-dideoxy-D-glucose (Qui3NFo), 3-formamido-3,6-dideoxy-D-galactose (Fuc3NFo), 4-formamido-4,6-dideoxy-D-glucose (Qui4NFo), and 4-formyl-D-perosamine as depicted in Figure 1b [4]. These unusual sugars have been found on such organisms as *Brucella abortus* [5], *Salmonella enterica* O60 [6], *Providencia alcalifaciens* O40 [7], *Francisella tularensis* [8], and *Campylobacter jejuni* [9]. Strikingly, all of the above organisms are extremely pathogenic. *B. abortus*, for example, is the causative agent of brucellosis [10]. *S. enterica* is a notorious human pathogen known to be a leading cause of hospitalizations and deaths due to the consumption of contaminated food [11]. *P. alcalifaciens* is an opportunistic organism associated with enteric diseases and was implicated in the 1996 food poisoning outbreak in Fukui, Japan [12]. *F. tularensis* is the causative agent of tularemia or “rabbit fever,” and because it can be produced as a highly infectious aerosol, it is classified as a select agent by the Centers of Disease Control in the United States [13]. Finally, *C. jejuni* is a major cause of gastroenteritis worldwide and, importantly, is now considered a triggering agent for the development of Guillain-Barré syndrome, a devastating acquired autoimmune peripheral neuropathy leading to severe muscle weakness and in some cases paralysis [14].

The genes encoding the enzymes required for the biosynthesis of such formylated sugars are typically located within clusters. The source of the formyl group is the cofactor N^{10} -formyltetrahydrofolate (Figure 1c). On the basis of bioinformatics, a pathway for the synthesis of one of these sugars, Qui4Fo, has been proposed as shown in Figure 2 [15]. Like most pathways for the biosynthesis of unusual sugars, the starting ligand, which in this case is glucose-1-phosphate, is activated by its attachment to a nucleoside monophosphate. The enzyme required for this reaction is a nucleotidyltransferase. The second step involves an oxidation of the C-4' carbon and removal of the C-6' hydroxyl group by an NAD^{+} -dependent 4,6-dehydratase. There is a subsequent amination of the sugar via a pyridoxal 5'-phosphate (PLP) dependent aminotransferase. The final step is the *N*-formylation of the

C-4' amino moiety by an enzyme requiring N^{10} -formyltetrahydrofolate for activity. Although the existence of N -formylated sugars was first reported in 1985 [16], it is only within the last several years that the overall architectures of the sugar N -formyltransferases have been defined in detail. This review focuses on our current understanding of the structures and functions of these intriguing enzymes.

First structure of a sugar N -formyltransferase

As indicated in Figure 1a, the lipid A component of the LPS contains phosphorylated sugars. These sugars, along with other phosphate moieties in the LPS, result in a negatively charged outer surface of the bacterium. As a first line of defense against pathogenic bacteria, host epithelial cells as well as circulating neutrophils and macrophages produce cationic antimicrobial peptides (CAMPS) that interact with the bacterial LPS ultimately resulting in cell death [17]. Strikingly, some human pathogens, such as *Salmonella typhimurium* and *Pseudomonas aeruginosa*, have been shown to alter their LPS composition by the addition of 4-amino-4-deoxy-L-arabinose to the lipid A component. The net result is that the negative charge of the LPS is reduced thereby leading to CAMPS resistance [18].

A key enzyme involved in the production of 4-amino-4-deoxy-L-arabinose (L-Ara4N) is ArnA [19]. It is a bifunctional enzyme with its N-terminal domain functioning as an N -formyltransferase and its C-terminal domain catalyzing an oxidative decarboxylation reaction. Formylation of the sugar is thought to be an obligatory step in the ultimate production of L-Ara4N-modified lipid A [19].

In 2005 the structure of the N-terminal domain of ArnA was reported by two independent research groups [20,21•]. These initial models defined the overall architecture of the N -formyltransferase domain of ArnA and provided details concerning the manner in which UMP and N^5 -formyltetrahydrofolate (shown in Figure 1c) are accommodated within the active site region. In addition, a characteristic signature sequence of HxSLLPKxxG motif was identified with the proline adopting a *cis* peptide conformation and the histidine residue functioning in catalysis [20]. Although informative, these structures did not reveal the manner in which ArnA binds a nucleotide-linked sugar in its N -formyltransferase domain.

Structure of an N -formyltransferase from *C. jejuni*

Structural analyses of the sugar N -formyltransferases took a hiatus following the published papers on ArnA in 2005 until 2013 when the first detailed investigation of the hypothetical protein C8J-1018 from *C. jejuni* 81116 was reported [9••]. This enzyme, hereafter referred to as WlaRD, was shown to catalyze the N -formylation of dTDP-Qui3N or dTDP-Fuc3N using N^{10} -formyltetrahydrofolate as the carbon source. Seven different crystal complexes of WlaRD were determined to 1.9 Å resolution or better for this investigation.

The first structure of WlaRD, with bound dTDP and N^5 -formyltetrahydrofolate, allowed for the molecular architecture of the enzyme to be revealed. Note that N^5 -formyltetrahydrofolate is often used in the study of N -formyltransferases because it is commercially available, and it is stable. It is not, however, catalytically competent. Shown in Figure 3a is a ribbon representation of the WlaRD dimer. The characteristic signature sequence first identified in

ArnA (HxSLLPKxxG) is slightly modified in WlaRD and corresponds to His 96 – Gly 105 (HxSALPKxxG). As in ArnA, the proline in the signature sequence adopts the *cis* conformation. The side chains of Lys 102 and Tyr 103 project towards the subunit:subunit interface. Each subunit of the WlaRD dimer adopts a bilobal type architecture with the N-terminal domain defined by Met 1 to Leu 197 and the C-terminal region formed by Val 198 to Lys 271. The N-terminal domain consists primarily of a six-stranded mixed β -sheet flanked on either side by α -helices whereas the C-terminal region contains a four-stranded antiparallel β -sheet. The active site is housed primarily within the N-terminal domain.

The two structures of WlaRD with bound N^5 -formyltetrahydrofolate and either dTDP-Qui3N or dTDP-Fuc3N revealed that the dTDP-sugar ligands are shifted in the active site by ~ 1 Å with respect to one another. Interestingly, there are no protein side chains involved in binding the pyranosyl moieties of these substrates. Kinetic analyses showed that the catalytic efficiencies of WlaRD for dTDP-Qui3N versus dTDP-Fuc3N were $1.1 \times 10^4 \text{ M}^{-1}\text{s}^{-1}$ and $5.6 \times 10^1 \text{ M}^{-1}\text{s}^{-1}$, respectively. The slightly different binding position of dTDP-Fuc3N versus dTDP-Qui3N in the active site is most likely the reason why WlaRD is catalytically less efficient with dTDP-Fuc3N as its substrate. The actual *in vivo* substrate for WlaRD is still not entirely clear, however. Due to the complexity and heterogeneity of the glycans produced by *C. jejuni* 81116 a complete structural characterization of its lipooligosaccharide by NMR has not yet been possible.

Of particular importance was the structure of WlaRD solved in the presence of N^{10} -formyltetrahydrofolate. This model revealed, for the first time, the manner in which any *N*-formyltransferase (or transformylase) binds the catalytically competent N^{10} -formyltetrahydrofolate cofactor. The structure of WlaRD with bound N^{10} -formyltetrahydrofolate demonstrated that there is a difference in rotation about the bridge that links the bicyclic ring to the phenyl moiety of the cofactor (Figure 3b). Due to this rotation, the C-9 carbons of the cofactors are displaced by ~ 2 Å within the active site region. Understanding this difference may prove important in the future design of antifolate-based therapeutics.

On the basis of the seven structures of WlaRD that were determined, it was possible to propose a model for the Michaelis complex as depicted in Figure 3c. Importantly, Asn 94, His 96, and Asp 132 are strictly conserved amongst the *N*-formyltransferases. Not surprisingly, the site-directed mutant proteins of WlaRD, namely N94A, H96N, and D132N, are completely inactive [9]. It is thought that His 96 in WlaRD serves as the catalytic base to remove a proton from the sugar amino group as it attacks the formyl carbonyl carbon of N^{10} -formyltetrahydrofolate to form a tetrahedral intermediate. In the WlaRD complexes, His 96 appears to be ideally positioned to serve such a role (Figure 3c). As suggested for glycnamide ribonucleotide transformylase, the conserved asparagine possibly functions in the formation of an oxyanion hole to stabilize the negatively charged oxygen of the tetrahedral transition state (schematically shown in Figure 4) [22,23]. Given that His 96 and Asp 132 in WlaRD are situated within ~ 3 Å of each other, it is possible that the proton transferred to His 96 from the sugar amino group is subsequently donated to the carboxylate of Asp 132. This aspartate, in turn, could protonate N10 of the cofactor thereby leading to the collapse of the tetrahedral intermediate. It is also possible that an intervening water

molecule, situated between His 96 and N10 of the cofactor, functions as a proton shuttle (Figure 4). Without the three-dimensional structure of WlaRD solved in the presence of a transition state mimic, however, the roles of the conserved aspartate and asparagine residues are still open to debate.

Structures of the *N*-formyltransferases from *F. tularensis* and *P. alcalifaciens* O30

Unlike WlaRD, the sugar *N*-formyltransferases, WbtJ and VioF from *F. tularensis* and *P. alcalifaciens* O30, respectively, function on dTDP-Qui4N rather than dTDP-Qui3N. The overall three-dimensional structures of WbtJ and VioF were reported in 2014 and 2015, respectively, and their quaternary structures are decidedly different from that observed for WlaRD as can be seen in Figure 5a [8,24•]. As in WlaRD, the N-terminal domains of WbtJ and VioF harbor the active site region. In WbtJ and VioF, however, the subunit:subunit interface is formed primarily by an extended stretch of polypeptide chain, an α -helix, and a β -hairpin motif. A comparison of the active sites for all three enzymes demonstrates that there are few interactions between the amino acid side chains of the respective proteins and the pyranosyl groups of the nucleotide-linked sugar substrates. Rather it appears that the residues lining the pyrophosphoryl binding pockets play pivotal roles in determining whether the enzyme functions on a dTDP-Qui4N or a dTDP-Qui3N substrate.

Structure of QdtF from *P. alcalifaciens* O40

The *P. alcalifaciens* O40 *N*-formyltransferase QdtF, like WlaRD, functions as a dimer and utilizes dTDP-Qui3N as its substrate. Strikingly, however, the overall fold of its subunit was shown to consist of three regions with the N-terminal and middle motifs followed by an ankyrin repeat domain as highlighted in Figure 5b [25••]. Whereas the ankyrin repeat is a common eukaryotic motif involved in protein:protein interactions, reports of its presence in prokaryotic enzymes have been limited [26,27]. The ankyrin repeat is composed of a helix-loop-helix motif of approximately 33 amino acid residues with the α -helices running antiparallel and an additional extended loop at the C-terminus that projects outward by $\sim 90^\circ$. Unexpectedly, in QdtF this ankyrin repeat houses a second binding pocket for dTDP-Qui3N, which is characterized by extensive interactions between the protein and the ligand (Figure 5c). To address the effects of this second binding site on catalysis, a site-directed mutant protein, W305A, was constructed. Kinetic analyses with dTDP-Qui3N demonstrated that the catalytic activity of the W305A variant was reduced by approximately sevenfold, and the structure of the mutant enzyme clearly showed that ligand binding in the ankyrin repeat had been completely abolished. The identity of the ligand that regulates QdtF activity *in vivo* is presently unknown. Regardless of its natural effector molecule, the structure of QdtF revealed for the first time that an ankyrin repeat is capable of binding small molecules. This expansion of the ankyrin repeat repertoire from simple protein:protein interactions to ligand binding and allosteric control is, indeed, intriguing.

Conclusions

At first glance, the diversity of prokaryotic carbohydrate structures observed in the O-antigens seems overwhelming. In fact, only about ten enzymatic reaction types are required for their syntheses (excluding the more exotic sugars). Eight of these reactions include nucleotidylylations, 4,6-dehydrations, 2,3-dehydrations, isomerizations, epimerizations, aminations, oxidations, and reductions. Due to the research efforts of laboratories worldwide the structures and functions of the enzymes required for these reactions have been well characterized both biochemically and structurally.

There are still important reactions in unusual sugar biosynthesis, however, that are just beginning to be unraveled such as the *N*-formylation of amino sugars. Indeed, the recent investigations of the sugar *N*-formyltransferases have provided fascinating results. For the first time, the structure of any *N*-formyltransferase bound to the catalytically competent cofactor *N*¹⁰-formyltetrahydrofolate has been revealed, and this may have important ramifications for the future design of antifolate chemotherapeutic agents. Moreover, the structure of one of the *N*-formyltransferases, QdtF, demonstrated that ankyrin repeats can be involved in small molecule binding, and thus ankyrin repeats can no longer be ascribed to protein:protein interactions alone. Much research remains, however. The structures of the enzymes involved in the biosynthesis of 3-formamido-3,6-dideoxy-D-galactose and 4-formyl-D-perosamine, for example, are presently unknown. There is absolutely no question that additional formylated sugars will be identified in the future, and the roles of these sugars in pathogenicity will be carefully explored as research techniques continue to improve.

Acknowledgments

Research in the Holden laboratory was supported in part by a grant from the National Institutes of Health (GM115921 to H. M. H.).

References and recommended reading

1. Lerouge I, Vanderleyden J. O-antigen structural variation: mechanisms and possible roles in animal/plant-microbe interactions. *FEMS Microbiol Rev.* 2001; 26:17–47.
2. Raetz CR, Whitfield C. Lipopolysaccharide endotoxins. *Annu Rev Biochem.* 2002; 71:635–700. [PubMed: 12045108]
3. Liu B, Knirel YA, Feng L, Perepelov AV, Senchenkova SN, Wang Q, Reeves PR, Wang L. Structure and genetics of *Shigella* O antigens. *FEMS Microbiol Rev.* 2008; 32:627–653. [PubMed: 18422615]
4. Knirel, YA.; Valvano, MA. *Bacterial Lipopolysaccharides*. Knirel, YA.; Valvano, MA., editors. New York: Springer-Verlag; 2011.
5. Wu AM, Mackenzie NE. Structural and immunochemical characterization of the O-haptens of *Brucella abortus* lipopolysaccharides from strains 19 and 2308. *Mol Cell Biochem.* 1987; 75:103–111. [PubMed: 3114617]
6. Perepelov AV, Liu B, Senchenkova SN, Shashkov AS, Feng L, Knirel YA, Wang L. Structure and gene cluster of the O-antigen of *Salmonella enterica* O60 containing 3-formamido-3,6-dideoxy-D-galactose. *Carbohydr Res.* 2010; 345:1632–1634. [PubMed: 20538266]
7. Ovchinnikova OG, Liu B, Guo D, Kocharova NA, Bialczak-Kokot M, Shashkov AS, Feng L, Rozalski A, Wang L, Knirel YA. Structural, serological, and genetic characterization of the O-antigen of *Providencia alcalifaciens* O40. *FEMS Immunol Med Microbiol.* 2012; 66:382–392. [PubMed: 23163869]

8. Zimmer AL, Thoden JB, Holden HM. Three-dimensional structure of a sugar *N*-formyltransferase from *Francisella tularensis*. *Protein Sci.* 2014; 23:273–283. [PubMed: 24347283]
- 9••. Thoden JB, Goneau MF, Gilbert M, Holden HM. Structure of a sugar *N*-formyltransferase from *Campylobacter jejuni*. *Biochemistry.* 2013; 52:6114–6126. Seven different crystal structures were reported in this paper allowing for a detailed description of the active site geometry of a sugar *N*-formyltransferase. Importantly, for the first time the manner in which any *N*-formyltransferase binds the catalytically competent N^{10} -formyltetrahydrofolate cofactor was revealed. [PubMed: 23898784]
10. Galinska EM, Zagorski J. Brucellosis in humans--etiology, diagnostics, clinical forms. *Ann Agric Environ Med.* 2013; 20:233–238. [PubMed: 23772567]
11. Scallan E, Hoekstra RM, Angulo FJ, Tauxe RV, Widdowson MA, Roy SL, Jones JL, Griffin PM. Foodborne illness acquired in the United States--major pathogens. *Emerg Infect Dis.* 2011; 17:7–15. [PubMed: 21192848]
12. Murata T, Iida T, Shiomi Y, Tagomori K, Akeda Y, Yanagihara I, Mushiaki S, Ishiguro F, Honda T. A large outbreak of foodborne infection attributed to *Providencia alcalifaciens*. *J Infect Dis.* 2001; 184:1050–1055. [PubMed: 11574921]
13. Rowe HM, Huntley JF. From the outside-in: the *Francisella tularensis* envelope and virulence. *Front Cell Infect Microbiol.* 2015; 5:94. [PubMed: 26779445]
14. Hughes RA, Cornblath DR. Guillain-Barré syndrome. *Lancet.* 2005; 366:1653–1666. [PubMed: 16271648]
15. Liu B, Chen M, Perepelov AV, Liu J, Ovchinnikova OG, Zhou D, Feng L, Rozalski A, Knirel YA, Wang L. Genetic analysis of the O-antigen of *Providencia alcalifaciens* O30 and biochemical characterization of a formyltransferase involved in the synthesis of a Qui4N derivative. *Glycobiology.* 2012; 22:1236–1244. [PubMed: 22661447]
16. Knirel YA, Vinogradov EV, Shashkov AS, Dmitriev BA, Kochetkov NK, Stanislavsky ES, Mashilova GM. Somatic antigens of *Pseudomonas aeruginosa*. The structure of the O-specific polysaccharide chains of lipopolysaccharides of *P aeruginosa* serogroup O4 (Lanyi) and related serotype O6 (Habs) and immunotype 1 (Fisher). *Eur J Biochem.* 1985; 150:541–550. [PubMed: 2410275]
17. LaRock CN, Nizet V. Cationic antimicrobial peptide resistance mechanisms of streptococcal pathogens. *Biochim Biophys Acta.* 2015; 1848:3047–3054. [PubMed: 25701232]
18. Band VI, Weiss DS. Mechanisms of antimicrobial peptide resistance in Gram-negative bacteria. *Antibiotics (Basel).* 2015; 4:18–41. [PubMed: 25927010]
19. Breazeale SD, Ribeiro AA, McClerren AL, Raetz CR. A formyltransferase required for polymyxin resistance in *Escherichia coli* and the modification of lipid A with 4-amino-4-deoxy-L-arabinose. Identification and function of UDP-4-deoxy-4-formamido-L-arabinose. *J Biol Chem.* 2005; 280:14154–14167. [PubMed: 15695810]
- 20•. Gatzeva-Topalova PZ, May AP, Sousa MC. Crystal structure and mechanism of the *Escherichia coli* ArnA (PmrI) transformylase domain. An enzyme for lipid A modification with 4-amino-4-deoxy-L-arabinose and polymyxin resistance. *Biochemistry.* 2005; 44:5328–5338. This paper, as well as reference 21, reported for the first time the structure of an *N*-formyltransferase that functions on nucleotide-linked sugars. [PubMed: 15807526]
21. Williams GJ, Breazeale SD, Raetz CR, Naismith JH. Structure and function of both domains of ArnA, a dual function decarboxylase and a formyltransferase, involved in 4-amino-4-deoxy-L-arabinose biosynthesis. *J Biol Chem.* 2005; 280:23000–23008. [PubMed: 15809294]
22. Smith GK, Mueller WT, Sliker LJ, DeBrosse CW, Benkovic SJ. Direct transfer of one-carbon units in the transformylations of *de novo* purine biosynthesis. *Biochemistry.* 1982; 21:2870–2874. [PubMed: 7104299]
23. Almasy RJ, Janson CA, Kan CC, Hostomska Z. Structures of apo and complexed *Escherichia coli* glycylamide ribonucleotide transformylase. *Proc Natl Acad Sci U S A.* 1992; 89:6114–6118. [PubMed: 1631098]
- 24•. Genthe NA, Thoden JB, Benning MM, Holden HM. Molecular structure of an *N*-formyltransferase from *Providencia alcalifaciens* O30. *Protein Sci.* 2015; 24:976–986. This paper

demonstrated that the quaternary structure for a sugar *N*-formyltransferase differs depending upon the identity of its nucleotide-linked substrate. [PubMed: 25752909]

- 25•. Woodford CR, Thoden JB, Holden HM. New role for the ankyrin repeat revealed by a study of the *N*-formyltransferase from *Providencia alcalifaciens*. *Biochemistry*. 2015; 54:631–638. This paper revealed an *N*-formyltransferase containing an ankyrin repeat domain. Importantly, the repeat domain harbored a binding site for a nucleotide-linked sugar. This structural analysis expanded the biological role of the ankyrin repeat motif from simple protein:protein interactions to small molecule binding. [PubMed: 25574689]
26. Voth DE. ThANKs for the repeat: Intracellular pathogens exploit a common eukaryotic domain. *Cell Logist*. 2011; 1:128–132. [PubMed: 22279611]
27. Jernigan KK, Bordenstein SR. Tandem-repeat protein domains across the tree of life. *PeerJ*. 2015; 3:e732. [PubMed: 25653910]

Highlights

The recent structural investigations on amino sugar N-formyltransferases are reviewed.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 1.

Gram-negative bacteria contain on their outermost surface a complex glycoconjugate referred to as the lipopolysaccharide or LPS. As schematically shown in (a), it is composed of a lipid A molecule, the core polysaccharide, and the O-specific polysaccharide. It is the O-specific polysaccharide or O-antigen that contributes to the wide species variations seen in nature. Quite unusual dideoxysugar sugars are sometimes found in the O-antigens including the formylated sugars depicted in (b). The *N*-formyltransferases that are involved in the biosynthesis of these formylated sugar employ *N*¹⁰-formyltetrahydrofolate as the carbon source (c). In many structural analyses the *N*⁵-formyltetrahydrofolate ligand is used because of its stability, but it is not catalytically competent.

Figure 2.

A possible biosynthetic pathway for the production of dTDP-4-formamido-4,6-dideoxy-D-glucose is shown. Steps 1, 2, 3, and 4 require dTTP, NAD⁺, PLP and L-glutamate, and N¹⁰-formyltetrahydrofolate, respectively. In step 2, the NAD⁺ is transiently reduced to NADH in the first step of the reaction. The hydride from NADH is subsequently transferred to the substrate in a subsequent step.

Figure 3.

The dimer structure of WlaRD is shown in (a) with the dTDP and N^5 -formyltetrahydrofolate ligands shown in a space-filling representation. The amino acids found in the signature sequence are depicted as sticks. The twofold rotational axis of the dimer is nearly perpendicular to the plane of the page. The difference in N^5 -formyltetrahydrofolate (green) versus N^{10} -formyltetrahydrofolate (blue) binding can be seen in (b). A model for the Michaelis complex of WlaRD is presented in (c) with protons added to the sugar amino group for clarity. It is thought that His 96 serves as the catalytic base to remove a proton from the sugar amino group as it attacks the carbonyl carbon of the formyl group.

Figure 4. Possible reaction mechanism for WlaRD. On the basis of the various complexes of WlaRD and site-directed mutagenesis experiments, it appears that His 96 serves as the catalytic base. The role of Asn 94 in the stabilization of the tetrahedral intermediate is not entirely clear, and the source of the proton required for the collapse of the tetrahedral intermediate is unknown.

Figure 5.

A ribbon representation of VioF is shown in (a). The bound ligands, dTDP and N^5 -formyltetrahydrofolate, are displayed as sticks. As can be seen those N -formyltransferases, such as VioF, that function on dTDP-Qui4N rather than dTDP-Qui3N substrates adopt significantly different quaternary structures than that observed for WlaRD. This was an unexpected result. More surprising was the ankyrin repeat domain observed in QdtF, which uses dTDP-Qui3N as a substrate. Indeed, the QdtF subunit contains similar N- and C-terminal domains as does WlaRD, but these domains are followed by three ankyrin repeats as highlighted in purple in (b). Perhaps the most unexpected aspect of the QdtF structure was the presence of a dTDP-Qui3N ligand bound tightly to the ankyrin repeat domain. Potential hydrogen bonds between the ligand and the protein are indicated by the dashed lines. Surprisingly, there are more interactions between the protein and the ligand in the ankyrin repeat than there are between the protein and the substrate in the active site.