Article

Crystal Structure of GenD2, an NAD-Dependent Oxidoreductase Involved in the Biosynthesis of Gentamicin

Natalia Cerrone de Araújo, Priscila dos Santos Bury, Mauricio Temotheo Tavares, Fanglu Huang, Roberto Parise Filho, Peter Leadlay, and Marcio Vinicius Bertacine Dias

ACS Chem. Biol., **Just Accepted Manuscript** • DOI: 10.1021/acschembio.9b00115 • Publication Date (Web): 17 Apr 2019

Crystal Structure of GenD2, an NAD-Dependent Oxidoreductase Involved in the Biosynthesis of Gentamicin

Natalia Cerrone de Araújo[1], Priscila dos Santos Bury[1], Maurício Themoteo Tavares[2], Fanglu Huang[3], Roberto Parise Filho[2], Peter Leadlay[3], Marcio Vinicius Bertacine Dias[1,4*].

[1]Department of Microbiology, Institute of Biomedical Science, University of São Paulo, Avenida Prof. Lineu Prestes 1374, 05508-900, São Paulo, Brasil.

[2]Department of Pharmacy, Faculty of Pharmaceutical Sciences, University of Sao Paulo, Prof. Lineu Prestes Avenue 580, 05508-900, Sao Paulo, Brazil.

[3]Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Cambridge, CB2 1GA, UK.

[4] Department of Chemistry, University of Warwick, Coventry, CV4 7AL, UK

*Corresponding Author: mvbdias@usp.br

Abstract. Gentamicins are clinically relevant aminoglycoside antibiotics produced by several *Micromonospora* species. Gentamicins are highly methylated and functionalized molecules and their biosynthesis include glycosyltransferases, dehydratase/oxidoreductases, aminotransferases, and methyltransferases. The biosynthesis of gentamicin A from gentamicin A2 involves three enzymatic steps that modify the hydroxyl group at position 3″ of the unusual garosamine sugar to provide its substitution for an amino group, followed by an *N*-methylation. The first of these reactions is catalyzed by GenD2, an oxidoreductase from the Gfo/Idh/MocA protein family, which reduces the hydroxyl at C3″ of gentamicin A to produce 3″-dehydro-3″-oxo-gentamicin A2 (DOA2). In this work, we solved the structure of GenD2 in complex with NAD+. Although the structure of GenD2 has a similar fold to other members of Gfo/Idh/MocA family, this enzyme has several new features, including a 3D-domain swapping of two β-strands that are involved in a novel oligomerization interface for this protein family. In addition, the active site of this enzyme also has several specialties which are possibly involved in the substrate specificity, including a number of aromatic residues and a negatively charged region, which is complementary to the polycationic aminoglycoside-substrate. Therefore, docking simulations provided insights into the recognition of gentamicin A2 and into the catalytic mechanism of GenD2. This is the first report describing the structure of an oxidoreductase involved in aminoglycoside biosynthesis and could open perspectives into producing new aminoglycosides derivatives by protein engineering.

Keywords: gentamicins; oxidoreductase; domain swapping; crystal structure.

INTRODUCTION

Aminoglycosides are antibiotics that usually target the bacterial ribosome causing decoding errors and consequently leading to inaccurate mRNA translation and/or premature termination of protein synthesis[1]. Gentamicins are among the most clinically relevant antibiotics of this class to treat severe infections caused by gram-negative bacteria[2]. The commercial form of gentamicin is composed of a mixture of five different compounds (gentamicin C1a, C2b, C2a, C2, and C1) that differ in the level of methylation in their unusual sugar rings[3]. Gentamicins have also been demonstrated to have effects on cancer cells by delaying the cell growth and inducing cell death, indicating that this antibiotic has other pharmacological use[4,5,6]. In spite of their benefits, gentamicins, similar to other aminoglycosides, also have several side effects, including reversible nephrotoxicity and irreversible hearing loss[7]. Gentamicins are highly modified trisaccharides that contains a central 2-deoxystreptamine (2-DOS) ring as an aminocyclitol aglycone (Ring I), which is further decorated by a purpurosamine (ring II) and garosamine (ring III) at positions C4 and C6, respectively[8,9] (Figure 1). Gentamicins are produced by several *Micromonospora* species, including *M. echinospora*[10], and their biosynthesis have been extensively studied[3,8,10-11], although the last steps for the gentamicin complex C are not yet completely understood[12]. A minimal set of genes responsible for the biosynthesis of gentamicin A, including the biosynthetic intermediates, was elucidated recently[3]. The biosynthesis of gentamicin A from gentamicin A2 occurs in three steps, which involves the dehydrogenation of the 3″-hydroxyl group on the ring III by GenD2 using nicotinamide adenine dinucleotide (NAD+) as a cofactor, the resulting C3″-keto group is transaminated by GenS2 followed by the methylation at the 3″-amine by GenN, an *N*-methyltransferase[3]. (Figure 1). The functions of these enzymes have been validated *in vivo* and *in vitro*[3]. The crystal structure

of GenN has been solved and its catalytic mechanism was proposed recently[13]. Interestingly, Huang et al.[3] were not able to prove the activity of any of these enzymes individually and the conversion of gentamicin A2 to gentamicin A was only possible when the three enzymes were incubated together in the presence of NAD+, L-glutamate and *S*-adenosyl methionine (SAM). It is possible to speculate that these enzymes might form a protein complex[13] or the activity of GenD2 and GenS2 could be inhibited by low levels of their own products and the activity of GenN could restore the activity of two upstream enzymes[3]. GenD2 is the 3″-oxidoreductase that catalyzes the oxidation of the C3″-hydroxyl group of gentamicin A to produce DOA2, which is the substrate for the transaminase GenS2[3]. GenD2 only shows high sequence identity to Sis12 (97%), TobD2 (56.6%) and KanD2 (48.7%) involved in the biosynthesis of the aminoglycosides sisomycin, tobramycin, and kanamycin, respectively[14]. On the other hand, these four enzymes show modest similarity to other dehydrogenases or oxidoreductases, including several ones that have been structurally characterized. GenD2 belongs to the Gfo/Idh/MocA family of oxidoreductases/dehydratases based on its sequence, in which the proteins have a fold composed of two domains, a Rossman-like or nucleotide binding domain and an oligomerization *C*-terminal domain[15]. However, most of Gfo/Idh/MocA family enzymes are involved in the oxidoreduction of monosaccharides. In contrast, gentamicin A2 is a pseudotrisaccharide and consequently, GenD2 may have a number of specializations for its substrate recognition. In addition, only a few oxidoreductases that act on position C3″-hydroxyl groups from sugars were so far characterized indicating a lacking on the understanding of these enzymes[16,17].

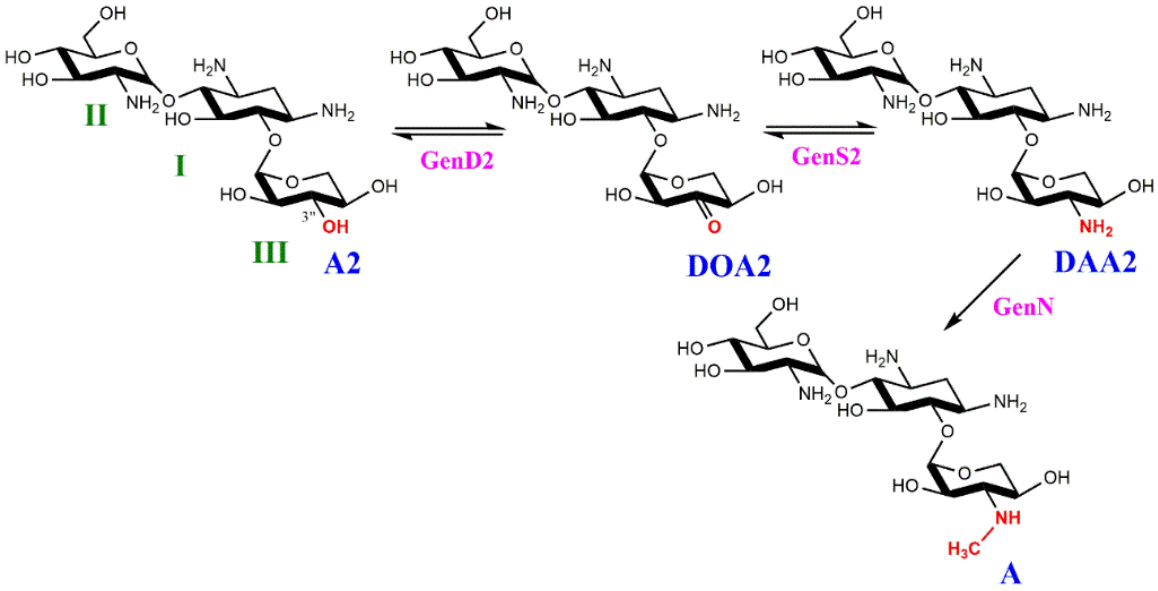


Figure 1. Biosynthesis of gentamicin A from gentamicin A2. The ring I is the 2-DOS (2-deoxystreptamine), II is the purpurosamine and, III is the garosimine. A2 (gentamicin A2); DOA2 (3″-dehydro-3″-oxo-gentamicin A2); DAA2 (3″-dehydro-3″-amino-gentamicin A2); A (gentamicin A). GenD2 is an oxidoreductase; GenS2 is an aminotransferase and, GenN is an *N*-methyltransferase.

In order to understand the structure of GenD2 and provide insights into its substrate specificity and catalytic mechanism, herein we have determined the 3D structure of the C3″-oxidoreductase GenD2 in complex with NAD+ and performed docking simulations. In addition, we have confirmed the preference of GenD2 for NAD+ using Isothermal Titration Calorimetry (ITC) and the stabilization of the GenD2 by different aminoglycoside-like substrates through Differential Scanning Fluorimetry (DSF).

RESUTLS AND DISCUSSION

Overall Structure

GenD2 was successfully expressed in BL21(DE3) cells and the enzyme had about 95% of purity after the gel filtration purification step, in which the enzyme eluted as a single peak (Figure S1, Supporting Information (SI)). Crystals for the enzyme were obtained in

several conditions, although the unique that was reproducible contained a high concentration of sodium formate.

GenD2 structure in complex with NAD+ was solved by molecular replacement using the MrBUMP server[18] at a resolution of 2.4 Å. Attempts to solve the structure using a single model of oxidoreductases from Protein Data Bank were unsuccessful. GenD2 crystals belong to the space group P 31 2 1 and have six protomers in the asymmetric unit, which although the proteins eluted as a tetramer by analytical gel filtration (figure not shown), these molecules do not form the canonical GFOR tetramer in the asymmetric unit. However, this tetrameric quaternary structure could be obtained by applying the symmetrical operations and this has a 222-point symmetry.

All six protomers of the asymmetric unit have almost the entire polypeptide chain (340 residues), with exception of the last alanine, and the RMSD between them is about 0.3 Å, indicating a rigid and restrained structure. Table 1 shows the X-ray data analyses, refinement and stereochemistry quality. The sequence of GenD2 has modest similarity to other oxidoreductase and dehydratase structures, with about 24% and 23% of identity with Glucose-Fructose Oxidoreductase (GFOR) from *Zymomonas mobilis*[19] and with oxidoreductase YvaA from *Bacillus subtilis* (PDB entry 3GFG, data not published), respectively (Figure 2). However, a search using the server DALI rendered the structure of oxidoreductase domain from *Kribbella flavida* (PDB entry 4H3V, data not published) and oxidoreductase WlbA from *Bordetella pertussis*[17], with a Z-score of 37.5 and 36.7 and RMSD of about 2.0 Å.

Table 1. Data Collection and Refinement Statistics

Data Collection	PDB:
PDB entry	6NOR
Wavelength (Å)	0.976260
Resolution range (Å)	49.95 - 2.40 (2.49 - 2.40)
Space group	P 31 2 1
Unit cell (Å)	124.0 124.0 272.6
Total reflections	1215767 (53058)
Unique reflections	95397 (9371)
Multiplicity	12.7 (11.4)
Completeness (%)	99.98 (99.86)
Mean I/sigma(I)	13.48 (3.38)
Wilson B-factor ($Å^2$)	33.93
R-merge	0.17 (1.05)
R-meas	0.19 (1.13)
R-pim	0.05 (0.32)
CC1/2	1 (0.78)
Refinement	
R-work	0.174 (0.224)
R-free	0.223 (0.287)
Number of non-hydrogen atoms	16099
macromolecules	15399
ligands	264
water	436
Protein residues	2049
RMS(bonds) (Å)	0.009
RMS(angles) (°)	1.17
Ramachandran favored (%)	96
Ramachandran outliers (%)	0.49
Clashscore	4.12
Average B-factor ($Å^2$)	33.7
macromolecules	33.7
ligands	34.40
solvent	33.40

All number in parenthesis refer to the high-resolution shell.

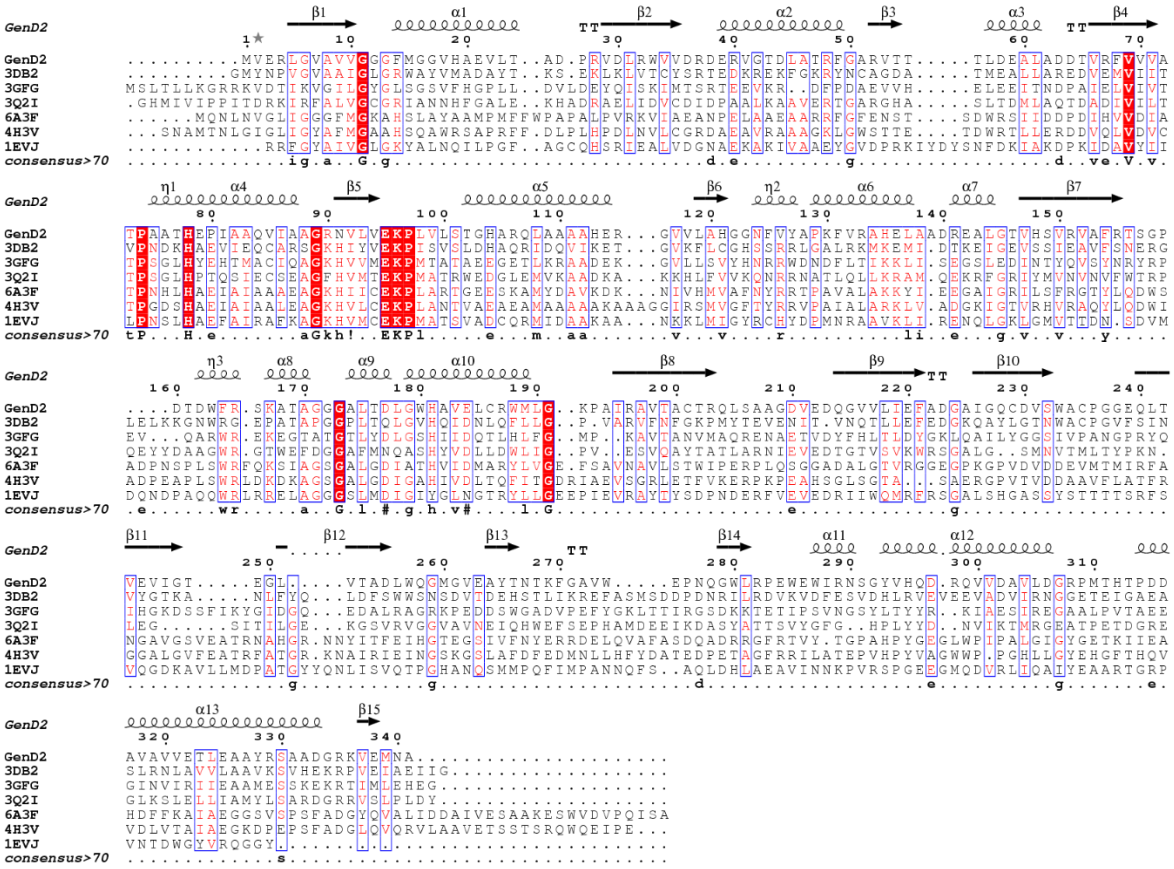


Figure 2. Sequence analysis of GenD2. Alignment of closest homologues of GenD2 rendered by DALI server[20]. Amino acids in red boxes are conserved in all sequences; Amino acids in red and inside blue line boxes are conserved in most of the sequences; Amino acids in black are not significantly conserved. The alignment was prepared using ENDScript server[21].

GenD2 protomers have a general folding similar to other NAD/NADP dependent oxidoreductases as described previously for GFOR and dihydrodipicolinate reductase (DHPR)[19]. GenD2 protomer has two domains, including an *N*-terminal domain, which is a Rossmann fold or the nucleotide binding domain and a *C*-terminal domain characterized to have a six-stranded β-sheet that participates in both active site and in oligomerization and further two α-helices. The *C*-terminal β-sheet is predominantly hydrophobic, and it is involved in the dimerization through of a back to back of a β-sheet from one protomer with the β-sheet from the adjacent protomer forming a two-fold symmetry axis. The α-

helices of the *C*-terminal domain is in another face of the *C*-terminal domain β-sheet and form an α/β sandwich. The *N*-terminal Domain is located adjacently to the α-helices of the *C*-terminal domain (Figure 3A). The superposition of GenD2 to other Gfo/Idh/MocA protein family members indicated that *N*-terminal domain is structurally conserved, while the *C*-terminal domain has more drastic conformational changes (Figure 3B). The most prominent structural difference is in the region between the residues Gly261 to Pro283, which includes the β-strands 7 and 8, that is flipped about 180° in comparison to other oxidoreductase/dehydrase structures[17,19,22]. These two β-strands are involved in a 3D domain swapping since they are involved in the 8-strand β-sheet commonly found in all member of this protein class. At least by our knowledge, this is the first description of a 3D domain swapping into the Gfo/Idh/MocA protein family involving two β-strands to form a quaternary structure. However, the description of swapping of other secondary elements has been reported for other members of this protein class, including for GFOR from *Zymomonas mobilis*[19]. The structure of this enzyme has an *N*-terminal arm which extends to another protomer of the tetramer and a truncated protein missing the first 31 residues rendered only dimers instead of tetramers and has an impact in the binding of co-factor, including the preference for NADP or NAD, indicating a key functional role of this extension[23]. Thus, although 3D domain swapping has been related to have a possible physiological relevance, including allosteric regulation, this role in GenD2 has not been studied in this work. In contrast, in other NAD(P) dependent oxidoreductases, as GFOR and WlbA, this region produces two extra strands of the *C*-terminal β-sheet to form the intradomain eight-stranded β-sheet. In GenD2, this region is also participating in the *C*-terminal domain β-sheet, but in an interdomain mode and consequently, it is involved in the contact with *C*-terminal domain β-sheet of the adjacent protomer to form a continuous β-sheet, with 6 β-strands from one monomer and 2 from the adjacent one. Unlike other

members of Gfo/Idh/MocA protein family, which all strains of *C*-terminal domain form a continuous and extensive 16-stranded β-sheet involving two protomers, the *C*-terminal domain of each GenD2 protomer is not involved to produce a 16-stranded β-sheet and rather than it is observed 4 individual 8-stranded β-sheet in the interface of the tetramer. Due to that, the 8-stranded β-sheets of each protomer of this dimer interface seem to be twisted about 30º in relation to each other in comparison to other proteins from the same family (Figure 3C and S2). The tetramer of GenD2 form a homooligomer of about 72x95x50 Å indicating that the quaternary structure of the protein has an approximately flattened shape (Figure 3D). The main contact for oligomerization is the *C*-terminal β-sheet, which are back to back between two protomers with extensive hydrophobic interaction and between the other two protomers through the 3D-domain swapping, described above and which is involved in the inter 8-stranded β-sheet.

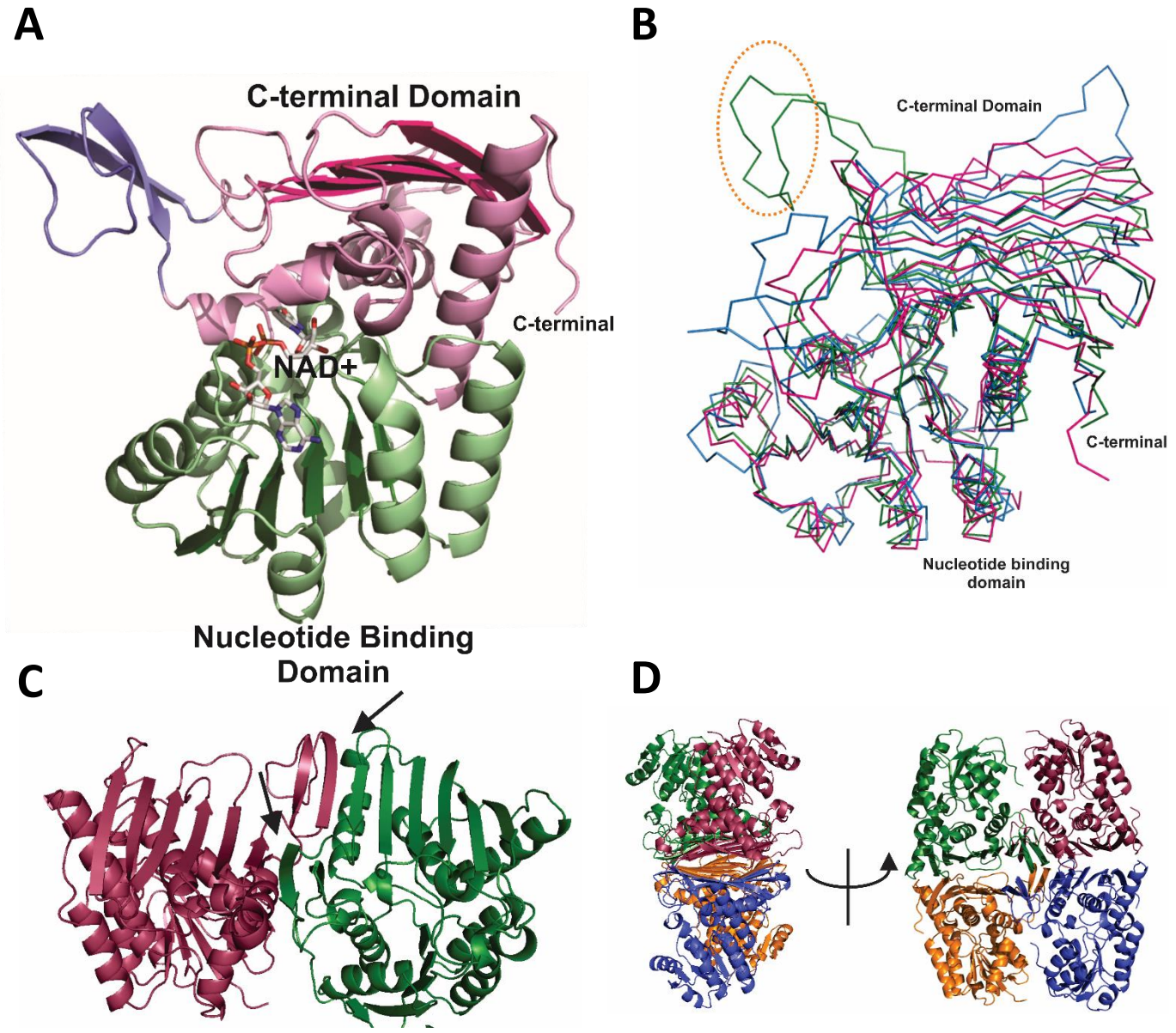


Figure 3. Tertiary and quaternary structure of GenD2. (A) The monomer of GenD2. In Green is represented the Nucleotide Binding Domain; In pink is represented the *C*-terminal domain; and in blue the two β-strand involved in the 3D-domain swapping. (B) The superposition of GenD2 with the two closest structural homologues from the Gfo/Idh/MocA family. In a dashed orange circle is represented the two β-strand involved in the domain swapping. GenD2 structure is in green; 3Q2I PDB entry (WlbA dehydrogenase from *Chromobacterium violaceum*) is in pink, and 4H3V PDB entry (oxidoreductase from *Kribbella flavida*) is in blue. (C) Dimer interface involving the 3D-domain swapping in GenD2. (D) Quaternary structure of GenD2 in two different perspectives with a rotation of 90º.

Analysis of the potential electrostatic surface of GenD2 tetramer did not reveal any evident surface for interaction with other enzymes involved in the modification of C3″ position of ring III of gentamicin (Figure S3, SI). Attempts to copurify GenD2 with GenS2 and GenN or to form the complex using cross-linkage assays failed. Whether the activity of GenD2 is dependent on protein complex formation remains elusive.

Cofactor preference of GenD2

In order to identify the preference of GenD2 to NAD+ or NADP+, we have measured the K_D using Isothermal Titration Calorimetry (ITC). GenD2 was titrated against both co-factors, in which only NAD+ indicated interaction with the protein, while NADP+ did not show any significant heat of interaction (Figure S4, SI)). The thermodynamic parameters obtained in the presence of NAD+ is shown in Table S1 and the obtained K_D was about 7.1±0.1 μM, which is consistent with the *in vitro* enzymatic assays[3] in which GenD2 is a NAD-dependent oxidoreductase.

The crystal structure of GenD2-NAD+ complex was obtained by incubating the protein in the presence of NAD+ during crystallization experiments. As expected, NAD+ binds to GenD2 in the *N*-terminal domain, sitting on the strands 1345 of the Rossmann fold and is highly buried in the protein between the *N*- and *C*-terminal domains (Figure 4A). The adenine-ribose moiety of NAD+ is relatively exposed to the solvent and does not form any hydrogen bond interaction with residues of the protein. The phosphoryl group is hydrogen bonded to the main chain nitrogens of Phe14 and Met15 and nitrogen indole of Trp163. The nicotinamide hydroxyl ribose groups hydrogen interacts with the main chain carbonyl groups of Thr73 and Lys96 and the imidazole side chain of His78. Finally, the nicotinamide group is positioned in the active site and interacts through a hydrogen bond with the side chain of Glu95, Tyr293, Asn124, the main chain carbonyl group of Gly122

and through a water molecule with the side chain of Glu95, Asn124, and Gln296 (Figure 4B).

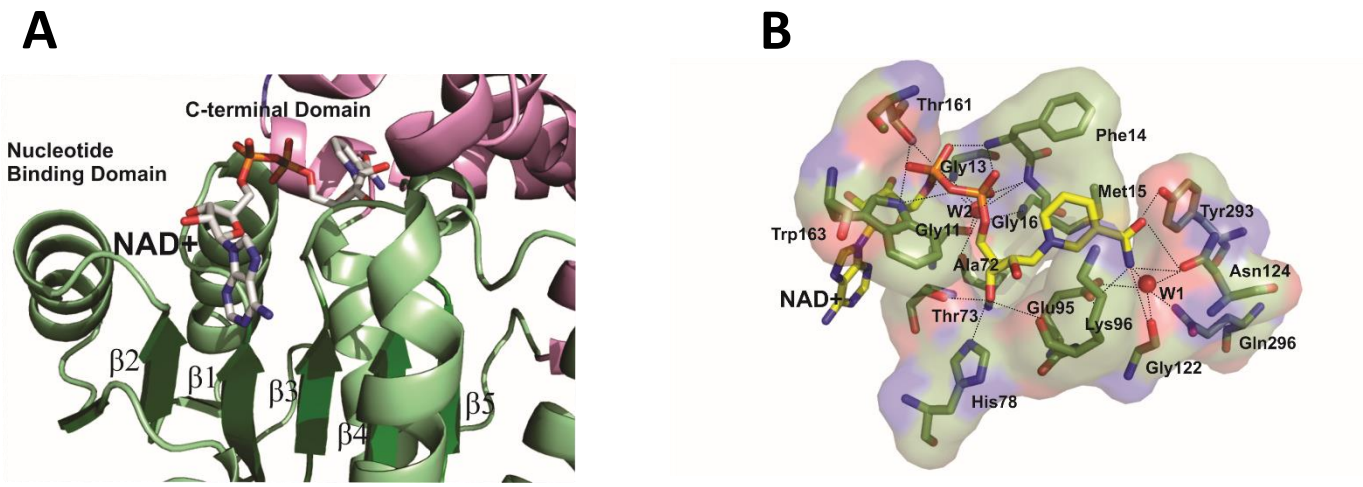


Figure 4. NAD+ binding site on GenD2. (A) The binding site of NAD+ indicating that NAD+ is sitting on the β1-β4 on GenD2 structure. (B) NAD+ binding site indicating the residues that hydrogen interact with NAD+.

Insights into the substrate binding and catalytic mechanism

Our attempts to obtain a crystal structure of GenD2 in complex with gentamicin homologues, kanamycin and tobramycin, through co-crystallization and soaking assays failed to observe any electron density corresponding to these molecules at the active site. However, we have performed a DSF assay using different aminoglycosides and most of the tested molecules promoted an increase in the melting temperature of GenD2 ranging from 4.5 °C for paromamine to 10 °C for gentamicin C, suggesting that interactions might occur between GenD2 and these aminoglycosides (Figure S5 and Table S2, SI). The lower increment of melting temperature observed for paromamine is expected since this compound does not have the ring III and only 2-DOS (ring I) and ring II contribute to the

increase in temperature, indicating a lower affinity of this ligand compared to the other ligands used.

Although various GFO/IDH/MocA family members have been structurally characterized, only a limited number of enzymes that act on C3-hydroxyl groups of sugars have been studied, including WlbA from *Bordetella pertussis* and *Chromobacterium violaceum*, which catalyze the oxidation of UDP-*N*-acetyl-D-glucosaminuronic acid (UDP-GlcNADcA) to UDP-2-acetamido-2-deoxy-3-oxo-D-glucuronic acid (UDP-3-keto-Glc-NAcA)[17]. The catalytic mechanism of the enzymes in this class involves a NAD(P)(H)-mediated hydride ion (H^-) transfer from (dehydrogenases) or to the substrate (reductases)[15]. GenD2 was proved to be a dehydrogenase acting on the gentamicin A2 at C3″ position of ring III to produce DOA2 during the biosynthesis of gentamicin A. However the activity of GenD2 was only observed in the presence of another two enzymes, aminotransferase GenS2 and *N*-methyltransferase GenN[11]. Similar to most of the other members of the Gfo/Idh/MocA protein family, the active site of GenD2 is a solvent-exposed region near the nicotinamide moiety of NAD+ between the *N*- and *C*-domain. In addition, GenD2 has the conserved motif EKP (Figure 2). The Lys96 in the motif interacts with side chains of Asp177 and His181 and could be the catalytic triad as suggested for other GFO/IDH/MocA family members[19,24]. Trp233 might also play a role in coordinating the position of Asp177 *via* the interaction between the nitrogen in the indole side chain to the side chain of Asp177 (Figure 5A). In addition, Lys96, as observed in other members of Gfo/Idh/MocA protein family oxidoreductases, is located in the *N*-terminal domain at end of β6. The superposition of GenD2 holoenzyme with WlbA dehydrogenase holoenzyme from *Chromobactrium violaceum* in complex with UDP-GlcNAcA suggests that the ring III of gentamicin A2 might adopt similar position and interactions as the glucose moiety of UDP-GlcNAcA and the two enzymes could have

similar catalytic mechanisms, although WlbA has an asparagine instead an aspartate in the catalytic triad (Figure 5A). Lys96 could act as a catalytic base, abstracting a proton from C3″-hydroxyl group, and a hydride ion from C3″ oxygen could be transferred to the C4 of the nicotinamide ring of NAD+. In addition, the position of the nicotinamide moiety of NAD+ is also conserved in GenD2 indicating that the ring III of gentamicin A2 may also bind at the *re* face of this moiety (Figure 5B). Residues Asp160 and Glu239 at the active site pocket may also be involved in the substrate recognition through their negatively charged side chains, which are complementary to the highly positively charged aminoglycoside substrate (Figure 5C). Furthermore, Asn124, Thr156, and Tyr293 could potentially form hydrogen bonds with the substrate. Finally, several hydrophobic amino acids, Phe14, Phe125, Phe154, Phe164, Trp257, Trp274, and Trp287 could form π-interaction with the substrate rings for the correct positioning at the catalytic cavity (Figure 5D).

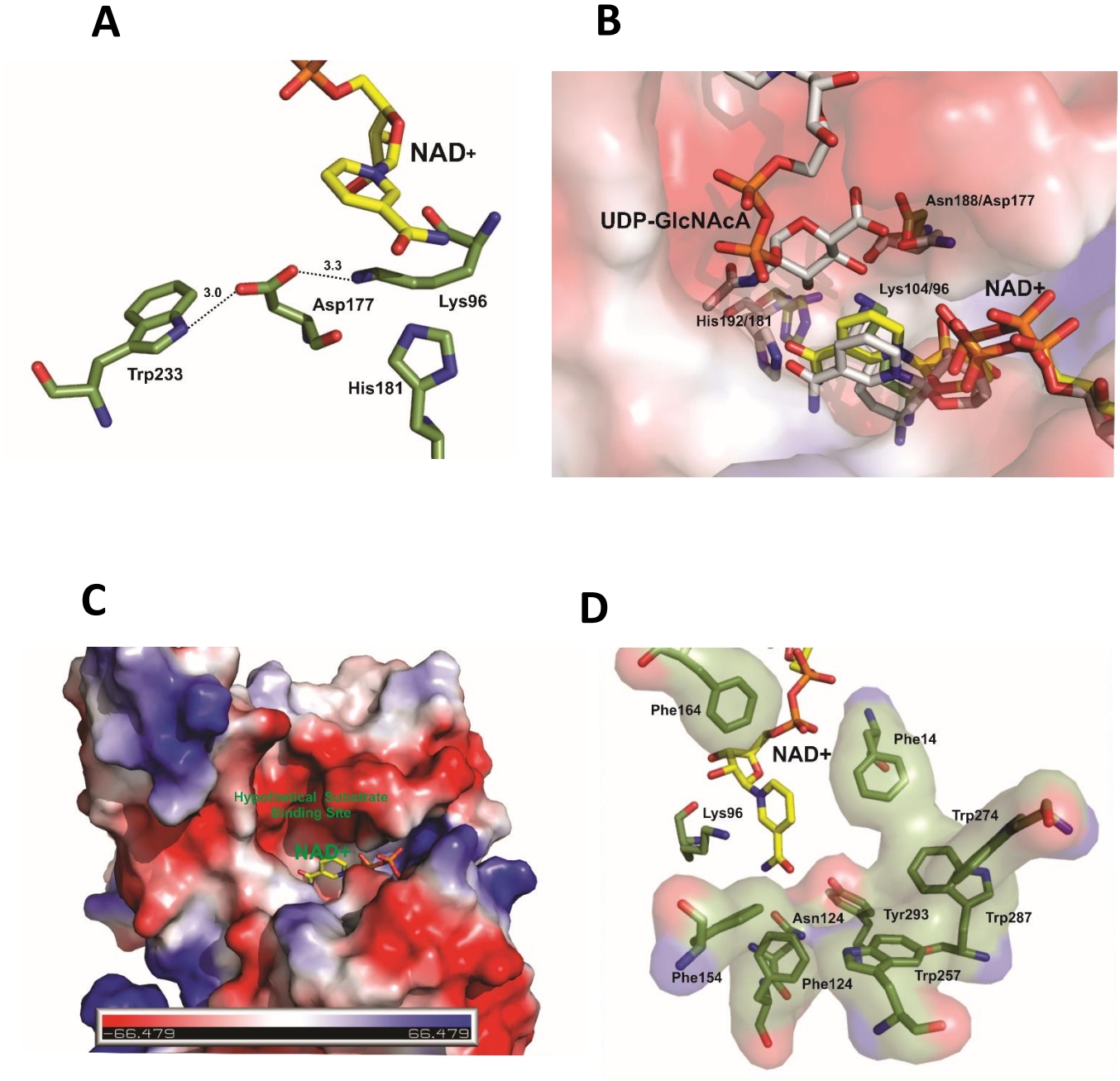


Figure 5. Hypothetical substrate binding site of GenD2. (A) The hypothetical catalytic triad that could be involved in the oxidoreductase reaction. (B) Superposition of GenD2 (carbons in yellow) and WlbA dehydrogenase from *Chromobacterium violaceum* in complex with NAD+ and UDP-GlcNAcA (carbons in white). The electrostatic potential surface is from GenD2. (C) An electrostatic potential surface of GenD2 indicating the highly negative charged hypothetical substrate binding site. (D) Hydrophobic residues that could be involved in the substrate orientation on the GenD2 active site.

Moreover, docking simulations using gentamicin A2 as a ligand corroborated the hypothetical catalytic mechanism of GenD2. The best-ranked pose of gentamicin A2

(Figure 6A) indicated that ring I interacts with GenD2 through two hydrogen bonds, which includes the side chains of Glu239 (not shown) and Tyr293. Ring II is stabilized by a hydrogen bond with Asp160 and further π-interactions with Trp274, Trp287, and Phe14 (not shown). Interestingly, the ring III of gentamicin A2 forms three hydrogen bonds with all amino acids of the proposed catalytic triad including Lys96, Asp177 and His181, which supports the involvement of these amino acids in the catalysis. The superposition of all docked poses indicates a well-conserved disposition of both ring I and III of gentamicin A2 into the narrow cavity of GenD2, while ring II, which is located in the solvent-exposed region, was able to adopt different conformations due to the entropic profile of the area (Figure 6B). Noteworthy, the C4 of the nicotinamide ring of NAD+ and the C3″-hydroxyl group of gentamicin A2 are faced each other inside the cavity (Figure 6C) besides are being surrounded by the proposed catalytic triad (Figure 6D). However, a structure of GenD2 holoenzyme in complex with the substrate or close analogs is required to further validate the proposed binding mode of gentamicin A2 into the GenD2 active site.

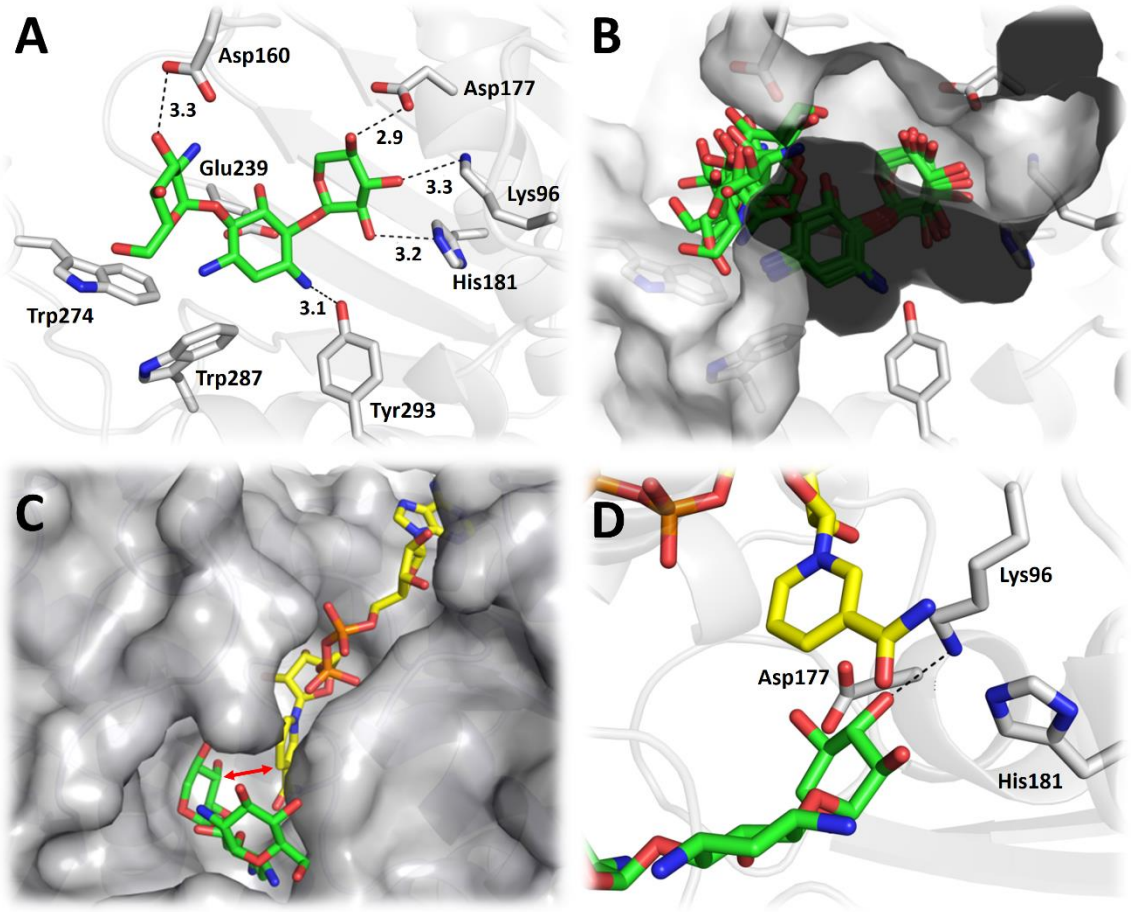


Figure 6. GenD2 docking findings. (A) The best-ranked pose of gentamicin A2 (carbons in green). (B) Superposed poses of the gentamicin A2 at the GenD2 binding cavity. (C) Merged structures of NAD+ (carbons in yellow) and gentamicin A2 (carbons in green) through the binding cavity of GenD2. The double-headed red arrow indicates the C3″ of ring III of gentamicin and the C4 of the nicotinamide ring of NAD+. (D) Docked ring III from gentamicin A2 surrounded by the suggested catalytic triad of GenD2 and the nicotinamide moiety from NAD+. Cartoon protein depicted in white. Carbons of GenD2 in white. Oxygen in red. Nitrogen in blue. Phosphorus in orange. Hydrogen bonds indicated as black dashed lines. Interatomic distances in angstroms (Å).

CONCLUSION

The substitution of the hydroxyl to amino groups on the sugar molecules is performed by two enzymatic steps, a dehydrogenase/oxidoreductase that converts the hydroxyl to a ketone group and this is further acted on by an aminotransferase that replaces the oxygen to an amino group. Several sets of dehydrogenation/transamination reactions are identified to act on different hydroxyl groups of aminoglycosides, but the modification at

position C3″ of ring III occurs only in gentamicins, sisomicin, tobramycin, and kanamycin and the biosynthetic gene clusters of these antibiotics all have the specific and conserved genes that act on this position. To investigate the structural basis of the function of GenD2, we have obtained its crystal structure in complex with NAD+ at 2.4 Å resolution. GenD2 belongs to the Gfo/Idh/MocA protein family of oxidoreductases and has the conserved catalytic residues proposed for deprotonation of the substrate to facilitate hydride abstraction from the substrate to the cofactor NAD+. ITC analysis has further confirmed the interaction between NAD+ and GenD2. Interestingly, GenD2 is the first member of Gfo/Idh/MocA protein family to have a 3D-domain swapping involving two β-strand of the *C*-terminal domain. Several features of GenD2 structure have implication on the substrate recognition including a very negatively charged active site for the binding of the alkaline substrate and several hydrophobic residues that could perform π-interaction with the unusual sugar rings to further stabilize the binding. *In silico* studies also suggested that Lys96, Asp177 and His181 could be involved in the C3″ oxidation of ring III of gentamicin A2 by GenD2. The characterization of the GenD2 structure could provide useful information for increasing the promiscuity of this enzyme in generating novel valuable aminoglycosides through both *in vivo* and *in vitro* approaches.

METHODS

Expression and Purification of Recombinant Proteins

The *genD2* gene from *Micromonospora echinospora* ATCC15835 was cloned into the plasmid pET28a (+) as described by Huang et al.[3]. Transformation of plasmids was performed by heat shock in *E. coli* BL21(DE3) competent cells. *E. coli* BL21(DE3) cells

carrying the recombinant plasmids were cultivated in LB (Luria Bertani) medium supplemented with kanamycin (50 µg/mL) at 37 °C until the cell density reached 0.6 at 600 nm. Overexpression of the protein was induced by β-D-1-thiogalactopyranoside (IPTG) (0.2 mM) at 18 °C with shaking at 200 rpm overnight. Cells were harvested by centrifugation (6000 x g, 4 °C, 30 min) and resuspended 1:1 in buffer (50 mM Tris-HCl, 100 mM NaCl, pH 7.7) with the addition of DNAse and lysozyme at a concentration of 50 µg/mL and phenylmethanesulfonyl fluoride (PMSF) (5 µM). Cell lysis was performed by sonication and the lysate was clarified by centrifugation (17000 x g, 4 °C, 1 h). The recombinant protein was purified by affinity chromatography using a 5 mL IMAC column (GE Healthcare), Ni^{2+} ion-charged His-Bind metal chelating resin (Novagen). GenD2 was eluted using a linear gradient of imidazole (up to 500 mM) on an Åkta Purifier system (GE Healthcare). To increase the purity, a gel filtration purification step was further performed using a HiLoad 16/60 Superdex 200 column (GE Healthcare) equilibrated with 50 mM Tris-HCl, 100 mM NaCl, pH 7.7. Fractions were collected and concentrated up to 15-20 mg/mL and the enzyme was stored at -80 °C. The purity of GenD2 was confirmed by SDS-PAGE 12% stained using Coomassie Brilliant Blue.

Protein crystallization

GenD2 crystallization trials were performed using the sitting-drop method using an Oryx 4 robot (Douglas Instruments) in 96 well plates (MRC2 Swissci). The drops had a total volume of 0.9 µL, 50% of protein solution at a concentration of 10 mg/mL with 5 mM cofactor NAD+ and 50% crystallization condition. After obtaining the first crystals hits, these conditions were optimized manually by hanging drop vapor diffusion method using 24 well VDX plates (Hampton Research). The best and most reproducible crystallization conditions for GenD2 contained 0.1 M sodium acetate pH 4.6, 2 M sodium formate at a

protein concentration of 15 mg/mL and the crystals appeared after approximately seven days in crystallization plates incubated at 30 ºC.

Diffraction and data collection

Data collection of GenD2 crystals was performed at Deutsche Elektronen-Synchrotron (DESy), PETRA III, P13 beamline, in Hamburg, Germany. The data analysis of X-ray diffraction images was performed by XDS[25]. The data was escalated by AIMLESS[26] from the program CCP4 suite[27].

Structure determination

The initial phases for the GenD2 structure were obtained by Molecular Replacement using the online server MrBUMP [18]. The refinement was carried out using Phenix.refine[28] from the Phenix crystallographic suite[29]. The visual inspections were performed using COOT[30] and the stereochemistry quality of the model was accessed using MolProbity[31]. All figures were prepared using PyMOL Molecular Graphics System, Version 1.8 (Schrödinger, LLC).

Differential scanning fluorimetry (DSF)

DSF experiments were performed using a Bio-Rad CFX connect Real-Time System with detection and excitation filters ranging from 450 to 580 nm. For the experiments, 10 μM of GenD2 in presence of 1 mM of NAD+ diluted in a buffer constituted by 50 mM HEPES, 100 mM NaCl, pH 7.0 was incubated with different aminoglycoside-like substrates (kanamycin, geneticin, gentamicin sulfate, apramycin, sisomicin, neomycin, amikacin, paromamine, and meropenem. This last one was used as negative control) at a concentration of 125 μM in the presence of the fluorescent dye SYPRO® Orange 5X (Life

Science). The volumes of 25 μL of protein, each ligand, and fluorescent dye were added to a well of the 96 well plates (Bio-Rad). The protein solutions were scanned from the temperature of 25 ºC to 95 ºC with an increment of 0.5º/minute. T_m values were estimated using Bio-Rad CFX Manager software or using the program Origin and the ΔT_m for each compound was calculated by the difference of protein T_m in the presence of ligand and, protein T_m in the absence of ligand.

Isothermal thermodynamic calorimetry (ITC)

ITC experiments were performed using an iT200 microcalorimetry (Malvern) at 25 ºC. Protein and co-factors (NAD+ and NADP+) were prepared using exactly the same buffer, which was constituted by 50 mM HEPES, 100 mM NaCl, pH 7.0. The protein concentration used in the cell was 150 μM and the ligand concentration in the syringe ranged from 2 to 10 mM. It was performed 18 injections of 10 μL of ligand against to the protein, spaced of 180 s and a stirring speed of 400 rpm. The ITC curves were fitted using nonlinear least-squares regression using the noninteraction one-site model from Origen 5.0.

Molecular docking studies

Gentamicin A2 structure was built and energy-minimized with the density functional theory (DFT), method Becke-3-Lee Yang Parr (B3LYP) with the standard 6-31G* basis set available in the Spartan'14 program (Wavefunction, Inc., Irvine, CA, USA). The experimental starting geometry used for the construction of gentamicin A2 was described previously (PDB entry 5U0N, gentamicin A)[13]. Molecular docking was performed with the GOLD 5.2 program (CCDC). The scoring function used was 'ASP', which is a default function of the GOLD 5.2 program. Hydrogen atoms were added to GenD2 according to

the data inferred by the GOLD 5.2 program on the ionization and tautomeric states. The docking interaction cavity in the protein was established with a 12 Å radius from the phenolic hydroxyl group (OH) of Tyr293. The number of genetic operations (crossover, migration, mutation) in each docking run that was used in the search procedure was set to 100,000. Docking simulations were performed six times. The figure of the best – and most frequent – docking pose for gentamicin A2 was generated by the PyMOL Molecular Graphics System, Version 1.8 (Schrödinger, LLC).

ASSOCIATED CONTENT

Supporting Information. The supporting information is available free of charge on the ACS Publication website at DOI:XXXX. GenD2 purification steps (Supporting Figure 1); GenD2 oligomerization interface (Supporting Figure 2); Electrostatic potential surface of GenD2 tetramer (Supporting Figure 3); Calorimetry parameters for GenD2 against NAD+ and NADP+ (Supporting Table 1); ITC titration of (A) NAD+ and (B) NADP+ against GenD2 (Supporting Figure 4); Differential Scanning Fluorimetry (DSF) for GenD2 in presence of different aminoglycoside-like substrates (Supporting Figure 5); Analysis of DSF for GenD2 in the presence of different aminoglycosides (Supporting Table 2).

ACKNOWLEDGMENT

This work was supported by FAPESP under grant 2010/15971-3, 2014/50324-0 and 2015/09188-8 to M.V.B.D. and fellowship 2014/07843-6 to P.S.B and a CAPES fellowship to N.C.A.; and by a project grant from the Medical Research Council, UK (G1001687) to P.F.L. We also would like to thank R. F. de Souza for helpful discussions.

AUTHOR CONTRIBUTION

N.C.A. performed the expression, purification, crystallization and structure refinement, ITC and DSF assays; P.S.B. determined the structure and performed the refinement. F.H. performed the cloning and determined the best expression conditions. P.F.L and M.V.B.D. conceived the work and M.V.B.D. and F.H. wrote the manuscript. M.T.T. and R.P. performed the docking simulations and reviewed the manuscript.

AUTHOR INFORMATION

Corresponding author: * 55 (0) 11 3091 7420. E-mail:mvbdias@usp.br; ORCID ID: 0000-0002-5312-0191

Notes

The authors declare no competing financial interest.

REFERENCES

1. Borovinskaya, M. A., Pai, R. D., Zhang, W., Schuwirth, B. S., Holton, J. M., Hirokawa, G., Kaji, H., Kaji, A., and Cate, J. H. D. (2007) Structural basis for aminoglycoside inhibition of bacterial ribosome recycling. *Nat. Struct. Mol. Biol. 14*, 727–732.

2. Hayward, R. S., Harding, J., Molloy, R., Land, L., Longcroft-Neal, K., Moore, D., and Ross, J. D. C. (2018) Adverse effects of a single dose of gentamicin in adults: A systematic review. *Br. J. Clin. Pharmacol. 84*, 223–238.

3. Huang, C., Huang, F., Moison, E., Guo, J., Jian, X., Duan, X., Deng, Z., Leadlay, P. F., and Sun, Y. (2015) Delineating the biosynthesis of gentamicin X2, the

common precursor of the gentamicin C antibiotic complex. *Chem. Biol. 22*, 251–261.

4. Codini, M., Cataldi, S., Ambesi-Impiombato, F. S., Lazzarini, A., Floridi, A., Lazzarini, R., Curcio, F., Beccari, T., and Albi, E. (2015) Gentamicin arrests cancer cell growth: The intriguing involvement of nuclear sphingomyelin metabolism. *Int. J. Mol. Sci. 16*, 2307–2319.

5. Cuccarese, M. F., Singh, A., Amiji, M., and O'Doherty, G. A. (2013) A novel use of gentamicin in the ROS-mediated sensitization of NCI-H460 lung cancer cells to various anticancer agents. *ACS Chem. Biol. 8*, 2771–2777.

6. Prokhorova, I., Altman, R. B., Djumagulov, M., Shrestha, J. P., Urzhumtsev, A., Ferguson, A., Chang, C.-W. T., Yusupov, M., Blanchard, S. C., and Yusupova, G. (2017) Aminoglycoside interactions and impacts on the eukaryotic ribosome. *Proc. Natl. Acad. Sci. U. S. A. 114*, E10899–E10908.

7. Fosso, M. Y., Li, Y., and Garneau-Tsodikova, S. (2014) New trends in aminoglycosides use. *Medchemcomm 5*, 1075–1091.

8. Kharel, M. K., Subba, B., Basnet, D. B., Woo, J. S., Lee, H. C., Liou, K., and Sohng, J. K. (2004) A gene cluster for biosynthesis of kanamycin from *Streptomyces kanamyceticus*: Comparison with gentamicin biosynthetic gene cluster. *Arch. Biochem. Biophys. 429*, 204–214.

9. Llewellyn, N. M., and Spencer, J. B. (2006) Biosynthesis of 2-deoxystreptamine-containing aminoglycoside antibiotics. *Nat. Prod. Rep. 23*, 864–874.

10. Unwin, J., Standage, S., Alexander, D., Hosted, T., Horan, A. C., and Wellington, E. M. (2004) Gene cluster in *Micromonospora echinospora* ATCC15835 for the

biosynthesis of the gentamicin C complex. *J. Antibiot. 57*, 436–445.

11. Guo, J., Huang, F., Huang, C., Duan, X., Jian, X., Leeper, F., Deng, Z., Leadlay, P. F., and Sun, Y. (2014) Specificity and promiscuity at the branch point in gentamicin biosynthesis. *Chem. Biol. 21*, 608–618.

12. Li, S., Guo, J., Reva, A., Huang, F., Xiong, B., Liu, Y., Deng, Z., Leadlay, P. F., and Sun, Y. (2018) Methyltransferases of gentamicin biosynthesis. *Proc. Natl. Acad. Sci. 115*, 1340–1345.

13. Bury, P. D. S., Huang, F., Li, S., Sun, Y., Leadlay, P. F., and Dias, M. V. B. (2017) Structural basis of the selectivity of GenN, an aminoglycoside *N*-methyltransferase involved in gentamicin biosynthesis. *ACS Chem. Biol. 12*, 2779–2787.

14. Kudo, F.; and Eguchi, T. (2009) Biosynthetic genes for aminoglycoside antibiotics. *J. Antibiot. 62*, 471–481.

15. Taberman, H., Parkkinen, T., and Rouvinen, J. (2016) Structural and functional features of the NAD(P) dependent Gfo/Idh/MocA protein family oxidoreductases. *Protein Sci. 25*, 778-786.

16. Kubiak, R. L., and Holden, H. M. (2011) Combined Structural and Functional Investigation of a C-3"- ketoreductase involved in the biosynthesis of DTDP-l-digitoxose. *Biochemistry 50*, 5905-5917.

17. Thoden, J. B., and Holden, H. M. (2011) Biochemical and structural characterization of WlbA from *Bordetella pertussis* and *Chromobacterium violaceum* : enzymes required for the biosynthesis of 2,3-diacetamido-2,3-dideoxy- D-mannuronic acid. *Biochemistry 50*, 1483–1491.

18. Keegan, R. M., and Winn, M. D. (2007) MrBUMP: An automated pipeline for

molecular replacement. *Acta Crystallographica Section D 64*, 119–124.

19. Kingston, R. L., Scopes, R. K., and Baker, E. N. (1996) The structure of glucose-fructose oxidoreductase from *Zymomonas mobilis*: an osmoprotective periplasmic enzyme containing non-dissociable NADP. *Structure 4*, 1413–1428.

20. Holm, L., and Rosenström, P. (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res. 38,* W545-549.

21. Robert, X., and Gouet, P. (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res. 42*, W320-324.

22. Sugiura, M., Nakahara, M., Yamada, C., Arakawa, T., Kitaoka, M., and Fushinobu, S. (2018) Identification, functional characterization, and crystal structure determination of bacterial levoglucosan dehydrogenase. *J. Biol. Chem.293*, 17375-17386.

23. Lott, J. S., Halbig, D., Baker, H. M., Hardman, M. J., Sprenger, G. A., and Baker, E. N. (2000) Crystal structure of a truncated mutant of glucose-fructose oxidoreductase shows that an N-terminal arm controls tetramer formation. *J. Mol. Biol. 304*, 575–584.

24. van Straaten, K. E., Zheng, H., Palmer, D. R. J., and Sanders, D. A. R. (2010) Structural investigation of *Myo*-inositol dehydrogenase from *Bacillus Subtilis*: implications for catalytic mechanism and inositol dehydrogenase subfamily classification. *Biochem. J. 432,* 237-247.

25. Kabsch, W. (2010) XDS. *Acta Crystallogr. D Biol. Crystallogr. 66,* 125–132.

26. Evans, P. R., and Murshudov, G. N. (2013) How good are my data and what is the resolution? *Acta Crystallogr. D Biol. Crystallogr. 69*, 1204–1214.

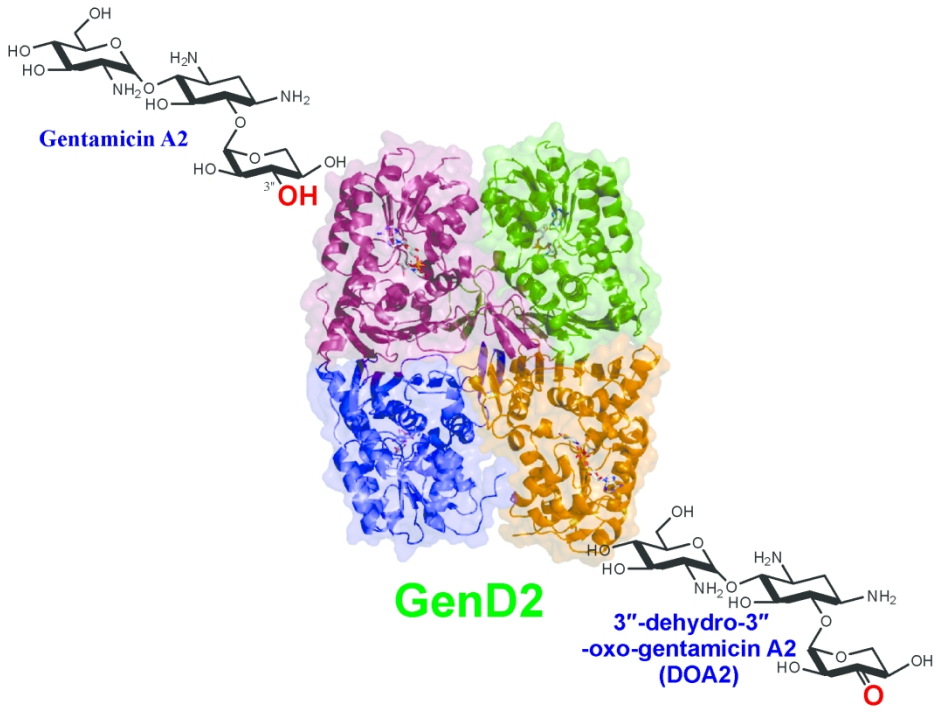
27. Collaborative Computational Project, N. 4. (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.50*, 760–763.

28. Afonine, P. V., Grosse-Kunstleve, R. W., Echols, N., Headd, J. J., Moriarty, N. W., Mustyakimov, M., Terwilliger, T. C., Urzhumtsev, A., Zwart, P. H., and Adams, P. D. (2012) Towards automated crystallographic structure refinement with Phenix.refine. *Acta Crystallogr. D Biol. Crystallogr. 68*, 352–367.

29. Terwilliger, T. C., Adams, P. D., Read, R. J., McCoy, A. J., Moriarty, N. W., Grosse-Kunstleve, R. W., Afonine, P. V., Zwart, P. H., and Hung, L. W. (2009) Decision-making in structure solution using bayesian estimates of map quality: the PHENIX autoSol wizard. *Acta Crystallogr. D Biol. Crystallogr. 65*, 582–601.

30. Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr. 60*, 2126–2132.

31. Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X., Murray, L. W., Arendall, W. B., Snoeyink, J., Richardson, J. S. (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res. 35*, W375-83.


Gentamicin A2
OH
GenD2
3″-dehydro-3″
-oxo-gentamicin A2
(DOA2)
O