Antibiotics

DOI: 10.1002/anie.201409792

Structural Characterization of O- and C-Glycosylating Variants of the Landomycin Glycosyltransferase LanGT2**

Heng Keat Tam, Johannes Härle, Stefan Gerhardt, Jürgen Rohr, Guojun Wang, Jon S. Thorson, Aurélien Bigot, Monika Lutterbeck, Wolfgang Seiche, Bernhard Breit, Andreas Bechthold,* and Oliver Einsle**

Abstract: *The structures of the O-glycosyltransferase LanGT2 and the engineered, C–C bond-forming variant LanGT2S8Ac show how the replacement of a single loop can change the functionality of the enzyme. Crystal structures of the enzymes in complex with a nonhydrolyzable nucleotide-sugar analogue revealed that there is a conformational transition to create the binding sites for the aglycon substrate. This induced-fit transition was explored by molecular docking experiments with various aglycon substrates.*

Glycosylated natural products predominate in antimicrobial and anticancer drug discovery. Often, the biological activity and pharmacological properties of the compounds are determined by attached sugar moieties[1] that also form part of the cellular defense strategy of antibiotic producers.[2] The stereo- and regiospecific glycosylation catalyzed by glycosyltransferases (GT) typically occurs during biosynthesis, as a multistep modification of complex core aglycons that may be secondary metabolites such as polyketides. The permissive nature of glycosyltransferases has allowed the differential O-glycosylation of a range of complex natural products and drugs through in vitro and in vivo strategies.[3] While such approaches have dramatically extended the structural diversity of a range of therapeutically important natural products, they have largely been restricted to O-glycosylation and, to a lesser extent, N-glycosylation.[4] The inclusion of C-glycosyltransferases, which are able to form C–C bonds, into this scheme would thus pave the way for the production of novel bioengineered metabolites.[5]

LanGT2 is an O-glycosyltransferase (O-GT) that catalyzes the initial glycosylation of the aglycon in the biosynthesis of landomycin A, and it shares high homology with the C-glycosyltransferase (C-GT) UrdGT2 from urdamycin A biosynthesis. Both enzymes transfer D-olivose from NDP-D-olivose to closely related angucycline-based aglycons to yield the respective O- or C-glycoside products (Figure 1). LanGT2 catalyzes the transfer of D-olivose to an aglycon C8-OH,

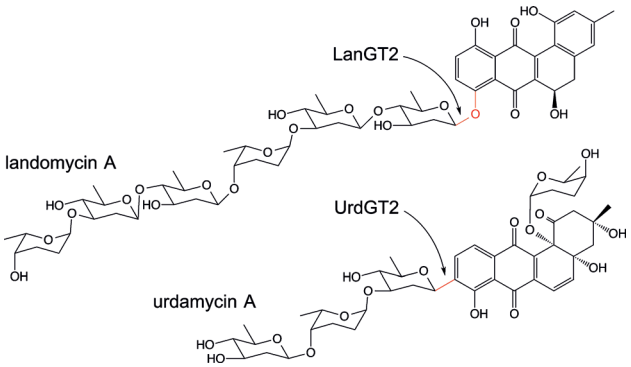


Figure 1. Structures of the angucycline antibiotics landomycin A and urdamycin A. The O-GT LanGT2 attaches D-olivose to the aglycon through an O-glycosidic bond, while the C-GT UrdGT2 links the same carbohydrate moiety through a C–C bond to the aglycon of urdamycin A.

while UrdGT2 attaches the same sugar directly to C9 of the aglycon, with inversion of the anomeric stereocenter in both cases. The aglycon substrate for UrdGT2 is 2-hydro-3-hydroxy-prejadomycin (UWM6),[6] whereas the exact physiological aglycon substrate for LanGT2 is not known. Expression of *urdGT2* in a *lanGT2*-deficient strain of *S. cyanogenus* S136 resulted in the formation of 9-C-olivosyltetrangulol, while the expression of *lanGT2* in the same mutant strain yielded 8-O-olivosyl-D-11-deoxylandomycinone.[7] LanGT2 and UrdGT2 share 53% sequence identity and were presumed to be similar in structure and mechanism, thus inspiring attempts to modify their substrate specificities by

[*] Dr. H. K. Tam, Dr. S. Gerhardt, Prof. Dr. O. Einsle
Institut für Biochemie, Albert-Ludwigs-Universität Freiburg
Albertstrasse 21, 79104 Freiburg (Germany)
E-mail: einsle@biochemie.uni-freiburg.de
Dr. J. Härle, Prof. Dr. A. Bechthold
Institut für Pharmazeutische Wissenschaften
Albert-Ludwigs-Universität Freiburg, 79104 Freiburg (Germany)
E-mail: andreas.bechthold@pharmazie.uni-freiburg.de
Prof. Dr. J. Rohr, G. Wang, Prof. Dr. J. S. Thorson
Center for Pharmaceutical Research and Innovation
University of Kentucky College of Pharmacy, Lexington, KY (USA)
Dr. A. Bigot, M. Lutterbeck, Dr. W. Seiche, Prof. Dr. B. Breit
Institut für Organische Chemie, Albert-Ludwigs-Universität Freiburg
Albertstrasse 21, 79104 Freiburg (Germany)
E-mail: bernhard.breit@chemie.uni-freiburg.de
Prof. Dr. O. Einsle
BIOSS Centre for Biological Signalling Studies
Schänzlestrasse 18, 79104 Freiburg (Germany)

[**] This study was supported in part by Deutsche Forschungsgemeinschaft (GSC-4, Spemann Graduate School of Biology and Medicine, SGBM), NIH (GM 105977-01A1 to J.R. and R37 AI52218 to J.S.T.), and by an EMBO Short Term Fellowship (ASTF 261-2012) to H.K.T. We thank the staff at the Swiss Light Source, Villigen, Switzerland, for assistance during data collection.

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201409792.

swapping relevant amino acid motifs between the enzymes.[8] A prime goal was to generate C-glycosylating variants of LanGT2, and this goal was eventually achieved by grafting residues ^{51}VATTDLPIRHFI^{62} of UrdGT2 into LanGT2.[7b] Furthermore, an S8A variant of the chimeric LanGT2 (LanGT2S8Ac), in which a signature residue (glycine or alanine) of the C-GTs UrdGT2, SimB7, HedL, and SsfS6[9] was introduced, further increased the C-glycosylation efficiency.[7b] Similarly, a recent mutagenesis study focusing on a *Pyrus communis* O-GT (*Pc*OGT) and an *Oryza sativa* C-GT (*Os*CGT) revealed a single mutation of *Pc*OGT, namely D118I, that led to engineered C-GT activity.[10]

The molecular features that distinguish O- and C-specific GTs remain poorly understood. Herein, we report the three-dimensional structures of LanGT2 and LanGTS8Ac in complex with the surrogate sugar nucleotide ligands TDP-carba-D-olivose and thymidine diphosphate. Our data suggest that both O- and C-GTs utilize conserved amino acid residues for general acid–base catalysis. The key feature that differentiates O-glycosylation from C-glycosylation is the specific orientation of the sugar nucleotide with respect to the nucleophile. The structural consequence of the mutations in LanGT2S8Ac is a reorientation of the aglycon to favor C-glycosylation. Our study helps to rationalize the altered specificity of the chimeric enzyme and provides the first structural template for understanding engineered C-GT activity.

Both LanGT2 and the engineered LanGTS8Ac were crystallized and their structures were determined by X-ray diffraction. They show the typical two-domain architecture of their class, with an N-terminal aglycon-binding domain and a C-terminal nucleotide-sugar-binding domain (Figure 2A). The structure of LanGT2S8Ac was highly similar to that of the wild type, with the major difference occurring in the grafted loop from residues 51–62 (Figure 2B). This region is structurally distinct in LanGT2 and UrdGT2,[8a] but within the LanGT2 background of the chimera, it bears a striking resemblance to the original UrdGT2 structure (Figure 2C).[8a] The conservation of the loop structure in this engineered C-glycosylating enzyme thus supports the hypothesis that the positioning of the aglycon is a primary determinant of the resulting mode of glycosylation.

The GT reaction follows a sequential bi–bi mechanism, wherein the nucleotide sugar is bound first, followed by binding of the aglycon and sugar transfer. We therefore tried to generate structures for LanGT2 and LanGT2S8Ac in complex with TDP-olivose in order to prepare the ground for aglycon binding studies. Unfortunately, the TDP-olivose complex proved to be unstable on the timescale of crystallization, thus resulting in structures that invariably only had the product TDP bound (Figure S1 in the Supporting Information). Therefore, we proceeded to synthesize a non-hydrolyzable analogue of the activated sugar substrate by replacing D-olivose by the corresponding carbasugar, carba-α-D-olivose, based on the assumption that the replacement of the ring oxygen by carbon atom would prevent glycosyl transfer (Scheme 1).

As a key step for the synthesis of α-D-carbaolivose thymidine diphosphate (**8**), we applied a one-pot desymmetrizing directed hydroformylation developed recently in our laboratory, followed by an intramolecular carbonyl–ene reaction.[11] After reductive removal of the chiral *o*-DPPF directing group and ozonolysis, a directed carbonyl reduction delivered the carbaolivose backbone with good overall yield and stereoselectivity. Standard functional group manipulations enabled the introduction of the thymidine diphosphate moiety to furnish (−)-**8** in enantiomerically pure form (see the Supporting Information).

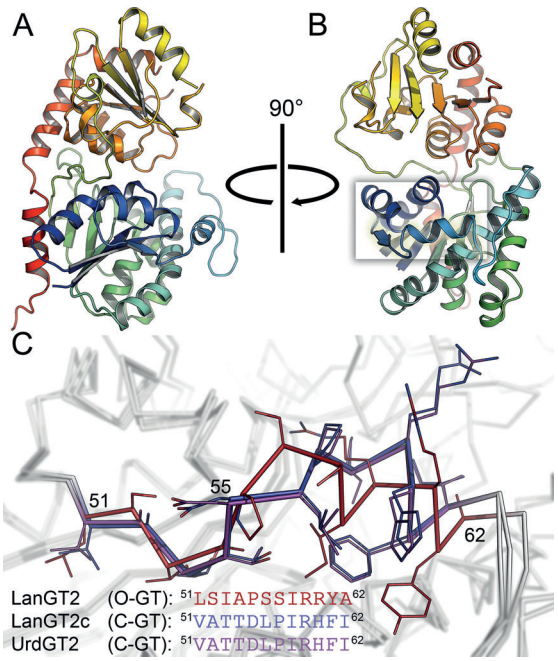
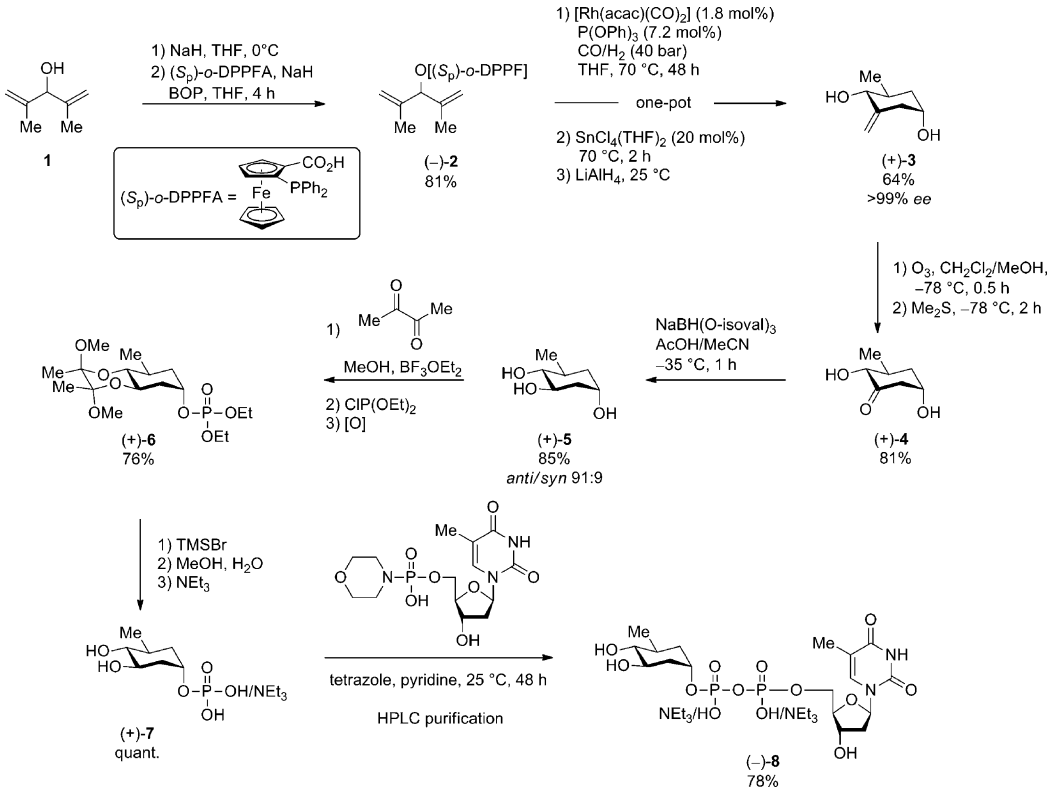


Figure 2. Three-dimensional structures of LanGT2. A) The LanGT2 monomer colored from blue at the N terminus to red at the C terminus. The protein is organized as a nucleotide-sugar-binding domain (top) and an aglycon-binding domain (bottom) that are connected through a flexible hinge. B) A 90° rotation highlights the grafted loop at the rim of the N-terminal domain. C) Detail of the grafted region in LanGT2 (red), LanGT2S8Ac (blue), and UrdGT2 (purple, PDB-ID 2P6P). The UrdGT2-derived region retains its conformation in the LanGTS8Ac chimera. The box in (B) highlights this region in LanGT2.

LanGT2 crystals were soaked with the synthetic TDP-carba-D-olivose (TcO) and LanGTS8Ac was cocrystallized with TcO according to the same protocols that merely yielded the TDP-bound structures when TDP-olivose was used. The structures of the two complexes were solved to 1.85 Å and 2.22 Å resolution, respectively (Table S1 in the Supporting Information), and they show a binding mode for TDP that is entirely consistent with the TDP complex and corresponds to the "tucked-under" conformation of the sugar described for other GTs.[12] Beside the additional electron density of the carbasugar moiety, the addition of the ligand also led to a change in the tertiary structures of the enzymes. The binding of TDP already induced a rotation of the two domains of the protein, which resulted in a closure of the interdomain cleft. With the TcO ligand, however, this closure was more



Scheme 1. Synthesis of α-D-carbaolivose thymidine diphosphate (**8**). acac = acetylacetonate, BOP = benzotriazol-1-yloxytris(dimethylamino)phosphonium, O-isoval = isovalerate, LAH = lithium aluminum hydride, THF = tetrahydrofuran, TMS = trimethysilyl.

pronounced, thus resulting in a 10° rotation of the C-terminal with respect to the N-terminal domain (Figure 3A). These observations indicate an induced-fit mechanism, as has been described for other GTs.[8b,13] In this "closed state", additional secondary structure elements of the protein were defined, in particular a large helical segment between residues 218 and 228 that was disordered in the absence of a ligand. In accordance with a sequential binding mechanism for the activated sugar and the aglycon, we thus conclude that the actual site for aglycon binding is only formed upon TDP-olivose binding and is stabilized by entropic effects. During the transition to the closed state, the loop region S217–F223, as well as the ^{283}HAGGVT^{288} motif in the adjacent loop, reorient to interact with the ligand, while residues S8–A12 and W136–R143 are displaced by carba-D-olivose (Figure S1 in the Supporting Information). Interestingly, helices α8a and α8b were poorly defined in the TDP-bound structure, thus indicating that the tightly closed structure of the enzyme had already relaxed towards an open conformation to release the product TDP (Figure S2).

Two distinct regions of the protein contribute to TDP-sugar binding. The loop region from W267–D271 mediates the recognition of the thymine base of TDP, with the indole side chain of residue W267 forming π-stacking interactions with the aromatic base, and three hydrogen bonds to the backbone peptide bonds of residues V268 and L270 providing the means to distinguish the thymine nucleobase from the similar cytosine. This arrangement is conserved among GT-1 family members, including the representative CalG1, CalG2, and SpnG.[14] On the side of the base distal from W267, the side chain of L270 is situated above the base, in a motif that is typically observed in TDP-sugar-utilizing GTs, such as GtfA, CalG3, SpnG, and SsfS6 (Figure 3B).[14,15] H283 in LanGT2 and LanGT2-S8Ac is conserved throughout the GT-1 family and has been proposed to stabilize the accumulation of charge on the phosphate during catalysis.[16] Additionally, the negative charges of the central pyrophosphate of the ligand are compensated by positioning this part of the molecule at the positively charged end of helix *h*10 of the enzyme. In the resulting binding pocket, multiple residues from the regions 217–220 and 283–288 interact with the two phosphate groups, thereby providing a stable binding mode that is identical to that observed for TDP alone (Figure 3B and Figure S3). As in other glycosyltransferases, recognition of the correct sugar moiety is achieved through hydrogen bonds. In LanGT2, only three residues form hydrogen bonds to the sugar, namely A284, G286 (5′-OH), and D137 (4′-OH; Figure 3B). Note that D137 is the only residue from the N-terminal aglycon-binding domain to interact with the activated sugar ligand, so this domain-spanning interaction may be essential for the observed conformational change upon ligand binding.

The conformational closure of the enzyme also alters the shape of the binding pocket, and here notable differences were observed between LanGT2 and LanGT2S8c. The resulting aglycon-binding pocket differs in size and shape between the two enzyme variants and this difference presumably forms the basis for the different outcome of the glycosyl transfer reaction.

Subsequent soaking and cocrystallization experiments with the TcO complexes of LanGT2 and LanGT2S8Ac in combination with various aglycon substrates (11-deoxylandomycinone, alizarin, anhydrolandomycinone, or tetrangulol) did not yield the desired ternary complexes, possibly because the specificity of the enzymes for the applied aglycons was too low. We therefore resorted to in silico ligand docking studies

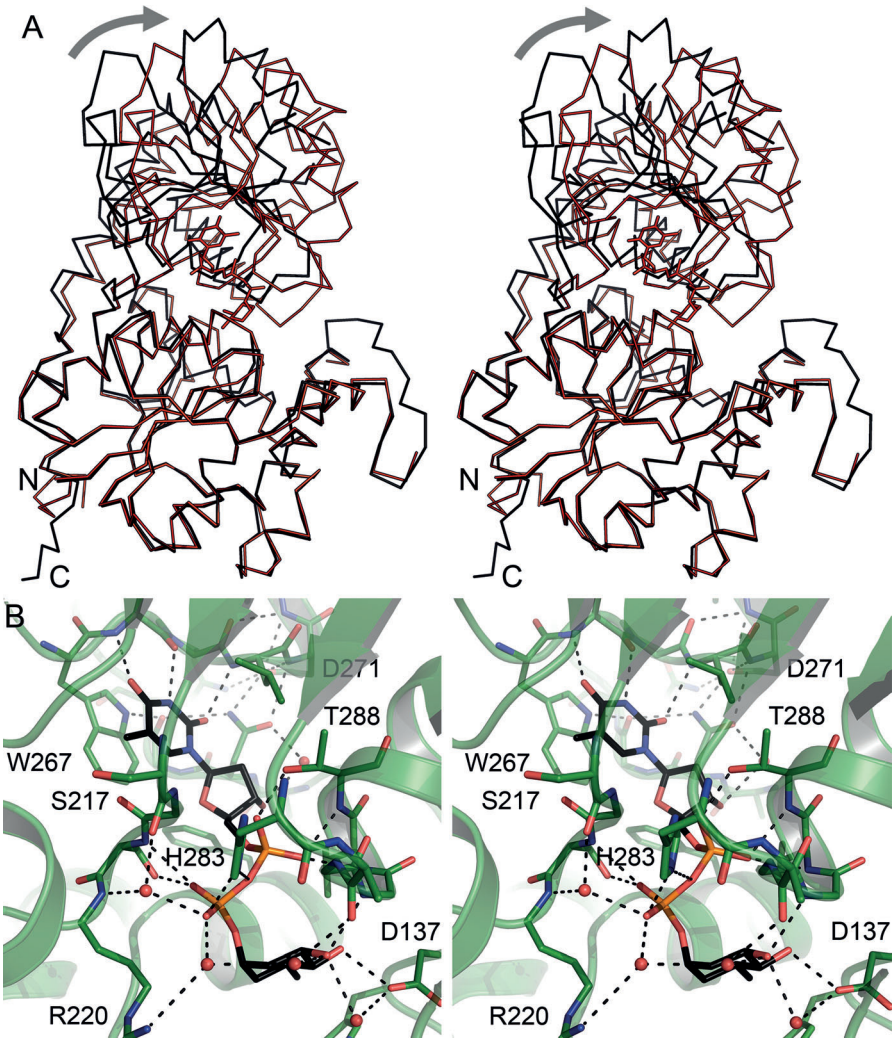


Figure 3. Ligand binding to LanGT2. A) Stereo image of a C_α-trace superposition of unbound LanGT2 (black) and the enzyme with the synthetic analogue TDP-carba-D-olivose (TcO, red) bound. The ligand induces a 10° rotation of the sugar-binding domain relative to the aglycon-binding domain (grey arrow). B) Stereo representation of the active site of LanGT2 with bound TcO. While the loop region 267–271 assures specific binding of the TDP moiety, the negatively charged phosphodiester is stabilized by the helix dipole of helix *h*10 (residues 286–295). The 4′ and 5′ hydroxy groups of olivose are recognized by G286 and D137, with the latter being the only residue from the N-terminal domain that is involved in binding.

based on the available crystal structures and a functional screening carried out previously.[7b] High-scoring docking solutions were obtained for 11-deoxylandomycinone and tetrangulol, where the aglycon substrates interacted with a series of hydrophobic residues (F83, W87, F88, M91, M116, and W136) in the interdomain cleft of the glycosyltransferases (Figure 4 and Figure S4). The aglycon-binding pocket of LanGT2S8Ac is considerably smaller than that of LanGT2, an effect that is mainly due to the conformation of helix α3 (in which residues I58 and I62 are oriented towards the cleft in the engineered C-GT). D137 is situated adjacent to the C8-OH group of the aglycon in both models, an observation consistent with previous studies that support the role of this residue as a catalytic base.[7b]

In the docking models, D137 is located 3.4 Å away from the C8-OH group, where it is suggested to recruit a water molecule to abstract a proton, thereby substantially increasing the nucleophilicity of the C8-hydroxy moiety. For the formation of an O-glycosidic bond by LanGT2, the C8-OH is closer to the sugar nucleotide donor than the *ortho* C9 atom (Figure 4A). By contrast, the altered orientation of the aglycon observed in the model for LanGT2S8Ac brings the C9 atom closer to the sugar nucleotide donor, thus resulting in a C–C coupling instead (Figure 4B). In the models, the distance between the C8-OH group (11-deoxylandomycinone) or the C9 atom (tetrangulol) and the C1′ atom of D-olivose is too large to give the productive orbital overlap necessary for glycosyl transfer. Catalysis may thus require a further rearrangement of the protein, possibly through the flexible loop region 64–76, which could act as a lid to cover the active site cleft.

Herein, we have shown a functional example for the concept of altering enzymatic specificity by exchanging specificity-determining regions between members of the same protein family. The synthesis and application of a nonhydrolyzable carbasugar analogue enabled us to probe the ordered sequential mechanism of the enzyme and has revealed an induced-fit mechanism to give the conformation that binds the aglycon compound. Grafting a loop region from UrdGT2 altered the binding pocket in LanGT2S8Ac so that docking simulations predict an entirely different mode of binding that results in the transformation of the O-glycosyltranferase into a C-glycosyltransferase.

Received: October 6, 2014
Published online: January 7, 2015

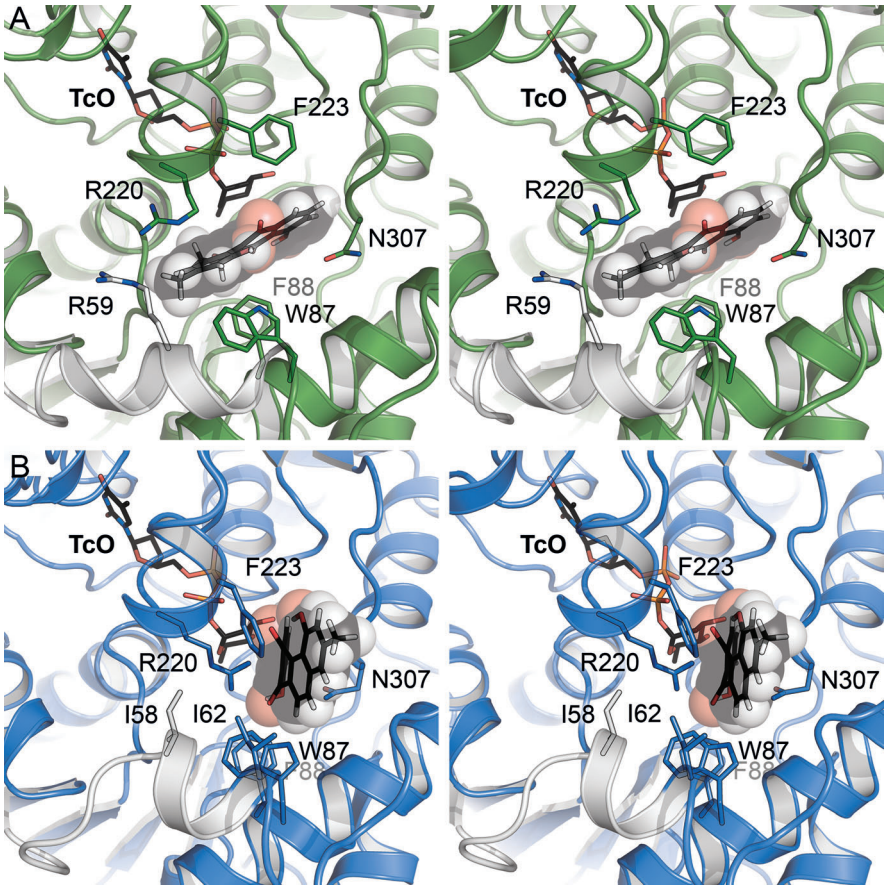


Figure 4. In silico analysis of aglycon binding. A) Stereo image of a docking solution for deoxylandomycinone and LanGT2. In the model, the ligand (shown as van der Waals spheres) displaces R220 and R59 and occupies the broad cleft formed by loop region 51–62 (white). B) In LanGT2S8Ac, the 51–62 loop was replaced by the UrdGT2 sequence (white), thus effectively closing the substrate-binding cleft. Molecular docking suggests that the ligand tetrangulol attains a different binding mode that reorients R220 towards the substrate and TDP-carbaolivose (TcO), thus resulting in a major repositioning of the ligand. In the orientation chosen, the catalytic base D137 is located behind the ligand.

Keywords: carbasugars · C-glycosylation · enzyme engineering · Friedel–Crafts alkylation · glycosyltransferases

[1] a) P. L. Hamilton, D. P. Arya, *Nat. Prod. Rep.* **2012**, *29*, 134 – 143; b) V. Kren, T. Rezanka, *Fems Microbiol. Rev.* **2008**, *32*, 858 – 889; c) A. C. Weymouth-Wilson, *Nat. Prod. Rep.* **1997**, *14*, 99 – 110.

[2] C. Vilches, C. Hernandez, C. Mendez, J. A. Salas, *J. Bacteriol.* **1992**, *174*, 161 – 165.

[3] a) C. Dürr, D. Hoffmeister, S. E. Wohlert, K. Ichinose, M. Weber, U. von Mulert, J. S. Thorson, A. Bechthold, *Angew. Chem. Int. Ed.* **2004**, *43*, 2962 – 2965; *Angew. Chem.* **2004**, *116*, 3022 – 3025; b) R. W. Gantt, R. D. Goff, G. J. Williams, J. S. Thorson, *Angew. Chem. Int. Ed.* **2008**, *47*, 8889 – 8892; *Angew. Chem.* **2008**, *120*, 9021 – 9024; c) D. Hoffmeister, B. Wilkinson, G. Foster, P. J. Sidebottom, K. Ichinose, A. Bechthold, *Chem. Biol.* **2002**, *9*, 287 – 295; d) S. H. Park, H. Y. Park, J. K. Sohng, H. C. Lee, K. Liou, Y. J. Yoon, B. G. Kim, *Biotechnol. Bioeng.* **2009**, *102*, 988 – 994; e) G. J. Williams, C. Zhang, J. S. Thorson, *Nat. Chem. Biol.* **2007**, *3*, 657 – 662; f) S. E. Wohlert, G. Blanco, F. Lombó, E. Fernández, A. F. Braña, S. Reich, G. Udvarnoki, C. Méndez, H. Decker, J. Frevert, J. A. Salas, J. Rohr, *J. Am. Chem. Soc.* **1998**, *120*, 10596 – 10601.

[4] a) A. Luzhetskyy, C. Méndez, J. A. Salas, A. Bechthold, *Curr. Top. Med. Chem.* **2008**, *8*, 680 – 709; b) A. P. Salas, L. L. Zhu, C. Sanchez, A. F. Braña, J. Rohr, C. Mendez, J. A. Salas, *Mol. Microbiol.* **2005**, *58*, 17 – 27; c) C. S. Zhang, E. Bitto, R. D. Goff, S. Singh, C. A. Bingman, B. R. Griffith, C. Albermann, G. N. Phillips, J. S. Thorson, *Chem. Biol.* **2008**, *15*, 842 – 853.

[5] a) R. W. Gantt, P. Peltier-Pain, J. S. Thorson, *Nat. Prod. Rep.* **2011**, *28*, 1811 – 1853; b) S. Singh, G. N. Phillips, J. S. Thorson, *Nat. Prod. Rep.* **2012**, *29*, 1201 – 1237.

[6] I. Baig, M. Kharel, A. Kobylyanskyy, L. L. Zhu, Y. Rebets, B. Ostash, A. Luzhetskyy, A. Bechthold, V. A. Fedorenko, J. Rohr, *Angew. Chem. Int. Ed.* **2006**, *45*, 7842 – 7846; *Angew. Chem.* **2006**, *118*, 8006 – 8010.

[7] a) A. Luzhetskyy, T. Taguchi, M. Fedoryshyn, C. Dürr, S. E. Wohlert, V. Novikov, A. Bechthold, *ChemBioChem* **2005**, *6*, 1406 – 1410; b) J. Härle, S. Günther, B. Lauinger, M. Weber, B. Kammerer, D. L. Zechel, A. Luzhetskyy, A. Bechthold, *Chem. Biol.* **2011**, *18*, 520 – 530; c) A. Mayer, T. Taguchi, A. Linnenbrink, C. Hofmann, A. Luzhetskyy, A. Bechthold, *ChemBioChem* **2005**, *6*, 2312 – 2315.

[8] a) M. Mittler, A. Bechthold, G. E. Schulz, *J. Mol. Biol.* **2007**, *372*, 67 – 76; b) P. K. Qasba, B. Ramakrishnan, E. Boeggeman, *Trends Biochem. Sci.* **2005**, *30*, 53 – 62.

[9] a) T. Bililign, C. G. Hyun, J. S. Williams, A. M. Czisny, J. S. Thorson, *Chem. Biol.* **2004**, *11*, 959 – 969; b) B. Faust, D. Hoffmeister, G. Weitnauer, L. Westrich, S. Haag, P. Schneider, H. Decker, E. Kunzel, J. Rohr, A. Bechthold, *Microbiology* **2000**, *146*, 147 – 154; c) L. B. Pickens, W. Kim, P. Wang, H. Zhou, K. Watanabe, S. Gomi, Y. Tang, *J. Am. Chem. Soc.* **2009**, *131*, 17677 – 17689; d) A. Trefzer, S. Pelzer, J. Schimana, S. Stockert, C. Bihlmaier, H. P. Fiedler, K. Welzel, A. Vente, A. Bechthold, *Antimicrob. Agents Chemother.* **2002**, *46*, 1174 – 1182.

[10] A. Gutmann, B. Nidetzky, *Angew. Chem. Int. Ed.* **2012**, *51*, 12879 – 12883; *Angew. Chem.* **2012**, *124*, 13051 – 13056.

[11] a) B. Breit, A. Bigot, *Chem. Commun.* **2008**, 6498 – 6500; b) B. Breit, D. Breuninger, *J. Am. Chem. Soc.* **2004**, *126*, 10244 – 10245; c) B. Breit, D. Breuninger, *Eur. J. Org. Chem.* **2005**, 3916 – 3929; d) B. Breit, D. Breuninger, *Synthesis* **2005**, 2782 – 2786.

[12] B. Schuman, J. A. Alfaro, S. V. Evans, *Top. Curr. Chem.* **2007**, *272*, 217 – 257.

[13] S. Moréra, L. Larivière, J. Kurzeck, U. Aschke-Sonnenborn, P. S. Freemont, J. Janin, W. Rüger, *J. Mol. Biol.* **2001**, *311*, 569 – 577.

[14] a) A. Chang, S. Singh, C. A. Bingman, J. S. Thorson, G. N. Phillips, *Acta Crystallogr. Sect. D* **2011**, *67*, 197 – 203; b) E. A. Isiorho, H. W. Liu, A. T. Keatinge-Clay, *Biochemistry* **2012**, *51*, 1213 – 1222; c) A. M. Mulichak, H. C. Losey, W. Lu, Z. Wawrzak, C. T. Walsh, R. M. Garavito, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 9238 – 9243.

[15] F. B. Wang, M. Q. Zhou, S. Singh, R. M. Yennamalli, C. A. Bingman, J. S. Thorson, G. N. Phillips, *Proteins Struct. Funct. Bioinf.* **2013**, *81*, 1277 – 1282.

[16] S. Ha, B. Gross, S. Walker, *Curr. Drug Targets Infect. Disord.* **2001**, *1*, 201 – 213.

