ChemComm

COMMUNICATION

Defining the molecular structure of teixobactin analogues and understanding their role in antibacterial activities

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

Anish Parmar [a], Stephen H. Prior[b], Abhishek Iyer[a, c], Charlotte S. Vincent[d], Dorien Van Lysebetten[c], Eefjan Breukink[e] , Annemieke Madder[c], Edward J. Taylor[d, *] and Ishwar Singh[a, *]

The discovery of the highly potent antibiotic teixobactin, which kills the bacteria without any detectable resistance, has stimulated interest in its structure activity relationship. However, a molecular structure-activity relationship has not been established so far for teixobactin. Moreover, the importance of the individual amino acids in terms of their L/D configuration and their contribution to molecular structure and biological activity are still unknown. For the first time, we have defined the molecular structure of seven teixobactin analogues through the variation of the D/L configuration of its key residues, namely N-Me-D-Phe, D-Gln, D-allo-Ile and D-Thr. Furthermore, we have established the role of the individual D amino acids and correlated this to the molecular structure and biological activity. Through extensive NMR and structural calculations, including molecular dynamics simulations we have revealed the residues for maintaining a reasonably unstructured teixobactin which is imperative for biological activity.

The discovery of the highly potent antibiotic teixobactin[1] to cope with the growing problem of antimicrobial resistance (AMR)[2] has provided a much needed impetus to antibiotic research. Although teixobactin does not mitigate all problems related to antibiotic resistance, it is a definite step in the right direction. In the short time since the antibiotic was discovered, several research groups, including our own, have made noteworthy contributions in the areas of total synthesis,[3, 4] syntheses and biological activities of teixobactin analogues[5, 6, 7, 8] and understanding the mechanisms of teixobactin.[9, 10, 11]

Replacing the enduracididine amino acid with its structurally closest natural amino acid arginine leads to Arg_{10} teixobactin analogue. The Arg_{10} teixobactin analogue follows a similar antibacterial activity trend as teixobactin (Figure 1). [3, 5, 6, 7] Therefore, conclusions drawn by synthesising analogues of the Arg_{10} teixobactin derivative should hold true for teixobactin as well.

Teixobactin is a cyclic depsipeptide and contains 11 amino acids, out of which 4 are D amino acids, namely N-Me-D-Phe, D-Gln, D-allo-Ile and D-Thr (Figure 1, marked in red). Peptides containing more D amino acids are generally less susceptible to enzymatic degradation[12] which may well be applicable for teixobactin.[1] In our previous work[6] we had already established the importance of the D-amino acids through the total synthesis of both D and L analogues of teixobactin. Replacing the D amino acids with their corresponding L configurations (except D-Thr_8) results in a 64-times decrease in antibacterial activity against *S. aureus.* The minimum pharmacophore of teixobactin, lipobactin[7] had also been published by Yang et al. indicating that residues 1-5 were probably responsible for membrane anchoring. It had also been shown that D-Thr_8 is critical for antibacterial activity and modification of the configuration of any residues in the core ring structure of teixobactin results in a significant decrease in activity.

In an interesting work, Yang et al, had showed that the complete enantiomeric teixobactin maintains full anti-bacterial activity (except in *Staphylococcus epidermis,* where it was reduced) indicating the possibility that only the relative configurations of amino acids and not their absolute configurations were important for antibacterial activity.

It is not yet known, however, if a molecular structure (three-dimensional structure)-activity relationship exists for teixobactin. To date, with the exception of the structural deduction of teixobactin published by Ling et. al.,[1] no molecular structural studies on teixobactin or its analogues

[a] School of Pharmacy, JBL Building, University of Lincoln, Beevor St. Lincoln LN67DL, UK. E-mail: isingh@lincoln.ac.uk
[b] School of Chemistry, University of Lincoln, Beevor St. Lincoln LN67DL, UK.
[c] Organic and Biomimetic Chemistry Research Group, Department of Organic Chemistry, Ghent University, Krijgslaan 281 (S4), B-9000 Ghent, Belgium
[d] School of Life Sciences, Joseph Bank Laboratories, University of Lincoln, Green Lane, Lincoln LN6 7DL. E-mail: etaylor@lincoln.ac.uk
[e] Department of Membrane Biochemistry and Biophysics, Institute of Biomembranes, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands

†Electronic Supplementary Information (ESI) available: Peptide synthesis, HPLC, LC-MS analysis, NMR spectra See DOI: 10.1039/x0xx00000x

have been reported. Moreover, the impact on the molecular structure and

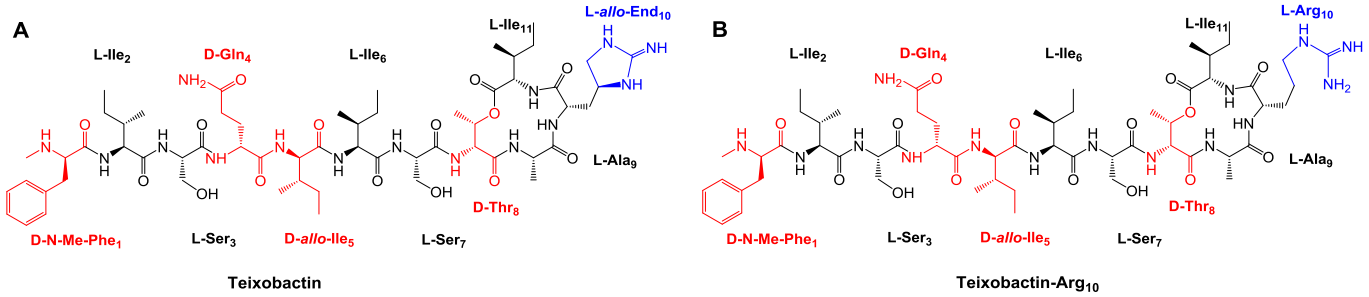

Figure 1. Structure of (A) Teixobactin and (B) Teixobactin-Arg_{10} with the D amino acids highlighted in red and the structural differences marked in blue.

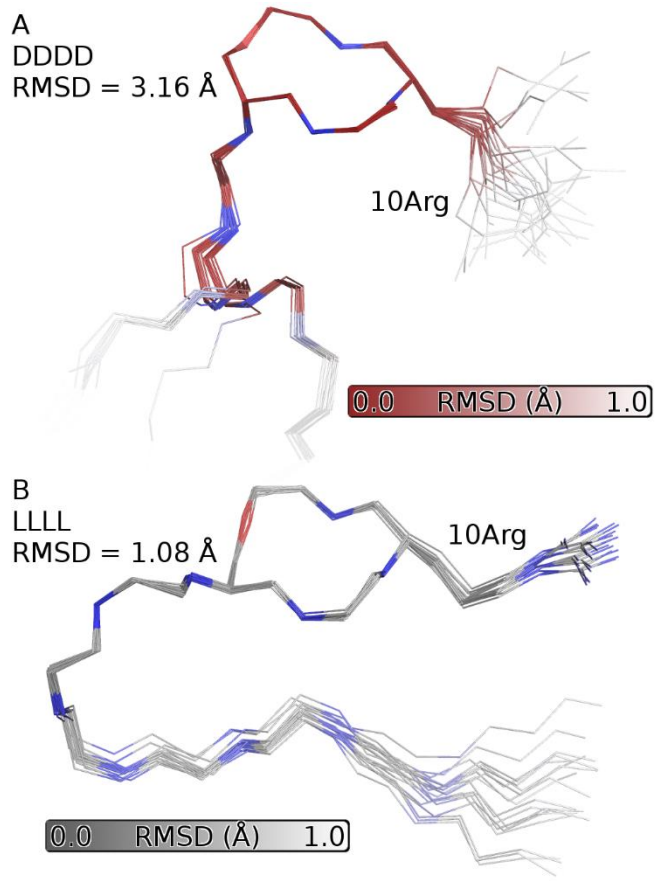

Figure 2. A. Structure of teixobactin analogue **1** (DDDD) exhibiting native stereochemistry. B. Structure of analogue **2** (LLLL) containing L-Thr8. Overlays of the 20 lowest energy structures aligned to the backbone atoms of the cyclic region are displayed. Atoms are colored according to their RMSD, fading to white at higher RMSD. Atom transparency also increases with RMSD, with more highly disordered regions fading to invisibility. For clarity, only backbone atoms and the non-hydrogen atoms of Arg_{10} are displayed.

activity of teixobactin by varying individual amino acids in terms of their L/D configuration is still unknown. This study aims to explain why L analogues of teixobactin are not active. We have selected the D amino acid residues and systematically replaced them with L configurations to understand the minimum number of D amino acid residues required to maintain biological activity. Furthermore, the low cost of L amino acids was expected to lower the financial constraint on teixobactin development. For the first time, we have determined molecular structures of seven teixobactin derivatives by changing D/L configurations from NMR, their antibacterial activity and lipid II binding. The results from our work will enable the better understanding of molecular structure activity relationship of teixobactins and their further development as drug like molecules.

In this work, we have investigated the role of each of the D amino acids, their impact on the molecular structure and activity of teixobactin and whether there exists a structure-activity relationship for the molecule. Through the syntheses of seven analogues of Arg_{10}-teixobactin (Table 1), followed by extensive NMR and structural calculations we have shown the importance of the D amino acids on the structure and activity of teixobactins.

Analogue No.	Code	Configuration of Amino Acids				
		N-Me-Phe	Gln	Ile	Thr	MIC[†]
1	DDDD	D	D	D	D	2
2	LLLL	L	L	L	L	GAW[‡]
3	DDLD	D	D	L	D	GAW
4	DLDD	D	L	D	D	GAW
5	LDDD	L	D	D	D	32-64
6	LLDD	L	L	D	D	GAW
7	LLLD*	L	L	L	D	128

Table 1. List of analogues of Arg10-teixobactin synthesized by varying the D/L configuration of the key residues N-Me-D-Phe_1, D-Gln_4, D-*allo*-Ile_5 and D-Thr_8.
[†] MIC values were measured against *S. aureus* and are given in μg/mL
[‡] Growth in all wells.
*This analogue contains an acetylated Phe instead of an N-Me-Phe residue.

Our initial efforts were focused towards understanding the pivotal role played by the stereochemistry of D-Thr_8 upon the gross structure of teixobactin. This was achieved through the synthesis of analogue **2** (LLLL) and subsequent comparison with analogue **1** (DDDD, Figure 2). Outside the cyclic region encompassing residues 8 to 11 the peptide was largely unstructured (Figure 2A). The altered topological environment available to the analogue **2** (LLLL) makes it energetically favourable to adopt a well-defined hairpin structure (figure 2B).

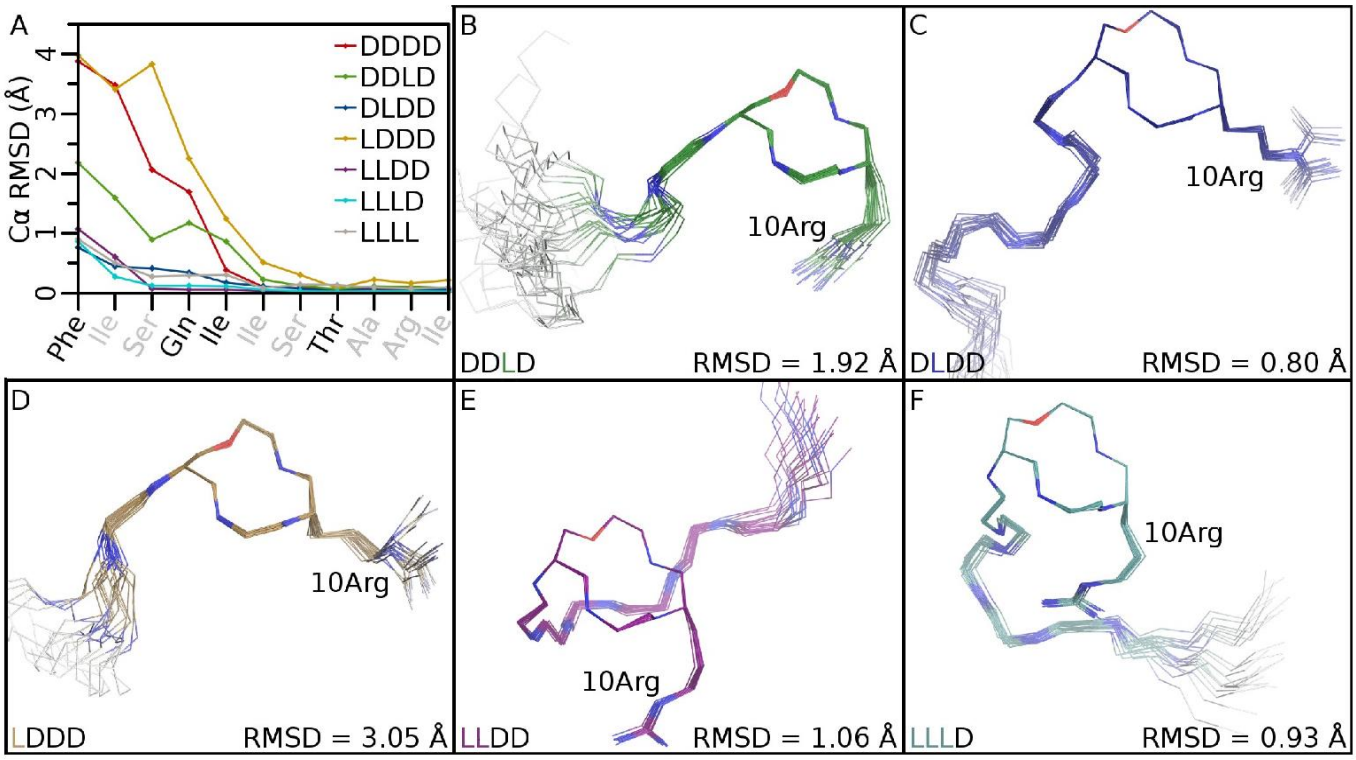

Figure 3. Effect of stereochemistry at different positions on the structure of teixobactin analogues. **A**. Majority of variation observed in structural ensembles accounted for by varying levels of disorder in the *N*-terminal residues. Positions labelled in *black type* have had their stereochemistry altered. The name of the analogues is as described in Table 1. **B-F**. Backbone traces of 20 lowest energy structures aligned to the backbone atoms of the cyclic region. Atom color and transparency as in Figure 1.

The teixobactin analogues routinely gave spectra of the highest quality (Figure S16), which made complete atomic assignment of each analogue possible (Table S2). Subsequent measurement of dipolar correlations allowed for full structural characterisation. The number of visible NOE cross-peaks (Figure S16) varied with different analogues: those with D-Gln$_4$ (for example figure S16 red contours) contained very few medium- and long-range correlations, which is typical for highly dynamic, unstructured peptides in which nuclei are seldom in close enough proximity for NOE build-up to occur. This flexibility is reflected in their high RMSD (Figure 2 A, Figure 3, Table S3).

In contrast, analogues possessing L-Gln$_4$ (**2**, **4**, **6**, **7**) (for example figure S16 blue contours) contained many more crosspeaks in their NOESY spectra, some of which are categorised as long-range (Table 2), a category entirely missing from the D-Gln$_4$ containing analogues (**1**, **3**, **5**). Having numerous medium- to long-range NOEs is characteristic of peptides adopting a higher degree of structure, and this situation is reflected in their sub- to low Angstrom convergence (Figure 2 B, Table S3).

Possessing an L-Ile$_5$ imparts a slight tendency for structure on the teixobactin analogues (Figure 3 A, B, F) but does not result in a high deviation from the original structure unlike that observed for L-Gln$_4$. For instance, the RMSD of analogue **1** (native stereochemistry) is ~3 Å (Figure 2 A, Figure 3 A) whereas that of **3** (DDLD), in which D-*allo*-Ile$_5$ has been replaced by L-Ile$_5$, is reduced to ~2 Å (Table 2, Figure 3 A, B). A similar slight reduction in RMSD is observed in **7** (LLLD, 0.93 Å; Table 2, Figure 3 A, F) when compared to **6** (LLDD, 1.06 Å; Table 2, Figure 3 A, E) where again, the only difference between them is the stereochemistry at position 5. The stereochemistry of position 1 has little to no effect on overall structure, as can be seen when one compares the RMSDs of **1** (DDDD) to **5** (LDDD, both ~3 Å; Figure 2A, Figure 3 A, D, Table S3). Although highly disordered, the N-terminal domains of both these analogues occupy a similar overall spatial region. Since the stereochemistry of position 1 is not important to the structure, it is likely that D configuration is important for slowing down enzyme degradation or for biological activity.

Taken together, this data allows us to conclude that the stereochemistry of position 4 is critical to structural stability. An ubiquitous structural characteristic of the L-Gln$_4$ containing analogues absent from the D-Gln$_4$ containing analogues is that Ile$_6$ packs against Ile$_{11}$. It is therefore possible that this long-range packing arrangement is made possible by the altered stereochemistry at position 4 and has the effect of stabilising the structures. We analysed the nature of this packing through extended molecular dynamics (MD) simulations on the microsecond time-scale. Trajectories were calculated for a total of 0.1 μs for analogue **1** and **7** in explicit solvent (Figure S17). It was immediately obvious from the simulations that the native analogue seldom visits this packing arrangement (Figure S17 B), and on the few occasions it does it is very short-lived. However, the sidechains of Ile$_6$ and Ile$_{11}$ remain in constant hydrophobic contact throughout the entire simulation in the case of **7** (LLLD, Figure S17 B).

Functionally, it was important to understand how this increased stability through packing might explain the MICs observed in all analogues except **5** (LDDD, Table 1). A plausible mechanism was also revealed through MD, in that the sidechain of Arg_{10} is less solvent exposed in **7** (LLLD, Figure S17 A). This interaction is stabilised through the formation of numerous transient intramolecular hydrogen bonds between Arg_{10} and other residues (Figure S17, bottom panel), a situation almost entirely lacking in the native analogue.

Therefore, it is possible to surmise that if the hydrogen bond donors of the guanidine group of Arg_{10} are spending a considerable proportion of their time in forming intra-molecular hydrogen bonds then they are unavailable for the formation of inter-molecular bonds. Thus, if Arg_{10} is unavailable for involvement in intermolecular recognition events, this could help to explain the decrease in MIC in case of analogue **7** (LLLD) as compared to **1** (DDDD, Table 1). The frequency of unfavoured backbone torsion angles present in the teixobactin analogues (Table S3) correlates strongly with D amino acid content. This could be due to limitations with the analysis algorithms used.

Figure S17 reveals that native teixobactin (Arg_{10}) is highly dynamic in solution, whereas an analogue containing L-Gln_4 is both more tightly packed and structurally stable. This increased burial (figure S17 A) is probably a result of the formation of a number of transient hydrogen bonds between the guanidine group of Arg_{10} and other polar sidechains in the analogue (*teal bars*). In contrast, Arg_{10} only forms two short-lived hydrogen bonds (*red bars*) throughout the whole simulation in the native form.

Complex formation of teixobactin analogues **1** (DDDD) and **2** (LLLL) with lipid II and geranyl pyrophosphate were performed as described previously using TLC (SI VIII) [1]. Both analogues bind to lipid II and geranyl pyrophosphate which is indicated by the reduction of lipid II/ geranyl pyrophosphate visible on TLC (figure S18 & S19). The binding of lipid II to LLLL analogue is unexpected as the LLLL is not biologically active. Therefore, it can be concluded that the binding assay on isolated lipid II does not necessarily reflect whole cell activities. A plausible explanation for this can be that lipid II is more freely available in solution than in the case of bacteria.

We have for the first time determined the 3D molecular structure of seven teixobactin analogues and established the importance of the individual amino acids in terms of their D/L configurations in maintaining a relatively unstructured teixobactin. Furthermore, we can directly correlate this propensity for structure with antimicrobial activity and have identified a possible mechanism by which this disorder is maintained. We have shown that the D- N-Me-Phe_{11} is not important from a structural standpoint. On the other hand, we have identified D-Gln_4 as being very essential and D-Ile_5 being important in maintaining the disordered structure of teixobactin, which is vital for its biological activity. We believe this work to be critical in understanding the structural-activity relationship and mechanism of teixobactin and its analogues and their further development as drug like molecules. The results presented in this work are of broad general interest and are expected to facilitate the future development of teixobactin derivatives and peptide based antibiotics for addressing the serious challenges posed by AMR.

Abhishek Iyer, Charlotte S. Vincent and Anish Parmar would like to thank the University of Lincoln for funding. Edward Taylor would like to thank the Royal Society for their kind support (grant number UF100116). Ishwar Singh would like to acknowledge the Royal Society for their kind support (grant number (RG130163) and Horizon 2020 (645684)). Dorien Van Lysebetten acknowledges the Ghent University Special Research Fund (BOF15/DOC/379) for financial support. Jan Goeman from Ghent University is thanked for the LC-MS/HRMS analysis.

Notes and references

1 L. L. Ling, T. Schneider, A. J. Peoples, A. L. Spoering, I. Engels, B. P. Conlon, A. Mueller, D. E. Hughes, S. Epstein, M. Jones, L. Lazarides, V. a Steadman, D. R. Cohen, C. R. Felix, K. A. Fetterman, W. P. Millett, A. G. Nitti, A. M. Zullo, C. Chen and K. Lewis, *Nature*, 2015, **517**, 455–459.

2 *http://amr-review.org/*, 2015.

3 A. M. Giltrap, L. J. Dowman, G. Nagalingam, J. L. Ochoa, R. G. Linington, W. J. Britton and R. J. Payne, *Org. Lett.*, 2016, **18**, 2788–2791.

4 K. Jin, I. H. Sam, K. H. L. Po, D. Lin, E. H. Ghazvini Zadeh, S. Chen, Y. Yuan and X. Li, *Nat. Commun.*, 2016, **7**, 12394.

5 Y. E. Jad, G. A. Acosta, T. Naicker, M. Ramtahal, A. El-Faham, T. Govender, H. G. Kruger, B. G. De La Torre and F. Albericio, *Org. Lett.*, 2015, **17**, 6182–6185.

6 A. Parmar, A. Iyer, C. S. Vincent, D. Van Lysebetten, S. H. Prior, A. Madder, E. J. Taylor and I. Singh, *Chem. Commun.*, 2016, **52**, 6060–6063.

7 H. Yang, K. H. Chen and J. S. Nowick, *ACS Chem. Biol.*, 2016, 1823–26.

8 S. A. H. Abdel Monaim, Y. E. Jad, G. A. Acosta, T. Naicker, E. J. Ramchuran, A. El-Faham, T. Govender, H. G. Kruger, B. G. de la Torre and F. Albericio, *RSC Adv.*, 2016, **6**, 73827–73829.

9 F. von Nussbaum and R. D. Süssmuth, *Angew. Chemie Int. Ed.*, 2015, **54**, 6684–6686.

10 W. Lee, K. Schaefer, Y. Qiao, V. Srisuknimit, H. Steinmetz, R. Müller, D. Kahne and S. Walker, *J. Am. Chem. Soc.*, 2016, **138**, 100–103.

11 T. Homma, A. Nuxoll, A. B. Gandt, P. Ebner, I. Engels, T. Schneider, F. Götz, K. Lewis and B. P. Conlon, *Antimicrob. Agents Chemother.*, 2016, 10.1128/AAC.01050-16.

12 R. Tugyi, K. Uray, D. Iván, E. Fellinger, A. Perkins and F. Hudecz, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 413–418.

