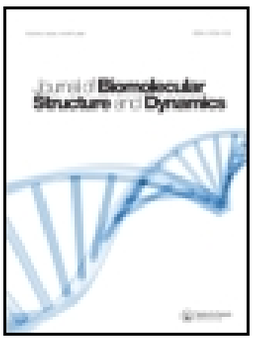


Taylor & Francis
Taylor & Francis Group

Journal of Biomolecular Structure and Dynamics

ISSN: 0739-1102 (Print) 1538-0254 (Online) Journal homepage: http://www.tandfonline.com/loi/tbsd20

Rational antibiotic design: in silico structural comparison of the functional cavities of penicillin-binding proteins and ß-lactamases

Mame Ndew Mbaye, Dimitri Gilis & Marianne Rooman

To cite this article: Mame Ndew Mbaye, Dimitri Gilis & Marianne Rooman (2017): Rational antibiotic design: in silico structural comparison of the functional cavities of penicillin-binding proteins and ß-lactamases, Journal of Biomolecular Structure and Dynamics, DOI: 10.1080/07391102.2017.1418678

To link to this article: https://doi.org/10.1080/07391102.2017.1418678

Accepted author version posted online: 18 Dec 2017.

Publisher: Taylor & Francis
Journal: *Journal of Biomolecular Structure and Dynamics*
DOI: http://doi.org/10.1080/07391102.2017.1418678

Rational antibiotic design: in silico structural comparison of the functional cavities of penicillin-binding proteins and ß-lactamases

Mame Ndew Mbaye[a,b], Dimitri Gilis[a], Marianne Rooman[a,*]

[a] 3BIO-BioInfo group, Université Libre de Bruxelles, CP 165/61, avenue F. Roosevelt 50, 1050 Bruxelles

[b] Université Cheikh Anta Diop de Dakar, Senegal

* Corresponding author, email: mrooman@ulb.ac.be

Rational antibiotic design: in silico structural comparison of the functional cavities of penicillin-binding proteins and ß-lactamases

Abstract

The class of ß-lactam antibiotics has proven highly efficient in targeting bacterial penicillin binding proteins (PBP) and leading to the blocking of the bacterial cell wall synthesis. However, the benefit of these drugs is limited because of bacterial resistance mechanisms; the most widespread resistance involves ß-lactamase enzymes (ßLACT) that inactivate ß-lactam-based molecules. We focused on PBPs and ßLACTs from enterobacteria, and performed a detailed in silico study of PBPs whose inactivation is lethal for the bacteria and of ßLACTs that have a PBP-type catalytic mechanism. The comparison of the sequences and structures of PBPs and ßLACTs shows an almost perfect conservation of the catalytic site, and a high spatial

resemblance of the whole functional cavity despite a very low overall sequence identity. Some notable differences in the functional cavity were observed in the vicinity of the catalytic site: four tyrosines are well conserved in the PBPs, whereas the residues occurring at equivalent positions in the ßLACT families present other physicochemical properties. These tyrosines are thus good candidates to be targeted in designing new antibiotic molecules with increased affinity and specificity for PBPs, with the goal of overcoming drug resistance. Our analysis also identified residues that have similar characteristics in most ßLACT families and different properties in PBPs; these are interesting targets for new ligands that specifically inhibit ßLACT proteins. The in silico approach presented here can be extended to other protein systems in view of guiding and improving rational drug design.

Keywords: Drug design, ß-Lactam antibiotic resistance, Enterobacteria, Protein sequence comparisons, Protein structure comparisons

List of abbreviations:

PBP: penicillin binding protein
ßLACT: ß-lactamase
3D: 3-dimensional
PDB: Protein Data Bank
rms: root mean square

1. Introduction

The discovery of penicillin in 1928 and its first medical use in the 40s marked the beginning of a successful era in the fight against bacterial infections (Hirsh, 1948), which has led to a drastic decrease in mortality from previously incurable infectious diseases. Although several other classes of antibiotic molecules have been developed over the years, ß-lactam antibiotics - among which penicillin - are still the most widely used (Van Boeckel, 2014; Bush & Bradford, 2016). These molecules share a ß-lactam ring, with various moieties attached to it. They have been extremely successful for many years, until the targeted bacteria started developing resistance. But even though the fate of ß-lactam antibiotics is sometimes questioned (Livermore, 2009), modified versions of the original molecules that escape resistance continue to be developed and to be clinically successful, which suggests that they still have a future (Page & Bush, 2014; Frère & Page, 2014; Qin *et al*. 2014).

The action of β-lactam antibiotics is well known (Wise & Park, 1965). They target specific bacterial receptors called penicillin-binding proteins (PBPs), which are involved in the biosynthesis of the bacterial cell wall by catalyzing the polymerization of the glycan strand and the cross-linking between glycan chains (Ghuysen, 1991). Their

natural substrate is D-Ala-D-Ala. The different synthesis steps involve several families of PBPs, referred to as PBP1-PBP12, which are classified based on their structure and catalytic activity and vary according to the type of bacteria (Sauvage *et al.*, 2008). The PBP catalytic site is formed by three sequence motifs in spatial proximity. One of these motifs contains an active serine that plays a key role in the catalysis. These three motifs are conserved among all PBP classes. Penicillin and other β-lactam antibiotics utilize their structural resemblance to the natural substrate to bind to the PBP active site (Ghuysen, 1991). Upon binding, they inactivate the PBP and block the synthesis of the cell wall. PBP inactivation is always deleterious and in some cases lethal for the bacteria. Its precise effect depends on the class of PBP and on the type of bacteria (Sauvage *et al.*, 2008).

Enterobacteriaceae are a class of gram-negative bacteria, which contain many genera among which *Escherichia, Yersinia*, *Klebsiella* and *Salmonella.* They cause a large series of diseases. Some are true pathogens, whereas others are opportunistic and cause secondary infections of, for example, wounds, the urinary and respiratory tracts and the circulatory system (Guentzel, 1996; Iredell *et al.,* 2016). Several classes of antibiotics have been developed against *enterobacteriaceae*, including ß-lactam antibiotics. It has been shown that the inactivation of the PBPs of type PBP2 or PBP3, or the simultaneous inactivation of both PBP1A and PBP1B, is lethal for the enterobacteria (Pfeifle *et al.,* 2000; Sauvage *et al.*, 2008; Derouaux *et al.*, 2008). The other types of PBPs play less important roles in this type of bacteria.

The major mechanism of *enterobacteriaceae* bacterial resistance against β-lactam antibiotics is the secretion of ßLACT enzymes, which catalyze the hydrolysis of β-lactams and hence degrade ß-lactam-based antibiotics (Fisher *et al.*, 2005; Cag *et al*., 2016; Bonomo, 2017). ßLACTs can be divided into two groups based on their enzymatic mechanism (Bush & Jacoby, 2010). The first (class B) consists of metallo-enzymes containing zinc ions that are involved in ß-lactams hydrolysis (Bebrone, 2007). Enzymes of the second group, containing classes A, C and D, hydrolyze the ß-lactams through a serine esterification mechanism (Ghuysen, 1991; Ruggiero, 2017). We focus in this study only on the latter group that contains active serine ßLACTs, as they have the same catalytic mechanism as PBPs. In particular, they contain the three characteristic motifs of PBPs.

The goal of this paper is to guide the rational design of ß-lactam antibiotics that target PBPs from *enterobacteriaceae*, escape the resistance due to ßLACTs, and have a lethal effect on the bacteria. Although these systems have been studied a lot, few studies have been devoted to the analysis of their three dimensional (3D) structures, and when they do, they are often restricted to the analysis of the catalytic site (Pratt, 2016). In contrast, we study the whole functional cavity and discuss how known ß-lactam-based antibiotic molecules could be extended to gain specificity for either PBPs or ßLACTs.

Among the different PBPs from enterobacteria, we chose to focus on PBP3, as it is the only PBP whose inactivation is lethal and whose 3D structure has been experimentally determined. We compared this protein to the families of active-site serine ßLACTs, which are found in *enterobacteriaceae* and for which at least one experimental structure is available.

We would like to stress that the in silico procedure developed here can be applied to many other protein systems, in view of designing in a rational manner inhibitor or activator molecules that are specific to one protein family and not to others.

2. Results and Discussion

2.1 Penicillin binding protein and ß-lactamase families

We focused on families of PBPs and active-site serine ßLACTs which are expressed in enterobacteria and for which at least one experimental X-ray structure is available in the Protein Data Bank (PDB) (Berman *et al.*, 2000). A single PBP family, namely PBP3, fulfills these criteria with the additional constraint that its inactivation is lethal. Among ßLACTs, nine families satisfy these conditions. Among these, seven are of class A (TEM, SHV, CTX, PER, GES, KPC, SME), one of class C (CMY), and one of class D (OXA); we used here the classification described in Bush & Jacoby (2010), based on sequence, specificity, and kinetics. These different proteins and their characteristics are listed in Table 1 and Table S1 of Supplementary Material.

PBP and ßLACT protein sequences were retrieved from the UniProt database (Leinonen *et al.*, 2004). We considered both the subset of manually annotated sequences from the Swiss-Prot database (Bairoch *et al.,* 2004) as well as the unreviewed sequences. The number of reviewed sequences is indeed very low, sometimes even zero (Table 1). Multiple sequence alignments were performed on all sequences belonging to the same PBP or ßLACT family, as described in Methods. We observed that the sequence conservation within each of the families is quite high (Table S2).

Table 1. Sequences and structures of PBPs and ß-LACTs used in the present analysis. The rows in bold correspond to the representative protein in each class. The number of sequences in parentheses indicate the subset of manually annotated sequences from the Swiss-Prot database.

Family		Number of sequences	Representative UniProt sequence	Number of structures	Representative PDB structure	Resolution (Å)	Number of residues in cavity	Cavity accessible surface area (Å^2)	Cavity molecular area (Å^2)
Penicillin Binding Proteins									
PBP	**PBP3**	**547 (5)**	**P0AD68**	**1**	**4bjp**	**2.50**	**50**	**2569**	**5050**
ß-Lactamases									
Class A	**KPC**	**46 (2)**	**A8DS27**	**3**	**3c5a**	**1.23**	**40**	**1895**	**4002**
	TEM	452 (3)	P62593	12	1m40	0.85	37	1453	3636
	SHV	287 (16)	P30896	5	1n9b	0.90	40	1614	3955
	GES	18 (0)	Q9KJY7	2	4gog	1.19	38	1287	3508
	CTX	566 (6)	C0L2W2	17	4hbt	1.10	41	1723	3790
	SME	5 (2)	Q54488	1	1dy6	2.13	39	1767	3748
	PER	16 (0)	P37321	1	4d2o	2.20	50	2221	5038
Class C	**CMY**	**66 (0)**	**Q99QC1**	**2**	**1zkj**	**1.55**	**52**	**1961**	**5392**
Class D	**OXA**	**85 (6)**	**P13661**	**5**	**1m6k**	**1.50**	**47**	**2012**	**4991**

The X-ray structures of ßLACTs and PBPs were retrieved from the PDB databank (Berman *et al.*, 2000) (Tables 1 and S1). The number of structures per family varies between 1 and 17. The representative structure of each family was chosen as the best resolution structure. These representative structures have all a resolution of 2.5 Å at most (Table S1).

Although the protein sequences within each family are quite well conserved, the sequence variability between proteins of different families is quite high. The pairwise sequence identity between the representative sequences of PBP and ßLACT proteins is insignificant, and so is the sequence identity between ßLACT proteins of different classes (Table S3). Only ßLACT families of class A have reasonably similar sequences, with a sequence identity ranging between 23 and 64%.

Despite the limited sequence similarity, PBP and ßLACTs adopt similar 3D structures (Table S4). The root mean square (rms) deviation between superimposed representative structures varies between 1 Å and 2 Å for ßLACT families of class A, and between 2 Å and 3 Å for PBPs and ßLACTs of different classes. We chose KPC as the representative protein of class A ßLACTs, given that it has the lowest average rms deviation with respect to the other class A proteins. The structural variability between ßLACTs of the different classes is illustrated in Fig. S1.

2.2 Delimitation and analysis of the functional cavities

For each PBP and ßLACT family, we identified the residues that are part of the functional cavity, using the procedure described in Methods. The cavity sequences of PBP and of the three ßLACT class representatives are shown in Fig. 1; the 3D structures

of the cavities of PBP and of the nine ßLACT families are depicted in Figs 2 and S2.

Figure 1. Residues that are part of the functional cavities in the representatives of the ßLACT classes and PBPs. The catalytic motifs, in blue, are aligned. The residues of PBP that differ from the equivalent residues in all ßLACTs, or all ßLACTs but one (see Table S5), are indicated in green and pink, respectively, and their spatial correspondence are marked by lines.

The cavity size is characterized by the number of residues, the solvent accessible surface area and the molecular surface area (Table 1). Class A ßLACTs but PER have by far the smallest cavity. Class C ßLACT has the largest cavity in terms of number of residues and molecular surface, and the PBP cavity is the largest in terms of solvent accessible surface area. This indicates that the PBP cavity is more open in the X-ray structure considered.

The ßLACTs and PBP functional cavities contain three well described motifs that constitute the catalytic site and are almost perfectly conserved (Sauvage *et al.*, 2008): SxxK, which contains the active serine, (S|Y)xN and (K|H)(S|T)G, with '|' denoting 'or' and 'x' any residue. Only in the OXA family is the asparagine of the second motif replaced by a valine. The conservation of these sequence motifs within each family is shown in Table S2.

The sequence similarity of the functional cavities is basically limited to these catalytic motifs (Fig. 1). Indeed, apart from these motifs, no global sequence alignment can be found, and the non-catalytic cavity residues seem to differ completely from one family to the other. This is not only true when comparing PBP and ßLACTS, but also ßLACTS of different classes. Note however that, within each of the subfamilies, the cavity residues are quite well conserved (Table S2). This result clearly demonstrates that sequence analyses are not sufficient to validly compare the cavities among PBPs and ßLACTs, but that we must use 3D structure representations and comparisons.

The cavities of PBP and all ßLACTs are depicted in Figs 2.a-b and Figs S2, with the residues colored according to their physicochemical properties. At first sight, the cavities differ, with PBP and class D ßLACT cavities being less polar and more aromatic

than class A and C ßLACT cavities. Note that some variability is also observed among class A ßLACTs, even though their sequences are well conserved.

To compare the functional cavities of PBP and ßLACTs in a more detailed and rigorous way, we first had to determine the spatial correspondences between the residues from different cavities. For that purpose, we superimposed the ßLACT cavity structures onto the PBP cavity, and identified the residues that are spatially close (see Methods). This yielded the corresponding PBP and ßLACT residues, which occupy equivalent spatial positions in their respective cavities. As shown in Fig. 1, the residue succession along the sequence is not maintained: the cavity residues of the different PBP and ßLACT classes seem to have been shuffled along the sequences.

2.3 Cavity characteristics specific to PBPs

Since the first objective is to help designing new ligand molecules that bind to PBP but not to any of the ßLACTs, or that bind to ßLACTs with lower affinity, we searched for the functional cavity features that are unique to PBP and do not appear in any of the nine ßLACT families. More specifically, we identified the residues in the PBP cavity whose physicochemical properties differ from those of the equivalent residues in the ßLACT families. We used for that purpose the spatial correspondence between ßLACT and PBP cavity residues using PBP as a reference, and the physicochemical properties listed in Methods; note that each PBP residue can have zero, one or several equivalent residues in each ßLACT. We used both a strict and a slightly relaxed criterion: the first involves identifying PBP residues whose physicochemical properties differ from all equivalent residues in all nine ßLACT families and the second, from all equivalent residues in all but one ßLACT families. The results are shown in Tables 1 and S5 and Figs 2.c-d.

To ensure that the physicochemical properties so identified are observed in most members of the protein families and not just in their representatives, we analyzed the amino acid conservation within each PBP and ßLACT family (section 2.1); the observed amino acids and their frequencies are specified in Table S5. To select a PBP-specific feature, we required it to occur in at least 75% of the PBPs, and in at most 5% of the ßLACTs.

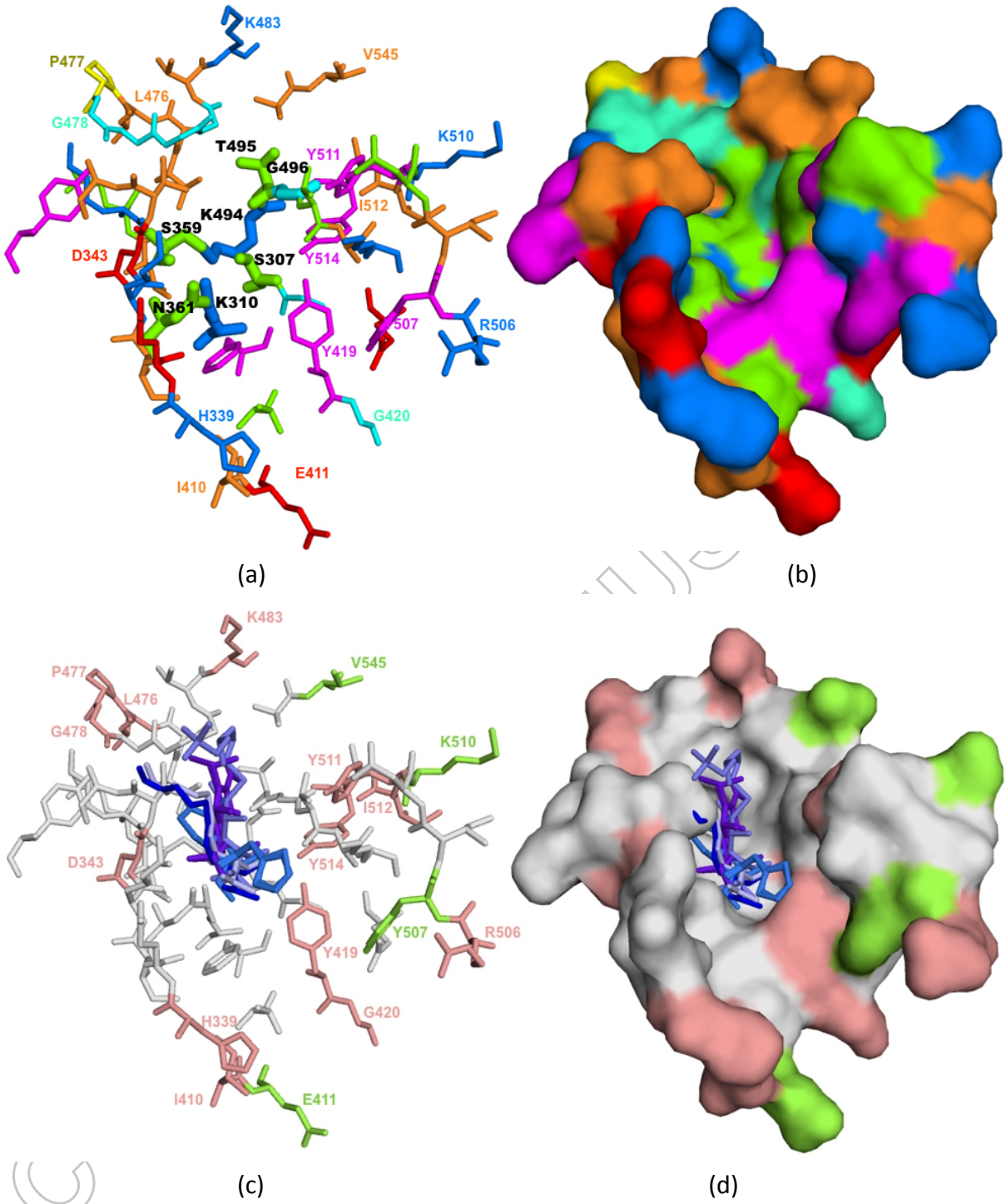


Figure 2. Stick (a,c) and surface (b,d) representations of the functional cavities of PBP, with similar spatial orientations. (a)-(b) PBP cavity colored according to the physicochemical properties of its residues. Positively and negatively charged residues are in blue and red, respectively; the polar, apolar and aromatic residues are in green, orange and magenta; glycines are in cyan and prolines in yellow. (a) The residues from the three catalytic motifs are in wide sticks, with bold labels. Normal labels indicate residues that have different physicochemical properties in PBP and ßLACTs. (c)-(d) PBP cavity; the residues that have a different physicochemical property than the

corresponding residues in all ßLACTs or all but one ßLACT are in green and salmon, respectively (see Table S5). The ligands present in the structures of ßLACT proteins are superimposed onto the PBP cavity and represented in blue sticks (see Table S1): imipenem (from 4GOG) in dark blue, imipenem (from 4H8R) purple blue, cefotixin (from 1YMX) in navy blue, doripenem (from 3ISG) in light blue, meropenem (from 2ZD8) in very light blue.

With this procedure, we found four PBP cavity residues that have a physicochemical property with no equivalent in any of the ßLACTs: E411, Y507, K510 and V545. As seen in Figs 2.c-d, the first three residues are located at the border of the PBP cavity and their side chains point away from the catalytic site. Therefore, they do not appear as good candidates to be targeted by newly designed ligands that would specifically recognize PBP. This is particularly clear in Figs 2.c-d, in which all ligands appearing in the X-ray structures of ßLACTs are mapped onto the PBP cavity.

Only the fourth residues, Y507, is situated in an internal part of the cavity. Its distance to the closest ligand atoms is about 8 Å. This residue could thus easily interact with some larger, rationally designed, ligand. The equivalent residues in ßLACTs are essentially negatively charged and polar residues, and thus do not present an aromatic moiety.

We also identified 13 cavity residues of PBP whose physicochemical properties differ from the properties of the corresponding residues in all but one ßLACT (Table S5). Most of them are situated at the border of the cavity, with often their side chain pointing outside the cavity (Figs 2.c-d). Only three residues are relatively close to the catalytic site, and are likely to be in contact with a well-designed ligand. These are the three tyrosines Y419, Y511 and Y514; for the latter, it is essentially the OH moiety of the Tyr that is close to the ligand binding site. The equivalent residues of Y419 in ßLACT families are mainly polar or hydrophobic, whereas the corresponding residues of Y511 and Y514 are glycines and hydrophobic or polar amino acids.

We performed a similar analysis by focusing on the capacity of amino acid side chains of being hydrogen bond donors or acceptors rather than on their physicochemical property. More precisely, we searched for PBP cavity positions with a peculiar H-bond feature that is different from that found at the corresponding positions of all ßLACTs, or of all but one ßLACT family.

We found no PBP cavity residues with an H-bond feature that differs from all nine ßLACT families, and five PBP residues with a different H-bond feature in all but one ßLACT family: L476, I512, V545, E411, and R506 (Table S6). The former three have no H-bond donor/acceptor atom on their side chain, whereas the corresponding ßLACT residues do have. The latter two are H-bond acceptor and donor, respectively, and have almost no equivalent ßLACT residues. However, these five residues are situated relatively far from the center of the cavity (Figs 2.c-d), which makes them poor

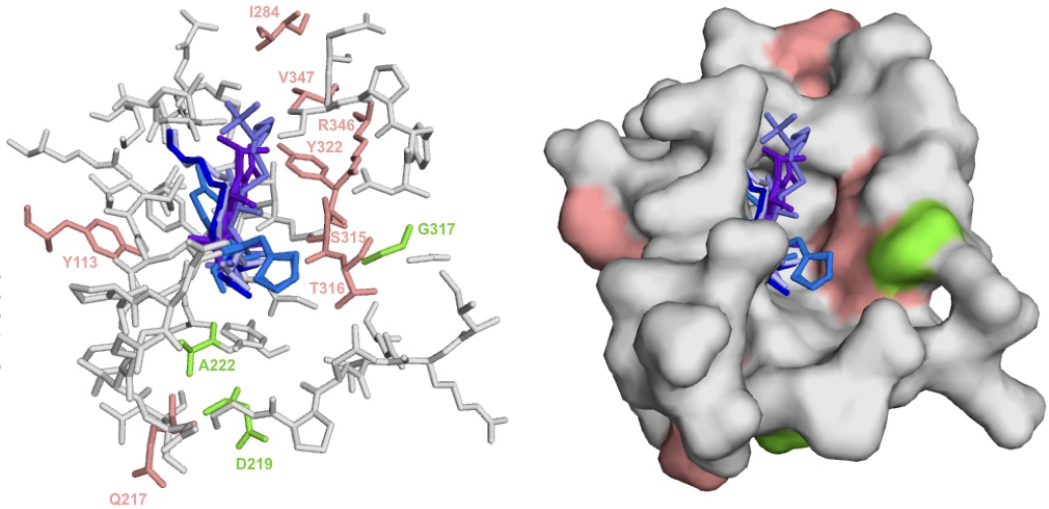
candidates to be targeted by reasonably small ligand molecules.

In conclusion, the residues that should be targeted when designing new ligands specific to PBPs are the four tyrosines Y507, Y419, Y511 and Y514. Note that it is their aromaticity that makes these residues PBP-specific and not their H-bond features. Indeed, some of their equivalent residues in ßLACT families are H-bond donors and/or acceptors just as they are, which explains that they do not appear in Table S6. Hence, a strategy to design new PBP-specific ligands consists in incorporating aromatic moieties that are unable to form H-bonds at well-defined positions of the ligand. Indeed, these moieties are expected to make stabilizing π-π interactions with the tyrosines in PBP and not with the corresponding ßLACT residues.

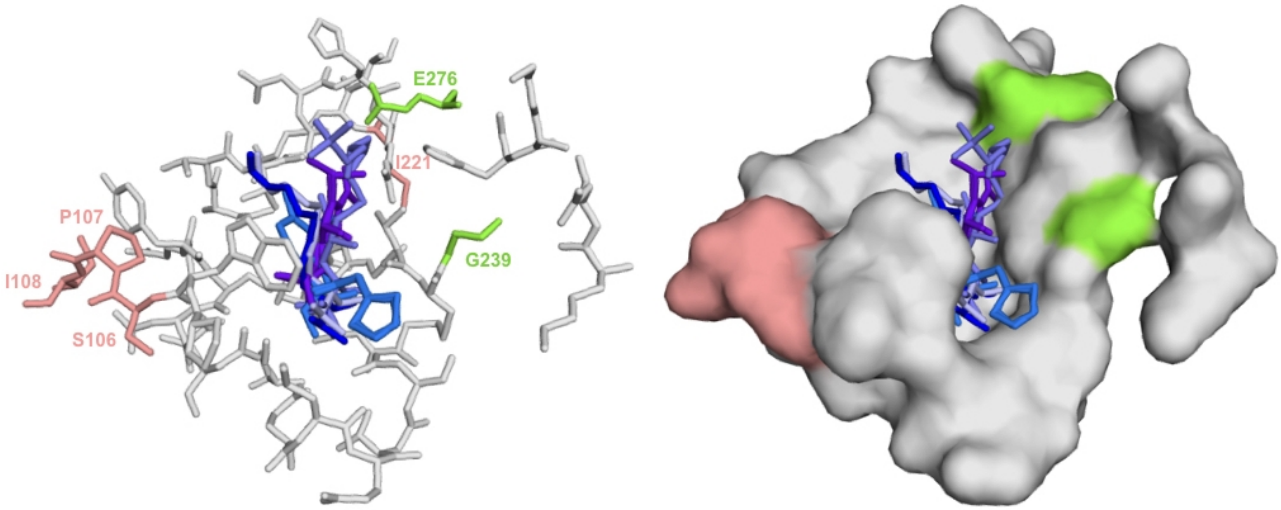
2.4 Cavity characteristics specific to ßLACTs

Our second objective is to find ßLACT cavity residues that share a common property which is different from that of the equivalent PBP residues. Our goal here is to identify features that could be used to build a ligand that has a larger specificity for ßLACTs than for PBPs, and could thus be used to inhibit ßLACTs.

Identifying the spatial correspondences between ßLACT and PBP residues is more difficult with ßLACT as a reference than with PBP, since the results depend on which ßLACT is used. To deal with this issue, we used in turn the representatives of class A, C and D ßLACTs, *i.e.* KPC, CMY and OXA. First, we identified for each residue in the KPC cavity the equivalent residues in the CMY, OXA and PBP cavities, and selected those for which the physicochemical property is the same in all ßLACT or all but one ßLACT, and is different in PBP. We repeated this procedure starting from CMY and OXA as a reference. Then we combined the results obtained from either KPC, CMY or OXA, and selected the residues that were commonly identified in the three runs.



(a) (b)



(c) (d)

Figure 3. Stick (a,c) and surface (b,d) representations of the functional cavities of ßLACTs, with similar spatial orientations as in Fig. 2. (a)-(b) CMY cavity; residues that share an identical physicochemical property in all, or in all but one, ßLACT class, and differ from the corresponding PBP residue properties, are in green and salmon (Table S7). (c)-(d) KPC cavity; ßLACT residues that share an identical H-bond feature in all or all but one ßLACT class, which is different from the features of equivalent PBP residues, are in green and salmon, respectively (Table S8). The ligands present in the structures of ßLACT proteins are superimposed onto the cavities and represented in blue sticks (see legend to Fig. 2 for details).

Three series of equivalent residues were found to share a common physicochemical property among all ßLACT classes which is different in PBP, and six series share a common property in two out of the three ßLACT classes (Table S7 and Figs 3.a-b). However, only four of them have their side chain pointing towards the cavity center. These are residues S315, T316, R346 and G317, using CMY numbering. We chose to discard the latter residue as its absence of side chain prevents it making specific interactions with a ligand. We also discarded the polar residues S315 and T316, as the equivalent residues in KPC are two cysteines, C69 and C238, which form a disulfide bridge. The only residue that we kept is R346, which it is very well conserved in KPC and CMY, but not in OXA where it is replaced by both polar and hydrophobic residues, like in PBP.

We carried out a similar analysis by focusing on the H-bond features that are common to KPC, CMY and OXA ßLACTs and different in PBP (Table S8, Figs 3.c-d). Among the selected residues, we disregarded the KPC residues S106, P107, I108 and I221 and the corresponding residues in CMY and OXA, as they point towards the cavity exterior. Two other series of residues are oriented towards the cavity interior: the KPC residues G239 and E276. We dropped the former as it has no side chain and thus cannot make

specific contacts. The latter, E276, corresponds to N343 in CMY and S258 in OXA. These three residues are H-bond acceptors, while the corresponding PBP residue, A544, is not. We thus selected it as a good candidate for ßLACT-specific ligand design.

In conclusion, the residues that should be targeted for designing ßLACT-specific molecules are N343 in CMY (E276 in KPC, and S258 in OXA) for its H-bond acceptor feature, and R346 in CMY (R220 in KPC) for its positive charge. Note that OXA resembles to the other ßLACTs for the former feature, and to PBPs for the latter. Targeting the positive charge would thus only yield ligands that bind specifically to class A and C ßLACTs.

3. Conclusion

The objective of this paper was to utilize 3D structural information to investigate whether the functional cavities of PBPs are significantly different from the cavities of active-serine ßLACTs, the most important source of resistance to ß-lactam-based antibiotics; a positive outcome means that such antibiotics still have a future, while a negative answer implies that we must definitely turn to other types of targets and drugs. The results obtained in this paper are rather positive, as we were able to detect a few significant differences that can be used for the rational design of modified antibiotics. Note that, although the active site of PBPs and ßLACTs has been well described earlier (Pratt, 2016), a detailed structural analysis of the full functional cavity was lacking.

More specifically, the present analysis led to the detection of the residues that should be targeted by new ß-lactam antibiotics, in view of making them specific for PBP binding and avoiding ßLACT resistance in enterobacteria. As summarized in Fig. 4.a, we identified a region in the PBP cavity that harbors four aromatic Tyr moieties, at positions 419, 507, 511 and 514, whereas the equivalent residues in ßLACTs are quite diverse: polar, negatively charged, small or hydrophobic but almost never aromatic. The aromaticity of this region can thus be exploited to design an extended ß-lactam ligand with one or several additional aromatic groups that make aromatic interactions with the Tyr residues in PBP, hence increasing their affinity for PBP and decreasing their affinity for ßLACTs.

We also identified two residues that share common features in class A and C ßLACTs, and differ in PBPs (Fig. 4.b). These residues occur in a region of the cavity next to the region containing the four tyrosines mentioned in the previous paragraph. It contains the positively charged R346 in CMY and R220 in KPC, whose equivalent residues in OXA and PBP are polar and hydrophobic, and the H-bond acceptors E276 in KPC, N243 in CMY and S258 in OXA, whose equivalent PBP residue is an alanine and thus neither an acceptor nor a donor. Hence, an extended ligand that would preferentially bind to ßLACTs with the aim of inactivating it should target the positive charge (R220 in KPC) and the H-bond acceptor (E276 in KPC).

In conclusion, we found very few residues with different physicochemical properties or H-bond features in ßLACTs and PBPs. The cavities are thus very similar, and it is quite difficult to design either ßLACT- or PBP-specific molecules. We yet identified some residues that are interesting targets for new, rationally extended, ligands, which could result in new drug candidates. These will involve a ß-lactam group, on which specific moieties will be attached at positions that allow them to specifically bind to either PBPs or ßLACTs. The correct positioning of the ligands in the PBP or ßLACT functional cavities will be checked using flexible docking simulations. The next stage will then consist in experimentally testing such new molecules.

Note that we filtered out the residues whose side chains point towards the cavity exterior, to prevent having to design ligand molecules that are too large to be able to penetrate the cells. However, the results of our analysis provide the necessary information to rationally design longer ßLACT- or PBP-specific molecules.

The present analysis is based on information obtained from X-ray structures and thus overlooks the role of protein dynamics. In this specific case, however, dynamics does not seem to play a crucial role in the affinity and specificity of PBPs and ßLACTs for ligands. Indeed, some of the structures analyzed were crystallized with a ligand and others without (Table S1), and the liganded and unliganded forms appear to have very similar structures (Table S4).

The approach described in this paper can be utilized as a general technique to guide in silico drug design. For example, it can be used for designing novel drugs, such as new antibiotic molecules that target PBPs or other bacterial proteins, or ligands for other systems where the specificity of binding to one protein class and not to another is important to prevent resistance or side effects.

As a last conclusion, we would like to stress the importance of taking 3D protein structures into account whenever possible. This is clear in the present analysis, where sequence analyses alone are incapable of detecting the similarities and dissimilarities between the functional cavities of ßLACTs and PBPs. More generally, limiting oneself to sequence information is tantamount to neglecting crucial information and leads to less valuable and reliable results and in silico predictions.

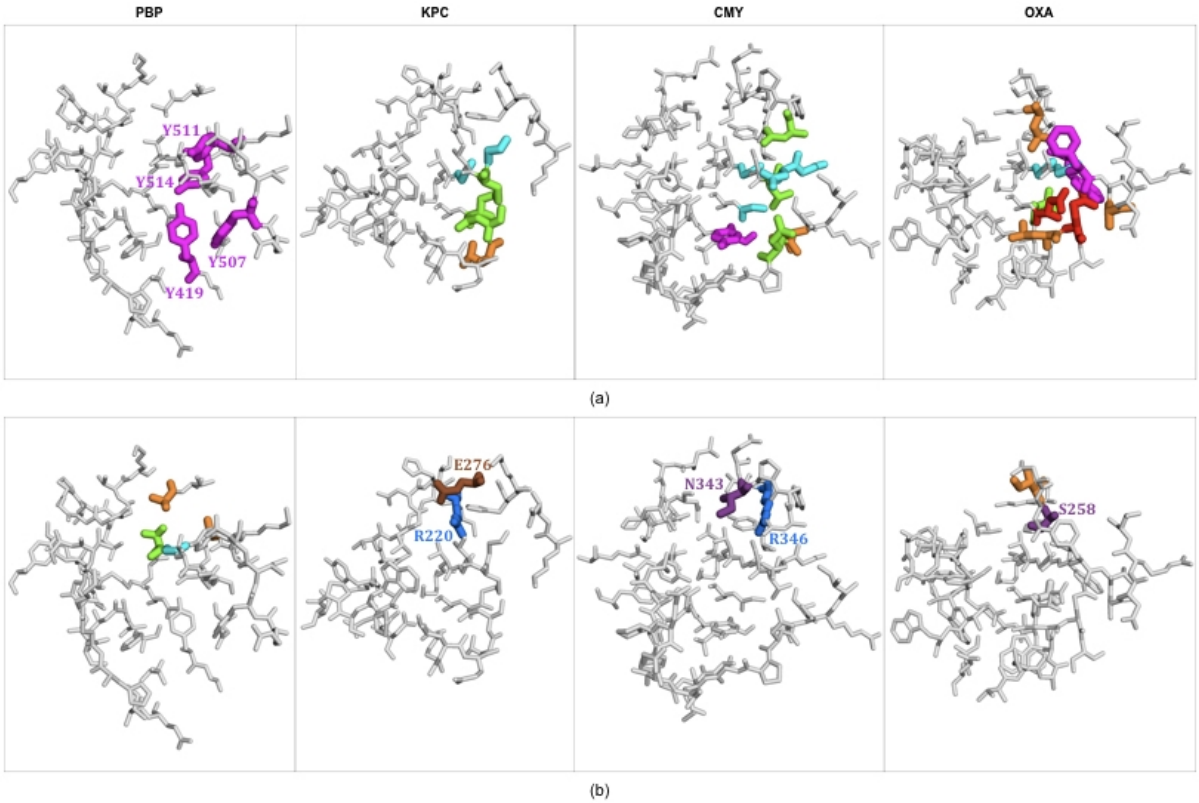


Figure 4. Residues in the functional cavities of PBP and the three ßLACT classes that should be targeted by extended ligands, shown in colored sticks. We used the same color code as in Fig. 2 (a)-(b) for the physicochemical properties; H-bond acceptors and donors/acceptors are in brown and purple, respectively. (a) Residues specific of the PBP cavity (with labels) with their equivalent in the ßLACTs. (b) Residues specific of ßLACT cavities (with labels) with their equivalent in PBP.

4. Methods

Sequence alignments. To compare the different protein sequences, we used the pairwise and multiple sequence alignment algorithm Clustal Omega (Sievers *et al*., 2014).

Structure alignments. For comparison of the different protein structures, we used the structure superimposition tool PDBeFold (Krissinel & Henrick, 2004).

Delimitation of the functional cavities**.** To identify the residues of each protein that belong to the functional cavity, we used the MetaPocket server (Huang 2009; Zhang *et al.* 2011). Among the cavities detected by MetaPocket, we selected the cavity that contains the well-described catalytic motifs from PBP and ßLACT. These are: SxxK, (S|Y)xN and (K|H)(S|T)G, where '|' denotes 'or' and 'x' any residue (Sauvage *et al.*, 2008).

For the ßLACT families for which several X-ray structures are available, we identified

the functional cavity in all the structures, superimposed them, and merged the selected residues. The delimitation of the so obtained consensus cavity was refined through visual inspection using the PyMOL program (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.).

Equivalent cavity residues. To identify residues that occupy equivalent spatial positions in the functional cavities of a set of proteins, we first chose among these cavities a reference cavity to which all other cavities must be compared with. All cavity structures are then superimposed onto the reference cavity. A sphere of 5 Å radius centered around the side chain geometric center of each residue of the reference cavity is then considered. The residues of the superimposed cavities whose side chain geometric center falls inside the sphere are identified and said to occupy equivalent spatial positions. In this way, we detected, for each residue in the reference cavity, the corresponding residue or residues in each of the other cavities.

Here we applied this procedure to detect residues from the ßLACT cavities that occupy equivalent positions to the PBP cavity residues, chosen as reference. Conversely, we also considered each of the representative ßLACT cavities as reference cavity and looked for the corresponding residues in PBP and the other ßLACTs.

Amino acid groups. To analyze the sequence conservation of the functional cavities, we defined amino acid classes according to their physicochemical properties: negatively charged (Glu and Asp), positively charged (Lys, Arg and His), aromatic (Trp, Phe and Tyr), polar (Ser, Thr, Cys, Asn, Gln), apolar (Ala, Val, Leu, Ile, Met); Pro and Gly are considered as separate classes.

Hydrogen bond donors and acceptors. We considered the H-bond donors and acceptors of the main chain which are not involved in an H-bond in the protein structure, and all the side chain H-bond donors and acceptors. Amino acids with an H-bond donor atom on their side chain are Arg (on the atoms NE1, NH1 and NH2), Gln (on NE2), Trp (on NE1), Lys (on NZ), His (on NE2), Asn (on ND2); amino acids with an H-bond acceptor atom on their side chain are Gln (on OE1), His (on ND1), Asn (on OD1), Glu (on OE1 and OE2), and Asp (on OD1 and OD2); amino acids with both a donor and an acceptor atom on their side chain are Tyr (on OH), Ser (on OG) and Thr (on OG1).

Acknowledgments

We thank Jean Marc Kwasigroch for help with the computer programs. MNM has a PhD grant from the Belgian Commission for Cooperation and Development (ARES-CCD). MR is Research Director at the Belgian Fund for Scientific Research (FNRS).

Supplementary material

Table S1. X-ray structures of the considered ßLACTs and PBPs available in the PDB database, with the co-crystallized ligands when present.

Table S2. Amino acids forming the functional cavities of the PBP and ßLACT families.

Table S3. Pairwise sequence identity between representative sequences of all considered ßLACT and PBP families.

Table S4. Root mean square deviation between the representative structures of the considered ßLACT and PBP families after structure alignment.

Table S5. PBP residues whose physicochemical properties differ from those of the corresponding residues in the ßLACT families.

Table S6. PBP residues whose hydrogen bond features differ from those of the corresponding residues in the ßLACT families.

Table S7. Residues that share an identical physicochemical property in the ßLACT classes, which differs from the physicochemical property of the equivalent PBP residue.

Table S8. Residues that share an identical hydrogen bond feature in the ßLACT classes, which differs from the H-bond feature in PBP.

Figure S1. Superposition of the representative structures of ßLACTs of class A, C and D

Figure S2. Stick and surface representations of the cavity residues in the nine ßLACT families.

References

Bairoch, A., Boeckmann, B., Ferro, S., & Gasteiger E. (2004) Swiss-Prot: juggling between evolution and stability. *Briefings in Bioinformatics, 5,* 39-55.

Bebrone, C. (2007). Metallo-ß-lactamases (classification, activity, genetic, organization, structure, zinc coordination) and their superfamily. *Biochemical pharmacology, 4,* 1686-1707.

Berman, H.M., Westbrook, Z., Feng, G., Gilliland, T.N., Bhat T.N., Weissig, I.N., Shindyalov P.E., & Bourne P.E. (2000). The Protein Data Bank. *Nucleic Acids Research, 28,* 235-242.

Bonomo, R.A. (2017). ß-Lactamases: A focus on current challenges. *Cold Spring Harbor Perspectives in Medicine*, 7.

Bush, K., & Bradford, P.A. (2016). ß-Lactams and ß-lactamase Inhibitors: an overview. *Cold Spring Harbor Perspectives in Medicine*, 6.

Bush, K., & Jacoby, G.A. (2010). Updated functional classification of beta-lactamases. *Antimicrobial Agents and Chemotherapy, 54*, 969-76.

Cag, Y., Caskurlu, H., Fan, Y., Cao, B., & Vahaboglu, H. (2016). Resistance mechanisms. *Annals of Translational Medicine, 4,* 326.

Derouaux, A., Wolf, B., Fraipont, C., Breukink, E., Nguyen-Distèche, M., Terrak, M. (2008) The monofunctional glycosyltransferase of Escherichia coli localizes to the cell division site and interacts with penicillin-binding protein 3, FtsW, and FtsN. *Journal of Bacteriology, 190,* 1831-4.

Fisher, J.F., Meroueh, S.O., & Mobashery, S. (2005). Bacterial resistance to beta-lactam antibiotics: compelling opportunism, compelling opportunity. *Chemical Reviews, 105*, 395-424.

Frère, J.M., & Page, M.G. (2014). Penicillin-binding proteins: evergreen drug targets. *Current Opinion in Pharmacology, 18*, 112-9.

Ghuysen, J.-M. (1991). Serine β-lactamases and penicillin-binding proteins. *Annual review Microbiology, 45*, 37–67.

Guentzel, M.N. (1996). Escherichia, Klebsiella, Enterobacter, Serratia, Citrobacter, and Proteus. In: Baron S, editor. Medical Microbiology. 4th edition. Galveston (TX): University of Texas Medical Branch at Galveston.

Hirsh, H.L. (1948). Penicillin; a review of the basic principles and their clinical application. *Medical Annals of the District of Columbia, 17*, 7-20.

Huang, B. (2009) MetaPocket: a meta approach to improve protein ligand binding site prediction. *OMICS, 13*, 325-30.

Iredell, J., Brown, J., & Tagg, K. (2016). Antibiotic resistance in Enterobacteriaceae: mechanisms and clinical implications. *British Medical Journal*, *352*, h6420.

Krissinel, E., & Henrick, K. (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta crystallographica*, *D60*, 2256-2268.

Leinonen, R., Diez, F.G., Binns, D., Fleischmann, W., Lopez, R., & Apweiler R. (2004) UniProt Archive. *Bioinformatics, 20*, 3236-3237.

Livermore, D.M. (2009). Has the era of untreatable infections arrived? *Journal of Antimicrobial Chemotherapy, 64 Suppl 1*, 29-36.

Page, M.G.P., Bush, K. (2014) Discovery and development of new antibacterial agents targeting Gram-negative bacteria in the era of pandrug resistance: is the future promising? *Current Opinion in Pharmacology, 18*, 91-97.

Pfeifle, D., Janas, E., & Wiedeman, B. (2000). Role of Penicillin-Binding Proteins in the Initiation of the AmpC β-Lactamase Expression in *Enterobacter cloacae. Antimicrobial agents and chemotherapy, 44*, 169-172.

Pratt, R.F. (2016). β‑Lactamases: Why and How. *Journal of Medicinal Chemistry. 2016, 59*, 8207-20.

Qin, W., Panunzio, M., & Biondi, S. (2014). β-Lactam antibiotics renaissance.

Antibiotics, 3, 193−215.

Ruggiero, M., Curto, L., Brunetti, F., Sauvage, E., Galleni, M., Power, P., Gutkind, G. (2017). Impact of Mutations at Arg220 and Thr237 in PER-2 β-Lactamase on Conformation, Activity, and Susceptibility to Inhibitors. Antimicrobial Agents and Chemotherapy, 61.

Sauvage, E., Kerff, F., Terrak, M., Ayala, J.A., Charlier, P. (2008). The penicillin-binding proteins: structure and role in peptidoglycan biosynthesis. *FEMS Microbiology Reviews, 32*, 234-58.

Sievers, F., & Higgins, D.G. (2014). Clustal omega. *Current Protocols in Bioinformatics, 48*, 3.13.1-16.

Van Boeckel, T.P., Gandra, S., Ashok, A., Caudron, Q., Grenfell, B.T., & Levin, S.A. (2014). Laxminarayan R. Global antibiotic consumption 2000−2010: An analysis of national pharmaceutical sales data. *Lancet Infectious Diseases, 14,* 742−750.

Wise, E.M., & Park, J.T. (1965). Penicillin: its basic site of action as an inhibitor of a peptide cross-linking reaction in cell wall mucopeptide synthesis. *Proceedings of the National Academy of Sciences U.S.A., 54*, 75−81.

Zhang, Z., Li, Y., Lin, B., Schroeder, M., & Huang, B. (2011). Identification of cavities on protein surface using multiple computational approaches for drug binding site prediction. *Bioinformatics*, *27*, 2083-2088.

Supplementary Material

Rational antibiotic design: in silico structural comparison of the functional cavities of penicillin-binding proteins and ß-lactamases

Mame Ndew Mbaye, Dimitri Gilis, Marianne Rooman

3BIO-BioInfo group, Université Libre de Bruxelles, CP 165/61, avenue F. Roosevelt 50, 1050 Bruxelles, and Université Cheikh Anta Diop de Dakar, Senegal

Table S1: X-ray structures of the considered ßLACTs and PBPs available in the PDB database, with the co-crystallized ligands when present.

Family	PDB code	Resolution (Å)	R-factor	Representative structure	Ligands
Penicillin Binding Proteins					
PBP3	4bjp	2.5	0.245	✓	–
ß-Lactamases					
KPC (Class A)	3c5a	1.23	0.196	✓	–
	3rxw	1.26	0.171		–
	2ov5	1.85	0.190		–
TEM (Class A)	1m40	0.85	0.112	✓	–
	1yt4	1.4	0.223		–
	1zg4	1.55	0.240		–
	1pzo	1.9	0.247		–
	1pzp	1.45	0.245		–

	1lhy	2.0	0.212		–
	1lio	2.5	0.260		–
	1li9	1.52	0.189		–
	1jwz	1.8	0.189		–
	1htz	2.4	0.261		–
	3gmw	2.1	0.234		–
	3p98	2.1	0.283		–
SHV (Class A)	1n9b	0.9	0.150	✓	–
	2zd8	1.05	0.166		meropenem
	4jpm	1.14	0.164		–
	1ong	1.1	0.186		–
	4gd6	1.53	0.194		–
GES (Class A)	4gog	1.1	0.168	✓	imipenem
	4h8r	1.25	0.170		imipenem
CTX (Class A)	4hbt	1.1	0.136	✓	–
	1ylw	1.74	0.206		–
	1yly	1.25	0.168		–
	1ylz	1.35	0.129		–
	1ym1	1.12	0.140		–
	1yms	1.60	0.193		–
	1ymx	1.70	0.197		cefoxitin
	3g2y	1.31	0.188		–
	3g2z	1.50	0.188		–
	3g30	1.80	0.245		–
	3g31	1.70	0.203		–

	3g32	1.31	0.182		–
	3g34	1.31	0.180		–
	3g35	1.41	0.182		–
	4dds	1.36	0.199		–
	4ddy	1.36	0.205		–
	4de0	1.12	0.166		–
SME (Class A)	1dy6	2.13	0.244	✓	-
PER (Class A)	4d2o	2.2	0.240	✓	–
CMY (Class C)	1zkj	1.55	0.235	✓	–
	1zc2	2.09	0.299		–
OXA (Class D)	1m6k	1.5	0.203	✓	–
	1k38	1.5	0.211		–
	3hbr	1.9	0.286		–
	3isg	1.4	0.206		doripenem
	4wmc	2.3	0.273		–

Table S2. Amino acids forming the functional cavities of the PBP and ßLACT families. The catalytic motifs are on an orange background. The residue numbering is that of the PDB file of the representative structure indicated in Table S1. The second row of each cell contains the amino acids in one letter code that are found at that position in the multiple sequence alignment of the proteins of the (sub)family, with "-" indicating an insertion. The frequency of occurrence of each residue is indicated as a subscript; only amino acids appearing in at least 1% of the sequences are included.

PBP3	KPC	TEM	SHV	GES	CTX	SME	PER	CMY	OXA
GLU-304 E_{100}	CYS-69 C_{100}	MET-69 $M_{90}L_4V_1I_1$	MET-69 $M_{98}I_{1\text{-}1}$	CYS-63 C_{100}	CYS-69 $C_{83}G_1M_{1\text{-}16}$	CYS-69 C_{100}	GLN-69 Q_{100}	GLU-62 E_{100}	ASP-66 $D_{81}C_{11}A_8$
GLY-306 G_{100}	SER-70 S_{100}	SER-70 $S_{97\text{-}3}$	SER-70 $S_{99\text{-}1}$	SER-64 S_{100}	SER-70 $S_{85\text{-}15}$	SER-70 S_{100}	SER-70 S_{100}	GLY-64 G_{100}	SER-67 S_{100}
SER-307 S_{100}	LYS-73 K_{100}	PHE-72 $F_{97}L_{1\text{-}2}$	LYS-73 K_{100}	LYS-67 K_{100}	LYS-73 $K_{86\text{-}13}$	LYS-73 K_{100}	PHE-72 F_{100}	SER-65 S_{100}	THR-68 $T_{89}S_{11}$
LYS-310 K_{100}	PRO-104 $P_{85}R_{15}$	LYS-73 $K_{97\text{-}2}$	ASP-104 D_{99}	GLU-98 $E_{56}K_{44}$	ASN-104 $N_{93}D_{1\text{-}6}$	TYR-104 Y_{100}	LYS-73 K_{100}	LYS-68 K_{100}	PHE-69 F_{100}
HIS-339 $H_{99}K_1$	TRP-105 $W_{96}R_2G_2$	GLU-104 $E_{87}K_{10}D_2$	TYR-105 Y_{100}	TRP-99 W_{100}	TYR-105 $Y_{94\text{-}6}$	HIS-105 $H_{80}Y_{20}$	THR-104 T_{100}	TYR-113 Y_{100}	KCX-70 K_{100}
GLU-340 $E_{93}Q_6L_1$	SER-106 S_{100}	TYR-105 $Y_{99\text{-}1}$	MET-129 $M_{97}V_3$	SER-100 S_{100}	TYR-129 $Y_{95}M_{1\text{-}4}$	SER-106 S_{100}	TRP-105 W_{100}	LEU-120 L_{100}	GLY-98 $G_{81}Y_{11}D_8$
ILE-341 I_{100}	PRO-107 P_{100}	SER-130 $S_{99}G_1$	SER-130 S_{100}	PRO-101 P_{100}	SER-130 $S_{96\text{-}4}$	PRO-107 P_{100}	SER-106 $A_{81}S_{19}$	GLN-121 Q_{100}	MET-99 $M_{81}N_{11}I_8$
LYS-342 $K_{98}R_2$	ILE-108 I_{100}	ASN-132 N_{100}	ASN-132	GLN-123	ASN-132 $N_{96\text{-}3}$	LEU-127	PRO-107	PHE-122 $I_{97}F_3$	ILE-101

			N_{100}	Q_{100}		L_{100}	P_{100}		$I_{81}R_{11}T_8$
ASP-343 D_{100}	ALA-126 A_{100}	GLU-166 $E_{99\text{-}1}$	GLU-166 E_{99}	LEU-124 L_{100}	GLU-166 $E_{95\text{-}5}$	TYR-129 Y_{100}	SER-128 S_{100}	PRO-123 P_{100}	TRP-102 $W_{89}T_{11}$
VAL-344 V_{100}	GLN-128 Q_{100}	PRO-167 $P_{97}T_{2\text{-}1}$	THR-167 T_{100}	SER-125 S_{100}	PRO-167 $P_{91}S_2T_{1\text{-}5}$	SER-130 S_{100}	HIS-129 H_{100}	GLU-124 $D_{98}E_2$	THR-110 $T_{89}L_{11}$
ALA-345 $A_{98}G_1$	TYR-129 Y_{100}	GLU-168 $E_{98\text{-}1}$	GLU-168 $E_{99}A_1$	ASN-127 N_{100}	THR-168 $T_{94}E_{1\text{-}5}$	ASN-132 N_{100}	SER-130 $S_{88}T_{12}$	GLU-125 $D_{97}Y_2E_2$	TRP-111 $W_{92}A_8$
TYR-347 $Y_{99}R_1$	SER-130 S_{100}	LEU-169 $L_{99\text{-}1}$	LEU-169 L_{99}	GLU-161 E_{100}	LEU-169 $L_{95\text{-}5}$	GLU-166 E_{100}	ASN-132 N_{100}	VAL-126 $V_{95}I_5$	MET-112 $M_{89}Q_{11}$
LEU-356 L_{99}	ASP-131 D_{100}	ASN-170 $N_{98\text{-}2}$	ASN-170 N_{100}	PRO-162 P_{100}	ASN-170 $N_{95\text{-}5}$	LEU-167 L_{100}	ASP-136 D_{100}	ASP-127 $R_{74}T_{24}D_2$	GLN-113 $Q_{81}S_{11}K_8$
GLN-357 Q_{99}	ASN-132 N_{100}	GLU-171 $E_{98\text{-}2}$	GLU-171 E_{100}	GLU-163 E_{100}	THR-171 $T_{93}S_1E_{1\text{-}5}$	GLU-168 E_{100}	GLU-166 E_{100}	ARG-149 $R_{98}H_2$	PHE-114 $F_{81}D_{11}Y_8$
LYS-358 K_{99}	GLU-166 E_{100}	ASP-214 $D_{96\text{-}4}$	ASP-214 D_{99}	MET-164 M_{100}	ALA-172 $A_{95\text{-}5}$	ASN-170 N_{100}	ALA-167 A_{100}	TYR-151 Y_{100}	SER-115 S_{100}
SER-359 S_{99}	LEU-167 L_{100}	LYS-215 $K_{95}R_{1\text{-}4}$	ARG-215 $R_{98}Q_1$	GLY-165 $G_{56}S_{44}$	ILE-173 $I_{94}L_{1\text{-}5}$	THR-171 T_{100}	GLN-168 Q_{100}	ASN-153 N_{100}	VAL-116 V_{100}
ASN-361 $N_{99}D_1$	GLU-168 E_{100}	VAL-216 $V_{96\text{-}4}$	VAL-216 V_{100}	ASP-166 D_{100}	THR-215 $T_{91}R_{1\text{-}8}$	ASN-214 N_{100}	MET-169 M_{100}	ILE-211 $V_{98}I_2$	VAL-117 V_{100}
ILE-410 $I_{93}L_{3\text{-}1}$	LEU-169 $L_{96}M_4$	ALA-217 $A_{93}Q_{3\text{-}4}$	ALA-217 A_{100}	ASN-167 N_{100}	THR-216 $T_{91}V_{1\text{-}8}$	THR-215 T_{100}	HIS-170 H_{100}	ARG-212 $H_{97}Y_2R_2$	TRP-118 $W_{92}P_8$
GLU-411 $E_{97}D_{2\text{-}1}$	ASN-170 N_{100}	GLY-218 $G_{92}D_{3\text{-}4}$	GLY-218 G_{100}	GLN-210 Q_{100}	GLY-217 $G_{90}A_{1\text{-}9}$	THR-216 T_{100}	ALA-171 A_{100}	VAL-213 $V_{94}A_3S_2G_2$	SER-120 $S_{92}Y_8$
THR-414 $T_{99\text{-}1}$	ASN-214 N_{100}	PRO-219 $P_{93}H_{2\text{-}4}$	PRO-219 P_{100}	THR-211 T_{100}	ALA-218 $A_{83}S_7G_{1\text{-}10}$	GLY-217 G_{100}	ASP-172 D_{100}	ASN-214 $S_{97}T_2N_2$	GLN-121 $Q_{99}H_1$
PHE-417 $F_{97}Y_{2\text{-}1}$	THR-215 T_{100}	LEU-220 $L_{94}Y_{1\text{-}5}$	LEU-220 $L_{99\text{-}1}$	GLY-212 G_{100}	ALA-219 $A_{89}P_{1\text{-}10}$	ASP-218 D_{100}	ASP-173 D_{100}	PRO-215 P_{100}	THR-157 $T_{92}D_8$
TYR-419 $Y_{99\text{-}1}$	THR-216 T_{100}	LYS-234 $K_{90\text{-}10}$	ILE-231 $I_{99\text{-}1}$	ASP-213 D_{100}	SER-220 $S_{87}R_1L_{1\text{-}11}$	ALA-219 A_{100}	GLN-174 Q_{100}	GLY-216 $G_{98}E_2$	GLU-158 $E_{81}Q_{11}S_8$
GLY-420 $G_{99\text{-}1}$	GLY-217 G_{100}	SER-235 $S_{87}T_{2\text{-}10}$	ALA-232 $A_{99\text{-}1}$	ALA-214 A_{100}	LYS-234 $K_{83\text{-}17}$	ARG-220 R_{100}	VAL-175 V_{100}	MET-217 $Q_{94}R_3M_2K_2$	ALA-159 $A_{81}S_{11}F_8$
VAL-474 $V_{99\text{-}1}$	ASN-218 N_{100}	GLY-236 $G_{90\text{-}10}$	ASP-233	THR-215 T_{100}	THR-235 $T_{83\text{-}17}$	LYS-234 K_{100}	GLN-176 Q_{100}	LEU-218 $L_{98}F_2$	TRP-160 W_{100}

			$D_{99\text{-}1}$						
ALA-475 $A_{99\text{-}1}$	HIS-219 H_{100}	ALA-237 $A_{87}T_{2\text{-}10}$	LYS-234 $K_{98}R_{1\text{-}1}$	LYS-229 K_{100}	GLY-236 $G_{82\text{-}17}$	THR-235 T_{100}	TYR-177 Y_{100}	ALA-219 $D_{98}A_2$	LEU-161 L_{100}
LEU-476 $L_{99\text{-}1}$	ARG-220 R_{100}	GLY-238 $G_{81}S_{8\text{-}10}$	THR-235 $T_{96}N_{2\text{-}1}$	THR-230 T_{100}	SER-237 $S_{81}A_{1\text{-}17}$	GLY-236 G_{100}	THR-214 T_{100}	ALA-222 A_{100}	GLU-162 $E_{81}M_{11}D_8$
PRO-477 $P_{99\text{-}1}$	ILE-221 I_{100}	GLU-240 $E_{81}K_7N_{1\text{-}11}$	GLY-236 $G_{98\text{-}2}$	GLY-231 G_{100}	GLY-238 $G_{81}C_1S_{1\text{-}17}$	SER-237 S_{100}	THR-215 T_{100}	TYR-223 Y_{100}	SER-163 $S_{92}G_8$
GLY-478 $G_{99\text{-}1}$	ARG-222 R_{100}	ARG-241 $R_{89\text{-}11}$	ALA-237 $A_{98\text{-}2}$	THR-232 T_{100}	GLY-239 $D_{42}G_{38}K_{1\text{-}18}$	CYS-238 C_{100}	THR-216 T_{100}	MET-265 M_{100}	LEU-200 $L_{81}Q_{11\text{-}8}$
GLY-479 $G_{99\text{-}1}$	ALA-223 A_{100}	GLY-242 $G_{88\text{-}11}$	SER-238 $G_{56}S_{40}A_1$	CYS-233 C_{100}	TYR-240 $Y_{81}R_{1\text{-}18}$	GLY-239 G_{100}	GLY-217 G_{100}	GLU-272 E_{100}	GLY-211 $G_{92}A_8$
GLY-480 $G_{99\text{-}1}$	LYS-234 K_{100}	SER-243 $S_{87}A_{1\text{-}12}$	GLU-240 $E_{65}K_{32\text{-}2}$	ALA-234 A_{100}	GLY-241 $G_{80\text{-}10}$	ALA-240 A_{100}	PRO-218 P_{100}	TYR-274 $L_{98}Y_2$	LYS-212 K_{100}
GLY-481 $G_{99\text{-}1}$	THR-235 T_{100}	ARG-244 $R_{85}S_2N_{1\text{-}17}$	ARG-241 $R_{98\text{-}2}$	ASN-235 N_{100}	THR-242 $T_{79}A_{1\text{-}20}$	TYR-241 $Y_{80}I_{20}$	GLN-219 $E_{81}Q_{19}$	LEU-284 $I_{98}L_2$	THR-213 $T_{89}S_{11}$
VAL-482 $V_{91}T_{7\text{-}1}$	GLY-236 G_{100}	ILE-246 $I_{87\text{-}13}$	ARG-244 $R_{97\text{-}3}$	GLY-236 G_{100}	THR-243 $T_{79}R_{1\text{-}20}$	GLY-242 G_{100}	ARG-220 R_{100}	ALA-285 $N_{98}A_2$	GLY-214 G_{100}
LYS-483 $K_{95}R_{3\text{-}1}$	THR-237 T_{100}	ALA-270 $A_{82}G_{1\text{-}17}$	GLY-245 $G_{97\text{-}3}$	GLY-237 $G_{89}A_{11}$	ASN-244 $N_{78}G_{1\text{-}21}$	THR-243 T_{100}	HIS-233 H_{100}	GLY-286 G_{100}	ALA-215 $A_{81}S_{11\text{-}8}$
LYS-494 $K_{99\text{-}1}$	CYS-238 C_{100}	THR-271 $T_{79}S_2N_{1\text{-}17}$	ILE-246 $I_{97\text{-}3}$	ARG-238 R_{100}	GLN-265 $G_{68}D_{1\text{-}31}$	LYS-270 K_{100}	LYS-234 K_{100}	ASN-287 $S_{98}N_2$	GLY-216 $G_{91}E_{1\text{-}8}$
THR-495 $T_{99\text{-}1}$	GLY-239 G_{100}	MET-272 $M_{80}Y_{1\text{-}18}$	SER-271 $S_{86}T_2E_{1\text{-}11}$	ALA-264 A_{100}	PRO-268 $P_{38}Q_{28}A_{1\text{-}33}$	ASP-271 D_{100}	THR-235 T_{100}	ILE-292 $A_{98}I_2$	PHE-217 $F_{81}W_{11}Y_8$
GLY-496 $G_{99\text{-}1}$	LYS-270 K_{100}	ARG-275 $R_{77}Q_1L_{1\text{-}21}$	MET-272 $M_{88\text{-}11}$	VAL-265 V_{100}	LYS-269 $K_{44}N_{21}S_{1\text{-}34}$	ASP-272 D_{100}	GLY-236 G_{100}	LEU-293 $L_{98}S_3F_{3\text{-}3}$	THR-219 $T_{81}L_{11}S_8$
THR-497 $T_{99\text{-}1}$	ASP-271 D_{100}	ASN-276 $N_{76}D_{2\text{-}21}$	ALA-273 $A_{87}G_{1\text{-}11}$	ARG-267 R_{100}	ALA-270 $A_{65}M_{1\text{-}34}$	LYS-273 K_{100}	THR-237 T_{100}	ASN-311 $S_{98}N_2$	ALA-220 $A_{82}R_{11}T_7$
ALA-498 $A_{98\text{-}1}$	LYS-273 K_{100}		ARG-275 $R_{87\text{-}12}$	ASP-268 D_{100}	GLU-271 $E_{64}A_{1\text{-}35}$	HIS-274 H_{100}	SER-238 S_{100}	LYS-312 K_{100}	ASN-221 $N_{91}R_7G_1Y_1$
LYS-499 $K_{98}R_{1\text{-}1}$	HIS-274 $H_{74}Y_{26}$		ASN-276 $N_{86\text{-}13}$		SER272 $S_{59}R_4N_{1\text{-}35}$	ASP-276 D_{100}	GLY-239 G_{100}	THR-313 T_{100}	ARG-222 $R_{81}E_{11}I_8$
ARG-506	GLU-		ILE-279		SER-274		ARG-	GLY-314	THR-223

$R_{83}K_{13}Q_1$	276 E_{100}		$I_{80\text{-}19}$		$R_{61}H_1N_{1\text{-}37}$		240 $K_{81}R_{19}$	G_{100}	$T_{81}S_{11}E_8$
TYR-507 $Y_{97}F_{1\text{-}1}$					LEU-277 $L_{60}I_{1\text{-}40}$		THR-241 T_{100}	SER-315 S_{100}	TRP-228 W_{100}
ILE-508 $I_{85}V_8M_5$							ALA-242 A_{100}	THR-316 T_{100}	PHE-229 $F_{92}W_8$
ASN-509 $N_{91}D_5S_1$							THR-244 T_{100}	ASN-317 $G_{98}N_2$	ASN-254 $N_{76}D_{11}A_8E_1$
LYS-510 $K_{87}R_{10}Q_1E_1$							LYS-266 K_{100}	GLY-318 G_{100}	LEU-255 $L_{76}I_{19}F_{1\text{-}4}$
TYR-511 $Y_{98\text{-}1}$							ASP-267 D_{100}	TYR-322 Y_{100}	SER-258 $S_{76}G_{11}E_{8\text{-}4}$
ILE-512 $I_{92}L_5V_{1\text{-}1}$							SER-268 S_{100}	ASN-340 $S_{95}N_5$	ILE-259 $I_{76}S_{11}V_8H_1$
ALA-513 $A_{96}S_{2\text{-}1}$							GLU-272 $S_{81}E_{19}$	TYR-341 Y_{100}	LYS-262 $K_{76}R_{11}Q_{9\text{-}4}$
TYR-514 $Y_{98\text{-}1}$							ARG-273 R_{100}	PRO-342 P_{100}	
ALA-544 $A_{93}G_4S_{1\text{-}2}$							ASN-275 N_{100}	ILE-343 $N_{98}I_2$	
VAL-545 $V_{95}I_{3\text{-}2}$							GLU-276 E_{100}	GLU-344 $P_{98}E_2$	
								ARG-346 R_{100}	
								ILE-347 $V_{98}I_2$	

Table S3. Pairwise sequence identity (in %) between representative sequences of all considered ßLACT and PBP families. The UniProt id of the sequences are mentioned in Table 1. The grey cells indicate class A ßLACTs.

	PBP3	KPC	TEM	SHV	GES	CTX	SME	PER	CMY	OXA
PBP3	0	8.5	9.5	11.1	9.6	10.0	9.7	4.2	12.0	8.3
KPC		0	38.2	39 .2	35.7	46.3	54.3	25.6	10.5	4.6
TEM			0	64.2	34.4	35.6	34.3	23.0	8.9	16.1
SHV				0	33.9	37.4	36.1	23.8	20.7	16.7
GES					0	33.3	36.5	24 .2	7.8	17.5
CTX						0	42.7	23.7	14.6	16.4
SME							0	26.1	15.1	1.9
PER								0	14.7	14.3
CMY									0	12.3
OXA										0

Table S4. Root mean square deviation (in Å) between the representative structures of the considered ßLACT and PBP families after structure alignment. The number of aligned residues over which the rms deviation is computed are given in parentheses. The grey cells indicate class A ßLACTs.

	PBP3	KPC	TEM	SHV	GES	CTX	SME	PER	CMY	OXA
PBP3	0	3.00 (195)	2.72 (210)	2.69 (191)	2.99 (204)	3.32 (207)	3.02 (200)	3.17 (216)	2.68 (195)	1.92 (212)
KPC		0	1.30 (243)	1.50 (242)	1.17 (244)	1.01 (258)	0.77 (257)	1.75 (234)	2.82 (190)	2.84 (192)
TEM			0	0.96 (256)	1.28 (257)	1.40 (252)	1.57 (251)	1.82 (246)	2.35 (185)	2.94 (214)
SHV				0	1.45 (253)	1.41 (243)	1.50 (242)	1.82 (235)	2.84 (197)	2.82 (195)
GES					0	1.09 (246)	1.23 (251)	1.76 (249)	2.88 (197)	2.81 (201)
CTX						0	1.06 (257)	1.73 (245)	2.68 (193)	2.70 (194)
SME							0	1.94 (248)	2.89 (199)	2.97 (197)
PER								0	2.43 (191)	2.97 (200)
CMY									0	3.07 (195)
OXA										0

.

Table S5. PBP residues whose physicochemical properties (as defined in Methods) differ from those of the corresponding residues in the 9 ßLACT families (in green) or in 8 of the 9 ßLACT families (in salmon). The amino acids occurring at a given position in the multiple sequence alignments of the family are given in parentheses; the most frequent amino acid is given first, followed by the others with their frequency (in %) as a subscript. For assessing the physicochemical properties, only ßLACT residues whose frequency exceeds 5%, and PBP amino acids whose frequency exceeds 75%, were considered. The residue number is the number from the representative PDB structure. The distance (in Å) is computed between the average side chain centroid of the PBP and ßLACT residues; this distance is given in parentheses.

PBP3	KPC	TEM	SHV	GES	CTX	SME	PER	CMY	OXA
$(H\|K_1)339$	E168 (4.3)	E168 (4.5)	$(E\|A_1)168$ (4.4)	E163 (4.8)			A171 (3.3)	$(Q\|R_3\|K_2\|M_2)217$ (3.7) $(G\|E_2)216$ (4.2) $(R\|T_{24}\|D_2)127$ (4.5)	
D343	P107 (3.7) D131 (4.5) S130 (5.0)	Y105 (3.0)	Y105 (2.6) N132 (5.0)	W99 (3.0) P101 (3.4)	Y105 (3.7) N132 (4.9)	$(H\|Y_{20})105$ (2.2) P107 (3.6)	W105 (2.8) P107 (3.7) $(S\|A_{19})106$ (4.5)	Y151 (4.0)	$(W\|T_{11})102$ (2.1) V116 (3.5) V117 (4.9)

(I|L_5)410	E168 (4.4)	E168 (4.3)	(E|A_1)168 (4.6)	E163 (3.7)		E168 (4.7)	A167 (2.4) Q168 (4.8)	(Q|R_3|K_2|M_2) 217 (2.5) (G|E_2)216 (4.3)	(T|D_8)157 (4.0) (E|Q_{11}|S_8)158 (4.4)
(E|D_2)411							Q168 (2.8)		
Y419	N170 2.6) (L|M_4)169 (3.2) C69 (4.6) C238 (4.6)	N170 (2.3) L169 (3.7)	N170 (2.8) L169 (3.2)	(G|S_{44})165 (1.4) M164 (3.3) C63 (4.2) C233 (4.8)	L169 (3.0) N170 (3.2) (C|M_1|G_1)69 (4.4)	N170 (3.2) C69 (5.0)	M169 (3.1) Q69 (3.2) Q176 (3.9)	Y223 (3.4) (V|A_3|S_2|G_2)213 (4.1) G64 (4.3) T316 (4.7)	L161 (3.4) (D|C_{11}|A_8)66 (4.2) (E|M_{11}|D_8)162 (4.5)
G420							V175 (4.7)		(A|S_{11}|F8)159 (3.6) (S|G_8)163 (3.8)
L476								(L|Y_2)274 (3.9)	
P477			P219 (4.2)						
G478	N218 (3.7)	(K|R_1)215 (4.2)	P219 (2.2)	D213 (3.5)	(T|R_1)215 (4.0)	D218 (4.1)	T215 (2.1)	G286 (3.9) M265 (4.2)	(Q|S_{11}|K_8)113 (4.4)
(K|R_3)483	H219 (3.6) A223 (3.9)	(P|H_2)219 (4.0)	L220 (4.1)	A214 (3.3)	(A|P_1)219 (2.6)	A219 (2.6)	(Q|E_{19})219 (1.2)	(I|L_2)284 (1.9) (V|I_2)347 (5.0)	
(R|K_{13}|Q_1)506									(N|R_7|G_1|Y1) 221 (3.5)
(Y|F_1)507		E171 (2.6) (E|K_7|N_1)240 (4.7)	E171 (1.7)	D166 (2.9) A234	(T|S_1|E_1)171 (1.3) (I|L_1)173	T171 (0.9)	D173 (3.1) Q176	(S|N_2|T_2)214 (3.3) (V|A_3|S_2|G_2)2	(A|R_{11}|T_7)220 (1.7)

				(4.9)	(4.0)		(4.2) Q174 (4.6)	13 (3.6)	
(K|R_{10}|Q_1)510					(Y|R_1)240 (4.6) (P|Q_{28}|A_1)268 (4.7)	(Y|I_{20})241 (4.3)			(T|S_{11}|E_8)223 (3.6)

Y511	G239 (2.1)	G242 (3.0)	(G|S_{40}|A_1)238 (3.8) (E|K_{32})240 (4.6)	G236 (3.1) A234 (4.7) R267 (5.0)	(G|C_1|S_1)238 (4.5) (D|G_{38}|K_1)239 (2.8) G241 (2.8)	A240 (3.3) G239 (3.5) G242 (3.8)	T241 (3.1) G239 (4.0) (R|K_{19})240 (4.0) S238 (4.2)	(S|N_5)340 (3.4) (G|N_2)317 (4.2) G318 (5.0)	(F|W_{11}|Y_8)217 (1.9) (G|E_1)216 (4.6) (L|I_{19}|F_1)255 (4.8)
(I|L_5|V_1)512		(S|A_1)243 (4.3)		(G|A_{11})237 (4.0) N167 (5.0)	(T|A_1)242 (3.8)	T243 (3.3)	K266 (3.8) D267 (4.7)		
Y514	G236 (3.9) C69 (4.4)	(M|L_4|V_1|I1)69 (3.2) G236 (3.9)	G245 (2.6) (M|I_1)69 (3.6) G236 (4.0)	G231 (4.2) C63 (4.7)	(N|G_1)244 (1.3) (C|M_1|G_1)69 (5.0)	G236 (3.9) C69 (4.4)	G236 (3.6)	G314 (2.8) G64 (4.7)	G214 (3.1) W228 (3.4) (T|S_{11})68 (4.5) (D|C_{11}|A8)66 (4.7)
(V|I_3)545			N276 (4.8)		(T|A_1)243 (4.5)		N275 (3.0)	P342 (3.0) Y341 (3.5) R346 (4.5)	

Table S6. PBP residues whose hydrogen bond features (as defined in Methods) differ from those of the corresponding residues in 8 out of the 9 ßLACT families. The amino acids occurring at a given position in the multiple sequence alignments of the family are given in parentheses; the most frequent amino acid is given first, followed by the others with their frequency (in %) as a subscript. For assessing the H-bond features, only ßLACT amino acids whose frequency exceeds 5%, and PBP amino acids whose frequency exceeds 75%, were considered. The residue number is the number in the representative PDB structure. The distance (in Å) is computed between the average side chain centroid of the PBP and ßLACT residues; this distance is given in parentheses. Hydrogen bond donors are in pink, acceptors in brown, and residues that are both donors and acceptors in purple.

PBP3	KPC	TEM	SHV	GES	CTX	SME	PER	CMY	OXA
(**E**|**D**$_2$)411							**Q**168 (2.8)		
L476								(L|**Y**$_2$)274 (3.9)	
(**R**|**K**$_{13}$|**Q**$_1$)506									(**N**|**R**$_7$|G$_1$|**Y**$_1$)221 (3.5)
(I|L$_5$|V$_1$)512		(**S**|A$_1$)243 (4.3)		(G|A$_{11}$)237 (4.0) **N**167 (5.0)	(**T**|A$_1$)242 (3.8)	**T**243 (3.3)	**K**266 (3.8) **D**267 (4.7)		
(V|I$_3$)545			**N**276 (4.8)		(**T**|A$_1$)243 (4.5)		**N**275 (3.0)	P342 (3.0) **Y**341 (3.5) **R**346 (4.5)	

Table S7. Residues that share an identical physicochemical property (as defined in Methods) in all (in green) or all but one (in salmon) ßLACT classes (KPC, CMY and OXA), which differs from the physicochemical property of the equivalent PBP residue. The amino acids occurring at that position in the family's multiple sequence alignments are given in parentheses; the most frequent amino acid is given first, and is followed by the others with their frequency (in %) as a subscript. For assessing the physicochemical properties, only ßLACT amino acids whose frequency exceeds 75%, and PBP amino acids whose frequency exceeds 5%, were considered. The residue number is the number from the representative PDB structure.

PBP3	KPC	CMY	OXA
D343	I108	Y113	V116 (W|P_8)118
(I|L_5)410		(Q|R_3|K_2|M_2)217	(T|D_8)157
T414 (F|Y_2)417	E166 E168	(D|A_2)219	(E|Q_{11}|S_8)158
(N|D_1)361 T414 (F|Y_2)417 Y419	(L|M_4)169	A222	V117 L161
(K|R_3)483	A223	(I|L_2)284 (V|I_2)347	(K|R_{11}|Q_9)262
Y419 A498	C69 C238	S315 T316	(D|C_{11}|A_8)66 (E|M_{11}|D_8)162

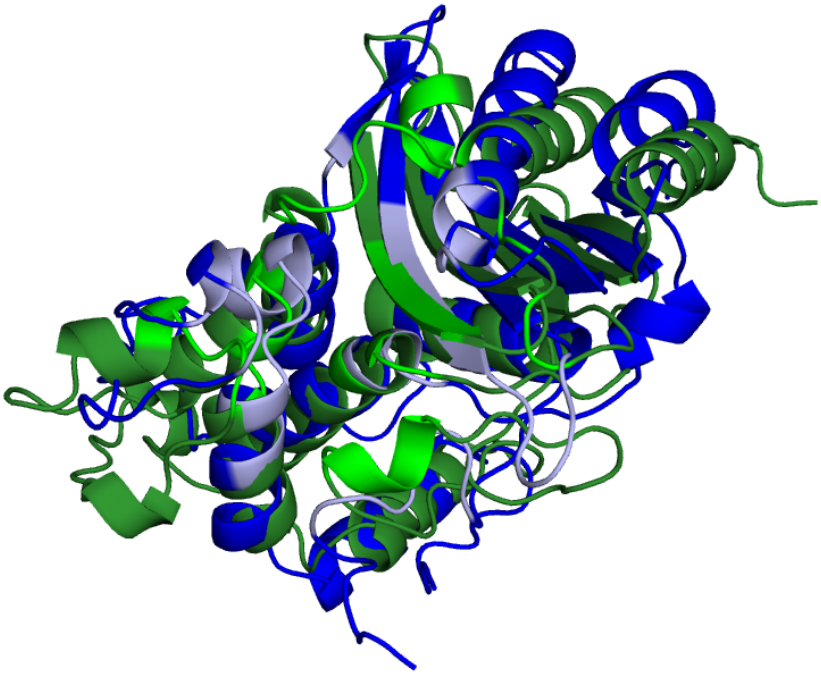
Y514			(G|E_1)216
T497 A498 (K|R_1)499 Y511	G239	(G|N_2)317	(G|E_1)216
T495	I221	Y322	(F|W_8)229
T495 G496 (A|S_2)513 (A|G_4|S_1)544	R220	R346	(S|G_{11}|E_8|K_1)258 (I|S_{11}|V_8|H_1)259

Table S8. Residues that share an identical hydrogen bond feature in at least 2 out of the 3 ßLACT classes KPC, CMY and OXA, which differs from the H-bond feature in PBP. The amino acids occurring at a given position in the multiple sequence alignments of the family are given in parentheses; the most frequent amino acid is given first, followed by the others with their frequency (in %) as a subscript. For assessing the H-bond features, only ßLACT amino acids whose frequency exceeds 75%, and PBP amino acids whose frequency exceeds 5%, were considered. The residue number is the number in the representative PDB structure. Hydrogen bond acceptors are in brown, donors in pink and donors/acceptors in purple.

PBP3	KPC	CMY	OXA
	S106		(**W**|P_8)118
D343 (**Y**|$\mathbf{R}_1$)347	P107 I108	**Y**113	V116
T495	I221	(**S**|$\mathbf{N}_2$)311 Y322	(F|$\mathbf{W}_8$)229
T497 (**K**|$\mathbf{R}_1$)499 **Y**511	G239	(G|$\mathbf{N}_2$)317	(G|$\mathbf{E}_1$)216
A544	**E**276	(**N**|I_2)343	(**S**|G_{11}|$\mathbf{E}_8$|$\mathbf{K}_1$)258

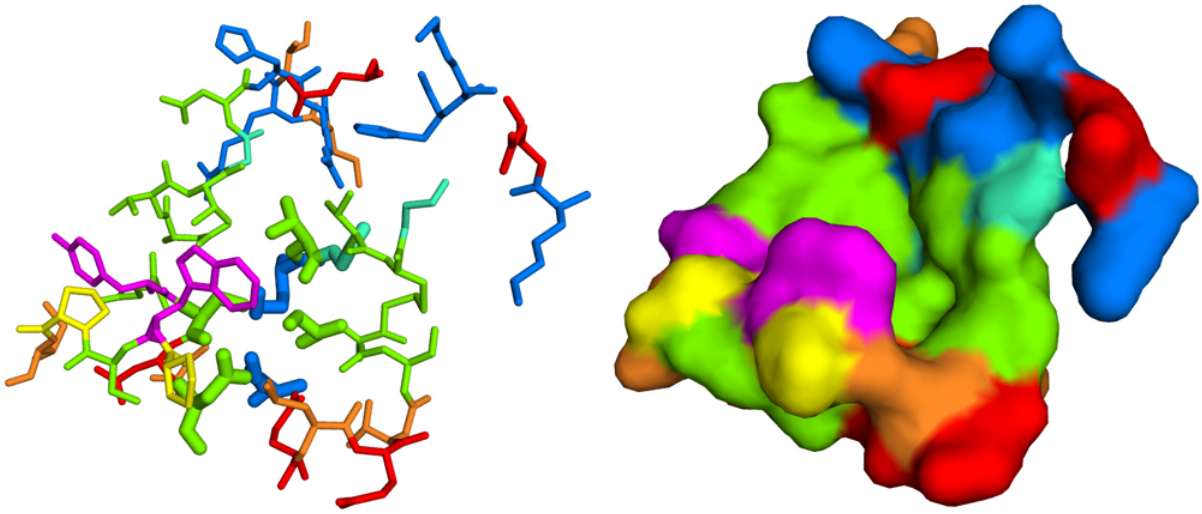
Figure S1. Superposition of the representative structures of ßLACTs of class A, C and D. The associated root mean square deviations are given in Table S4. (a) KPC (class A representative, PDB code 3C5A) in green and CMY (class C representative, PDB code 1ZKJ) in red; the residues that belong to the functional cavity are in light green and salmon. (b) KPC (class A representative, PDB code 3C5A) in green and OXA (class D representative, PDB code 1M6K) in blue; the residues that belong to the functional cavity are in light green and light blue.

(a)

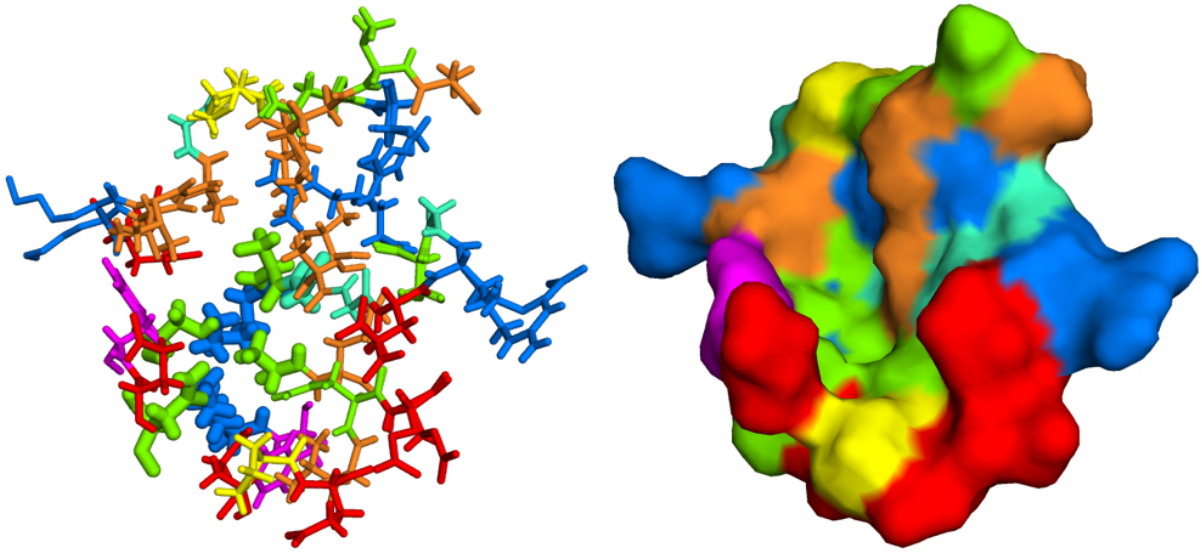


(b)

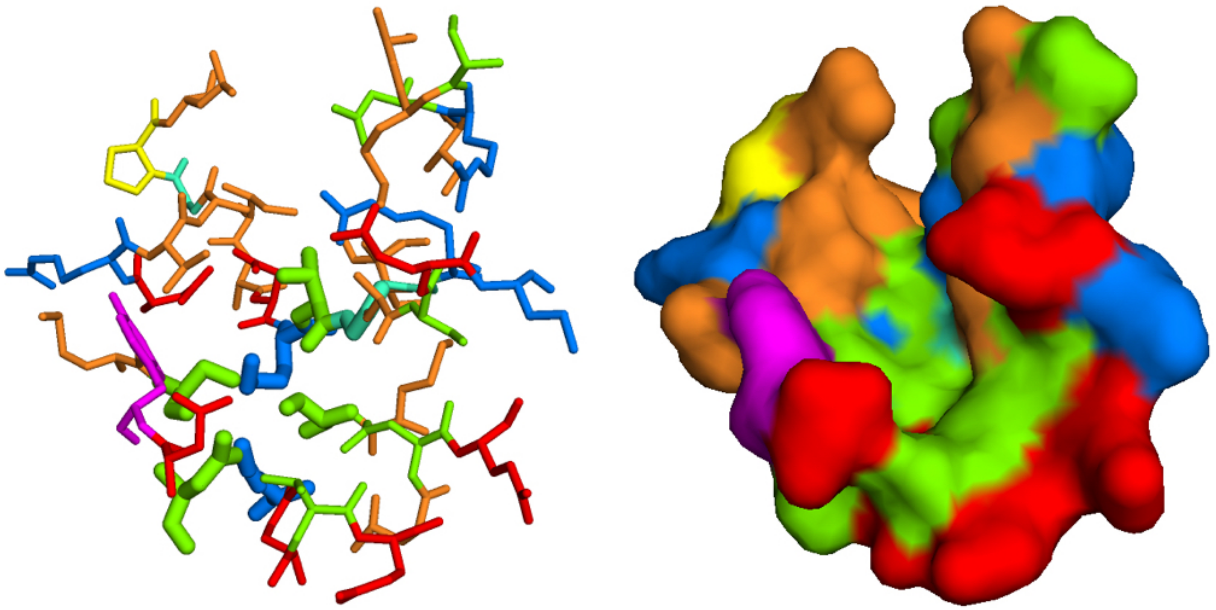
Figure S2. Stick and surface representations of the cavity residues in the 9 ßLACT families. The positively and negatively charged residues are in blue and red, respectively. The polar, apolar and aromatic residues are in green, orange and magenta. Glycines are in cyan and prolines in yellow. The residues from the three catalytic motifs are in large sticks. (a): KPC. (b): TEM. (c): SHV. (d): GES. (e): CTX. (f): SME. (g): PER. (h): CMY. (i): OXA.



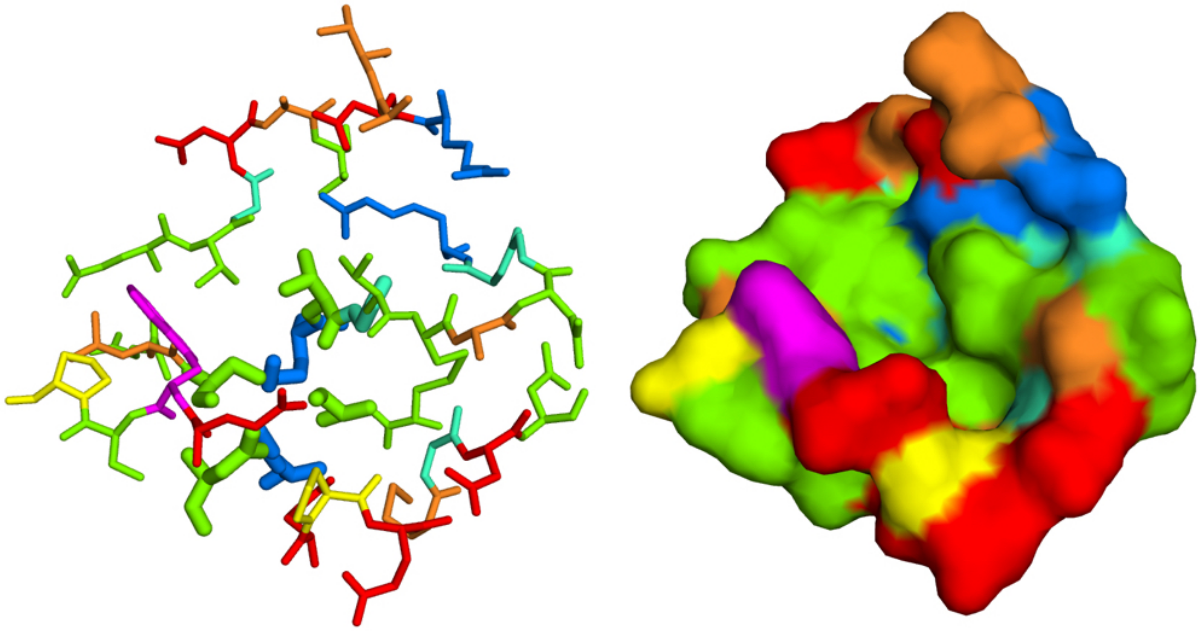
(a)



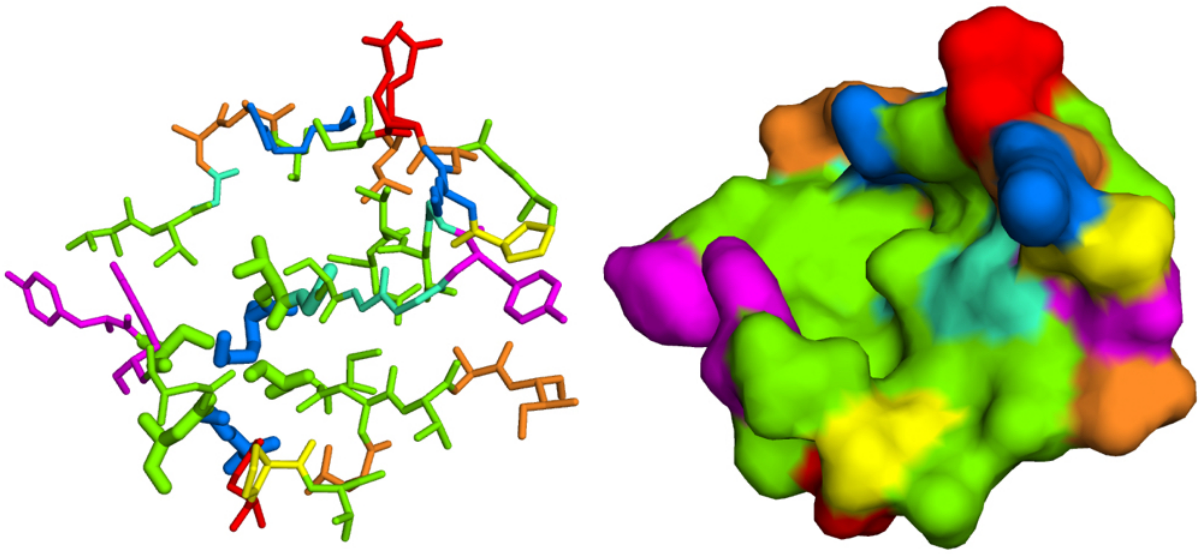
(b)



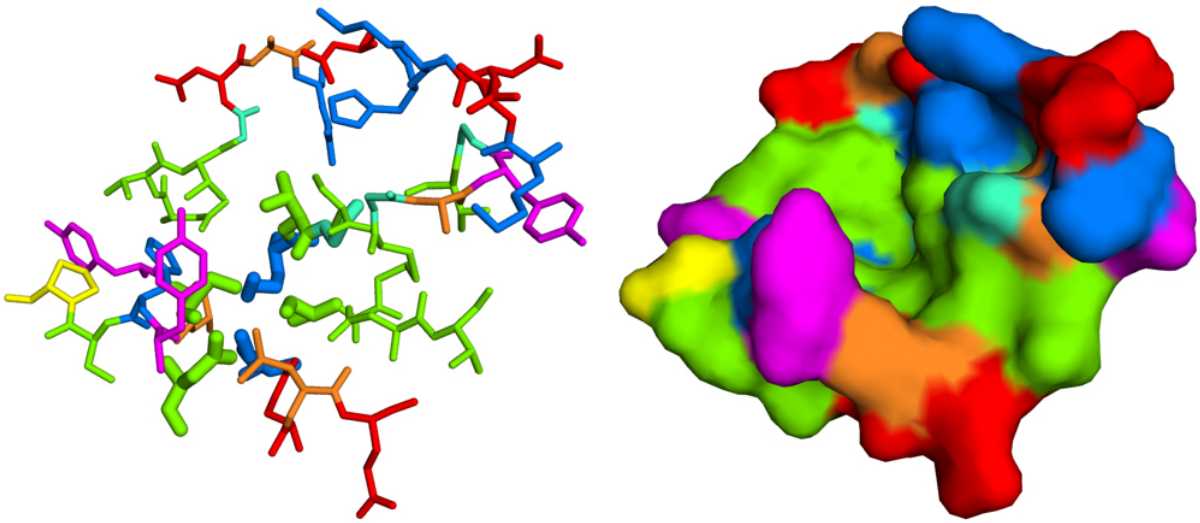
(c)



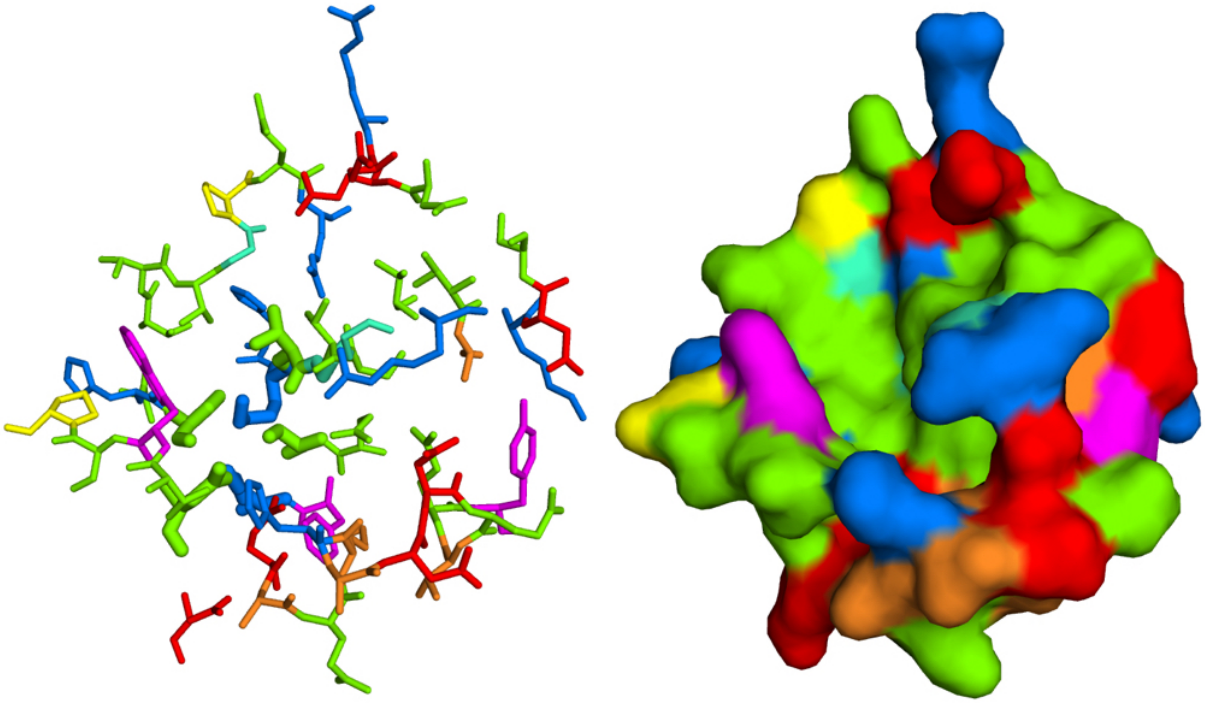
(d)



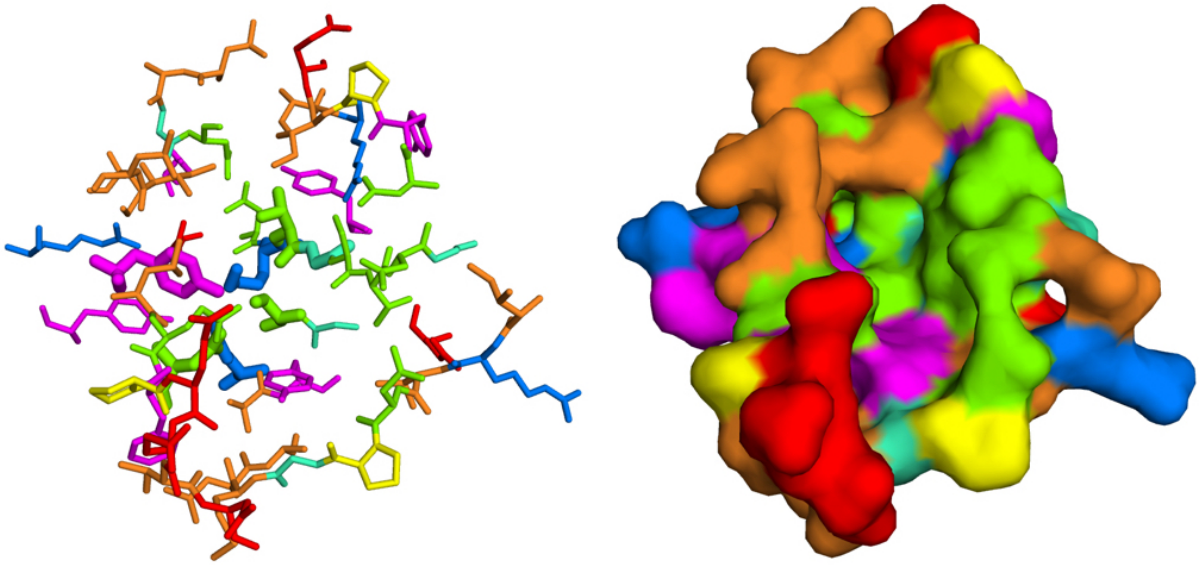
(e)



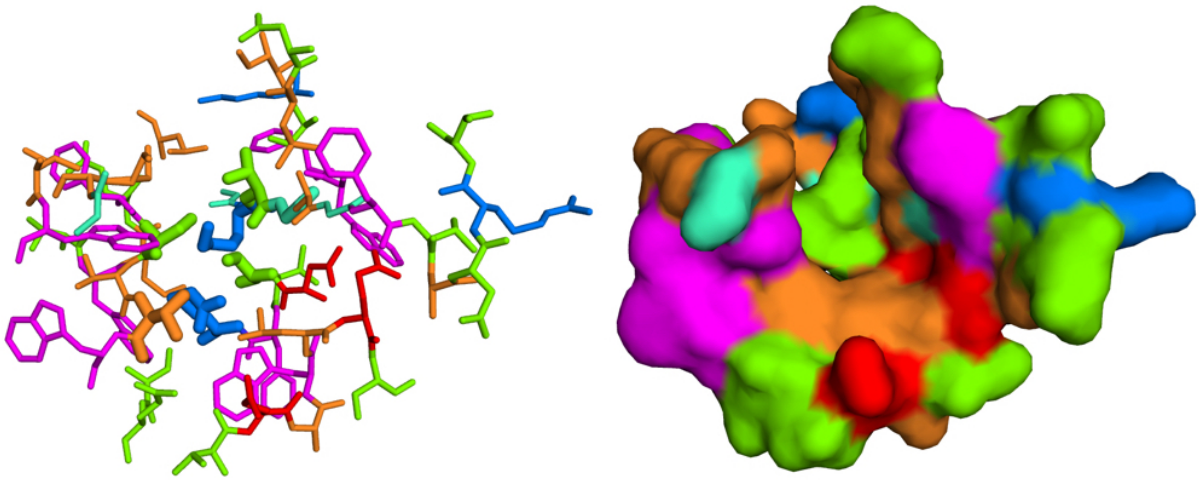
(f)



(g)



(h)



(i)