Supporting Information

Molecular features of cephalosporins important for activity against antimicrobial-resistant *Neisseria gonorrhoeae*

Jonathan M. Turner[1], Kristie L. Connolly[2], Kate E. Aberman[3^], Joseph C. Fonseca II[3#], Avinash Singh[1], Ann E. Jerse[2], Robert A. Nicholas[3,4] and Christopher Davies[1]*

From the [1]Department of Biochemistry and Molecular Biology, Medical University of South Carolina, Charleston, South Carolina 29425; [2]Department of Microbiology and Immunology, Uniformed Services University, Bethesda, MD 20814; [3]Departments of Pharmacology and [4]Microbiology and Immunology, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599

*Corresponding author: Department of Biochemistry & Molecular Biology, University of South Alabama, 5795 USA Drive North, Mobile, AL 36688. Tel +1 (651) 460-6659; E-mail address: cdavies@southalabama.edu;; Present addresses: ^National Institute of Allergy and Infectious Diseases, Immunobiology Section, 33 North Drive, Bethesda MD 20814 and [#]Morehouse College, 830 Westview Drive SW, Atlanta, GA 30314

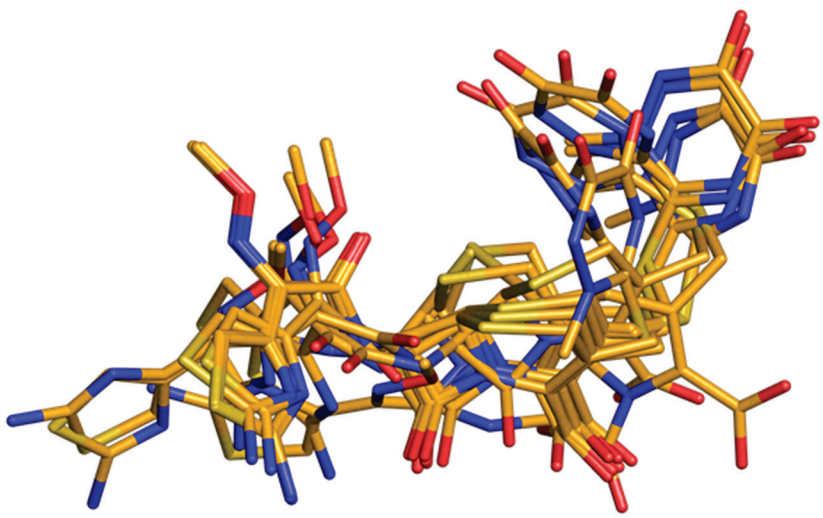


Figure S1: *Superimposition of all docked poses.* Top ten poses for ceftriaxone

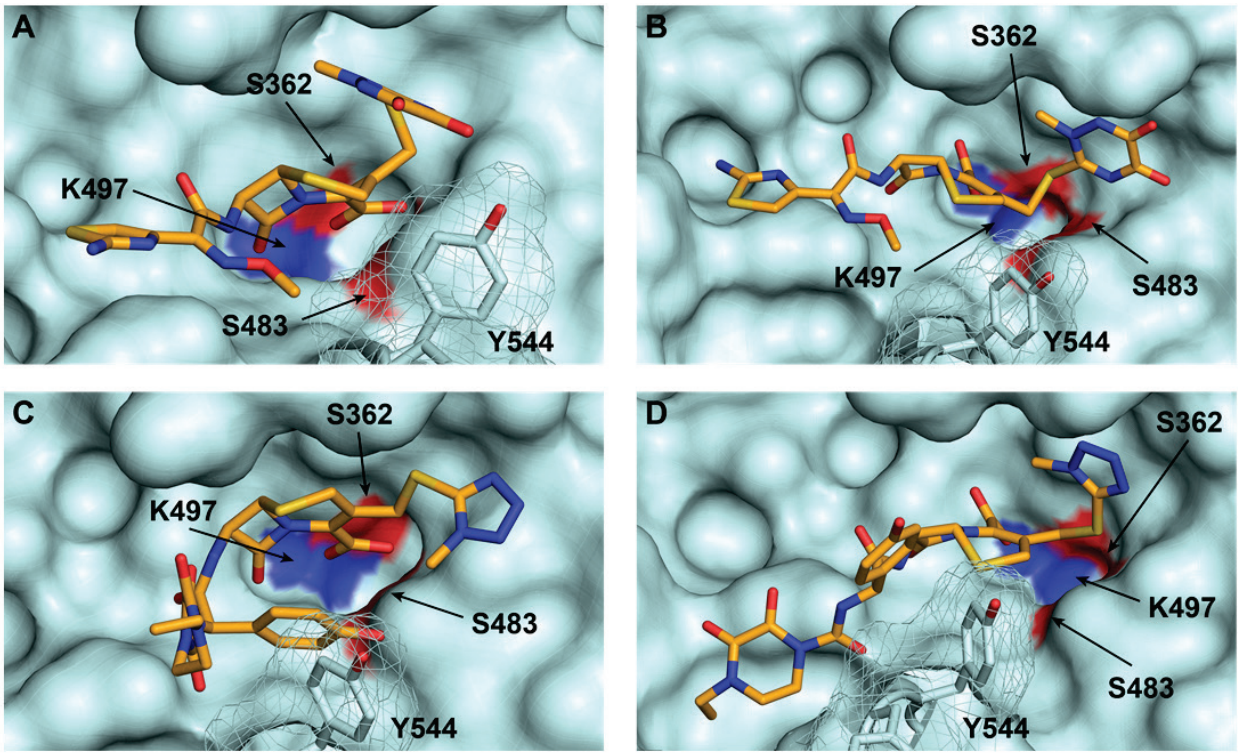


Figure S2: *Surface rendering of ceftriaxone and cefoperazone binding poses.* **A.** In the major pose predicted by the pharmacophore-constrained docking protocol, the C4 carboxylate of ceftriaxone (orange) fits into a small pocket bordered by the side chains of S362, S483, and K497 (surface colorized by atom type). **B.** In the minor pose, ceftriaxone is more extended, spanning the active site. **C.** Major pose for cefoperazone (orange). **D.** Minor pose for cefoperazone. Y544 is shown with a transparent mesh surface to visualize the binding pocket.

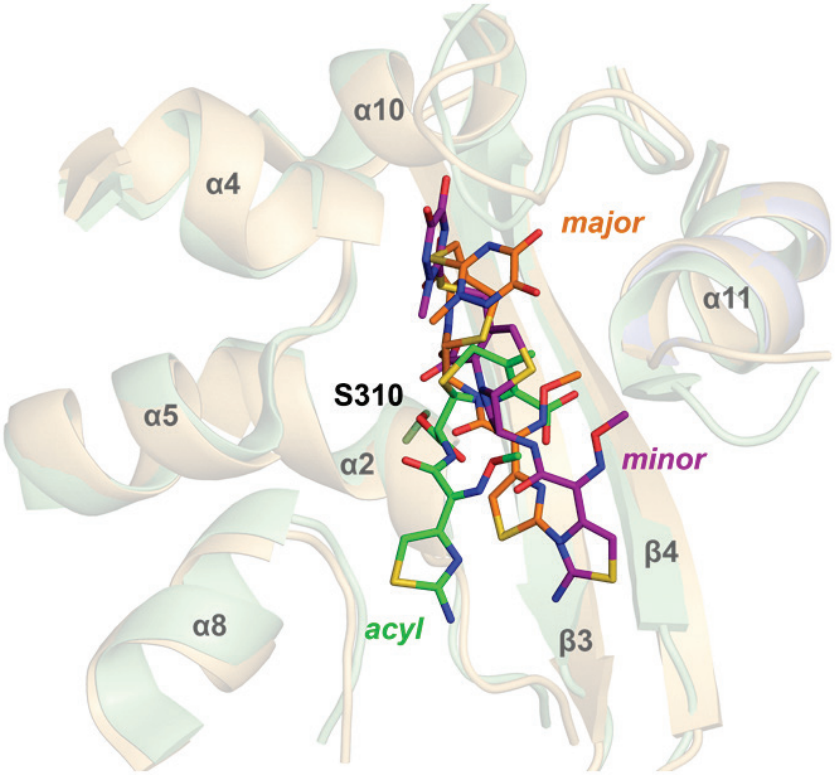


Figure S3: *Overlay of the major and minor docked poses with the structure of tPBP2^{H041} acylated by ceftriaxone (CRO).* Both poses of docked precovalent complexes show minimal overlap with the published acylated structure. The major pose from the docking analysis is shown in orange, the minor pose is shown in purple, and the acylated structure is shown in green.

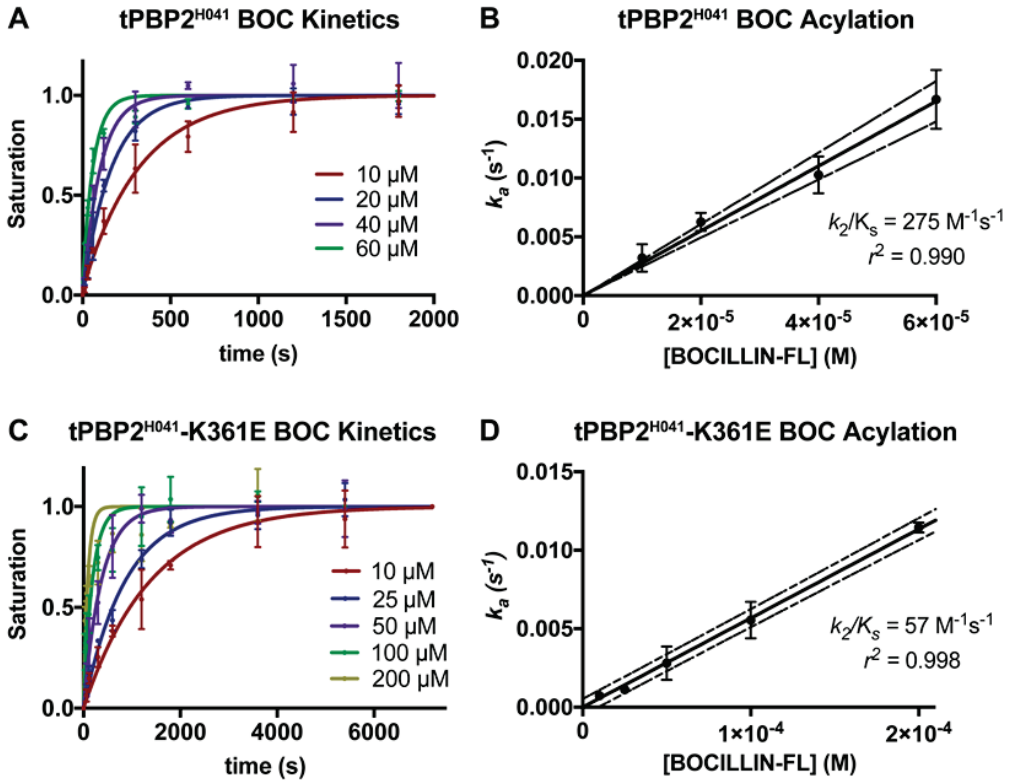


Figure S4: *Second-order rate of acylation of Bocillin-FL against tPBP2^{H041}* and its K361E mutant. **A.** Time-dependent acylation of tPBP2^{H041} at various Bocillin-FL concentrations. Each curve was used to derive the pseudo first-order rate constant, k_a. **B.** Plot of k_a against the concentration of Bocillin-FL, where the slope is the second order acylation rate constant (k_2/K_s). **C.** Acylation curves for acylation of tPBP2^{H041}-K361E by Bocillin-FL. **D.** Derivation of the second order acylation rate constant (k_2/K_s). For both experiments, a minimum of three kinetic experiments were completed for each concentration.

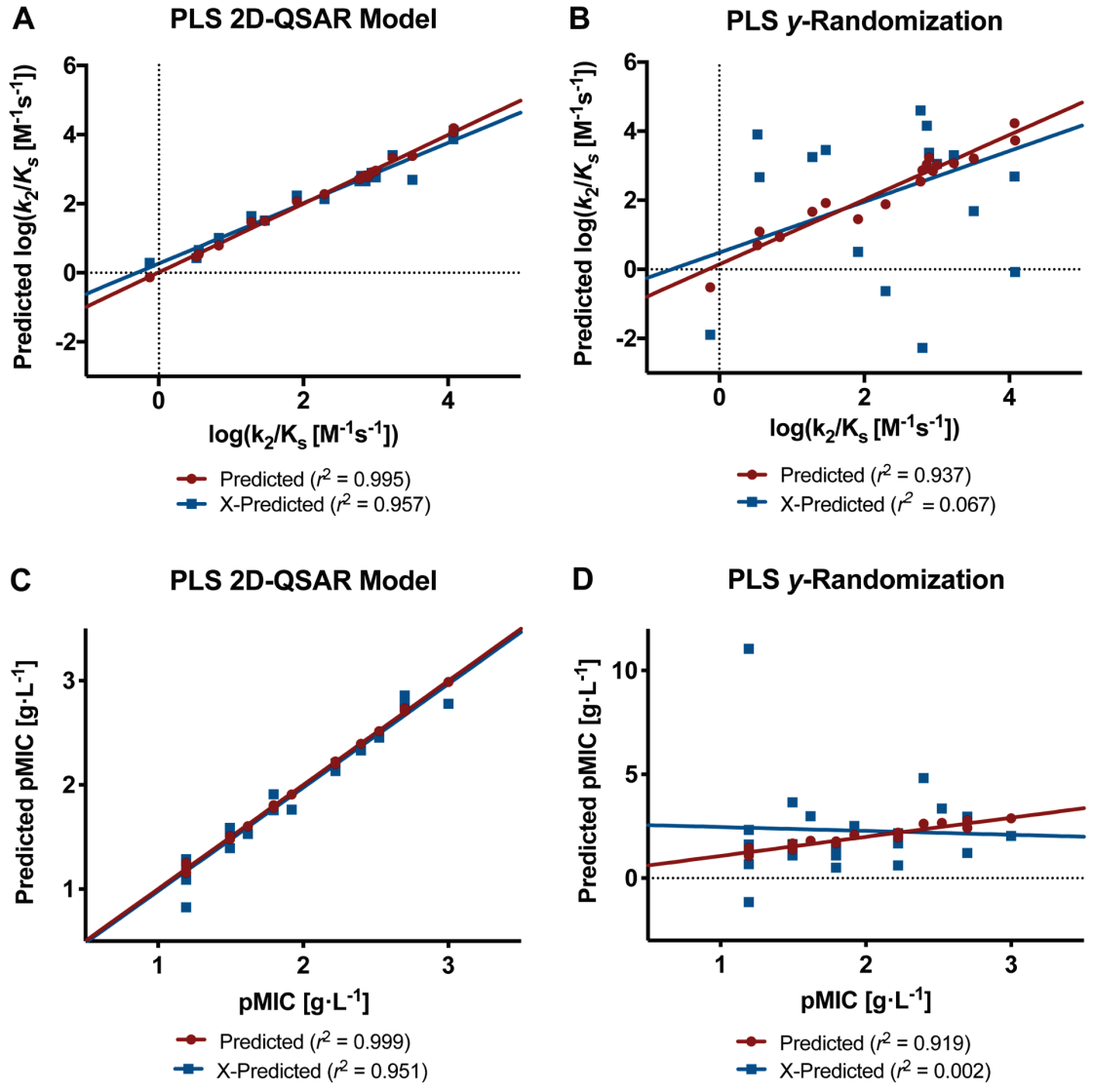


Figure S5: *Quantitative structure activity relationships of cephalosporins against N. gonorrhoeae.* **A.** Partial least squares QSAR model of tPBP2^{H041} acylation rate constant data, generated using the physicochemical and structural descriptors shown in **Table 2**. The model exhibits high internal validity by direct application to the training set (predicted $r^2 = 0.995$) and by leave-one-out cross validation (X-predicted $r^2 = 0.957$). **B.** Random assignment of second-order acylation rate constants to physicochemical and structural descriptors yields a poor PLS model of tPBP2^{H041}-inhibitory activity (X-$r^2 = 0.067$). **C.** Partial least squares QSAR model of antimicrobial activity against *Neisseria gonorrhoeae* H041, generated using the descriptors shown in **Table 3**. The model exhibits high internal validity by direct application to the training set (predicted $r^2 = 0.999$) and by leave-one-out cross validation (X-predicted $r^2 = 0.951$). **D.** Random assignment of MIC values to physicochemical and structural descriptors yields a poor PLS model of tPBP2^{H041}-inhibitory activity (X-predicted $r^2 = 0.002$).

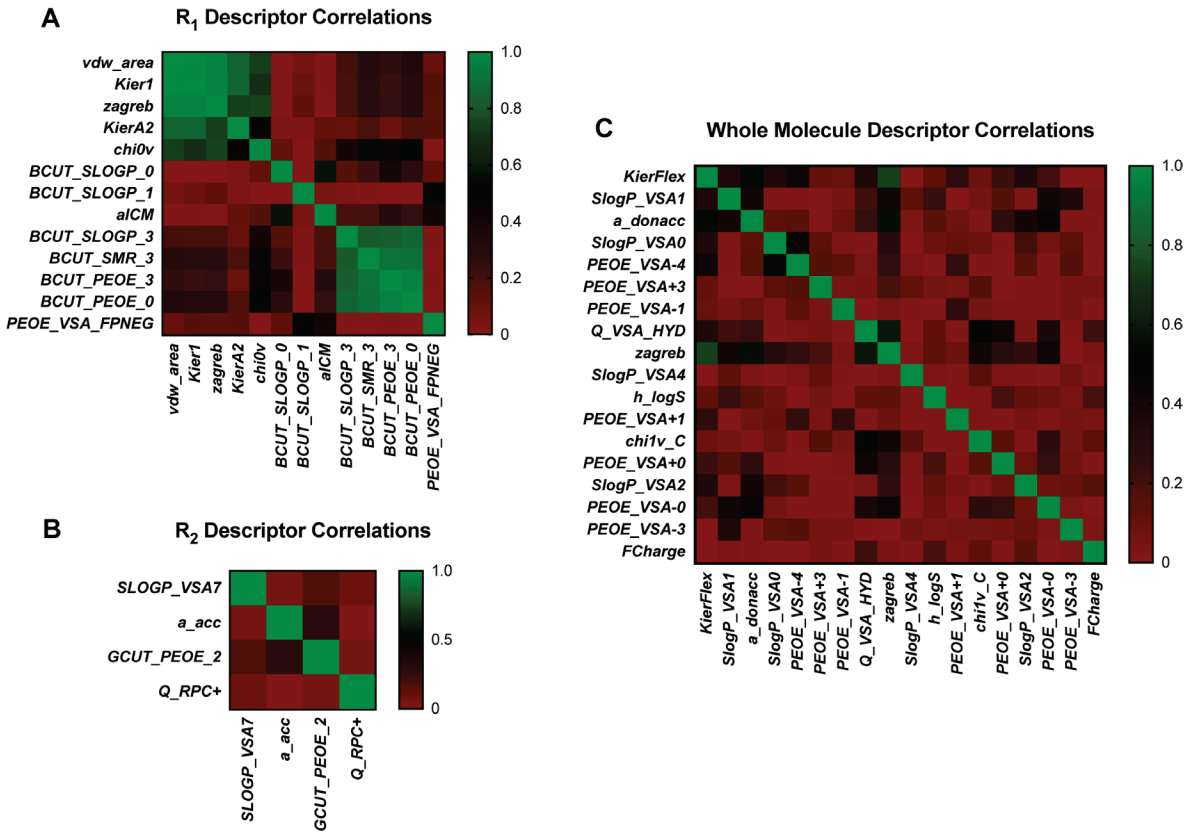


Figure S6: *Heat maps of the correlation matrices between input physicochemical descriptors.* To determine the extent of input variable collinearity, correlation coefficients were calculated for each unique pair of descriptors using their values for the cephalosporins examined in these studies. The scale represents the square of the Pearson correlation coefficient (r^2). **A.** Correlation coefficients for R_1 descriptors used in the PLS QSAR model for second-order acylation rate. **B.** Correlation coefficients for R_2 descriptors used in the PLS QSAR model for second-order acylation rate. **C.** Correlation coefficients for whole-molecule descriptors used in the PLS QSAR model for antimicrobial activity.

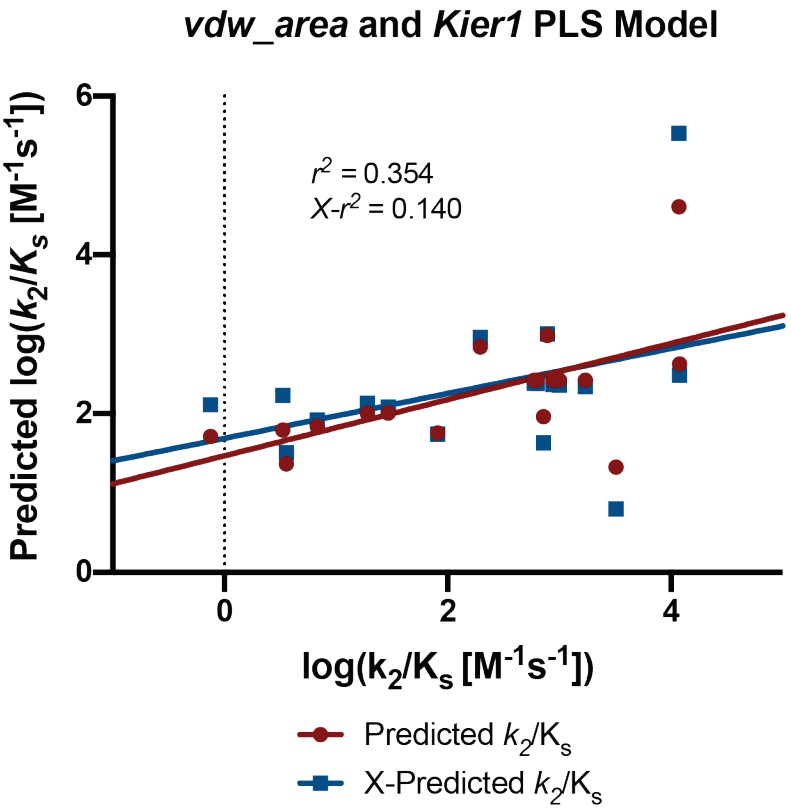


Figure S7: *Partial least squares QSAR model of tPBP2^{H041} acylation rate constant data, generated using only the vdw_area and Kier1 descriptors.* In comparison to the 17-descriptor model presented in **Figure S5**, the minimal model exhibits relatively lower internal validity by direct application to the training set (predicted r^2 = 0.354) and by leave-one-out cross validation (X-predicted r^2 = 0.140).

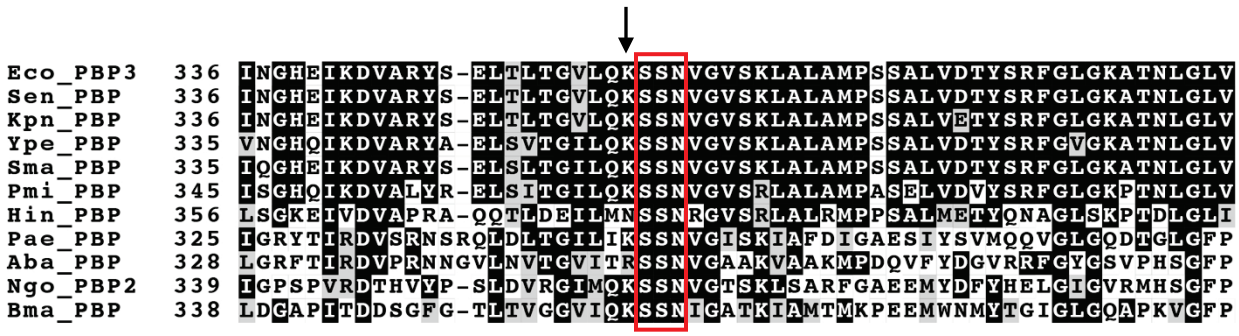


Figure S8: *Sequence alignment of PBP2 homologues from several Gram-negative bacteria.* The red box marks the SXN motif of PBPs and the arrow denotes the position corresponding to Lys361 in *N. gonorrhoeae* PBP2. (Key: Eco, *Escherichia coli*; Sen, *Salmonella enterica*, Kpn, *Klebsiella pneumoniae*; Ype, *Yersinia pestis*; Sma, *Serratia marcescens*; Pmi, *Proteus mirabilis*; Hin, *Haemophilus influenzae*; Pae, *Pseudomonas aeruginosa*; Aba, *Acinetobacter baumannii*; Ngo, *Neisseria gonorrhoeae* and Bma, *Burkholderia mallei*. The sequences were aligned in ClustalX and the figure made using the Boxshade server (https://embnet.vital-it.ch/software/BOX_form.html).

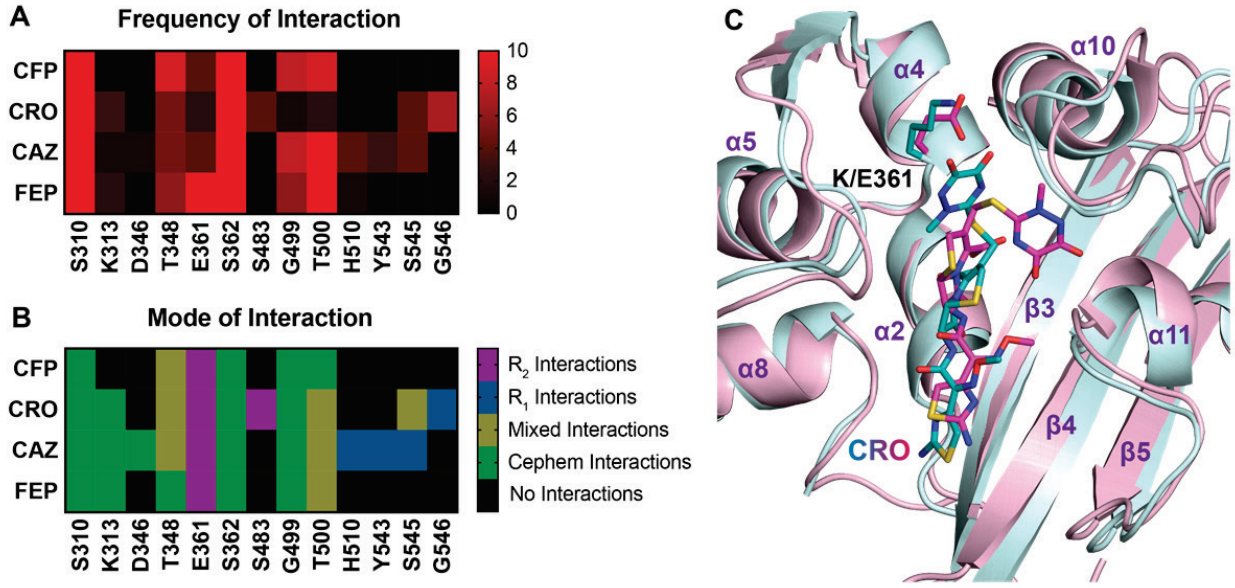


Figure S9: *Pharmacophore-constrained induced fit docking of select cephalosporins to tPBP2^{H041}-K361E.* **A.** Frequency of interaction with residues in the active site among docking poses. **B.** Modes of interactions with residues in the active site among docking poses. While interactions with S310 and T500 are conserved, less frequent interactions are seen with residue 361. **C.** A representative pose of ceftriaxone (magenta) shows the R_2 group occupying a different position compared to that observed in the acylated tPBP2^{WT} crystal structure (light blue) [1].

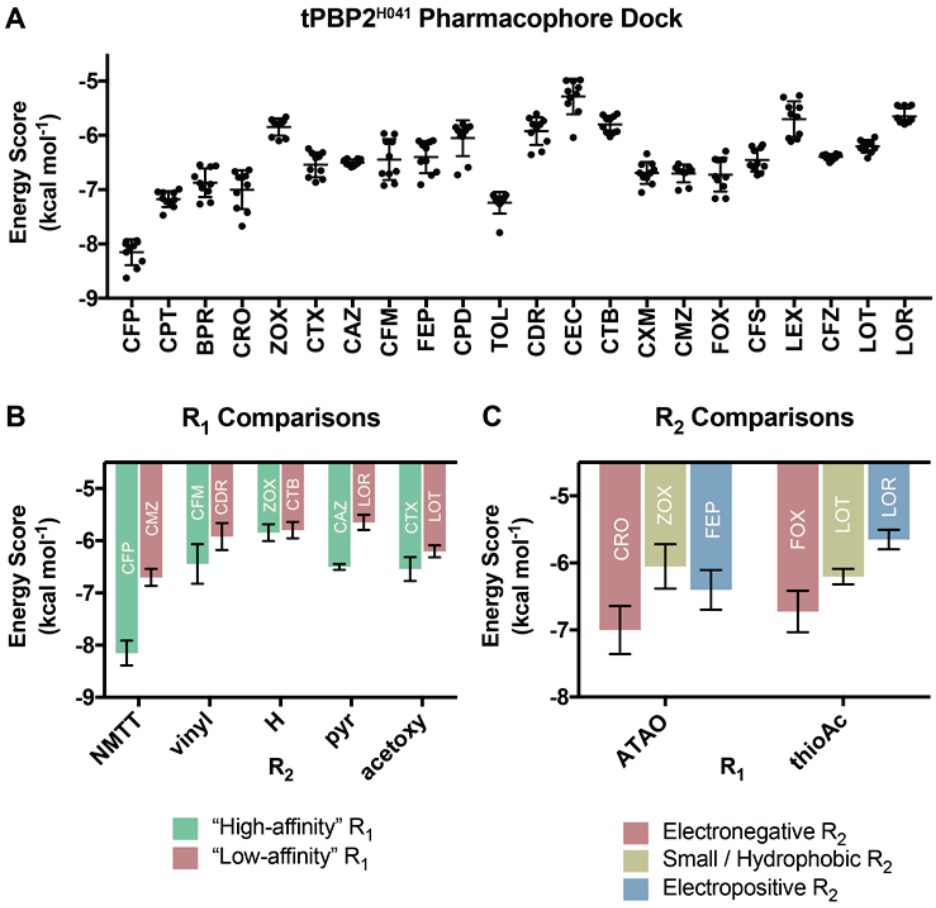


Figure S10: *Docking energy score analyses.* **A**. Overall energy scores for the cephalosporins tested, rank ordered by second order acylation rate (n = 10 poses), show a correlation of computational affinity with activity. **B**. Comparisons of energy scores for cephalosporins with different R_1 and identical R_2 groups "High-affinity" indicates the molecule of the pair exhibiting more rapid acylation (*e.g.,* cefoperazone). "Low-affinity" indicates the molecule of the pair exhibiting slower acylation (*e.g.,* cefmetazole). **C**. Comparisons of energy scores for cephalosporins with similar R_1 groups and different R_2 show a preference for negative electrostatics. NMTT = *N*-methyltetrazole, pyr = pyridyl, ATAO = 2-(2-aminothioazol-4-yl)-2-(alkoxyimino)acetyl and thioAc = thiophen-2-ylacetyl.

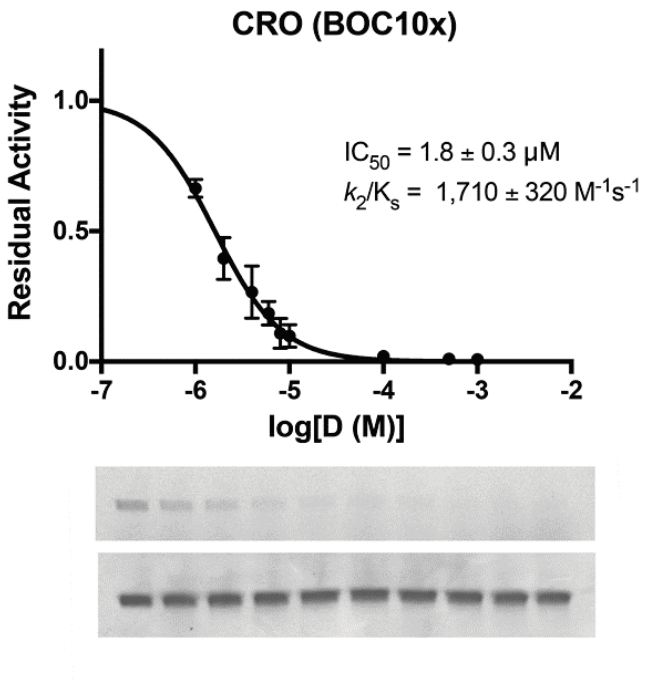
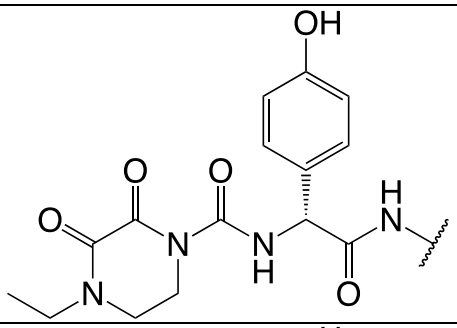
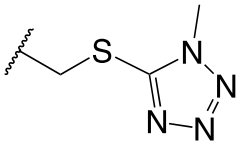
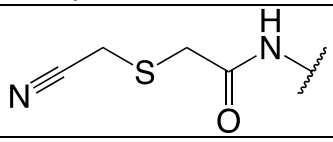
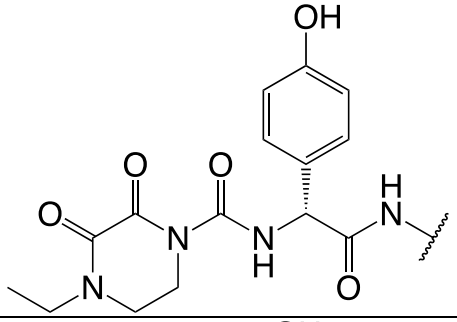
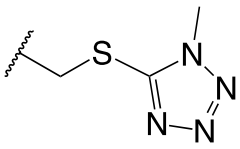
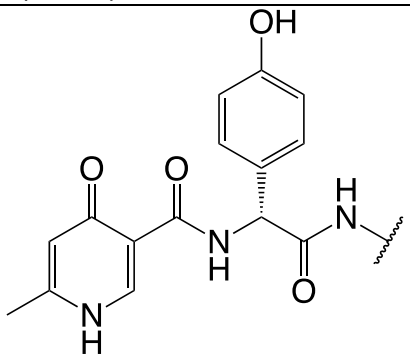
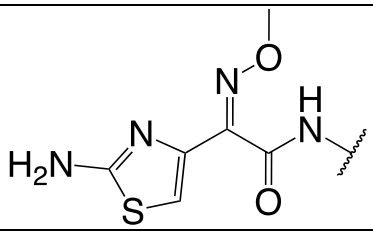
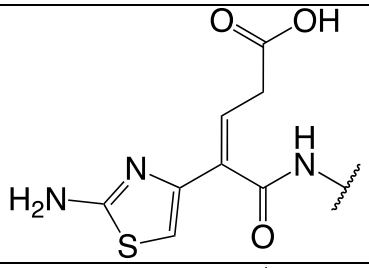
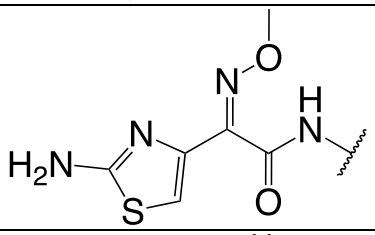
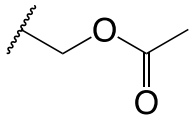
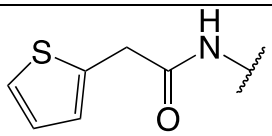
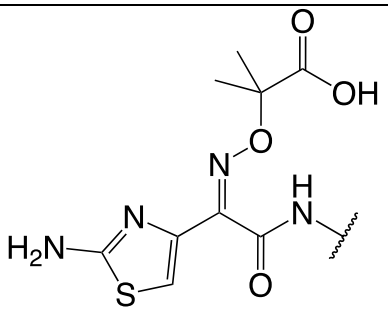
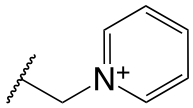
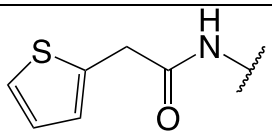
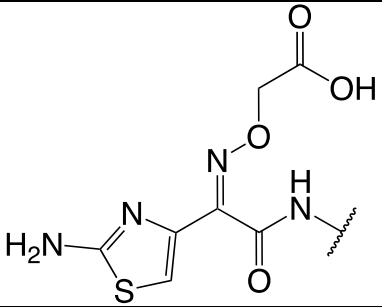
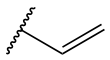
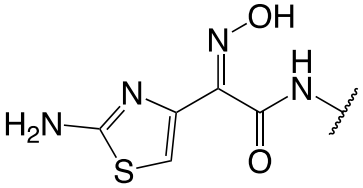
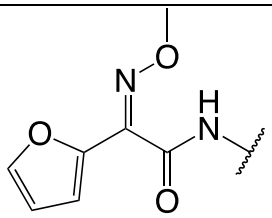
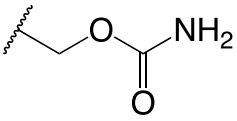
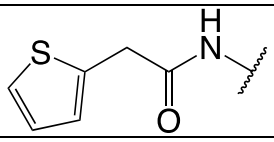


Figure S11: *Representative half-maximal inhibitory concentration determination*. Data are shown for the measurement of ceftriaxone inhibition of tPBP2^{H041}. UV gel imaging for Bocillin-FL detection and a grayscale image of Coomassie staining are shown. Error bars are standard deviation (n = 5, r^2 = 0.977). BOC10x: co-incubation with 10 µM Bocillin-FL.

Table S1: *Pairwise comparisons of cephalosporins with identical R_2.*

Cephalosporin	R_1	R_2	k_2/K_s ($M^{-1}s^{-1}$)[a]	Fold-Change[b]
CFP			11,800 ± 1,300	3,300
CMZ[†]			3.6 ± 0.1	
CFP			11,800 ± 1,300	98
CPM			120 ± 14	
ZOX			1,000 ± 210	52

CTB			19.1 ± 0.2	
CTX			880 ± 120	∞
LOT			~ 0‡	
CAZ			790 ± 150	∞
LOR			~ 0‡	
CFM			720 ± 60	8.8
CDR			82 ± 7	

CFM			6.8 ± 0.2	2.1
FOX†			3.3 ± 0.6	

[a]Second-order acylation rate constants taken from **Table 1**. [b]Quotient of second-order acylation rate constants for fast- versus slow-acylating cephalosporin. †Cephamycin – possesses C7-OMe group. ‡Inhibition plot could not be fit due to minimal inhibition seen at the highest concentration used.

Table S2: *Application of PLS QSAR to a test set of cephalosporins.*

Cephalosporin	k_2/K_s **($M^{-1}s^{-1}$)[a]**	**Predicted k_2/K_s ($M^{-1}s^{-1}$)[b]**
cefpiramide (CPM)	120 ± 14	3.2
cephalexin (LEX)	~ 0.3†	2.6
cefazolin (CFZ)	~ 0‡	0.1
cephalothin (LOT)	~ 0‡	1.4
cephaloridine (LOR)	~ 0‡	24.6

[a]Second-order acylation rates for native $tPBP2^{H041}$, as reported in **Table 1**. [b]Second-order acylation rates computed using the model reported in **Table 2** and **Figure 7**. Test set computations exhibit an overall root mean square error of 12.4 $M^{-1}s^{-1}$ compared to true acylation rate.

Table S3: *Physicochemical descriptor definitions.*

Descriptor	Definition
a_acc	Number of hydrogen bond acceptor atoms (excluding acidic atoms, but including atoms that function as both hydrogen bond donors and acceptors)
a_donacc	Number of hydrogen bond donor atoms plus number of hydrogen bond acceptors atoms
a_ICM	Mean atom information content / entropy of element distribution. Sum of $p_i \log p_i$ *for* all *i,* where $p_i = n_i/n$, n_i is the number of occurrences of atomic number *i* in the molecule, and *n* is the total number of atoms.
BCUT_PEOE_0	Smallest eigenvalue of a modified adjacency matrix in which each *ij* entry takes the value 1/sqrt(b_{ij}), where b_{ij} is the formal bond order between bonded atoms *i* and *j*. The diagonal takes the value of PEOE partial charges.
BCUT_PEOE_3	Largest eigenvalue of a modified adjacency matrix in which each *ij* entry takes the value 1/sqrt(b_{ij}), where b_{ij} is the formal bond order between bonded atoms *i* and *j*. The diagonal takes the value of PEOE partial charges.
BCUT_SLOGP_0	Smallest eigenvalue of a modified adjacency matrix in which each *ij* entry takes the value 1/sqrt(b_{ij}), where b_{ij} is the formal bond order between bonded atoms *i* and *j*. The diagonal takes the value of atomic contribution to logP(o/w).
BCUT_SLOGP_1	Second eigenvalue of a modified adjacency matrix in which each *ij* entry takes the value 1/sqrt(b_{ij}), where b_{ij} is the formal bond order between bonded atoms *i* and *j*. The diagonal takes the value of atomic contribution to logP(o/w).
BCUT_SLOGP_3	Largest eigenvalue of a modified adjacency matrix in which each *ij* entry takes the value 1/sqrt(b_{ij}), where b_{ij} is the formal bond order between bonded atoms *i* and *j*. The diagonal takes the value of atomic contribution to logP(o/w).
BCUT_SMR_3	Largest eigenvalue of a modified adjacency matrix in which each *ij* entry takes the value 1/sqrt(b_{ij}), where b_{ij} is the formal bond order between bonded atoms *i* and *j*. The diagonal takes the value of atomic contribution to molar refractivity.
chi0v	Atomic valence connectivity index (order 0) from Hall, calculated as the sum of 1/sqrt(v_i) over all heavy atoms *i* with $v_i > 0$, where $v_i = (p_i - h_i)/(Z_i - p_i - 1)$ and p_i is the number of *s* and *p* valence electrons.

chi1v_C	Carbon valence connectivity index (order 1), calculated as the sum of 1/sqrt($v_i v_j$) over all bonds between carbon atoms i and j where $i < j$, where $v_i = (p_i\text{-}h_i)/(Z_i\text{-}p_i\text{-}1)$ and p_i is the number of s and p valence electrons.
FCharge	Total formal charge.
GCUT_PEOE_1	Second eigenvalue of a modified distance adjacency matrix in which each ij entry takes the value 1/sqrt(d_{ij}), where d_{ij} is the graph distance between atoms i and j. The diagonal takes the value of PEOE partial charges.
GCUT_PEOE_2	Third eigenvalue of a modified distance adjacency matrix in which each ij entry takes the value 1/sqrt(d_{ij}), where d_{ij} is the graph distance between atoms i and j. The diagonal takes the value of PEOE partial charges.
h_logS	log of aqueous solubility (mol/L) using a 7-parameter model based on Hueckel Theory with $r^2 = 0.83$, RMSE 0.85 on 1,708 molecules.
Kier1	First kappa shape index: $(n\text{-}1)^2/m^2$, where n is the number of atoms in a hydrogen-suppressed molecular graph and m is the number of bonds.
KierA2	Second alpha modified shape index: $s(s\text{-}1)^2/m^2$ where $s = n+a$, where a is a correction factor derived from the covalent radii of bonded atoms.
KierFlex	Kier molecular flexibility index: (*KierA1*)(*KierA2*)/*n*.
PEOE_VSA_FPNEG	Fractional negative polar van der Waals surface area: sum of approximate van der Waals surface area v_i (Å^2) for each atom i with partial charge $q_i < -0.2$, divided by the total van der Waals surface area.
PEOE_VSA-1	Sum of approximate van der Waals surface area v_i (Å^2) for each atom i whose contribution to *PEOE* q_i is in the range [-0.1, 0.05).
PEOE_VSA-3	Sum of approximate van der Waals surface area v_i (Å^2) for each atom i whose contribution to *PEOE* q_i is in the range [-0.2, -0.15).
PEOE_VSA-4	Sum of approximate van der Waals surface area v_i (Å^2) for each atom i whose contribution to *PEOE* q_i is in the range [-0.25, -0.2).
PEOE_VSA+0	Sum of approximate van der Waals surface area v_i (Å^2) for each atom i whose contribution to *PEOE* q_i is in the range [0, 0.05).
PEOE_VSA+1	Sum of approximate van der Waals surface area v_i (Å^2) for each atom i whose contribution to *PEOE* q_i is in the range [0.05, 0.1).
PEOE_VSA+3	Sum of approximate van der Waals surface area v_i (Å^2) for each atom i whose contribution to *PEOE* q_i is in the range [0.15, 0.2).

PEOE_VSA+4	Sum of approximate van der Waals surface area v_i (Å^2) for each atom *i* whose contribution to *PEOE* q_i is in the range [0.2, 0.25).
PEOE_VSA+6	Sum of approximate van der Waals surface area v_i (Å^2) for each atom *i* whose contribution to *PEOE* $q_i > 0.3$.
Q_RPC+	Relative positive partial charge: the largest positive q_i divided by the sume of positive q_i.
SlogP_VSA0	Sum of approximate van der Waals surface area v_i (Å^2) for each atom *i* whose contribution to *logP(o/w)* $L_i \leq -0.4$.
SlogP_VSA1	Sum of approximate van der Waals surface area v_i (Å^2) for each atom *i* whose contribution to *logP(o/w)* L_i is in the range (-0.4, 0.2].
SlogP_VSA2	Sum of approximate van der Waals surface area v_i (Å^2) for each atom *i* whose contribution to *logP(o/w)* L_i is in the range (-0.2, 0].
SlogP_VSA4	Sum of approximate van der Waals surface area v_i (Å^2) for each atom *i* whose contribution to *logP(o/w)* L_i is in the range (0.1, 0.15]
SlogP_VSA7	Sum of approximate van der Waals surface area v_i (Å^2) for each atom *i* whose contribution to *logP(o/w)* L_i is in the range (0.25, 0.30].
SMR_VSA3	Sum of approximate van der Waals surface area v_i (Å^2) for each atom *i* whose contribution to molar refractivity R_i is in the range (0.35, 0.39].
SMR_VSA4	Sum of approximate van der Waals surface area v_i (Å^2) for each atom *i* whose contribution to molar refractivity R_i is in the range (0.4, 0.44].
vdw_area	area of van der Waals surface (Å^2) calculated using connection table approximation.
zagreb	Zagreb index: the sum of squares of heavy atom degree d_i over all heavy atoms *i.*

References

1. Singh, A.; Tomberg, J.; Nicholas, R. A.; Davies, C., Recognition of the β-lactam carboxylate triggers acylation of *Neisseria gonorrhoeae* penicillin-binding protein 2. *J Biol Chem* **2019,** *294* (38), 14020-14032.