Article type: Original article

IN VITRO EVALUATION OF THE ANTIMICROBIAL POTENTIAL OF BETULINIC ACID AND ANALYSIS OF MECHANISMS OF ACTION WITH MOLECULAR MODELING

Gabriela Cristina Soares Rodrigues[1], Mayara dos Santos Maia[1], Thalisson Amorim de Souza[2], Edeltrudes de Oliveira Lima[3], Luiz Eduardo C. G. dos Santos[3], Shellygton Lima Silva[3], Marcelo Sobral da Silva[2], José Maria Barbosa Filho[2]; Valnês da S. Rodrigues-Junior[4,5], Luciana Scotti[1], Marcus Tullius Scotti[1].

[1]*Laboratory of Cheminformatics, Program of Natural and Synthetic Bioactive Products (PgPNSB), Health Sciences Center, Federal University of Paraíba, João Pessoa-PB, Brazil*

[2] *Multi-User Characterization and Analysis Laboratory, Program of Natural and Synthetic Bioactive Products (PgPNSB), Health Sciences Center, Federal University of Paraíba, João Pessoa-PB, Brazil.*

[3]*Program of Natural and Synthetic Bioactive Products (PgPNSB), Health Sciences Center, Federal University of Paraíba, 58051-900, João Pessoa-PB, Brazil*

[4]*National Institute of Science and Technology in Tuberculosis (INCT-TB), Brazil*

[5]*Program of Biotechnology, Center for Biotechnology, Federal University of Paraíba, Campus I, Cidade Universitária, João Pessoa 58051-900, PB, Brazil.*

Correspondence: luciana.scotti@gmail.com; Tel.: +55 83 99624 5075

Table S1. Access codes of selected protein sequences obtained from the GenBank database.

	DNA gyrase	beta-lactamase	PBP
Staphylococcus aureus	BAR07518.1	KSA62127.1	BBA24197.1
S. epidermidis	KSZ68723.1	RUN10939.1	EJE17909.1
Pseudomonas aeruginosa	OHQ53593.1	KJJ11413.1	TEF07458.1
Escherichia coli	KHJ14714.1	KXG94406.1	ARB43848.1
Mycobacterium tuberculosis	AIH53390.1	KMY17933.1	QRX86232.1
	CYP51	SAP-2	DHFR
Candida albicans	AIQ80983.1	AAM21051.1	AAC05610.1
C. tropicalis	QXT50470.1	AAD33216.1	XP_002550914.1
C. glabrata	KTB23913.1	-	KTB19277.1

Aspergillus flavus	QOR32183.1	-	RMZ41744.1
Penicillium citrinum	-	-	-
Trichophyton rubrum	-	-	OAL62276.1
Microsporum canis	-	-	EEQ31600.1

Table S2. PDB codes of template proteins and degree of identity proteins of interest.

Species	Protein	ID PDB molde	Identity
S. epidermidis	DNA gyrase	5NPP	93.83%
	beta-lactamase	1XA1	94.90%
	PBP	3VSL	83.18%
	SAP-2	2H6S	51.20%
C. tropicalis	DHFR	1AOE	81.25%
A. flavus	CYP51	6CR2	64.03%
Trichophyton rubrum	DHFR	6DRS	48.97%

Table S3. Percentage of amino acids present in the allowed and favored regions of the Ramachandran plot for each model and percentage of the degree of compatibility between the 3D structure and the 1D amino acid sequence, based on the models generated by Verify 3D.

Species	Protein	Ramachandran	Verify 3D
S. epidermidis	DNA gyrase	99.3%	89.63%
	beta-lactamase	97.8%	100%
	PBP	98.5%	90.52%
	SAP-2	98.7%	93.15%
C. tropicalis	DHFR	100%	92.23%
A. flavus	CYP51	98.8%	93.02%
Trichophyton rubrum	DHFR	100%	97.51%

Table S4. RMSD values for the proteins selected in the study.

Species	Protein	PDB ligand	RMSD
S. aureus	DNA gyrase	E32	-
	beta-lactamase	benzylpenicillin	0.29
	PBP	cefotaxime	0.30
P. aeruginosa	DNA gyrase	EZ6	0.25
	beta-lactamase	imipenem	0.32
	PBP	imipenem	0.44
E. coli	DNA gyrase	6G9	0.25
	beta-lactamase	ácido borônico	0.23
	PBP	-	-
M. tuberculosis	DNA gyrase	MDP	0.29
	beta-lactamase	amoxicilin	0.42
	PBP	ampicilin	0.30
C. albicans	CYP	posaconazole	0.83
	SAP-2	benzamidine	0.19

	DHFR	N22	0.61
C. tropicalis	CYP	-	-
C. glabrata	CYP	-	-
A. flavus	DHFR	H8A	0.32

Figure S1. Alignment of DNA gyrase enzyme sequences from selected bacteria in the study. The gray regions correspond to non-similar and non-identical amino acids. The red regions correspond only to identical amino acids. The yellow regions are similar amino acids. The black boxes represent the active site amino acids.

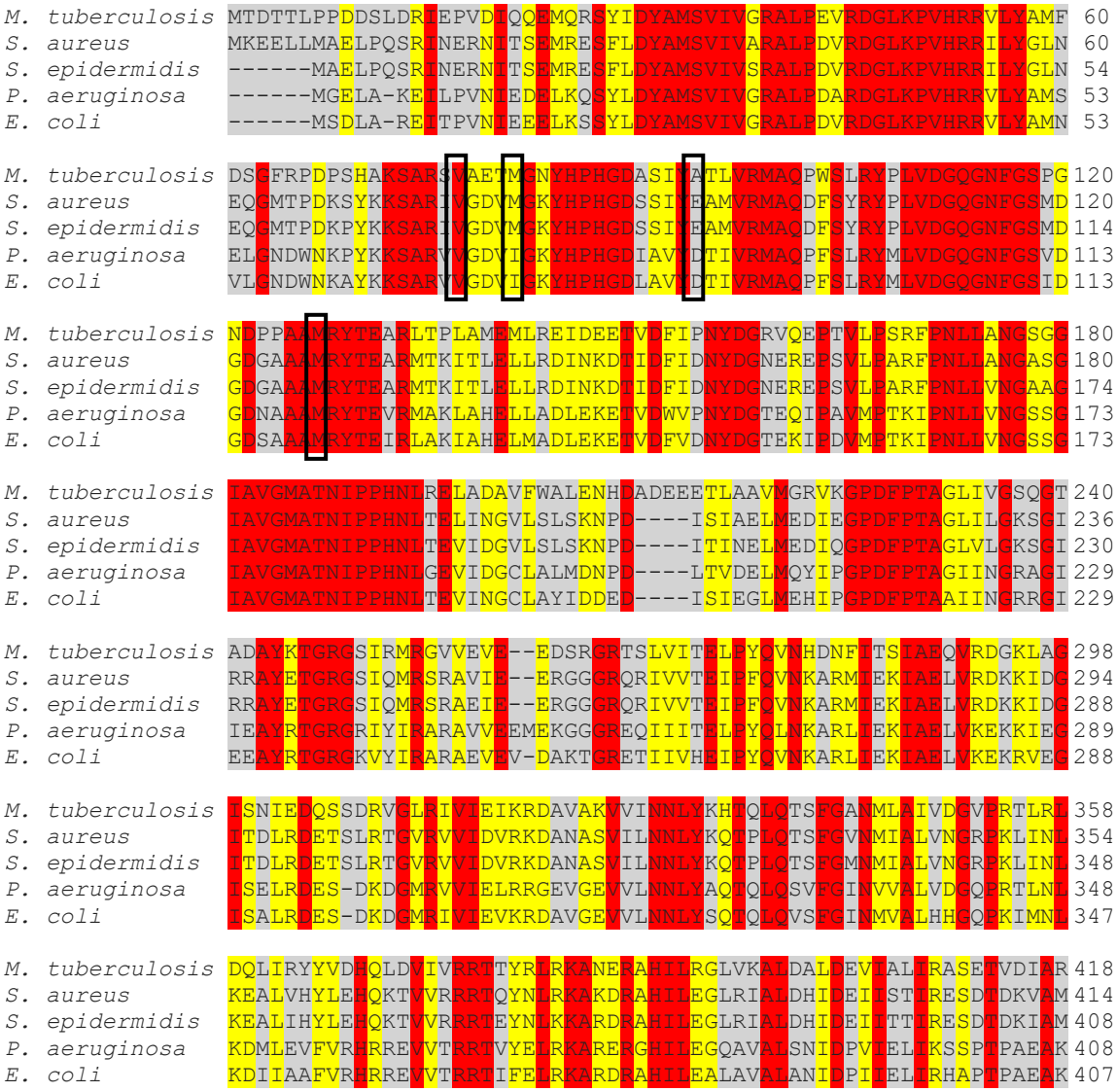
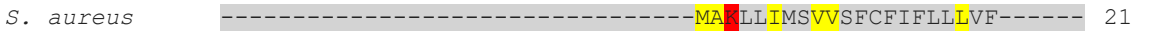


Figure S2. Alignment of beta-lactamase enzyme sequences from selected bacteria in the study. The gray regions correspond to non-similar and non-identical amino acids. The red regions correspond only to identical amino acids. The yellow regions are similar amino acids.



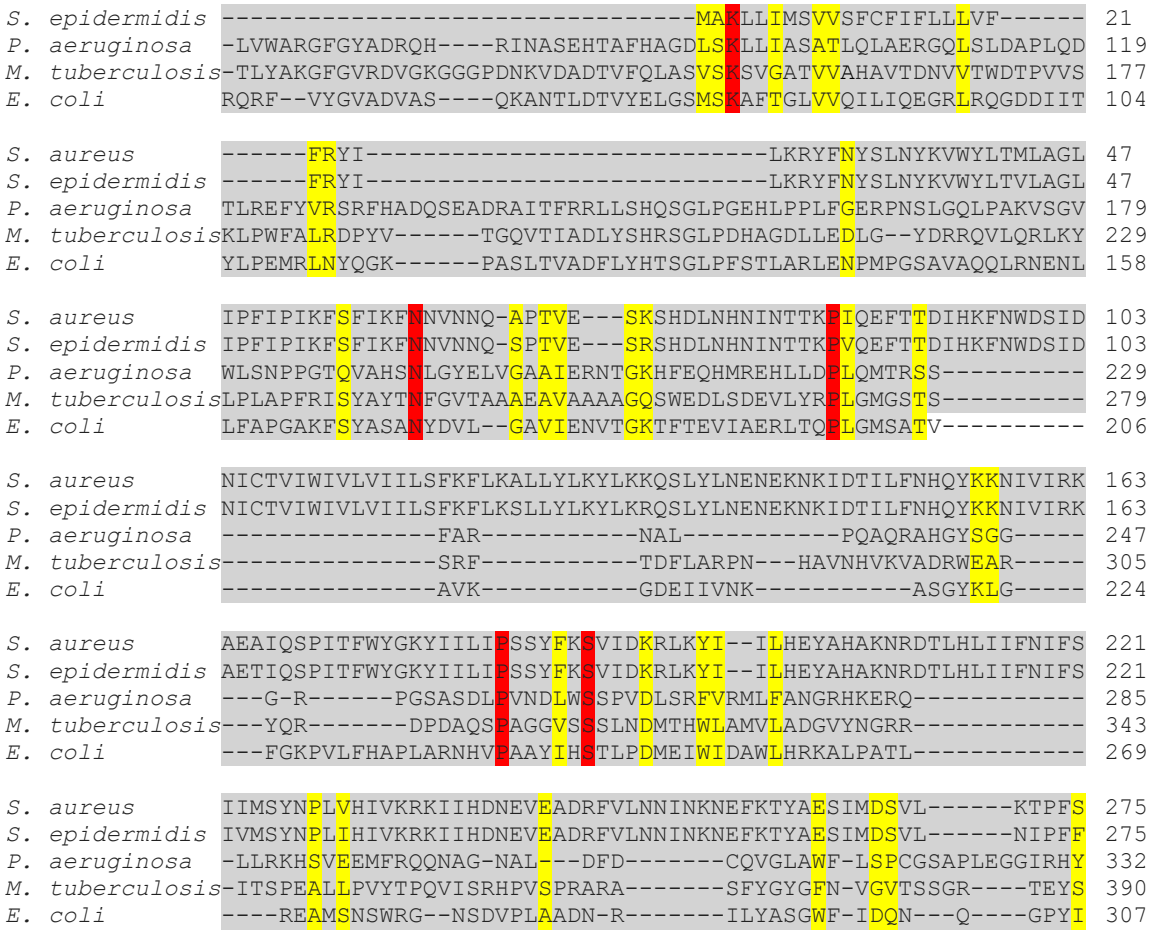


Figure S3. Alignment of PBP enzyme sequences from selected bacteria in the study. The gray regions correspond to non-similar and non-identical amino acids. The red regions correspond only to identical amino acids. The yellow regions are similar amino acids.

Figure S4. Alignment of CYP51 enzyme sequences from selected fungi in the study. The gray regions correspond to non-similar and non-identical amino acids. The red regions correspond only to identical amino acids. The yellow regions are similar amino acids. The black boxes represent the active site amino acids.

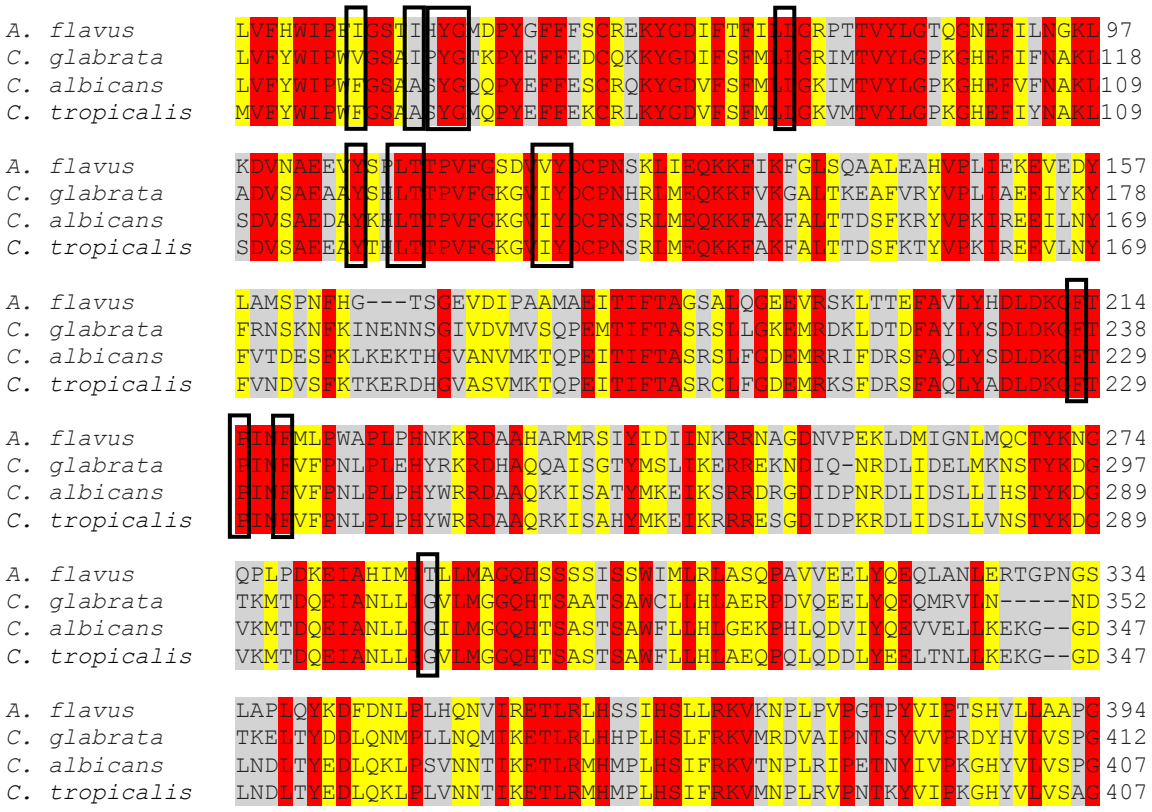
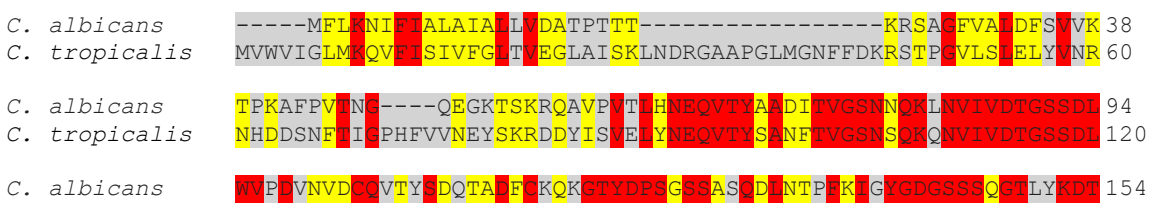


Figure S5. Alignment of SAP-2 enzyme sequences from selected fungi in the study. The gray regions correspond to non-similar and non-identical amino acids. The red regions correspond only to identical amino acids. The yellow regions are similar amino acids.



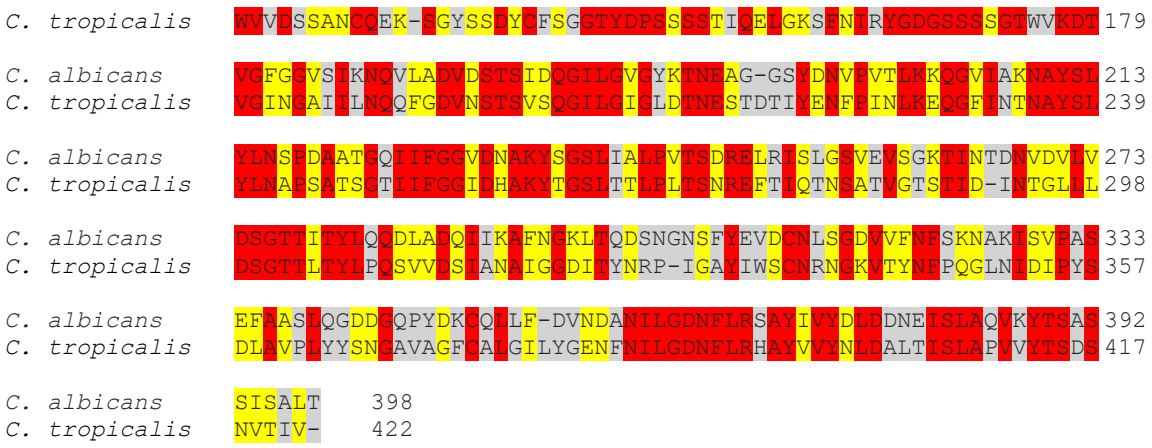


Figure S6. Alignment of DHFR enzyme sequences from selected fungi in the study. The gray regions correspond to non-similar and non-identical amino acids. The red regions correspond only to identical amino acids. The yellow regions are similar amino acids.

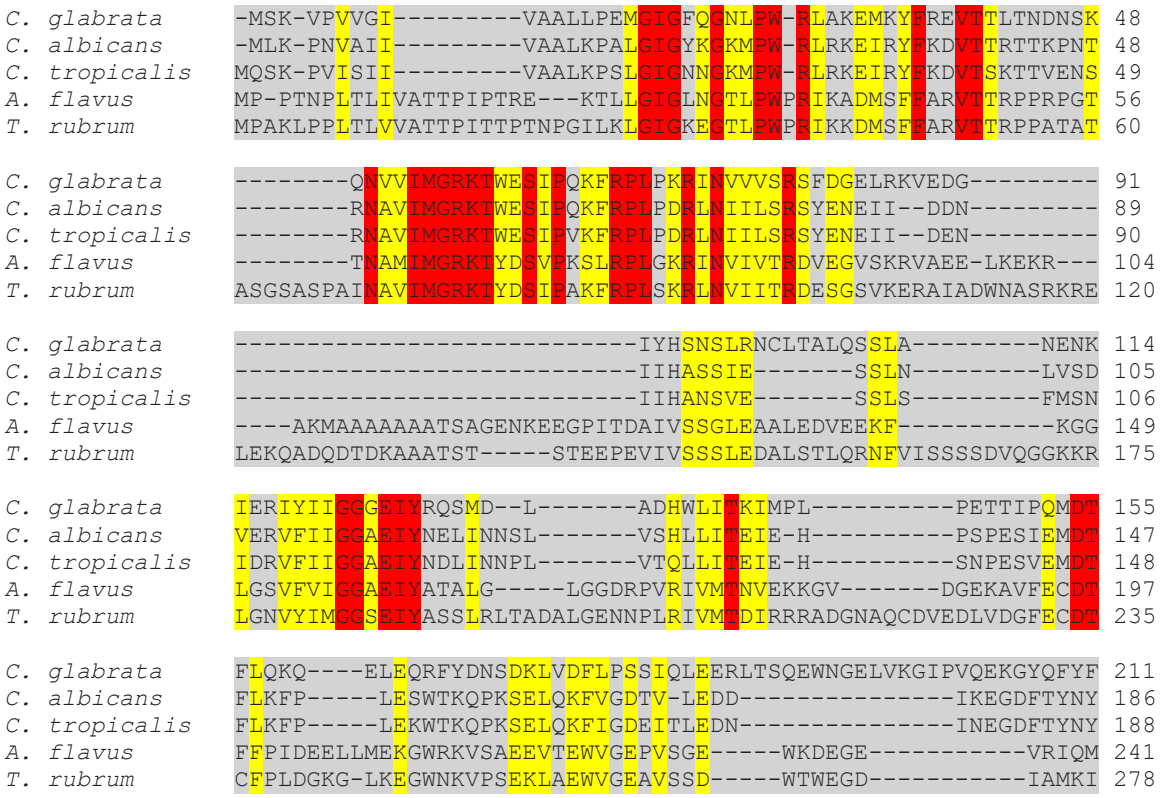
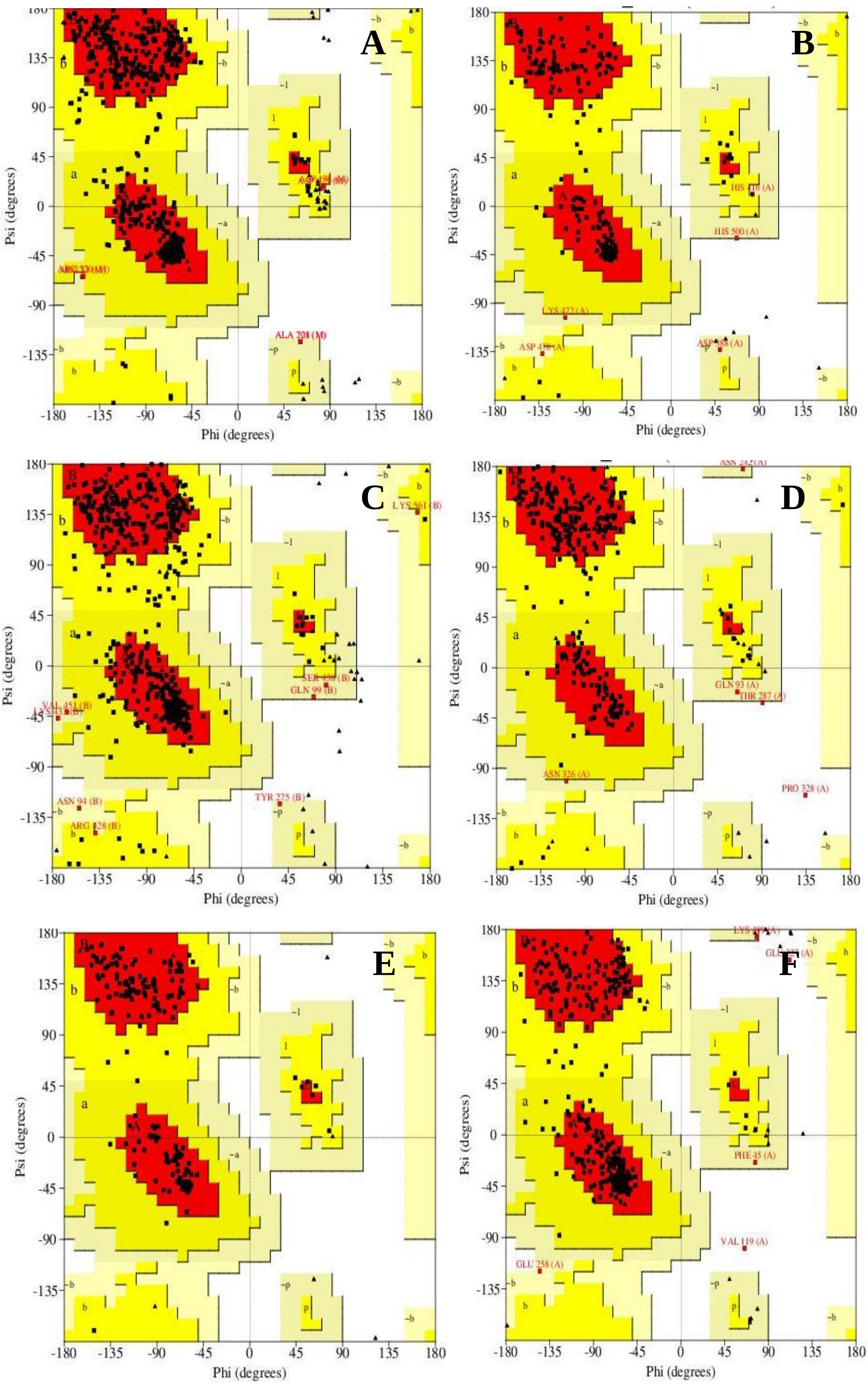
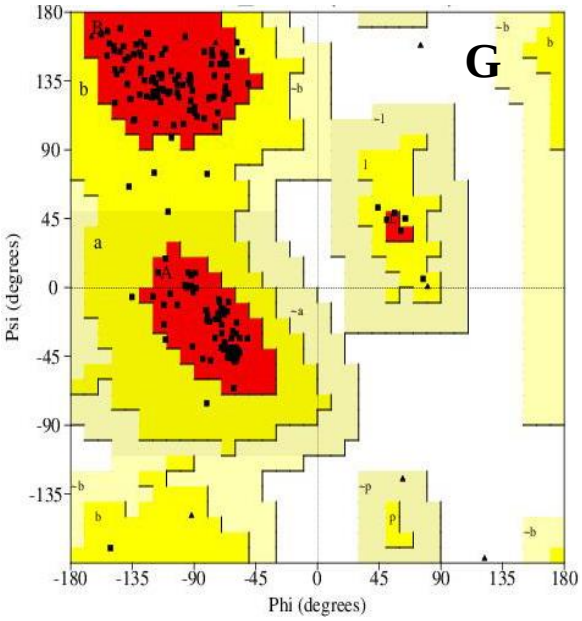


Figure S7. Ramachandran plot of homology models generated for (A) *S. epidermidis* DNA gyrase, (B) *S. epidermidis* beta-lactamase, (C) *S. epidermidis* PBP, (D) *S. tropicalis* SAP-2, (E) DHFR from *S. tropicalis*, (F) CYP51 from *A. flavus* and (G) DHFR from *T. rubrum*. The colored regions represent the allowed and favored regions of secondary structures and the white regions represent the forbidden regions.


A
B
C
D
E
F
Psi (degrees)
Phi (degrees)
180
135
90
45
0
-45
-90
-135
-180
LYS 472 (A)
ASP 478 (A)
HIS 416 (A)
HIS 500 (A)
LYS 561 (B)
SER 430 (B)
GLN 99 (B)
VAL 451 (B)
ASN 94 (B)
ARG 128 (B)
TYR 275 (B)
GLN 93 (A)
THR 287 (A)
ASN 326 (A)
PRO 328 (A)
PHE 45 (A)
VAL 119 (A)
GLU 258 (A)


G
Psi (degrees)
Phi (degrees)
180
135
90
45
0
-45
-90
-135
-180
-135
-90
-45
0
45
90
135
180
B
b
a
A
~b
~l
l
~a
~p
p