

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 Rodrigues¹, Mayara dos Santos Maia¹, Thalisson Amorim de Souza², Edeltrudes de Oliveira Lima³, Luiz Eduardo C. G. dos Santos³, Shellygton Lima Silva³, Marcelo Sobral da Silva², José Maria Barbosa Filho²; Valnês da S. Rodrigues-Junior^{4,5}, Luciana Scotti¹, Marcus Tullius Scotti¹.

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

<i>Aspergillus flavus</i>	QOR32183.1	-	RMZ41744.1
<i>Penicillium citrinum</i>	-	-	-
<i>Trichophyton rubrum</i>	-	-	OAL62276.1
<i>Microsporum canis</i>	-	-	EEQ31600.1

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

Species	Protein	ID PDB molde	Identity
<i>S. epidermidis</i>	DNA gyrase	5NPP	93.83%
	beta-lactamase	1XA1	94.90%
	PBP	3VSL	83.18%
	SAP-2	2H6S	51.20%
<i>C. tropicalis</i>	DHFR	1AOE	81.25%
<i>A. flavus</i>	CYP51	6CR2	64.03%
<i>Trichophyton rubrum</i>	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
<i>S. epidermidis</i>	DNA gyrase	99.3%	89.63%
	beta-lactamase	97.8%	100%
	PBP	98.5%	90.52%
	SAP-2	98.7%	93.15%
<i>C. tropicalis</i>	DHFR	100%	92.23%
<i>A. flavus</i>	CYP51	98.8%	93.02%
<i>Trichophyton rubrum</i>	DHFR	100%	97.51%

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

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

	DHFR	N22	0.61
<i>C. tropicalis</i>	CYP	-	-
<i>C. glabrata</i>	CYP	-	-
<i>A. flavus</i>	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.

<i>M. tuberculosis</i>	MTDTTLPPDDSLDRIEPVDFQQEMQRSYIDYAMSIVVGRALFEVRDGLKPVHRRVLYAMF	60
<i>S. aureus</i>	MKEELLMAELPQSRINERNTSSEMRSESFLDYAMSIVVARALFDVARDGLKPVHRRILYGLN	60
<i>S. epidermidis</i>	-----MAELPQSRINERNTSSEMRSESFLDYAMSIVVSRALFDVARDGLKPVHRRILYGLN	54
<i>P. aeruginosa</i>	-----MGELA-KETLPPVNEDELKOSYLDYAMSIVVGRALFDARDGLKPVHRRVLYAMS	53
<i>E. coli</i>	-----MSDLA-REITPPVNEELKSSYLDYAMSIVVGRALFDVARDGLKPVHRRVLYAMN	53
<i>M. tuberculosis</i>	DSGFRPDPSHAKSARSVAETMNYYHPHGDASIATLVRMACPWSLRYPLVDGQGNFGSPG	120
<i>S. aureus</i>	EQGGMTPDKSYKKSAFIVGDVMSKYHPHGDSSIPEAMVRMACQDFSYRYPLVDGQGNFGSMD	120
<i>S. epidermidis</i>	EQGGMTPDKPKYKKSAFIVGDVMSKYHPHGDSSIPEAMVRMACQDFSYRYPLVDGQGNFGSMD	114
<i>P. aeruginosa</i>	ELGNDWNKPYKKSAFIVGDVMSKYHPHGDIAVDTIVRMACQPFSLRMLVVDGQGNFGSVD	113
<i>E. coli</i>	VLGNDWNKAYKKSAFIVGDVMSKYHPHGDLAVDTIVRMACQPFSLRMLVVDGQGNFGSID	113
<i>M. tuberculosis</i>	NDPPAAAMRYTEARLTPLAMEMLREIDEETVDFIPNYDGRVQEFTVILPSRFPNLLANGSGG	180
<i>S. aureus</i>	GDGAAAMRYTEARMTKITLLELLRDINKDTIDFIDNYDGNEREFSVILPARFPNL LANGASG	180
<i>S. epidermidis</i>	GDGAAAMRYTEARMTKITLLELLRDINKDTIDFIDNYDGNEREFSVILPARFPNLIVNGAAG	174
<i>P. aeruginosa</i>	GDNAAAMRYTEVMAKLAHELLIADLEKETVWDVPNYDGTEQIPAVMPTKIPNLIVNGSSG	173
<i>E. coli</i>	GDSAAAMRYTEIELAKIAHELLMADLEKETVDFVDNYDGTEKIEDVMPTEKIPNLIVNGSSG	173
<i>M. tuberculosis</i>	IAVGMATNIPPHNLTRELADAVFWALENHDADEEEETVLAAMGRVKGPDFPTAGLIVLSQGT	240
<i>S. aureus</i>	IAVGMATNIPPHNLTTELINGVILSLSKNPD----ISIAELIMEDIEGPDFPTAGLILGKSGI	236
<i>S. epidermidis</i>	IAVGMATNIPPHNLTTEVIDGVILSLSKNPD----ITINELIMEDIQGPDFPTAGLVLGKSGI	230
<i>P. aeruginosa</i>	IAVGMATNIPPHNLTGEVIDGCLALMDNP----LTVDLMQYIIPGPDFPTAGIINGRAGI	229
<i>E. coli</i>	IAVGMATNIPPHNLTTEVINGCLAYIDDED----ISIEGLMEHIIPGPDFPTAAIINGRRGI	229
<i>M. tuberculosis</i>	ADAYKTGRGSIRMFGVVEVE--EDSRGRTSLVITELPYQVNHDNFITSIAFQYRDGKLAG	298
<i>S. aureus</i>	RRAYETGRGSIQMPSRSRAVIE--ERGGGRQRIVVTEIPFCQVNKARMIEKIAELVRDKKIDG	294
<i>S. epidermidis</i>	RRAYETGRGSIQMPSRAEIE--ERGGGRQRIVVTEIPFCQVNKARMIEKIAELVRDKKIDG	288
<i>P. aeruginosa</i>	IEAYRTGRGRIYIRARAVVSEMEKGGREREQIIITELPYQLNKARLIEKIAELIKEKKIEC	289
<i>E. coli</i>	EEAYRTGRGKVYIRARAEV--DAKTRGETIIVHEIPYQVNKARLIEKIAELIKEKRVEG	288
<i>M. tuberculosis</i>	ISNIEDQSSDRVGLIVIVIEIKRDAVAKVVINNNLYKHTQLQTSFGAMMLIVDGVPRTLRL	358
<i>S. aureus</i>	ITDLRDETSLRTGVVVV1DVRKDANASVILNNLYKQTPLQTSFGVMMLLVNCRPKLINL	354
<i>S. epidermidis</i>	ITDLRDETSLRTGVVVV1DVRKDANASVILNNLYKQTPLQTSFGMMMLLVNCRPKLINL	348
<i>P. aeruginosa</i>	ISELRDES-DKDGMEVVI1ELRRGEVGEVVLNNLYAQTQLQSVFGINVVA1LVDGQPRTLNL	348
<i>E. coli</i>	ISALRDES-DKDGMEVVI1EVKRDADVGEVVLNNLYSQTQLQVSFGINMMVALHHGQPKIMNL	347
<i>M. tuberculosis</i>	DQLIRYYVDHQLDIVVRRITYRURKANERAHILRGLVKALDALDEVIALRASETVDIAR	418
<i>S. aureus</i>	KEALVHYLEHQKTVVRRRTQYNLRKAKDRAHILEGLRLIAIDHIDEIISTRESDDTKVAM	414
<i>S. epidermidis</i>	KEALIHYLEHQKTVVRRRTYEYLKKARDRAHILEGLRLIAIDHIDEIITRESDDTKIAM	408
<i>P. aeruginosa</i>	KDMLEVFVRHRREVVTTRRTVYELRKARERGHILEGQAVALSNIIDPVIELKSSPTPAEAK	408
<i>E. coli</i>	KDIIIAAFVRHRREVVTTRRTIFEYLRKARDRAHILEALAVALANILPTELERHAPTPAEAK	407

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.

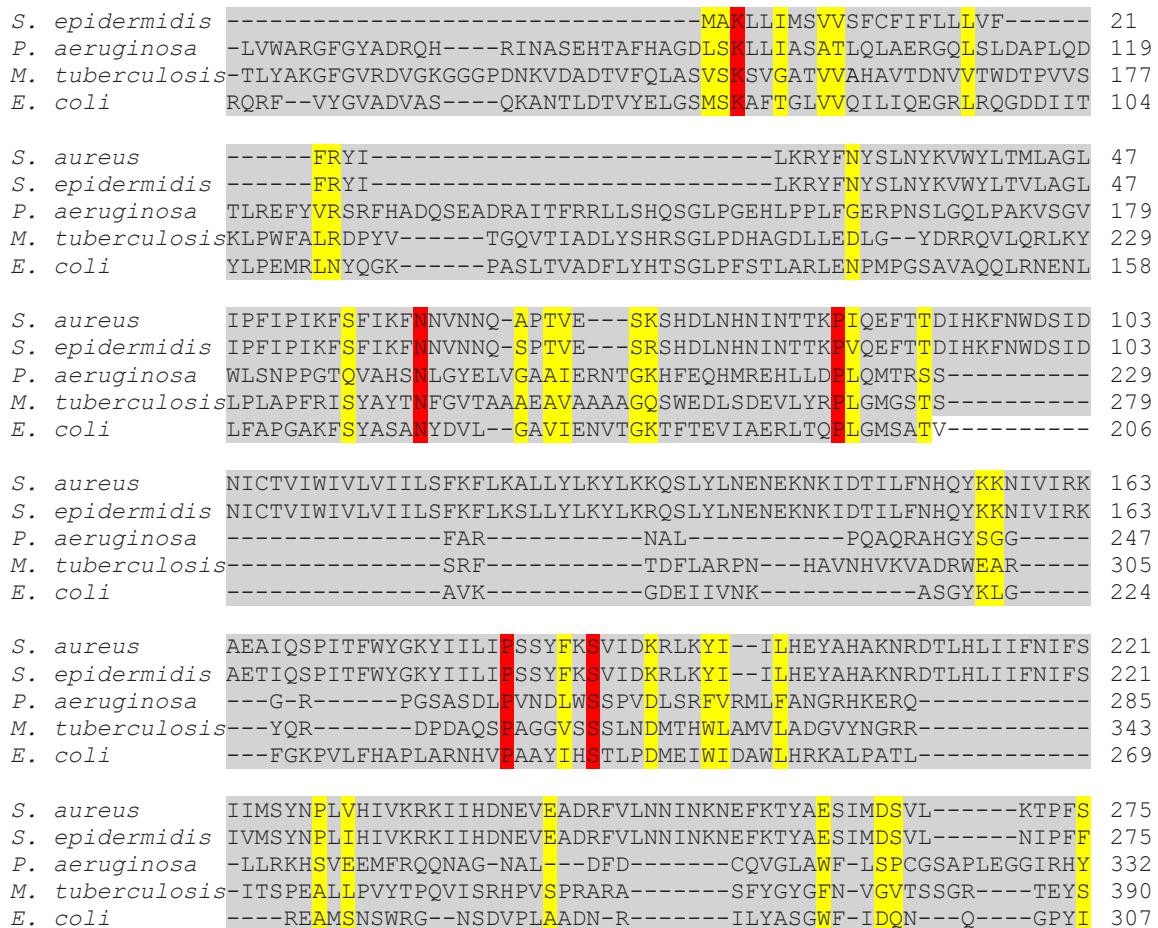
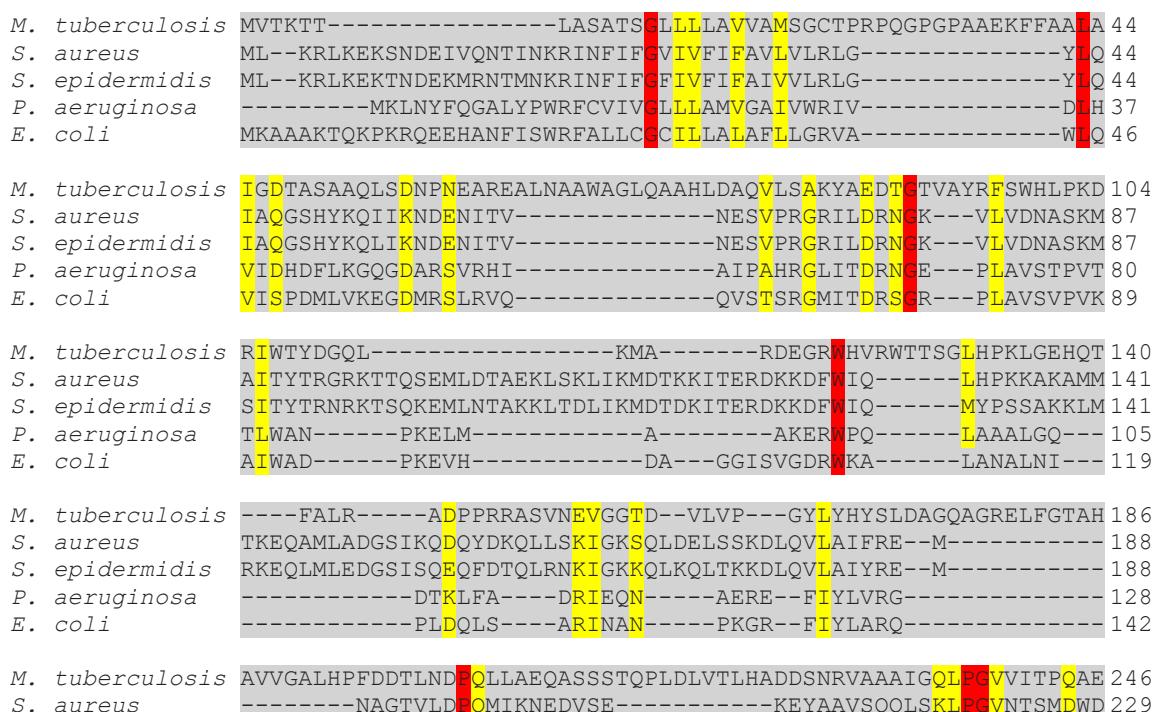


Figure S3. Alignment of PBP enzyme sequences from selected bacteria in the study. The gray regions correspond to non-similar and non-identical amino acids. The red regions correspond only to identical amino acids. The yellow regions are similar amino acids.



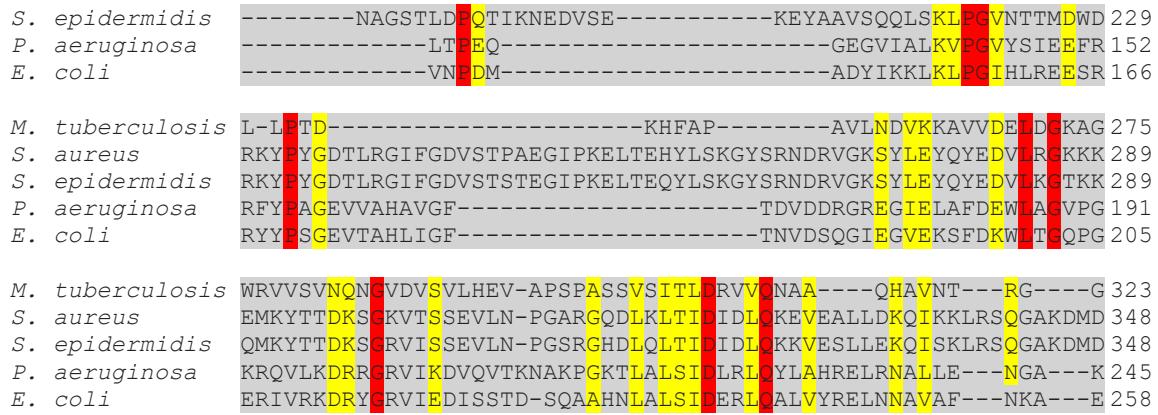


Figure S4. Alignment of CYP51 enzyme sequences from selected fungi in the study. The gray regions correspond to non-similar and non-identical amino acids. The red regions correspond only to identical amino acids. The yellow regions are similar amino acids. The black boxes represent the active site amino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

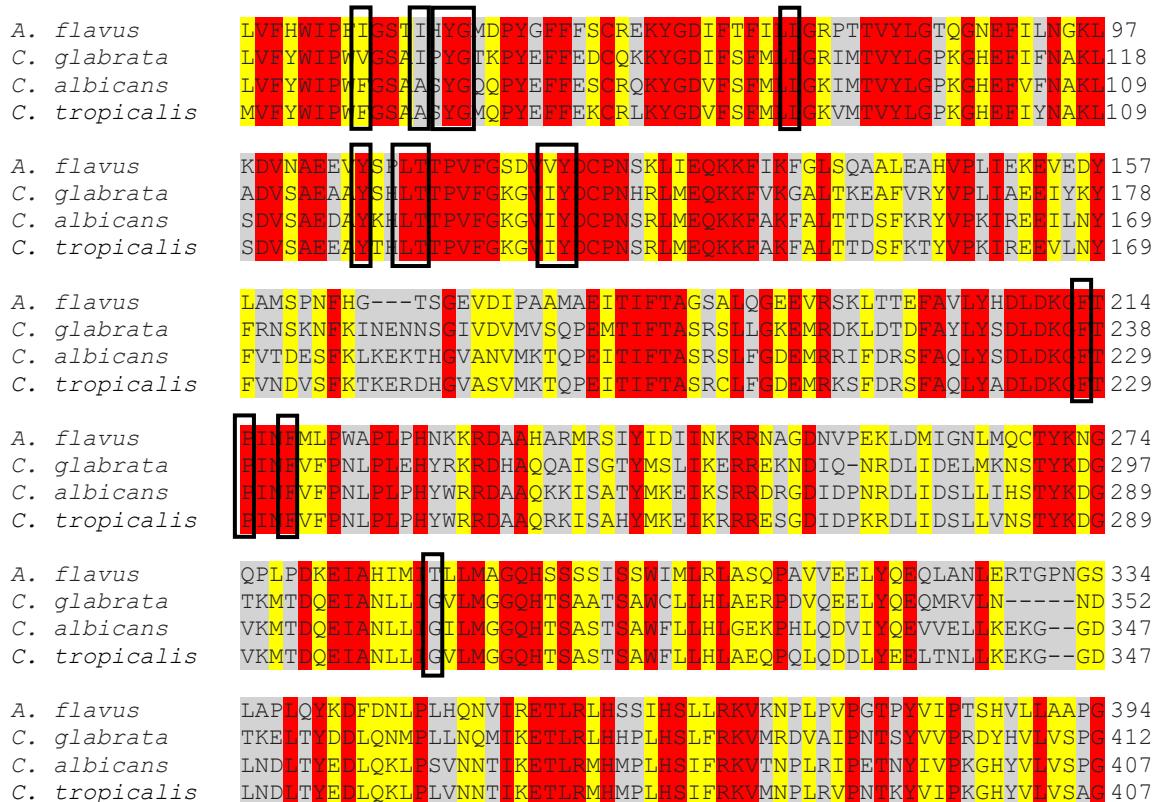
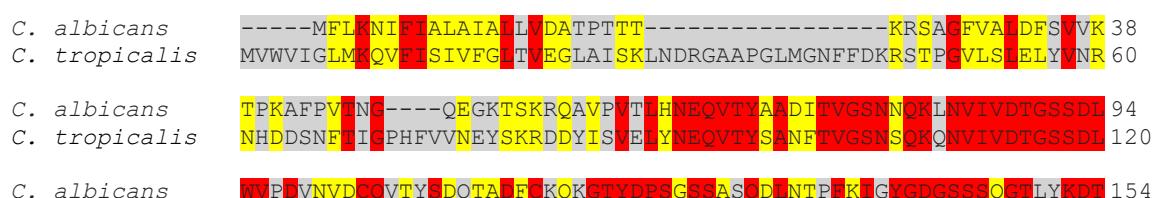


Figure S5. Alignment of SAP-2 enzyme sequences from selected fungi in the study. The gray regions correspond to non-similar and non-identical amino acids. The red regions correspond only to identical amino acids. The yellow regions are similar amino acids.



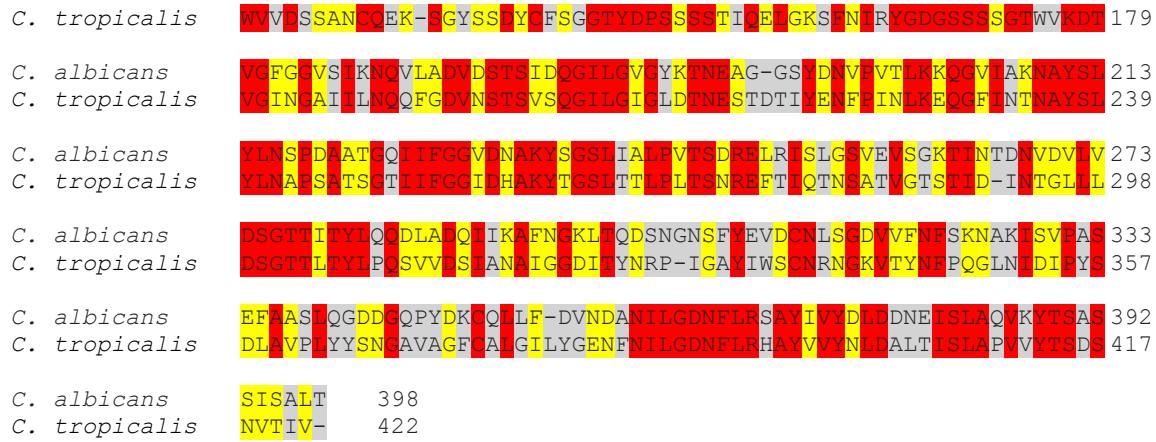


Figure S6. Alignment of DHFR enzyme sequences from selected fungi in the study. The gray regions correspond to non-similar and non-identical amino acids. The red regions correspond only to identical amino acids. The yellow regions are similar amino acids.

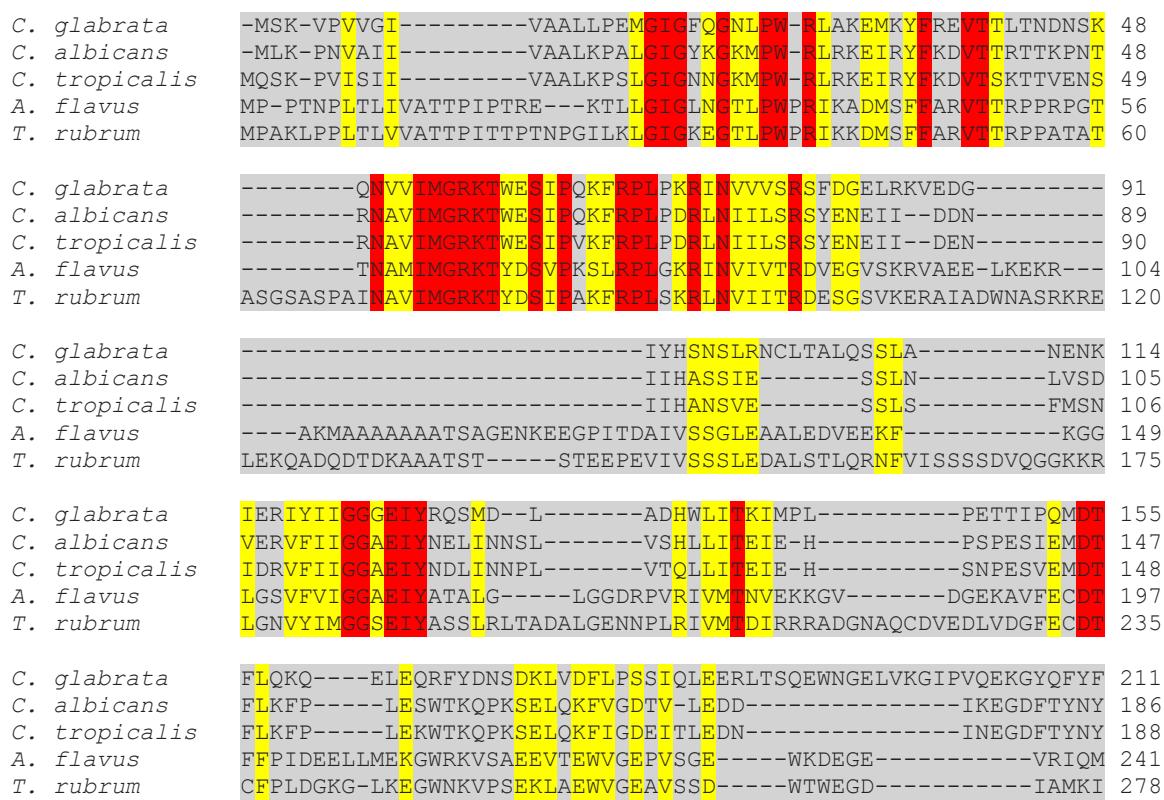


Figure S7. Ramachandran plot of homology models generated for (A) *S. epidermidis* DNA gyrase, (B) *S. epidermidis* beta-lactamase, (C) *S. epidermidis* PBP, (D) *S. tropicalis* SAP-2, (E) DHFR from *S. tropicalis*, (F) CYP51 from *A. flavus* and (G) DHFR from *T. rubrum*. The colored regions represent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

