

N-肉豆蔻酰基转移酶家族的进化踪迹分析研究

盛春泉¹, 朱杰², 张万年^{1*}, 徐辉¹, 缪震元¹, 姚建忠¹, 张珉¹

(第二军医大学 1. 药学院, 2. 科研部, 上海 200433)

摘要: 为了阐明抗真菌药物作用新靶点——*N*-肉豆蔻酰基转移酶(NMT)活性位点中的关键功能残基分布, 指导设计特异性抑制剂, 开展了 NMT 家族的蛋白多元序列联配研究, 并构建了蛋白系统进化树。在此基础上, 采用进化踪迹分析技术识别得到了 NMT 活性位点中肉豆蔻酰 CoA 结合位点、催化反应中心和抑制剂结合位点的重要功能残基。通过对白色念珠菌 NMT 活性位点中药物结合位点的研究发现, Trp126, Asn175 和 Thr211 是 NMT 家族的高度保守残基, 而且不与已有的抑制剂发生直接的相互作用, 因此将是发现新型 NMT 抑制剂的重要药物结合位点。亚家族特异性残基 Pro338, Leu350, Ile352 和 Ala353 可作为已有抑制剂结构优化的重要位点, 据此设计的新化合物将有望进一步提高对真菌 NMT 的选择性。进化踪迹分析结果为进一步阐明 NMT 结构-功能关系和新型 NMT 抑制剂类抗真菌药物的设计提供重要信息。

关键词: *N*-肉豆蔻酰基转移酶; 进化踪迹分析; 药物结合位点

中图分类号: R916.693 文献标识码: A 文章编号: 0513-4870(2007)02-0157-09

Evolutionary trace analysis of *N*-myristoyltransferase family

SHENG Chun-quan¹, ZHU Jie², ZHANG Wan-nian^{1*}, XU Hui¹, MIAO Zhen-yuan¹,
YAO Jian-zhong¹, ZHANG Min¹

(1. School of Pharmacy, 2. Department of Science and Research, Second Military Medical University, Shanghai 200433, China)

Abstract To clarify the important functional residues in the active site of *N*-myristoyltransferase (NMT), a novel antifungal drug target and to guide the design of specific inhibitors, multiple sequence alignments were performed on the NMT family and thus evolutionary trace was constructed. The important functional residues in myristoyl CoA binding site, catalytic center and inhibitor binding site of NMT family were identified by ET analysis. The trace residues were mapped onto the active site of CaNMT. Trp126, Asn175 and Thr211 are highly conserved trace residues and do not interact with current NMT inhibitors, which are potential novel drug binding sites for the novel inhibitor design. Pro338, Leu350, Ile352 and Ala353 are class-specific trace residues, which are important for the optimization of current NMT inhibitors. The trace residues identified by ET analysis are of great importance to study the structure-function relationship and also to guide the design of specific inhibitors.

Key words *N*-myristoyltransferase; evolutionary trace analysis; drug binding site

N-肉豆蔻酰基转移酶(*N*-myristoyltransferase, NMT)是一种细胞溶质单体酶, 它催化将肉豆蔻酸(14个碳原子的饱和脂肪酸)从肉豆蔻酰 CoA 转移至某些真核细胞蛋白或病毒蛋白 *N* 端的甘氨酸上,

形成 *N*-肉豆蔻酰化蛋白^[1,2]。目前已发现有 100 多种可肉豆蔻酰化的蛋白, 主要包括 GDP 结合蛋白、激酶、磷酸酶和钙结合蛋白等。它们分布在细胞的不同位置并具有不同的生物学功能, 主要是参与二级信号转导、囊泡转运、病毒成熟和肿瘤形成等一些重要的生物学过程^[3,4]。NMT 具有 I 型和 II 型同工酶, 分别位于内质网和细胞质中。NMT 催化反应的历程是一个“有序的双-双反应机制”(ordered birbi

收稿日期: 2006-06-26

基金项目: 国家自然科学基金资助项目(30400567)。

* 通讯作者 Tel / Fax: 86-21-25074460

E-mail: zhangwnk@hotmail.com

reaction mechanism)^[5-6]。首先, NMT特异性识别肉豆蔻酰 CoA, 并与其结合形成 NMT-肉豆蔻酰 CoA 二元复合物。肉豆蔻酰 CoA 与 NMT 结合后共同组成了一个功能性的肽类底物结合位点, 然后识别特异性肽类(蛋白)底物形成 NMT-肉豆蔻酰 CoA-肽类底物的三元复合物。最后, 肉豆蔻酸部分被催化转移至肽类(蛋白)底物 N 端的甘氨酸上, 自由的 CoA 首先释放, N-肉豆蔻酰化产物随之释放。

大量的研究已经证实 NMT 是白色念珠菌和新型隐球菌等临床重要致病真菌生存所必须的酶, 基于 NMT 开发新一代的抗真菌药物有望克服现有抗真菌药物疗效有限、毒副作用大和耐药问题严重等缺陷^[7]。目前真菌 NMT 抑制剂的研究尚处于起步阶段, 主要集中在对苯并咪唑类化合物进行结构修饰^[8-9], 难以取得突破。因此, 如果能深入分析真菌 NMT 活性部位的功能位点和特异性残基, 不仅可以阐明它们在蛋白质进化中的意义, 而且可以发现新的药物作用位点, 为设计新结构类型、高活性和高特异性的先导化合物提供思路。

白色念珠菌 NMT (CaNMT) 和酿酒酵母 NMT (ScNMT) 的晶体结构已被测定^[10-11], 这为基于靶酶活性位点开展合理药物设计奠定了坚实基础。Lichtarge 等^[12]发展的进化踪迹分析 (evolutionary trace analysis, ET) 技术是一种识别蛋白质特异性功能残基进而发现活性位点、蛋白功能表面等的重要工具。对蛋白质家族进行进化踪迹分析可以深入了解蛋白质结构-功能关系, 并指导进行合理药物设计。目前, 进化踪迹技术已经成功应用于识别 G 蛋白偶联受体和拓扑异构酶^[13-15]等药物作用靶点活性位点的重要功能残基。本文首先开展了 NMT 家族的多元序列匹配研究, 然后通过 ET 技术识别得到了 NMT 家族的保守残基和真菌亚家族特异性残基, 并探讨了它们的生物学意义, 进而识别得到新的药物结合位点, 为开展 NMT 抑制剂的全新设计和虚拟高通量筛选提供重要信息。

进化踪迹分析原理

在蛋白质的进化过程中, 对蛋白质结构和功能有重要作用的残基发生突变的几率很低, 即使这些功能残基发生突变也往往是为了寻求蛋白亚家族功能的特异性。目前, 可以通过 X 射线衍射技术测定蛋白质晶体结构和定点突变等实验技术来发现和确证这些重要功能残基。然而, 仅仅依靠实验方法研

究这些功能残基不仅费时费力, 而且效率低下。进化踪迹分析方法则通过研究蛋白质进化过程中氨基酸序列的保守性和蛋白质结构中保守残基的三维排列来发现蛋白质家族中重要的功能残基及其结合表面。进化踪迹研究结果可以为定点突变残基的选取和基于蛋白质结构的抑制剂设计提供重要信息。

Lichtarge 等发展的进化踪迹分析方法^[12]的计算原理主要分成以下 3 个步骤: 首先, 将蛋白质家族的氨基酸序列进行多元序列匹配, 在此基础上构建蛋白质进化树。通过定义序列间相似性的阈值可以将蛋白质家族分成不同的亚家族, 阈值越高蛋白质亚家族数目越少, 功能的特异性越强。其次, 基于多元序列匹配结果, 发现在整个蛋白质家族中保守的残基 (conserved residues) 和仅在蛋白亚家族内保守的亚家族特异性残基 (class-specific residues)。一般情况下, 保守残基对维持蛋白功能是必需的, 而亚家族特异性残基则与亚家族的功能特异性相关。保守残基和亚家族特异性残基统称为踪迹残基 (trace residues)。定义不同的序列等同率阈值 (percentage identity cutoff, PIC), 可识别得到不同的踪迹残基, 通常采用的 PIC 数值为 30%, 40%, 50%, 60%, 70%, 80% 和 90%。最后, 通过聚类分析方法将踪迹残基映射到蛋白质家族中一个已知结构的蛋白上, 并在三维空间加以聚类, 得到踪迹残基簇 (trace residue cluster), 反映了维持蛋白质结构所必需的一些重要结构特征。

材料与方法

序列与结构 所有 NMT 的氨基酸序列取自 NCBI (<http://www.ncbi.nlm.nih.gov>), 按照其来源分成以下 4 种类型。真菌 NMT: *Candida albicans* (AA26048), *Saccharomyces cerevisiae* (AAB67436), *Candida glabrata* (AA26048), *Histoplasma capsulatum* (AAA17549), *Aspergillus fumigatus* (BAA8765), *Schizosaccharomyces pombe* (O43010), *Cryptococcus neoformans* (AAA17547); 哺乳动物 NMT: human I (NP066565), *Bus taurus* I (AAF31460), *Mus musculus* I (NP032733), human II (NP004799), *Bus taurus* II (AAF36406), *Mus musculus* II (NP032734); 原生生物 NMT: *Drosophila melanogaster* (AAD27855); *Caenorhabditis elegans* (P46548), *Plasmodium falciparum* (AAF18461); 植物 NMT: *Arabidopsis thaliana* I (AAF60968)。CaNMT 与苯并咪唑类抑制剂复合物

<i>C. albicans</i>	-----MSGDNTGNKNSAPSKEIE	19
<i>S. cerevisiae</i>	-----MSEEDKAKKLE	11
<i>C. glabrata</i>	-----MSEKKIE	7
<i>H. capsulatum</i>	MSEQEGNQSEHQSEHVGESEGLPNETPTTSQSTNASTGTAGKGEKSSDGDPAANPAT	60
<i>A. fumigatus</i>	-----MSDSKDRKGAPEGQSSEKKGAV	24
<i>S. pombe</i>	-----MDNENKNTKNSQQDSFSEGGIR	24
<i>C. neoformans</i>	-----MSSDNKAATDEEIR	15
Human I	MADESETAVKPPAPPLPQMMEGNGNGHEHCSDCENEEDNSYNRGLSPANDTGAKKKKKK	60
<i>B. taurus</i> I	MGDESETAVKPPAPPLPQMMEGNGNGHEHCSDCENEEDNSYNRGLSPANDTGAKKKKKK	60
<i>M. musculus</i> I	MADESETAVKLPAPSLPLMMEGNGNGHEHCSDCENEEDNSHNRSGLSPANDTGAKKKKKK	60
Human II	MAEDSESAASQQSLELDDQDTCG-----IDGDNEEETEHAKEG---SPGGYLGAKKKKKK	51
<i>B. taurus</i> II	MAEDSESAASQQSLELDDQDTCG-----IDGDNEEETEHAKEG---SPGGYLGAKKKKKK	51
<i>M. musculus</i> II	MAEDSESAASQQSLELDDQDTCG-----IDGDNEEETEHAKEG---SPGGDLGAKKKKKK	51
<i>D. melanogaster</i>	MPNENAEDLSGQELKQKAKEVAD-----AS-----EAMLEKVVAG	35
<i>C. elegans</i>	MSHGSHDGPCCGGHGGDGGAG-----SRPSVNDVQAL	34
<i>P. falciparum</i>	-----MN	2
<i>A. thaliana</i> I	-----MADNNSPPGSVEKQADQIVEANP	23
<i>C. albicans</i>	ELLKLLAMG-----	28
<i>S. cerevisiae</i>	NLLKLLQLNDDT-----	24
<i>C. glabrata</i>	ELLKLLSMNNGDM-----	20
<i>H. capsulatum</i>	KLTPSMAESLLELNPALRSE-----	80
<i>A. fumigatus</i>	NITPQMAESLLENNPALRNE-----	44
<i>S. pombe</i>	ELLDRLALRSLTE-----	37
<i>C. neoformans</i>	RALKAADLMKILDG-----	29
Human I	QKKKKEKGETDSAQDQP-----	78
<i>B. taurus</i> I	QKKKKEKGETDSAQDQP-----	78
<i>M. musculus</i> I	QKKKKEKGSDEMESTQDQP-----	78
Human II	QKRKKEKPNSGGTKSDSASDSQEIKIQQPS-----	81
<i>B. taurus</i> II	QKRKKEKPNSGGTKSDSASDSQEIKIQQPS-----	81
<i>M. musculus</i> II	QKRKKEKPNSGGTKSDSASDSQEIKIQQSSKHNAIWQQISAGAAMGGDTMEGEWIDL RMY	111
<i>D. melanogaster</i>	LNIQDTASTNAAGNEDAE-----	53
<i>C. elegans</i>	VDQLRLAGVDVSNMNPIN-----	52
<i>P. falciparum</i>	DDKKDFVG-----	10
<i>A. thaliana</i> I	LVKDDTSLETTVRR-----	37
<i>C. albicans</i>	-----QELSPAQQK---EMKDYKFWKTQPVPVSLSETVTEEG---	61
<i>S. cerevisiae</i>	-----SKFTQEQQK---AMKDHKFWRTQPVKDFDEKVVVEEG---	57
<i>C. glabrata</i>	-----SKLTANQRK---EMKEYKFWKTQPVTKFDEEVKEEG---	53
<i>H. capsulatum</i>	-LAGMDKEKATEALRQMNISDLLTGLSVNPKNQKDMASFQFWQTQPVIRFDDRESESPD--	138
<i>A. fumigatus</i>	-TAGMDKDKAAEAMRKMNAEELLTGLSVSGKNQKDMASYKFWQTQPVPRFDETSDDTG---	101
<i>S. pombe</i>	-----K---EAAAAPP-----KTYEDFKFWKTQPVPKFDECTQEG---	70
<i>C. neoformans</i>	-----KMALGNKS-GTKNLGEHKFWKTQPVPQITGSGASAPMEE	67
Human I	VKMNSLPAERIQEIQKAIELFSVGGQPAKTMEEASKRSYQFWDTPVPVKLGEVVNT----H	135
<i>B. taurus</i> I	VKMNSLPAERIQEIQKAIELFSVGGQPAKTMEEASKRSYQFWDTPVPVKLGEVVNT----H	135
<i>M. musculus</i> I	VKMTSLPAERIQEIQKAIELFSVGGQPAKTMEEASKRSYQFWDTPVPVKLGEVVNT----H	135
Human II	-KNPSVPMQKLQDIQRAMELLSACQGPARNIDEAAKHRYQFWDTPVPVKLDEVITS----H	137
<i>B. taurus</i> II	-KNPSVPMQKLQDIQRAMELLSACQGPARNIDEAAKHRYQFWDTPVPVKLDEVITS----H	137
<i>M. musculus</i> II	HKNPTIPIQKLQDIQRAMELLSACQGPARNIDEATKRRYQFWDTPVPVKLNEVITS----H	168
<i>D. melanogaster</i>	-QPDGAKNEASVSANASKQALLQAVSDAMASTRQMAKFAFVSTQPVTKLDEQVTT----N	109
<i>C. elegans</i>	-----TAPRDMDEARSKSFQFVSTQPVPQMDTEVPADV---N	86
<i>P. falciparum</i>	-----RDLYQLIR-----NAKDKIKIDYKFWYTPVPKINDEFDENY---N	48
<i>A. thaliana</i> I	-----FQDSMSEAKTHKFWETQPVGQFKDIGDTSPL-E	69
<i>C. albicans</i>	-PIDKLTPTDVPNDPLPLISDFEWSTLIDIDNLQLELYKLLYDNYVEDIDATFRFKYSH	121
<i>S. cerevisiae</i>	-PIDKPTPEDISDKPLPLSSFEWCSIDVDNKKQLEDVFLNENYVEDRDAGFRFN YTK	117
<i>C. glabrata</i>	-PIHEEKT PADIPDEPLPLPDFEWC AIDVDEKQLEDVFLNENYVEDRDASFRFN YTR	113
<i>H. capsulatum</i>	-GPIKIVELDQVSREP IPLVDGFVWVTLIDDEADVKEFYELLANHYVEDGSAMFRFN YSP	198
<i>A. fumigatus</i>	-GPIKIIDPEKVSKEPDALLEGFEWATLDTNETELQELWDLTYHYVEDDNAMFRFN YSQ	161
<i>S. pombe</i>	-PIDPNTDINQVREPYRLKKEFEWATIDVTNDNELSEVHELLTENYVEDATAMLRFA YIS	130
<i>C. neoformans</i>	GPIDDPKTPADV KQEPGVLPA GF EWSTIDINDEEQSKEYVYLLCENYVEDDDAMFRFN YSR	128
Human I	GPVEPDK---DNIRQEPYTL PQGFTWDALDLGDRGV LKELYTLLNENYVEDDDNMFRFDY SP	194
<i>B. taurus</i> I	GPVEPDK---DNIRQEPYTL PQGFTWDALDLGDRGV LKELYTLLNENYVEDDDNMFRFDY SP	194
<i>M. musculus</i> I	GPVEPDK---DNIRQEPYTL PQGFTWDALDLGDRGV LKELYTLLNENYVEDDDNMFRFDY SP	194
Human II	GAIEPDK---DNVRQEPYSL PQGF MWDTLDLSDAEVLKELYTLLNENYVEDDDNMFRFDY SP	196
<i>B. taurus</i> II	GAIEPDK---DNVRQEPYSL PQGF MWDTLDLSDAEVLKELYTLLNENYVEDDDNMFRFDY SP	196
<i>M. musculus</i> II	GAIEPDK---DNIRQEPYSL PQGF MWDTLDLSNAEVLKELYTLLNENYVEDDDNMFRFDY SP	227
<i>D. melanogaster</i>	ECIEPNKEISEIRALPYTLPGGFKWVTLDLNDANDL KELYTLLNENYVEDDDAMFRFDY QP	170
<i>C. elegans</i>	CAIEENIALDKVRAEPPSLPAGFRWSNVDLSDEEQLNELYNLLTRNYVEDDDSMFRFDY SA	147
<i>P. falciparum</i>	EPFISDNKVEDVRKEEYKLP SGYAWCVCDITKENDRSDIYNLLTDNYVEDDDNVFRFN YSS	109
<i>A. thaliana</i> I	GPIEPATPLSEVKQEPYNLPSVYEWTTCDMNSDDMCSEYVYLLKNNYVEDDENMFRFN YSK	130

<i>C. albicans</i>	EFFQWALKPPGWKRDWHVGVVVKSTGKLVAFIAATPVTFKLNKSNKVIDSVEINFLCIHKK	182
<i>S. cerevisiae</i>	EFFNWALKSPGWKDDWHIGVVRVKE TQKLVAFISAIPVTLGVR—GKQVPSVEINFLCVHKQ	176
<i>C. glabrata</i>	EFFNWALKSPGWTPDWHIGVVRVKAASKL IAFISAIPVRLRVR—AKVIDSVEINFLCVHKQ	172
<i>H. capsulatum</i>	AFLNWALKAPGWKREWHVGVVRASKSGKLVASICGVPAEIAVR—GKSLK—VTEINFLCVHKK	257
<i>A. fumigatus</i>	SFLHWALMSPGWKKEWHVGVVRATKSRKLVASICGVPTENVR—NQKVK—VVEINFLCIHKK	220
<i>S. pombe</i>	EFLRWALMPPGYVKEWHVGVVRVKSRRKLVAFISAVPLSIRVR—DKI IKKCAEVNFLCIHKK	190
<i>C. neoformans</i>	EFLWALTPAGYLPDWHIGVVRVQTKKLVAFISGIKIDIRVR—AKTFPAAEINFLCVHKK	187
<i>Human I</i>	EFLWALRPPGWLPQWHCGVVRVSSRKL VGFISAIPANIHI—YDTEKKMVEINFLCVHKK	253
<i>B. taurus I</i>	EFLWALRPPGWLPQWHCGVVRVSSRKL VGFISAIPANIHI—YDTEKKMVEINFLCVHKK	253
<i>M. musculus I</i>	EFLWALRPPGWLPQWHCGVVRVSSRKL VGFISAIPANIHI—YDTEKKMVEINFLCVHKK	253
<i>Human II</i>	EFLWALRPPGWLLQWHCGVVRVSSNKKLVGFISAIPANIRI—YDSVKKMVEINFLCVHKK	255
<i>B. taurus II</i>	EFLWALRPPGWLLQWHCGVVRVSSNKKLVGFISAIPANIRI—YDSVKKMVEINFLCVHKK	255
<i>M. musculus II</i>	EFLWALRPPGWLLQWHCGVVRVSSNKKLVGFISAIPANIRI—YDSVKKMVEINFLCVHKK	255
<i>D. melanogaster</i>	EFLKWSLQPPGWKRDWHVGVVRVEKSGKLVGFISAIPSKLKS—YDKVLKVV D INFLCVHKK	229
<i>C. elegans</i>	DFLKWALQVPGFRPEWHCGVRADSNRLLAFI GAVPQTVRV—YDKTVNMVEINFLCVHKN	206
<i>P. falciparum</i>	EFLWALSSPNYKWNHIGVKYESTNKL VGFISAIPIDMCV—NKNIKMAEVNFLCVHKS	168
<i>A. thaliana I</i>	EFLRWALRPPGYYSWHIGVRAKTSKLL VAFISGVPARIRVR—DEVVKMAEINFLCVHKK	189
<i>C. albicans</i>	LRNKRLAPVLIKEITRRVKNQNIWQALY TGGSI LPTPLTTCRYQHRPINWSKLDHVGFSHL	243
<i>S. cerevisiae</i>	LRSKRLTPVLIKEITRRVKNCDIWHALY TAGIVLPAPVSTCRYTHRPLNWKLYEVDFTGL	237
<i>C. glabrata</i>	LRSKRLTPVLIKEITRRVKNRNIWHALY TAGVVLPAVSTCRYAHRPLNWDKLYEVQFTDL	233
<i>H. capsulatum</i>	LRSKRLTPVLIKEITRRCYLNGIYQAIY T VGI MLPTPV SACRYYHRALDWLKLHEVGF SPL	318
<i>A. fumigatus</i>	LRSKRLTPVLIKEITRRCYLNGIYQAIY TAGVVLPTPVSSCRYYHRPLDWLKLHEVGF SPL	281
<i>S. pombe</i>	LRSKRLTPLLKEVTRRCHLENVQAVYTAG VLLSPVLSRYMHRSLNWKLYDIGFAPF	251
<i>C. neoformans</i>	LRSKRLAPVLIKEVTRRVNLTNIWQAIY TAGVILPTPIGTCRYFHRNLNPPKLV DIGFSPL	248
<i>Human I</i>	LRSKRVAPVLIREITRRVHLEGI FQAVYTAGV VLPKPVGTCRYWHRSLNPRKLEVKFSHL	314
<i>B. taurus I</i>	LRSKRVAPVLIREITRRVHLEGI FQAVYTAGV VLPKPVGTCRYWHRSLNPRKLEVKFSHL	314
<i>M. musculus I</i>	LRSKRVAPVLIREITRRVHLEGI FQAVYTAGV VLPKPVGTCRYWHRSLNPRKLEVKFSHL	314
<i>Human II</i>	LRSKRVAPVLIREITRRVNLEGI FQAVYTAGV VLPKPIATCRYWHRSLNPRKLEVKFSHL	316
<i>B. taurus II</i>	LRSKRVAPVLIREITRRVNLEGI FQAVYTAGV VLPKPIATCRYWHRSLNPRKLEVKFSHL	316
<i>M. musculus II</i>	LRSKRVAPVLIREITRRVNLEGI FQAVYTAGV VLPKPIATCRYWHRSLNPRKLEVKFSHL	347
<i>D. melanogaster</i>	LRSKRVAPVLIREITRRVNL TGI FQAA YTAGV VLPVATCRYWHRSLNPKKLV D VRF SHL	290
<i>C. elegans</i>	LRSRVAPVLIREITRRVNV TGI FQAA FTAGV IVPKPVSVCRYYHRSLNPRKLDVRF SHL	267
<i>P. falciparum</i>	LRSKRLAPVLIKEITRRINLES I WQAIY TAGV VLPKPISTARYFHRSINVKKLEI DIGFSCL	229
<i>A. thaliana I</i>	LRSKRLAPVMIKEVTRRVHLENIWQAA YTAGV ILPPTITTCQYWHRSLNPKKLDVGF SRL	250
<i>C. albicans</i>	PPNQTKSSMVASYTL PNNPKL KGLRPM TGKDVSTVLSLLYKYQERFDIVQLFTEE—EFKHW	303
<i>S. cerevisiae</i>	PDGHTTEEDMIAENALPAKTKTAGLRK LK KEDIDQVFELFKRYQSRFELIQIFTK E—EFEHN	297
<i>C. glabrata</i>	PPNATKAEMVAKYTL PKATKTAGLREL RLEDVDQALALFNRYQSRFDIVQEFTKE—EFIHW	293
<i>H. capsulatum</i>	PIGSTKSRQVTRNHLPGHTS TPGLRPMQSKD IAVQD LLNRYLKRFDLSQIFSRK—EVDHL	378
<i>A. fumigatus</i>	PAGSTKARQITKNHLPSTTS TPGLRPMPEK DITVHDL LQRYLSRFALNQAFTR E—EVDHW	341
<i>S. pombe</i>	PLGSTKKEKETA KYHLPPNTQTPGLRPMELKDVP AVQSLLSQYMERFELAHLFSEE—EVRHW	311
<i>C. neoformans</i>	PRGSTIARLVQQYSVP SHRIPGFREMKKEDVPQV GALLRRYLD RFDVAQAFKDDDEVEHW	309
<i>Human I</i>	SRNMTMQRTMKLYR LPETPKTAGLRPMETKDIPV VHQLLTRYL KQFHLTPVMS—QEEVEHW	374
<i>B. taurus I</i>	SRNMTMQRTMKLYR LPETPKTAGLRPM EKKDIPV VHQLLSRYL KQFHLTPVMS—QEEVEHW	374
<i>M. musculus I</i>	SRNMTMQRTMKLYR LPETPKTAGLRPM EKKDIPV VHQLLSRYL KQFHLTPVMN—QEEVEHW	374
<i>Human II</i>	SRNMTLQRTMKLYR LPDVTKTSGLRPM EPKDIKSVRELINTY LKQFHLAPVMD—EEEVAHW	376
<i>B. taurus II</i>	SRNMTLQRTMKLYR LPDVTKTSGLRPM EPKDIKSVRELINTY LKQFHLAPVMD—EEEVAHW	376
<i>M. musculus II</i>	SRNMTLQRTMKLYR LPDVTKTSGLRPM EPKDIKAVRELINTY LKQFHLAPVMD—DAEVAHW	407
<i>D. melanogaster</i>	ARNMTMQRTMKLYKLPDQPKTKGYRRI TAKMDKAHKLLEDY LKRFQLSPVFS—KEEFRHW	350
<i>C. elegans</i>	SAKMTMARTIKLYKLP EETARNLREMKSTDPVQVFKLLT TSLKQYSLAPVYNSEEELAHA	328
<i>P. falciparum</i>	NTRLTMSRAIKLYRIDDTLNIKNLR LMKKKDIDGLQKLLNEHLKQYNLHATFS—KEDVAHW	289
<i>A. thaliana I</i>	GARMTMSRTIKLYKLPDAPITPGFRKMEPRD VPAVTRLLRNYLSQFGVATDFD—ENDVEHW	310
<i>C. albicans</i>	MLGHDENS-----DSNVVKS YVVEDENGIITDYFSYLLPFTVLDNAQHDELGIAYLFYY	358
<i>S. cerevisiae</i>	FIDGEESLPL-----DKQVIFS YVVEQPDGKITDFFSFYSLPFTILNNTKYKDLGIGLYYY	353
<i>C. glabrata</i>	FINDKNVVE-----QDKRVVFSYVVESE GKVTDFFSFYSLPFTILNNSRYKDLGIGLYYY	349
<i>H. capsulatum</i>	LLHKEKPGA-----EQIVWSYVAEEP GTHRITDFAAFYSLESSVLQNSKHKNVKAAYLYYY	434
<i>A. fumigatus</i>	L VH K P E T V K-----EQVVWAYVVEDPETHKITDFFSFYNLESTVIQNPKH DNVR AAYLYYY	397
<i>S. pombe</i>	FLYTDKVSS-----GPVVWSYVVENPESKKITDFFSFYSLPSTVIGNPKYKDIQAAYLYYY	367
<i>C. neoformans</i>	FLSGQGKEVGGRRVEQVVWAYVVEDPTTHRITDLISFYALPSTIMKHPKHNLNAA YMFYY	370
<i>Human I</i>	FYP-----QENIIDTFVVENANGEVTD FLSFYTL PSTIMNHPTHKSLKAAYSFYN	424
<i>B. taurus I</i>	FYP-----QENIIDTFVVENANGEVTD FLSFYTL PSTIMNHPTHKSLKAAYSFYN	424
<i>M. musculus I</i>	FYP-----QENIIDTFVVENANGEVTD FLSFYTL PSTIMNHPTHKSLKAAYSFYN	424
<i>Human II</i>	FLP-----REHIIDTFVVESPNGLTDFLSFYTL PSTVMHHPAHKSLKAAYSFYN	426
<i>B. taurus II</i>	FLP-----REHIIDTFVVESPNGLTDFLSFYTL PSTVMHHPAHKSLKAAYSFYN	426
<i>M. musculus II</i>	FLP-----REHIIDTFVVENPSGKLTDFLSFYTL PSTVMHHPAHKSLKAAYSFYN	457
<i>D. melanogaster</i>	FTP-----KEGIIDCFVVADEKGNITDL TSYCLPSSVMHHPVHKTVRAAYSFYN	400
<i>C. elegans</i>	LVP-----KKGVVYSYVAENQNGKITDF VSFYSLPSTVMGHTTHKTIYAA YLYYY	378
<i>P. falciparum</i>	FTP-----IDQVIITYVNE—ENGEIKDL ISFYSLPSKVLGNKNYINLNAAF SFYN	338
<i>A. thaliana I</i>	LLPR-----EDVVD SYLVESPETHDVTDFCSFYTL PSTILGNPNYTTLKAAYS YYN	361

<i>C. albicans</i>	ASDS-----FEKPN-----YKKRLNELITDALITSKKFG	387
<i>S. cerevisiae</i>	ATDADFQFKDRFDPKATKA-----LKTRLCELIYDACILAKNAN	392
<i>C. glabrata</i>	ASDADFQFKFDRFDKEGTSL-----LKQRLSTLVQDACIIAAQNK	388
<i>H. capsulatum</i>	ATETAFAEKEKG-----LKERLLMLINDVILAKKER	466
<i>A. fumigatus</i>	ATETAFTNNMKA-----LKERLLMLMNDALILAKKAH	429
<i>S. prombe</i>	ASDSCPDLSSSQLA-----FVERCKLIVNDALILAKKPH	403
<i>C. neoformans</i>	ATDVIFPPSSSAHSADVDNAGESVVAAVGTGGEDAKTKKKLETRLNALTADLLIIAKQAG	431
Human I	VHTQTP-----LLDLMSDALVLAKMKG	446
<i>B. taurus I</i>	VHTQTP-----LLDLMSDALVLAKMKG	446
<i>M. musculus I</i>	VHTQTP-----LLDLMSDALVLAKMKG	446
Human II	IHTETP-----LLDLMSDALILAKSKG	448
<i>B. taurus II</i>	IHTETP-----LLDLMSDALILAKSKG	448
<i>M. musculus II</i>	IHTETP-----LLDLMSDALILAKSKG	448
<i>D. melanogaster</i>	VSTKTP-----WLDLMSDALISARNVQ	422
<i>C. elegans</i>	VAGSVT-----PKQLINDSLILANREK	400
<i>P. falciparum</i>	ITTTTT-----FKNLIQDAICLAKRNN	360
<i>A. thaliana I</i>	VATQTS-----FLQLMNDALIVSKQKG	383
<i>C. albicans</i>	VDVFNCLTCQDNTYFLKDCFKFGSGDGF-LNYLNFNYRTFPMDDGIDKKTKEVVEDQTSIGIV	448
<i>S. cerevisiae</i>	MDVFNALTSQDNTLFLDDLKFGPGDGF-LNFYLFNYRAKPIITGGLNPDNSNDIKRRSN-VGV	452
<i>C. glabrata</i>	MDVFNALSSQDNTLFLLEDLKFPGDGF-LNFYLFNYRTFPIITGGLTEDQHFDTEHRSN-VGV	448
<i>H. capsulatum</i>	FDVMNALTLDHNPFLFLEQLKFGAGDGG-LHYLNFNYRTAPIAGGVNDKNLPDERKRGG-VGV	526
<i>A. fumigatus</i>	FDVFNALTLDHNPFLFLEQLKFGAGDGG-LHFYLYNYRTAPVPGGVNEKNLPDEKRMGG-VGI	489
<i>S. prombe</i>	FDVFNAVTVLDNNFLKDLKFGEGDGF-LNYIYNYNCPKIPGGIDASKSVDSRPSG-MGF	463
<i>C. neoformans</i>	FDVFNALTLLDNNMFLQEQLKFGPGDGY-LNYLYNWNCAPIDGGQ---HSTTAKQGSKIGV	488
Human I	FDVFNALDLMENKTFLEKLFKFGIGDGN-LQYYLYNWKCPSMGAEK-----VGL	493
<i>B. taurus I</i>	FDVFNALDLMENKTFLEKLFKFGIGDGNLQYYLYNWKCPSMGAEK-----VGL	494
<i>M. musculus I</i>	FDVFNALDLMENKTFLEKLFKFGIGDGN-LQYYLYNWKCPSMGAEK-----VGL	493
Human II	FDVFNALDLMENKTFLEKLFKFGIGDGN-LQYYLYNWRCPGTDSEK-----VGL	495
<i>B. taurus II</i>	FDVFNALDLMENKTFLEKLFKFGIGDGN-LQYYLYNWRCPGTDSEK-----VGL	495
<i>M. musculus II</i>	FDVFNALDLMENKTFLEKLFKFGIGDGN-LQYYLYNWRCPGTDSEK-----VGL	526
<i>D. melanogaster</i>	MDVYNALDLMENKTFLEKLFKFGAGDGN-LQYYLYNWRCPSMQPEE-----IAL	469
<i>C. elegans</i>	FDVFNALDLMHNEKTFSDLFKFGKGDGN-LQYYLYNWKCADMKPSQ-----IGL	447
<i>P. falciparum</i>	FDVFNALVMDNYSVFDLKFGEKGDGS-LQYYLYNWKCASCPSK-----IGI	407
<i>A. thaliana I</i>	FDVFNALDVMHNEKTFLEKLFKFGPGDGG-LHYLYNYRRLKSALKPAE-----LGL	431
<i>C. albicans</i>	VLL	451
<i>S. cerevisiae</i>	VML	455
<i>C. glabrata</i>	VML	451
<i>H. capsulatum</i>	ILV	529
<i>A. fumigatus</i>	VML	492
<i>S. prombe</i>	VMI	466
<i>C. neoformans</i>	VML	491
Human I	VLQ	496
<i>B. taurus I</i>	VLQ	497
<i>M. musculus I</i>	VLQ	496
Human II	VLQ	498
<i>B. taurus II</i>	VLQ	498
<i>M. musculus II</i>	VLQ	529
<i>D. melanogaster</i>	ILM	472
<i>C. elegans</i>	VLQ	450
<i>P. falciparum</i>	VLL	410
<i>A. thaliana I</i>	VLL	434

Figure 1 The multiple sequence alignment of NMT family. The residues in the active site are grey shaded and highly conserved residues across the NMT family are red colored

的晶体结构取自 PDB 数据库 (PDB 登录号: 1YL)。

计算方法 采用 InsightII/Homologv 模块中的 Align123 程序进行多元序列联配, 参数设计为: Blnsum 打分矩阵, Gap_open 为 10, Gap_extension 为 1。然后基于 NMT 的晶体结构特征^[16], 对匹配结果进行微调, 尽量避免在 CaNMT 晶体结构中的二级结构保守区引入 gap。基于多元序列联配结果, 构建

蛋白序列系统进化树。基于不同的 PIC 阈值, 识别得到 NMT 家族的保守残基和家族特异性残基, PIC 水平分别为 30% ~ 90%, 彼此间隔 10%。基于识别得到的踪迹残基的侧链原子的三维空间距离, 采用等级聚类分析创建踪迹残基三维结构簇。由于蛋白内部的残基不容易在进化中发生突变, 因此为了避免假阳性, 计算并选取溶剂可及表面中亲水表

面超过 15% 的所用踪迹残基, 并映射到 CaNMT 的三维结构上, 分析活性位点中潜在的药物结合位点。

结果与讨论

1 多元序列匹配与进化树构建

NMT 家族的多元序列匹配结果见图 1。在此基础上, 基于序列等同率构建 NMT 家族的蛋白序列进化树 (图 2)。由图 2 可见, NMT 家族主要分成 3 个较大的亚家族。真菌 NMT 之间的同源性一般在 47% 以上, 构成了一个较大的亚家族。然而, 新型隐球菌 (*Candida neoformans*) NMT 与拟南芥 (*Arabidopsis thaliana*) NMT 之间具有 50% 左右的同源性, 两者构成了一个较小的亚家族, 推测真菌和植物 NMT 之间的进化距离较为接近。NMT 分为 I 型和 II 型, 哺乳动物 I 型和 II 型 NMT 之间同源性在 70% 左右, 而同型 NMT 之间的同源性非常高, 均在 90% 以上。原生生物 NMT 之间的同源性一般在 50% 左右, 它们与哺乳动物 NMT 一同构成了一个较大的亚家族, 推测两者具有较为接近的进化距离。

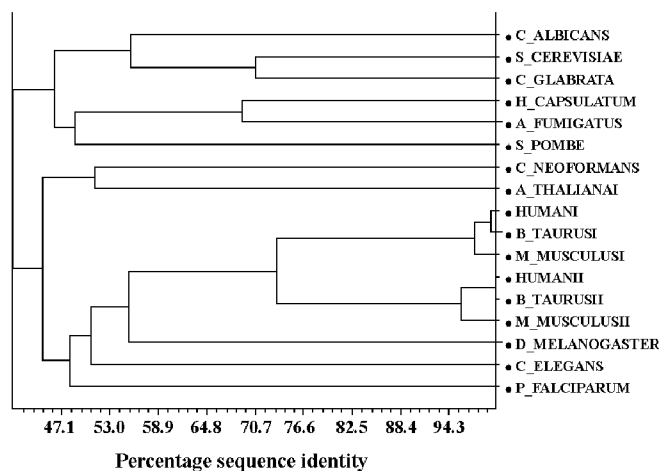


Figure 2 The dendrogram of 17 sequences of NMT family

2 NMT 活性位点踪迹残基的识别

NMT 的晶体结构显示, 其三维结构具有内部对称性, 鞍状的 β -片层结构跨越蛋白质的核心区域, 其两侧被一些 α -螺旋所包围。在蛋白质表面具有一个狭长的“凹槽”, 横跨两个相似的对称单位, 凹槽的中心为一个较深的口袋, 是小分子抑制剂的结合区域。凹槽中位于蛋白质的 N 端区域为肉豆蔻酰 CoA 结合区域, C 端区域和中心口袋一起组成肽类抑制剂的结合区域^[16]。识别 NMT 活性位点的踪迹残基会为研究 NMT 结构与功能关系和基于结构的合理药物设计提供重要信息。由于不同种属

NMT 之间的同源性相对较高, 在 PIC 值为 30% 上识别得到的踪迹残基均为 NMT 家族的高度保守残基, 这些残基对酶的结构和催化功能十分重要。随着 PIC 值的升高, 整个家族经过聚类分析后被划分为多个亚家族, 亚家族内踪迹残基的特异性增强。定义晶体结构 1NYL 中苯并咪唑类抑制剂 benzofuran-5 分子周围 0.8 nm 范围内的所有残基活性为药物结合位点, 在不同 PIC 值水平上识别得到的活性位点的踪迹残基见表 1。

目前已报道有来源于白色念珠菌和酿酒酵母 NMT 的晶体结构, 结合定点突变实验研究结果已经确证了部分活性位点的重要残基。将这些残基的突变信息映射到 CaNMT 的三维结构上, 可以发现它们主要分布在 3 个区域: 肉豆蔻酰 CoA 结合区域, 催化反应中心区域和与抑制剂结合区域, 下面对其进行逐一讨论。

2.1 NMT 肉豆蔻酰 CoA 结合位点的踪迹残基识别 NMT 必需首先与肉豆蔻酰 CoA 结合后, 才能进一步识别特异性的肽类底物。定点突变实验研究发现^[17-19], Leu103, Ala208, Asn421 和 Gly447 突变后会引起 NMT 对肉豆蔻酰 CoA 亲和力的降低, 其中 Ala208 同时还会影响 NMT 与肽类底物的结合。进化踪迹分析结果显示, 在 30% 的水平上成功识别得到保守性踪迹残基 Leu103, Ala208 和 Asn421。Gly447 在除果蝇 (*Drosophila melanogaster*) NMT 外的其余 16 种 NMT 中也是保守的, 它在 PIC 值为 50% 的水平上被识别得到。

2.2 NMT 活性位点参与催化反应的踪迹残基识别 蛋白质的 N 肉豆蔻酰化反应是一个亲核加成-消除的反应过程。ScNMT 的晶体结构显示^[10], NMT 活性位点存在三大结构要素促使催化反应的顺利进行: 首先, Phe170 和 Leu171 的主链酰胺基团组成一个“氧阴离子穴” (oxyanion hole), 将肉豆蔻酰 CoA 的硫酯羰基极化; 其次, ScNMT 的 C 端 Leu455 的羧基负离子将肽类或蛋白质底物上 N 端 Gly 的氨基阳离子去质子化形成中性的 $-NH_2$, 然后亲核进攻肉豆蔻酰 CoA 的硫酯羰基, 形成四面体过渡态; 再次, ScNMT 活性位点的 Asn169 和 Thr205 与肽类和/或蛋白质底物上 N 端的 Gly 形成氢键相互作用, 使得底物的 Gly 处于活性位点正确的位置, 有利于催化反应的顺利进行。对 ScNMT 的定点突变研究进一步证实, 将 Phe170, Leu171, Leu455, Asn169 和 Thr205 突变后会导致催化反应速率的大幅降低^[20], 它们分别对应于 CaNMT 中的 Phe176, Leu177, Leu451,

Table 1 The identified conserved residues and class-specific residues in drug binding sites of CaNMT by ET analysis

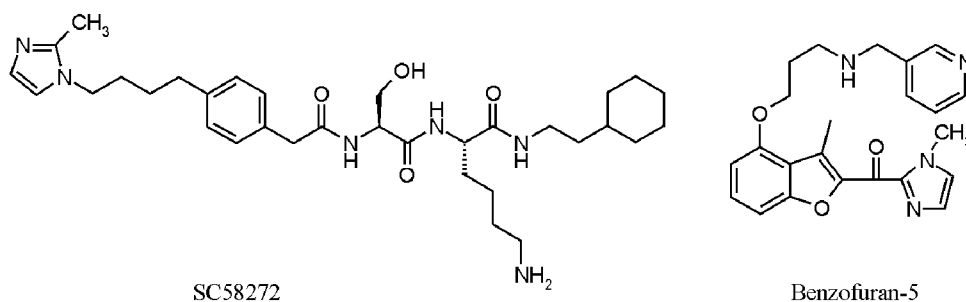
Sequence identity	Conserved residue	Class-specific residue
30%	<u>Leu103^a</u> , Tyr107, Val108, Glu109, Asp110, Arg116, <u>Phe117^c</u> , <u>Tyr119^c</u> , Trp126, Leu128, <u>Asn175^b</u> , <u>Phe176^b</u> , <u>Leu177^b</u> , Cys178, <u>Ala208^{a,c}</u> , <u>Thr211^b</u> , <u>Tyr225^c</u> , <u>His227^c</u> , Arg228, <u>Phe240^c</u> , Tyr335, Leu337, Asp389, Val390, Asn392, Lys407, Phe408, Gly409, Gly411, Asp412, Gly413, Leu415, <u>Asn421^a</u>	None
40%	<u>Leu103^a</u> , Tyr107, Val108, Glu109, Asp110, Arg116, <u>Phe117^c</u> , <u>Tyr119^c</u> , Trp126, Leu128, <u>Asn175^b</u> , <u>Phe176^b</u> , <u>Leu177^b</u> , Cys178, <u>Ala208^{a,c}</u> , <u>Thr211^b</u> , <u>Tyr225^c</u> , <u>His227^c</u> , Arg228, <u>Phe240^c</u> , Tyr335, Leu337, Asp389, Val390, Asn392, Lys407, Phe408, Gly409, Gly411, <u>Asp412^c</u> , Gly413, Leu415, <u>Asn421^a</u>	None
50%	<u>Leu103^a</u> , Tyr107, Val108, Glu109, Asp110, <u>Phe115^c</u> , Arg116, <u>Phe117^c</u> , <u>Tyr119^c</u> , Trp126, Leu128, Ile174, <u>Asn175^b</u> , <u>Phe176^b</u> , <u>Leu177^b</u> , Cys178, <u>Ala208^{a,c}</u> , <u>Thr211^b</u> , Cys223, <u>Tyr225^c</u> , <u>His227^c</u> , Arg228, <u>Phe240^c</u> , Tyr335, Leu337, <u>Tyr354^c</u> , Asp389, Val390, Asn392, Leu394, Lys407, Phe408, Gly409, Gly411, <u>Asp412^c</u> , Gly413, Leu415, <u>Asn421^a</u> , <u>Gly447^a</u>	Phe124, Phe152, Pro338, Phe339, Ile352
60%	<u>Leu103^a</u> , Tyr107, Val108, Glu109, Asp110, <u>Phe115^c</u> , Arg116, <u>Phe117^c</u> , <u>Tyr119^c</u> , Trp126, Ala127, Leu128, Ile174, <u>Asn175^b</u> , <u>Phe176^b</u> , <u>Leu177^b</u> , Cys178, <u>Ala208^{a,c}</u> , <u>Thr211^b</u> , Cys223, Arg224, <u>Tyr225^c</u> , <u>His227^c</u> , Arg228, Val238, <u>Phe240^c</u> , Tyr335, Leu337, <u>Tyr354^c</u> , Asp389, Val390, Asn392, Leu394, Lys407, Phe408, Gly409, Gly411, <u>Asp412^c</u> , Gly413, Leu415, <u>Asn421^a</u> , <u>Gly447^a</u>	Ala113, Ser120, Phe124, Phe152, Pro338, Phe339, Leu350, Ile352, Ala353, Leu355, Cys393
70%	<u>Leu103^a</u> , Tyr107, Val108, Glu109, Asp110, <u>Asp112^c</u> , <u>Phe115^c</u> , Arg116, <u>Phe117^c</u> , <u>Tyr119^c</u> , Trp126, Ala127, Leu128, Ile174, <u>Asn175^b</u> , <u>Phe176^b</u> , <u>Leu177^b</u> , Cys178, <u>Ala208^{a,c}</u> , <u>Thr211^b</u> , Cys223, Arg224, <u>Tyr225^c</u> , <u>His227^c</u> , Arg228, Val238, <u>Phe240^c</u> , Tyr335, Leu337, <u>Tyr354^c</u> , Asp389, Val390, Asn392, Leu394, Lys407, Phe408, Gly409, Gly411, <u>Asp412^c</u> , Gly413, Leu415, <u>Asn421^a</u> , <u>Gly447^a</u>	Ala113, Ser120, Phe124, Phe152, Pro338, Phe339, Leu350, Ile352, Ala353, Leu355, Cys393, <u>Leu451^{b,c}</u>
80% ^d	<u>Leu103^a</u> , Tyr107, Val108, Glu109, Asp110, <u>Asp112^c</u> , <u>Phe115^c</u> , Arg116, <u>Phe117^c</u> , <u>Tyr119^c</u> , Trp126, Ala127, Leu128, Ile174, <u>Asn175^b</u> , <u>Phe176^b</u> , <u>Leu177^b</u> , Cys178, <u>Ala208^{a,c}</u> , <u>Thr211^b</u> , Cys223, Arg224, <u>Tyr225^c</u> , <u>His227^c</u> , Arg228, Val238, <u>Phe240^c</u> , <u>Tyr256^c</u> , Tyr335, Leu337, <u>Tyr354^c</u> , Asp389, Val390, Asn392, Leu394, Lys407, Phe408, Gly409, Gly411, <u>Asp412^c</u> , Gly413, Leu415, <u>Asn421^a</u> , <u>Gly447^a</u>	Ala113, Ser120, Phe124, Phe152, Pro338, Phe339, Leu350, Ile352, Ala353, Leu355, Cys393, <u>Leu451^{b,c}</u>

^a Identified important functional residues in myristoyl CoA binding site; ^b Identified important functional residues for the catalytic activity of NMT; ^c Identified important functional residues in the inhibitor binding site; ^d Identified trace residues at PIC level of 80% and 90% are identical

Asn175和 Thr211。进化踪迹分析结果显示,在 30%的水平上成功识别得到保守性踪迹残基 Phe176, Leu177, Asn175和 Thr211。CaNMT的 C端 Leu451在 NMT家族中并不高度保守,但该残基起作用的是其主链羧基负离子^[11],与氨基酸的类型关系并不大,因此 Leu451没有被识别为保守性踪迹残基,而是作为亚家族特异性残基在 PIC为 70%的水平上被识别得到。

2.3 NMT活性位点与抑制剂结合的踪迹残基识别 CaNMT与肽类抑制剂 SC58272和苯并呋喃类小分子抑制剂 benzofuran-5(图 3)复合物的晶体结

构揭示了两大类抑制剂与 NMT活性位点的相互作用模式,为设计选择性作用于真菌 NMT的抑制剂提供了重要信息。进一步的定点突变实验证实^[11], Phe115, Phe117, Tyr225和 Phe240, Tyr256和 Leu451影响与肽类抑制剂 SC58272的结合, Asp112, His227和 Asp412影响 NMT对 SC58272的识别; Phe117, Tyr119, Tyr225, His227, Tyr256, Tyr354和 Leu451影响 NMT与小分子抑制剂 benzofuran-5的结合。进化踪迹分析结果显示,在 PIC值为 30%的水平上成功识别得到保守性踪迹残基 Phe117, Tyr119, Tyr225, His227, Phe240和 Asp412。Phe115和 Tyr354在绝

**Figure 3** Structures of SC58272 and benzofuran-5

大多数 NMT 中也高度保守,它们在 PIC 值为 50% 的水平上被识别得到。Asp112 和 Tyr256 在哺乳动物和原生生物 NMT 中高度保守,但在真菌 NMT 中保守性不强,它们分别在 PIC 值为 70% 和 80% 的水平上被识别得到。

3 CaNMT 活性位点的药物结合位点分析

在进化踪迹分析识别得到 NMT 家族蛋白进化踪迹残基的基础上,计算各残基的溶剂可及表面,选取亲水性表面超过 15% 的蛋白质残基作为潜在的药物结合位点进行重点研究。图 4 显示了苯并咪唑类抑制剂 benzofuran-5 分子周围 0.8 nm 范围内的所有保守性残基(紫红色)和亚家族特异性残基(蓝色)。

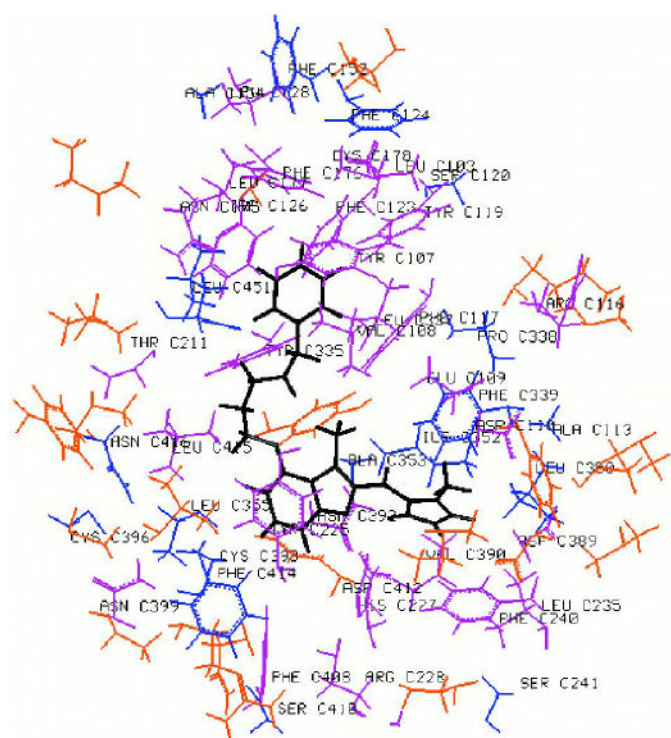


Figure 4 Drug binding sites of CaNMT. Black represents compound benzofuran-5, purple represents conserved residues, blue represents class-specific residues

在 benzofuran-5 分子咪唑环 4 位侧链的吡啶环周围主要是 NMT 家族高度保守的残基,以疏水性残基为主。吡啶 N 原子和 Tyr119 形成氢键,吡啶环与附近的保守残基 Leu103, Tyr107, Val108, Phe123, Phe176, Leu177, Leu337 形成疏水相互作用。在吡啶环的 4 位和 5 位附近有 NMT 家族高度保守残基 Asp175 和 Trp126, 它们并不与 benzofuran-5 形成直接的相互作用,因此这两个残基有望成为发现新型 NMT 抑制剂的重要结合位点。Benzofuran-5 分子 4

位侧链的烷基部分附近有保守残基 Tyr335, Leu415 和 Thr211, 其中 Thr211 对 NMT 发挥催化功能具有重要作用,因此也是新型 NMT 抑制剂设计的重要位点。Benzofuran-5 分子咪唑环位于 CaNMT 活性位点的核心区域,它周围有保守残基 Tyr225 和 His227, 其中咪唑环的氧原子与 His227 形成氢键。CaNMT 活性位点的亚家族特异性残基主要分布在 benzofuran-5 分子咪唑环 2 位侧链周围,亚家族特异性残基 Phe339 和保守残基 Phe240, Phe115 组成了强疏水环境与 2 位侧链形成疏水相互作用,保守残基 Asp392 与 2 位侧链的咪唑 N 原子形成氢键。上述这些残基是该类化合物产生高 NMT 抑制活性的重要分子基础。Benzofuran-5 分子 2 位侧链周围的亚家族特异性残基 Pro338, Leu350, Ile352 和 Ala353 不与已有的抑制剂产生相互作用,在抑制剂设计中考虑上述残基将有助于提高对真菌 NMT 的选择性。

结论

本文基于 NMT 家族的多元序列联配,构建了该家族的进化树,并在此基础上通过踪迹分析技术识别得到了 NMT 家族的保守性和亚家族特异性残基。经晶体结构和定点突变实验所证实的 NMT 家族重要功能残基均得到了准确识别,证明了 ET 方法的可靠性。对 CaNMT 活性位点中药物结合位点的分析发现,苯并咪唑类抑制剂 benzofuran-5 分子 4 位侧链被 NMT 家族的保守残基所包围,而亚家族特异性残基主要分布在 2 位侧链附近,说明 4 位侧链所处区域是化合物产生 NMT 抑制活性的首要因素,而 2 位侧链所处区域对提高化合物抑酶活性和选择性起着关键作用。Trp126, Asp175 和 Thr211 是 NMT 家族的高度保守残基,而且不与已有的抑制剂发生直接的相互作用,因此将是发现新型 NMT 抑制剂的重要药物结合位点。亚家族特异性残基 Pro338, Leu350, Ile352 和 Ala353 可作为苯并咪唑类抑制剂结构优化的重要靶点,据此设计的新化合物将有望进一步提高对真菌 NMT 的亲力和选择性。上述研究结果为研究 NMT 结构-功能关系和新型 NMT 抑制剂类抗真菌药物的设计提供重要信息。

References

[1] Boutin JA. Myristoylation [J]. Cell Signal 1997, 9: 15-35
 [2] Farazi TA, Waksman G, Gordon JI. The biology and enzymology of protein N-myristoylation [J]. J Biol

- Chem, 2001, 276: 39501–39504
- [3] Resh MD. Regulation of cellular signalling by fatty acid acylation and prenylation of signal transduction proteins [J]. Cell Signal, 1996, 8: 403–412
- [4] Zha J, Weiler S, Oh KJ, et al. Posttranslational *N*-myristoylation of BID as a molecular switch for targeting mitochondria and apoptosis [J]. Science, 2000, 290: 1761–1765
- [5] Rudnick DA, Møller CA, Rocque WJ, et al. Kinetic and structural evidence for a sequential ordered Bi Bi mechanism of catalysis by *Saccharomyces cerevisiae* myristoyl-CoA: protein *N*-myristoyltransferase [J]. J Biol Chem, 1991, 266: 9732–9739
- [6] Bhatnagar RS, Futterer K, Farazi TA, et al. Structure of *N*-myristoyltransferase with bound myristoyl-CoA and peptide substrate analogs [J]. Nat Struct Biol, 1998, 5: 1091–1097
- [7] Georgopapadakou NH. Antifungals targeted to protein modification focus on protein *N*-myristoyltransferase [J]. Expert Opin Investig Drugs, 2002, 11: 1117–1125
- [8] Masubuchi M, Ebike H, Kawasaki K, et al. Synthesis and biological activities of benzofuran antifungal agents targeting fungal *N*-myristoyltransferase [J]. Bioorg Med Chem, 2003, 11: 4463–4478
- [9] Yanazaki K, Kaneko Y, Suwa K, et al. Synthesis of potent and selective inhibitors of *Candida albicans* *N*-myristoyltransferase based on the benzothiazole structure [J]. Bioorg Med Chem, 2005, 13: 2509–2522
- [10] Farazi TA, Waksman G, Gordon JI. Structures of *Saccharomyces cerevisiae* *N*-myristoyltransferase with bound myristoyl-CoA and peptide provide insights about substrate recognition and catalysis [J]. Biochemistry, 2001, 40: 6335–6343
- [11] Sogabe S, Masubuchi M, Sakata K, et al. Crystal structures of *Candida albicans* *N*-myristoyltransferase with two distinct inhibitors [J]. Chem Biol, 2002, 9: 1119–1128
- [12] Lichtarge O, Boume HR, Cohen FE. An evolutionary trace method defines binding surfaces common to protein families [J]. J Mol Biol, 1996, 257: 342–358
- [13] Madabushi S, Gross AK, Philippi A, et al. Evolutionary trace of G protein-coupled receptors reveals clusters of residues that determine global and class-specific functions [J]. J Biol Chem, 2004, 279: 8126–8132
- [14] Chakravarty S, Hutson AM, Estes MK, et al. Evolutionary trace residues in noroviruses: importance in receptor binding, antigenicity, virion assembly, and strain diversity [J]. J Virol, 2005, 79: 554–568
- [15] Song YL, Qi YP, Zhang WN, et al. Evolutionary trace analysis of eukaryotic DNA topoisomerase I superfamily: identification of novel antitumor drug binding site [J]. Sci Sin C (中国科学 C 辑), 2005, 48: 375–384
- [16] Weston SA, Camble R, Colls J, et al. Crystal structure of the antifungal target *N*-myristoyltransferase [J]. Nat Struct Biol, 1998, 5: 213–221
- [17] Duronio RJ, Reed SJ, Gordon JI. Mutations of human myristoyl-CoA: protein *N*-myristoyltransferase cause temperature-sensitive myristic acid auxotrophy in *Saccharomyces cerevisiae* [J]. Proc Natl Acad Sci, 1992, 89: 4129–4133
- [18] Kishore NS, Wood DC, Mehta PP, et al. Comparison of the acyl chain specificities of human myristoyl-CoA synthetase and human myristoyl-CoA: protein *N*-myristoyltransferase [J]. J Biol Chem, 1993, 268: 4889–4902
- [19] Johnson DR, Duronio RJ, Langner CA, et al. Genetic and biochemical studies of a mutant *Saccharomyces cerevisiae* myristoyl-CoA: protein *N*-myristoyltransferase (m72pLeu99 → Pro) that produces temperature-sensitive myristic acid auxotrophy [J]. J Biol Chem, 1993, 268: 483–494
- [20] Farazi TA, Manchester JK, Waksman G, et al. Pre-steady-state kinetic studies of *Saccharomyces cerevisiae* myristoyl-CoA: protein *N*-myristoyltransferase mutants identify residues involved in catalysis [J]. Biochemistry, 2001, 40: 9177–9186