

吡咯赖氨酸：第 22 种参与蛋白质生物合成的氨基酸*

凌晨^{1,2)} 王恩多^{1)**}

(¹⁾中国科学院上海生命科学研究院, 生物化学与细胞生物学研究所, 分子生物学国家重点实验室, 上海 200031;

²⁾中国科学院研究生院, 上海 200031)

摘要 吡咯赖氨酸在产甲烷菌的甲胺甲基转移酶中发现, 是目前已知的第 22 种参与蛋白质生物合成的氨基酸, 与标准氨基酸不同的是, 它由终止密码子 UAG 的有义编码形成. 与之对应的在产甲烷菌中也含有特异的吡咯赖氨酰-tRNA 合成酶 (PylRS) 和吡咯赖氨酸 tRNA (tRNA^{Pyl}). tRNA^{Pyl} 具有不同于经典 tRNA 的特殊结构. 产甲烷菌通过直接途径和间接途径这两种途径生成吡咯赖氨酰-tRNA^{Pyl}(Pyl-tRNA^{Pyl}), 它还可能通过 mRNA 上的特殊结构以及其他还未发现的机制, 控制 UAG 编码成为终止密码子或者吡咯赖氨酸. 比较了吡咯赖氨酸与另一种非标准氨基酸, 第 21 种氨基酸——硒代半胱氨酸的相似点与不同点.

关键词 吡咯赖氨酸, 吡咯赖氨酸 tRNA, 吡咯赖氨酰-tRNA 合成酶, 吡咯赖氨酸插入元件, 硒代半胱氨酸

学科分类号 Q71

蛋白质的生物合成是生命活动的重要组成部分, 通常蛋白质由 20 种氨基酸组成. 存储于 DNA 上的遗传信息通过转录传递至 mRNA, mRNA 上的核苷酸序列作为蛋白质合成的模板, 决定蛋白质的氨基酸排列顺序. 由于 mRNA 上每三个核苷酸决定一个氨基酸, 所以被称为三联体密码 (triplet code). 自然界传统的三联体密码共有 64 个, 其中 61 个用来编码 20 种标准氨基酸, 另外 3 个: UAA, UAG, UGA 为终止密码子, 指导蛋白质翻译的终止. 三联体密码在自然界中近乎完全通用.

在从 mRNA 上的遗传密码 (genetic code) 翻译成为氨基酸的过程中, tRNA 起着十分重要的作用. 每一种 tRNA 上都有反密码子 (anticodon) 序列, 用来识别对应的三联体密码. 在翻译过程中, tRNA 3' 端的腺苷酸上核糖的羟基与对应氨基酸的羧基共价连接, 通过 tRNA 反密码子与 mRNA 上密码子的正确配对, 将氨基酸依照 mRNA 上密码子的序列传送到肽链上的正确位置. 这个共价连接过程, 称为 tRNA 的氨基酰化反应, 由特异性的氨酰-tRNA 合成酶 (aaRS) 催化. 这种酶具有极高的专一性, 对应特定的氨基酸和 tRNA, 从而保证蛋白质合成的精确性^[1]. 20 种 aaRS 的一级结构和高级结构各不相同, 根据极有限的序列同源性和活性中心的高级结构, 大肠杆菌中 20 种 aaRS 被分为 2 类 (class I 和 class II), 每类各 10 种 aaRS^[1]. 但是这种

分类也不是绝对的, 例如: 大肠杆菌中的赖氨酰-tRNA 合成酶 (LysRS) 带有第二类 aaRS 的结构特点, 属于第二类 aaRS; 但是在一些古细菌中 LysRS 除了以第二类 aaRS 的形式出现, 还有一种 LysRS 具有第一类 aaRS 的结构特点, 属于第一类 aaRS^[1].

1986 年, 第 21 种参与蛋白质生物合成的氨基酸——硒代半胱氨酸 (selenocysteine) 被发现^[2]. 不同于 20 种标准氨基酸, 它由终止密码子 UGA 直接编码. 2002 年, 由另外一个终止密码子 UAG 编码的第 22 种氨基酸吡咯赖氨酸 (pyrrolysine) 也被发现^[3].

1 古细菌中发现 UAG 被编码为吡咯赖氨酸

古细菌——巴氏甲烷八叠球菌 (*Methanosarcina barkeri*, *M. barkeri*) 是一种产甲烷菌, 甲烷生成的底物分别是单甲胺 (monomethylamine, MMA), 二甲胺 (dimethylamine, DMA) 和三甲胺 (trimethylamine, TMA), 它们被对应的 3 种甲胺甲基转移酶——单甲胺甲基转移酶 (MtmB)、二甲胺甲

*国家自然科学基金资助项目(30330180).

** 通讯联系人.

Tel: 021-54921241, Fax: 021-54921011

E-mail: edwang@sibs.ac.cn

收稿日期: 2004-12-21, 接受日期: 2005-01-28

基转移酶(MtbB)、三甲胺甲基转移酶(MttB)分别催化,产生甲烷.最先在编码 MtmB 的基因 *mtmB1* 中发现阅读框内的 UAG 琥珀终止子,它没有终止蛋白质的翻译^[4].后来在编码 MtbB 和 MttB 的基因 *mtbB* 和 *mttB* 中也发现了同样的现象^[5],虽然这 3 个基因在序列上不存在同源性.噬乙酸甲烷八叠球菌(*Methanosarcina acetivorans*)^[6]和马氏甲烷八叠球菌(*Methanosarcina mazei*)^[7]的基因组序列表明目前所有已知的产甲烷菌,它们的甲胺甲基转移酶基因内都存在阅读框内的 UAG.

M.barkeri 的 MtmB 氨基酸序列测定表明,*mtmB1* 中的 UAG 被编码为赖氨酸(lysine)^[8].由于蛋白质测序过程中条件比较剧烈,因此不排除 UAG 被编码为赖氨酸的某种修饰形式的可能性.进一步的晶体结构证明,*mtmB1* 基因的 UAG 被编码为吡咯赖氨酸(pyrrrolysine)而非赖氨酸^[9,10].吡咯赖氨酸与赖氨酸的结构如图 1 所示.

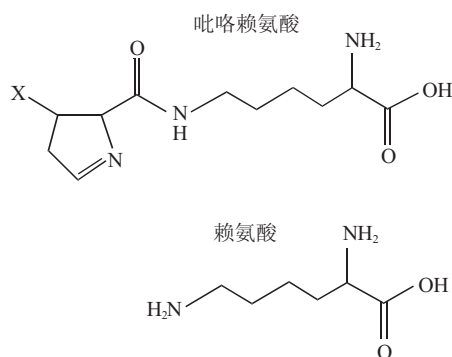


Fig.1 The structure of pyrrolysine and lysine

图 1 吡咯赖氨酸和赖氨酸的结构

X 代表甲基

2 吡咯赖氨酸 tRNA 和吡咯赖氨酰-tRNA 合成酶

M.barkeri 的基因组中 *mtmB1* 基因附近存在 *pylT*, *pylS*, *pylB*, *pylC* 和 *pylD* 基因簇^[11]. *PylT* 基因编码不寻常的吡咯赖氨酸 tRNA ($tRNA^{Pyl}$ 或 $tRNA_{CUA}$),它拥有 CUA 反密码子,用于编码 UAG 成为有义密码子.它的二级结构和经典的 tRNA 二级结构有着很大差异.反密码茎由 6 个而非 5 个碱基对组成,因此它的可变环含有 3 个而非 4 个碱基, D 环比通常的 tRNA 小.虽然在 $tRNA^{Pyl}$ 中可以找到许多 tRNA 的保守碱基,但是广泛保守的 D 环的 G18 G19 序列和 T 环的 T ψ C 序列却不存在^[11]. $tRNA^{Pyl}$ 的特殊性质看起来与它编码 UAG 琥珀终止

子的特殊功能密切相关. $tRNA^{Pyl}$ 的二级结构见图 2.

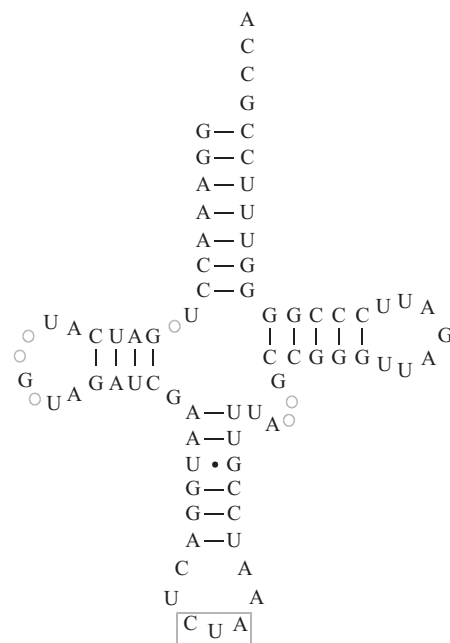


Fig.2 The cloverleaf structure of *M. barkeri* $tRNA^{Pyl}$

图 2 *M. barkeri* $tRNA^{Pyl}$ 的三叶草结构

比典型 tRNA 结构缺少的碱基用空心圈表示.

PylS 基因编码一种与第二类赖氨酰-tRNA 合成酶(Class II LysRS)有一些类似结构特征的 aaRS^[11],但是这种酶不属于 *M.barkeri* 的两类 LysRS——LysRS1、LysRS2 中的任何一种^[12],它可能编码了吡咯赖氨酰-tRNA 合成酶(*PylRS*).

基因 *pylB*, *pylC* 和 *pylD* 据推测应该与吡咯赖氨酰-tRNA^{Pyl} 的形成相关,但是目前尚无相关的实验证明它们的性质与功能.

3 吡咯赖氨酸如何进入蛋白质

吡咯赖氨酸可能以两种途径进入蛋白质:一种是翻译后成熟蛋白质的赖氨酸残基被修饰成为吡咯赖氨酸,另一种是吡咯赖氨酸在翻译过程中直接进入蛋白质. *PylT* 和 *pylS* 的存在以及 UAG 的特殊编码使得人们推测吡咯赖氨酸可能是翻译中进入蛋白质而非翻译后.那么又出现了另一个问题: $Pyl-tRNA^{Pyl}$ 是怎样形成的?这也有两种可能途径:吡咯赖氨酸直接与 $tRNA^{Pyl}$ 结合形成 $Pyl-tRNA^{Pyl}$ 或者先形成赖氨酰-tRNA^{Pyl}(Lys-tRNA^{Pyl}),再通过翻译前修饰形成 $Pyl-tRNA^{Pyl}$.

目前的实验结果已经证实了通过两种途径可以形成 $Pyl-tRNA^{Pyl}$. $tRNA^{Pyl}$ 在 *M.barkeri* 的第一类和第二类 LysRS 的共同作用下,可以与赖氨酸结合

形成 Lys-tRNA^{Pyl}, 然后通过对赖氨酸的修饰形成 Pyl-tRNA^{Pyl} [13]. 第一类或第二类 LysRS 单独存在的情况下, 该反应不能进行, 两类 LysRS 都存在的情况下, 酶无法识别吡咯赖氨酸, 不能直接形成 Pyl-tRNA^{Pyl}. 另一方面, tRNA^{Pyl} 可以被 PylRS 直接催化形成 Pyl-tRNA^{Pyl}, 这一反应在体外实验和体内实验中都已经得到证实 [12,14]; 赖氨酸和 tRNA^{Lys} 都不是 PylRS 的底物. Pyl-tRNA^{Pyl} 的两种形成途径如图 3 所示.

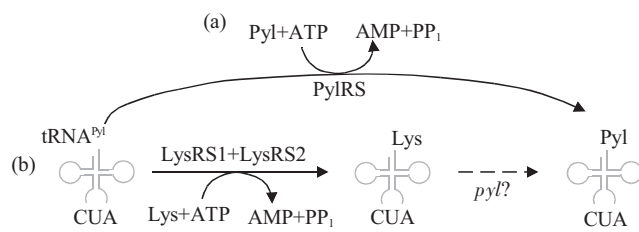


Fig.3 Two ways of tRNA^{Pyl} aminoacylation

图 3 形成 Pyl-tRNA^{Pyl} 的两种途径

(a) tRNA^{Pyl} 被吡咯赖氨酰 tRNA 合成酶催化直接与吡咯赖氨酸结合; (b) tRNA^{Pyl} 在第一类和第二类赖氨酰 tRNA 合成酶的共同作用下先与赖氨酸结合形成 Lys-tRNA^{Pyl}, 再通过某种还未确定的 *pyl?* 基因的修饰作用形成 Pyl-tRNA^{Pyl}.

4 UAG 有义编码和无义编码的控制机制

目前我们知道不仅在产甲烷菌的甲胺甲基转移酶中阅读框内的 UAG 被编码成为吡咯赖氨酸, 这种情况还可能存在于细菌哥本哈根脱亚硫酸菌 (*Desulfotobacterium hafniense*) 中 [11] 和其他我们还未鉴定的基因中. 那么, 是什么机制控制 UAG 在某些情况下编码成为吡咯赖氨酸, 而在另一些情况下编码成为终止信号呢?

近期的实验曾将 PylRS, tRNA^{Pyl} 和 *M. barkeri* 的甲胺甲基转移酶基因共同转化进入大肠杆菌 (*E.coli*) 中, 当吡咯赖氨酸被加入培养基中, 它能够进入甲胺甲基转移酶肽链上的相应位置, 效率高达 75%. 虽然该实验并未进一步验证, 吡咯赖氨酸是否可以进入其他含有阅读框内 UAG 人造基因的相应肽链, 但是转化对宿主菌大肠杆菌的生长没有明显毒性, 说明吡咯赖氨酸并未进入大肠杆菌 UAG 编码的相应部位. 由此看来, 尽管在 PylRS 和 tRNA^{Pyl} 存在的情况下, 吡咯赖氨酸仍然需要其他信息来决定是否进入肽链.

第 21 种氨基酸, 硒代半胱氨酸, 也曾存在类

似的问题. 在 UGA 编码为硒代半胱氨酸的生物中, 绝大部分 UGA 仍然编码为终止信号, 只有小部分成为硒代半胱氨酸. 后来发现硒代半胱氨酸通过 mRNA 上的特殊结构——硒代半胱氨酸插入元件 (selenocysteine inserting element), 以及特殊的延伸因子 (SelB) 和其他辅助因子来解决这个问题. 所以推测吡咯赖氨酸也是同样情况, 并在 *M. barkeri* 甲胺甲基转移酶基因的有义 UAG 密码子附近找到了类似硒代半胱氨酸插入元件的序列 [15]. 图 4 显示了与两种特殊氨基酸插入蛋白质相关的 mRNA 结构.

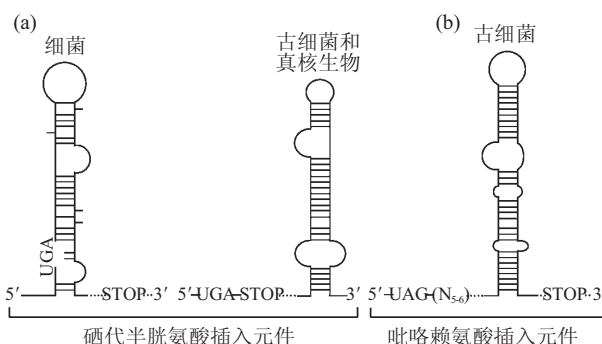


Fig.4 mRNA structure concerning the insertion of selenocysteine and pyrrolysine

图 4 与硒代半胱氨酸和吡咯赖氨酸插入蛋白质相关的 mRNA 结构

(a) 硒代半胱氨酸插入元件. 在细菌中, SelB 结合硒代半胱氨酰-tRNA^{Sec} 和 SECIS, 使得 UGA 不编码终止信号而是硒代半胱氨酸 (真核生物和古细菌中, SelB 的功能由两种蛋白质共同完成). (b) *M. barkeri* 中推测的吡咯赖氨酸插入元件. 它的作用推测与 SECIS 类似.

5 吡咯赖氨酸和硒代半胱氨酸的比较

作为目前仅知的两种非标准氨基酸, 吡咯赖氨酸和硒代半胱氨酸有许多相似之处, 例如它们都是由终止密码子的有义编码生成, 它们的 tRNA 都具有不寻常的二级结构, 但是它们也有一些不同之处. 由于目前吡咯赖氨酸的研究都是来源于 *M. barkeri*, 而硒代半胱氨酸的研究在 *E.coli* 中最为清楚, 因此我们以这两种细菌为代表, 比较这两种特殊的氨基酸.

a. aaRS 和 tRNA: 硒代半胱氨酸不存在对应的 aaRS, 丝氨酰-tRNA 合成酶 (SerRS) 在硒代半胱氨酰-tRNA^{Sec} 的形成过程中起重要作用, 吡咯赖氨酸有特定的 aaRS. tRNA^{Sec} 和 tRNA^{Pyl} 的结构都与经典 tRNA 不同, 但是它们彼此的结构也不相同.

tRNA^{Sec} 的最显著结构是接受茎有 8 个碱基对, tRNA^{Pyl} 的结构见图 2.

b. 氨基酰-tRNA 的合成途径: 硒代半胱氨酸-tRNA^{Sec} 的形成只有一种间接的途径, 如图 5 所示^[6]. Pyl-tRNA^{Pyl} 的生成可以通过 PylRS 直接作用, 也可以通过第一类和第二类 LysRS 的共同间接作用 (图 3).

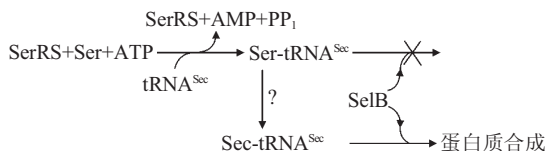


Fig.5 The indirect formation of selenocysteinyl-tRNA

图 5 硒代半胱氨酸-tRNA^{Sec} 的间接合成

c. 分布: 已经发现硒代半胱氨酸广泛存在于自然界的古细菌, 细菌和真核生物中; 吡咯赖氨酸目前仅发现在产甲烷菌和细菌 *Desulfotobacterium hafniense* 中存在. 当然吡咯赖氨酸的发现时间较晚, 随着研究的不断深入, 或许我们可以在越来越多的生物中发现它的踪迹.

d. 氨基酰-tRNA 与延伸因子的相互作用: 细菌中的硒代半胱氨酸有特异的延伸因子 SelB, 丝氨酸-tRNA^{Sec} 和硒代半胱氨酸-tRNA^{Sec} 与细菌常见的延伸因子 Tu 作用极弱, 而且 SelB 仅与硒代半胱氨酸-tRNA^{Sec} 而非丝氨酸-tRNA^{Sec} 结合, 保证了硒代半胱氨酸精确地进入肽链. Pyl-tRNA^{Pyl} 与延伸因子的作用机制还未明了, 但是已有实验证明 tRNA^{Pyl} 与大肠杆菌的延伸因子 Tu 可以结合^[7].

6 小 结

吡咯赖氨酸是近年来发现的非标准氨基酸, 它和硒代半胱氨酸一样, 都是由终止密码子的有义编码形成. 它们有一些相同的性质, 但是目前的研究表明, 它们的不同之处更多. 我们有可能在更多的物种中发现吡咯赖氨酸, 可以对吡咯赖氨酸编码的控制机制以及吡咯赖氨酸插入元件等问题进行更深入和更全面的研究, 吡咯赖氨酸也可能在肽链的起始, 延伸和终止方面都有不同于普通氨基酸的特殊性质. 它和硒代半胱氨酸为我们提供了研究蛋白质生物合成的特例, 使我们对遗传密码有更多的认识.

为什么自然界会有吡咯赖氨酸这种特殊编码氨

基酸的存在呢? 一种可能的解释是它是远古的基因编码形式, 在某种自然选择压力下逐渐减少, 因此现在极少发现; 另一种解释则完全相反, 可以认为它是遗传密码后期的一种进化, 为了适应某种改变或获得某种未知的新功能而进行的改进, 在自然界中不是逐渐降低, 而是逐渐增加. 这两种猜测都需要进一步的实验数据来验证.

参 考 文 献

- 1 Ibba M, Söll D. Aminoacyl-tRNA synthesis. *Annu Rev Biochem*, 2000, **69**: 617~650
- 2 Bock A, Forchhammer K, Heider J, *et al.* Selenocysteine: the 21st amino acid. *Mol Microbiol*, 1991, **5** (3): 515~520
- 3 Atkins J F, Gesteland R. The 22nd amino acid. *Science*, 2002, **296** (5572): 1409~1410
- 4 Burke S A, Lo S L, Krzycki J A. Clustered genes encoding the methyltransferases of methanogenesis from monomethylamine. *J Bacteriol*, 1998, **180** (13): 3432~3440
- 5 Paul L, Ferguson D J, Krzycki J A. The trimethylamine methyltransferase gene and multiple dimethylamine methyltransferase genes of *Methanosarcina barkeri* contain in-frame and read-through amber codons. *J Bacteriol*, 2000, **182** (9): 2520~2529
- 6 Galagan J E, Nusbaum C, Roy A, *et al.* The genome of *M. acetivorans* reveals extensive metabolic and physiological diversity. *Genome Res*, 2002, **12** (4): 532~542
- 7 Deppenmeier U, Johann A, Hartsch T, *et al.* The genome of *Methanosarcina mazei*: evidence for lateral gene transfer between bacteria and archaea. *J Mol Microbiol Biotechnol*, 2002, **4** (4): 453~461
- 8 James C M, Ferguson T K, Leykam J F, *et al.* The amber codon in the gene encoding the monomethylamine methyltransferase isolated from *Methanosarcina barkeri* is translated as a sense codon. *J Biol Chem*, 2001, **276** (36): 34252~34258
- 9 Hao B, Gong W, Ferguson T K, *et al.* A new UAG-encoded residue in the structure of a methanogen methyltransferase. *Science*, 2002, **296** (5572): 1462~1466
- 10 Hao B, Zhao G, Kang P T, *et al.* Reactivity and chemical synthesis of L-pyrrolysine - the 22nd genetically encoded amino acid. *Chem Biol*, 2004, **11** (9): 1317~1324
- 11 Srinivasan G, James C M, Krzycki J A. Pyrrolysine encoded by UAG in archaea: charging of a UAG-decoding specialized tRNA. *Science*, 2002, **296** (5572): 1459~1462
- 12 Blight S K, Larue R C, Mahapatra A, *et al.* Direct charging of tRNA^{CUA} with pyrrolysine *in vitro* and *in vivo*. *Nature*, 2004, **431** (7006): 333~335
- 13 Polcarpo C, Ambrogelly A, Ruan B, *et al.* Activation of the pyrrolysine suppressor tRNA requires formation of a ternary complex with class I and class II lysyl-tRNA synthetases. *Mol Cell*, 2003, **12** (2): 287~294
- 14 Polcarpo C, Ambrogelly A, Berube A, *et al.* An aminoacyl-tRNA

- synthetase that specifically activates pyrrolysine. Proc Natl Acad Sci USA, 2004, **101** (34): 12450~12454
- 15 Schimmel P, Beebe K. Genetic code seizes pyrrolysine. Nature, 2004, **431** (7006): 257~258
- 16 Praetorius-Ibba M, Ibba M. Aminoacyl-tRNA synthetase in archaea: different but not unique. Molecular Microbiology, 2003, **48** (3): 631~637
- 17 Theobald-Dietrich A, Frugier M, Giege R, *et al.* Atypical archaeal tRNA pyrrolysine transcript behaves towards EF-Tu as a typical elongator tRNA. Nucleic Acids Res, 2004, **32** (3): 1091~1096

Pyrrolysine: The 22nd Amino Acid*

LING Chen^{1,2)}, WANG En-Duo^{1)**}

(¹)State Key Laboratory of Molecular Biology, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, The Chinese Academy of Sciences, Shanghai 200031, China;

(²)Graduate School of The Chinese Academy of Sciences, Shanghai 200031, China)

Abstract Pyrrolysine, known as the 22nd amino acid, is found in *Methanosarcina barkeri* (*M.barkeri*) methylamine methyltransferases. It comes from the sense decoding of the UAG amber stop codon. It has specific pyrrolysyl-tRNA synthetase and tRNA^{Pyl}. tRNA^{Pyl} has noncanonical secondary structure. *M.barkeri* has two routes: direct and indirect routes to synthesize pyrrolysyl-tRNA^{Pyl}. The special structure in the mRNA and other unknown mechanisms may control the decoding of UAG as pyrrolysine or termination signal. Pyrrolysine was compared with the 21st amino acid: selenocysteine.

Key words pyrrolysine, tRNA^{Pyl}, pyrrolysyl-tRNA synthetase, pyrrolysine inserting element, selenocysteine

*This work was supported by a grant from The National Natural Sciences Foundation of China (30330180).

**Corresponding author . Tel: 86-21-54921241, Fax: 86-21-54921011, E-mail: edwang@sibs.ac.cn

Received: December 21, 2004 Accepted: January 28, 2005