

血色红假单胞菌 Form II Rubisco 基因的克隆及序列分析

吕 红 王 竞 杜翠红 周集体 安利佳

(大连理工大学环境与生命学院 大连 116023)

摘要: 根据 GenBank 中已登录的 Form II Rubisco (*cbbM*) 基因序列, 设计一对引物 R_1 和 R_2 , 并在引物 5'端分别加上 *Sma* I 和 *Not* I 位点。以血色红假单胞菌 1.2352 的染色体为模板, 通过 PCR 扩增目的片段, 克隆后测序。核苷酸序列测定结果表明, 该菌株的 *cbbM* 基因全长为 1,386 bp, 共编码 461 个氨基酸, 推导的氨基酸序列与其它生物相应序列的同源性在 76% ~ 90% 之间。表明了 *cbbM* 基因在进化上的保守性。

关键词: 血色红假单胞菌, *cbbM*, 测序, 同源比较

中图分类号: Q785 文献标识码: A 文章编号: 0253-2654 (2004) 02-0076-06

CLONING AND SEQUENCE ANALYSIS OF RHODOPSEUDOMONAS RUTILA FORMII RUBISCO

Liu Hong WANG Jing DU Cui-Hong ZHOU Ji-Ti AN Li-Jia

(School of Environmental and Biological Science and Technology, Dalian University of Technology, Dalian 116023)

Abstract: The *cbbM* gene of *Rhodopseudomonas rutila* 1.2352 was obtained by PCR from the chromosome of the strain with primers designed according to the *cbbM* genes in GenBank database. The result of sequencing shows that 1,386 bp of *cbbM* gene covers the complete open frame, encoding 461 amino acid residues. There is homology of 76% ~ 90% at deduced amino acid level. It reveals that the FormII Rubisco is conserved throughout evolution.

Key words: *Rhodopseudomonas rutila*, *CbbM*, Sequencing, Homology comparison

1, 5-二磷酸核酮糖羧化酶/加氧酶 (简称 Rubisco) 是卡尔文循环的关键酶, 该酶既催化 CO₂ 固定反应, 又能催化 CO₂ 氧合反应, 放出 CO₂。可见, Rubisco 在这两种反应中起着重要的枢纽作用, 它调节着细胞中碳的流向。特别是在植物当中, 它影响了 CO₂ 的固定效率和作物的产量, 并且它是植物可溶性蛋白中含量最高的蛋白质^[1]。因此, 对于 Rubisco 的研究既有理论意义又有应用价值。

Rubisco 一般分为两种类型^[2]: Form I Rubisco 和 Form II Rubisco。Form I Rubisco 广泛存在于植物和一些原核生物中。它是由 8 个大亚基和 8 个小亚基组成的 L₈S₈ 结构, 分子量约为 550kD。而 Form II Rubisco 是由大亚基组成的 L_x 结构, 分子量范围为 110 ~ 450kD, 最初是在深红红螺菌 (*Rhodospirillum rubrum*) 中发现的。它除了含有与催化相关的保守氨基酸残基如 Lys-329、Lys-191 和 Lys-166 等^[3]外, 所有亚基氨基酸序列基本上不同于 Form I 的大亚基。研究发现许多自养细菌中都含有 Rubisco 的两种类型, 血色红假单胞菌 (*Rhodopseudomonas rutila*) 就是其中之一。本文以血色红假单胞菌 (AS1.2352) 的染色体为模板, 通过 PCR 扩增到 Form II Rubisco (*cbbM*) 基因序列, 克隆后测序, 并与已发表的相应序列进行比较。为深入研究 Rubisco 基因结构及进一步构建

固定CO₂的基因工程菌提供理论依据。

1 材料与方法

1.1 菌种及质粒

血色红假单胞菌(AS1.2352)购自中国科学院菌种保藏中心,大肠杆菌(*E.coli*)BL21和载体pGEM4T-1购自宝生物工程(大连)有限公司。

1.2 生化试剂

Taq酶、dNTP、T4连接酶、限制性内切酶均购自宝生物工程(大连)有限公司,RNaseA、溶菌酶、氨苄青霉素购自上海华美生物公司。

1.3 菌株1.2352的异养培养

用带有橡皮塞的血清瓶作为培养容器,30℃、光照(光照强度约2,000 Lx)、静止培养。异养培养基参见文献[4]。

1.4 菌株1.2352染色体的提取

将菌株1.2352异养培养至对数中期,离心收集菌体,放入-20℃冰箱过夜。取出于37℃迅速融化,按每克湿菌体加入25mL提取液(50mmol/L Tris-HCl, pH8.0; 10mmol/L EDTA)悬浮菌体。加入溶菌酶至1mg/mL,37℃水浴1h,加入10%SDS至终浓度1.6%,55℃轻摇15min,加入5mol/L KAc至终浓度0.5mol/L,置于冰浴2h,加入等体积氯仿/异戊醇(24:1),混匀,8,000 r/min离心10min。吸取上清,向上清液中再加入等体积的苯酚/氯仿/异戊醇(25:24:1),氯仿/异戊醇,各抽提1次。加入2倍体积的预冷无水乙醇,-20℃冰箱放置4h,离心取沉淀,自然晾干。加入适量纯水溶解,弃去不溶物,加入RNaseA至终浓度10μg/mL,37℃水浴30min,冷却后备用。

1.5 PCR引物的设计

根据GenBank中已登录的Form II Rubisco(*cbbM*)基因序列,设计一对引物R₁和R₂。上游引物R₁: 5'-TACCCGGGATGGGACCAAGTCGAACCGCTAC-3',下游引物R₂: 5'-TCAGCGGCCGCTTACGCCCTGCGGCTTCAG-3',为克隆方便,引物的5'端分别设计有*Sma*I和*Not*I位点,由宝生物工程(大连)有限公司合成设计引物。

1.6 PCR扩增

以菌株1.2352的染色体为模板进行PCR扩增,PCR扩增条件:94℃,1min,进入循环:94℃,30s;55℃,30s;72℃,1min;30个循环;72℃,7min。PCR结束后电泳,将含扩增产物的琼脂块切下,用宝生物DNA凝胶回收试剂盒回收。

1.7 PCR产物的克隆和鉴定

将扩增产物和pGEX4T-1载体分别用*Sma*I/*Not*I双酶切后,前者乙醇沉淀后回收,后者电泳回收,用宝生物DNA连接试剂盒进行连接反应,然后电转化至大肠杆菌BL21感受态中。涂布LB(内含50mg/L氨苄青霉素)平板37℃过夜培养,挑选阳性菌落进行摇瓶培养,碱裂解法^[5]提取质粒,命名为LR。将质粒LR用*Sma*I/*Not*I双酶切及PCR扩增来鉴定。

1.8 *cbbM*基因的序列测定和同源比较

序列测定由宝生物工程(大连)有限公司完成,分析软件为GeneDoc,所用数据库为GenBank、PDB数据库。

2 结果

2.1 菌株 1.2352 染色体提取和 PCR 扩增

如上所述, 将提取物进行电泳(见图1), 可以看到分子量很大的一条DNA带, 即染色体。以它为模板进行PCR扩增, 所得扩增产物为一条分子量约为1.4kb的DNA带, 没有其它非特异DNA带出现(见图2)。而且其分子量与GenBank中已登录的 $cbbM$ 基因分子量很相近。

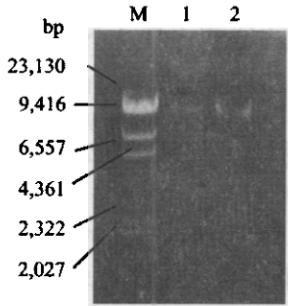


图1 菌株1.2352的染色体

M λDNA/Hind III, 1, 2 均为染色体

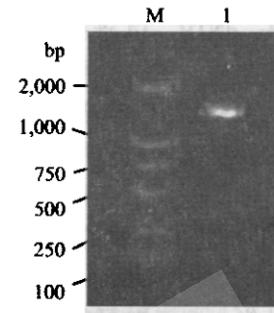


图2 $cbbM$ 基因的PCR扩增

M DL2000, 1 PCR扩增产物

2.2 重组质粒的鉴定

将重组质粒LR进行 Sma I单酶切及 Sma I/ Not I双酶切(见图3), 结果表明该质粒LR比载体分子量大, 双酶切后有一条分子量约为1.4kb的DNA带, 以重组质粒LR为模板进行PCR扩增, 反应条件如上, 在电泳图上也呈现出一条分子量约为1.4kb的DNA带(见图4), 表明克隆成功。

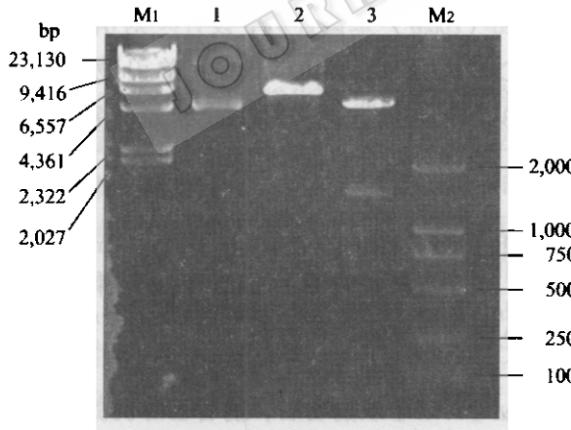


图3 重组质粒LR的酶切谱图

M1 λDNA/Hind III, 1 重组质粒LR/ Sma I, 2 重组质粒LR/ Sma I + Not I, M2 DL2000

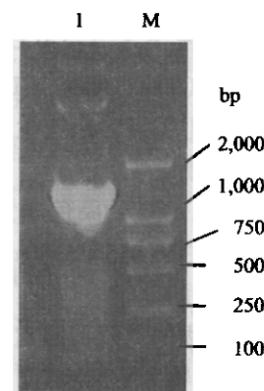


图4 重组质粒LR的PCR扩增

M DL2000, 1 PCR扩增产物

2.3 菌株1.2352 $cbbM$ 基因的核苷酸序列及推导的氨基酸序列

在PCR引物的设计中, 由于上游引物含启动密码子ATG(下划线标出), 下游引物含终止密码子TTA(下划线标出), 因此PCR扩增产物包含基因的完整读码框(图5)。核苷酸序列分析表明, PCR扩增产物长约1,400个核苷酸, 含有一个完整的 $cbbM$ 基因

序列，长为1,386个核苷酸，共编码461个氨基酸，G+C含量为66.4%，单个大亚基的分子量理论值为50.492 kD。

TGGATCCCCGAATCCCGG	ATG GAC CAG TCG AAC CCC TAC CCC AAG CTC AAC CTC	57
M D Q S N R Y A N L N L		12
AAA GAG AGC GAG CTG ATC GCG GGC CGA CGG CAC GTG CTC TGC TCC GCC TAC ATC ATC AAG CGG		117
K E S E L I I A G G C R H V L C A Y I M K P		32
AAG GCC GCC TTC CCT AAC TTC ATC CAG ACC CCC GCC CAC TTC GCC CCC GAA TCC TCG ACC		177
K A G F G N F I Q T A A H F A A E S S T		52
GGC ACC AAC CTC GAA GTG TCC ACC ACC GAC TTC ACC CGC GGC GTC GAC GCG CTG GTG		237
C T N V E V S T T D D F T R G V D A L V		72
TAC GAG ATC GAC GAG GCC AAG GGC CTG ATC AAG ATC CCG TAC CCG ATC GAG CTG TTC GAT		297
Y E I D E A K G L M K I A Y P I E L F D		92
CGC AAC CTC ATC GAC GCC CGC CGC ATG ATC GCC TCG TTC CTG ACC CTG AGC ATC GGC AAC		357
R N V I D G R A M I A S F L T L T I G N		112
AAC CAG GGC ATG CGC GAC GTC GAA TAC GCC AAG ATG TAC GAC TTC TAC GTC CCG CCG CCT		417
N Q G M G D V E Y A K M Y D F Y V P P A		132
TAC TTG AAG CTG TTC GAC GCC CCG TCG AGC ACG ATT CCG GAT CTG TGG CGC GTG CTC CGC		477
Y L K L F D G P S T T I R D L W R V L G		152
CCC CGC GTC AAC GGC CCC TTC ATC GTC CGC ACC ATC ATC AAG CCG AAG CTC GGC CTG		537
R P V V N G G F I V G T I I K P K L G L		172
CGC CGC CAG CGC TTC GCC AAT GCC TGC TAC GAT TTC TGG CTG CGC GGC GAC TTC ATC AAG		597
R P Q P F A N A C Y D F W L G G D F I K		192
AAC GAC GAG CGG CAG GGC AAC CAG GTG TTC CGG CGG TTC AAC GAC ACC GTC CCC CCC GTC		657
N D E P Q G N Q V F A P F K D T V R A V		212
GCC GAC GCC ATG CGC CCC CGC CAC GAC AAG ACC CCC GAA CCC AAC CTG TTC TCG TTC AAC		717
A D A M R R A Q D K T C E A K L F S F N		232
ATC ACC GCC CAC CAT CAC TAC GAC ATG CTC CGG CGC GGT GAG TTC ATC CTC GAG ACC TTC		777
I T A D D H Y E M L A R C E F I L E T F		252
GCC GAC AAC GCC GAC CAC ATC GCG TTC CTG GTC GAC GCC TAC GTT GCC GGC CCG GGC GCG		837
A D N A D H I A F L V D G Y V A G P A A		272
GTG ACC ACC GCC CGC CGC GCG TTC CCG AAG CAG TAC CTG CAC TAC CAT CGC CCC GGC CAC		897
V T T A R R A F P K Q Y L H Y H R A G H		292
GGT CGG GTG ACC TCG CGG CAG TCC AAG CGC CGG TAG ACC GCC TTC GTC CTG TCG AAG ATG		957
G A V T S P Q S K R G Y T A F V L S K M		312
GGC CGC TTG CAG GGC GCC TCG GGC ATC CAC ACC CGC ACC ATG GGC TTC GGC AAG ATG GAA		1017
A R L Q G A S G I H T G T M G F G K M E		332
GCC GAA GCC CGC GAC CGC CGC ATC GCC TAC ATG ATC ACC GAG CAT TCG GCC GAC CGT CGG		1077
G E A A D R A I A Y M I T E D S A D G P		352
TAC TTC CAC CAC GAG TGG CTC CGC CTC AAC CCG ACC ACC CGG ATC ATC TCC GGC GGC ATG		1137
Y F H Q E W L G L N P T T P I I S G G M		372
AAC CGG CTG CGG ATG CGG GCC TTC TCC GAC AAC CTC CGC CAC TCC AAC CTG ATC ATG ACC		1197
N A L R M P G F F D N L G H S N L I M T		392
GGC CGC GGT CGG CGG TTC CGT CAC GTC GAC CGC CGG CCC CCT GCT GCC AAG TCG CTG CGT		1257
A G C G A F G H V D G C A A G A K S L R		412
CAG CCC CAA CAG TCC TCG AAC CAG CGC CGC GAT CGG GTC GAG TTC CCC AAG GAT CAC CGC		1317
Q A E Q C W K Q G A D P V E F A K D H R		432
CAA TTC GCC CGC CGC TTC GAG AGC TTC CCG CAG GAC GCC GAC AAG CTG TAT CCG AAC TGG		1377
E F A R A F E S F P Q D A D K L Y P N W		452
CGC CGC AAC CTG AAG CCC CAG GCG CGG TAA CGGGCGCCATCTCACTGACTCA		1430
R A K L K P Q A A * 461		

图5 菌株1.2352 cbbM基因的核苷酸序列及推导的氨基酸序列

2.4 cbbM基因的氨基酸序列的同源比较

由菌株1.2352 cbbM基因推导的氨基酸序列与其它生物的相应序列进行了同源比较（见表1和图6），结果表明，血色红假单胞菌cbbM基因的氨基酸序列与同是红假单胞菌属的*R. globiformis*^[6]和*R. capsulatus*的相应序列同源性最高，都是90%。与其它菌属的*R. rubrum*、*T. denitrificans*、*H. marinus*^[7]的相应序列同源性依次为86%、84%和76%。与*R. pachyptila symbiont*的相应序列同源性则为87%。可见它们的Form II Rubisco的同源性很高，表明其在进化上的保守性。

表 1 *cbbM* 基因的氨基酸序列的同源比较

同源性	R.glo	R.nub	R.cap	H.mar	T.den	R.pac	
R.rutia	90%	86%	90%	76%	84%	87%	
R.rut	--HDQSNRYANINLKESE	20	--	* 40	*	60	
R.glo	--MDQSNRYARLDLQ	KEADLTIAGCRHVLCA	YMKF	KAGF	ENFIQTA	AEE	
R.rub	--MDQSNRYANLALKKD	LTIAGCKHVLCA	MKFK	GAGFC	AAE	STGTNVEV	
R.cap	--MDQSNRYARLDLKEA	DLLTIAGCRHVLCA	YMKF	KAGY	AAE	STGTNVEV	
T.den	--MDQSNRYARLDLKEA	DLLTIAGCRHVLCA	YMKF	KAGY	AAE	STGTNVEV	
H.mar	--MDQSNRYARLDLKEA	DLLTIAGCRHVLCA	YMKF	KAGY	AAE	STGTNVEV	
R.pac	MDALDTQNRYSDEL	SLKEDELIAS	EDYVLCA	YMKF	SYG-Y	AAE	
	6DQ3 RY L L e	6a G h6L AY	6kpk G5	56	AAHFAAEE	STGTNVEV	
R.rut	--DDFTRGVDALVYEID	80	--	* 100	*	120	
R.glo	--DDFTRGVDALVYEID	GBK	--	--BIMKIA	PWLE	EDFDRNIV	
R.rub	--DDFTRGVDALVYEID	PEKAK	--	--ELTKIA	PWLE	EDFDRNIV	
R.cap	--DDFTRGVDALVYEID	PEK	--	--EIMKIA	PWLE	EDFDRNIV	
T.den	--DDFTKGVDALVYYIDEAS	--	--EDM	RIAYPLE	DFDRNIV	TDGKAMIA	
H.mar	--DDFTKGVDALVYYIDEAS	--	--EAA	FGDKGGI	MLVSEL	TDGKAMIA	
R.pac	--DDFTKGVDALVYYIDEAS	--	--EIMKIA	PWLE	DFDRNIV	TDGKAMIA	
	DDFT4GVDALVYYe6D		m4IAYP6	LFD N6 DG	m	sf63L 6GNNQGMGD6	
R.rut	--SYAKMMDYDFYVE	140	--	* 160	*	180	
R.glo	--SYAKMMDYDFYVE	BAYLKLFDG	PTT	I	RDLW	VRLGR	
R.rub	--SYAKMMDYDFYVE	BCYLR	LDG	PSMNIA	ADLWR	VRLGR	
R.cap	--SYAKMMDYDFYVE	BCYLR	LDG	PSMNIA	ADLWR	VRLGR	
T.den	--EHAKMIDFVYERCI	QMF	DCG	ATDIS	SNLWR	ILGR	
H.mar	--KGLKMLDFVLEKVM	K	F	DFG	SDLMW	VLGR	
R.pac	--SYAKMMDYDFYVE	BCYLR	LDG	PSMNIA	ADLWR	VRLGR	
	e aKM DFy6P	5DGP	I	6w r6LGrp	1GG 6	GTTIIPKLGRLP	
R.rut	--DFWLGGDFIKNDPQ	200	--	* 220	*	240	
R.glo	--DFWLGGDFIKNDPQ	CNC	PPADL	KDTI	IRI	ADAMKRAQD	
R.rub	--DFWLGGDFIKNDPQ	CNC	PPADL	KDTI	IRI	ADAMKRAQD	
R.cap	--DFWLGGDFIKNDPQ	CNC	PPADL	KDTI	IRI	ADAMKRAQD	
T.den	--QFWLGGDFIKNDPQ	GNQ	CE	CKLKVPL	PLVY	DKAQQD	
H.mar	--DFWLGGDFIKNDPQ	GNQ	CE	CKLKVPL	PLVY	DKAQQD	
R.pac	--DFWLGGDFIKNDPQ	GNQ	CE	CKLKVPL	PLVY	DKAQQD	
	FWLGGd IKNDPQ	NQ F	k	6 VadAM	RAQD	TG2AKLF8	
R.rut	--SFILETPAD-NADHIA	260	--	* 280	*	300	
R.glo	--SFILETPCE-NADHVA	LVDG	VIA	GAAPVTT	TARRAF	PKOYLHVR	
R.rub	--SFILETPCE-NADHVA	LVDG	VIA	GAAPVTT	TARRCF	DRPLHYR	
R.cap	--SFILETPCE-NADHVA	LVDG	VIA	GAAPVTT	TARRSF	PROFLHYR	
T.den	--DYALEVEGR-DADKLA	E	LVDG	VGG	GMVTT	AROYEGOY	
H.mar	--BYVLGEPAKYGEKHVA	LVDG	V	PGAV	ITSRRAFF	DTYLH	
R.pac	--SFILETPCE-EASQVA	LVDG	VAG	GA	TA	ARRFNQEL	
	e5 L F	a	6AFLVD	5V	Gp	ttaRR 5	
R.rut	--VLSKSMARIQGASGI	320	--	* 340	*	360	
R.glo	--VLSKSMARIQGASGI	HTGTMCF	GKME	GEAAD	RAIAYM	ITEDSAD	
R.rub	--VLSKSMARIQGASGI	HTGTMCF	GKME	GEAAD	KIMAYM	ITEDSAD	
R.cap	--VLSKSMARIQGASGI	HTGTMCF	GKME	GEAAD	KIMAYM	ITEDSAD	
T.den	--VIAKMSRIQGASGI	HTGTMCF	GKME	GE	GDRIAYM	ITEDSAD	
H.mar	--CYMKLARIMGASGI	HTGTMCF	GKME	GE	NDHE	DEAOGP	
R.pac	--VHSKMTTRI	QGASGMHV	GTMGY	GKME	GEAS	RLIAYMI	
	v K6 RL	GasG6H	GTMG5	GKMBG	d4 6AYM6	ERDSADGP	
R.rut	--80	*	400	*	420	*	440
R.glo	--GMNALRMPGF	FDNLGH	SNLIM	TAGG	GAPGH	VDFG	ADP
R.rub	--GMNALRMPGF	FDNLGH	SNVII	QT	GGGAF	GHDG	YF
R.cap	--GMNALRMPGF	FDNLGH	SNVII	QT	GGGAF	GHDG	YF
T.den	--GMNALRMPGF	FDNLGH	SNVII	QT	GGGAF	GHDG	YF
H.mar	--GMDALR	LPGEFFENL	GHG	NT	AGG	SYGHIDS	PAAGC
R.pac	--GMNALR	LPGEFFENL	GHG	NT	AGG	SYGHIDS	PAAGC
	GM1ALR6PGFF	NLgh N6I	T	GGG 5GH6D	AGa	SLrQ	wk G d 6 a h4E
R.rut	--PARAFES	EQDAD	KLY	ENT	RAK	--KPQAA	--
R.glo	--PARAFES	EQDAD	KLY	ENT	RAK	--H	--RAA
R.rub	--PARAFES	EQDAD	KLY	ENT	RAK	--H	--RAA
R.cap	--PARAFES	EQDAD	KLY	ENT	RAK	--H	--RAA
T.den	--PARAFES	EQDAD	KLY	ENT	RAK	--H	--RAA
H.mar	--PARAFES	EQDAD	KLY	ENT	RAK	--H	--RAA
R.pac	--PARAFES	EQDAD	KLY	ENT	RAK	--H	--RAA
	PARAFES	EF	DAD	IP	PCW	WRD	KLVH

图 6 *cbbM* 基因的氨基酸序列的同源比较

3 讨论

研究表明 Form II Rubisco 的同源性很高, 这便于人们通过核酸杂交及 PCR 来发现其它生物的 Form II Rubisco。

由于人们最初是在 *R. rubrum* 中发现的 Form II Rubisco, 因此对其研究的比较深入。研究者通过亲合标记和 X 射线等方法对其活性位点已经了解的比较清楚, 比如 Lys-329、Gly328、Lys-191、Ile-164 和 Lys-166 等^[3,8,9] 位点。这些位点中, 以 Lys 残基居多, 研究表明 Rubisco 活性区 Lys 残基的 ε 氨基与 CO₂ 作用形成氨基碳酸基团, 然后与 Mg²⁺ 键合。Rubisco 反应平衡及动力学研究表明, ε 氨基的 CO₂ 化对形成有活性的 Rubisco 有重要作用。从图 6 中可以看到, 在这些氨基酸及其附近的区域, 上述生物都具有相同的序列, 表明这些位点有可能也是血色红假单胞菌 1.2352 Form II Rubisco 的活性位点, 这有待进一步证实。

血色红假单胞菌 1.2352 *cbbM* 基因的克隆和测序, 为进一步进行该基因的表达和定点诱变奠定了基础。

参 考 文 献

- [1] 陈为钩, 赵贵文, 顾月华. 生物化学与生物物理进展, 1999, 26 (5): 433~436.
- [2] 董晓丽, 周集体, 杜翠红, 等. 高技术通讯, 2001, (12): 95~97.
- [3] George H L. Chen Y R, Hartman F C. Biochem, 1993, 32: 9018~9024.
- [4] 钱存荣, 黄仪秀编. 微生物学实验教程. 北京: 北京大学出版社, 1999. 211~212.
- [5] Sambrook J, Fritsch E F, Maniatis T. Molecular cloning, a laboratory manual second edition. New York: Cold Spring Harbor Laboratory Press, 1989.
- [6] Janeil L G, Deane L F, Tabita F R. J Biol Chem, 1991, 266: 14646~14653.
- [7] Yaguchi T, Chung S Y, Igarashi Y, et al. Biochi Biotechnol Biochem, 1994, 58 (9): 1733~1737.
- [8] Tabita F R, McFadden B A. J Biol Chem, 1974, 249: 3459~3464.
- [9] Chene P, Day A G, Fersht A R, et al. Biochem Biophys Res Commun, 1997, 232 (2): 482~486.