

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.



[12] 发明专利说明书

C12N 15/62 (2006.01)

C07K 14/22 (2006.01)

C07K 19/00 (2006.01)

专利号 ZL 01808738.8

[45] 授权公告日 2007年9月26日

[11] 授权公告号 CN 100339482C

[22] 申请日 2001.2.28 [21] 申请号 01808738.8

[30] 优先权

[32] 2000.2.28 [33] GB [31] 0004695.3

[32] 2000.11.13 [33] GB [31] 0027675.8

[86] 国际申请 PCT/IB2001/000420 2001.2.28

[87] 国际公布 WO2001/064920 英 2001.9.7

[85] 进入国家阶段日期 2002.10.28

[73] 专利权人 启龙股份公司

地址 意大利锡耶纳

[72] 发明人 M·B·阿科 M·科曼杜西

C·加莱蒂 V·马斯格阿尼

M·M·朱利亚尼 M·皮扎

[56] 参考文献

US5547670A 1996.8.20

WO9936544A 1999.7.22

WO9216643A 1992.10.1

审查员 曹克浩

[74] 专利代理机构 上海专利商标事务所有限公司

代理人 徐迅

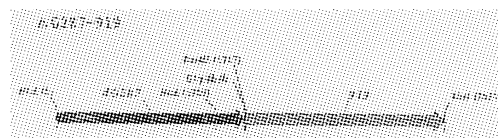
权利要求书 3 页 说明书 45 页 附图 5 页

[54] 发明名称

奈瑟球菌蛋白质的杂交表达

[57] 摘要

以单个杂交蛋白形式表达两种或多种奈瑟球菌的蛋白质(如 A 和 B), 其可以式 $\text{NH}_2 - \text{A} - \text{B} - \text{COOH}$ 简单地表示。



1. 一种同时异源表达第一种蛋白和第二种蛋白质的方法，其特征在于：

(a) 第一种蛋白和第二种蛋白翻译成一条多肽链；

(b) 第一种蛋白：

(i) 是氨基酸序列 Δ 287：

```

SPDVKS ADTLSKPAAP VVSEKETEAK EDAPQAGSQ QGAPSAQGSQ DMAAVSEENT
GNGGAVTADN PKNEDEVAQN DMPQNAAGTD SSTPNHTPDF NMLAGNMENQ ATDAGESSQP
ANQPDMANAA DGMQGDDPSA GGQNAAGTAA QGANQAGNNQ AAGSSDPIPA SNPAPANGGS
NFGRVDLANG VLIDGPSQNI TLTHCKGDSC SGNNFLDEEV QLKSEFEKLS DADKISNYKK
DGKNDKFBVL VADSVQMRGI NQYIIFYKPK PTSFARFRS ARSRRSLPAE MPLIPVNQAD
TLIVDGEAVS LTGHSGNIFA PEGNYRYLTY GAEKLPGGSY ALRVQGEPAK GEMLAGAAYV
NGEVLHFHTE NGRPYPTRGR FAAKVDFGSK SVDGIIDSGD DLHMGTQKFK AAIDGNGFKG
TWFENGSGDV SGKPYGPAGE EVAGKYSYRP TDAEKGGFGV FAGKKEQD*

```

或

(ii) 是氨基酸序列 Δ G741：

```

VAAD IGAGLADALT APLDHKDKGL
QSLTLDQSVR KNEKLKLAQ GAEKTYGNGD SLMTGKLNK KVSRRDFIRQ
IEVDGQLITL ESGEFQVYKQ SHSALTAFQT BQIQDSEHSG KMAKRQFRI
GDIAGEHTSF DKLPEGGRAT YRGTAFGSDD AGGKLTYYTID FAAKQGNKGI
EHLKSPELNV DLAAADIKPD GKRHAVISGS VLYNQAEKGS YSLGIFGGKA
QEVAGSAEVK TVNGIRHIGL AAKQ*

```

2. 如权利要求 1 所述的方法，其特征在于，第二种蛋白质的氨基酸序列选自：

(i) WO99/24578 公开的偶数 SEQ ID NO:2-892；(ii) WO99/36544 公开的偶数 SEQ ID NO:2-90；(iii) WO99/57280 公开的偶数 SEQ ID NO:2-3020；(iv) WO99/57280 公开的偶数 SEQ ID NO:3040-3114；(v) WO99/57280 公开的偶数 SEQ ID NO:3115-3241；(vi) 蛋白质 NMB0001-NMB2160。

3. 如权利要求 1 或 2 所述的方法，其特征在于，第二种蛋白质是氨基酸序列 ORF46：

```

1  LGISRKISLI LSILAVCLPM HAHASDLAND SFIRQVLDRO HFEPDGKYHL
51 FGSRGELAER SGHIGLGKIQ SHQLGNLMIQ QAAIKGNIGY IVRFSDHGHE
101 VHSPFDNHAS HSDSDEAGSP VDGFSLYRIH WDGYEHHPAD GYDGPQGGGY

```

```

151 PAPKGARDIY SYDIKGVAQN IRLNLTDNRS TGQRLADRFH NAGSMLTQGV
201 GDGFKRATRY SPELDRSGNA AEAFTNGTADI VKNIIGAAGE IVGAGDAVQG
251 ISEGSNIAVM HGLGLLSTEN KMARINDLAD MAQLKDYAAA AIRDWAVQNP
301 NAAQGIKAVS NIFMAAIPK GIGAVRGKYG LGGITAHPIK RSQMGAIALP
351 KGKSAVSDNF ADAAYAKYPS PYHSRNIKSN LEQRYGKANI TSSTVPPSNG
401 KNVKLADQRH PKTGVVFDGK GFPNFEKHVK YDTKLDIQEL SGGGIKAKP
451 VSDAKPRWEV DRKLNKLTTR EQVEKNVQEI RGNKNSNFS QHAQLEREIN
501 KLKSADEINF ADGMGKFTDS MNDKAFSRLV KSVKENGFTN PVVEYVEING
551 KAYIVRGNNR VFAAEYLGRH HELKPKKVDV PVPNTSWKNP TDVLESNGNV
601 KRPRYRSK*

```

4.如权利要求 1 或 2 所述的方法，其特征在于，第二种蛋白质是氨基酸序列 Δ G983:

```

NSRATTAKSA AVSYAGIKNE MCKDRSMLCA GRDDVAVTDR DAKINAPPPN
LHTGDFPNPN DAYKNLINLK PAIEAGYTGR GVEVGIVDTG ESVGSISFPE
LYGRKEHGYN ENYKNYTAYM RKEAPEDGGG KDIEASFDD E AVIETEAKPT
DIRHVKEIGH IDLVSHIIGG RSVVDGRPAGG IAPDATLHIM NTNDETKNEM
MVAAIRNAWV KLGGERGVRIV NNSFGTTSRA GTADLFQIAN SEEQYRQALL
DYSGGDKTDE GIRLMQQSDY GNLSYHIRNK NMLFIFSTGN DAQAQNPNTYA
LLPFYKDAQ KGIITVAGVD RSGEKFKREM YGEPGTEPLE YGSNHCGITA
MWCLSAPYEA SVRFTRTNPI QIAGTSFSAP IVTGTAALLL QKYPWMSNDN
LRTTLLTTAQ DIGAVGVDSK FGWGLLDAGK AMNGPASFPF GDFTADTKGT
SDIAYSFRND ISGTGGLIKK GGSQQLQHLGN NTYTGKTIIE GSSLVLYGNN
KSDMRVETKG ALIYNGAASG GSLNSDGIVY LADTDQSGAN ETVHIKGSLO
LDGKGTLYTR LGKLLKVDGT AIIGGKLYMS ARGKGAGYLN STGRRVPFLS
AAKIGQDYSF FTNIETDGL LASLDSVEKT AGSEGDLSY YVRRGNAART
ASAAAHSAPA GLKHAVEQGG SNLENLMVEL DASESSATPE TVETAADRT
DMPGIRPYGA TFRAAAVQH ANAADGVRIE NSLAATVYAD STAAHADMQG
RRLKAVSDGL DHNGTGLRVI AQTQDDGGTW EQGGVEGKMR GSTQTVGIAA
KTGENTTAAA TLGMGRSTWS ENSANAKTDS ISLFAIRHD AGDIGYKGL
FSYGRYKNSI SRSTGADEHA EGSVNGTLMQ LGALGGVNVF FAATGDLTVE
GGLRYDLLKQ DAFAEKGSAL GWSGNSLTEG TLVGLAGLKL SQPLSDKAVL
FATAGVERDL NGRDYTVTGG FTGATAATGK TGARNMPHTR LVAGLGADVE
FGNGWNLAR YSYAGSKQYG NHSGRVGVGY RF*

```

5.如权利要求 1 或 2 所述的方法，其特征在于，第二种蛋白质是氨基酸序列 919:

```

CQSKSIQTFPQDTSVINGPDRPVGIPDPAGTTVGGGAVYTVVPHLSLPHWAAQDFAKGLQSFRLGCANLKNRQGWQDVCQAQAFQTFVHSF
QAKQFFERYFTPHQVANGSLAGTVTGYEYVPLKGDPRRTAQARFPIYGIPIDFISVPLPAGLNSGJALVRIQRTGKNSGTIDNTGGTHIAD
LSRFPITARTTAIKRPFEGSRFLPYHTRNQINGGALDCKAPITLCYAEDEFVELFMHIOGSGELKTFSGKYIRIGYADKNEHPYVSIGRYMD
KGYLKLQTSMQGIKAYMRONPQRLAEVLCQNPYSYIFPRELAGSSNDGPVGLGTFIMGEYAGAVDRHYITLGAPLFVATAHPVTRKALNRL
LMAQDTGSAIKGAVRVDYFWGYGDEAGELAGKQKTTGYVWQLLFGMGKPEYRP

```

6. 如权利要求 1 或 2 所述的方法，其特征在于，第二种蛋白质是氨基酸序列 953:

```

ATYKVDEYHANARFAIDHFMSTNVGGFYGLTGSVDFDQAKRDKIDITIPVANLQSGSQHFTDHLKSADI PDAAQYPIRIFVSTKGFNFKG
KLVSVGDNLTMHGKTAPVKLKAERFNCYQSPMAKTEVCGGDFSTTIDRTKNGVDYLVNVGMTKSVRIDIQIEAAKQ

```

7. 如权利要求 1 或 2 所述的方法，其特征在于，第二种蛋白质是氨基酸序列 961:

```

ATNDDVKKAAATVAIAAAYNNGQENGFKAGETIYDIDEDGTITKIDATAADVEADDFGLGLKQVVTNLTQVNVENKQNVDAKVKAAESEI
EKLTKLADTDAALADTDAALDATTNALNKLGENITTTFAETKINI VKIDEKLEAVADTVDKHAEAFND IADSLDETNTKADERVKTANEAK
QTAEETKQNVDAKVKAAETAAGKAEAAAGTANTAADKAEVAKVTDIKADIAINKONIAKQANSADVYTRRESDSKFRVIDGLNATTEKLD
TRLASAEKS LADHDTRLNGLDKTVSDLRKBTROGLAEQALSGLFQYVNGRFFVTRAVNGYKSESAVAIGTGFPRFTENFAAKAGVAVGTSS
GSSAAYHVGVNYEW

```

8.如权利要求 1 或 2 所述的方法，其特征在于，第一种蛋白质在多肽链中第二种蛋白质的上游。

9. 如权利要求 1 或 2 所述的方法，其特征在于，第一种蛋白质在多肽链中第二种蛋白质的下游。

10. 如权利要求 1 或 2 所述的方法，其特征在于，表达在大肠杆菌宿主中。

11. 式 NH₂-A-B-COOH 的杂交蛋白质，其中 A 和 B 是不同的，分别是：

(i)氨基酸序列 Δ G287:

SPDVKS ADTLSKPAAP VVSEKETEAK EDAPQAGSQ QGAPSAQGSQ DMAAVSEENT
 GNNGAVTADN PKNEDEVAQN DMPQNAAGTD SSTPNHTFDP NMLAGNMENQ ATDAGESSQP
 ANQPDMANAA DGMQGDPSA GGQNAAGTAA QGANQAGNNQ AAGSSDFIPA SNPAPANGGS
 NFGRVDLANG VLIDGPSQNI TLTHCKGDSC SGNNFLDEEV QLKSEFEKLS DADKISNYKK
 DGKNDKFVGL VADSVQMKGI NQYIIFYKPK PTSFARFRRS ARSRRSLPAE MPLIPVNQAD
 TLIVDGEAVS LTGHSGNIFA PEGNYRYLTY GAEKLPGGSY ALRVQGEPAK GEMLAGAAYV
 NGEVLHFHTE NGRPYPTRGR FAAKVDFGSK SVDGIIDSGD DLHMGTOQKFK AAIDGNGFGK
 TWTENSGSDV SGKPYGPAGE EVAGKYSYRP TDAEKGGFGV FAGKKEQD*

或

(ii)氨基酸序列 Δ G741:

VAAD IGAGLADALT APLDHKDKGL
 QSLTLDQSVR KNEKLLAAQ GAEKTYGNGD SLNTGKLEND KVSRLFIRQ
 IEVDGQLITL ESGEFQVYKQ SHSALTAFQT EQIQDSEHSG KMAKRFRI
 GDIAGEHTSF DKLPEGGRAT YRGTAFGSDD AGGKLTYYTID FAAKQNGKI
 EHLKSPFLNV DLAAADIKPD GKRHAVISGS VLYNQAEKGS YSLGIFGGKA
 QEVAGSAEVK TVNGIRHIGL AAKQ*

12.如权利要求 11 所述的蛋白质，其特征在于，A 和 B 直接连接。

13.如权利要求 11 所述的蛋白质，其特征在于，A 和 B 通过接头肽连接。

14.如权利要求 13 所述的蛋白质，其特征在于，所述接头肽是聚甘氨酸接头。

15.一种同时异源表达第一种蛋白和第二种蛋白质的方法，其特征在于：

-将第一种蛋白和第二种蛋白翻译成单多肽链；

-第一和第二蛋白选自：

(i)第一蛋白是‘919’，第二蛋白是‘Δ G287’；

(ii)第一蛋白是‘953’，第二蛋白是‘Δ G287’；

(iii)第一蛋白是‘ORF46.1’，第二蛋白是‘Δ G287’；

(iv)第一蛋白是‘ORF1’，第二蛋白是‘ORF46.1’；

(v)第一蛋白是‘919’，第二蛋白是‘ORF46.1’；

(vi)第一蛋白是‘919’，第二蛋白是‘519’；

(vii)第一蛋白是‘ORF97’，第二蛋白是‘225’；

-第一和第二蛋白的氨基酸序列如下：

蛋白质‘919’是 SEQ ID NO:4 的氨基酸 412-832，蛋白质‘Δ 287’是 SEQ ID NO:2，蛋白质‘953’是 SEQ ID NO:6 的氨基酸 412-579，蛋白质‘ORF46.1’是 SEQ ID NO:18 的氨基酸 1058-1466，蛋白质‘ORF1’是 WO99/24578 的 SEQ ID NO:650，蛋白质‘519’是 WO99/57280 的 SEQ ID NO:358，而蛋白质‘225’是 WO99/57280 的 SEQ ID NO:802。

16.权利要求 15 所述的方法，其特征在于，第二种蛋白质是‘Δ G287’，其中第二种蛋白质位于第一种蛋白质的 C-末端。

17.权利要求 15 所述的方法，其中表达是在大肠杆菌宿主中进行的。

奈瑟球菌蛋白质的杂交表达

本文将所引用的全部文献都纳入作为参考。

技术领域

本发明涉及蛋白质表达的领域。具体说，本发明涉及奈瑟球菌(如淋病奈瑟球菌或较佳地为脑膜炎奈瑟球菌)的蛋白质的异源表达。

背景技术

国际专利申请 WO 99/24578、WO99/36544、WO99/57280 和 WO 00/22430 公开了脑膜炎奈瑟球菌(*Neisseria meningitidis*)和淋病奈瑟球菌(*Neisseria gonorrhoeae*)的蛋白质。这些蛋白质一般是以 N-末端 GST-融合体或 C-末端 His-标记融合体在大肠杆菌中表达的(即异源表达)，虽然也公开了其它表达系统(包括在天然的奈瑟球菌中的表达)。

本发明的目的是提供这些蛋白质的异源表达的其它或改进方法。这些方法通常影响表达的水平、纯化的简易程度、表达的细胞内定位和/或表达的蛋白质的免疫学特性。

发明的公开

在本发明中，以单杂交蛋白表达本发明的两种或多种(如 3、4、5、6 或更多)蛋白质。较佳地，不使用非奈瑟球菌的融合配体(如 GST 或聚-His)。

这有两个优点。其一，可能不稳定或本身表达很差的蛋白质可以通过加入合适的克服该问题的杂交配体予以帮助。其二，简化工业生产，即制备两种分别有用的蛋白质只需要一次表达和纯化。

因此，本发明提供了同时异源表达本发明的两种或多种蛋白质的方法，其中所述的本发明的两种或多种蛋白质是融合的(即，它们是作为单多肽链翻译的)。

该方法通常包括如下步骤：获得编码本发明第一种蛋白质的第一核酸；获得编码本发明第二种蛋白质的第二核酸；连接第一和第二核酸。将得到的核酸插入表达载体中，或已作为表达载体的一部分。

当仅连接两种蛋白质时，可以式 $\text{NH}_2\text{-A-B-COOH}$ 简单地表示杂交蛋白。A 和 B 各选自任何奈瑟球菌蛋白质，尤其是那些以 SEQ#1-4326 表示的。该方法非常适合表达蛋白质 orf1、orf4、orf25、orf40、Orf46/46.1、orf83、233、287、292L、564、687、741、907、919、953、961 和 983。

由式 $\text{NH}_2\text{-A-B-COOH}$ 在下表中以 'X' 表示的 42 个杂交体是优选的：

↓A	B→	ORF46.1	287	741	919	953	961	983
ORF46.1			X	X	X	X	X	X
287	X			X	X	X	X	X
741	X	X			X	X	X	X
919	X	X	X			X	X	X
953	X	X	X	X			X	X
961	X	X	X	X	X	X		X
983	X	X	X	X	X	X	X	

因此，优选的以杂交体表达的蛋白质是 ORF46.1, 287, 741, 919, 953, 961 和 983。它们可以全长形式或多-甘氨酸缺失(ΔG)形式使用(如 $\Delta\text{G-287}$ 、 ΔGTbp2 、 ΔG741 、 ΔG983 等)、或以截短形式使用(如 $\Delta\text{1-287}$ 、 $\Delta\text{2-287}$ 等)，或以结构域缺失的形式使用(如 287B、287C、287BC、ORF46₁₋₄₃₃、ORF46₄₃₃₋₆₀₈、ORF46、961c 等)等。

特别优选的是：(a) 含 919 和 287 的杂交蛋白；(b) 含 953 和 287 的杂交蛋白；(c) 含 287 和 ORF46.1 的杂交蛋白；(d) 含 ORF1 和 ORF46.1 的杂交蛋白；(e) 含 919 和 ORF46.1 的杂交蛋白；(f) 含 ORF46.1 和 919 的杂交蛋白；(g) 含 ORF46.1、287 和 919 的杂交蛋白；(h) 含 919 和 519 的杂交蛋白；和(i) 含 ORF97 和 225 的杂交蛋白。

附图中显示了其它优选例，它们包括 $\Delta\text{G287-919}$ 、 $\Delta\text{G287-953}$ 、 $\Delta\text{G287-961}$ 、 $\Delta\text{G983-ORF46.1}$ 、 $\Delta\text{G983-741}$ 、 $\Delta\text{G983-961}$ 、 $\Delta\text{G983-961C}$ 、 $\Delta\text{G741-961}$ 、 $\Delta\text{G741-961C}$ 、 $\Delta\text{G741-983}$ 、 $\Delta\text{G741-ORF46.1}$ 、ORF46.1-741、ORF46.1-961、ORF46.1-961C、961-ORF46.1、961-741、961-983、961C-ORF46.1、

961C-741、961C-983、961CL-ORF46.1、961CL-741 和 961CL-983。

当使用 287 时，其优先位于杂交体的 C-末端；如果在 N-末端使用它，则优先使用 287 的 ΔG 形式(如与 ORF46.1、919、953 或 961 杂交的杂交体的 N-末端)。

当使用 287 时，其优先为菌株 2996 或菌株 394/98 的。

当使用 961 时，其优先在 N-末端。可使用 961 的结构域形式。

WO 99/66741 公开了 ORF46、287、919 和 953 的多态形式的排序。本发明可以使用这些多态形式中的任一形式。

较佳地，本发明杂交蛋白中的组成蛋白质(A 和 B)来源于是同一菌株。

杂交体中的融合蛋白可以是直接相连的，或者是通过接头肽相连的，如通过聚-甘氨酸接头(即，G_n，其中 n=3、4、5、6、7、8、9、10 或更多)或通过协助克隆的短肽序列相连。显然不宜将 ΔG 蛋白质连接于聚-甘氨酸接头的 C-末端。

融合蛋白可以缺失天然前导肽或可以包含 N-末端融合配体的前导肽序列。

宿主

较佳地，利用异源宿主。异源宿主可以是原核的或真核的。优先为大肠杆菌，但其它合适的宿主包括枯草芽孢杆菌(*Bacillus subtilis*)、霍乱弧菌 (*Vibrio cholerae*)、伤寒沙门氏菌 (*Salmonella typhi*)、鼠伤寒沙门氏菌(*salmonella typhimurium*)、脑膜炎奈瑟球菌、淋病奈瑟球菌、乳糖奈瑟球菌(*Neisseria lactamica*)、灰色奈瑟球菌(*Neisseria cinerea*)、分枝杆菌(如结核分枝杆菌 (*M.tuberculosis*)、酵母等。

载体、宿主等

如上所述的方法，本发明提供了(a)用于这些方法的核酸和载体；(b)含所述载体的宿主细胞；(c)可用这些方法表达的或可表达的蛋白质；(d)包含这些蛋白质的组合物，其可能适合作为疫苗、或例如诊断剂或免疫原性的组合物；(e)用作药物(如疫苗)或诊断剂的组合物；(f)这些组合物在制备以下物质中的用途(1)用于治疗或预防奈瑟球菌引起的感染的药物(2)检测奈瑟球菌或由奈瑟球

菌引起的抗体存在与否的诊断剂，和/或(3)能产生抗奈瑟球菌抗体的药物；和(g)治疗患者的方法，其包括对该患者施用治疗有效量的这些组合物。

序列

本发明还提供了具有以下实施例中所列出的任何的蛋白质或核酸。本发明还提供了具有与这些序列是序列相同性的蛋白质和核酸。如上所述，“序列相同性”的程度最好大于50%(如60%、70%、80%、90%、95%、99%或更大)。

本文的命名

本文参考在 WO 99/24578、WO 99/36544 和 WO 99/57280 中公开的 2166 个蛋白质序列，并将它们编号为如下的 SEQ#:

申请	蛋白质序列	本文的 SEQ#
WO 99/24578	偶 SEQ ID 2-892	SEQ# 1-446
WO 99/36544	偶 SEQ ID 2-90	SEQ#447-491
WO 99/57280	偶 SEQ ID 2-3020	SEQ#492-2001
	偶 SEQ ID 3040-3114	SEQ#2002-2039
	SEQ ID 3115-3241	SEQ#2040-2166

除了这种 SEQ#编号外，本文还使用了 WO 99/24578、WO 99/36544 和 WO 99/27280 中的命名约定(如 WO99/24578 和 WO 99/36544 中用的 ‘ORF4’、‘ORF40’、‘OFR40-1’等；WO 99/57280 中用的 ‘m919’、‘g919’和 ‘a919’等)。

在本文中 Tettelin 等 [Science(2000)287:1809-1815]中的从 NMB0001 到 NMB2160 的 2160 个蛋白质称为 SEQ#2167-4326[参见 WO 00/66791]。

本文采用的术语“本发明的蛋白质”指包含以下的蛋白质：

- (a)SEQ#1-4326 中的一个序列；或
- (b)与 SEQ#1-4326 中的一个序列相同的序列；或
- (c)SEQ#1-4326 中的一个序列的片段。

(b)中的“序列相同性”的程度最好大于50%(如60%、70%、80%、90%、95%、99%或更大)。其包括突变体和等位基因变体[如，参见 WO 00/66741]。相同性最好用 Smith-Waterman 同源性搜寻算法确定，如在 MPSRCH 程序

(Oxford Molecular)中执行的,采用参数“缺口罚分(gap open penalty)”为 12,“缺口延伸罚分(gap extension penalty)”为 1 进行缺口仿射搜索。通常,将两种蛋白质之间 50%或更高的相同性视为功能等效的指示。

(c)中的“片段”应包含 SEQ#1-4326 中一个序列的至少 n 个连续的氨基酸,且根据具体的序列 n 为 7 或更高(如 8、10、12、14、16、18、20、25、30、35、40、50、60、70、80、90、100 或更高)。较佳地,片段包含 SEQ#1-4326 中一个序列的表位。优选的片段是在 WO 00/71574 和 WO 01/04316 中公开的那些。

本发明优选的蛋白质是在脑膜炎奈瑟球菌血清群 B 中发现的。

根据本发明使用的优选蛋白质是血清群 B 脑膜炎奈瑟球菌菌株 2996 或菌株 394/98(新西兰菌株)。除非特别指出,本文所述的蛋白质是脑膜炎奈瑟球菌菌株 2996 的蛋白质。但是,应该理解通常本发明并不受菌株的限制。参考具体的蛋白质(如‘287’、‘919’等)可以包括任何菌株的该蛋白质。

应该理解术语“核酸”包括 DNA 和 RNA,以及它们的类似物,如含有修饰骨架的那些类似物,还包括肽核酸(PNA)等。

附图简述

图 1-26 显示本发明的杂交蛋白。

本发明的进行模式

实验详述

克隆策略和寡核苷酸设计

用以脑膜炎奈瑟球菌 B MC58 的基因组序列为基础设计的寡核苷酸,通过 PCR 扩增编码感兴趣的抗原的基因。除非特别指出,通常将菌株 2996 的基因组 DNA 用作 PCR 反应的模板,将扩增的片段克隆入表达载体 pET21b+(Novagen),从而以 C-末端 His 标记的产物形式表达该蛋白,或将其克隆入 pET-24b+(Novagen)以‘未标记的’形式(如 Δ G287K)表达该蛋白。

不用融合配体和用其自身前导肽(如果存在时)表达蛋白质,进行开放读框(ATG 到终止密码子)的扩增。

当蛋白质以‘未标记的’的形式表达时,通过从预定的前导序列设计 5'-端扩增引物下游除去前导肽。

用于 PCR 的引物的解链温度取决于整个引物中杂交核苷酸的数量和类型，并用以下公式确定：

$$T_{m1} = 4(G+C) + 2(A+T) \quad (\text{除去尾部的})$$

$$T_{m2} = 64.9 + 0.41(\%GC) - 600/N \quad (\text{完整的引物})$$

对整个低聚物而言，所选寡核苷酸的解链温度通常为 65-70℃，仅针对杂交区域，解链温度为 50-60℃。

用 Perkin Elmer 394 DNA/RNA 合成仪合成寡核苷酸，将其从柱上洗脱到 2.0ml NH₄OH 中，在 56℃ 培育 5 小时以去保护。加入 0.3M 乙酸钠和 2 体积的乙醇沉淀寡聚物。将样品离心，将沉淀物重悬浮于水中。

		序列	限制酶切位点
fu (961)-	Fwd	CGCGGATCC -GGAGGGGGTGGTGTCC	BamHI
741(MC58)-His	Rev	CCCGCTCGAG-TTGCTTGGCGGCAAGGC	XhoI
fu (961)-983-His	Fwd	CGCGGATCC - GGCGGAGGCGGCACTT	BamHI
	Rev	CCCGCTCGAG-GAACCGGTAGCCTACG	XhoI

fu (961)- Orf46.1- His	Fwd	CGCGGATCCGGTGGTGGTGGT- TCAGATTTGGCAAACGATTC	BamHI
	Rev	CCC <u>GCTCGAG</u> -CGTATCATATTTACGTGC	XhoI
fu (961 c-L)- 741(MC58)	Fwd	CGCGGATCC -GGAGGGGGTGGTGTCTG	BamHI
	Rev	CCC <u>GCTCGAG</u> -TTATTGCTTGGCGGCAAG	XhoI
fu (961c-L)-983	Fwd	CGCGGATCC - GGCGGAGGCGGCACTT	BamHI
	Rev	CCC <u>GCTCGAG</u> -TCAGAACCGGTAGCCTAC	XhoI
fu (961c-L)- Orf46.1	Fwd	CGCGGATCCGGTGGTGGTGGT- TCAGATTTGGCAAACGATTC	BamHI
	Rev	CCC <u>GCTCGAG</u> -TTACGTATCATATTTACGTGC	XhoI
fu-(ΔG287)-919- His	Fwd	CGCGGATCCGGTGGTGGTGGT- CAAAGCAAGAGCATCCAAACC	BamHI
	Rev	CCC <u>AAGCTT</u> -TTCGGGCGGTATTCGGGCTTC	HindIII
fu-(ΔG287)-953- His	Fwd	CGCGGATCCGGTGGTGGTGGT- GCCACCTACAAAGTGGAC	BamHI
	Rev	GCCC <u>AAGCTT</u> -TTGTTGGCTGCCTCGAT	HindIII
fu-(ΔG287)-961- His	Fwd	CGCGGATCCGGTGGTGGTGGT-ACAAGCGACGACG	BamHI
	Rev	GCCC <u>AAGCTT</u> -CCACTCGTAATTGACGCC	HindIII
fu-(ΔG287)- Orf46.1-His	Fwd	CGCGGATCCGGTGGTGGTGGT- TCAGATTTGGCAAACGATTC	BamHI
	Rev	CCC <u>AAGCTT</u> -CGTATCATATTTACGTGC	HindIII
fu-(ΔG287-919)- Orf46.1-His	Fwd	CCC <u>AAGCTT</u> GGTGGTGGTGGTGGT- TCAGATTTGGCAAACGATTC	HindIII
	Rev	CCC <u>GCTCGAG</u> -CGTATCATATTTACGTGC	XhoI
fu-(ΔG287- Orf46.1)-919-His	Fwd	CCC <u>AAGCTT</u> GGTGGTGGTGGTGGT- CAAAGCAAGAGCATCCAAACC	HindIII
	Rev	CCC <u>GCTCGAG</u> -CGGGCGGTATTCGGGCTT	XhoI
fu ΔG287(394.98)- ...	Fwd	CGCGGATCCGCTAGC-CCCGATGTTAAATCGGC	NheI
	Rev	CGGGGATCC-ATCCTGCTCTTTTTGCCGG	BamHI
fu Orf1-(Orf46.1)- His	Fwd	CGCGGATCCGCTAGC-GGACACACTTATTTCCGGCATC	NheI
	Rev	CGCGGATCC-CCAGCGGTAGCCTAATTTGAT	
fu (Orf1)-Orf46.1- His	Fwd	CGCGGATCCGGTGGTGGTGGT- TCAGATTTGGCAAACGATTC	BamHI
	Rev	CCC <u>AAGCTT</u> -CGTATCATATTTACGTGC	HindIII
fu (919)-Orf46.1- His	Fwd1	GCGGCGT <u>CGAC</u> GGTGGCGGAGGCACTGGATCCTCAG	Sall
	Fwd2	GGAGGCACTGGATCCTCAGATTTGGCAAACGATTC	
	Rev	CCC <u>GCTCGAG</u> -CGTATCATATTTACGTGC	XhoI
Fu (orf46)-287-His	Fwd	CGGGGATCCGGGGGCGGGCGGTGGCG	BamHI
	Rev	CCC <u>AAGCTT</u> ATCCTGCTCTTTTTGCCGGC	HindIII
Fu (orf46)-919-His	Fwd	CGCGGATCCGGTGGTGGTGGTCAAAGCAAGAGCATCCA AACC	BamHI
	Rev	CCC <u>AAGCTT</u> CGGGCGGTATTCGGGCTTC	HindIII

Fu (orf46-919)-287-His	Fwd	CCCCAAGCTTGGGGGCGGCGGTGGCG	HindIII
	Rev	CCCGCTCGAGATCCTGCTCTTTTTGCCGGC	XhoI
Fu (orf46-287)-919-His	Fwd	CCCCAAGCTTGGTGGTGGTGGTGGTCAAAGCAAGAGCAT CCAAACC	HindIII
	Rev	CCCGCTCGAGCGGGCGGTATTTCGGGCTT	XhoI
(ΔG741)-961c-His	Fwd1	GGAGGCACTGGATCCGCAGCCACAAACGACGACGA	XhoI
	Fwd2	GCGGCCTCGAG-GGTGGCGGAGGCACTGGATCCGCAG	
	Rev	CCCGCTCGAG-ACCCAGCTTGTAAAGGTTG	XhoI
(ΔG741)-961-His	Fwd1	GGAGGCACTGGATCCGCAGCCACAAACGACGACGA	XhoI
	Fwd2	GCGGCCTCGAG-GGTGGCGGAGGCACTGGATCCGCAG	
	Rev	CCCGCTCGAG-CCACTCGTAATTGACGCC	XhoI
(ΔG741)-983-His	Fwd	GCGGCCTCGAG- GGATCCGGCGGAGGCGGCACITCTGCG	XhoI
	Rev	CCCGCTCGAG-GAACCGGTAGCCTACG	XhoI
(ΔG741)-orf46.1-His	Fwd1	GGAGGCACTGGATCCTCAGATTTGGCAAACGATTC	Sall
	Fwd2	GCGGCGTTCGACGGTGGCGGAGGCACTGGATCCTCAGA	
	Rev	CCCGCTCGAG-CGTATCATATTTACGTGC	XhoI
(ΔG983)-741(MC58)-His	Fwd	GCGGCCTCGAG-GGATCCGGAGGGGGTGGTGTCCGC	XhoI
	Rev	CCCGCTCGAG-TTGCTTGGCGGCAAG	XhoI
(ΔG983)-961c-His	Fwd1	GGAGGCACTGGATCCGCAGCCACAAACGACGACGA	XhoI
	Fwd2	GCGGCCTCGAG-GGTGGCGGAGGCACTGGATCCGCAG	
	Rev	CCCGCTCGAG-ACCCAGCTTGTAAAGGTTG	XhoI
(ΔG983)-961-His	Fwd1	GGAGGCACTGGATCCGCAGCCACAAACGACGACGA	XhoI
	Fwd2	GCGGCCTCGAG-GGTGGCGGAGGCACTGGATCCGCAG	
	Rev	CCCGCTCGAG-CCACTCGTAATTGACGCC	XhoI
(ΔG983)-Orf46.1-His	Fwd1	GGAGGCACTGGATCCTCAGATTTGGCAAACGATTC	Sall
	Fwd2	GCGGCGTTCGACGGTGGCGGAGGCACTGGATCCTCAGA	
	Rev	CCCGCTCGAG-CGTATCATATTTACGTGC	XhoI

*将该引物用作反向引物，将所有 287 的 C 末端融合于 His-标记。

§ 与 287-His 反向引物联用的正向引物。

NB-所有 PCR 反应使用菌株 2996，除非特别指出(如菌株 MC58)。

在所有以 ATG 起始不跟随唯一 NheI 位点的构建物中，ATG 密码子是用于克隆的 NdeI 位点的一部分。在 5'端用 NheI 作为克隆位点制备的构建物(如所有那些在 N-末端包含 287 的)另有两个融合于抗原的编码序列的密码子(GCT AGC)。

染色体 DNA 模板的制备

在 100ml GC 培养基中使脑膜炎奈瑟球菌前株 2996、MC58、394.98、1000 和 BZ232(及其它)生长至指数期，离心收获，并重悬浮于 5ml 缓冲液(20% w/v 蔗糖、50mM Tris-HCl、50mM EDTA、pH 8)中。在冰上培育 10 分钟后，加入

10ml 裂解液(50mM NaCl、1%十二烷基肌氨酸钠(Na-Sarkosyl)、50 μ g/ml 蛋白酶 K)裂解细菌，悬浮液在 37 $^{\circ}$ C 培育 2 小时。进行 2 次苯酚提取(平衡至 pH8)和一次 CHCl₃/异戊醇(24: 1)提取。加入 0.3M 乙酸钠和 2 体积乙醇沉降 DNA，并离心收集。用 70%(v/v)乙醇洗涤沉淀物 1 次，并重溶解于 4.0ml TE 缓冲液(10mM Tris-HCl、1mM EDTA、pH 8.0)。读取 OD₂₆₀ 测定 DNA 浓度。

PCR 扩增

标准的 PCR 过程：进行如下在 40 μ M 各寡核苷酸引物、400-800 μ M dNTP 溶液、1x PCR 缓冲液(包含 1.5mM MgCl₂)、2.5 单位 TaqI DNA 聚合酶(用 Perkin-Elmer AmpliTaq, Boehringer Mannheim ExpandTM 长模板)存在时，将 200ng 2996、MC581000、或 BZ232 菌株的基因组 DNA 或 10ng 重组克隆的质粒 DNA 制备物用作模板。

将整个混合物在 95 $^{\circ}$ C 初步培育 3 分钟后，每份样品进行 2-步的扩增：用除去引物(T_{m1})的限制酶尾部的杂交温度进行前 5 轮。然后按全长低聚物(T_{m2})计算的杂交温度进行 30 轮。根据要扩增的 Orf 的长度，在 68 $^{\circ}$ C 或 72 $^{\circ}$ C 进行的延伸时间各不相同。对 Orf1 而言，自 3 分钟开始的延伸时间每轮递增 15 秒。以在 72 $^{\circ}$ C 10 分钟延伸步骤完成循环。

将扩增的 DNA 直接加到 1%琼脂糖凝胶上。按制造商的说明，用 Qiagen 凝胶提取试剂盒纯化相应于正确大小条带的 DNA 片段。

PCR 片段和克隆载体的消化

用适合的限制酶消化相应于扩增片段的纯化 DNA，从而克隆入 pET-21b+、pET22b+、或 pET-24b+。用 QIAquick PCR 纯化试剂盒(按制造商的说明)纯化消化的片段，用 H₂O 或 10mM Tris(pH 8.5)洗脱。用适合的限制酶消化质粒载体，加到 1.0%琼脂糖凝胶上，用 Qiagen AIQquick 凝胶提取试剂盒纯化相应于消化的载体的条带。

克隆

将预先消化和纯化的、相应于各基因的片段连接到 pET21b+、pET22b+或

pET-24b+中。将在连接缓冲液(由制造商提供)中的 T4 DNA 连接酶用于摩尔比为 3: 1 的片段/载体。

通过在冰上培育连接酶反应溶液和细菌 40 分钟,然后在 37°C 培育 3 分钟,将重组质粒转化入感受态大肠杆菌 DH5 或 HB101 中。

然后添加 800 μ l LB 肉汤,并在 37°C 培育 20 分钟。在 Eppendorf 微型离心机中以最高速度离心这些细胞,并重悬浮于约 200 μ l 的上清液中,并涂布在 LB 氨苄青霉素(100mg/ml)琼脂上。

在 4.0ml LB 肉汤+100 μ g/ml 氨苄青霉素中培育随机选择的集落过夜,进行重组集落的筛选。使细胞沉淀,并按制造商的说明用 Qiagen QIAprep Spin Miniprep 试剂盒提取质粒 DNA。用适合的限制酶消化约 1 μ g 的各微型制备物,将消化物加到 1-1.5% 琼脂糖凝胶(取决于预计的插入大小)与分子量标记(1kb DNA Ladder, GIBCO)平行。根据插入的大小选择阳性克隆。

表达

各基因克隆入表达载体后,将重组质粒转化入适合表达重组蛋白的大肠杆菌菌株中。如上所述,用 1 μ l 各构建物转化大肠杆菌 BL21-DE3。将单重组集落接种入 2ml LB+Amp(100 μ g/ml)中,在 37°C 培育过夜,然后在 100ml 烧瓶中用 20ml LB+Amp(100 μ g/ml)以 1: 30 稀释,使 OD₆₀₀ 在 0.1-0.2 之间。将烧瓶置于旋转式水浴摇床中于 30°C 或 37°C 培育,直到 OD₆₀₀ 显示适合诱导表达的指数生长期(0.4-0.8OD)。加入 1.0mM IPTG 诱导蛋白质表达。在 30°C 或 37°C 培育 3 小时后,测定 OD₆₀₀ 并检测表达。用微型离心机离心 1.0ml 各样品,将沉淀物重悬浮于 PBS 中,用 SDS-PAGE 和考马斯蓝染色分析。

His-标记的蛋白质的纯化

从菌株 2996 和 MC58 中克隆了 287 的各种形式。用 C-末端 His 标记的融合体进行构建,其包括成熟形式(aa18-427)、含缺失(Δ 1、 Δ 2、 Δ 3 和 Δ 4)的构建物和由 B 或 C 结构域组成的克隆。对以 His-融合体纯化的各克隆而言,划线接种单集落,并且在 37°C LB/Amp(100 μ g/ml)琼脂板上培育过夜。将从该平板上分离的集落接种到 20ml LB/Amp(100 μ g/ml)的液体培养基中,并在 37°C 振荡

生长过夜。以 1:30 将过夜培养物稀释到 1.0L LB/Amp(100 μ g/ml)液体培养基中, 让其在最佳温度(30 或 37 $^{\circ}$ C)生长, 直到 OD₅₅₀ 达到 0.6-0.8。添加 IPTG(终浓度为 1.0mM)诱导重组蛋白的表达, 再培育培养物 3 小时。于 4 $^{\circ}$ C, 以 8000g 离心 15 分钟收获细菌。将细菌沉淀物重悬浮于 7.5ml(i)冷缓冲液 A(300mM NaCl、50mM 磷酸盐缓冲液、10mM 咪唑、pH 8.0), 用于可溶性蛋白质; 或(ii)缓冲液 B(10mM Tris-HCl、100mM 磷酸盐缓冲液, pH 8.8 和任选地 8M 尿素), 用于不溶性蛋白质。以可溶形式纯化的蛋白质包括 287-His、 Δ 1、 Δ 2、 Δ 3 和 Δ 4287-His、 Δ 4287MC58-His、287c-His 和 287cMC58-His。蛋白质 287bMC58-His 是不溶的并相应地纯化。用 Branson Sonifier 450, 在冰上以 40W、30 秒超声处理破坏细胞 4 次, 并于 4 $^{\circ}$ C 以 13000xg 离心 30 分钟。对于不溶蛋白质, 将沉淀物重悬浮于 2.0ml 缓冲液 C(6M 盐酸胍、100mM 磷酸盐缓冲液、10mM Tris-HCl、pH 7.5), 并用 Dounce 匀浆器处理 10 次。以 13000g 离心匀浆 30 分钟并保留上清液。将可溶的和不可溶制备物的上清液与 150 μ l Ni²⁺-树脂(先用缓冲液 A 或缓冲液 B 平衡)混合, 并在室温温和振荡培育 30 分钟。树脂是按制造商说明制备的 Chelating Sepharose Fast Flow(Pharmacia)。于 4 $^{\circ}$ C, 以 700g 离心分批制备物 5 分钟, 弃去上清液。用 10ml 缓冲液 A 或 B 洗涤树脂 2 次(分批)10 分钟, 重悬浮于 1.0ml 缓冲液 A 或 B 中, 加到一次性柱上。用(i)缓冲液 A(4 $^{\circ}$ C)或(ii)缓冲液 B(室温)持续洗涤树脂, 直到流出物的 OD₂₈₀ 达到 0.02-0.01。再用(i)冷缓冲液 C(300mM NaCl、50mM 磷酸盐缓冲液、20mM 咪唑、pH 8.0)或(ii)缓冲液 D(10mM Tris-HCl、100mM 磷酸盐缓冲液、pH 6.3 和任选地 8M 尿素)进一步洗涤树脂, 直到流出物的 OD₂₈₀ 达到 0.02-0.01。加入 700 μ l(i)冷洗脱缓冲液 A(300mM NaCl、50mM 磷酸盐缓冲液、250mM 咪唑、pH 8.0)或(ii)洗脱缓冲液 B(10mM Tris-HCl、100mM 磷酸盐缓冲液、pH 4.5 和任选地 8M 尿素)洗脱 His-融合蛋白, 收集组分直到 OD₂₈₀ 显示获得了所有重组蛋白。用 SDS-PAGE 分析 20 μ l 量的各洗脱组分。用 Bradford 试验法计算蛋白质浓度。

变性的 His-融合蛋白的复性

需要变性以稳定 287bMC8, 因此在免疫接种前需进行复性步骤。将甘油加到上述得到的变性组分中, 使终浓度为 10% v/v。用透析缓冲液 I(10% v/v 甘油,

0.5M 精氨酸、50mM 磷酸盐缓冲液、5.0mM 还原的谷胱甘肽、0.5mM 氧化的谷胱甘肽、2.0M 尿素, pH 8.8)将蛋白质稀释至 200 μ g/ml, 用相同的缓冲液在 4 $^{\circ}$ C 透析 12-14 小时。于 4 $^{\circ}$ C, 用缓冲液 II(10% v/v 甘油, 0.5M 精氨酸、50mM 磷酸盐缓冲液、5.0mM 还原的谷胱甘肽、0.5mM 氧化的谷胱甘肽、pH 8.8)再进行透析 12-14 小时。用以下公式计算蛋白质的浓度:

$$\text{蛋白质(mg/ml)} = (1.55 \times \text{OD}_{280}) - (0.76 \times \text{OD}_{260})$$

免疫接种

在第 0、21 和 35 天, 用抗原免疫接种 Balb/C 小鼠, 在第 49 天分析血清。

血清分析-ELISA

将不包囊的 MenB M7 和包囊的菌株置于巧克力琼脂板上, 在 37 $^{\circ}$ C、5% CO₂ 中培育过夜。用无菌 Dracon 刷从琼脂板上收集细菌菌落, 并接种到含有 0.25% 葡萄糖的 Mueller-Hinton 肉汤(Difco)中。每 30 分钟监测一次细菌的生长, 随后测定 OD₆₂₀。让细菌生长直到 OD 达到 0.4-0.5。将培养物以 4000rpm 离心 10 分钟。弃去上清液, 用 PBS 洗涤细菌 2 次, 重悬浮于含有 0.025% 甲醛的 PBS 中, 并在 37 $^{\circ}$ C 培育 1 小时, 然后在 4 $^{\circ}$ C 搅拌培育过夜。在 96 孔 Greiner 板的各孔中添加 100 μ l 细菌细胞, 并在 4 $^{\circ}$ C 培育过夜。然后用 PBT 洗涤缓冲液(0.1% Tween-20 的 PBS 溶液)冲洗这些孔 3 次。在各孔中加入 200 μ l 饱和缓冲液(2.7% 聚乙烯吡咯烷酮 10 的水溶液), 将这些平板在 37 $^{\circ}$ C 培育 2 小时。用 PBT 冲洗这些孔 3 次。在各孔中加入 200 μ l 稀释的血清(稀释缓冲液: 1% BSA、0.1% Tween-20、0.1% NaN₃ 的 PBS 溶液), 将这些平板在 37 $^{\circ}$ C 培育 2 小时。用 PBT 冲洗这些孔 3 次。在各孔中加入 100 μ l HRP-缀合的兔抗-小鼠(Dako)血清(用稀释缓冲液以 1: 2000 稀释), 将这些平板置于 37 $^{\circ}$ C 培育 90 分钟。用 PBT 缓冲液洗涤这些孔 3 次。在各孔中加入 100 μ l HRP 的底物缓冲液(25ml 柠檬酸缓冲液 pH 5, 10mg 邻-苯二胺(phenildiamine)和 10 μ l H₂O₂), 并将平板在室温中静置 20 分钟。在各孔中加入 100 μ l 12.5% H₂SO₄, 随后测定 OD₄₉₀。计算 ELISA 滴定度, 即高于预先免疫血清稀释度的、OD₄₉₀ 值为 0.4 的血清稀释度。当 OD₄₉₀ 值为 0.4 的血清稀释度高于 1: 400 时, 将 ELISA 视为阳性。

血清分析-FACS 扫描细菌结合分析

将不包囊的 MenB M7 菌株置于巧克力琼脂板上，在 37℃、5%CO₂ 中培育过夜。用无菌 Dracon 刷从琼脂板上收集菌落，并接种到 4 支装有 0.25% 葡萄糖的 8ml Mueller-Hinton 肉汤(Difco)试管中。每 30 分钟监测细菌的生长，然后测定 OD₆₂₀。让细菌生长直到 OD 达到 0.35-0.5。将培养物以 4000rpm 离心 10 分钟。弃去上清液，用封阻缓冲液(1% BSA 的 PBS 溶液，0.4% NaN₃)重悬浮沉淀物，并以 4000rpm 离心 5 分钟。将细胞重悬浮于封阻缓冲液，使 OD₆₂₀ 达到 0.05。在 96 孔 Costar 板的各孔中添加 100μl 细菌细胞。在各孔中加入 100μl 稀释的 (1:100、1: 200、1: 400)血清(以封阻缓冲液配制)，并将这些平板在 4℃ 培育 2 小时。以 4000rpm 离心细胞 5 分钟，吸出上清液，在各孔中加入 200μl/孔封阻缓冲液洗少涤细胞。在各孔中加入 100μl R-Phicoerytrin 缀合的 F(ab)₂ 山羊抗-小鼠(以 1: 100 稀释)，将平板置于 4℃ 培育 1 小时。以 4000rpm 离心 5 分钟以沉淀细胞，添加 200μl/孔封阻缓冲液洗涤细胞。吸出上清液，细胞重悬浮于 200μl/孔 PBS、0.25% 甲醛。将样品转移到 FACScan 管中并读数。FACScan(Laser Power 15mW)的条件设定为：FL2 开；FSC-H 阈值：92；FSC PMT 电压：E01；SSC PMT:474；Amp. Gains 6.1；FL-2 PMT:586；补偿值：0。

血清分析-杀菌试验

于 37℃，在 5% CO₂、在巧克力琼脂板(以冷冻原液起始)上培育脑膜炎奈瑟球菌菌株 2996 过夜。收集菌落，并将其接种到 7ml 含有 0.25% 葡萄糖的 Mueller-Hinton 肉汤中，使 OD₆₂₀ 达到 0.05-0.08。在 37℃ 振荡培育约 1.5 小时，直到 OD₆₂₀ 达到 0.23-0.24。用 50mM 磷酸盐缓冲液(pH7.2，含 10mM MgCl₂、10mM CaCl₂ 和 0.5%(w/v)BSA(分析缓冲液))以 10⁵CFU/ml 工作稀释度稀释细菌。最终反应混合物的总体积为 50μl，其中 25μl 连续 2 倍稀释的测试血清，12.5μl 工作稀释度的细菌，12.5μl 幼兔补体(终浓度 25%)。

对照包括：用补体血清培育的细菌、用细菌培育并在 56℃ 加热 30 分钟补充灭活的免疫血清。在加入幼兔补体后，用斜置方法立即将 10μl 对照置于 Mueller-Hinton 琼脂板上(0 时间)。37℃ 旋转培育 96 孔板 1 小时。将每份样品

的 7 μ l 涂布在 Mueller-Hinton 琼脂板上作为斑点，而用斜置方法将 10 μ l 对照涂布在 Mueller-Hinton 琼脂板上(1 时间)。37 $^{\circ}$ C 培育琼脂板 18 小时，计算相应于 0 时间和 1 时间的菌落数量。

血清分析-Western 印迹法

将纯化的蛋白质(500ng/泳道)、外膜小泡(5 μ g)和 MenB 菌株 2996 衍生的全细胞提取物(25 μ g)加到 12% SDS-聚丙烯酰胺凝胶上，并转移到硝基纤维素膜上。用转化缓冲液(0.3% Tris 碱，1.14% 甘氨酸、20%(v/v)甲醇)在 4 $^{\circ}$ C、150mA 进行转化 2 小时。于 4 $^{\circ}$ C，在饱和缓冲液(10%脱脂乳、0.1% Triton X100 的 PBS 溶液)中培育过夜使膜饱和。用冲洗缓冲液(3%脱脂乳、0.1% Triton X100 的 PBS 溶液)冲洗膜 2 次，并在 37 $^{\circ}$ C 与用洗涤缓冲液以 1: 200 稀释的小鼠血清一起培育 2 小时。冲洗膜 2 次，与 1: 2000 稀释的辣根过氧化物酶标记的抗-小鼠 Ig 一起培育 90 分钟。用 0.1% Triton X100 的 PBS 溶液冲洗膜 2 次，用 Opti-4CN 底物试剂盒(Bio-Rad)显色。加入水终止反应。

如下制备 OMV: 于 37 $^{\circ}$ C、5% CO₂，在 5 个 GC 板上让脑膜炎奈瑟球菌 2996 生长过夜，用接种环收获，并重悬浮于 10ml 20mM Tris-HCl pH 7.5、2mM EDTA 中。在 56 $^{\circ}$ C 热灭活 45 分钟，在冰上超声处理破碎细胞 5 分钟(50%负载循环(duty cycle)、50%输出，Branson 超声仪 3mm 微型针头)。以 5000g 离心 10 分钟除去未被破坏的细胞，回收含有完整细胞包膜组分的上清液，于 4 $^{\circ}$ C 以 50000g 进一步离心过夜。将含有膜的沉淀物重悬浮于 2%二烷基肌氨酸钠、20mM Tris-HCl pH 7.5、2mM EDTA 中，在室温培育 20 分钟以溶解内膜。以 10000g 离心上清液 10 分钟除去聚集体，以 50000g 进一步离心上清液 3 小时。用 PBS 冲洗含有外膜的沉淀物，并重悬浮于相同的缓冲液中。用 BSA 作为标准，由 D.C.Bio-Rad 蛋白分析(改进的 Lowry 方法)测定蛋白质的浓度。

如下制备全细胞提取物: 使脑膜炎奈瑟球菌在 GC 板上培育过夜，用接种环收获，并重悬浮于 1ml 20mM Tris-HCl 中。在 56 $^{\circ}$ C 热灭活 30 分钟。

实施例 1-ORF46 的杂交体

脑膜炎奈瑟球菌(血清群 B, 菌株 2996)的完整 ORF46 蛋白质具有如下序列:

1 LGISRKISLI LSILAVCLPM HAHASDLAND SFIRQVLDRO HFEFDPGKYHL
 51 FGSRGELAER SGHIGLGKIQ SHQLGNLMIQ QAAIKGNIGY IVRFSDHGHE
 101 VHSPFDNHAS HSDSDEAGSP VDGFSLYRIH WDGYEHHPAD GYDGPQGGGY

151 PAPKGARDIY SYDIKGV AQN IRLNLTDNRS TGQRLADRFH NAGSMLTQGV
 201 GDGFKRATRY SPELDRSGNA AEA FNGTADI VKNIIGAAGE IVGAGDAVQG
 251 ISEGSNIAMV HGLGLLSTEN KMARINDLAD MAQLKDYAAA AIRDWAVQNP
 301 NAAQGI EAVS NIFMAA IPIK GIGAVRGKYG LGGITAHPIK RSOMGAIALP
 351 KGKSAVSDNF ADAAYAKYPS PYHSRNIRSN LEQRYGKENI TSSTVPPSNG
 401 KNVKLADQRH PKTGVPFDGK GFPNFEKHVK YDTKLDIQEL SGGGIPKAKP
 451 VSDAKPRWEV DRKLNKLTTR EQVEKNVQEI RINGNKNSNFS QHAQLEREIN
 501 KLKSADEINF ADGMGKFTDS MNDKAFSRLV KSVKENGFTN PVVEYVEING
 551 KAYIVRGNNR VFAEYLGRI HELKFKKVDV PVPNTSWKNP TDVNLNESGNV
 601 KRPRYRSK*

在前导肽下加下划线。

可以在 WO 00/66741 中发现其它菌株的 ORF46 的序列。

ORF46 在其 C-末端和 N-末端与 287、919 和 ORF1 融合。该杂交蛋白通常是不溶解的，但产生一些良好的 ELISA 和杀菌结果(针对同源 2996 菌株)：

蛋白质	ELISA	杀菌 Ab
Orf1-Orf46.1-His	850	256
919-Orf46.1-His	12900	512
919-287-Orf46-His	n.d.	n.d.
Orf46.1-287His	150	8192
Orf46.1-919His	2800	2048
Orf46.1-287-919His	3200	16384

为了比较，构建了 ORF46.1、287(以 GST 融合体或 Δ G287 的形式)和 919 的“三”杂交体，并针对各种菌株(包括同源 2996 菌株)将其与三种抗原的简单混合物相比。FCA 用作佐剂：

	2996	BZ232	MC58	NGH38	F6124	BZ133
混合物	8192	256	512	1024	>2048	>2048
ORF46.1-287-919his	16384	256	4096	8192	8192	8192
ΔG287-919-ORF46.1his	8192	64	4096	8192	8192	16384
ΔG287-ORF46.1-919his	4096	128	256	8192	512	1024

同样，这些杂交体显示相当等的或更佳的免疫活性。

针对各种异源菌株，将两种蛋白质(菌株 2996)的杂交体与单种蛋白质相比：

	1000	MC58	F6124 (MenA)
ORF46.1-His	<4	4096	<4
ORF1-His	8	256	128
ORF1—ORF46.1-His	1024	512	1024

再次，这些杂交体显示相等的或更佳的免疫活性。

实施例 2- Δ G287 的杂交体

发现 287 中(Gly)₆ 序列的缺失对蛋白质的表达有显著影响。将缺失 N-末端氨基酸多达 GGGGGG 的蛋白质称为 ' Δ G287'。在菌株 MC58 中，它的基本序列(前导肽有下划线)为：

```

SPDVKS ADTLSKPAAP VVSEKETEAK EDAPQAGSQ QGAPSAQGSQ DMAAVSEENT
GNGGAVTADN PKNEDEVAQN DMPQNAAGTD SSTPNHTPDP NMLAGNMENQ ATDAGESSQP
ANQPDMANAA DGMQGGDDPSA GGQNAGNTAA QGANQAGNNQ AAGSSDPIPA SNPAPANGGS
NFGRRVLANG VLIDGFSQNI TLTHCKGDSC SGNNFLDEEV QLKSEFEKLS DADKISNYKK
DGKNDKFGVL VADSVQMKGI NQYIIFYKPK PTFARFRRS ARSRRLPAE MPLIPVNQAD
TLIVDGEAVS LTGHSGNIFA PEGNYRYLTY GAEKLPGGSY ALRVQGEPAK GEMLAGAAVY
NGEVLHFHTE NGRFYPTRGR FAAKVDGSK SVDGIIDSGD DLHMGTKQFK AAIDGNGFKG
TWTENGSGDV SGKFGYPAGE EVAGKYSYRP TDAEKGFGV FAGKKEQD*

```

与 ' Δ 287-His' 或 ' Δ 287^{未标记的}' 相比，有或无 His-标记的 Δ G287(分别为 ' Δ G287-His' 和 ' Δ G287K') 以很好的水平表达。

在基因变异性数据的基础上，从许多 MenB 菌株(尤其是从菌株 2996、MC58、100 和 BZ232)的大肠杆菌中表达 Δ G287-His 的变体。结果也好-它们都具有很高的 ELISA 滴定度，且血清杀菌滴定度 > 8192。由 pET-24b 表达的 Δ G287K 在 ELISA 和血清杀菌实验中有极佳滴定度。

还将聚-Gly 序列的缺失应用于 Tbp2(NMB0460)、741(NMB1870) 和 983(NMB1969)。在不编码其前导肽的序列且没有聚-Gly(即以 " Δ G 形式")，在 pET 载体中克隆并在大肠杆菌中表达时，观察到相同的作用-在携带聚-甘氨酸段缺失的克隆中表达很好，若在表达的蛋白中存在甘氨酸时则表达差或不表达。

将 Δ G287 直接融合于 919、953、961(如下所示的序列)和 ORF46.1 的符合读框的上游:

 Δ G287-919

```
ATGGCTAGCCCCGATGTTAAATCGGGCGGACACGCTGTCAAACCAGCGCCCTCCTGTTGTTGCTGAAAAAGAGACAGAG
GTAAAAGAAGATGCGCCACAGGCAGGTTCTCAAGGACAGGGCGCGCCATCCACACAAGGCAGCCAAGATATGGCGGCA
GTTTCGGCAGAAAAACAGGCAATGCGCGTTCGGCAACAACGGACAAACCCAAAAATGAAGACGAGGGACCGCAAAAT
GATATGCCCGCAAAATCCGCGCAATCCGCAAAATCAAACAGGGAACAACAACCCCGCGATTCTTCAGATTCCGCCCCC
CGGTCAAACCCCTGCACCTGCGAATGGCGGTAGCAATTTTGAAGGGTTGATTGGCTAATGGCGTTTGGATTGATGGG
CCGTGCAAAATATAACGTTGACCCACTGTAAAGGCGATTCTTGTAAATGGTGATAATTATTTGGATGAAGAAGCACCG
TCAAAATCAGAATTTGAAAAATTTAAATGAGTCTGAACGAATTGAGAAATATAAGAAAGATGGGAAAAGCGATAAATTT
ACTAATTTGGTTGCGACAGCAGTTCAAGCTAATGGAACTAACAAATATGTCATCATTTATAAAGACAAGTCCGCTTCA
TCTTCATCTGCGCGATTTCAGGCGTTCTGCACGGTTCGAGGAGGTGCTTCCTGCGGAGATGCCGCTAATCCCCGTCAAT
CAGGCGGATACGCTGATTTGTCGATGGGGAAGCGGTCAAGCTGACGGGGCATTCGCGCAATATCTTCGCGCCCGAAGGG
AATFACCGGTATCTGACTTACGGGGCGGAAAAATGCCCCGGCGATCGTATGCCCTCCGTGTGCAAGGCGAACCGGCA
AAAGGCGAAATGCTTGTGCTGGCACGGCCGTGTACAACGGCGAAGTGTGTCATTTTCATACGGAAAACGGCCGTCCGTAC
CCGACTAGAGGCAGGTTTGCCTGCAAAAGTCGATTTCCGCGAGCAAAATCTGTGGACGGCATTATCGACAGCGGCGATGAT
TTGCATATGGGTACGCAAAATTTCAAAGCCGCCATCGATGGAAACGGCTTTAAGGGGACTTGGACGGAAAAATGGCGGC
GGGATGTTTCCGGAAGGTTTTACGGCCCGGCCGGCGAGGAAGTGGCGGAAAAATACAGCTATCGCCCGACAGATGCG
```

GAAAAGGGCGGATTCGGCGTGTTCGGCGGCAAAAAAGAGCAGGATGGATCCGGAGGAGGAGGATGCCAAAGCAAGAGC
 ATCCAAACCTTTCGCAACCCGACACATCCGTCATCAACGGCCCGGACCGGCGGTCCGGCATCCCCGACCCCGCCGGA
 ACAGCGGTCCGGCGGCGGGCCGCTCTATACCGTGTACCGCACCTGTCCCTGCCCACTGGGCGGCGCAGGATTC
 GCCAAAAGCCTGCAATCTTCGCGCTCGGCTGCGCCAAATTTGAAAAACCGCCAAGGCTGGCAGGATGTGTGCGCCAA
 GCCTTTCAAAACCCCGTCCATTCCTTTCAGGCAAAACAGTTTTTTTGAACGCTATTTTCACGCCGTGGCAGGTTGCAGGC
 AACGGAAGCCTTTCGCGGTACGGTTACCGGTATTCAGAGCCGGTGTGAAGGGCGACGACAGGCGGACGGCACAAGCC
 CGCTTCCCGATTTACGGTATTCGCGACGATTTATCTCCGTCGCCCTGCCTGCCGGTTTCGCGGAGCGGAAAAGCCCTT
 GTCGCGATCAGGCAGACGGGAAAAACAGCGGCACAATCGACAATACCGCGGCACACATACCGCGGACCTCTCCCGA
 TTCCCATCACCGCGGCACAACGGCAATCAAAGGCAGGTTTGAAGGAAGCCGCTTCTCCCTACCACACGCGCAAC
 CAAATCAACGGCGGCGCGCTTGACGGCAAAGCCCCGATCTCGGTTACGCCGAAGACCCCGTTCGAATTTTTTTTATG
 CACATCCAAGGCTCGGGCCCTTGAAAACCCCGTCCGGCAAATACATCCGCATCGGCTATGCCGACAAAACGAACAT
 CCTACGTTTCCATCGGACGCTATATGGCGGACAAAGGCTACCTCAAGCTCGGGCAGACCTCGATGCAGGGCATCAA
 GCCTATATGCCGCAAAATCCGCAACGCCCTCGCCGAAGTTTTGGGTCAAACCCAGCTATATCTTTTTCCGCGAGCTT
 GCCGGAAGCAGCAATGACGGTCCCGTCCGCGCACTGGGCGCCGTTGATGGGGGAATATGCCGCGCAGTCCGACCGG
 CACTACATTACCTTGGGCGCGCCCTTATTTGTCCGACCCGCCATCCGGTTACCCGCAAAGCCCTCAACCGCCTGAT
 ATGGCGCAGGATACCGCGAGCGGATTAAGGCGCGGTGCGCGTGGATTATTTTTGGGGATACGGCGACGAAGCCGGC
 GAACTTTCGGCGCAAACAGAAAACACGGGTACGCTCGCAGCTCCTACCCAACGGTATGAAGCCCGAATACCGCCCG
 TAACTCGAG

1 MASPDVKSAD TLSKPAAPV AVAEKETEVEKED APQAGSQGQ APSTQGSQDM
 51 AAVSAENTGN GGAATTDKPK NEDEGPQNDM PQNSAESANQ TGNNQPADSS
 101 DSAPASNPA ANGGSNFGRV DLANGVLIDG PSQNTITLTHC KGDSCNGDNL
 151 LDEEAPSKSE FENLNESERI EKYYKDKGSD KFTNLVATAV QANGTNKYVI
 201 IYDKSASSS SARFRRSARS RRSPLAEMPL IPVNQADTLI VDGEAVSLTG
 251 HSGNIFAPEG NYRYLTYGAE KLPGGSYALR VQGEPAKSEM LAGTAVYNGE
 301 VLHFHTENGR PYPTRGRFAA KVDGFSKSD GIIDSGDDLH MGTQKFKAAI
 351 DGNFGKGTWT ENGGDVSGR FYGPAGEEVA GKYSYRPTDA EKGGFVGFAG
 401 KKEQDSSGGG GCQSKSIQTF PQPDTSVING PDRFVGIPTD AGTTVGGGGA
 451 VYTVVPHLSL PHWAAQDFAK SLQSFRLGCA NLKNRQWQD VCAQAFQTFV
 501 HSFQAKQFFE RYFTFWQVAG NGSLAGTVTG YYEPVLKGD RRTAQARFPI
 551 YGIPDDFISV PLPAGLRSGK ALVRIROTGK NSGTIDNTGG THTADLSRFP
 601 ITARTTAIKG RFEGSRFLPY HTRNQINGGA LDGKAPILGY AEDFVELFEM
 651 HIQSGRLKT PSGKYRIGY ADKNEHPYVS IGRYMADKGY LKLGQTSMQG
 701 IKAYMRQNPQ RLAEVLQNP SYIFFRELQAG SSNDGPFVGL GTFPLMGEYAG
 751 AVDRHYITLG APLFVATAHP VTRKALNRLI MAQDTGSAIK GAVRVDYFWG
 801 YGDEAGELAG KQRTTGYVWQ LLPNGMKPEY RP*

AG287-953

ATGGCTAGCCCCGATGTAAATCGCGGACACGCTGTCAAACCGGCGCTCCTGTTGTTGCTGAAAAAGAGACAGAG
 GTAAAAGAAGATGCCACAGGCAGGTTCTCAAGGACAGGGCGCCATCCACACAAGGCAGCCAAGATATGGCGGCA
 GTTTCGGCAGAAAATACAGGCAATGGCGGTGCGGCAACAACGGACAAACCCAAAATGAAGACGAGGGACCGCAAAAT
 GATATGCCGCAAAATTCGCGCAATTCGCAAAATCAAACAGGGAACAACCAACCCCGCGATTCCTTCAGATTCGCGCCCC
 CGCTCAAACCTGCACCTGCGAATGGCGGTAGCAATTTGGAAGGGTTGATTTGGCTAATGGCGTTTTGATGATGGG
 CCGTCGCAAAATATAAGTTGACCCACTGTAAAGCGGATTTCTGTAATGGTGAATTTATTTGATGAAGAACCG
 TCAAATCAGAAATTCGAAAATTTAAATGAGTCTGAACGAATTTGAGAAATATAAGAAAGATGGGAAAAGCGATAAATTT
 ACTAATTTGGTTGCGACAGCAGTTCAAGCTAATGGAATAACAAATATGTTCATCATTATAAAGACAAGTCCGCTTCA
 TCTTCATCTGCGCGATTTCAGGCGTCTGACCGGTGAGGAGGTGCGTTCCTGCGGAGATGCCCTAATCCCGTCAAT
 CAGGCGGATACGCTGATTTGTCGATGGGGAAGCGGTGACGCTGACGGGGCATTCGGCAATATCTTCGCGCCGAAGG
 AATTACCGGTATCTGACTTACGGGGCGGAAAAATTCGCGCGGATCGTATGCCCTCCGTTGCAAGGCGAACCCGCA
 AAAGGCGAAATGCTTGTGGCACGGCGGTGTAACAACGGCAAGTCTGCAATTTTCATACGGAAAACGGCGTCCGTAC
 CCGACTAGAGGCAGGTTTCGCGCAAAAGTTCGATTTCCGCGAGCAAAATCTGTGGACGGCATTATCGACAGCGCGGATGAT
 TTGCATATGGGTACGCAAAAATTCAAAGCCGCATCGATGGAACCGGCTTTAAGGGGACTTGGACGGAAAATGGCGGC
 GGGGATGTTTCCGGAAGGTTTACGGCCCGGCGGCGAGGAAGTGGCGGAAAATACAGCTATCGCCCGACAGATGCG
 GAAAAGGGCGGATTCGGCGTGTTCGCGGCAAAAAAGAGCAGGATGGATCCGGAGGAGGAGGCCACCTACAAAGTG
 GACGAATATCACGCCAACGCCCCGTTTCGCCATCGACCATTTCAACACCAGCACCAACGTCGCGCGGTTTTACGGTCTG
 ACCGTTCCGTCGAGTTTCGACCAAGCAAAACGCGGATAAAATCGACATCACCATCCCCGTTGCCAACCTGCAAAGC
 GGTTCGCAACACTTTACCAGCACCTGAAATCAGCCGACATCTTCGATGCCGCAATATCCGGACATCCGCTTTGTT
 TCCACCAAATCAACTTCAACGGCAAAAACCTGGTTTCCGTTGACGGCAACCTGACCATGCACGGCAAAACCGCCCC
 GTCAAACCTCAAAGCCGAAAATTCAACTGCTACCAAAGCCCGATGGCGAAAACCGAAGTTTTCGCGGGCGACTTCAGC
 ACCACCATCGACCGCAAAATGGGGCTGGACTACCTCGTTAACGTTGGTATGACCAAAAGCGTCCGCATCGACATC
 CAAATCGAGGCAGCCAAACAATAACTCGAG

1 MASPDVKSAD TLSKPAAPV AVAEKETEVEKED APQAGSQGQ APSTQGSQDM
 51 AAVSAENTGN GGAATTDKPK NEDEGPQNDM PQNSAESANQ TGNNQPADSS
 101 DSAPASNPA ANGGSNFGRV DLANGVLIDG PSQNTITLTHC KGDSCNGDNL
 151 LDEEAPSKSE FENLNESERI EKYYKDKGSD KFTNLVATAV QANGTNKYVI

201 IYKDKSASS SARFRRSARS RRS LPAEMPL IPVNQADTLI VDGEAVSLTG
 251 HSGNIFAPEG NYRYLTYGAE KLP GGSYALR VQGEPAKGEM LAGTAVYNGE
 301 VLHFHTENGR PYPTRGRFAA KVDFGSKSVD GIIDSGDDLH MGTQKPKAAI
 351 DGNFGKGTWT ENGGGDVSGR FYGPAGEEVA GKYSYRPTDA EKGGFVVFAG
 401 KKEQDGS GGG GATYKVD EYH ANARFAIDHF NTSTNVGGFY GLTGSVEFDQ
 451 AKRDGKIDIT IPVANLQSGS QHFTDHLKSA DIFDAAQYPD IRFVSTKFNF
 501 NGKKLVSDG NLTMHGKTAP VKLKA EKFN C YQSPMAKTEV CGGDFSTTID
 551 RTRKWGDVYLV NVG MTKSVRI DIQIEAAKQ*

AG287-961

ATGGCTAGCCCGGATGTTAAATCGGCGGACACGCTGTCAAACCGGCCGCTCCTGTTGTTGCTGAAAAAGAGACAGAG
 GTAAAAGAAGATGCGCCACAGGCAGGTTCTCAAGGACAGGGCGGCCATCCACACAAGGCAGCCAAGATATGGCGGCA
 GTTTCGGCAGAAAATACAGGCAATGGCGGTGCGGCAACAACGGACAAACCCAAAAATGAAGACGAGGGACCGCAAAAT
 GATATGCCGCAAAAATTCGCCGAATCCGCAAAATCAAACAGGGAACAACCAACCCGCCGATTTCTCAGATTCGCCCCCC
 GCGTCAAACCTGCACCTGCGAATGGCGGTAGCAATTTTGGAAAGGGTTGATTTGGCTAATGGCGTTTGTATTGATGGG
 CCGTCGCAAAATATAACGTTGACCCACTGTAAAGGCGATTCTTGTAAATGGTGATAATTTATTGGATGAAGAAGCACCG
 TCAAAAATCAGAATTTGAAAATTTAAATGAGTCTGAACGAATTTGAGAAATATAAGAAAGATGGGAAAAGCGATAAATTT
 ACTAATTTGGTTGCGACAGCAGTTCAAGCTAATGGAAC TAACAAATATGTCATCATTTATAAAGACAAGTCCGCTTCA
 TCTTCATCTGCGCGATTTCAGGCGTTTGCACGGTCGAGGAGTTCGCTTCCGCGGAGATGCCGCTAATCCCGCTCAAT
 CAGGCGGATACGCTGATGTCGATGGGGAAGCGGTCAGCCTGACGGGGCATTCGGGCAATATCTTCGCGCCCGAAGGG
 AATTACCGGTATCTGACTTACGGGGCGGAAAAATGCCCCGGCGGATCGTATGCCCTCCGCTGCAAGGCGAACCGGCA
 AAAGGCGAAATGCTTGTGCGCACGGCCGTGTACAACGGCGAAGTCTGCAATTTTCATACGGAAAACCGCCGTCGCTAC
 CCGACTAGAGGCAGGTTTGC CGCAAAAGTFCGATTTCCGCGACCAATCTGTGGACGGCATTATCGACAGCGCGATGAT
 TTGCATATGGGTACGCAAAAATTCAAAGCCGCATCGATGGAAACGGCTTTAAGGGGACTTGGACGGAAAATGGCGGC
 GGGGATGTTTCCGGAAGGTTTACGGCCCGCCGGCGAGGAAGTGGCGGGAAAATACAGCTATCGCCCGACAGATGCG
 GAAAAGGGCGGATTCGGCGTGTTCGCGCAAAAAGAGCAGGATGGATCCGAGGAGGAGGCCACAAACGACGAC
 GATGTTAAAAAGCTGCCACTGTGGCCATGCTGCTGCCTACAACAATGGCCAAAGAAATCAACGGTTTCAAAGCTGGA
 GAGACCATCTACGACATGATGAAGACGGCACAATTACCAAAAAGACGCAACTGCAGCCGATGTTGAAGCCGACGAC
 TTTAAAGGCTGCGGTCTGAAAAAGTCTGACTAACCTGACCAAAAACCGTCAATGAAAACAAAACGCTCGATGCC
 AAAGTAAAAGCTGCAGAATCTGAAAATGAAAAAGTAACTAACCAAGTTAGCAGACACTGATGCCGCTTTAGCAGATACT
 GATGCCGCTCTGGATGCAACCACCAACGCCCTTGAATAAATTTGGGAGAAAATATAACGACATTTGCTGAAGAGACTAAG
 ACAATATCGTAAAAATGATGAAAAATTAGAAGCCGTGGCTGATACCGTCGACAAGCATGCCGAAGCATTCAACGAT
 ATCGCCGATTCATTTGGATGAAACCAACACTAAGGCAGACGAAGCCGTCAAAACCGCCAATGAAGCCAAACAGACGGCC
 GAAGAAAACCAACAAAACGTCGATGCCAAAGTAAAAGTGCAGAACTGCAGCAGGCAAGCCGAAGCTGCCGCTGGC
 ACAGCTAATATGTCAGCCGACAAAGCCGAAGCTGTGCTGCAAAAAGTTACCGACATCAAAGCTGATATCGCTACGAAC
 AAAGTAAATATTGCTAAA AAAAGCAAAACAGTCCGACGCTGACACCAGAGAAGAGTCTGACAGCAAATTTGTCAGAATT
 GATGGTCTGAACGCTACTACC GAAAAATGGACACACGCTTGGCTTCTGCTGAAAAATCCATTTGCCGATCAGGATACT
 CGCCTGAACGGTTTGGATAAAACAGTGTGACACCTGCGCAAAAGAAAACCCGCCAAGGCCCTTCAGAAACAAGCCGCTC
 TCCGCTCTGTTCCAACCTTACAACGTGGGTCCGTTCAATGTAACGGCTGCAGTCGGCGGCTACAAATCCGAATCGGCA
 GTCGCCATCGGTACCGGCTTCCGCTTTACCGAAAACCTTTGCCGCCAAAGCAGGCGTGGCAGTCCGGCACTTCTGCTCCGCT
 TCTTCCGACGCTACCATGTGCGCGTCAATTACGAGTGGTAACTCGAG

1 MASPDVKSAD T LSKPAAPVV AEKETE VKED APQAGSQGQ APSTQGSQDM
 51 AAVSAENTGN GGAATTDKPK NEDEGPONDM PQNSAESANQ TGNQPADSS
 101 DSAPASNPA ANGGSNFGRV DLANGVLIDG PQNITLTHC KGDSCNGDNL
 151 LDEEAPSKSE FENLNE SERI EKYKDKSD KFTNLVATAV QANGTNKYVI
 201 IYKDKSASS SARFRRSARS RRS LPAEMPL IPVNQADTLI VDGEAVSLTG
 251 HSGNIFAPEG NYRYLTYGAE KLP GGSYALR VQGEPAKGEM LAGTAVYNGE
 301 VLHFHTENGR PYPTRGRFAA KVDFGSKSVD GIIDSGDDLH MGTQKPKAAI
 351 DGNFGKGTWT ENGGGDVSGR FYGPAGEEVA GKYSYRPTDA EKGGFVVFAG
 401 KKEQDGS GGG GATNDDVVK AATVAIAAAY NNGQEIINGFK AGETIYDIDE
 451 DGTITKDAT AADVEADDFK GLGLKVVVN LTKTVNENKQ NVDKVKAAE
 501 SELEKLTTL ADTDAALADT DAALDATTNA LNKLGENTTT FAEFTKTNIV
 551 KIDEKLEAVA DTVDKHAEAF NDIADSLDET NTKADEAVKT ANEAKQTAE E
 601 TRQNVDAKVK AETAAGKAE AAAGTANTAA DKAEVAARKV TDIKADIATN
 651 KDNIAKKANS ADVYTREESD SKFVRIDGLN ATTEKLDTRL ASAEKSIADH
 701 DTRLNGLDKT VSDLRKETRO GLAEQAALSG LFQPYNVGRF NVTAAVGGYK
 751 SESAVAIGTG FRFTENFAAK AGVAVGTSSG SSAAYHVGUN YEW*

	ELISA	杀菌
△G287-953-His	3834	65536
△G287-961-His	108627	65536

对 919 和 ORF46.1 而言, 将针对杂交蛋白产生的抗体的杀菌效力(同源菌株)与针对组分抗原(用 287-GST)的简单混合物产生的抗体相比:

	与 287 的混合物	与△G287 的杂交体
919	32000	128000
ORF46.1	128	16000

还获得了针对异源 MenB 菌株和针对血清型 A 和 C 的杀菌活性的数据:

	919		ORF46.1	
菌株	混合物	杂交体	混合物	杂交体
NGH38	1024	32000	-	16384
MC58	512	8192	-	512
BZ232	512	512	-	-
MenA (F6124)	512	32000	-	8192
MenC (C11)	>2048	>2048	-	-
MenC (BZ133)	>4096	64000	-	8192

因此, 在 N-末端与△G287 的杂交蛋白在免疫学上优于与△G287-ORF46.1 的简单混合物, 即使针对异源菌株也特别有效。可以在 pET-24b 中表达△G287-ORF46.1。

用新西兰菌株 394/98, 而非 2996, 制备相同的杂交蛋白:

△G287NZ-919

ATGGCTAGCCCGATGTCAGTCGGCGGACACGCTGTCAAACCTGCCGCCCTGTTGTTTCTGAAAAAGAGACAGAG
 GCAAAGGAAGATGCGCCACAGGCAGGTCTCAAGGACAGGGCGCGCCATCCGCACAAAGCCGGTCAAGATATGGCGGCG
 GTTTCGGAAGAAAATACAGGCAATGGCGGTGCGGCAGCAACGGACAAACCCAAAAATGAAGACGAGGGGCGCAAAAT
 GATATGCCGCAAAATGCCGCCGATACAGATAGTTGACACCGAATCACACCCCGGCTTCGAATATGCCGGCCGGAAT
 ATGAAAACCAAGCACCGGATGCCGGGAATCGGAGCAGCCGGCAAACCAACCGGATATGGCAAATACGGCGGACGGA
 ATGCAGGGTGACGATCCGTCGGCAGGCGGGGAAAATGCCGGCAATACGGCTGCCCAAGGTACAAATCAAGCCGAAAAC
 AATCAAACCGCCGGTTCTCAAATCCTGCCTCTTCAACCAATCCTAGCGCCACGAATAGCGGTGGTGATTTTGGAAAGG
 ACGAACCTGGGCAATCTGTGTGATTGACGGGCGGTGCGAAAATATAACGTTGACCCACTGTAAAGGCGATTTCTGT
 AGTGGCAATAATTTCTTGGATGAAGAGTACAGCTAAAATCAGAATTTGAAAAATTAAGTGATGCAGACAAAATAAGT
 AATTACAAGAAAGATGGGAAGAATGACGGGAAGAATGATAAATTTGTCGGTTTGGTTGCCGATAGTGTGACAGATGAAG
 GGAATCAATCAATATATATCTTTTATAAACCTAAACCCACTTCATTTGCGCGATTTAGGCGTTCGACGGTTCGAGG
 CGGTGCTTCCGGCCGAGATGCCGCTGATTCCTGCAATCAGGCGGATACGCTGATGTCGATGGGAAGCGGTGAGC
 CTGACGGGCGATTCGGCAATATCTTCGCGCCCGAAGGGAAATACCGGTATCTGACTTACGGGGCGGAAAAATGCCC
 GGCGGATCGTATGCCCTCCGTGTTCAAAGGCGAACCTTCAAAGGCGAAATGCTCGCGGGCACGGCAGTGTACAACGGC
 GAAGTGTGCAATTTTCATACGGAACCGCCGTCGTCCTCCCGTCCAGAGGCAGGTTTGGCCGCAAAAGTCGATTTCCGGC
 AGCAAAATCTGTGGACCGCATTATCGACAGCGGCGATGGTTTGCATATGGGTACGCAAAATTCAAAGCCGCAATCGAT
 GGAAACGGCTTTAAGGGGACTTGGACGGAATGGCGGGGATGTTTCCGGAAAGTTTACGGCCCGCCGGCGAG

GAAGTGGCGGGAAAATACAGCTATCGCCCAACAGATGCGGAAAAGGGCGGATTCGGCGTGTTCGCGGCAAAAAAGAG
 CAGGATGGATCCGGAGGAGGAGGATGCCAAAGCAAGAGCATCCAAACCTTCCGCAACCCGACACATCCGTCATCAAC
 GGCCCGGACCGGCCGTCGGCATCCCGACCCCGCGGAACGACCGTCCGCGCGGGCGGGCGGCTATACCGGTTGTA
 CCGCACTGTCCCTGCCCACTGGGCGGCGCAGGATTCGCCAAAAGCCTGCAATCCTTCCGCTCGGCTGCGCAAT
 TTGAAAAACCGCAAGGCTGGCAGGATGTGTGCGCCCAAGCCTTTCAAACCCCGTCCATTCTTTCAGGCAAAACAG
 TTTTGTGAACGCTATTTACGCGCTGGCAGGTTGCAGGCAACGGAAGCCTTCCCGGTACGGTTACCGGCTATTACGAG
 CCGGTGCTGAAGGGCGACGACAGGCGGACGACAGCCCGCTTCCCGATTTACGGTATTCCCGACGATTTATCTCC
 GTCCCTTGCCTGCGGTTTTCGCGAGCGGAAAAGCCCTTGTCCGCATCAGGCAGACGGGAAAAACAGCGGCACAATC
 GACAATACCGGCGGCACACATACCGCGGACCTCTCCGATTCCCATCAGCGCGCACACGGAATCAAGCGGAGG
 TTTGAAGGAAGCCGCTTCCCTCCCTACCACACGCGCAACCAATCAACGCGCGCGCGCTTGACGGCAAAGCCCGATA
 CTCGGTTACGCGGAAGACCCCGTCAACTTTTTTTTATGCACATCCAAGGCTCGGGCGCTGTGAAAACCCCGTCCGGC
 AAATACATCCGCATCGGCTATGCCGACAAAAACGAACATCCCTACGTTTCCATCGGACGCTATATGGCGGACAAAGGC
 TACCTCAAGCTCGGCGAGACCTCGATGCAGGCGATCAAGCCTATATGCGGCAAAATCCGCAACGCCCTCGCCGAAGTT
 TTGGGTCAAACCCAGCTATATCTTTTCCGCGAGCTTCCCGGAAGCAGCAATGACCGTCCCGTCCGCGCAGTGGGC
 ACGCCGTTGATGGGGGAATATGCGGCGCAGTCGACCGGCACTACATTACCTTGGGCGCGCCCTTATTTGTGCGCCAC
 GCCCATCCGTTTACCCGCAAAGCCCTCAACCCGCTGATTATGGCGCAGGATACCGGCGAGCGGATTAAGGCGCGGTTG
 CCGTGGATTATTTTGGGATACGGCGACGAAGCCGGCGAACTTCCCGGCAAACAGAAAACACGGGTTACGTTCTGG
 CAGCTCTACCAACGGTATGAAGCCCGAATACCGCCCGTAAAAGCTT

- 1 MASPDVKSAD TLSKPAAPVV SEKETEAKED APQAGSQGG APSAQQGDM
- 51 AAVSEENIGN GGAATDKPK NEDEGAQNDM PQNAADTDSL TPNHTPASNM
- 101 PAGNMENQAP DAGESEQPAN QPDMANTADG MQGDDPSAGG ENAGNTAAQG
- 151 TNQAENNQTA GSQNPASSTN PSATNSGGDF GRTNVGNSVV IDGPSQNTL
- 201 THCKGDCSG NNFLEEVQL KSEFEKLSDA DKISNYKKG KNDGKDKFV
- 251 GLVADSVQMK GINQYIIFYK PKPTSFAFR RSARSRRSLP AEMPLIPVNO
- 301 ADTLIVDGEA VSLTGHSGNI FAPEGNYRVL TYGAEKLPFG SYALRVQGEF
- 351 SKGEMLAGTA VYNGEVLHFM TENGRRPSPSR GRFAAKVDFG SKSVDGIIDS
- 401 GDGLHMGTK FKAALDNGF KGTWTEGGG DVSGRFYGPA GEEVAGKYSY
- 451 RPTDAEKGGF GVFAGKKEQD GSGGGGCQSK SIQTFPQPD SVINGEDRPFV
- 501 GIPDPAGTTV GGGAVYTVV PHLSLPHWAA QDFAKSLQSF RLGCANLKNR
- 551 QGWQDVCAQA FQTFVHSFQA KOFFERYFPP WQVAGNGSLA GTVIGYYEPV
- 601 LKGDERRTAQ ARFPYIGIPD DFISVPLPAG LRSKALVRI RQTGKNSGTI
- 651 DNTGGHTHAD LSRFPITART TAIKGRFEGS RFLPYHTRNQ INGGALDGA
- 701 PILGYAEDFV ELFFMHIOGS GRLKTPSGKY IIRIGYADKNE HPVVISGRYM
- 751 ADKGYLKLQ TSMQGIKAYM RONPQRLAEV LGQNPYSYIFF RELAGSSNDG
- 801 PVGALGTPLM GEYAGAVDRH YITLGAPLFV ATAHVTRKA LNRLIMAQDT
- 851 GSAIKGAVRV DYFWGYGDEA GELAGKQKTT GYVWQLLPNG MKPEYRP*

AG287NZ-953

ATGGCTAGCCCCGATGTCAAGTCCGCGGACACGCTGTCAAACCTGCCGCCCTGTTGTTTCTGAAAAAGAGACAGAG
 GCAAAGGAAGATGCGCCACAGGCAGGTTCTCAAGGACAGGGCGCCATCCGCACAAGGCGGTCAAGATATGGCGGCG
 GTTTCGGAAGAAAATACAGGCAATGGCGTCCGCGCAGCAACCGGCAACCCAAAAATGAAGACGAGGGGGCGCAAAAT
 GATATGCCGCAAAATGCCGCGGATACAGATGTTTGACACCGAATCACACCCCGCTTCAATATGCGCGCGGCAAT
 ATGGAACAAGCACCAGGATGCCGGGAATCGGAGCAGCCGCAACCAACCGGATATGGCAAAATACGGCGGACGGA
 ATGCAGGGTACGATCCGTCGGCAGGCGGGGAAAATGCCGCAATACGGCTGCCAAGGTACAAATCAAGCCGAAAAC
 AATCAAACCGCGGTTCTCAAATCTCGCTCTTCAACCAATCTTAGCGCCACGAATAGCGGTGGTGTATTTGGAAGG
 ACGAAGCTGGGCAATCTGTGTGATGACGGGCGTCGCAAAATATAACGTTGACCCACTGTAAAGGCGATTCTTGT
 AGTGGCAATAATTTCTTGGATGAAGAAGTACAGCTAAAAATCAGAAATTTGAAAAATTAAGTGTGCAGACAAAATAAGT
 AATTACAAGAAAGATGGGAAGAATGACGGGAAGAATGATAAATTTGTCGGTTTGGTTGCCGATAGTGTGCAGATGAAG
 GGAATCAATCAATATATATCTTTTATAAACCTAAACCCACTTCAATTTGCGCGATTAGGCGTTCTGCACGGTTCGAGG
 CGGTCCGTTCCGGCCGAGATGCCGCTGATTCCCGTCAATCAGGCGGATACGCTGATTGTGATGGGGAAGCGGTACGC
 CTGACGGGCAATTCGCGCAATATCTTCCGCGCCGAAGGGAATACCGGTATCTGACTTACGGGGCGGAAAAATGCC
 GCGGATCGTATGCCCTCCGTTTCAAGCGAACCTTCAAAGGCGAAATGCTCGCGGCGCAGGCAAGTGTACACCGG
 GAAGTGTGCATTTTCATACGGAACCGCCGTCCTCCCGTCCAGAGGCAGGTTTCCCGCAAAAGTCCGATTTCCG
 AGCAAATCTGTGGACGGCAATATCGACAGCGGCGATGGTTTGCATATGGGTACGCAAAAATTCAAAGCCGCCATCGAT
 GGAACCGCTTTAAGGGGACTTGGACGGAATGCGCGCGGGGATGTTCCGGAAGTTTTACGGCCCGCGCGGCGAG
 GAAGTGGCGGAAAATACAGCTATCGCCCAACAGATGCGGAAAAGGGCGGATTCGGCGTGTTCGCGGCAAAAAGAG
 CAGGATGCGCGGAGGAGGAGCCACTACAAGTGGACGAATACAGCACAACCGCCGTTTCGCAATCGACCAT
 TTCAACACCAGCACAACGTCGGCGGTTTTTACGCTGACCGGTTCCGTCGAGTTTCGACCAAGCAACCGGCGG
 AAAATCGACATCACCATCCCGTTGCCAACCTGCAAAGCGGTTCCGCAACACTTACCGACCCTGAAATCAGCCGAC
 ATCTTCGATGCCGCCAATATCCGGACATCCGCTTGTGTTCCACCAATCAACTTCAACGCAAAAACCTGGTTTCC
 GTTACGGCAACCTGACCATGACGGCAAAACCGCCCGTCAAACCTCAAAGCCGAAAAATTCAACTGCTACCAAAGC
 CCGATGGCGAAAACCGAAGTTTTCGCGCGGCGACTTCAAGCACCACATCGACCGCAACAAATGGGGCGTGACTACCTC
 GTTAACGTTGGTATGACCAAAAAGCGTCCGCATCCAAATCGAGGCGCAAAACAATAAAGCTT

- 1 MASPDVKSAD TLSKPAAPVV SEKETEAKED APQAGSQGG APSAQQGDM

51 AAVSEENTGN GGAAATDKPK NEDEGAQNDM PQNAADTDSL TPNHTPASNM
 101 PAGNMENQAP DAGESEQPAN QPDMANTADG MQGDDPSAGG ENAGNTAAQG
 151 TNQAENNQTA GSONPASSTN PSATNSGGDF GRTNVGNSVV IDGPPSQNITL
 201 THCKGSDCSG NNFLDEEVQL KSEFEKLSDA DKISNYKKDG KNDGKNDKPV
 251 GLVADSVQMK GINQYIIFYK PKPTSPARFR RSARSRRSLP AEMPLIPVNO
 301 ADTLIVDGEA VSLTGHSGNI FAPBGNRYRL TYGAEKLP GG SYALRVQGEF
 351 SKGEMLAGTA VYNGEVLHFFH TENGRPSPSR GRFAAKVDFG SKSVDGIIDS
 401 GDGLHMGTOK FKAALDNGF KGTWTENGGG DVSGKPYGPA GEEVAGKYSY
 451 RPTDAEKGGF GVFAGKKEQD GSGGGGATYK VDEYHANARF AIDHFNTSTN
 501 VGGFYGLTGS VEFDQAKRDG KIDITIPVAN LQSGSQHFTD HLKSADIFDA
 551 AQYPIRFVS TKFNFNKKL VSV DGNLTMH GKTAPVKLKA EKFNICYQSPM
 601 AKTEVCGGDF STTIDRTKVG VDYLNVNVMG KSVRIDIQIE AAKQ*

AG287NZ-961

ATGGCTAGCCCCGATGTC AAGTCGGCGGACACGCTGTCAAACCTGCCGCCCTGTTGTTTCTGAAAAAGAGACAGAG
 GCAAAGGAAGATGCGCCACAGGCAGGTTCTCAAGGACAGGGCGGCCATCCGCACAAGGCGGTCAAGATATGGCGGCG
 GTTTCGGAAGAAAATACAGGCAATGGCGGTGCGGCAGCAACGGACAAAACCCAAAAATGAAGACGAGGGGGCGCAAAAT
 GATATGCCGCAAAATGCCGCCGATACAGATAGTTTGACACCGAATCACACCCCGGCTTCGAATATGCCGCGGAAAT
 ATGGAAAACCAAGCAGCGGATGCCGGGAAATCGGAGCAGCCGGCAACCAACCGGATATGGCAAATACGGCGGACGGA
 ATGCAAGGTTGACGATCCGTCGGCAGGGCGGAAAATGCCGGCAATACCGCTGCCAAGGTACAAATCAAGCCGAAAAC
 AATCAAACCCGGTTCTCAAATCCTGCCTCTTCAACCAATCCTAGCGCCACGAATAGCGGTGGTGTATTTGGAAAGG
 ACGAACGTTGGCAATTTCTGTTGTGATTGACGGCCGTCGCAAAAATAACGTTGACCCACTGTAAGGCGGATTTCTGT
 AGTGGCAATAATTTCTTGGATGAAGAAGTACAGCTAAATCAGAATTTGAAAAAATTAAGTGATGCAGACAAAATAAGT
 AATTACAAGAAAGATGGGAAGAATGACGGGAAGAATGATAAATTTGTCGGTTTGGTTGCCGATAGTGTGCAGATGAAG
 GGAATCAATCAATATATATCTTTATAAACCTAAACCCACTTCATTTGCCGATTTAGGCGTTCTGCACGGTCGAGG
 CGGTCCGTTCCGGCCGAGATGCCGCTGATCCCGTCAATCAGGCGGATACGCTGATTTGCATGGGGAAGCGGTCAGC
 CTGACCGGGGCATTCGGCAATATCTTCGGCGCCGAAGGAAATTACCGGTATCTGACTTACGGGGCGGAAAAATGCCC
 GCGGATCGTATGCCCTCCGTTCAAGGCGAACCTTCAAAGGCGAAATGCTCGCGGGCACGGCAGTGTACAACGGC
 GAAGTGTGCATTTTCATACGGAAAACGGCCGTCGTCGCCGTCAGAGGCGGTTTGGCCGAAAAGTCGATTTCCGGC
 AGCAAATCTGTGACGGCATTATCGACAGCGCGGATGTTTGCATATGGGTACGCAAAAATTCAAAGCCGCCATCGAT
 GGAACCGGCTTAAAGGGGACTTGGACGGAAAATGGCGGCGGGGATGTTTCCGGAAGTTTACGGCCCGCCGGCGGAG
 GAAGTGGCGGAAAATACAGCTATCGCCCAACAGATCGGAAAAGGGCGGATTCGGCGTGTTCGCCGGCAAAAAGAG
 CAGGATGGATCCGGAGGAGGAGGCCACAACAGCAGCAGATGTTAAAAAAGCTGCCACTGTGGCCTGCTGCTGCC
 TACAACAATGGCCAAGAAATCAACGGTTTCAAAGCTGGAGAGACCATCTACGACATTGATGAAGACGGCACAAATACC
 AAAAAAGACGCAACTGCAGCCGATGTTGAAGCCGACGACTTAAAGGTCTGGGTCTGAAAAAAGTCGTGACTAACCTG
 ACCAAAACCGTCAATGAAAACAAAACAAACGTCGATGCCAAAGTAAAAGCTGCAGAAATCTGAAATAGAAAAGTTAAACA
 ACCAAGTTAGCAGACACTGATGCCGCTTTAGCAGACTGATGCCGCTCTGGATGCAACCACCAACGCCTTGAATAAA
 TTGGGAGAAAATATAACGACATTTGCTGAAGAGACTAAGACAAATATCGTAAAAATTTGATGAAAAATTAGAAGCCGTG
 GCTGATACCGTCGACAAGCATGCCGAAGCATTCAACGATATCGCCGATTCATTGGATGAAACCAACACTAAGGCGAGC
 GAAGCCGTCAAACCGCCAATGAAGCCAAACAGACGGCCGAAGAAACCAACAAAACGTCGATGCCAAAGTAAAAGCT
 GCAGAACTGCAGCAGGCAAAGCCGAAGCTGCCGCTGGCACAGCTAATACTGCAGCCGACAAGGCCGAAGCTGCTCGT
 GCAAAGTTACCGACATCAAAGCTGATATCGCTACGAAACAAAGATAATATTTGCTAAAAAGCAAACAGTCCGACGCTG
 TACACCAGAGAAGAGTCTGACAGCAAATTTGTGAGAATTTGATGGTCTGAACGCTACTACCAGAAAATTTGACACACGC
 TTGGCTTCTGCTGAAAAATCCATTGCCGATCACGATACTCGCTGAACGGTTTGGATAAAAACAGTGTGACACCTGCCG
 AAGAAAACCCGCAAGGCCTTGCAGAACAGCCGCGCTCTCCGGTCTGTTCCAACCTTACAACGTTGGTCCGTTCAAT
 GTAACGGCTGCAGTCCGGCGGTACAAATCCGAATCGGCAGTCCCATCGGTACCGCTTCCGCTTTACCGAAAACCTT
 GCCGCAAGCAGGCGTGGCAGTCCGCACTTCGTCGGTTTCTTCCGACGCTACCATGTCCGCTCAATTACGAGTGG
 TAAAAGCTT

1 MASPDVKSAD TLSKPAAPVV SEKETEAKED APQAGSQGQ APSAQGGQDM
 51 AAVSEENTGN GGAAATDKPK NEDEGAQNDM PQNAADTDSL TPNHTPASNM
 101 PAGNMENQAP DAGESEQPAN QPDMANTADG MQGDDPSAGG ENAGNTAAQG
 151 TNQAENNQTA GSONPASSTN PSATNSGGDF GRTNVGNSVV IDGPPSQNITL
 201 THCKGSDCSG NNFLDEEVQL KSEFEKLSDA DKISNYKKDG KNDGKNDKPV
 251 GLVADSVQMK GINQYIIFYK PKPTSPARFR RSARSRRSLP AEMPLIPVNO
 301 ADTLIVDGEA VSLTGHSGNI FAPBGNRYRL TYGAEKLP GG SYALRVQGEF
 351 SKGEMLAGTA VYNGEVLHFFH TENGRPSPSR GRFAAKVDFG SKSVDGIIDS
 401 GDGLHMGTOK FKAALDNGF KGTWTENGGG DVSGKPYGPA GEEVAGKYSY
 451 RPTDAEKGGF GVFAGKKEQD GSGGGGATND DDVKKAAATVA IAAAYNNGQE
 501 INGPKAGETI YDIDEDGTTT KKDATAADVE ADDFKGLGLK KVVNLTNKTIV
 551 NENQNVDK VKAABSEIEK LTTKLADTDA ALADTDAALD ATTNALNKLK
 601 ENITTFAEET KTNIVKIDEK LEAVADTVDK HAEAFNDIAD SLDETNTKAD
 651 EAVKTANEAK QTABETKQNV DAKVKAABTA AGKABAAAGT ANTAADKAEA
 701 VAAKVTDIKA DIATNKDNLA KKANSADVYT REESDSKPFVR IDGLNATTEK
 751 LDTRLASAEK SIADHDTRLN GLDKTVSDLR KETRQGLABQ AALSGLFPQY
 801 NVGRFNVTAA VGGYKSESAV AIGTGFRFTE NFAAKAGVAV GTSSGSSAAY
 851 HVGVNYEW*

实施例 3-△G983 的杂交体

蛋白质 983 具有如下序列:

```

983                                     ➡ AG983
1  MRTTPTFPTK TFKPTAMALA VATTLSACLG GGGGTSAPD FNAGGTGIGS
51 NSRATTAKSA AVSYAGIKNE MCKDRSMLCA GRDDVAVTDR DAKINAPPPN
101 LHTGDFPNPN DAYKNLINLK PAIEAGYTGR GVEVGIVDTG ESVGSISFPE
151 LYGRKEHGYN ENYKNYTAYM RKEAPEDGGG KDIEASFDDE AVIETEAKPT
201 DIRHVKEIGH IDLVSHIIGG RSVDRPAGG IAPDATHLIM NTNDETKNEM
251 MVAAIRNAWV KLGERGVRIV NNSFGTTSRA GTADLFQIAN SEEQYRQALL
301 DYSGGDKTDE GIRLMQOSDY GNLSYHIRNK NMLFIFSTGN DAQAQPNTYA
351 LLPFYEKDAQ KGIITVAGVD RSGEKFKREM YGEPGTEPLE YGSNHCGITA
401 MWCLSAPYEA SVRFTRTNPI QIAGTSFSAP IVTGTAALLL QKYPWMSNDN
451 LRTLLTTAQ DIGAVGVDSK FGWGLLDAGK AMNGPASFPF GDFTADTKGT
501 SDIAYSFRND ISGTGGLIKK GGSQQLHGN NTYTGKTIIE GGSVLVLYGNN
551 KSDMRVETKG ALIYNGAASG GSLNSDGIVY LADTDQSGAN ETVHIKGSLO
601 LDGKGTLYTR LGKLLKVDGT AIIGGKLYMS ARGKGAGYLN STGRRVFFLS
651 AAKIGQDYSF FTNIETDGG LSLDSVEKT AGSEGDLSY YVRRGNAART
701 ASAAHSAPA GLKHAVEQGG SNLENLMVEL DASESSATPE TVETAADRT
751 DMPGIRPYGA TFRAAAVQH ANAADGVRIF NSLAATVYAD STAAHADMQG
801 RRLKAVSDGL DHNGTGLRVI AQTQDGGTW EQGGVEGKMR GSTQTVGIAA
851 KTGENTTAAA TLGMGRSTWS ENSANAKTDS ISLFAGIRHD AGDIGYLGKL
901 FSYGRYKNSI SRSTGADEHA EGSVNGTLMQ LGALGGVNPV FAATGDLTVE
951 GGLRYDLLKQ DAFAEKGSAL GWSGNSLTEG TLVGLAGLKL SQPLSDKAVL
1001 FATAGVERDL NGRDYTVTGG FTGATAATGK TGARNMPHTR LVAGLGADVE
1051 FGNGWNLAR YSYAGSKQYG NHSGRVGVGY RF*

```

因此△G983 具有如下基本序列:

```

                                     TSAPD FNAGGTGIGS
NSRATTAKSA AVSYAGIKNE MCKDRSMLCA GRDDVAVTDR DAKINAPPPN
LHTGDFPNPN DAYKNLINLK PAIEAGYTGR GVEVGIVDTG ESVGSISFPE
LYGRKEHGYN ENYKNYTAYM RKEAPEDGGG KDIEASFDDE AVIETEAKPT
DIRHVKEIGH IDLVSHIIGG RSVDRPAGG IAPDATHLIM NTNDETKNEM
MVAAIRNAWV KLGERGVRIV NNSFGTTSRA GTADLFQIAN SEEQYRQALL
DYSGGDKTDE GIRLMQOSDY GNLSYHIRNK NMLFIFSTGN DAQAQPNTYA
LLPFYEKDAQ KGIITVAGVD RSGEKFKREM YGEPGTEPLE YGSNHCGITA
MWCLSAPYEA SVRFTRTNPI QIAGTSFSAP IVTGTAALLL QKYPWMSNDN
LRTLLTTAQ DIGAVGVDSK FGWGLLDAGK AMNGPASFPF GDFTADTKGT
SDIAYSFRND ISGTGGLIKK GGSQQLHGN NTYTGKTIIE GGSVLVLYGNN
KSDMRVETKG ALIYNGAASG GSLNSDGIVY LADTDQSGAN ETVHIKGSLO
LDGKGTLYTR LGKLLKVDGT AIIGGKLYMS ARGKGAGYLN STGRRVFFLS
AAKIGQDYSF FTNIETDGG LSLDSVEKT AGSEGDLSY YVRRGNAART
ASAAHSAPA GLKHAVEQGG SNLENLMVEL DASESSATPE TVETAADRT
DMPGIRPYGA TFRAAAVQH ANAADGVRIF NSLAATVYAD STAAHADMQG
RRLKAVSDGL DHNGTGLRVI AQTQDGGTW EQGGVEGKMR GSTQTVGIAA
KTGENTTAAA TLGMGRSTWS ENSANAKTDS ISLFAGIRHD AGDIGYLGKL
FSYGRYKNSI SRSTGADEHA EGSVNGTLMQ LGALGGVNPV FAATGDLTVE
GGLRYDLLKQ DAFAEKGSAL GWSGNSLTEG TLVGLAGLKL SQPLSDKAVL
FATAGVERDL NGRDYTVTGG FTGATAATGK TGARNMPHTR LVAGLGADVE
FGNGWNLAR YSYAGSKQYG NHSGRVGVGY RF*

```

将△G983 作为在其 C-末端与 ORF46.1、741、961 或 961c 的杂合体表达:

ΔG983-ORF46.1

```

ATGACTTCTGCGCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCA
GCAGTATCTTACGCGGTATCAAGAACGAAATGTGCAAAAGACAGAAGCATGCTCTGTGCGCGGTCCGGATGACGTTGCG
GTTACAGACAGGGATGCCAAAATCAATGCCCCCCCGGAATCTGCATACCGGAGACTTTCCCAAACCCAAAATGACGCA
TACAAGAATTTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTTCGAC
ACAGGCGAATCCGTCCGCAGCATATCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAA
AACTATACGGCGTATATGCGGAAGGAAGCGCCTGAAGACGGAGCGGTAAGACATTGAAGCTTCTTTTCGACGATGAG
GCCGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAGAAAATCGGACACATCGATTGGTCTCCCAT

```

ATTTATTGGCGGGCGTTCCCGTGGACGGCAGACCTGCAGGGCGGTATTGCGCCCGATGCGACGCTACACATAATGAATACG
AATGATGAAACCAAGAACGAAATGATGGTTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTTGGCGTGCGC
ATCGTCAATAACAGTTTTTGAACAACATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTCGGAGGAGCAG
TACCGCCAAGCGTTGCTCGACTATTCCGGCGGTGATAAAACAGACGAGGGTATCCGCCTGATGCAACAGAGCGATTAC
GGCAACCTGTCTTACCACATCCGTAATAAAAAATGCTTTTCATCTTTTCGACAGGCAATGACGCACAAGCTCAGCCC
AACACATATGCCCTATTGCCATTTTATGAAAAGACGCTCAAAAAGGCATTATCACAGTCGCAGGCGTAGACCCGCAAT
GGAGAAAAGTTCAAACGGGAAATGTATGGAGAACCGGGTACAGAACCCTTGAGTATGGCTCCAACCATTGCGGAATT
ACTGCCATGTGGTGCCTGTGGCACCCTATGAAGCAAGCGTCCGTTTCAACCCGTACAAACCCGATTCAAATGCGCGGA
ACATCCTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTTGATGAGCAACGACAAAC
CTGCGTACCACGTTGCTGACGACGGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCCGGCTGGGGACTGCTG
GATGCGGGTAAGGCCATGAACGGACCCGCGCTTTCCGTTFCGGCGACTTTACCGCCGATACGAAAGGTACATCCGAT
ATTGCTACTCCTTCCGTAACGACATTTTACGGCACGGGGCGCTGATCAAAAAGGGCGGCACTGCAACTGCAC
GGCAACAACACCTATACGGGCAAAACCATTATCGAAGGCGGTTCCGTTGGTGTGTACGGCAACAACAAATCGGATATG
CGCGTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGCGCGCAGCTGAACAGCGACGGCATTGTCTAT
CTGGCAGATACCGACCAATCCGGCGCAAACGAAACCGTACACATCAAAGGCAGTCTGCAGCTGGACGGCAAAGGTACG
CTGTACACACGTTTGGGCAAACTGCTGAAAGTGGACGGTACGGCGATTATCGGCGGCAAGCTGTACATGTTCGGCACGC
GGCAAGGGGGGCGACTTCTCAACAGTACCGGACGCGTGTTCCTTCTGAGTCCGCCAAAATCGGGCAGGATTAT
TCTTTCTTCAAAAACATCGAAACCGACGGCGGCTGCTGGCTTCCCTCGACAGCGTCGAAAAACAGCGGGCAGTGAA
GGGACACGCTGTCTTATATGTCCGTGCGCGCAATGCGGCACGGACTGCTTCGCGCAGCGGCACATTCCGCGCCCGCC
GGTCTGAAACACCGCGTAGAACAGGGCGGCGCAATCTGGAAAAACCTGATGGTGAACCTGGATCCCTCCGAATCATCC
GCAACCCCGAGACGGTTGAAACTGCGGCAGCCGACCGCAGATATGCCGGCATCCGCCCTTACGGCGCAACTTTTC
CGCGCAGCGCGGCGTACAGCATGCGAATGCCCGCCGACGCTGTACGATCTTCAACAGTCTCGCGCTACCGTCTAT
GCCGACAGTACCGCGCCCATGCCGATATGCAGGACGCGCGCTGAAAGCCGTATCGGACGGGTTGGACCACAACGGC
ACGGGTCTGCGGTCATCGCGCAAAACCAACAGGACGGTGAACGTTGGAAACAGGGCGGTGTGAAGGCAAAAATGCCG
GGCAGTACCAAAACCGTCCGCATTCGCCGAAAACCGGGGAAAATACGACAGCAGCCGCCACACTGGGCGATGGGACGC
AGCAGTGGAGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATCGGCAGTACGGCAGTCCGGC
GATATCGGCTATCTCAAAGGCTGTCTTCTTACCGGACGCTACAAAACAGCAGCATCAGCCGACAGCCGGTGCAGGAA
CATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTCCGCAACG
GGAGATTTGACGGTGAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCGGAAAAGGCGAGTGTCTTG
GGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTCCGACTCGCGGGTCTGAAGCTGTGCAACCCCTTGAGCGAT
AAAGCCGCTCTGTTTGAACGGCGGGCGTGAACCGCACCTGAACGGACGCGACTACACGCTAACGGCGGGCTTTACC
GGCGCAGCTGCAGCAACCGCAAGACGGGGGCGCAATATGCGCGCACACCCGCTCTGGTTGCGCGCCTGGGCGCGGAT
GTCGAAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCGGTTCCAAAACAGTACGGCAACCCAGCGGA
CGAGTCGGCGTAGGCTACCGGTTCCCTGACGGTGGCGGAGGCACTGGATCCTCAGATTTGGCAAACGATTCCTTTTATC
CGGCAGGTTCTCGACCGTCAACGATTTGAAACCCGACGGGAAATACCACCTATTCGGCAGCAGGGGGAACTTCCCGAG
CGCAGCGCCATATCGGATTTGGAAAAATACAAAGCCATCAGTTGGGCAACCTGATGATTCAAACAGGCGGCAATTA
GGAAATATCGGCTACATTTGCTCCGCTTTTCCGATCACGGGCACGAAAGTCCATTTCCCTTCGACAACCATGCCTCACAT
TCCGATTTGATGAAGCCGGTAGTCCCGTTGACGGATTTAGCCTTTACCGCATCCATTTGGGACGGATACGAACACCAT
CCCGCGACGGCTATGACGGGCCACAGGGCGGGCTTATCCGCTCCCAAAGGCGGAGGGATATATACAGCTACGAC
ATAAAAGGCGTTGCCAAAATATCCGCTCAACCTGACCGCAACCGCAGCACCAGGCAACGGCTTGGCGACCGGTTTC
CACAAATGCGCGTAGTATGCTGACGCAAGGAGTAGGCGCAGGATTCAAACGCGCCACCCGATACAGCCAGCTGGAC
AGATCGGGCAATGCCCGCAAGCCTTCAACGGCACCTGCAGATATCGTTAAAAACATCATCGGCGCGGACAGGAAAT
GTCGGCGCAGGGATGCCGTCAGGGCATAAGCGAAGGCTCAAACATGCTGTGCATGCACGGCTTGGGTCTGCTTTCC
ACCGAAAACAAGATGGCGGCATCAACGATTTGGCAGATATGGCGCAACTCAAAGACTATGCCCGCAGCAGCCATCCGC
GATTTGGCAGTCCAAAACCCCAATGCCGCAAGGCATAGAAGCCGTCAGCAATATCTTTATGGCAGCCATCCCCATC
AAAGGGATTTGGAGCTGTTCGGGGAAAATACGGCTTGGGCGGCATCAGGCACATCCTATCAAGCGGTTCGAGATGGGC
GCGATCGCATTCGCCGAAAGGAAATCCGCGTCAAGGCAATTTTGGCGATGCGCGCATACGCCAAATACCCGTTCCCT
TACCATTCCCAGAAATATCCGTTCAAACCTGGAGCAGGTTACGGCAAAGAAAACATCACCTCTCAACCGTGCCGCGC
TCAAACGGCAAAAATGTCAAACCTGGCAGACCAACGCCACCCGAAGACAGGCGTACCGTTTGACGGTAAAGGTTTCCG
AATTTTGAAGAAGCACGTTGAAATATGATACGCTCGAGCACCACCACCACCACTGA

1	MTSAPDFNAG	GTGIGSNSRA	TTAKSAVSY	AGIKNEMCKD	RSMLCAGRDD
51	VAVTDRDAKI	NAPFPNLHTG	DFPNPNDAYK	NLNLKPAIE	AGYTGRGVEV
101	GIVDTGESVG	SISFPELYGR	KEHGYNENYK	NYTAYMRKEA	PEDGGGKDI
151	ASFDDEAVIE	TEAKPTDIRH	VKEIGHIDLIV	SHIIGGRSVD	GRPAGGLAPD
201	ATLHIMNTND	ETKNEMMVA	IRNAWVKLGE	RGVRIVNNSF	GTTSRAGTAD
251	LFQIANSEEQ	YRQALLDYSG	GDKTDEGIRL	MQQSDYGNLS	YHIRNKNMLF
301	IFSTGNDAQA	QPNFYALLPF	YEKDAQRGI	TVAGVDRSGE	KFKREMYGEP
351	GTEPLEYGSN	HCGITAMWCL	SAPYEASVRF	TRTNPIQIAG	TSFSAPIVTG
401	TAALLLQKYP	WMSNDNLRRT	LLTTAQDIGA	VGVDKFWGW	LLDAGKAMNG
451	PASFPFGDFT	ADTRKGTSDIA	YSFRNDISGT	GGLIKRGGSQ	LQLHGNNTYT
501	GRTIIEGGSL	VLYGNKSDM	RVETKCALIY	NGAASGGSLN	SDGIVYLADT
551	DQSGANETVH	IKGSLQLDGE	GTLYTRLGKL	LKVDGTALIG	GKLYMSARGK
601	GAGYLNSTGR	RVPFLSAAKI	GQDYSFFTNI	ETDGGLLASL	DSVEKTAGSE
651	GDTLSEYVRR	GNAARTASAA	AHSAPAGLKH	AVEQGGSNLE	NLMVELDASE
701	SSATPETVET	AAADRTDMPG	IRPYGATFRA	AAAVQHANA	DGVRIFNLSA
751	ATVYADSTAA	HADMQRRLK	AVSDGLDHNG	TGLRVIAQTQ	QDGGTWEQGG

801	VEGKMRGSTQ	TVGLAAKTGE	NTTAAATLGM	GRSTWSENSA	NAKTDSISLF
851	AGIRHDAGDI	GYLKGLFSYG	RYKNSISRST	GADEHAEGSV	NGTLMQLGAL
901	GVVNPFFAAT	GDLTVEGGLR	YDLLKQDAFA	EKGSALGWSG	NSLTEGTLVG
951	LAGLKLSQPL	SDKAVLFATA	GVERDLNGRD	YTVTGGFTGA	TAATGKTGAR
1001	NMPHTRLVAG	LGADVEFGNG	WNGLARYSYA	GSKQYGNHSG	RVGVGYRFLD
1051	GGGGTGSDDL	ANDSFIRQVL	DRQHFEVDGK	YHLFGSRGEL	AERSCHIGLG
1101	KIQSHQLGNL	MIQQAALKGN	IGYIVRFSDH	GHEVHSFFDN	HASHSDSDEA
1151	GSPVDGFSLY	RIHWDGYEHH	PADGYDGPQG	GGYPAPKGR	DIYSYDIRGV
1201	AQNIRLNLTD	NRSTGQRLAD	RPHNAGSMLT	QGVGDGFKRA	TRYSPELDRS
1251	GNAAEAFNGT	ADIVKNIIGA	AGEIVGAGDA	VQGISSEGSNI	AVMHGLGLLS
1301	TENKMARIND	LADMAQLKDY	AAAAIRDWAV	QNPNAAQGIE	AVSNIFMAAI
1351	PIKGIGAVRG	KYGLGGITAH	PIKRSQMGAI	ALPKGKSAVS	DNFADAAYAK
1401	YPSPHYSRNI	RSNLEQRYGK	ENITSSTVPP	SNGKNVKLAD	QRHPKTGVVF
1451	DGKGFPNFEK	HVKYDTLEHH	HHHH*		

AG983-741

ATGACTTTCGCGCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACACAGCGAAATCAGCA
GCAGTATCTTACGCCGTATCAAGAACGAAATGTGCAAGACAGAAGCATGCTCTGTGCCGGTCCGGATGACGTTGCG
GTTACAGACAGGGATGCCAAAATCAATGCCCCCCCGAATCTGCATACCGGAGACTTTCAAAACCCAAATGACGC
TACAAGAAATTTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGAC
ACAGGCGAATCCGTCGCGCAGCATATCCTTTCGCCAAGTGTATGGCAGAAAAGAACACGGCTATAACGAAAATACAAA
AATATACGGCGTATATGCGGAAGGAAGCGCTGAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTCGACGATGAG
CCGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGCCAGTAAAAGAAATCGGACACATCGATTTGGTCTCCCAT
ATTATTGCGCGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTCGCGCCGATGCGACGCTACACATAATGAATACG
AATGATGAAACCAAGAACGAAATGATGGTTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAAGCTGGCGTGGC
ATCGTCAATAACAGTTTGGAAACAACATCGAGGGCAGGCACTGCCGACCTTTCAAAATAGCCAAATTCGGAGGAGCAG
TACCGCCAAGCGTTGCTCGACTATTCGCGCGGTGATAAAAACAGACGAGGGTATCCGCCCTGATGCAACAGAGCGATTAC
GGCAACCTGTCTTACCACATCCGTAATAAAAACATGCTTTTTCATCTTTCGACAGGCAATGACGCCAAGCTCAGCCC
AACACATATGCCCATTGTCATTTATGAAAAGACGCTCAAAAAGGCATTATCACAGTTCGACGGCTAGACCGGATGAG
GGAGAAAAGTTCAAACGGGAAATGTATGGAGAACC GGGTACAGAACCGCTTGAGTATGGCTCCAACCCATTGCGGAATT
ACTGCCATGTGGTGCCTGTGCGCACCCCTATGAAGCAAGCGTCCGTTTCACCCGTACAAAACCCGATTCAAATGCGCGGA
ACATCCTTTTCGCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAAATACCCGTGGATGAGCAACGACAAC
CTGCGTACCACGTTGCTGACGACGGCTCAGGACATCGGTGCAGTCCGCGTGGACAGCAAGTTCCGGCTGGGGACTGCTG
GATCGCGGTAAAGCCATGAACCGGACCCGCTCCTTTCCGTTCCGGGACTTACCGCCGATACGAAAGGTACATCCGAT
ATTGCCTACTCCTTCGTAACGACATTTTCAGGCACGGCGGCTGATCAAAAAGGCGCGCACCCAACTGCAACTGCAC
GGCAACACACCTATACGGGCAAACCAATATCGAAGGCGGTTCCGCTGGTGTGTACGGCAACAACAATCGGATATG
CGCGTCAAACCAAAGGTGCGCTGATTTATAACGGGCGGGCATCCGGCGGAGCCTGAACAGCGACGGCATTGTCTAT
CTGGCAGATACCGACCAATCCGGCGCAAACGAAACCGTACACATCAAAGGCAGTCTGCAGCTGGACGGCAAAGGTACG
GTGATCACACGTTTGGGCAAACCTGCTGAAAGTGGACCGTACCGCGATTATCGGCGGCAAGCTGTACATGTCCGCACGC
GGCAAGGGGCGAGGCTATCTCAACAGTACCGGACGAGCTGTTCCTTCCCTTCCCTGAGTGCCGCAAAAATCGGGCAGGATTAT
TCTTTCTTCAACAACATCGAAACCGACGGCGGCTGCTGGCTTCCCTCGACAGCGTCAAAAACAGCGGGCAGTGAA
GGCGACAGCTGTCTATTATGTCGCTCGCGGCAATCGGGCACGGACTGCTTCGGCAGCGGCACATTCCGCGCCCGCC
GGTCTGAAACACGCCGTAGAACAGGGCGGCGAGCAATCTGGAAACCTGATGGTCAACTGGATGCTCCGAATCATCC
GCAACACCCGAGACGGTTGAAACTGCGGCAGCCGACCGCACAGATATGCCGGCATCCGCCCCCTACGGCGCAACTTTC
CGCGCAGCGGACGCGTACAGCATCGGAATGCCCGCCAGCGTGTACGCATCTTCAACAGTCTCGCCGCTACCGCTTAT
GCCGACAGTACCGCCGCCATGCCGATATGCAGGGACGCGCTGAAAGCCGTATCGGACGGGTTGGACCAACAACGGC
ACGGGTCTGCGCGTCATCGCGCAAACCCAAACAGGACGGTGAACGTGGGAACAGGGCGGTGTTGAAGGCAAAAATGCCG
GGCAGTACCCAAACCGTCCGGCATTGCCCGGAAAACCGCGGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGC
AGCACATGGAGCGAAAACAGTGCAAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAGGATGCGGGC
GATATCGGCTATCTCAAAGGCTGTCTCTTACGGACGCTACAAAACAGCATCAGCCGAGCACCAGTCCGGTGGCGGAA
CATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGGCGCAACG
GGAGATTTGACGGTCAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCGGAAAAGGCAGTGTCTTG
GGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGCTGCGGATCGCGGGTCTGAAGCTGTGCGAACCCCTTGAGCGAT
AAAGCCGCTCTGTTGCAACGGCGGGCGTGGAAACGGACCTGAAACGGACGCGACTACACGGTAAACGGGCGGCTTTACC
GGCGGACTGCAGCAACCGGCAAGACGGGGCACGCAATATGCCGCACACCCGCTGTTGCGGGCCTGGGCGCGGAT
GTGCAATTCGGCAACCGGCTGGAACGGCTTGGCACGTTACAGCTACCGCGGTTCCAAAACAGTACGGCAACCAACAGCGGA
CGAGTCCGGCTAGGCTACCGGTTCTCGAGGGATCCGGAGGGGTTGGTGTGCGCGCCGACATCGGTGCGGGGCTTGCC
GATGCATAACCGCACCGCTCGACCATAAAGACAAAGGTTTGCAGTCTTTGACGCTGGATCAGTCCGTCAGGAAAAC
GAGAACTGAAGCTGGCGGCACAAGGTGCGGAAAACCTTATGAAAACGTTGACAGCCTCAATACGGGCAAAATGAAG
AACGACAAGGTGAGCGTTTCGACTTTATCCGCCAAATCGAAGTGGACGGGACGCTCATTACCTTGGAGAGTGGAGAG
TTCCAAATATACAAAAGCCATTCGCGCTTAACCGCTTTTCAGACCGAGCAAAATACAAGATTCGGAGCATTCCGGG
AAGATGGTTGCGAAACGCCAGTTTCAAGATCGGCGACATAGCGGGCGAACAATACATCTTTTGACAAGCTTCCCGAAGGC
GGCAGGGGACATATCGCGGACGGCGTTCCGTTTTCAGACGATGCCGGCGGAAAACGACCTACACCATAGATTTGCC
GCCAAGCAGGGAAACGGCAAATCGAACATTTGAAATCGCCAGAACTCAATGTGACCTGGCCGCGCCGATATCAAG
CCGGATGGAACCGCATGCGCTCATCAGCGGTTCCGCTCTTACAACCAAGCCGAGAAAGGCAGTTACTCCCTCGGT

ATCTTTGGCGGAAAAGCCAGGAAGTTGCCGGCAGCGGGAAGTGAAAACCGTAAACGGCATAACGCCATATCGGCCTT
GCCGCCAAGCAACTCGAGCACCACCACCACCACCCTGA

1 MTSAPDFNAG GTGIGSNSRA TTAKSAAVSY AGIKNEMCKD RSMLCAGRDD
51 VAVTDRDAKI NAPPNHLTG DFPNPNDAYK NLINLKPAIE AGYTGRGVEV
101 GIVDTGESVG SISFPELYGR KEHGYNENYK NYTAYMRKEA PEDGGGKDIE
151 ASFDDEAVIE TEAKPTDIRH VKEIGHIDLK SHLIIGRSVD GRPAGGIAPD
201 ATLLHMTND ETKNEMMVAA IRNAWVKLGE RGVRIVNNSE GTTSRAGTAD
251 LFQIANSEEQ YRQALLDYSG GDKTDEGIRL MQQSDYGNLS YHIRNKNMLF
301 IFSTGNDQA QPNTYALLPF YEKDAQKGI TVAGVDRSGE KFKREMYGEP
351 GTEPLEYGSN HCGITAMWCL SAPYEASVRF TRINPIQIAG TSFSAPIVTG
401 TAALLLQKYP WMSNDNLRIT LLTTAQDIGA VGVDSKFGWG LLDAGKAMNG
451 PASFPFGDFT ADTKGTS DIA YSFRNDISGT GGLIKKGGSQ LQLHGNNTYT
501 GKTIIEGGSV VLYGNKSDM RVETKALYI NGAASGGSLN SDGIVYLADT
551 DQSGANETVH IKGSLQLDGL GTLYTRLGKL LKVDGTALIG GKLYMSARGK
601 GAGYLNSTGR RVPFLLSAKI GDYSFFTNI ETDGGLLASL DSVEKTAGSE
651 GDTLSYYVRR GNAARTASAA AHSAPAGLKH AVEQGGSNLE NLMVELDASE
701 SSATPETVET AAADRTDMPG IRPYGATFRA AAAVQHANA DGVRIFNSLA
751 ATVYADSTAA HADMQRRLL AVSDGLDENG TGLRVIAQTQ QDGGTWEQGG
801 VEGKMRGSTQ TVGLAARTGE NTTAAATLGM GRSTWSENSA NAKTDSISLF
851 AGIRHDAGDI GYLKGLFSYG RYKNSISRST GADEHARGSV NGTLMQLGAL
901 GGVNVFFAAT GDLTVEGGLR YDLLKQDAFA EKGSAWGWSG NSLTFEGLTVG
951 LAGLKLQPL SDKAVLFATA GVERDLNDRD YTVTGGPTGA TAATGKTGAR
1001 NMPETRLVAG LGADVEFGNG WNLARYSYA GSKQYGNHSG RVGVGYRFLK
1051 GSGGGVAAD IGAGLADALT APLDHRDKGL QSLTLDQSVR KNEKLLKAAQ
1101 GAERTYNGND SLNTGKLNKND KVSRRDFIRQ IEVDGQLITL ESGEFQVYKQ
1151 SHSALTAFTQ EQIQDSEHSG KMKVAKRQFRI GDIAGEHTSF DKLPEGGRAT
1201 YRGTAFGSDD AGGKLTYYTID FAAKQNGKI EHLKSPELNV DLAAADIKPD
1251 GKRHAIVSGS VLYNQAEKGS YSLGIFGGKA QEVAGSARVK TVNGIRHIGL
1301 AAKQLEHHHH HH*

AG983-961

ATGACTTCTGCGCCCGACTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCA
GCAGTATCTTACCGCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCCGCTCGGGATGACGTTGGC
GTTACAGACAGGGATGCCAAAATCAATGCCCCCCCGAATCTGCATACCGGAGACTTTCCAAAACCAAATGACGCA
TACAAGAATTTGATCAACCTCAAACCTGCAATGAAGCAGGCTATACAGGACCGGGGTAGAGGTAGGTATCGTCGAC
ACAGGCGAATCCGTCCGCAGCATATCTTTCCCGAACTGTATGGCAGAAAAGAACACGGCTATAACGAAAATTACAAA
AATATACCGCGTATATGCCGAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATGAAGCTTTCTTCGACGATGAG
GCCGTTATAGAGACTGAAGCAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATTTGGTCTCCCAT
ATTATTGGCGGCGTTCCTGGAGCGGACCTGCAGGCGGTATTGCCCGGATGCGACGCTACACATAATGAATACG
AATGATGAAACCAAGAACGAAATGATGGTTGCAGCCATCCGCAATGCATGGGTCAAGCTGGCGAACCTGGCGTCCG
ATCGTCAATAACAGTTTGGAAACAACATCGAGGCGAGGCACTGCCGACCTTTTCCAAATAGCCAATTCGGAGGAGCAG
TACCGCCAAGCGTTGCTCGACTATTCGGCGGTGATAAAAACAGACGAGGGTATCCGCTGATGCAACAGAGCGATTAC
GGCAACCTGTCTACCACATCCGTAATAAAAACATGCTTTTCATCTTTTCGACAGGCAATGACGCACAAGCTCAGCCC
AACACATATGCCCTATTGCCATTTTATGAAAAGACGCTCAAAAAGGCATTATCACAGTCCGAGGCGTAGACCGCAGT
GGAGAAAAGTTCAAACGGGAAATGTATGGAGAACCGGTACAGAACCGCTTGAGTATGGCTCCAACCATTTGCCGAAT
ACTGCCATGTGGTGCCTGTCCGCACCCTATGAAGCAAGCGTCCGTTTCAACCCGTACAAAACCGATTCAAATTCGGGA
ACATCCTTTTCCGACCCATCGTAACCGGCACGGCGGCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAAC
CTGGTACCAGTGTGCTGACGACGGCTCAGGACATCGGTGCAGTCCGGCTGGACAGCAAGTTCCGCTGGGACTGCTG
GATCGGGTAAAGCCATGAACGGACCCCGCTCTTTCCGTTCCGGCGACTTTACCGCCGATACGAAAGGTACATCCGAT
ATTGCCTACTCCTTCCGTAACGACATTTACGGCACGGGCGGCTGATCAAAAAGGCGGCAGCCAACGTCAACTGCAC
GGCAACAACCTTATACGGGCAAAACCATTTATCGAAGGCGGTTCCGTTGGTGTGTACGGCAACAACAATCGGATATG
CGCTCGAAACCAAAGGTGCGCTGATTTATAACGGGCGGCATCCGGCGGACGCTGAACAGCGACGGCATTGTCTAT
CTGGCAGATACCGACCAATCCGGCGCAACGAAACCGTACACATCAAAGGCAAGTCTGCAGCTGGACGGCAAAAGTACG
CTGTACACACGTTTGGGCAAACTGCTGAAAGTGGACGGTACGGCGATTATCGGGCGCAAGCTGTACATGTCCGGCAGC
GGCAAGGGGGCAGGCTATCTCAACAGTACCGGACGACGTTTCCCTTCTGAGTCCGCCAAAATCGGGCAGGATTAT
TCTTTCTTCAAAAACATCGAAACCGACGGCGGCTGCTGGCTTCCCTGACAGCGCTCGAAAACAGCGGGCAGTGAA
GGCGACAGCTGTCTATTATGTCCGTCGGGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATTCGGCGCCCGC
GGTCTGAAACACGCCGTAGAACAGGGCGGACGAACTGGA AAAACCTGATGCTTGA AACTGCAATGCCCTCGCAATCAGC

CATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTGCCGCAACG
GGAGATTTGACGGTCAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTGCGCCGAAAAGGCAGTGTGTTG
GGCTGGAGCGGCAACAGCCTCACTGAAGGCACGCTGGTTCGGACTCGCGGGTCTGAAGCTGTGCGCAACCCCTGAGCGAT
AAAGCCGTCCTGTTTGCACGGCGGGCGTGAACCGGACCTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACC
GGCGGACTGCAGCAACCGGCAAGACGGGGCACGCAATATGCCGCACACCCGTCGTTGGTTCGCCGCGCTGGGCGGGAT
GTCGAATTCGGCAACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGTTCCAAACAGTACGGCAACCACAGCGGA
CGAGTCGGCGTAGGCTACCGGTTCCCTCGAGGGTGGCGGAGGCACTGGATCCGCCACAAACGACGACGATGTTAAAAA
GCTGCCACTGTGGCCATTGCTGCTGCCTACAACAATGGCCAAAGAAATCAACGGTTTCAAAGCTGGAGAGACCATCTAC
GACATTGATGAAGACGGCACAATTACCAAAAAGACGCAACTGCAGCCGATGTTGAAGCCGACTTTAAAGGTCTG
GGTCTGAAAAAAGTCTGACTAACCTGACCAAAACCGTCAATGAAAAACAAACAAACGTCGATGCCAAAGTAAAGCT
GCAGAATCTGAAATAGAAAAGTTAACAACCAAGTTAGCAGACACTGATGCCGCTTTAGCAGATACTGATGCCGCTCTG
GATGCAACCACCAACGCTTTGAATAAATTTGGGAGAAAATATAACGACATTTGCTGAAGAGACTAAGACAAATATCGTA
AAAATTGATGAAAAATTTAGAAGCCGTTGGCTGATACCGTCGACAGCAAGCATGCCGAAGCATTCAACGATATCGCCGATTC
TTGGATGAAACCAACTAAGGCAGACGAAGCCGTCAAAACCGCCAAATGAAGCCAAACAGACGGCCGAAGAAACCAA
CAAAACGTCGATGCCAAAGTAAAGCTGCAGAACTGCAGCAGGCAAGCCGAAGCTGCCGCTGGCACAGCTAATACT
GCAGCCGACAAGGCCGAAGCTGTGCTGCAAAAGTTACCGACATCAAAGCTGATATCGCTACGAACAAAGATAATATT
GCTAAAAAGCAACAGTGGCCGCTGTACACCAGAGAGAGTCTGACAGCAAATTTGTGAGAATTGATGGTCTGAAC
GCTACTACCGAAAAATTTGGACACAGCTTGGCTTCTGTGAAAAATCCATTGCCGATCAGGATACTCGCCTGAACGGT
TTGGATAAAACAGTGTTCAGACTGCGCAAAGAAACCCGCAAGGCCCTGTCAGAAACAAGCCGCGCTCTCCGGTCTGTTC
CAACCTTACAACGTTGGTTCGTTCAATGTAACGGCTGCAGTCCGCGCTACAATCCGAATCCGCGCTCTCCGGTCTGTTC
ACCGGCTTCCGCTTTACCGAAAACCTTTGCCCCAAAGCAGGCGTGGCAGTCCGCACTTCGTCGGTTCTTCCGAGCC
TACCATGTCGGCGTCAATTACGAGTGGCTCGAGCACCACCACCACCACCCTGA

1	MTSAPDFNAG	GTGIGSNSRA	TTAKSAAVSY	AGIKNEMCKD	RSMLCAGRDD
51	VAVTDRDAKI	NAPPPNLHTG	DFPNPNDAYK	NLNLKPAIE	AGYTGRGVEV
101	GIVDTGESVG	SISFPELYGR	KEHGYNENYK	NYTAYMRKEA	PEDGGGKDIE
151	ASFDDEAVIE	TEAKPTDIRH	VKEIGHIDLIV	SHIIGGRSVD	GRPAGGIAPD
201	ATLHLMFTND	ETKNEMMVA	IRNAWVKLGE	RGVRIVNNNSF	GTTSRAGTAD
251	LFQLANSEBQ	YRQALLDYSG	GDKTDEGIRL	MQQSDYGNLS	YHIRKNMLF
301	IFSTGNDAQA	QPNFYALLPF	YEKDAQGLII	TVAGVDRSGE	KFKREMYGEP
351	GTEPLEYGSN	HCGITAMWCL	SAPYEASVRF	TRTNPQIAG	TSFSAPIVTG
401	TAALLLQKYP	WMSNDNLRTT	LLTTAQDIGA	VGVDKFGWG	LLDAGKAMNG
451	PASFPPGDFT	ADTKGTSDLA	YSFRNDISGT	GGLIKKGGSQ	LQLHGNNTYT
501	GKTIIEGSSL	VLVGNKSDM	RVETKCALIY	NGAASGGLN	SDGIVYLADT
551	DQSGANETVH	IKGSLQLDGK	GTLYTRLGKL	LKVDGTAIIG	GKLYMSARGK
601	GAGYLNSTGR	RVPFLSAAKI	GQDYSFFTNI	ETDGLLASL	DSVEKTAGSE
651	GDTLSYVRR	GNAARTASAA	AHSAPAGLKH	AVEQGSNLE	NLMVELDASE
701	SSATPETVET	AAADRTDMPG	IRPYGATFRA	AAAVQHANA	DGVRIFFNSLA
751	ATVYADSTAA	HADMQRRRK	AVSDGLDHNG	TGLRVIAQTQ	QDGGTWEQGG
801	VEGKMRGSTQ	TVGLIAKTGE	NTTAAATLGM	GRSTWSENSA	NARTDSISLF
851	AGIRHDAGDI	GYLKGLFSYG	RYKNSISRST	GADEHAEGSV	NGTLMQLGAL
901	GGVNVFFAAT	GDLTVEGGLR	YDLLKQDAFA	EKGSALGWSG	NSLTEGTLVG
951	LAGLRLSQPL	SDKAVLFATA	GVERDLNGRD	YTVTGGFTGA	TAATGKTGAR
1001	NMPHTRLVAG	LGADVEFGNG	WNGLARYSYA	GSKQYGNHSG	RVGVGYRFLF
1051	GGGGTGSATN	DDVVKAAATV	ALAAAYNNGQ	EINGFKAGET	IYDIDEDGTI
1101	TKKDATAADV	EADDFKGLGL	KKVVTNLTKT	VNENKQNVDA	KVKAASEEIE
1151	KLTTKLADTD	AALADTDAAL	DATTNALNKL	GENITTFABE	TKTNIVKIDE
1201	KLEAVADTVD	KHABAFNDIA	DSLDEFNTKA	DEAVKTANEA	KQTAERTKQN
1251	VDARVKAET	AAGKAEAAAG	TANTAADKAE	AVAAKVTDIK	ADIATNKDNI
1301	AKKANSADV	TRESDSKFFV	RIDGLNATTE	KLDTRLASAE	KSIADHDTRL
1351	NGLDKTVSDL	RKETROGLAE	QAALSGLFQP	YNVGRFNVTA	AVGYKSESA
1401	VAIGTGFRFT	ENFAAKAGVA	VGTSSGSSAA	YHVGVNVEWL	EHHHHHH*

AG983-961c

ATGACTTCTGCGCCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCA
GCAGTATCTTACGCGGTATCAAGAACGAAATGTGCAAAGACAGAAGCATGCTCTGTGCGCGGTGGGATGACGTTGCC
GTTACAGACAGGGATGCCAAAATCAATGCCCCCCCGAATCTGCATACCGGAGACTTTCAAAACCAAATGACGCA
TACAAGAATTTGATCAACCTCAAACCTGCAATTGAAGCAGGCTATACAGGACCGGGGTAGAGGTAGGTATCGTCGAC
ACAGGCGAATCCGTCGGCAGCATATCCTTTCCCGAAGTGTATGCGAGAAAGAACACGGCTATAACGAAAATTAACAA
AATATAACGGCTATATGCGGAAGGAAGCGCCTGAAGACGGAGGCGGTAAGACATGAAAGCTTCTTTCGACGATGAG
GCCGTTATAGAGACTGAAGCAAAGCCGACGGATATCCGACCCTAAAAGAAATCGGACACATCGATTTGGTCTCCCAT
ATTTATGGCGGGCGTTCGCTGGACGGCAGACCTGCAGCCGTTATTCGCGCGGATGCGACGCTACACATAATGAATACG
AATGATGAAACCAAGAACGAAATGATGGTTGCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTTGGCCTGCC
ATCGTCAATAACAGTTTGAACAACATCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAAATTCGGAGGAGCAG
TACCGCCAAGCGTTGCTCGACTATTCCGCGGTGATAAAACAGACGAGGATATCCGCTGATGCAACAGAGCGATTAC
GGCAACCTGTCTTACCACATCCGTAATAAAAACATGCTTTTCATCTTTTCGACAGGCAATGACGCACAAGCTCAGCC

AACACATATGCCCTATTGCCATTTTATGAAAAAGACGCTCAAAAAGGCATTATCACAGTCGCGAGGCGTAGACCCGAGT
GGAGAAAAGTTCAAACGGGAAATGTATGGAGAACCAGGTTACAGAACCGCTTGAGTATGGCTCCAACCATTCGGGAATT
ACTGCCATGTGGTGCCTGTTCGGCACCCCTATGAAGCAAGCGTCCGTTTACCCCGTACAAAACCCGATTCAAAATTGCCGGA
ACATCCTTTTCCGCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAAAC
CTGCGTACCACGTTGCTGACGACCGCTCAGGACATCGGTGCAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTGCTG
GATGCGGGTAAGGCCATGAACGGACCCCGCTCCTTCCGTTCGGCGACTTACCGCCGATACGAAAGGTACATCCGAT
ATTGCCACTCCTTCCGTAACGACATTTACGGCACGGGCGGCTGATCAAAAAGGCGGCAGCCAATGCAACTGCAC
GGCAACAAACCTTATACGGGCAAAACCATTATCGAAGGCGGTTCCGTGGTGTGTACGGCAACAAATCCGATATG
CGCGTCGAAACCAAAGGTGCGCTGATTTATAACGGGGCGGCATCCGCGGCAGCCTGAACAGCGACGGCATTGTCTAT
CTGGCAGATACCGACCAATCCGGCGCAACGAAACCGTACACATCAAAGGCAGTCTGCAGCTGGACGGCAAAGGTACG
CTGTACACACGTTTCGGCAAACCTGCTGAAAGTGGACCGTACGGCGATTTATCGGCGGCAAGCTGTACATGTCCGACGC
GGCAAGGGGGCAGGCTATCTCAACAGTACCGGACGACGTTCCCTTCCCTGAGTGCAGCCAAAATCGGGCAGGATTTAT
TCTTTCTTCAACAACATCGAAACCGACGGCGGCTGCTGGCTTCCCTCGACAGCGTCGAAAAACAGCGGGCAGTGAA
GGCGACACGCTGTCTTATTATGTCCGTTCGGCGCAATGGCGCACGGACTGCTTCGGCAGCGGCACATTCGGCCCGCC
GGTCTGAAACACGCCGTAGAACAGGGCGGCAGCAATCTGAAAACCTGATGGTCGAACTGGATGCCTCCGAATCATCC
GCAACACCCGAGACGGTTGAAACTGCGGGCAGCCGACCGCACAGATATGCCGGGCATCCGCCCTACGGCGCAACTTTC
CGCGCAGCGGCAGCCGTACAGCATGCGAATGCCGCCGAGCGGTGTACGCATCTTCAACAGTCTCCGCGCTACCGTCTAT
GCCGACAGTACCGCCGCCATGCGGATATGCAAGGACGCGCCTGAAAGCCGTATCGGACGGGTTGGACCACAACGGC
ACGGGTCTGCGCGTATCGCGCAACCCAAACAGGACGGTGGAAACGTTGGAAACAGGGCGGTGTGAAGGCAAAATGCGC
GGCAGTACCCAAACCGTTCGGCATTCGGCGGAAAACCGCGCAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGC
AGCACATGGAGCGAAACAGTGCBAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGC
GATATCCGCTATCTCAAAGCCGTGTTCTCCTACGGACGCTACAAAACAGCATCAGCCGACGACCGGTGCGGACGAA
CATGCGGAAGGCAGCGTCAACGGCACGCTGATGCAGCTGGGCGCACTGGGCGGTGTCAACGTTCCGTTTCCGCAACG
GGAGATTTGACGGTGAAGGGCGTCTGCGCTACGACCTGCTCAACAGGATGCATTCGCGGAAAAGGCAAGTGTG
GGCTGGAGCGGCAACAGCCTCACTGAAGGCAGCTGGTGGACTCGCGGTCTGAAGCTGTGCAACCCCTTGGAGCGAT
AAAGCCGCTCTGTTTGAACGGCGGGCGTGGAAACGGCAGCTGAACGGACGCGACTACACGGTAACGGGCGGCTTTACC
GGCGGACTGCAGCAACCGGCAAGACGGGGCACGCAATATGCCGCACACCCGCTCGGTTGCCGGCCTGGGCGCGGAT
GTCGAATTCGGCAACCGCTGGAACCGCTTGGCACGTTACAGCTACGCCGTTTCCAAAACAGTACGGCAACCACAGCGGA
CGAGTCGGCGTAGGCTACCGGTTCTCTGAGGGTGGCGGAGGCATGGATCCGCCACAACAGCAGCAGATGTTAAAAAA
GCTGCCACTGTGGCCATGCTGCTGCCTACAACAATGGCCAAGAAATCAACCGTTTCAAAGCTGGAGAGACCATCTAC
GACATGTGTAAGACGGCACAATTACCAAAAAGACGCAACTGCAGCCGATGTTGAAGCCGACGACTTTAAAGGTCTG
GGTCTGAAAAAGTCTGACTAACCTGACAAAACCGTCAATGAAAACAAAACAGTGCATGCCAAAAGTAAAAGCT
GCAGAACTGAAATGAAAAAGTTAACAACCAAGTTAGCAGACACTGATGCCGCTTTAGCAGATACTGATGCCGCTCTG
GATGCAACCACCAACCGCTTGAATAAATGGGAGAAAATATAACGACATTTGCTGAAGAGACTAAGACAAAATATCGTA
AAAAATGATGAAAAATTAGAAGCCGTGGCTGATACCCGTCGACAAGCATGCCGAAGCATTCAACGATATCGCCGATTC
TTGGATGAAACCAACACTAAGGCAGACGAAGCCGTCAAACCCGCAATGAAGCCAAACAGACGGCCGAA/AAACAAA
CAAAACGTCGATGCCAAAGTAAAAGCTGCAGAACTGCAGCAGGCAAGCCGAAGCTGCCGCTGCCACAGCTAATACT
GCAGCCGACAAGGCCAAGCTGTCCGTGCAAAAGTTACCGACATCAAAGCTGATATCGCTACGAACAAGATAATATT
GCTAAAAAAGCAAAACAGTGCAGCAGCTGTACACCAGAGAAGAGTCTGACAGCAAAATTTGTGAGAAATGATGGTCTGAAC
GCTACTACCGAAAATTTGGACACACGCTTGGCTCTGCTGAAAATCCATTTGCCGATCACGATACCTCGCTGCAAGCGGT
TTGGATAAAAACAGTGTACAGCTGCGCAAGAAAACCGCAAGGCCCTGCGAACAAGCCGCGCTCTCCGCTCTGTTTC
CAACCTTACAACGTGGGTCTCGAGCACCACCACCACCCTGA

1	MTSAPDFNAG	GTGIGSNSRA	TTAKSAAVSY	AGIKNEMCKD	RSMLCAGRDD
51	VAVTDRDAKI	NAPPPNLHTG	DFPNPNDAYK	NLINLKPALIE	AGYTGRGVEV
101	GIVDTGESVG	SISFPELYGR	KEHGYNENYK	NYTAYMRKEA	PEDGGGKDIE
151	ASFDDEAVIE	TEAKPTDIRH	VKEIGHIDLIV	SHIIGGRSVD	GRPAGGIAPD
201	ATLHIMNTND	ETKNEMMVAA	IRNAWVKLGE	RGVRIVWNSF	GTTBRAGTAD
251	LFQIANSEEQ	YRQALLDYSG	GDKTDEGIRL	MQQSDYGNLS	YHIRNKNMLF
301	IFSTGNDQAQ	QPNTYALLPF	YEKDAQKGI	TVAGVDRSGE	KFKREMYGEP
351	GTEPLEYGSN	HCGITAMWCL	SAPYEASVRF	TRTNPIQIAG	TSFSAPIVTG
401	TAALLLQKYP	WMSNDNLRTT	LITTAQDIGA	VGVDKFPWGW	LLDAGKAMNG
451	PASFPFGDFT	ADTRGTSIDIA	YSFRNDISQT	GGLIKKGGSQ	LQLHGNNTYT
501	GKTIIEGGS	VLYGNKSDM	RVETKGALY	NGAASGGSLN	SDGIVYLADT
551	DQSGANETVH	IKGSLQLDGG	GTLYTRLGKL	LKVDGTAIIG	GKLYMSARGK
601	GAGYLNSTGR	RVPFLSAAKI	GQDYSFFFTNI	ETDGGLLASL	DSVEKTAGE
651	GDTLSYVYRR	GNAARTASAA	AHSAPAGLKH	AVEQGGSNLE	NLMVELDASE
701	SSATPETVET	AAADRTDMPG	IRPYGATFRA	AAAVQHANA	DGVRIFNSLA
751	ATVYADSTAA	HADMQRRLK	AVSDGLDHNG	TGLRVIAQTQ	QDGGTWEQGG
801	VEGKMRGSTQ	TVGIAAKTGE	NTTAAATLGM	GRSTWSENSA	NAKTDSISLF
851	AGIRHDAGDI	GYLKLFSYG	RYKNSISRST	GADEHAEQSV	NGTLMQLGAL
901	GGVNVFFAAT	GDLTVEGGLR	YDLLKQDAPA	EKGSALGWSG	NSLTEGTLVG
951	LAGLKLSQL	SDKAVLFATA	GVERDLNGRD	YTVTGGFTGA	TAAATGKTGAR
1001	NMPHTRLVAG	LGADVFEFNG	WNGLARYSYA	GSKQYGNHSG	RVGVGYRFLF
1051	GGGGTGSATN	DDDVKKAATV	AIAAAYNNGQ	EINGFKAGET	IYDIDEDGTI
1101	TKKDATAADV	EADDFKGLGL	KKVVTNLTKT	VNENKQNVDA	KVKAASEIE
1151	KLTKRLADTD	AALADTDAAL	DATNALNLK	GENITTFAE	TRTNIVKIDE

1201 KLEAVADTVD KHAEAFNDIA DSDLDETNTKA DEAVKTANE AQTAEETKQM
 1251 VDAKVKAET AAGKAEAAAG TANTAADKAE AVAAKVTDIK ADIATNKDNI
 1301 AKKANSADVY TREESDSKFV RIDGLNATTE KLDTRLASAE KSIADHDTRL
 1351 NGLDKTVSDL RKETRQGLAE QAALSGLFQP YNVGLEHHHH HH*

实施例 4- Δ G741 的杂交体

蛋白质 741 具有如下序列:

1 VNRTAFCCLS LTTALILTAC SSGGGVVAAD IGAGLADALT APLDHKDKGL
 51 QSLTLDQSVR KNEKLLAAQ GAERTYGNVD SLNTGKLNK KVSRLFDFIRQ
 101 IEVDGQLITL ESSEFQVYKQ SHSALTAFQT EQIQDSEHSG KMVAKRQFRI
 151 GDIAGEHTSF DKLPEGGRAT YRGTAFGSD AGGKLYTYTID FAAKQNGKI
 201 EHLKSPELNV DLAAADIKPD GKRHAVISGS VLYNQAEKGS YSLGIFGGKA
 251 QEVAGSAEVK TVNGIRHIGL AAKQ*

因此, Δ G741 具有如下基本序列:

VAAD IGAGLADALT APLDHKDKGL
 QSLTLDQSVR KNEKLLAAQ GAERTYGNVD SLNTGKLNK KVSRLFDFIRQ
 IEVDGQLITL ESSEFQVYKQ SHSALTAFQT EQIQDSEHSG KMVAKRQFRI
 GDIAGEHTSF DKLPEGGRAT YRGTAFGSD AGGKLYTYTID FAAKQNGKI
 EHLKSPELNV DLAAADIKPD GKRHAVISGS VLYNQAEKGS YSLGIFGGKA
 QEVAGSAEVK TVNGIRHIGL AAKQ*

将 Δ G741 直接融合于蛋白质 961、961c、983 和 ORF46.1 的符合读框的上游:

AG741-961

ATGGTCGCCCGGACATCGGTGCGGGGCTTGCCGATGCACCTAACCGCACCGCTCGACCATAAAGACAAAGGTTTGCGAG
 TCTTTGACGCTGGATCAGTCCGTGAGAAAACGAGAACTGAAGCTGGCGGCACAAGGTGCGGAAAAACTTATGGA
 AACGGTGACAGCCTCAATACGGGCAATTTGAAGAACGACAAGGTGAGCGGTTTCGACTTTATCCGCCAAATCGAAGTG
 GACGGGCAGCTCATTACCTTGGAGAGTGGAGAGTTCCAAGTATACAAACAAAGCCATTCCGCCTTAACCGCCTTTTCAG
 ACCGAGCAAATACAAGATTTCGGAGCATTCGGGAAGATGGTTGCGAAAACGCCAGTTTCAGAAATCGGCGACATAGCGGGC
 GAACATACATCTTTTGACAAGCTTCCCGAAGGCGGCGAGGGCGACATATCGCGGGACGGCGTTCCGGTTTCAGACGATGCC
 GCGGAAAACTGACCTACACCATAGATTTTCGCCCGCAAGCAGGAAACCGGCAAAATCGAACATTTGAAATCGCCAGAA
 CTCAATGTCGACCTGGCCGCGCCGATATCAAGCCGGATGGAAAACGCCATGCCGTCATCAGCGGTTCCGTCCTTAC
 AACCAAGCCGAGAAAGGCAGTTACTCCCTCGGTATCTTTGGCGGAAAAGCCAGGAAGTTGCCGGCAGCGCGGAAGTG
 AAAACCGTAAACCGCATACGCCATATCGGCCCTTGCCGCAAGCAACTCGAGGGTGGCGGAGGCACTGGATCCGCCACA
 AACGACGACGATGTTAAAAAGCTGCCACTGTGGCCATTGCTGCTGCTTACAACAATGGCCAAAGAAATCAACGGTTTC
 AAAGCTGGAGAGACCTACGACATTTGATGAAGAGCGGCAATTAACAAAAAGAGCGCAACTGCAGCCGATGTTGAA
 GCCGACGACTTTAAAGGCTCGGCTGTAAGAAAGTCTGACTAACCTGACCAAAACCGTCAATGAAAAACAACAACAA
 GTCGATGCCAAAGTAAAGCTGCAGAACTGAAATAGAAAAGTTAACAACCAAGTTAGCAGACACTGATGCCGCTTTTA
 GCAGTACTGATGCCGCTCTGGATGCAACCACCAACGCCCTTGAATAAATTTGGGAGAAAATATAACGACATTTGCTGAA
 GAGACTAAGACAAATATCGTAAAAATTTGATGAAAAATTAGAAGCCGTGGCTGATACCGTCGACAAGCATGCCGAAGCA
 TTCAACGATATCGCCGATTCATTGGATGAAACCAACTAAGGCAGACGAAGCCGTCAAAACCGCCAATGAAGCCAAA
 CAGACGGCCGAAGAAACCAAAACGCTCGATGCCAAAAGTGAAGAGCTGCAGAAACTGCAGCAGGCAAAAGCCGAAGCT
 GCCGCTGGCACAGCTAATACTGCAGCCGACAAGGCCGAAGCTGTCCGTGCAAAAAGTTACCGACATCAAAGCTGATATC
 GCTACGAACAAAGATAATATGCTAAAAAGCAACAGTCCGACGCTGACACCAGAGAAGAGTCTGACAGCAAAATTT
 GTCAGAAATGGTCTGAACGCTACTACCGAAAAATTTGACACACGCTTGGCTTCTGCTGAAAAATCCATTGCCGAT
 CACGATACCTCGCTGACCGGTTTGGATAAAACAGTGTGACAGCTGCGCAAAGAAACCGCCAAGGCCCTTGCAGAACAA
 GCCGCGCTCTCCGTTCTGTTCCAACCTTACAACGTTGGTTCGGTTCAATGTAACGCTGCAGTCCGCGGCTACAAATCC
 GAATCGGCAGTCCGATCGGTACCGGCTTCCGCTTTACCGAAAACTTTGCCGCAAGCAAGCCGCTGGCAGTCCGCACT
 TCGTCCGGTTCTTCCGCGACCTACCATGTCCGCGTCAATTACGAGTGGCTCGAGCACCACCACCACCACCCTGA

1 MVAADIGAGL ADALTAFLDH KDKGLQSLTL DQSVRKNEK KLAAQGAERT
 51 YNGDSLNTG KLKNDKVSFR DFIRQIEVDG QLITLESSEF QVYKQSHSAL
 101 TAFQTEQIQD SEHSGKRVAK RQFRIGDIAG EHTSFDKLPE GGRATYRGT
 151 FGSDDAGGKL TYTIDFAAQ GNGKIEHLKS PELNVDLAAA DIKPDGKRHA
 201 VISGSVLYNQ AEKGSYSLGI FGGKAQEVAG SAEVKTVNGI RHIGLAAKQL
 251 EGGGGTGSAT NDDDVKKAAT VALAAAYNNG QEINGFKAGE TIYDIDEDGT
 301 ITKKDATAAD VEADDFKGLG LKKVVTNLTK TVNENKQNV D ARVKAASEI
 351 EKLTTKLDAT DAALADTDAE LDATTNALNK LGENITTFAE ETKTNIVRLD
 401 EKLEAVADTV DKHAEAFNDI ADSDLDETNTK ADEAVKTANE AKQTAEETKQ

451 NVDKVKAAE TAAGKAEAAA GTANTAADKA EAVAARKVTDI KADIATNKDN
 501 IAKKANSADV YTREESDSKF VRIDGLNATF EKLDTRLASA EKSIADHDTR
 551 LNGLDKTVSD LRKETRQGLA EQAALSGLFQ PYNVGRFNVF AAVGGYKSES
 601 AVAIGTGFRF TENFAAKAGV AVGTSSGSSA AYHVGVNYEW LEHHHHHH*

AG741-961c

ATGGTCGCCCGACATCGGTGCGGGGCTTGCCGATGCACTAACCGCACCGCTCGACCATAAAGACAAAGGTTGCGAG
 TCTTTGACGCTGGATCAGTCCGTCAGGAAAAACGAGAACTGAAGCTGGCGGCACAAGGTGCGGAAAAAAGTTATGGA
 AACGGTGACAGCCTCAATACGGGCAAATTAAGAAACGACAAGGTCAGCCGTTTCGACTTTATCCGCCAAATCGAAGTG
 GACGGGCAGCTCATTACCTTGGAGAGTGGAGAGTTCAGTATACAAACAAAGCCATTCCGCCCTAACCGCCTTTTCAG
 ACCGAGCAAATACAAGATTCCGGAGCATTCCGGGAAGATGGTTGCGAAAACGCCAGTTTCAGAAATCGGGCAGATAGCGGGC
 GAACATACATCTTTTGACAAGCTTCCCAGAGCGGCAGGGCGACATATCGCGGGACGGCGTTTCGGTTTCAGACGATGCC
 GCGGAAAACTGACCTACACCATAGATTTCCGCCCAAGCAGGGAACCGGCAAATCGAACATTGAAATCGCCAGAA
 CTCAATGTCGACCTGGCCCGCCGATATCAAGCCGGATGGAACCGCCATGCCGTCATCAGCGGTTCCGTCCCTTAC
 AACCAAGCCGAGAAAGGCAGTTACTCCCTCGGTATCTTTGGCGGAAAAGCCAGGAAGTTGCGGGCAGCGCGGAAGTG
 AAAACCGTAAACGGCATAACCCATATCGGCCCTTGCCGCCAAGCAACTCGAGGGTGGCGGAGGCAGTGGATCCGCCACA
 AACGACGAGATGTTAAAAAAGCTGCCACTGTGGCCATTGCTGCTGCTTACAACAATGGCCAAGAAATCAACGGTTTC
 AAAGCTGGAGAGACCATCTACGACATTTGATGAAGACGGCACAATTACCAAAAAAGACGCAACTGCAGCGGTTCCGTCCCTTAC
 GCCGACGACTTTAAAGGTTCTGGGTCTGAAAAAAGTCTGACTAACCTGACCAAAACCGTCAATGAAAAACAAACAAAAC
 GTCGATGCCAAAGTAAAGCTGCAGAACTGAAATAGAAAAGTTAAACAACCAAGTTAGCAGACACTGATGCCGCTTTA
 GCAGATACTGATGCCGCTCTGGATGCAACCACCAACGCTTGAATAAATTTGGAGAAAATATAACGACATTTGCTGAA
 GAGACTAAGACAAATATCGTAAAAATTTGATGAAAAATTAAGAAGCCGTGGCTGATACCGTCGACAAGCATGCCGAAGCA
 TTCAACGATATCGCCGATTCATTGGATGAAACCAACACTAAGGCAGACGAAGCCGTCAAAACCGCAATGAAGCCAAA
 CAGACGGCCGAAGAAACCAACAAACCGTTCGATGCCAAAGTAAAGCTGCAGAAACTGCAGCAGGCAAGCCGAAGCT
 GCCGCTGGCACAGCTAATACTGCAGCCGACAAGCCGAAGCTGTGCTGCAAAAAGTTACCGACATCAAAGCTGATATC
 GCTACGAACAAAGATAATATTGCTAAAAAAGCAACAGTGCAGGCTGTACACCAGAGAAGAGTCTGACAGCAAATTT
 GTCAGAAATGATGGTCTGAACGCTACTACCGAAAAATTTGACACACGCTTGGCTTCTGCTGAAAAATCCATTGCCGAT
 CACGATACTCGCCCTGAACGGTTTGGATAAAACAGTGTGCAGCCTGCGCAAAGAAACCCGCCAAGCCCTTGCAGAAACA
 GCCGCGCTCTCCGGTCTGTTCCAACCTTACAACGTGGGTCTCGAGCACACCACCACCACCACTGA

1 MVAADIGAGL ADALTAFLDH KDKGLQSLTL DQSVRKNEL KLAAQGAEKT
 51 YNGDLSLNTG KLKNDKVSFR DFIRQIEVDG QLITLESGEF QVYKQSHSAL
 101 TAFQTEQIQD SEHSGKMKVAK RQFRIGDIAG EHTSFDKLEPE GGRATYRGT
 151 FGSDDAGGKL TYTIDFAAKQ GNGKIEHLKS PELNVDLAAA DIKPDGKRHA
 201 VISGSVLYNQ AEKGSYSLGI FGGKAQEVAG SAEVKTVNGI RHIGLAAKQL
 251 EGGGGTGSAT NDDDVKKAAT VALAAAYNNG QEINGFKAGE TIYDIDEDGT
 301 ITKKDATAAD VEADDFKGLG LKKVVTNLTK TVNENKQNV DAKVKAASEI
 351 EKLTTKLADT DAALADTDAA LDATTNALNK LGENITTFAB ETKTNIVKID
 401 EKLEAVADTV DKHAEAFNDI ADLDETNTK ADEAVKTANE AKQTAETKQ
 451 NVDKVKAAE TAAGKAEAAA GTANTAADKA EAVAARKVTDI KADIATNKDN
 501 IAKKANSADV YTREESDSKF VRIDGLNATF EKLDTRLASA EKSIADHDTR
 551 LNGLDKTVSD LRKETRQGLA EQAALSGLFQ PYNVGLHHH HHH*

AG741-983

ATGGTCGCCCGACATCGGTGCGGGGCTTGCCGATGCACTAACCGCACCGCTCGACCATAAAGACAAAGGTTGCGAG
 TCTTTGACGCTGGATCAGTCCGTCAGGAAAAACGAGAACTGAAGCTGGCGGCACAAGGTGCGGAAAAAAGTTATGGA
 AACGGTGACAGCCTCAATACGGGCAAATTAAGAAACGACAAGGTCAGCCGTTTCGACTTTATCCGCCAAATCGAAGTG
 GACGGGCAGCTCATTACCTTGGAGAGTGGAGAGTTCAGTATACAAACAAAGCCATTCCGCCCTAACCGCCTTTTCAG
 ACCGAGCAAATACAAGATTCCGGAGCATTCCGGGAAGATGGTTGCGAAAACGCCAGTTTCAGAAATCGGGCAGATAGCGGGC
 GAACATACATCTTTTGACAAGCTTCCCAGAGCGGCAGGGCGACATATCGCGGGACGGCGTTTCGGTTTCAGACGATGCC
 GCGGAAAACTGACCTACACCATAGATTTCCGCCCAAGCAGGGAACCGGCAAATCGAACATTGAAATCGCCAGAA
 CTCAATGTCGACCTGGCCCGCCGATATCAAGCCGGATGGAACCGCCATGCCGTCATCAGCGGTTCCGTCCCTTTAC
 AACCAAGCCGAGAAAGGCAGTTACTCCCTCGGTATCTTTGGCGGAAAAGCCAGGAAGTTGCGCGGACCGCGGAAGTG
 AAAACCGTAAACGGCATAACCCATATCGGCCCTTGCCGCCAAGCAACTCGAGGGATCCGGCGGAGGCGGCACCTTCTGCG
 CCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTAC
 GCCGGTATCAAGAACGAAATGTGCAAGACAGAAGCATGCTCTGTGCGCGTTCGGGATGACGTTGCGGTTACAGACAGG
 GATGCCAAAAATCAATGCCCCCCCCCGAATCTGCATACCGGAGACTTCCAAACCCAAATGACGCATACAAGAAATTC
 ATCAACCTCAAACCTGCAATTAAGCAGGCTATACAGGACCGCGGGTAGAGGTAGGTATCGTCGACACAGGCGAATTC
 GTCGGCAGCATATCTTTCCCGAACTGTATGCGAGAAAAGAACCGGCTATAACGAAAATTAACAAAACATATACGGCG
 TATATGCGGAAGGAAGCCCTGAAGACGGAGCGGTAAGACATTGAAGCTTCTTTTCGACGATGAGGCCGTTATAGAG
 ACTGAACCAAAGCCGACGGATATCCGCCACGTAAGAAATCGGACACATCGATTGGTCTCCATATATTATGGCGGG
 CGTTCCTGGACGGCAGACCTGCAGGCGGTATGCGCCGATGCGACGCTACACATAATGAATACGAATGATGAAACC
 AAGAACGAAATGATGGTTGACGCAATCCGCAATGCAAGCTCAAGCTGGGCGAACGTTGGCGTGCAGTACGTCATAAC
 AGTTTGGAAACAACATCGAGGGCAGGCACCTGCCGACTTTCCAAATAGCCAATTCGGAGGAGCAGTACGCCCAAGC
 TTGCTCGACTATTCGGCGGTTGATAAAACAGACAGGAGGTATCCGCTGATGCAACAGAGCGATTCAGGCAACCTGTCC

TACCACATCCGTAATAAAAAACATGCTTTTTCATCTTTTCGACAGGCAATGACGCACAAGCTCAGCCCAACACATATGCC
 CTATTGCCATTTTATGAAAAAGACGCTCAAAAAGGCATTATCAGAGTCGAGGCGTAGACCCGAGTGGAGAAAAGTTTC
 AAACGGGAAATGTATGGAGAACCGGTACAGAACCGCTTGAGTATGGCTCCAACCAATTGCGGAATTAAGTCCATGTGG
 TGCTGTGGCCACCTATGAAGCAAGCGTCCGTTTCACCCGTACAAACCCGATTCAAATTCGCCGAACATCCTTTTCC
 GCACCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGACAACCTGCGTACCAG
 TTGCTGACGACGGCTCAGGACATCGGTGCAAGTCGGCGTGGACAGCAAGTTCGGCTGGGGACTGTGGATGCGGGTAAG
 GCCATGAACGGACCCGCGTCTTCCGTTTCGGCGACTTTTACCAGGATACGAAAGGTACATCCGATATGCGCTACTCC
 TTCCGTAACGACATTTTACGGCACGGGCGGCTGATCAAAAAAGGCGGCGAGCCAACTGCAACTGCACGGCAACAACACC
 TATACGGGCAAAACCATTATCGAAGCGGTTCCGTTGGTGTGTACGGCAACAACAATCGGATATGCGCGTCAAAACC
 AAAGGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGCGAGCCTGAACAGCGACGGCATTGTCTATCTGGCAGATACC
 GACCAATCCGGCCAAACGAAACCGTACACATCAAAAGGCAGTCTGCAGCTGGACGGCAAAGGTACGCTGTACACACGT
 TTGGGCAAACCTGCTGAAAGTGGACGGTACGGCGATATCGCGGGCAAGCTGTACATGTCCGACCGGCAAGGGGGCA
 GGCTATCTCAACAGTACCGGACGACGTGTTCCCTTCCGTGAGTGCAGCCAAAATCGGGCAGGATATTCTTTCTTCCACA
 AACATCGAAACCGACGGCGGCTGCTGGCTCCCTCGACAGCGTCAAAAAACAGCGGGCAGTGAAGGCGACACGCTG
 TCCATTATGTCCGTCGGGCAATCGCGCACGGACTGCTTCGGCAGCGGCACATTCGCGCCCGCGGCTGTGAAACAC
 ACCGTGAAACCGGGCGGAGCAATCTGGAAAACCTGATGGTFCGAACCTGGATGCCTCCGAATCATCCGCAACACCCGAG
 ACGGTTGAAACTGCGGCAGCCGACCGCACAGATATGCGGGCATCCGCCCTACGGCGCAACTTTCCGCGCAGCGGCA
 GCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACGATACC
 GCCGCCATGCCGATATGCAAGGACGCCGCCGTGAAAGCCGATCGGACGGGTTGGACCACAACGGCAGCGGCTGTGGC
 GTCATCGCGCAAACCCAAACAGGACGGTGGAACTGGGAACAGGGCGGTTGTGAAGGCAAATGCGCGGCAGTACCCAA
 ACCGTCCGCAATTCGCCCGGAAACCGCGGAAAATACGACAGCAGCCGACACTGGGCATGGGACGCGACACATGGAGC
 GAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGCTAT
 CTCAAAGGCTGTTCTCTACGGACGCTACAAAACAGCATCAGCCGACGACCCGGTGGCGACGAAACATGCCGGAAGGC
 AGCGTCAACGGCACGCTGATGCAAGTGGGGCGACTGGGGCGGTGCAACGTTCCGTTTGCAGCAACGGGAGATTTGACG
 GTCGAAGGCGGCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAGGCAGTGTCTTGGGCTGGAGCGGC
 AACAGCCTCACTGAAGGACCGCTGCTCGGACTCGCGGCTGGAAGTGTGCAACCCCTTGAGCGATAAAGCCGCTCTG
 TTTGCAACGGCGGCGTGGAAACCGGACCTGAACGGACGCGACTACACGGTAACGGGGCGGCTTACCGGCGGACTGCA
 GCAACCGGCAAGACGGGGCACGCAATATGCCGACACCCGCTGTTGTTGCCGCGCTGGGCGGGATGTGCAATTCGGC
 AACGGCTGGAACGGCTTGGCACGTTACAGCTACGCCGTTCCAAAACAGTACGGCAACCACAGCGGACGAGTCCGCGTA
 GGCTACCGGTTCTCGAGCACCACCACCACCCTGA

1	MVAADIGAGL	ADALTAFLDH	KDKGLQSLTL	DQSVRKNEKL	KLAAQGAETK
51	YNGDLSLNTG	KLKNDKVSFR	DFIROIEVDG	QLITLESGEF	QVYKQSHSAL
101	TAFQTEQIQD	SEHSKGMVAK	RQFRIGDIAG	EHTSFDKLE	GGRATYRGTA
151	FGSDDAGGKL	TYTIDFAAKQ	GNGKIEHLKS	PELNVDLAA	DIKPDGKRHA
201	VISGSVLYNQ	AEKGSYSLGI	FGGKAQEVAG	SAEVKTVNGI	RHIGLAAKQL
251	EGSGGGGTS	PDFNAGGTGI	GSNSRATTAK	SAAVSYAGIK	NEMCKDRSML
301	CAGRDDVAVT	DRDAKINAPP	PMLHTGDFPN	PNDAYKNLIN	LKPAIEAGYT
351	GRGVEVGVVD	TGESVGSISF	PELYGRKEHG	YNENYKNYTA	YMRKEAPEDG
401	GGKDIRASF	DEAVIETFEAK	PTDIRHVKEI	GHIIDLVSII	GGRSVDGRPA
451	GGIAPDATLH	IMNTNDETKN	EMMVAAIRNA	WVKLGERGVR	IVNNSFGTTS
501	RAGTADLFQI	ANSEBQYRQA	LLDYSGGDKT	DEGLIRLMQOS	DYGNLSYHIR
551	NKNMLFIFST	GNDAAQAPNT	YALLPFYEKD	AQKGIITVAG	VDRSGEKFKR
601	EMYGEPGTEP	LEYGSNHCGI	TAMWCLSAPY	EASVRFTRTN	PIQLAGTSFS
651	APIVTGTAAL	LLQKYPWMSN	DNLRTTLLT	AQDIGAVGVD	SKFGWGLLDA
701	GKAMNGPASF	PFGDFTADTK	GTSDIAYSFR	NDISGTGLI	KKGGSQQLHL
751	GNNTYTGTI	IEGGLVLYG	NNKSDMRVET	KGALIYNGAA	SGGSLNSDGI
801	VYLADTDQSG	ANETVHIKGS	LQLDGRGTY	TRLGKLLKVD	GTAIIGGKLY
851	MSARGKGAGY	LNSTGRRVFP	LSAAKIGQDY	SFFTNIETDG	GLLASLDSVE
901	RTAGSEGLTL	SYVRRGNAA	RTASAAHSA	PAGLKHAVEQ	GGSNLENLMV
951	ELDASESAT	PETVETAAD	RTDMPGIRPY	GATFRAAAAV	QHANAADGVR
1001	IFNSLAATVY	ADSTAAHADM	QGRRLKAVSD	GLDHNGTGLR	VIAQTQDDG
1051	TWEQGGVEGK	MRGSTQTVGI	AAKTGENTTA	AATLGMGRST	WSENSANAKT
1101	DSISLFFAGIR	HDAGDIGYLK	GLFSYGRYKN	SISRSTGADE	HAEGSVNGTL
1151	MQLGALGGVN	VFFAATGDLT	VEGGLRYDLL	KQDAFAEKG	ALGWSGNSLT
1201	EGTLVGLAGL	KLSQPLSDKA	VLFATAGVER	DLNDRDYTVT	GGFTGATAAT
1251	GKTGARNMPH	TRLVAGLGAD	VEFGNGWNL	ARYSYAGSKQ	YGNHSGRVGV
1301	GYRFLHEHHH	HH*			

AG741-ORF46.1
 ATGGTCCGCCGACATCGGTGCGGGGCTTCCGATGCACTAACCCGACCGCTCGACCATAAAGACAAAGGTTTGCAG
 TCTTTGACGCTGGATCAGTCCGTGAGGAAAAACGAGAAACTGAAAGCTGGCGGCACAAGGTGCGGAAAAAACTTATGGA
 AACGGTGACAGCCTCAATACGGGCAATGAAGAACGACAAGGTGAGCCGTTTCGACTTTATCCGCCAAATCGAAGTG
 GACGGGCGAGCTCATTACCTTGGAGAGTGGAGAGTCCAAAGTATACAAACAAGCCATTCGGCCTTAACCGCCTTTTCCAG
 ACCGAGCAAATACAAGATTCCGAGCATTCCGGGAAGTGGTTCGAAACCGCAGTTTCAAGATTCGGGACATAGCGGGC
 GAACATACATCTTTTGACAAGCTTCCCGAAGGCGGACGGCGACATATCGGGGACGGCGTTCGGTTCAGACGATGCC

```

GGCGGAAAACCTGACCTACACCATAGATTTCCGCCCAAGCAGGGAAAACGGCAAATCGAACATTTGAAATCGCCAGAA
CTCAATGTGACCTCGCCGCCGATATCAAGCCGGATGGAAAACGCCATGCCGTTCATCAGCGGTTCCGTCCCTTAC
AACCAAGCCGAGAAAGGCAGTTACTCCCTCGGTATCTTTGGCGGAAAAGCCAGGAAGTTGCCGGCAGCGCGGAAGTG
AAAACCGTAAACGGCATAACGCATATCGGCCCTTGGCCCAAGCAACTCGACGGTGGCGGAGGCACTGGATCCTCAGAT
TTGGCAAACGATTCTTTTATCCGGCAGGTTCTCGACCGTCAGCATTTCGAACCCGACGGGAAAATACCACCTATTCGGC
AGCAGGGGGAACTTGCCGAGCGCAGCGGCCATATCGGATTGGGAAAAATACAAAGCCATCAGTTGGGCAACCTGATG
ATTCACACAGCGGCCATTAAAGGAAATATCGGCTACATTTGTCCGCTTTTCCGATCACGGGCACGAAGTCCATCCCCC
TTCGACAACCATGCCCTCACATTCGGATTCTGATGAAGCCGGTAGTCCCGTTGACGGATTAGCCTTTACCGCATCCAT
TGGGACGGATACGAACACCATCCCGCCGACGGCTATGACGGGCCACAGGGCGGGGCTATCCCGCTCCCAAAGGCCG
AGGGATATATACAGCTACGACATAAAAGGCGTTGCCCAAATAATCCGCTCAACCTGACCGACAACCGCAGCACCGGA
CAACGGCTTGGCCGACCGTTTCCACAATGCCGGTAGTATGCTGACGCAAGGAGTAGGGCAGCGATTCAAACCGGCCACC
CGATACAGCCCCGAGCTGGACAGATCGGGCAATGCCCGCAAGCCTTCAACGGCACTGCAGATATCGTTAAAAACATC
ATCGGCGCGGCAGGAGAAATTTGCGGCGCAGCGCATGCCGTGCCAGGGCATAAGCGAAGGCTCAAACATTTGCTGTCTG
CACGGCTTGGGTCTGCTTTCCACCAGAAAACAAGATGGCGCGCATCAACGATTTGGCAGATATGGCGCAACTCAAAGAC
TATGCCCGCAGCAGCCATCCGCGATTGGGCGAGTCCAAAACCCCAATGCCCGCACAAGGCATAGAAGCCGTGAGCAATATC
TTTATGGCAGCCATCCCATCAAAGGGATTGGAGCTGTTCCGGGAAAATACGGCTTGGCGGCATCACGGCACATCCT
ATCAAGCGGTTCGAGATGGGCGGATCGCATTTGCCGAAAGGAAAATCCCGCTCAGCGACAATTTTGGCCGATGGCGCA
TACGCCAAATACCGCTCCCTTACCATTCCCGAAATATCCGTTCAAACCTGGAGCAGCGTTACGGCAAAGAAAACATC
ACCTCCTCAACCGTCCCGCCGTCAAACGGCAAATAATGTCAAACCTGGCAGACCAACGCCACCCGAAGACAGGCGTACCG
TTTGCAGGTAAGGGTTTCCGAAATTTGAGAAGCACGTGAAATATGATACGCTCGAGCACCACCACCACCACCTGA

```

```

1   MVAADIGAGL ADAL/TAPLDH KDKGLQSL/TL DQSVRKNEKL KLAAQGAEKT
51  YNGDLSLNTG KLKNDKVSFR DFIRQIEVDG QLITLESGEF QVYKQSHSAL
101 TAFQTEBQIQD SEHSGKMVAK RQFRIGDIAG EHTSFDKLEP GGRATYRGT
151 FGSDDAGGKL TTYTIDFAAQ GNGKIEHLKS PELNVDLAAA DIKPDGKRHA
201 VISGSVLYNQ AEKGSYSLGI FGGKAQEVAG SAEVKTVNGI RHIGLAAKQL
251 DGGGGTGSDD LANDSFIRQV LDRQHFEPDG KYHLFGSRGE LAERSGHIGL
301 GKIQSHQLGN LMIQQAALIK NIGYIVRFS D HGHEVHSPFD NHASHSDSDE
351 AGSPVDGFSL YRIHWDGYEH HPADGYDGPQ GGGYPAPKGA RDIYSYDIKG
401 VAQNIRLNL T DNRSTGQRLA DRPHNAGSML TQGVGDGPKR ATRYSPELDR
451 SGNAAEAFNG TADIVKNIIG AAGEIVGAGD AVQGISSEGN IAVMHGLGLL
501 STENKMARIN DLADMAQLKD YAAAAIRDWA VQNPNAAQGI EAVSNIFMAA
551 IPIKGIGAVR GKYLGGITA HPIKRSQMG A LALFKGKSAV SDNFADAAYA
601 KYPSPYHSRN IRSNLEQRYG KENITSSTVP PSNGKNVCLA DQRHPKTGVP
651 FDGKGFNFE KHVKYDTLEH HHHHH*

```

实施例 5-287 的杂交体

以全长、具有 C-末端 His-标记或无其前导肽但具有 C-末端 His-标记的 287 的表达水平非常低。用 N-末端 GST-融合实现了较好的表达。将 GST 用作 N-末端融合配体的替代方法，将 287 置于蛋白质 919 的 C-末端(‘919-287’)、蛋白质 953 的 C-端(‘953-287’)和蛋白质 ORF46.1 的 C-末端(‘ORF46.1-287’)。在这两种方法中，前导肽都是缺失的，且杂交体是直接的符合读框的融合。

为了制备 953-287 杂交体，通过设计正向引物自各序列前导区的下游，除去两种蛋白质的前导肽；在 953 反向引物中除去终止密码子序列，但在 287 反向引物中包含终止密码子序列。对 953 基因而言，用于扩增的 5'和 3'引物分别包括 NdeI 和 BamHI 限制酶切位点，而对 287 基因的扩增而言，5'和 3'引物分别包括 BamHI 和 XhoI 限制酶切位点。用这种方法，用 NdeI-BamHI(克隆第一个基因)且随后用 BamHI-XhoI(克隆第二个基因)，可以实现 pET21b+中两个基因的顺序定向克隆。

通过将编码 287 的成熟部分的序列克隆入 pET21b+ 中的 919-His 克隆的 3'-端的 XhoI 位点, 可以得到 919-287 杂交体。设计用于扩增 287 基因的引物, 从而在 PCR 片段的 5'-引入 SalI 限制酶切位点并在 3'-引入 XhoI 位点。因为由 SalI 和 XhoI 限制酶产生的粘性末端是相容的, 因此可以将由 SalI-XhoI 消化的 287PCR 产物插入到由 XhoI 切割的 pET21b-919 克隆中。

类似地得到 ORF46.1-287 杂交体。

将针对杂交蛋白产生的抗体的杀菌效力(同源菌株)与针对组分抗原的简单混合物产生的抗体相比:

	与 287 的混合物	与 287 的杂交体
919	32000	16000
953	8192	8192
ORF46.1	128	8192

对 919-287 和 953-287 而言, 还获得了针对异源 MenB 菌株和针对血清型 A 和 C 的杀菌活性的数据:

菌株	919		953		ORF46.1	
	混合物	杂交体	混合物	杂交体	混合物	杂交体
MC58	512	1024	512	1024	-	1024
NGH38	1024	2048	2048	4096	-	4096
BZ232	512	128	1024	16	-	-
MenA (F6124)	512	2048	2048	32	-	1024
MenC (C11)	>2048	n.d.	>2048	n.d.	-	n.d.
MenC (BZ133)	>4096	>8192	>4096	<16	-	2048

还构建了 ORF46.1 和 919 的杂交体。在 N-末端用 919 得到了最佳结果(高 4 倍的滴定度)。

还测试了杂交体 919-519His、ORF97-225His 和 225-ORF97His。它们的 ELISA 滴定度和杀菌抗体应答结果中等。

两种蛋白质 A 和 B 的杂交体可以是 NH₂-A-B-COOH 或 NH₂-B-A-COOH, 用△G287 还制得了在 N-末端与 287 “反向”杂交的杂交体。使用一系列菌株, 包括同源菌株 2996。FCA 用作佐剂:

菌株	287 & 919		287 & 953		287 & ORF46.1	
	$\Delta G287-919$	919-287	$\Delta G287-953$	953-287	$\Delta G287-46.1$	46.1-287
2996	128000	16000	65536	8192	16384	8192
BZ232	256	128	128	<4	<4	<4
1000	2048	<4	<4	<4	<4	<4
MC58	8192	1024	16384	1024	512	128
NGH38	32000	2048	>2048	4096	16384	4096
394/98	4096	32	256	128	128	16
MenA (F6124)	32000	2048	>2048	32	8192	1024
MenC (BZ133)	64000	>8192	>8192	<16	8192	2048

通常在 N-末端用 287 观察到较好的杀菌滴定度。

当融合于蛋白质 961[如上所示的 NH_2 - $\Delta G287$ -961-COOH-序列]时,得到的蛋白质是不溶解的,纯化必须将其变性和复性。在复性后,发现约 50%蛋白质仍不溶解。比较可溶和不溶蛋白质,用可溶蛋白质得到更好的杀菌滴定度(FCA 作为佐剂):

	2996	BZ232	MC58	NGH38	F6124	BZ133
可溶的	65536	128	4096	>2048	>2048	4096
不溶的	8192	<4	<4	16	n.d.	n.d.

但用明矾佐剂替代,可以改善不溶形式的蛋白质的滴定度:

不溶的	32768	128	4096	>2048	>2048	2048
-----	-------	-----	------	-------	-------	------

还在杂交蛋白中使用 961c(见上)。由于 961 及其结构域变体指导有效的表达,它们非常适合作为杂交蛋白的 N-末端部分。

实施例 23-其它杂交体

附图中显示本发明的其它杂交蛋白,它们具有如下所示的序列。当与单独蛋白质相比时,这此杂交蛋白是有利的:

ORF46.1-741

```

ATGTCAGATTTGGCAAACGATTCCTTTATCCGGCAGGTTCTCGACCGTCAGCATTTCGAACCCGACGGGAAATACCAC
CTATTCGGCAGCAGGGGGAACTTGCCGAGCGCAGCGGCCATATCGGATTTGGGAAAATACAAAGCCATCAGTTGGGC
AACCTGATGATTC AACAGGCGCCATTAAAGGAAATATCGGCTACATTGTCCGCTTTTCCGATCACGGGCACGAAGTC
CATTCCCCCTTCGACAACCATGCCTCACATTCCGATTCCTGATGAAGCCGGTAGTCCCGTTGACGGATTTAGCCTTTAC
CGCATCCATTTGGACGGATACGAACACCATCCCGCGACGGCTATGACGGGCCACAGGGCGGGCGCTATCCCGCTCCC
AAAGGCGCGAGGGATATATACAGCTACGACATAAAAGGCGTTGCCCAAATATCCGCTCAACCTGACCGACAACCGC
AGCACCGGACAACGGCTTGCCGACCGTTCCACAATGCCGGTAGTATGCTGACGCAAGGAGTAGGCGACGGATTCAAA
CGCGCCACCCGATACAGCCCCGAGCTGGACAGATCGGGCAATGCCGCCGAAGCCTTCAACGGCACTGCAGATATCGTT
AAAAACATCATTCGGCGCGCAGGAGAAATTGTCCGGCAGGGCATGCCGTGCAGGGCATAAGCGAAGGCTCAAACATT

```

GCTGTCAATGCACGGCTTGGGCTGCTTTCCACCGAAAACAAGATGGCGCGCATCAACGATTTGGCAGATATGGCGCAA
 CTCAAAGACTATGCCGAGCAGCCATCCGCGATTTGGGCGAGTCCAAAACCCCAATGCCGACAAAGGCATAGAAGCCGTC
 AGCAATATCTTTATGGCAGCCATCCCATCAAAGGGATTGGAGCTGTTCCGGGAAAATACGGCTTGGCGCGCATCACG
 GCACATCCTATCAAGCGGTCCGAGATGGCGCGCATCGCATTTCCGAAAGGAAATCCGCCGTACGCGACAATTTTGGC
 GATGCGGCATACGCCAAATACCCGTCCTTACCATTTCCGAAATATCCGTTCAAACCTGGAGCAGCGTTACGGCAA
 GAAAACATCACCTCTCAACCGTCCCGCGTCAAACGGCAAAAATGTCAAACCTGGCAGACCAACGCCACCCGAAGACA
 GCGTACCGTTTGACGGTAAAGGTTTCCGAATTTGAGAAGCACGTGAAATATGATACGGGATCCGGAGGGGTTGGT
 GTCGCGCCGACATCGGTGCGGGCTTGCCTGATGCACTAACCGCACCGCTCGACCATAAAGCAAAGGTTTGCAGTCT
 TTGACCGTGGATCAGTCCGTGAGAAAACGAGAACTGAAGCTGGCGGCACAAGTGGGAAAACCTTATGGAAAC
 GGTGACAGCCTCAATACGGGCAAATGAAGAACGACAAGTCCAGCCGTTTCGACTTTATCCGCCAAATCGAAGTGGAC
 GGGCAGCTCATTACCTTGGAGAGTGGAGAGTTCCAAGTATACAAACAAGCCATTCGCCCTTAACCGCTTTCAGACC
 GAGCAAATACAAGATTCGGAGCATTCCGGGAAGATGGTTGCGAAACGCCAGTTCAGAATCCGGCAGATAGCGGGCGAA
 CATACATCTTTTGACAAGCTTCCCGAAGCGCGCAGGGCCGACATATCGCGGACCGCGTTCCGTTTCAGACGATGCGCGC
 GGAAAACCTGACCTACACCATAGATTTCCGCGCCAAAGCAGGGAAAACGGCAAAAATCGAACATTTGAAATCGCCAGAACTC
 AATGTCGACCTGGCCGCGCGGATATCAAGCCGATGGAAAACGCCATGCCGTTCATCAGCGGTTCCGTCCTTTACAAC
 CAAGCCGAGAAAGCAGTTACTCCCTCGGTATCTTTGGCGGAAAAGCCAGGAAATGGCCGCGAGCGGAAAGTGA
 ACCGTAAACGGCATAAGCCATATCGGCCTTGCCTCAAGCAACTCGAGCACCACCACCACCACCCTGA

- 1 MSDLANDSFI RQVLDLQHFPE PDGKYHLFGS RGELAERSGH IGLGKIQSHQ
- 51 LGNLMIQAA IKGNIGYIVR FSDHGHEVHS PFDNHASHSD SDEAGSPVDG
- 101 FSLYRLHWDG YEHHPADYD GPQGGYPAP KGARDIYSYD IKGVAQNIRL
- 151 NLTDNRSTGQ RLADRFHNAG SMLTQGVGDG FKRATRYSPF LDRSGNAEA
- 201 FNGTADIVKN IIGAAGEIVG AGDAVQGISE GSNLAVMHGL GLLSTENKMA
- 251 RINDLADMAQ LKDYAAAIR DWAVQNPNA QGIEAVSNIF MAIPIKIGIG
- 301 AVRKYVGLGG ITAHPIKRSQ MGATLPLKPK SAVSDNFADA AYAKYPSPYH
- 351 SRNIRSNLEQ RYKKNITSS TVPPSNGKNV KLADQRHPKT GVPFDGKGF
- 401 NFEKHVKYDT GSGGGVVAAD IGAGLADALT APLDHDKDKGL QSLTLDQSVR
- 451 KNEKLLKLAQ GAERTYNGD SLNTGKLNKND KVSRRDFIRQ IEVDGQLITL
- 501 ESGEFQVYKQ SHSALTAFT BQIQDSEHSG KMVAKRQFRI GDIAGEHTSF
- 551 DKLPBGGRT YRGTAFGSDD AGGKLYTID FAAKQNGKI EHLKSPFLNV
- 601 DLAAADIKPD GKRHAVISGS VLYNQAEKGS YSLGIFGGKA QEVAGSAEVK
- 651 TVNGIRHIGL AAKQLEHHH HH*

ORF46.1-961

ATGTCAGATTTGGCAAACGATCTTTTTATCCGGCAGGTTCTCGACCGTCAGCATTTCGAACCCGACGGGAAATACCAC
 CTATTCGGCAGCAGGGGGAACTTGCAGGCGCAGCGCCATATCGGATTTGGGAAAATACAAAGCCATCAGTTGGGC
 AACCTGATGATTCAACAGGCGCCATTAAAGGAATATCGGCTTACATGTTCCGCTTTCCGATCACGGGCAGAAAGTC
 CATTCCCTTCGACAACCATGCTCACATTCGATTTCTGATGAAGCCGTTAGTCCCGTTGACGGATTTAGCCTTTAC
 CGCATCCATTGGGACGGATACGAACACCATCCCGCCAGCGCTATGACGGGCCACAGGGCGGCGCTATCCCGCTCCC
 AAAGCGCGAGGGATATATACAGCTACGACATAAAGGCGTTGCCCAAAATATCCGCTCAACCTGACCCGACAAACCGC
 AGCACCGGACAACGGCTTGCAGCCGTTTCCACAATGCCGGTAGTATGCTGACGCAAGGAGTAGGCGACGGATTCAA
 CGCGCCACCCGATACAGCCCCGAGCTGGACAGATCGGGCAATGCCCGGAAGCCTTCAACGGCACTGAGATATCGTT
 AAAACATCATCGCGCGGCGAGGAAATTTGCGCGCAGCGGATGCCGTGCGAGGCGATAAGCGAAGGCTCAAACATT
 GCTGTCAATGCACGGCTTGGGCTGCTTTCCACCGAAAACAAGATGGCGCGCATCAACGATTTGGCAGATATGGCGCAA
 CTCAAAGACTATGCCGAGCAGCCATCCGCGATTGGGCGAGTCCAAAACCCCAATGCCGCAAAAGGCATAGAAGCCGTC
 AGCAATATCTTTATGGCAGCCATCCCATCAAAGGGATTGGAGCTGTTCCGGGAAAATACGGCTTGGGCGGCATCACG
 GCACATCCTATCAAGCGGTCCGAGATGGCGCGCATCGCATTTCCGAAAGGAAATCCGCCGTACGCGACAATTTTGGC
 GATGCGGCATACGCCAAATACCCGTCCTTACCATTCCGAAATATCCGTTCAAACCTTGGAGCAGCGTTACGGCAA
 GAAAACATCACCTCTCAACCGTCCCGCTCAAACGGCAAAAATGTCAAACCTGGCAGACCAACGCCACCCGAAGACA
 GCGTACCGTTTGACGGTAAAGGTTTCCGAATTTGAGAAGCACGTGAAATATGATACGGGATCCGGAGGAGGAGGA
 GCCCAAACGACGACGATGTTAAAAAGCTGCCACTGTGGCCATTGCTGCTGCCTACAACAATGGCCAAGAAATCAAC
 GGTTCAAAGCTGGAGAGACCATCTACGACATTGATGAAGACGGCACAAATACCAAAAAGACGCAACTGCAGCCGAT
 GTTGAAGCCGACGACTTTAAAGGCTGGGTCTGAAAAGAGTCTGACTAACCTGACCAAAAACCGTCAATGAAAACAAA
 CAAAACGTCGATGCCAAAGTAAAGCTGCAGAATCTGAAATAGAAAAGTTAACAACCAAGTTAGCAGACACTGATGCC
 GCTTTAGCAGATACTGATGCCGCTCTGGATGCAACCACCAACGCTTTGAATAAATGGGAGAAAATATAACGACATTT
 GCTGAAGAGACTAAGCAAATATCGTAAAAATGATGAAAAATAGAAGCCGTTGGCTGATACCGTCGACAAGCATGCC
 GAAGCATCAACGATATCGCCGATTCATTGGATGAAACCAACTAAGGCAGACGAAAGCCGTCAAAACCCGCAATGAA
 GCCAAACAGACGGCCGAGAAACCAACAAACGTCGATGCCAAAGTAAAGCTGCGAAGAACTGACGAGCAAGGAGCC
 GAAGCTGCCGCTGGCAGCACTAATGCTGACCGGACAAAGCCGAAAGCTGTGCTGCAAAAAGTTACCGACATCAAAGCT
 GATATCGCTACGAACAAAGATAATATTGCTAAAAAGCAAACAGTGCAGACCTGCGCAAAGAAACCCGCAAGCCCTTGA
 GAACAAGCCGCTCTCCGGCTGTTCCAACTTACAACCTGAGTGGCTGAGTTCCGTTCAATGTAACGGCTGCAGTCCGGCGT
 AAATCCGAATCCGCTCGCTCGGATCCGCTTCCGCTTACCAGAAAACCTTTGCCGCAAAGCAGCGTGGCAGTCC
 GGCATTCGTCGGTTCTTCCGAGCCTACCATGTCGGCGTCAATFACGAGTGGCTCGAGCACCACCACCACCACC
 TGA

```

1   MSDLANDSFI RQVLDRQHFE PDGKYHLFGS RGELAERSGH IGLGKIQSHQ
51  LGNLMIQQAA IKGNIGYIVR FSDHGHEVHS PFDNHASHSD SDEAGSPVDG
101 FSLYRIHWDG YEHHPADGYD GPQGGYPAP KGARDIYSYD IKGVAQNIRL
151 NLTDNRSTGQ RLADRFHNAG SMLTQGVGDG FKRATRYSP E LDRSGNAAEA
201 FNGTADIVKN IIGAAGEIVG AGDAVQGISE GSNIAVMHGL GLLSTENKMA
251 RINDLADMAQ LKDYAAAAIR DWAVQNPNA QGIEAVSNIF MAAIPIKGIG
301 AVRKYGLGG ITAHPIKRSQ MGALALPKGK SAVSDNFADA AYAKYPSPYH
351 SRNIRSNLEQ RYKKNITSS TVPPSNGKNV KLADQRHPKT GVPFDGKGF
401 NFERHKVYDT GSGGGGATND DDVKKAATVA IAAAYNNGQE INGFKAGETI
451 YDIDEDGTIT KKDATAADVE ADDFKGLGLK KVVTNLTKTV NENKQNVDAK
501 VKAAESEIEK LTTKLADTDA ALADTDAALD ATTNALNKLK ENITTFAEET
551 KTNIVKIDEK LEAVADTVDK HAEAFNDIAD SLDETNTKAD EAVKTANEAK
601 QTAEETKQNV DAKVKAETA AGKAEAAAGT ANTAADKAEA VAAKVTDIKA
651 DIATNKDNIA KRANSADVYT REESDSKFVR IDGLNATTEK LDTRLASAEK
701 SIADHDTRLN GLDKTVSDLR KETROGLAEQ AALSGLFQPY NVGRFNVTA
751 VGGYKSESAV AIGTGFRFTE NFAAKAGVAV GTSSGSSAAY HVGVNVEWLE
801 HHHHHH*

```

ORF46.1-961c

```

ATGTCAGATTTGGCAAACGATTCTTTTATCCGGCAGGTTCTCGACCCTCAGCATTTCGAACCCGACGGGAAATACCCAC
CTATTCGGCAGCAGGGGGAACTTGCCGAGCGCAGCGCCCATATCGGATGGGAAAATAACAAGCCATCAGTTGGGC
AACCTGATGATCAACAGCGCGCATTAAGGAAATACGGCTACATTGTCCGCTTTCCGATCACGGGCACCGAAGTC
CATCCCCCTTCGACAACCATGCCTCACATTCTGATGAAGCCGGTAGTCCCGTTGACGGATTAGCCTTTAC
CGCATCCATTGGGACGGATACGAACACCATCCCGCCGACGGCTATGACGGGCCACAGGGCGGGCTATCCOGTCCC
AAAGCGCGAGGGATATATACAGCTACGACATAAAGGCGTTGCCCAAATATCCGCCTCAACCTGACCACAACCGC
AGCACCCGACAACGGCTTGCCGACCGTTTCCACAATGCCGGTAGTATGCTGACGCAAGGAGTAGGCGACGGATTCAA
CGCGCCACCAGTACAGCCCGAGCTGGACAGATCGGGCAATGCCCGGAAGCCTTCAACGGCACTGCAGATATCGTT
AAAAACATCATCCGCGCGGAGAGAAATTTGTCGGCGCAGCGATGCCGTGCAGGCATAAGCGAAGGCTCAAACATT
GCTGTCATGCACGGCTTGGGTCTGCTTCCACCGAAAACAAGATGGCGCGCATCAACGATTTGGCAGATATGGCGCAA
CTCAAAGACTATGCCGAGCAGCCATCCGCGATTGGGCAGTCCAAAACCCCAATGCCGCACAAGGCATAGAAGCCGTC
AGCAATATCTTTATGGCAGCCATCCCCATCAAAGGGATGGAGCTGTTTCGGGGAAAATACGGCTTGGGCGGCATCAG
GCACATGCTTATCAAGCGGTGCGAGATGGGCGCGATCGCATTCGCCGAAAAGGAAAATCCCGCTGAGCGCAATTTTGGC
GATGCCGCATACGCCAAATACCCGTCCCTTACCATTCCCGAAAATATCCGTTCAAACTTGGAGCAGCGTTACGGCAA
GAAAACATCACCTCCTCAACCGTGCCGCGTCAAACGGCAAATATGTCAAACTGGCAGACCAACGCCACCCGAAGACA
GGCGTACCCTTTGACGGTAAAGGGTTTCCGAATTTTGAGAAACACGTAAGAAATATGATACGGGATCCGGAGGAGGGA
GCCACAAACGACGACGATGTTAAAAAGCTGCCACTGTGGCCATTGCTGCTGCCTACAACAATGGCCAAGAAATCAAC
GGTTTCAAAGCTGAGAGACCATCTACGACATTGATGAAGACGGCACAAATTACAAAAAAGACGCAACTGCAGCCGAT
GTTGAAGCCGACGACTTTAAAGGTCTGGGTCTGAAAAAAGTCTGACTAACCTGACCAAAACCGTCAATGAAAAACAA
CAAACCGTCGATGCCAAAGTAAAGCTGCAGAATCTGAAATAGAAAAGTTAAACAACCAAGTTAGCAGACACTGATGCC
GCTTTAGCAGATACTGATGCCGCTCTGGATGCAACCACCAACCGCTTGAATAAATTTGGGAGAAAATATAACGACATTT
GCTGAAGAGACTAAGACAAATATCGTAAAAAATGATGAAAAATTAGAAGCCGTGGCTGATACCGTCGACAAGCATGCC
GAAGCATCTCAACGATATCCCGGATTCATTGGATGAAACCAACTAAGGCAGACGAAAGCCGTCAAAACCCCAATGAA
GCAAAACAGCAGCGGCGAAGAAACCAAAACAGTTCGATGCCAAAGTAAAGCTGCAGAAAACCTGCAGCAGCAAGCC
GAAGCTGCCGCTGGCACAGCTAATACTGCAGCCGACAAGGCCGAAGCTGTCCGTGCAAAAGTTACCGACATCAAAGCT
GATATCCCTACGAACAAAGATAATATTGCTAAAAAGCAAAACAGTGCAGCAGCTGTACACCAGAGAAGAGTCTGACAGC
AAATTTGTGAGAAATGATGGTCTGAACGCTACTACCGAAAAATGGACACACGCTTGGCTTCTGCTGAAAAATCCATT
GCCGATCACGATACTCGCCTGAACGGTTTGGATAAAACAGTGTGACACCTGCGCAAAGAAACCCGCCAAGGCCTTGA
GAACAAACCCGCGCTCTCCGCTCTGTTCCAACCTTACAACGTTGGGTCTCGAGCACCACCACCACCACCTGA

```

```

1   MSDLANDSFI RQVLDRQHFE PDGKYHLFGS RGELAERSGH IGLGKIQSHQ
51  LGNLMIQQAA IKGNIGYIVR FSDHGHEVHS PFDNHASHSD SDEAGSPVDG
101 FSLYRIHWDG YEHHPADGYD GPQGGYPAP KGARDIYSYD IKGVAQNIRL
151 NLTDNRSTGQ RLADRFHNAG SMLTQGVGDG FKRATRYSP E LDRSGNAAEA
201 FNGTADIVKN IIGAAGEIVG AGDAVQGISE GSNIAVMHGL GLLSTENKMA
251 RINDLADMAQ LKDYAAAAIR DWAVQNPNA QGIEAVSNIF MAAIPIKGIG
301 AVRKYGLGG ITAHPIKRSQ MGALALPKGK SAVSDNFADA AYAKYPSPYH
351 SRNIRSNLEQ RYKKNITSS TVPPSNGKNV KLADQRHPKT GVPFDGKGF
401 NFERHKVYDT GSGGGGATND DDVKKAATVA IAAAYNNGQE INGFKAGETI
451 YDIDEDGTIT KKDATAADVE ADDFKGLGLK KVVTNLTKTV NENKQNVDAK
501 VKAAESEIEK LTTKLADTDA ALADTDAALD ATTNALNKLK ENITTFAEET
551 KTNIVKIDEK LEAVADTVDK HAEAFNDIAD SLDETNTKAD EAVKTANEAK
601 QTAEETKQNV DAKVKAETA AGKAEAAAGT ANTAADKAEA VAAKVTDIKA
651 DIATNKDNIA KRANSADVYT REESDSKFVR IDGLNATTEK LDTRLASAEK
701 SIADHDTRLN GLDKTVSDLR KETROGLAEQ AALSGLFQPY NVGLEHHHHH
751 H*

```


961-ORF46.1

ATGGCCACAAACGACGACGATGTTAAAAAGCTGCCACTGTGGCCATTGCTGCTGCCTACAACAATGGCCAAGAAATC
AACGGTTTCAAAGCTGGAGAGACCATCTACGACATTGATGAAGACGGCACAATTACCAAAAAGACGCAACTGCGACC
GATGTTGAAGCCGACGACTTAAAGGTCTGGGTCTGAAAAAGTCGTGACTAACCTGACCAAAACCGTCAATGAAAAC
AAACAAAACGTCGATGCCAAAAGTAAAAGCTGCAGAATCTGAAATAGAAAAGTTAACAAACCAAGTTAGCAGACACTGAT
GCCGCTTTAGCAGATACTGATGCCGCTCTGGATGCAACCACCAACGCCTTGAATAAATTTGGGAGAAAATATAACGACA
TTTGCTGAAGAGACTAAGACAAATATCGTAAAAATTTGATGAAAAATTAGAAGCCGTGGCTGATACCGTCGACAAGCAT
GCCGAAGCATTCAACGATATCGCCGATTCATTTGGATGAAACCAACACTAAGGCAGACGAAGCCGTCAAAACCGCCAAT
GAAGCCAAACAGACGGCCGAAGAAACCAACAAAACGTCGATGCCAAAAGTAAAAGCTGCAGAACTGCAGCAGGCAAA
GCCGAAGCTGCCGTCGGCACAGCTAATACTGCAGCCGACAAGGCCGAAGCTGTCGCTGCAAAAAGTTACCGACATCAAA
GCTGATATCGCTACGAACAAAGATAATATTGCTAAAAAAGCAAACAGTGCAGCGTGTACACCAGAGAAGAGTCTGAC
AGCAAATTTGTCAGAATTTGATGGTCTGAACGCTACTACCGAAAAATTTGGACACACGCTTTGGCTTCTGCTGAAAAATCC
ATTGCCGATCACGATCTCGCCTGAACGGTTTGGATAAAAACAGTGTGACAGCTGCGCAAGAAAACCCGCCAAGGCCTT
GCAGAAACAGCCGCGCTCTCGGTCTGTTCACACCTTACAACGTGGGTTCGGTTCATGTAACGGCTGCAGTCCGGCGG
TACAAATCCGAATCGGCAGTCCCATCGGTACCGGCTTCCGCTTTACCGAAAACCTTTGCCCGCAAAAGCAGCGTGGCA
GTCGGCACTTCGTCCGTTCTTCCGACGCTACCATGTTCGGCGTCAATTACGAGTGGGGATCCGGAGGAGGAGTCA
GATTTGGCAAACGATTCCTTTATCCGGCAGGTTCTCGACCGTCAGCATTTGAAACCCGACGGGAAATACCACCTATTC
GGCAGCAGGGGGAACTTGGCCGAGCGCAGCGCCATATCGGATTTGGGAAAAATACAAAGCCATCAGTTGGGCAACCTG
ATGATTCACAGGCGGCCATTAAAGGAAATATCGGCTACATTTGTCGCTTTCCGATCACGGGCACGAAGTCCATTCC
CCCTTCGACAAACCATCCCTCACATTCGATTTGATGAAGCCGGTAGTCCGTTGACGGATTTAGCCTTTACCGCATC
CATTTGGGACCGATACGAACACCATCCCGCCGACGGCTATGACGGCCACAGGGCGGCGCTATCCGCTCCCAAGGC
GCGAGGGATATATACAGCTACGACATAAAAGGCGTTGCCCAAAATATCCGCTCAACCTGACCGACAACCGCAGCACC
GGACAACGGCTTGCAGCAGCTTCCACAATGCCGGTAGTATGCTGACGCAAGGAGTAGCGGACGGATTCAAACCGGCC
ACCCGATACAGCCCGAGCTGGACAGATCGGGCAATGCCGCGGAGCCTTCAACGGCACTGCAGATATCGTTAAAAAC
ATCATCGGCGCGGAGGAGAAATTTGTCGGCGCAGGCGATGCCGTGCAGGGCATAAGCGAAGGCTCAAACATTTGCTGTC
ATGCACGGCTTGGGTCTGCTTCCACCGAAAACAAGATGGCCGCGCATCAACGATTTGGCAGATATGGCGCAACTCAA
GACTATGCCCGCAGCAGCCATCCGCGATTGGGCGAGTCCAAAACCCCAATGCCGCAAGGCAATAGAAGCCGTCAGCAAT
ATCTTTATGCGAGCCATCCCATCAAAGGATTTGGAGTGTTCGGGGAAAATACGGCTTTGGGCGGCATCACGGCACAT
CCTATCAAGCGGTTCGAGATGGGCGGATCGCATTTCCGAAAGGAAATCCGCGCTCAGCGACAATTTTGGCCGATGCG
GCATACGCCAAATACCCGTCCTTACCATTCCCGAAATATCCGTTCAAACCTTGGAGCAGCGTTACGGCAAGAAAAC
ATCACCTCTCAACCGTCCCGCTCAAACCGCAAAAATGTCAAACCTGGCAGACCAACGCCACCCGAAAGACAGGCGTA
CCGTTTACGGTAAAGGTTTCCGAATTTGAGAAGCACGTGAAATATGATACGCTCGAGCACCACCACCACCACCAC
TGA

Table with 6 columns: Line number (1-801), and six columns of amino acid sequences. The sequences are aligned in blocks of 50 lines. The last line (801) contains 'HHHHH*'.

961-741

ATGGCCACAAACGACGACGATGTTAAAAAGCTGCCACTGTGGCCATTGCTGCTGCCTACAACAATGGCCAAGAAATC
AACGGTTTCAAAGCTGGAGAGACCATCTACGACATTGATGAAGACGGCACAATTACCAAAAAGACGCAACTGCGACC
GATGTTGAAGCCGACGACTTAAAGGTCTGGGTCTGAAAAAGTCGTGACTAACCTGACCAAAACCGTCAATGAAAAC
AAACAAAACGTCGATGCCAAAAGTAAAAGCTGCAGAATCTGAAATAGAAAAGTTAACAAACCAAGTTAGCAGACACTGAT
GCCGCTTTAGCAGATACTGATGCCGCTCTGGATGCAACCACCAACGCCTTGAATAAATTTGGGAGAAAATATAACGACA
TTTGCTGAAGAGACTAAGACAAATATCGTAAAAATTTGATGAAAAATTAGAAGCCGTGGCTGATACCGTCGACAAGCAT
GCCGAAGCATTCAACGATATCGCCGATTCATTTGGATGAAACCAACACTAAGGCAGACGAAGCCGTCAAAACCGCCAAT
GAAGCCAAACAGACGGCCGAAGAAACCAACAAAACGTCGATGCCAAAAGTAAAAGCTGCAGAAAAGTGCAGCAGGCAAA
GCCGAAGCTGCCGCTGGCACAGCTAATACTGCAGCCGACAAGGCCGAAGCTGTCGCTGCAAAAAGTTACCGACATCAAA
GCTGATATCGCTACGAACAAAGATAATATTGCTAAAAAAGCAAACAGTGCAGCGTGTACACCAGAGAAGAGTCTGAC
AGCAAATTTGTCAGAATTTGATGGTCTGAACGCTACTACCGAAAAATTTGGACACACGCTTTGGCTTCTGCTGAAAAATCC
ATTGCCGATCACGATACTCGCTGAACGGTTTGGATAAAAACAGTGTGACAGCTGCGCAAGAAAACCCGCCAAGGCCTT

GCAGAACAAGCCGCGCTCTCCGGTCTGTTCACACCTTACAACGTGGGTCGGTTCATGTAAACGGCTGCAGTCGGCGGC
TACAAATCCGAATCGGCAGTCGCCATCGGTACCGGCTTCCGCTTTACCGAAAACCTTTCGCCCAAAGCAGGCGTGGCA
GTCCGACCTTCGTCGGTCTTCCGCGAGCCTACCATGTCCGGCTCAATTACGAGTGGGGATCCGGAGGGGGTGGTGT
GCCGCCACATCGGTGCGGGGCTTCCGATGCACTAACCGCACCGCTCGACCATAAAGACAAAGGTTTGCAGTCTTTG
ACGCTGGATCAGTCCGTCAGGAAAAACGAGAACTGAAGTTCGGCGGCACAAGGTTCGGGAAAAAACTTATGGAAACGGT
GACAGCCTCAATACGGGCAAATGAAGAACGACAAGGTTCAGCCGTTTCGACTTATCCGCCAAATCGAAGTGGACGGG
CAGCTCAATTACCTTGGAGAGTGGAGAGTCCCAAGTATACAAACAAAGCCATTCCGCTTAACCGCCTTTCAGACCGAG
CAAATACAAGATTCCGGAGCATTCCGGGAAGATGGTTCGGAAAACGCCAGTTCAGAATCGGCGACATAGCGGGCGAACAT
ACATCTTTTGACAAGCTTCCCAGAGGCGGCAGGGCGACATATCGCGGGACGGCGTTCGGTTCAGACGATGCCGGCGGA
AACTGACCTACACCATAGATTTCCCGCCAAAGCAGGGAACCGGCAAAATCGAACATTTGAAATCGCCAGAACTCAAT
GTCGACCTGGCCCGCCGATATCAAGCCGGATGGAAAACGCCATGCCGTATCAGCGGTTCCGCTTTTACAACCAA
GCCGAAAAGCAGTTACTCCCTCGGTATCTTTGGCGGAAAAGCCAGGAAAGTTGCCGGCAGCGCGGAAAGTGA AAC
GTAAACGGCATAACCCATATCGGCCCTTCCGCGCAAGCAACTCGAGCACACCACCACCACCCTGA

1 MATNDDVVK AATVAIAAAY NNGQEINGFK AGETIYDIDE DGTITKKDAT
51 AADVEADDFK GLGLKVVVN LTRVFNENKQ NVDKVKAAE SEIEKLTTKL
101 ADTDAALADT DAALDATTNA LNKLGENTT FAEEFKTNIV KIDEKLEAVA
151 DTVDKHAEAF NDIADSLDET NTKADEAVKT ANEAKQTAE TKQNVDAKV
201 AAETAAGKAE AAAGTANTAA DRABAVA KV TDIKADIATN KDNI AKKANS
251 ADVYTREESD SKFVRIDGLN ATTEKLDTRL ASAEKSIADH DTRNLGLDKT
301 VSDLRKFTRQ GLAEQAALSG LFPYVNGRF NVTAAVGGYK SESAVAIGTG
351 FRPTENFAAK AGVAVGTBSG SSAAYHVGVN YEWGSGGGV AADIGAGLAD
401 ALTAPLDHKD KGLQSLTLDQ SVRKNEKLLK AAQGAEKTYG NGDSLNTGKL
451 KNDKVS RFD IRQIEVDGQL ITLESGEFQV YKQSHSALTA FQTEQIQDSE
501 HSGK MVAKRQ FRIGDLAGEH TSPDKLPEGG RATYRGTAFG SDDAGGKLT
551 TIDFAAKQGN GKIEHLK SPE LNVDLAAADI KPDGKRHAVI SGSVLYNQAE
601 KGSYSLGIFG GKAEVAGSA EVKTVNGIRH IGLAARQLEH HHHH*

961-983

ATGGCCACAAACGACGACGATGTTAAAAAGCTGCCACTGTGGCCATTGCTGCTGCCCTACAACAATGGCCAAAGAAATC
AACCGTTTCAAAGCTGGAGAGACCATTACGACATGATGAAGACGGCACAATTACCAAAAAGACGCAACTGCAGCC
GATGTTGAAGCCGACGACTTAAAGGTCTGGGTCTGAAAAAGTCTGACTAACCTGACCAAAACCGTCAATGAAAAC
AAACAAAACGTCGATGCCAAAGTAAAAGCTGCAGAACTGAAATAGAAAAGTTAACAAACCAAGTTAGCAGACACTGAT
GCCCTTTAGCAGATACGATGCCGCTCTGGATGCAACCACCAACCGCTTGAATAAATGGGAGAAAATATAACGACA
TTTGCTGAAGAGACTAAGACAAATATCGTAAAAATGATGAAAAATAGAAGCCGTGGCTGATACCGTCGACAAGCAT
GCCGAAGCATTCAAGATATCGCCGATTCATGGATGAAACCAACACTAAGGCAGACGAAGCCGTCAAACCCGCAAT
GAAGCCAAACAGACGGCCGGAAGAAACCAAAACAAACCTGATGCCAAAGTAAAAGCTGCAGAACTGCAGCAGGCAAA
GCCGAAGCTGCCGCTGGCACAGCTAATACTGCAGCCGACAAGGCCGAAGCTGTCCGTGCAAAAAGTTACCGACATCAA
GCTGATATCGCTACGAACAAGATAATATGCTAAAAAGCAAAACAGTCCGACGCTGTACACCAGAGAAGAGTCTGCAC
AGCAAATTTGTCAGAAATGATGGTCTGAACGCTACTACGAAAATTTGGACACACGCTTGGCTTTCGCTGAAAAATCC
ATTTCCGATACAGTACTCGCTGAACGGTTTGGATAAAAACAGTGTGACACCTGCGCAAAGAAACCCGCAAGCCCTT
GCAGAAACAGCCGCTCTCCGGTCTGTTCACACTTACAACGTGGGTCGGTCAATGTAAACGGCTGCAGTCGGCGGC
TACAAATCCGAATCGGCAGTCGCCATCGGTACCGGCTTCCGCTTTACCGAAAACCTTTCGCCCAAAGCAGGCGTGGCA
GTCGGCACTTCTCCGGTCTTCCGCGACCTACCATGTCCGGCTCAATTACGAGTGGGGATCCGGCGGAGCGGCACT
TCTGCGCCGACTTCAATGCAGGCGGTACCGGTATCGGCAGCAACAGCAGGCAACACAGCGAAATCAGCAGCAGTA
TCTTACGCGCGTATCAAGAACGAAATGTGCAAAAGCAGAAAGTCTGCTGCGCGTTCGGGATGACGTTGCGGTTACA
GACAGCGGATGCCATAAATCAATGCCCCCGAACTTGCATACCGGAGACTTTCCAAACCCAAATGACGCATACAAG
AATTTGATCAACCTCAAACCTGCAATTTGAAGCAGGCTATACAGGACGCGGGGTAGAGGTAGGTATCGTCGACACAGGC
GAATCCGTCGGCAGCATATCTTTCCGAACTGTATGGCAGAAAAGAACAGGCTATAACGAAAATACAAAAACTAT
ACGGCGTATATGCCGAAGGAAGCGCCTGAAGACGGAGGCGGTAAAGACATTGAAGCTTCTTTTCGACGATGAGGCCGTT
ATAGAGACTGAAGCAAAGCCGACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATTGGTCTCCCATATTATT
GGCGGGCGTTCCGTGGACGGCAGACCTGCAGGCGGTATTCGCGCCGATGCGACGCTACACATAATGAATACGAATGAT
GAAACCAAGAACGAAATGATGGTTCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACGTTGGCGTCCGATCGTC
AATAACAGTTTTGGAACAACATCGAGGGCAGGCACCTGCCAGCTTTTCCAAATAGCCAATTCGGAGGAGCAGTACCGC
CAAGCGTTGCTCGACTATTCCGGCGGTGATAAAACAGACGAGGTATCCGCTGATGCAACAGAGCGATTACGGCAAC
CTGTCTACCACATCCGTAATAAAAAACATGCTTTTTCATCTTTTCGACAGGCAATGACGCACAAGCTCAGCCCAACACA
TATGCCCTATTGCCATTTTATGAAAAAGACGCTCAAAGAGCCATATCAAGTCCGAGGCGTAGACCCGAGTGGAGAA
AAGTTCAAACGGGAAATGATGAGAAACCGGTACAGAACCGCTTGAGTATGGCTCCAACCATTCGGGAAATTAAGTCC
ATGTGGTGCCTGTCCGCACCCATGAAGCAAGCGTCCGTTTACCCGTACAAACCCGATTCAAATTCGCCGAACATCC
TTTTCCGCACCCATCGTAACCGGCACGGCGCTCTGCTGCTGCAAAAATACCCGTTGATGAGCAACGCAACCTGCGT
ACCAGTTGCTGACGACGGCTCAGGACATCGGTGCAGTCCGCGTGGACAGCAAGTTCGGCTGGGACTGCTGGATCGG
GGTAAGGCCATGAACGGACCCGCTCTTTCCGTTCCGCGACTTTACCGCCGATACGAAAGGTACATCCGATATTGCC
TACTCTTCCGTAACGACATTTACGGCACGGCGGCTTATCAAAGAGCGGCGAGCCAACTGCAACTGCACGGCAAC
AACACCTATACGGGCAAAACCAATATCGAAGGCGTTCGCTGGTGTGTACGGCAACAACAATCGGATATCGCGCTC
GAAACCAAGGTGCGCTGATTTATAACGGGCGGCATCCGGCGGAGCCTGAACAGCGACGGCATGTCTATCTGGCA
GATACCGACCAATCCGGCGCAAACGAAACCGTACACATCAAAGGCGTCTGCAGCTGGACGGCAAAGGTACGCTGTAC

ACACGTTTGGGCAAACCTGCTGAAAGTGGACGGTACGGCGATTATCGGGCGCAAGCTGTACATGTCGGCACGGGCAAG
 GGGCAGGCTATCTCAACAGTACCGGACGACGTGTTCCCTTCTGAGTCCGCCAAAATCGGGCAGGATTATTCTTTC
 TTCACAAACATCGAAACCGACGGCGGCTGCTGGCTTCCCTCGACAGCGTCGAAAAACAGCGGGCAGTGAAGGCGAC
 ACGCTGTCCTATTATGTCGGTTCGGCGCAATGCGGCACGGACTGCTTCGGCAGCGGCACATCCGCGCCCGCGGCTG
 AAACACGCGGTAGAACAGGGCGGCAATCTGGAAAACCTGATGGTGAACCTGGATGCCTCCGAATCATCCGCAACA
 CCCGAGACGGTTGAAACTGCGGCAGCGACCGCAGATATGCGGGCATCCGCCCTACGGCGCAACTTTCGGCGCA
 GCGGCAGCCGTACAGCATGCGAATGCCGCCGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGAC
 AGTACCGCCGCCATGCCGATATGCAAGGACGCGCCTGAAAGCCGATACGGACGGGTGGACCAACAAGGCACGGGT
 CTGCGCGTATCGCGCAAACCAACAGGACGGTGGAAACAGGGCGGTGTTGAAGGCAAAAATGCGCGGCAGT
 ACCCAAACCGTCCGCATTGCGCGAAAACCGGCGAAAATACGACAGCAGCCGCACACTGGGCATGGGACGCGACACA
 TGGAGCGAAAACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAGCATGCCGGGATATC
 GGCTATCTCAAAGCCCTGTTCTCTACGGACGCTACAAAAACAGCATCAGCCGACACCGGTGGGACGAACATGCG
 GAAGCAGCGTCAACGGCAGCGTGCAGCTGGGCGCAGTGGGCGGTGTCACGTTCCGTTTGCCTGCAACGGGAGAT
 TTGACGGTCAAGGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTCGCCGAAAAGGCAGTGCCTTGGGCTGG
 AGCGGCAACAGCCTCACTGAAGGCAGCTGGTCCGACTCGGGGTCTGAAGCTGTCGCAACCCCTTGAGCGATAAAGCC
 GTCTGTTTGCACAGCGCGGCGTGGAAACGCGACCTGAACGGACCGGACTACCGGTAACGGCGGCTTACCGGCGG
 ACTGCAGCAACCGGCAAGACGGGGCAGCAATATGCCGCACACCCGCTGCGTGGTTCGCCGCTGGGCGCGGATGTCGAA
 TTCGGCAACGGTGGAAACGGCTTGGCAGTTACAGCTACCGCGGTTCCAAACAGTACGGCAACCACAGCGGACGAGTC
 GCGTAGGCTACCGGTTCCCTCGAGCACCACCACCACCCTGA

1	MATNDDVKK	AATVALAAAY	NNGQEINGFK	AGETIYDIDE	DGTITTKDAT
51	AADVEADDFK	GLGLKVVVN	LTRFVNENKQ	NVDAKVKAAE	SEIEKLTTKL
101	ADTDAALADT	DAALDATMA	LNKLGENITT	FAEETKTNIV	KIDFKLEAVA
151	DTVDKHAEAF	NDIADSLDET	NTKADRAVKT	ANEAKQTAE	TKQNVDAKV
201	AAETAAGKAE	AAAGTANTAA	DKAEVAAKV	TDIKADIATN	KDNLAKKANS
251	ADVTFREESD	SKFVRIDGLN	ATTEKLDTRL	ASAESLADH	DTRLNGLDKT
301	VSDLRKETRO	GLAEQAALSG	LFQPNVGRF	NVTAAVGGYK	SESAVAIGTG
351	FRFTENFAAK	AGVAVGTSSG	SSAAYHVGVN	YEWGSGGGT	SAPDFNAGGT
401	GIGSNSRATT	AKSAAVSYAG	IKNEMCKDRS	MLCAGRDDVA	VTDRDAKINA
451	PPPNLHTGDF	PNPNDAYKML	INLKPATBAG	YTGGRVVEGI	VDTGESVSGI
501	SFPELYGRKE	HGYNENYKNY	TAYMRKBAPE	DGGGKDIEAS	FDDEAVIETE
551	AKPTDIRHVK	EIGHIDLVSF	IIGGRSVDGR	PAGGIAPDAT	LHDMNTNDET
601	KNEMVVAAIR	NAWVKLGERG	VRIVNSFGT	TSRAGTADLF	QLANSEEQYR
651	QALLDYSGED	KTDEGIRLMQ	QSDYGNLSYH	IRNKNMLFIF	STGNDAQAQP
701	NTYALLPFYE	KDAQGIITV	AGVDRSGEKF	KREMYGEPGT	EPLEYGSNHK
751	GITAMWCLSA	PYEASVRFTR	TNPIQIAGTS	PSAPIVTGTA	ALLLQKYPWM
801	SNDNLRTTLL	TTAQDIGAVG	VDSKFGWLL	DAGKAMNGPA	SFPFGDFTAD
851	TKGTSDIAYS	FRNDISGTGG	LIKKGGSQIQ	LHGNNFTYTK	TILEGGSVLV
901	YGNNSDMRV	ETKALTYNG	AASGGSLSND	GIVYLADTDQ	SGANETVHIK
951	GSLQLDGKGT	LYTRLGKLLK	VDGTAIIGGK	LYMSARGKGA	GYLNSTGRRV
1001	PFLSAAKIGQ	DYSFFTNIET	DGGLLASLDS	VEKTAGSEGD	TLSYVVRGN
1051	AARTASAAAH	SAPAGLKHAV	EQGGSNLENL	MVELDASESS	ATPETVETAA
1101	ADRTDMPGIR	PYGATFRAAA	AVQHANAADG	VRIFNSLAT	VYADSTAHA
1151	DMQGRRLKAV	SDGLDHNGTG	LRVLAQTQOD	GGTWEQGGVE	GKMRGSTQTV
1201	GIAAKTGENT	TAAATLGMGR	STWSSENSANA	KTDSISLFFAG	IRHDAGDIGY
1251	LKGLFSYGRY	KNSISRSTGA	DEHAEGSVNG	TLMQLGALGG	VNVPPAATGD
1301	LTVEGGLRYD	LLKQDAFAEK	GSALGWSGNS	LTEGTLVGLA	GLKLSQFLSD
1351	KAVLFATAGV	ERDLNGRDYT	VTGGFTGATA	ATGKTGARNM	PHTRLVAGLG
1401	ADVEFENGWN	GLARYSYAGS	KQYGNHSGRV	GVGYRFLEHH	HHHH*

961c-ORF46.1

ATGGCCACAAAACGACGACGATGTTAAAAAGCTGCCACTGTGGCCATTGCTGCTGCCTACAACAATGGCCAAGAAATC
 AACGGTTTCAAAGCTGGAGAGACCATCTACGACATTTGATGAAGACGGCACAATACCAAAAAAGACGCAACTGCAGCC
 GATGTTGAAGCCGACGACTTTAAAGGTCTGGGTCTGAAAAAGTTCGTGACTAACCTGACCAAAAACCGTCAATGAAAAA
 AAACAAAACGTCGATGCCAAAGTAAAGCTGCAGAACTGAAATAGAAAAAGTTAACAAACCAAGTTAGCAGACACTGAT
 GCCGCTTTAGCAGATACTGATGCCGCTCTGGATGCAACCACCAACGCCCTTGAATAAATTTGGGAGAAAATATAACGACA
 TTTGCTGAAGAGACTAAGACAAATATCGTAAAAATTTGATGAAAAATTAGAAGCCGTGGCTGATACCGTCGACAAGCAT
 GCCGAAGCATTTCAACGATATCGCCGATTCAATGGATGAAACCAACACTAAGGCAGACGAAGCCGTCAAAACCGCCAAT
 GAAGCCAAAACGACCGCCGAAGAAACCAAAACAAACCTCGATGCCAAAGTAAAAGCTGCAGAACTGCAGCAGGCAAA
 GCCGAAGCTGCCCGCTGGCACAGCTAATACTGCAGCCGACAAGCCGAAGCTGTCGCTGCAAAAGTTACCGACATCAA
 GCTGATATCGCTACGAAACAAAGATAAATTTGCTAAAAAGCAAAACAGTGCAGCGTGTACACCAGAGAAGAGCTGTGAC
 AGCAAAATTTGTGAGAAATTTGATGGTCTGAAACGCTACTACCGAAAAATTTGGACACACGCTTTGGCTTCTGCTGAAAAATCC
 ATTGCCGATCAGGACTCGCTGAAACGGTTTGGATAAAACAGTGTGACAGCTGCGCAAAAGAAACCCGCCAAGGCCCTT
 GCAGAACAGCCGCGCTCTCCGCTCTGTTCCAACCTTACAACCTGGGTGGATCCGGAGGAGGAGGATCAGATTTGGCA
 AACGATCTTTTATCCGGCAGGTTCTCGACCGTCAAGCTTTCGAAACCCGACGGGAAATACCACCTATTCCGCAGCAGG
 GGGAACTTGCAGCGCAGCGCCATATCGGATTTGGAAAAATACAAAGCCATCAGTTGGGCAACCTGATGATTCAA

CAGGCGGCCATTAAAGGAAATATCGGCTACATTGTCCGCTTTTCCGATCACGGGCACGAAGTCCATTCCCCCTTCGAC
AACCATGCCTCACATTCGGATTCTGATGAAGCCGGTAGTCCCCTGACGGATTAGCCCTTACCGCATCCATTGGGAC
GGATACGAACACCATCCCGCCGACGGCTATGACGGGCCACAGGGCGGGCTATCCCGCTCCCAAAGCGCGAGGGAT
ATATACAGCTACGACATAAAAGCGTTGCCAAAATATCCGCTCAACCTGACCGACAACCGCAGCACCGGACAACGG
CTTGCCGACCGTTTCCCAATGCCGGTAGTATGCTGACGCAAGGAGTAGGCGACGGATTCAAACGCCACCCGATAC
AGCCCCGAGCTGGACAGATCGGGCAATGCCGCCGAAGCCTTCAACGGCACTGCAGATATCGTTAAAAACATCATCGGC
GCGCAGGAGAAATGTGCGGCGCAGGCGATGCCGTGACGGGCATAAGCGAAGGCTCAAACATTTGCTGTATGCACGGC
TTGGGTCTGCTTTCCACCGAAAACAAGATGGCGCGCATCAACGATTTGGCAGATATGGCGCAACTCAAAGACTATGCC
GCAGCAGCCATCCGCGATTGGGCAGTCCAAAACCCCAATGCCGCACAAGGCATAGAAGCCGTCAGCAATATCTTTATG
GCAGCCATCCCATCAAAGGGATTGGAGCTGTTCCGGGAAAATACGGCTTGGGCGGCATCACGGCACATCCCTATCAAG
CGGTCGAGATGGGCGCGATCGCATTTGCCGAAAGGGAAATCCGCGTCAGCGACAATTTTGGCGATGCGGCATACGCC
AAATACCCGTCCTTACCATTCCGAAAATATCCGTTCAAACCTGGAGCAGCGTTACGGCAAAGAAAACATCACCTCC
TCAACCGTCCGCGCTCAAACGGCAAAAATGTCAAACCTGGCAGACCAACGCCACCCGAAGCAGGGCTACCGTTTGAC
GGTAAAGGGTTTCCGAATTTGAGAAGCACGTGAAATATGATACGCTCGAGCACCACCACCACCACCCTGA

- 1 MATNDDDVKK AATVAIAAAY NNGQEINGFK AGETIYDIDE DGTITKKDAT
51 AADVEADDFK GLGLKRVVFN LTKTVNENKQ NVDAKVKAAE SEIEKLTTKL
101 ADTDAALADT DAALDATTNA LNKLGENITF FAEETRTNIV KIDEKLEAVA
151 DTVDKHAEAF NDIADSLDET NTKADEAVKT ANEAKQTAE TKQNVDAKVK
201 AAETAAGKAE AAAGTANTAA DKAEVAARKV TDIKADLATN KDNIARKANS
251 ADVYTREESD SKFVRIDGLN ATTEKLDTRL ASAEKSIADH DTRLNGLDKT
301 VSDLRKETRQ GLAEQAALSG LFQPYNVGGS GGGGSDLAND SFIRQVLDRO
351 HFEPDGYKYL FGSRGELAER SGHIGLKIQ SHQLGNLMIQ QAAIKGNIGY
401 IVRFSHDGHE VHSPFDNHAS HSDSDEAGSP VDGFSLYRIH WDGVEHHPAD
451 GYDGPQGGY PAKPGARDIY SYDIKVAQN IRLNLDNRN TGQRLADRFH
501 NAGSMLTQGV GDGFKRATRY SPELDRSGNA ARAFNGTADI VKNLIGAAGE
551 IVGAGDAVQG ISEGSNIAVM HGLGILLSTEN KMARINDLAD MAQLKDYAAA
601 AIRDWAVQNP NAAQGIKAVS NIFMAAIIPIK GIGAVRGKYG LGGITAHPIK
651 RSQMGAIALP KGKSAVSDNF ADAAYAKYPS PYHSRNRSN LEQRYGKINI
701 TSSTVPPSNG KNVKLADQRH PKTGVVFDGK GFNPFKHKVH YDTLEHHHHH
751 H*

961c-741

ATGGCCACAAACGACGACGATGTTAAAAAGCTGCCACTGTGGCCATTGCTGCTGCCTACAACAATGGCCAAGAAATC
AACGGTTTCAAAGCTGGAGAGACCATCTACGACATTTGATGAAGACGGCACAATTACCAAAAAGACGCAACTGCAGCC
GATGTTGAAGCCGACGACTTTAAAGTCTGGGTCTGAAAAAGTCCGTGACTAACCTGACCAAAAACCGTCAATGAAAAC
AAACAAAACGTCGATGCCAAAGTAAAAGCTGCAGAATCTGAAATAGAAAAGTTAACAACCAAGTTAGCAGACACTGAT
GCCGCTTTAGCAGATACTGATGCCGCTCTGGATGCAACCACCAACGCCCTTGAATAAATTTGGGAGAAAATATAACGACA
TTTGCTGAAGAGACTAAGACAATATCTGTAATAATTTAGAAAATTTAGAAGCGTGGCTGATACCGTCGACAAGCAT
GCCGAAGCAITCAAAGATATCGCCGATTCATTGGATGAAACCAACTAAGGCAGACGAAGCGTCAAACCCGCCAAT
GAAGCCAAACAGACGGCCGAAGAAACCAAAACAAACCGTCGATGCCAAAGTAAAAGCTGCAGAAACTGCAGCAGGCAA
GCCGAAGCTGCCGCTGGCAGACTAATACTGCAGCCGACAAAGCCGAAGCTGTGCTGCAAAAAGTTACCGACATCAA
GCTGATATCGCTACGAACAAGATAATATTTGCTAAAAAAGCAAACAGTCCGACGCTGTACACCAGAGAGAGTCTGAC
AGCAAATTTGTCAGAATTTGATGGTCTGAACGCTACTACCGAAAATTTGGACACACGCTTGGCTTCTGTGAAAAATCC
ATTTCCGATCAGATACTCCCTGAACGTTTGGATAAAACAGTGTGACAGCTGCCCAAAGAAACCCGCAAGCCCTT
GCAGAACAAAGCCGCTCTCCGGTCTGTTCCAACCTTACAACGTGGGTGGATCCGGAGGGGGTGGTGTGCGCCCGGAC
ATCGTGGCGGGCTTGCCGATGCACTAACCCGACCGCTCGACCATAAAGACAAAGTTTGCAGTCTTTGACGCTGGAT
CAGTCCGTGAGAAAACGAGAACTGAAGCTGGCGGCACAAGGTGCCGAAAAAATTTATGAAAACGGTGCAGCCTC
AATACGGGCAAAATGAAGAACGACAAGGTCAGCCGTTTCGACTTTATCCGCCAAATCGAAGTGGACGGGCAGCTCAT
ACCTTGGAGAGTGGAGAGTTCCAAGTATACAAACAAGCCATTCGGCTTAAACCGCTTTTCAGACCCGAGCAAATACAA
GATTCGGCAATTCGGGAAAGATGGTTGCGAAACGCCAGTTTCAAGATCGCGACATAGCGGGCGAACATACATCTTTT
GACAAGCTTCCCGAAGGCGGACGGCGACATATCGCGGACGGCGTTCCGTTTTCAGACGATGCCGGCGGAAAACCTGAC
TACACCATAGATTTCCGCGCAAGCAGGGAAAACGGCAAAATCGAACATTTGAAATCGCCGAAACTCAATGTGACCTG
GCCGCGCCGATATCAAGCCGGATGAAAACGCCATGCCGTGATCAGCGGTTCCGCTCTTTACAACCAAGCCGAGAAA
GGCAGTTACTCCCTCGGTATCTTTGGCGGAAAAGCCAGGAAAGTTGCCGCGCAGCGGGAAGTGAACCCGTAACCGG
ATACGCCATATCCGCTTCCGCAAGCAACTCGAGCACCACCACCACCACCCTGA

- 1 MATNDDDVKK AATVAIAAAY NNGQEINGFK AGETIYDIDE DGTITKKDAT
51 AADVEADDFK GLGLKRVVFN LTKTVNENKQ NVDAKVKAAE SEIEKLTTKL
101 ADTDAALADT DAALDATTNA LNKLGENITF FAEETRTNIV KIDEKLEAVA
151 DTVDKHAEAF NDIADSLDET NTKADEAVKT ANEAKQTAE TKQNVDAKVK
201 AAETAAGKAE AAAGTANTAA DKAEVAARKV TDIKADLATN KDNIARKANS
251 ADVYTREESD SKFVRIDGLN ATTEKLDTRL ASAEKSIADH DTRLNGLDKT
301 VSDLRKETRQ GLAEQAALSG LFQPYNVGGS GGGGVAADIG AGLADALTAP
351 LDHKDRGLQS LTLDDQSVRKN EKLKLAQGA ERTYGNGLSL NTGKLANDKV
401 SRFDFIRQIE VDGQLITLES GEFOVYKQSH SALTAFOQTEQ IQDSEHSRGM

451 VAKRQFRIGD IAGEHTSFDK LPEGGRATYR GTAFGSDDAG GKLTYTIDFA
 501 AKQNGKIEH LKSPELNVDL AAADIKPDGK RHAVISGSVL YNQAERGSYS
 551 LGIFGGKAQE VAGSAEVKTV NGIRHIGLAA KQLEHHHHH *

961c-983

ATGGCCACAAACGACGACGATGTTAAAAAGCTGCCACTGTGGCCATTGCTGCTGCCCTACAACAAATGGCCAAGAAATC
 AACGGTTTCAAAGCTGGAGAGACCATCTACGCATTTGATGAAGACGGCACAATTACCAAAAAAGACGCAACTGCAGCC
 GATGTTGAAGCCGACGACTTTAAAGGTCTGGGTCTGAAAAAGTCTGACTAACCTGACCAAAACCGTCAATGAAAAC
 AAACAAAACGTCGATGCCAAAGTAAAGCTGCAGAATCTGAAATAGAAAAGTTAACAAACCAAGTTAGCAGACACTGAT
 GCCGCTTTAGCAGATACTGATGCCGCTCTGGATGCAACCACCAACCGCTTGAATAAATGGGAGAAAATATAACGACA
 TTTGCTGAAGAGACTAAGACAAAATATCGTAAAAAATTTGATGAAAAATFAGAAGCCGTGGCTGATACCGCTCGACAAGCAT
 GCCGAAGCATTCAACGATATCGCCGATTCAATTTGGATGAAACCAACACTAAGGCAGACGAAGCCGTCAAACCGCCAAT
 GAAGCCAAACAGACGGCCGAAGAACCAACAAAACGTCGATGCCAAAGTAAAGCTGCAGAAAATGCAGCAGGCAAAA
 GCCGAAGCTGCCGCTGCCACAGCTAATACTGCAGCCGACAAAGCCGAAGCTGTGCTGCAAAAAGTTACCGCATCAAA
 GCTGATATCGCTACGAACAAAGATAATATTTGCTAAAAAGCAACAGTGGCCGAGTGTACACCAGAGAAGAGTCTGAC
 AGCAAAATTTGTCAGAATTTGATGGTCTGAACGCTACTACCGAAAAATTTGGACACACGCTTGGCTTCTGCTGAAAAATCC
 ATTGCCGATCAGATACTCGCCTGAACGGTTTGGATAAAAACAGTGTGACACCTGCGCAAAAGAAACCCGCCAAGGCCCTT
 GCAGAACAAGCCGCTCTCCGCTCTGTTCCAACTTACACGTTGGTGGATCCGGCGGAGGCGGCACCTTCTGCGCCC
 GACTTCAATGCAGGCGTACCGGTATCGGCAGCAACAGCAGAGCAACAACAGCGAAATCAGCAGCAGTATCTTACGCC
 GGTATCAAGAAGCAATTTGCAAAAGACAGAAGCATGCTCTGTGCCGGTCCGGATGACGTTTGGGTTACAGACAGGGAT
 GCCAAAATCAATGCCCCCCCGAATCTGCATACCGGAGACTTTCAAACCCAAATGACGCATACAAGAATTTGATC
 AACCTCAAACCTGCAATGAAGCAGGCTATACAGGACCGGGGTAGAGGTAGGTATCGTCGACACAGGCCAATCCGTC
 GGCAGCATATCTTTCCCGAATGATGTCAGAAAAGAACACCGGCTATAACGAAAATTAACAAAATATACGGCGTAT
 ATGCGGAAGGAAGCCCTGAAGACGGAGGCGGTAAAGACATTTGAAGCTTCTTTTCGACGATGAGGCGGTATAGAGACT
 GAAGCAAGCCGACGGATATCCGCCAGTAAAGAAATCGGACACATCGATTTGGTCTCCCATATTTATTTGGCGGGCGT
 TCCGTGGACCGCAGACTGCAGGCGGTATTGCGCCCGATGCGCAGCTACACATAATGAATACGAATGAAACCAAG
 AACGAAATGATGGTTGCAGCCATCCGCAATGCATGGGTCAAGCTGGCGAACCTGGCGTCCGATCGTCAATAACAGT
 TTTGGAACAACATCGAGGGCAGGCACTGCCGACTTTTCCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCGTTG
 CTCGACTATTTCCGGCGGTGATAAAACAGACGAGGGTATCCGCTGATGCAACAGAGCGATACCGCAACCTGTCTTAC
 CACATCCGTAATAAAAACATGCTTTTCATCTTTTCGACAGGCAATGACGCACAAGCTCAGCCCAACACATATGCCCTA
 TTTGCCATTTTATGAAAAAGACGCTCAAAAAGGCATTTACACAGTGCAGGCGTAGACCGGCAAGAAGTTTCAA
 CGGAAATGTATGAGAACCGGGTACAGAACCGCTTGTGATGTTGCTTCAACCATTTGCGGAATTTACTGCCATGTTGGTGC
 CTGTCCGCCACCTATGAAGCAAGCGTCCGTTTCAACCGTACAAACCCGATTCAAATTTGCCGGAACATCTTTTCCGCA
 CCCATCGTAACCGGCACGGCGGCTCTGCTGCTGCAGAAATACCCGTGGATGAGCAACGCAACCTCGGTACCAGTTG
 CTGACGACGGCTCAGGACATCGGTGCAGTCCGGTGGACAGCAAGTTCCGGCTGGGGACTGCTGGATGCGGGTAAAGCC
 ATGAACGGACCCCGCTCTTTCCGTTCCGGCAGCTTTACAGCTGCAGGCTACGAAAGGTFACATCCGATATTGCCACTCTCTC
 CGTAACGACATTTTAGGACGGCGGCTGATCAAAAAGGCGGCAGCCAACCTGCAACTGCACGGCAACAACACCTAT
 ACGGGCAAAACCATTTATCGAAGGCGGTTCCGCTGGTGTGATCGGCAACAACAATTCGGATATGCGCGTGAACCAAA
 GGTGCGCTGATTTATAACGGGGCGGCATCCGGCGGACGCTGAACAGCGACGGCATTTGCTATCTGGCAGATACCGAC
 CAATCCGGCGCAAACGAAACCGTACACATCAAAAGGCAGTCTGCAGCTGGACGGCAAAGGTACGGCTGTACACGCTTTG
 GCAACCGTCTGAAAGTGGACGGTACGGCGATTATCGCGGGCAAGCTGTACATGTCGGCAGCGCAAGGAGGGCAGGC
 TATCTCAACAGTACCGGACGACGTTTCCCTTCTGAGTCCCGCCAAAATCGGGCAGGATTAATCTTTCTTCAAAAAC
 ATCGAAACCGACGGCGGCTGCTGGCTTCCCTCGACAGCGTTCGAAAAACAGCGGGCAGTGAAGGGACACCGCTGTCC
 TATTTATGTCGCTCGCGGCAATGCGGCACGGACTGCTTCCGACAGCGGCACATTCGCGCCCGCCGGTCTGAAACACGCC
 GTAGAAACAGGGCGGCAGCAATCTGGAAAACCTGATGGTTCGAATGGATGCCTCCGAATCATCCGCACACCCGAGACG
 GTTGAACACTGCGGCAGCGACCGCACAGATATGCCGGCATCCGCCCTACGGCGCAACTTTTCGCGCAGCGGCAGCC
 GTACAGCATCGGAATGCCCGGACCGGTGATCGCATCTTCAACAGTCTCGCCGCTACCGTCTACCGTACAGTACCGCC
 GCCCATGCCGATATGCAGGGACCGCCCTGAAAGCCGATTCGGACGGGTTGGACCACAACGGCACGGGTCTGCGCGTC
 ATCGCGCAAAACCAACAGGACGGTGAACGTTGGAAACAGGGCGGTGTTGAAGGCAAAATCGCGGGCAGTACCCAACCC
 GTCGGCATTTCCCGCAAAACCGGCGAAAATACGACAGCAGCCGCCACACTGGGCATGGGACGCAGCATGGAGCGAA
 AACAGTGCAAATGCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCACGATGCGGGCGATATCGGCTATCTC
 AAAGCCCTGTTCTCTACGGACGCTACAAAACAGCATCAGCCGACAGCCCGTGCAGGACGAAACATGCGGAAGGCAGC
 GTCAACGGCACGCTGATGCAGCTGGGCGCACGGCGGTGTCACGTTCCGTTTCCGCAACGGGAGATTTGACGGTC
 GAAGCGGTCTGCGCTACGACCTGCTCAAACAGGATGCATTTCCCGAAAAGGAGTGTCTTTGGGCTGGAGCGGCAAC
 AGCCTCACTGAAGGCACGCTGGTCCGACTCGCGGTCTGAGCTGTGCAACCTTTGAGCGATAAAGCCGTCTGTTT
 GCAACGGCGGGCGTGAACCGGACCTGAACGGACCGCATACCGGTAACGGGCGGCTTTACCGCGCGACTGCAGCA
 ACCGGCAAGACGGGGCACGCAATATGCCGCACACCCGCTGTTGTTGCCGCGCTGGGCGCGGATGTCGAATTCGGCAAC
 GGCTGGAACGGCTTTGGCAGTTACAGCTACGCCGTTTCCAAACAGTACGGCAACCACAGCGGACGAGTGGCGTAGGC
 TACCGGTTCTCGAGCACCACCACCACCACCTGA

1 MATNDDVVK AATVAIAAY NNGQEINGFK AGETIYDIDE DGTITKDAT
 51 AADVEADDFK GLGLKVVVN LTRTVNENKQ NVDKVKAAE SEIEKLTTKL
 101 ADTDAALADT DAALDATNA LNKLGENTIT FAEFRTNIV KIDEKLEAVA
 151 DTVDKHBAF NDIADSLDET NTKADEAVKT ANEAKQTAE TKQNVDAKV
 201 AAFTAAGKAE AAAGTANTAA DKAEVAARKV TDIKADLATN KDNIAKKANS
 251 ADVVTRRES D SKFVRIDGLN AITTEKLDTRL ASAEKSIADH DTRLNGLDKT

301	VSDLRKETRO	GLAEQAALSG	LFQPYNVGGS	GGGGSAPDF	NAGGTGIGSN
351	SRATTAKSAA	VSYAGIKNEM	CKDRSMLCAG	RDDVAVTDRD	AKINAPPPNL
401	HTGDFPNPND	AYKNLNLKLP	AIEAGYTPRG	VEVGIVDTGE	SVGSISFPPEL
451	YGRKEHGYNE	NYKNYTAYMR	KEAPEDEGGK	DIEASFDEEA	VIETFAKPTD
501	IRHVKEIGHI	DLVSHIIGGR	SVDGRPAGGI	APDATLHLMN	TNDETKNEMM
551	VAAIRNAWVK	LGERGVRIVN	NSFGTTSRAG	TADLFQIANS	EEQYRQALLD
601	YSGGDKTDEG	IRLMQQSDYG	NLSYHIRNKN	MLFIFSTGND	AQAQPNTYAL
651	LPPFEKDAQK	GIITVAGVDR	SGEKFKREMY	GEPGTEPLEY	GSNHCGITAM
701	WCLSAPYEAS	VRFTRTNPIQ	IAGTFSFSAPI	VTGTAALLLQ	KYPWMSNDNL
751	RT'TLLT'TAQD	IGAVGVDSKF	GWGLLDAGKA	MNGPASFFPG	DFTADTKGTS
801	DIAYSFRNDI	SGTGGLIKKG	GSQQLQHGNN	TYTGKTIIEG	GSLVLYGNKN
851	SDMRVETKGA	LIYNGAASGG	SLNSDGIIVYL	ADTDQSGANE	TVHKGSLQL
901	DGKGTLYTRL	GKLLKVDGTA	IIGGKLYMSA	RKKGAGYLNS	TGRRVFFLSA
951	AKIGQDYSFF	TNIETDGGLL	ASLDSVEKTA	GSEGD'FLSY	VRRGNAARTA
1001	SAAAHSAPAG	LKHAVEQGG	NLENLMVELD	ASESSATPET	VETAAADR'TD
1051	MPGIRPYGAT	FRAAAVQHA	NAADGVRIFN	SLAATVYADS	TAHADMQGR
1101	RLKAVSDGLD	HNGTGLEVIA	QTQODGGTWE	QGGVEGKMRG	STQTVGIAAK
1151	TGENTTAAAT	LGMGRSTWSE	NSANAKTDSI	SLFAGIRHDA	GDIGYLGKLF
1201	SYGRYKNSIS	RSTGADEHAE	GSVNGTLMQL	GALGGVNVFF	AATGDLTVEG
1251	GLRYDLLKQD	AFAEKGSALG	WSGNSL'LEGT	LVGLAGLKL	QPLSDKAVLF
1301	ATAGVERDLN	GRDYTVTGGF	TGATAATGKT	GARNMPHTRL	VAGLGADVEF
1351	GNGWNLARY	SYAGSKQYGN	HSGRVGVGYR	FL'EH'HH'HHH*	

961cL-ORF46.1

ATGAAACACTTTCCATCCAAAGTACTGACCACAGCCATCCTTGCCACTTTCTGTAGCGGCGCACTGGCAGCCACAAC
GACGACGATGTTAAAAGCTGCCACTGTGGCCATTGTCTGCTGCCTACAACAATGGCCAAGAAATCAACGGTTTCAA
GCTGGAGAGACCTACTACGACATTGATGAAGACGGCACAATTACCAAAAAGACGCAACTGCAGCCGATGTTGAAGCC
GACGACTTTAAAGGTCTGGGTCTGAAAAAGTCTGTACTAACCTGACCAAAACCGTCAATGAAAACAACAAAACGTC
GATGCCAAAGTAAAAGCTGCAGAATCTGAAATAGAAAAGTTAACAAACCAAGTTAGCAGACACTGATGCCGCTTTAGCA
GATACTGATCCGCTCTGGATGCAACCACCAACGCCTTGAATAAATTGGGAGAAAATATAACGACATTTGCTGAAGAG
ACTAAGACAAATATCGTAAAATTGATGAAAATTAGAAGCCGTGGCTGATACCGTCGACAAAGCATGCCGAAGCATTC
AACGATATCGCCGATTCATGGATGAAACCAACACTAAGGCAGACGAAGCCGTCAAACCGCCAATGAAAGCCAAACAG
ACGGCCGAAAGAACCAACAAAACCGTCGATGCCAAAGTAAAAGCTGCAGAACTGCAGCAGGCAAAGCCGAAGCTGCC
GCTGGCACAGCTAATACTGCAGCCGACAGGCCGAAAGCTGTCTGCTGCAAAAGTTACCGACATCAAAGCTGATATCGCT
ACGAACAAAGATAATATTTGCTAAAAAGCAAACAGTCCCGACGCTGTACACCAGAGAAGAGTCTGACAGCAAATTTGCT
AGAATGATGGTGTGAACGCTACTACCGAAAATTTGACACAGCTTGGCTTCTGCTGAAAATCCATTTGCCCGATCAC
GATACTGCCTGAAACGGTTTGGATAAAAACAGTGTCTAGACCTCGCCAAAGAAACCCGCAAGCCCTTGCAGAACAGCC
GGCTCTCCGGTCTGTTCCAACCTTACAACGTGGGTGGATCCGGAGGAGGAGATCAGATTTGGCAAACGATTTCTTTT
ATCCCGCAGGTTCTCGACCGTCAGCATTTCGAACCCGACGGGAAATACCACCTATTCGGCAGCAGGGGGAACTTGCC
GAGCCGAGCGCCATATCGGATTTGGGAAAATACAAAGCCATCAGTTGGGCAACCTGATGATTCACAGCGCGCCATT
AAAGGAAATATCGGCTACATTTGTCGCTTTTCCGATCACGGGCAGAAAGTCCATTTCCCTTCGACAACCATGCGCTCA
CATTCGATTTCTGATGAAGCCGGTAGTCCCGTTGACGGATTTAGCCTTTTACCGCATCCATTGGGACGGATACGAACAC
CATCCCGCCGACGGCTATGACGGGCCACAGGGCGGGGCTATCCCGCTCCCAAAGGCGGAGGGATATATACAGCTAC
GACATAAAGCGGTTGCCCAAATATCCGCTCAACCTGACCGACAACCGCAGCACCGGACAACGGCTTGCCGACCGT
TTCCACAATGCCGCTAGTATGCTGACGCAAGGAGTAGCGCAGCGATTCAAACCGGCCACCGGATACAGCCCGAGCTG
GACAGATCGGGCAATGCCCGGAAGCCTTCAACGGCACTGCAGATATCGTTAAAACATCATCGGCGGGCAGGAGAA
ATTGTCGGCGCAGCATGCGCTGCAGGGCATAAGCGAAGCTCAAACATTTGCTGTATGCATGCAGCGCTTGGGTCTGCTT
TCCACCGAAAACAAGATGGCGCGCATCAACGATTTGGCAGATATGGCGCAACTCAAAGACTATGCCGAGCAGCCATC
CGCGATTTGGGCGATCCAAAACCCCAATGCCGCACAAGGCATAGAAGCCGTCAGCAATATCTTTATGGCAGCCATCCCC
ATCAAAGGATTTGGAGCTGTTCCGGGAAAATACGGCTTGGGCGGCATCACGGCACATCTATCAAGCGGTCGAGATG
GGCGGATCGCATTTGCCGAAAGGAAATCCGCGCTCAGCGACAATTTGCCGATGCCGCATACGCCAAATACCCGCTCC
CCTTACCATTTCCGAAATATCGTTCAAACCTTGGAGCAGCGTTACGGCAAAGAAAACATCACCTCCTCAACCGTGCCG
CCGTCAAACGGCAAATAATGTCAAACCTGGCAGACCAACGCCACCCGAAAGACAGGCGTACCGTTTGGACGGTAAAGGGTTT
CCGAATTTTGAAGACAGTGAATATGATACGTAACCTCGAG

1	MKHFPKSVLT	TAILATPCSG	ALAA'TNDDDV	KKAATVAIAA	AYNNGQEING
51	FKAGETIYDI	DEDGTITRKD	ATAADVEADD	FKGLGLKQV	TNLTKIVNEN
101	KQNVDAKVK	AESEIEKLT	KLADTDAALA	D'FDAALDATT	NALNKLGENI
151	TTFAEETKT	IVKIDEKLEA	VADTVDKHAE	AFNDIADSLD	BTNTRADEAV
201	KTANEAKQTA	EETKQNVDAK	VKAAETAAGK	AREAAAGTANT	AADKAEAVAA
251	KVTDIKADIA	TNKDNIAKKA	NSADV'TREE	SDSKFVRIDG	LNATTEKLDT
301	RLASAERKIA	DHDTRLNGLD	KTVSDLRKET	RQGLAEQAAL	SGLFPYNVG
351	GSGGGGSDLA	NDSFIRQVLD	RQHFEPDGKY	HLFGSRGELA	ERSGHIGLGK
401	IQSHQLGNLM	IQQAAIKGN	GYIVRFSDEH	HEVHSPFDNH	ASHSDSDEAG
451	SFVDGFSLYR	IHWGVEEHP	ADGYDGPQGG	GYPAPKAGARD	IYSYDIKQVA
501	QNI'RLN'LDN	RSTGQRLADR	PHNAGSMLTQ	GVGDGFKRAT	RYSPELDRSG

551 NAAEAFNGTA DIVKNIIGAA GEIVGAGDAV QGISEGSNIA VMHGLGLLST
601 ENKMARINDL ADMAQLKDYA AAAIRDWAVQ NPNAOQIEA VSNIFMAAIP
651 IKGIGAVRGK YLGGGITAHP IKRSQMGATA LPKGSAVSD NFADAAYAKY
701 PSFYHSRNIR SNLEQRYGKE NITSSVTPPS NGKNVKLADQ RHPKTVPFDD
751 GKGFNPFEXH VKYDT*

961cL-741

ATGAAACACTTTCCATCCAAAGTACTGACCACAGCCATCCTTGCCACTTTCGTAGCGGGCACTGGCAGCCACAAC
GACGACGATGTTAAAAAGCTGCCACTGTGGCCATTGCTGCTGCCTACAACAATGGCCAAGAAATCAACGGTTTCAAA
GCTGGAGAGACCATCTACGACATTGATGAAGACGGCACAATTACCAAAAAGACGCAACTGCAGCCGATGTTGAAGCC
GACGACTTTAAAGGCTGGGTCTGAAAAAAGTCGTGACTAACCTGACCAAAACCGTCAATGAAAAACAACAACCGTC
GATGCCAAAGTAAAGCTGCAGAATCTGAAATAGAAAAGTTAACAAACCAAGTTAGCAGACACTGATGCCGCTTTAGCA
GATACTGATGCCGCTCTGGATGCAACCACCAACGCCCTTGAATAAATTTGGGAGAAAATATAACGACATTTGCTGAAGAG
ACTAAGACAAATATCGTAAAAATTTGATGAAAAATTTAGAAGCCGTGGCTGATACCGTCGACAAGCATGCCGAAGCATTC
AACGATATCGCCGATTCATTGGATGAAACCAACACTAAGGCAGACGAAGCCGTCAAAACCGCCAAATGAAGCCAAACAG
ACGGCCGAAGAAACCAACAACCGTTCGATGCCAAAGTAAAGCTGCAGAAACTGCAGCAGGCAAGCCGAAGCTGCC
GCTGGCACAGCTAATACTGCAGCCGACAAGGCCGAAGCTGTGCTGCAAAAAGTTACCGACATCAAAGCTGATATCGCT
ACGAACAAGATAATATTGCTAAAAAGCAACAGTCCGACGCTGTACACCAGAGAAGAGTCTGACAGCAAATTTGTC
AGAATTGATGGTCTGAACGCTACTACCGAAAAATTTGACACACGCTTGGCTTCTGCTGAAAAATCCATTGCCGATCAC
GATACTCGCCTGAACGGTTTGGATAAAAACAGTGTGACACCTGGCCAAAGAAACCCGCCAAGGCCCTTGCAGAACAAAGCC
CGCTCTCCGGTCTGTTCCACCTTACAACGTTGGGTGGATTCGGAGGGGGTGGTGTGCGCCGCGACATCGGTGGGGG
CTTGGCGATGCACTAACCGCACCGCTCGACCATAAAGACAAAAGTTTGCAGTCTTTGACGCTGGATCAGTCCGTCAGG
AAAAACGAGAAACTGAAGCTGGCGGCACAAGGTGCCGAAAAAATTTATGGAAACGGTGACAGCCTCAATACGGGCAA
TTGAAGAACGACAAGGTCAGCCGTTTTCGACTTTATCCGCCAAATCGAAGTGGACGGGCAGCTCAATACCTTGGAGAGT
GGAGAGTTCCAAGTATACAAACAAGCCATTCCGCCTTAACCGCTTTCAGACCGAGCAAATACAAGATTCGGAGCAT
TCCGGGAAGATGGTTCGGAACGCCAGTTCAGAATCGCGACATAGCGGGGGAACATACATCTTTTGACAAGCTTCCC
GAAGGGCGCAGGGCGACATATCGCGGGACCGCTTCCGTTTTCAGACGATGCCGGCGGAAAACTGACCTACACCATAGAT
TTCCGCCCAAGCAGGGAAACCGCAAATCGAACATTTGAAATCGCCAGAACTCAATGTCGACCTGGCCGCGCCGAT
ATCAAGCCGGATGGAAAACGCCATGCCGTCATCAGCGGTTCCGTCCTTTACAACCAAGCCGAGAAAGGCAGTTACTCC
CTCGGTATCTTTGGCGGAAAAGCCAGGAAGTTGCCGGCAGCGGGAAGTGAACCCGTAAACGGCATACGCCATATCC
GGCTTGGCCCAAGCAACTCGAGCACCACCACCACCACCTGA

1 MKHFPSKVL TAILATFCSG ALAATNDDDV KKAATVAIAA AYNNGQEING
51 FKAGETIYDI DEDGTITKDD ATAADVEADD FKGLGLKKV VTNLTKTVNEN
101 KQNVDAKVA AESEIEKLTT KLADTDAALA DTDALDATT NALNKLGENI
151 TTFAEETKTN IVKIDEKLEA VADTVDKHAE AFNDIADSLD ETNTRKAEAV
201 KTANEAKQTA EETKQNVDAK VKAAETAAGK AEAAAGTANT AADKAEAVAA
251 KVTDIKADIA TNKDNIAKKA NSADVITREE SDSKFVRIDG LNATTEKLDI
301 RLASAEKSIA DHDTRLNGLD KTVSDLRKET RQGLAEQAAL SGLFQFVNVQ
351 GSGGGGVAAD IGAGLADALT APLDHKDRGL QSLTLDSQSVR KNEKLLKLAQ
401 GAERTYGNED SLNTGKLNND KVSRLFDFIRQ IEFVDQLITL ESGEFQVYKQ
451 SHSALTAFQT BQIQDSEHSG KMVAKRQFRI GDLAGHSTF DKLPEGGRAT
501 YRGTAFGSDD AGGKLYTID FAAKQNGKI EHLKSPELNV DLAAADIKPD
551 GKRHAVISGS VLYNQAEKGS YSLGIFGGKA QEVAGSAEVK TVNGIRHIGL
601 AAKQLEHHHH HH*

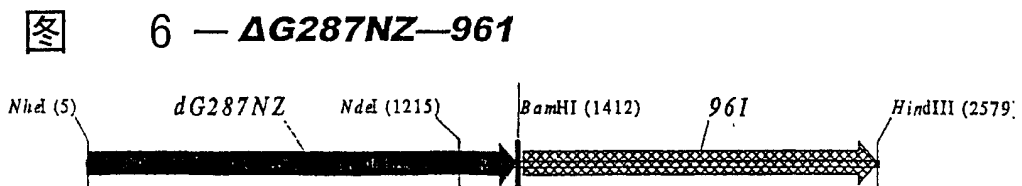
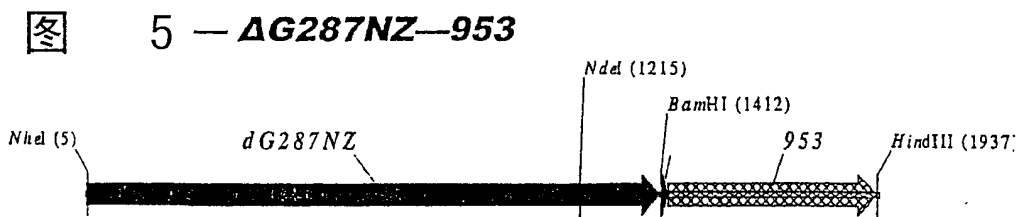
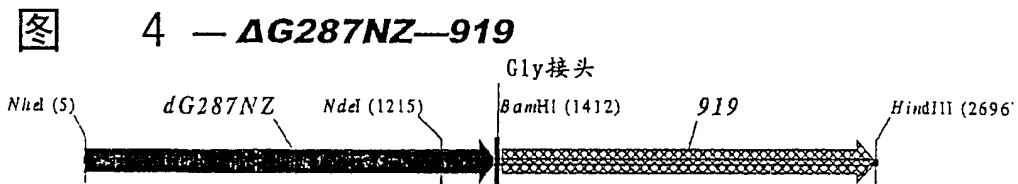
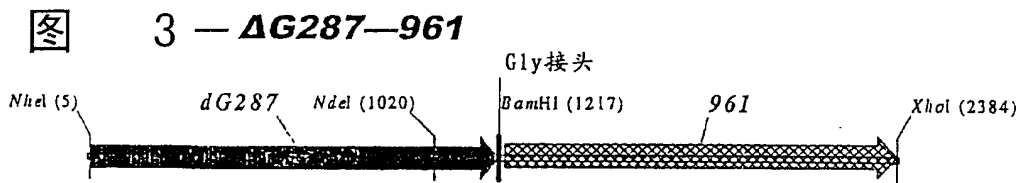
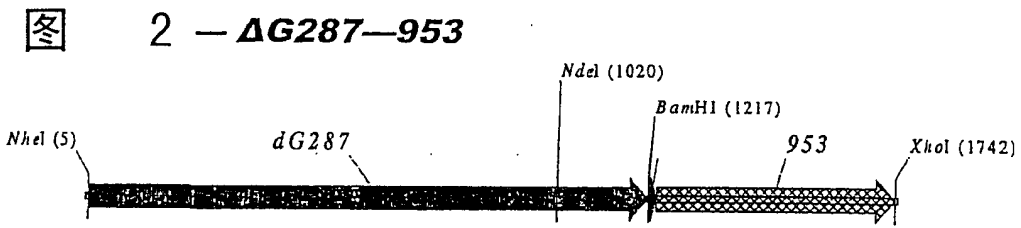
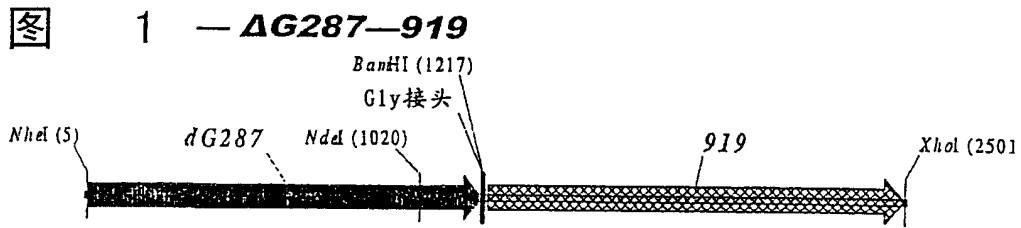
961cL-983

ATGAAACACTTTCCATCCAAAGTACTGACCACAGCCATCCTTGCCACTTTCGTAGCGGGCACTGGCAGCCACAAC
GACGACGATGTTAAAAAGCTGCCACTGTGGCCATTGCTGCTGCCTACAACAATGGCCAAGAAATCAACGGTTTCAAA
GCTGGAGAGACCATCTACGACATTGATGAAGACGGCACAATTACCAAAAAGACGCAACTGCAGCCGATGTTGAAGCC
GACGACTTTAAAGGCTGGGTCTGAAAAAAGTCGTGACTAACCTGACCAAAACCGTCAATGAAAAACAACAACCGTC
GATGCCAAAGTAAAGCTGCAGAATCTGAAATAGAAAAGTTAACAAACCAAGTTAGCAGACACTGATGCCGCTTTAGCA
GATACTGATGCCGCTCTGGATGCAACCACCAACGCCCTTGAATAAATTTGGGAGAAAATATAACGACATTTGCTGAAGAG
ACTAAGACAAATATCGTAAAAATTTGATGAAAAATTTAGAAGCCGTGGCTGATACCGTCGACAAGCATGCCGAAGCATTC
AACGATATCGCCGATTCATTGGATGAAACCAACACTAAGGCAGACGAAGCCGTCAAAACCGCCAAATGAAGCCAAACAG
ACGGCCGAAGAAACCAACAACCGTTCGATGCCAAAGTAAAGCTGCAGAAACTGCAGCAGGCAAGCCGAAGCTGCC
GCTGGCACAGCTAATACTGCAGCCGACAAGGCCGAAGCTGTGCTGCAAAAAGTTACCGACATCAAAGCTGATATCGCT
ACGAACAAGATAATATTGCTAAAAAGCAACAGTCCGACGCTGTACACCAGAGAAGAGTCTGACAGCAAATTTGTC
AGAATTGATGGTCTGAACGCTACTACCGAAAAATTTGACACACGCTTGGCTTCTGCTGAAAAATCCATTGCCGATCAC
GATACTCGCCTGAACGGTTTGGATAAAAACAGTGTGACACCTGGCCAAAGAAACCCGCCAAGGCCCTTGCAGAACAAAGCC
CGCTCTCCGGTCTGTTCCACCTTACAACGTTGGGTGGATTCGGCGGAGGGCGCACTTCTGCGCCGACTTCAATGCA
GGCGTACCGGTATCGGCAGCAACAGCAGAGCAACAACCGGAAATCAGCAGCAGTATCTTACGCGGTATCAAGAAC
GAAATGTGCAAGACAGAAGCATGCTCTGTGCCGGTCCGGATGACGTTGCCGTTACAGACAGGGATGCCAAAATCAAT
GCCCCCCCCCCGAATCTGCATACCGGAGACTTCCAAACCCAAATGACGCATACAAGAATTTGATCAACCTCAAACCT
GCAATGAAGCAGGCTATACAGGACCGGGGTAGAGGTAGGTATCGTGCACACAGGCAATCCGTCGGCAGCATATCC

TTTCCGGAACGTATGGCAGAAAAGAACACGGCTATAACGAAATACAAAACTATACGGCGTATATGCGGAAGGAA
 GCGCTGAAGACGGAGGCGGTAAAGACATTTGAAGCTTCTTTCCAGCATGAGGCCGTTATAGAGACTGAAGCAAAGCCG
 ACGGATATCCGCCACGTAAAAGAAATCGGACACATCGATTTGGTCTCCCATATTATTGGCGGGCGTTCCGTGGACGGC
 AGACCTGCAGGCGGTATTCGCCCCGATGCGACGCTACACATAATGAATACGAATGATGAAACCAAGAACGAATGATG
 GTTCAGCCATCCGCAATGCATGGGTCAAGCTGGGCGAACCTGGCGTGGCAGTTCGTCATAACAGTTTGGAAACAACA
 TCGAGGGCAGGCACTGCCGACCTTTTCCAAATAGCCAATTCGGAGGAGCAGTACCGCCAAGCGTTGCTCGACTATCC
 GCGGTGATAAAAACAGACGAGGGTATCCGCTGATGCAACAGAGCGATTACGGCAACCTGTCTACCACATCCGTAAT
 AAAAAATGCTTTTCATCTTTTCGACAGGCAATGACGACAGCTCAGCCCAACACATATGCCCTATTGCCATTTTAT
 GAAAAAGACGCTCAAAAAGGCATTTATCACAGTCGACGGCGTAGACCCGAGTGGAGAAAAGTTCAACCGGGAATGTAT
 GGAGAACCGGGTACAGAACCCTTGAGTATGGCTCCAACCATTCGCGAATTAAGTGCATGTGGTGCCTGTCCGCCACC
 TATGAAGCAAGCGTCCGTTTACCCGTACAAACCCGATTCAAATTCGCGGAACATCCTTTTCCGCCACCCATCGTAACC
 GGCACGGCGGCTCTGCTGCTGCAGAAAATACCCGTGGATGAGCAACGACAACTGCGTACCACGTTGCTGACGACGGCT
 CAGGACATCGGTGCAGTCCGCGTGGACAGCAAGTTCGGCTGGGGACTGCTGGATGCGGGTAAGGCCATGAACGGACCC
 GCGTCTTTCCGTTCCGGGACTTTACCGCCGATAGAAAGGTACATCCGATATTGCCCTACTCCTTCCGTAACGACATT
 TCAGGCACGGGCGGCTGATCAAAAAGGCGGCGAGCAACTGCAACTGCACGGCAACAACACCTATACGGGCAAAAACC
 ATTTATCGAAGGCGGTTCCGCTGGTGTGTGACGGCAACAACAATCGGATATGCGCGTGAACCAAGGTCGCTGAT
 TATAACGGGGCGGCATCCGCGCGGACGCTGAACAGCGGACCGCATTTGCTTATCTGGCAGATACCACCAATCCGCGCA
 AACGAAAACCGTACACATCAAAGGCAGTCTGCAGCTGGACGGCAAAGGTACGCTGTACACAGTTTGGGCAAACTGCTG
 AAAGTGGACGGTACGGCGATTATCGGCGCAAGCTGTACATGTCCGACGCGGCAAGGGGGCAGGCTATCTCAACAGT
 ACCGACGACGCTGTTCCCTTCCCTGAGTCCGCCAAAATCGGGCAGGATTTATCTTTCTTACAAAACATCGAAACCGAC
 GCGCGCTGCTGGCTTCCCTCGACAGCGTCAAAAACAGCGGGCAGTGAAGGGCAGACGCTGTCTATTATGTCCGT
 CGCGCAATGCGGCAAGGACTGCTTCGGCAGCGGCACATTCGCGCGCCCGCGGTCTGAAAACAGCGCCGTAGAACGGGC
 GGCAGCAATCTGAAAACCTGATGGTGAACCTGGATGCCTCCGAATCATCCGCAACACCCGAGACGGTTGAAAACCTGCG
 GCAGCCGACCGCACAGATATGCCGGCATCCGCCCCACGGCGCAACTTTCCGCGCAGCGGCAGCCGTACAGCATGCG
 AATGCCCGGACGGTGTACGCATCTTCAACAGTCTCGCCGCTACCGTCTATGCCGACAGTACCGCCGCCATGCCGAT
 ATGCAGGGACCGCCGCTGAAAGCCGATTCGGACGGGTTGGACCACAACGGCAGGGTCTGCGCGTATCGCGCAAAACC
 CAACAGGACGGTGGGAACAGGGCGGTGTTGAAGCAAAAATGCGCGGCAAGTACCCAAAACCGTCCGCAATGCC
 GCGAAAACCGGCAAAAATACGACAGCAGCCGCACACTGGGCATGGGACGACAGCATGGAGCGAAAACAGTGA
 GCAAAAACCGACAGCATTAGTCTGTTTGCAGGCATACGGCAGTATGCGGGCGATATCGGCTATCTCAAAGGCCTGTTT
 TCCTACGGACGCTACAAAACAGCATCAGCCGACGACCCGTTGCGGACGAAACATGCGGAAGGCAGCGTCAACGGCAGC
 CTGATGCAGCTGGGCGACTGGGCGGTGTCACGTTCCGTTTGGCCGCAACGGGAGATTTGACGGTCAAGGGCGGTCTG
 CGTACGCACTGCTCAAACAGGATGCATTCGCCGAAAAGGCAAGTGTCTTTGGGCTGGAGCGGCAACAGCTCCGTA
 GGCACGCTGGTCCGACTCGCGGCTGTAAGCTGTCCGCAACCCCTGAGCGATAAAGCCGCTCTGTTTCAACGGCGGGC
 GTGGAACGGACCTGAACGGACGGACTACCGGTAACGGGCGGCTTTACCGGCGGACTGCAGCAACCGGCAAGACG
 GGGGACGCAATATGCCGACACCCGCTGCGTTGCCGGCCTGGCGCGGATGTGCAATTCGGCAACGGCTGGAACGGC
 TTGGCACGTTACAGCTACGCCGTTCCAAAACAGTACGGCAACCACAGCGGACGAGTCCGGCTAGGCTACCGGTTCTGA
 CTCGAG

1	MKHFPSKVL	TAILATFC	SG	ALAATNDD	VDV	KKAATVA	IAAA	AYNNGQE	ING	
51	FKAGETIY	DI	DEDGTIT	KKD	ATAADVE	ADD	FKGLQLK	KVV	TNLTKTV	NEN
101	KQNVDAK	VKA	AESEIEK	LIT	KLADTDA	ALA	DTDAALD	AT	NALNKL	GENI
151	TTFAEETK	TN	IVKIDEK	LEA	VADTVDK	HAE	AFNDIAD	SLD	ETNTKAD	EAV
201	KTANEAKQ	TA	EETKQNV	DAK	VKAETAAG	K	AEAAAGT	ANT	AADKAE	AVAA
251	KVTDIKAD	IA	TNKDNI	LAKA	NSADVYT	REE	SDSKFVR	IDG	LNATTE	KLDT
301	RLASAEKS	IA	DHDTRLN	GLD	KTVSDLR	KET	RQGLAEQ	AAL	SGLFPQ	PYVNG
351	GSGGGGTS	AP	DFNAGGT	GIG	SNSRATT	AKS	AAVSYAG	IKN	EMCKDR	SMLC
401	AGRDDVAV	TD	RDakin	APP	NLHTGDF	PNP	NDAYKNL	INL	KPAIEAG	YTG
451	RGYVGI	VDT	GESVGS	ISFP	ELYGRKE	HGY	NENYKNY	TAY	MRKEAP	EDGG
501	GKDIEASF	DD	EAVIET	EAKP	TDIRHV	KEIG	HIDLIVS	HIIG	GRSVDGR	EPAG
551	GIAPDATH	I	MNTNDET	KNE	MMVAAIL	RAW	VKLGERS	VRI	VNNSFG	TFSR
601	AGTADLFQ	IA	NSEEQYR	QAL	LDYSGGD	KTD	EGIRLMQ	QSD	YGNSLY	HIRN
651	KNMLFIF	STG	NDAQAQ	PNTY	ALLPFY	EKDA	QKGIITV	AGV	DRSGEK	FKRE
701	MYGEPGTE	PL	EYGSNH	CGIT	AMWCLS	APYE	ASVRFTR	TNP	IQIAGT	SFSA
751	PIVTGTA	AALL	LQKYPW	MSND	NLRTTLL	TIA	QDIGAVG	VDS	KFGWGL	LDDAG
801	KAMNGPAS	FP	FGDPTAD	TKG	TSDIAYS	FRN	DISGTG	GGLIK	KGGSQ	LQLHG
851	NNTYTGT	II	EGGSLV	LYGN	NKSDMR	VETK	GALIYNG	AAS	GGSLNS	DGIV
901	YLADTDQ	SGA	NEFVHI	KGSL	QLDGKGT	LYT	RLGKLL	KVDG	TAIIGK	KLYM
951	SARGKGAG	YL	NSTGRR	VFFL	SAKIGQ	DYS	FFTNIET	DGG	LLASLD	SVKE
1001	TAGSEGT	LS	YYVRRG	NAAR	TASAAA	HSAP	AGLKHA	VEQG	GSNLEN	LMVE
1051	LDASESS	ATP	ETVETA	AADR	TDMPGIR	PHYG	ATFRAAA	AVQ	HANAAD	GVRI
1101	FNSLAAT	VYA	DSTAHA	ADMQ	GRRLKAV	SDG	LDHNGT	GLRV	IAQTQD	CGDT
1151	WEQGGV	EGRM	RGSTQT	VGIA	ARTGENT	TAA	ATLGMGR	STW	SENSAN	AKTD
1201	SISLFA	GIRH	DAGDIG	YLKG	LFSYGRY	KNS	ISRSTGA	DEH	AEGSVN	GILM
1251	QIGALG	VNV	FFAATG	DLTV	EGGLRY	DLK	QDAFAE	KGSA	LWWSGN	SLTE
1301	GTLVGL	LAGLK	LSQPLS	DKAV	LFATAG	VERD	LNGRDY	TVTGT	GFTGATA	ATG
1351	KTGARN	MPH	RLVAGL	GADV	EPGN	GWGLA	RYSYAG	SKQY	GNHSGR	VGVG
1401	YRF*									

可以理解本发明仅以实施例的方式进行描述，在本发明的范围和精神内还可进行改变。例如，设想可以使用其它菌株的蛋白质[如，参见 WO 00/66741，ORF4、ORF40、ORF46、225、235、287、519、726、919 和 953 的多态序列]。



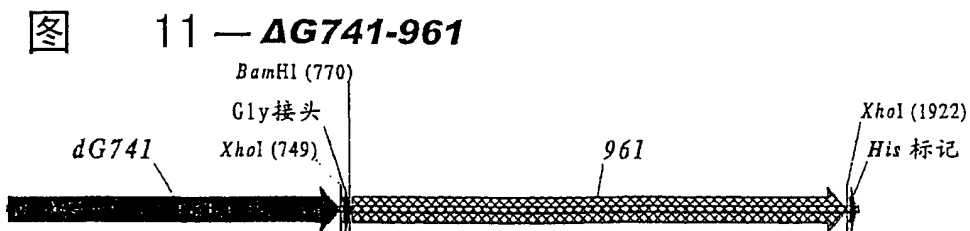
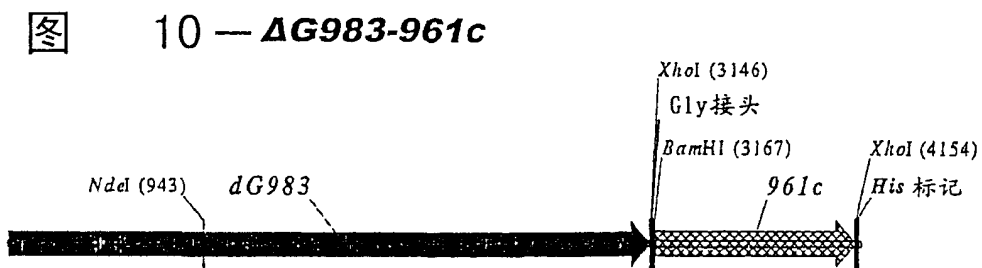
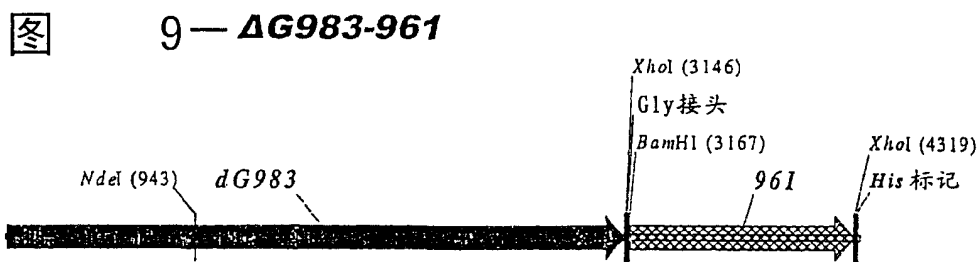
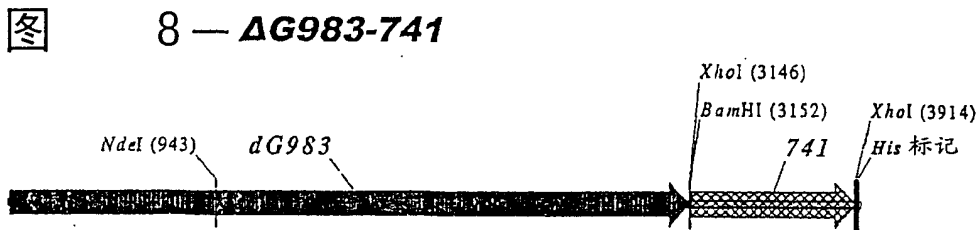
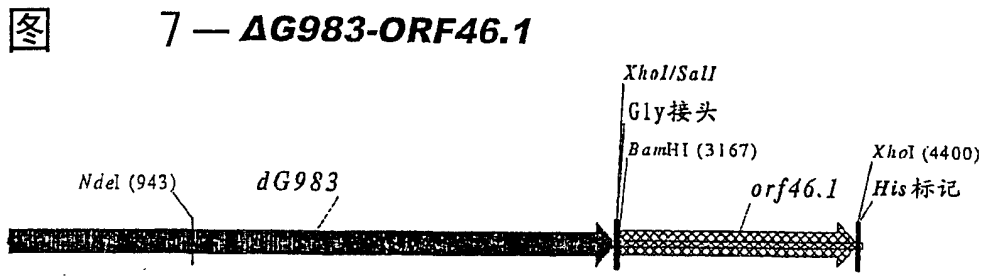


图 12 — $\Delta G741-961c$

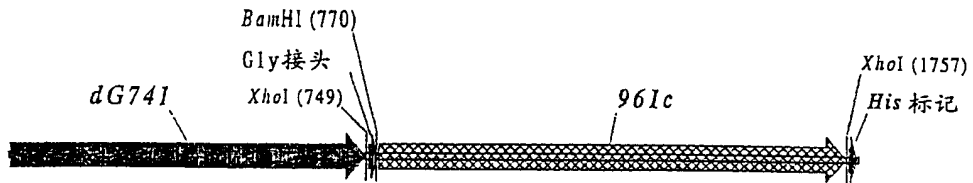


图 13 — $\Delta G741-983$

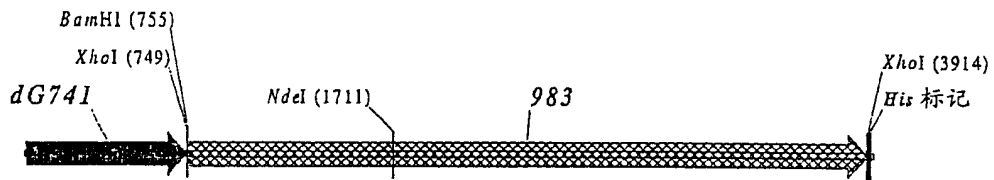


图 14 — $\Delta G741-ORF46.1$



图 15 — $ORF46.1-741$

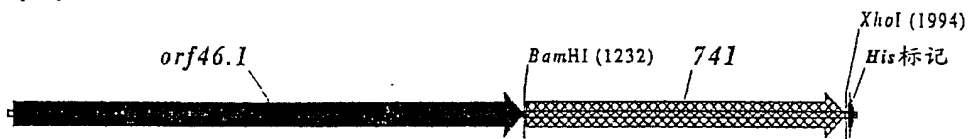


图 16 — $ORF46.1-961$



图 17 — **ORF46.1-961c**

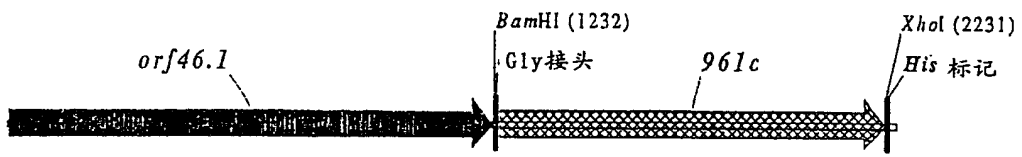


图 18 — **961-ORF46.1**



图 19 — **961-741**

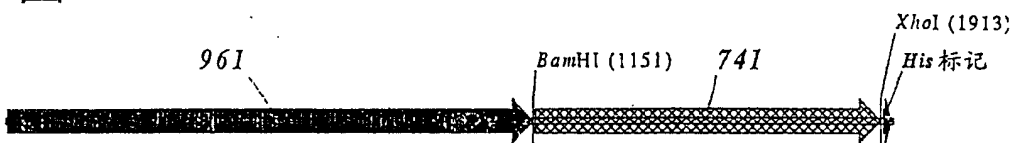


图 20 — **961-983**

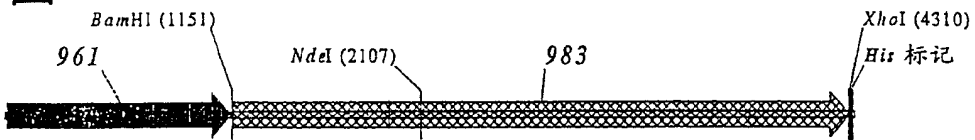


图 21 — **961c-ORF46.1**



图 22 — **961c-741**

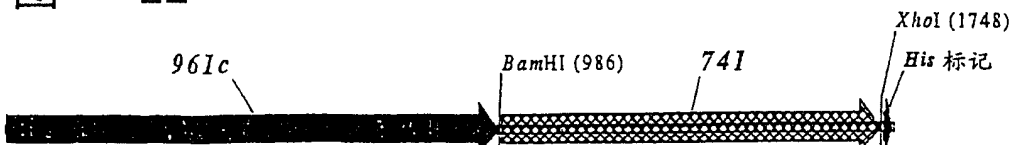


图 23 — 961c-983

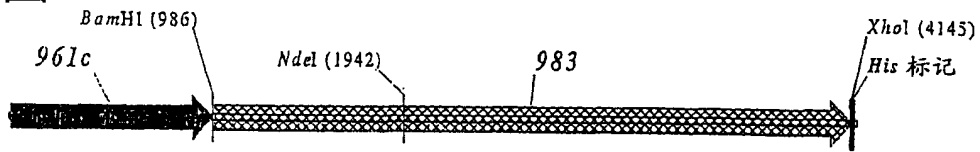


图 24 — 961cL-ORF46.1

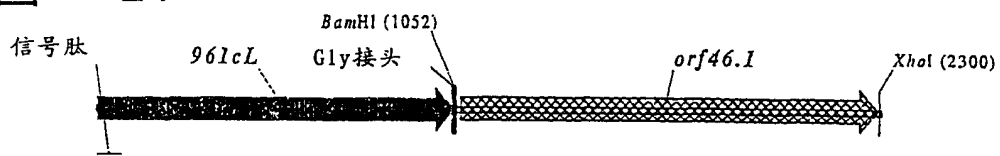


图 25 — 961cL-741

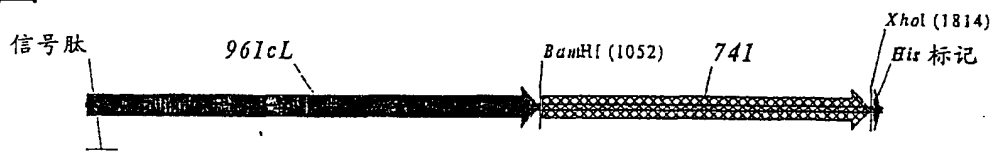


图 26 — 961cL-983

