西瓜噬酸菌*ftsH*基因功能预测

Functional Prediction of the Gene *ftsH* in *Acidovorax Citrulli*

汇报人：　季苇芹

汇报组：　G06

组员：田雪力、季苇芹、曹步暐、李欢

2018.11.11

《实用生物信息技术》期末总结交流报告会

目　录

研究背景

1. 瓜类细菌性果斑病 (Bacterial Fruit Blotch, BFB)

- 检疫性种传病害，主要为害西瓜、甜瓜等葫芦科作物。
- 在我国十多个省份地区均有分布。
- **危害广、损失重、防治难**

图1.1 瓜类细菌性果斑病在西瓜及甜瓜上的为害症状

赵廷昌等.植保技术与推广,2001

研究背景

病原菌：西瓜噬酸菌（*Acidovorax citrulli*）

革兰氏阴性菌，不产荧光，严格好氧，菌体短杆状，单根极生鞭毛，能在41℃下生长，但不能在4℃生长。

图1.2 菌体在KB平板上菌落形态。
直径1-2mm，圆形，全缘光滑，乳白色

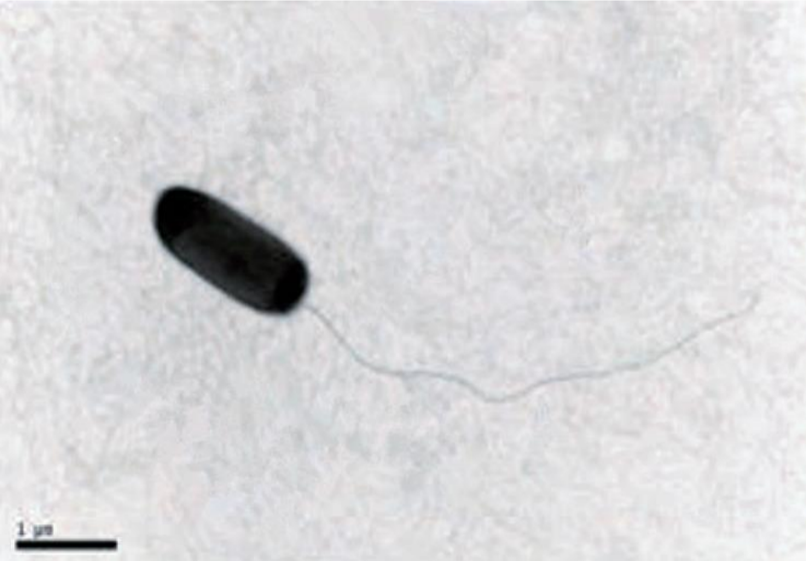


图1.3 透射电子显微镜观察西瓜噬酸菌鞭毛形态

赵廷昌等.植保技术与推广,2001

研究背景

2. AAA ATPase

AAA ATPase是功能多样的大型蛋白家族，有一个约230个氨基酸残基的高度保守结构域，**存在于所有生物**中，在细胞中的**功能必不可少**。参与如DNA复制、蛋白降解、膜融合、微管、过氧化物酶体生物合成、信号转导和基因表达调控等过程。

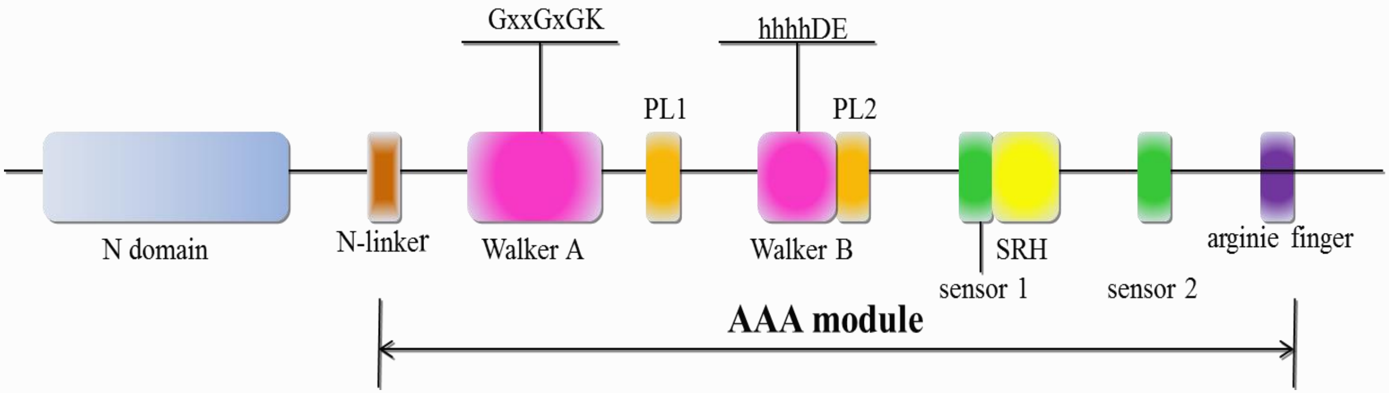


图1.4 AAA结构域结构

AAA module: 230个左右高度保守的氨基酸残基

Walker A和Walker B: ATP结合；

SRH: ATP水解

N-linker: 转移ATP水解产生能量；

PL1和PL2: 底物识别

Jan P. Erzberger and James M. Rev. Biophys. Biomol. Struct.2006.

研究背景

果斑菌中的AAA ATPase

研究表明，AAA ATPase基因的缺失会降低病原菌致病能力。在大肠杆菌的5种AAA蛋白酶（ClpXP、ClpAP、HslUV、Lon和FtsH）中只有FtsH是必需的。

表1 果斑菌中**AAA ATPases 研究进展**

	基因功能	**功能**
Lon	***Aave_1457*** ***Aave_3667***	**热激蛋白，参与蛋白质量控制，可降解富含特殊标签（如非极性N末端）的蛋白（NCBI注释）**
RuvB	***Aave_4237***	**霍利迪结DNA解旋酶（李晶 等.,2016）**
FtsH	***Aave_2617***	**依赖Zn^{2+}和ATP的特殊蛋白，原核生物中是生长必需的能量依赖型蛋白酶（NCBI注释）**
ClpB	***Aave_1482***	**VI型分泌系统，不影响种传苗能力、种子表面定殖能力以及生物膜形成能力(Tian *et al*., 2015)**
HrcN	***Aave_0463***	**III型分泌系统，致病力明显减弱，生长能力明显下降，运动能力下降，生物膜形成能力增强(严婉荣 等., 2015)**
HslVU	***Aave_0811***	**热激蛋白，参与蛋白质量控制及应激反应（NCBI注释）**
MoxR	***Aave_2163*** ***Aave_1063***	**参与调节酸酐的运输（李晶 等.,2016）**

研究背景

3. FtsH蛋白

FtsH蛋白（filamentation temperature-sensitive H），是属于AAA蛋白酶家族的**ATP 和Zn^{2+}依赖型金属蛋白酶**，在原核和真核生物中广泛存在（表2）。

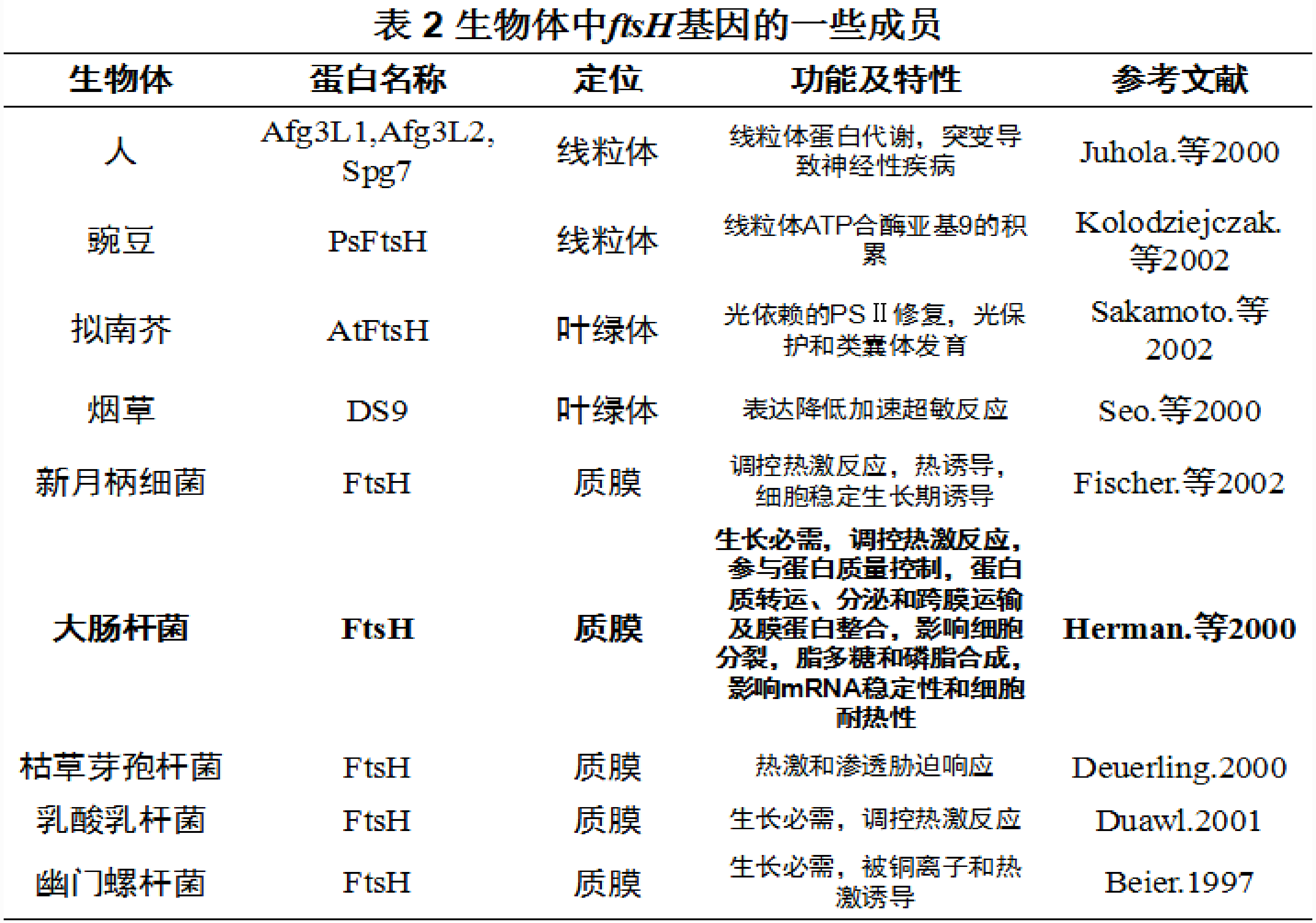


表 2 生物体中*ftsH*基因的一些成员

生物体	蛋白名称	定位	功能及特性	参考文献
人	Afg3L1,Afg3L2, Spg7	线粒体	线粒体蛋白代谢，突变导致神经性疾病	Juhola.等2000
豌豆	PsFtsH	线粒体	线粒体ATP合酶亚基9的积累	Kolodziejczak.等2002
拟南芥	AtFtsH	叶绿体	光依赖的PS II 修复，光保护和类囊体发育	Sakamoto.等2002
烟草	DS9	叶绿体	表达降低加速超敏反应	Seo.等2000
新月柄细菌	FtsH	质膜	调控热激反应，热诱导，细胞稳定生长期诱导	Fischer.等2002
大肠杆菌	**FtsH**	**质膜**	**生长必需，调控热激反应，参与蛋白质量控制，蛋白质转运、分泌和跨膜运输及膜蛋白整合，影响细胞分裂，脂多糖和磷脂合成，影响mRNA稳定性和细胞耐热性**	**Herman.等2000**
枯草芽孢杆菌	FtsH	质膜	热激和渗透胁迫响应	Deuerling.2000
乳酸乳杆菌	FtsH	质膜	生长必需，调控热激反应	Duawl.2001
幽门螺杆菌	FtsH	质膜	生长必需，被铜离子和热激诱导	Beier.1997

研究背景

在大肠杆菌中，由*ftsH*编码的FtsH蛋白的结构包括**N端跨膜域**、**AAA结构域**、**锌离子结合模块**等，且其N末端与C末端同在膜内侧，有两个跨膜域(TM) 。 FtsH蛋白具有**ATP酶活性**、**蛋白水解活性**和**分子伴侣活性**，参与**蛋白质量平衡控制**、**逆境胁迫应激反应**、 **蛋白的装配**、**转运和折叠**等过程。

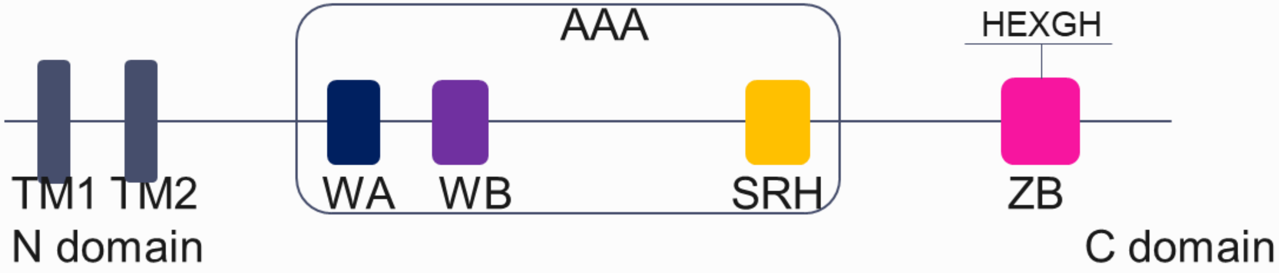


图1.5 FtsH蛋白结构

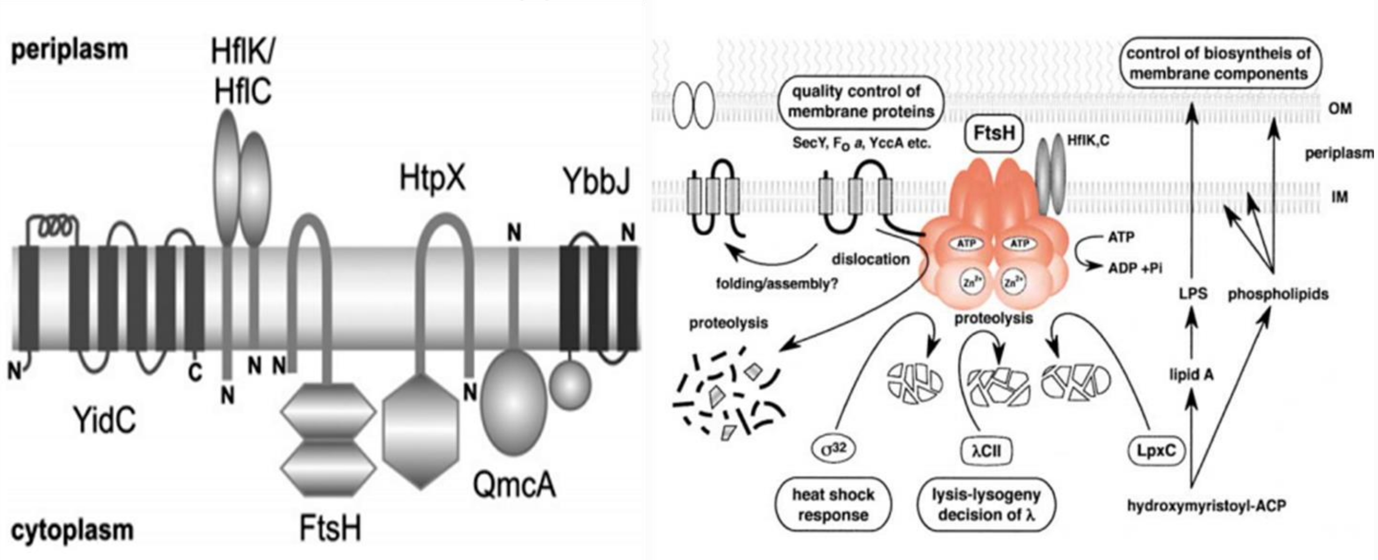


图1.6 FtsH蛋白在大肠杆菌中作用图

Dutta B, Gitaitis R, Smith S, *et al*. PLoS One, 2014, 9(6): e99215；Jiang C H, Wu F, Yu Z Y, *et al*. Microbiological Research, 2015, 170: 95-104.

研究背景

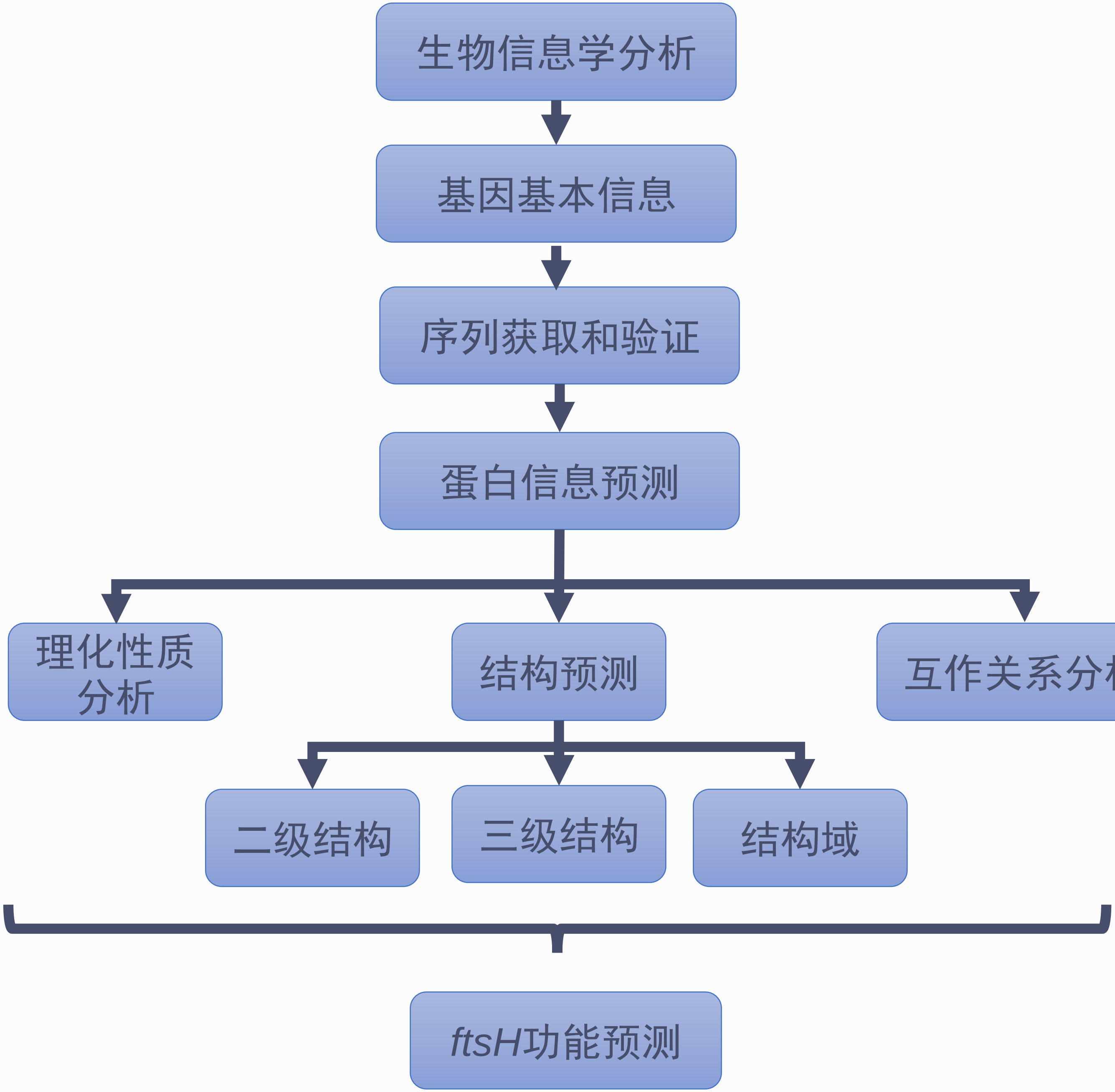
目前，由于瓜类细菌性果斑病的具体致病机理尚未明确，防治主要依靠传统的化学防治手段，防治效果不理想，缺乏针对性的防治措施。因此亟需对西瓜噬酸菌的致病机理进行研究，从而为果斑病的针对性防治提供理论指导。

由于FtsH蛋白在大肠杆菌等细菌的各项生命活动中发挥重要作用，且其在西瓜噬酸菌中的作用尚未明确，因而我们选取FtsH研究对象，希望通过研究它的功能来进一步明确果斑病菌的致病机理，以期为生产上的防治提供分子基础。

在研究前，**利用生物信息学手段对FtsH蛋白的功能进行预测，有利于后续实验更有目的性的展开**。

赵廷昌,孙福在,王兵万. 植保技术与推广,2001,3:37-38-36; 王叶筠. 中国西瓜甜瓜,2003,5:32-34；许勇,张兴平,宫国义,张海英. 中国西瓜甜瓜,2003, 6:36-37；黄月英. 福建农业科技,2008,2:62-64；商明清,张德满,杨勤民. 植物检疫,2014,28(3):93-96.

技术路线



生物信息学分析

基因基本信息

序列获取和验证

蛋白信息预测

理化性质分析

结构预测

互作关系分析

二级结构

三级结构

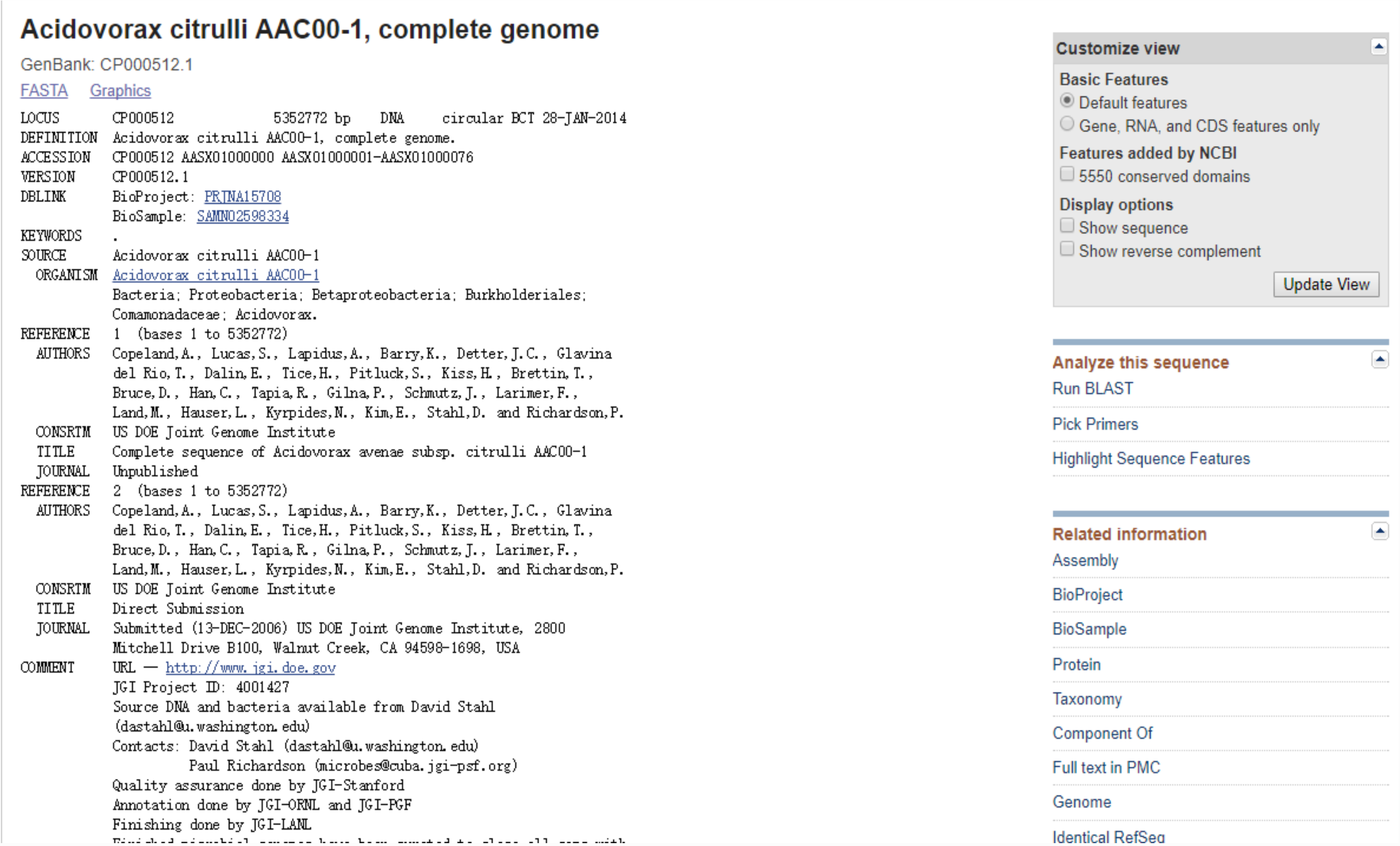
结构域

*ftsH*功能预测

基因基本信息及序列获取

1. *ftsH*基本信息

登录NCBI，查找西瓜噬酸菌AAC00-1菌株全基因组数据库（https://www.ncbi.nlm.nih.gov/nuccore/120587178），在库中搜索FtsH。



Acidovorax citrulli AAC00-1, complete genome

GenBank: CP000512.1

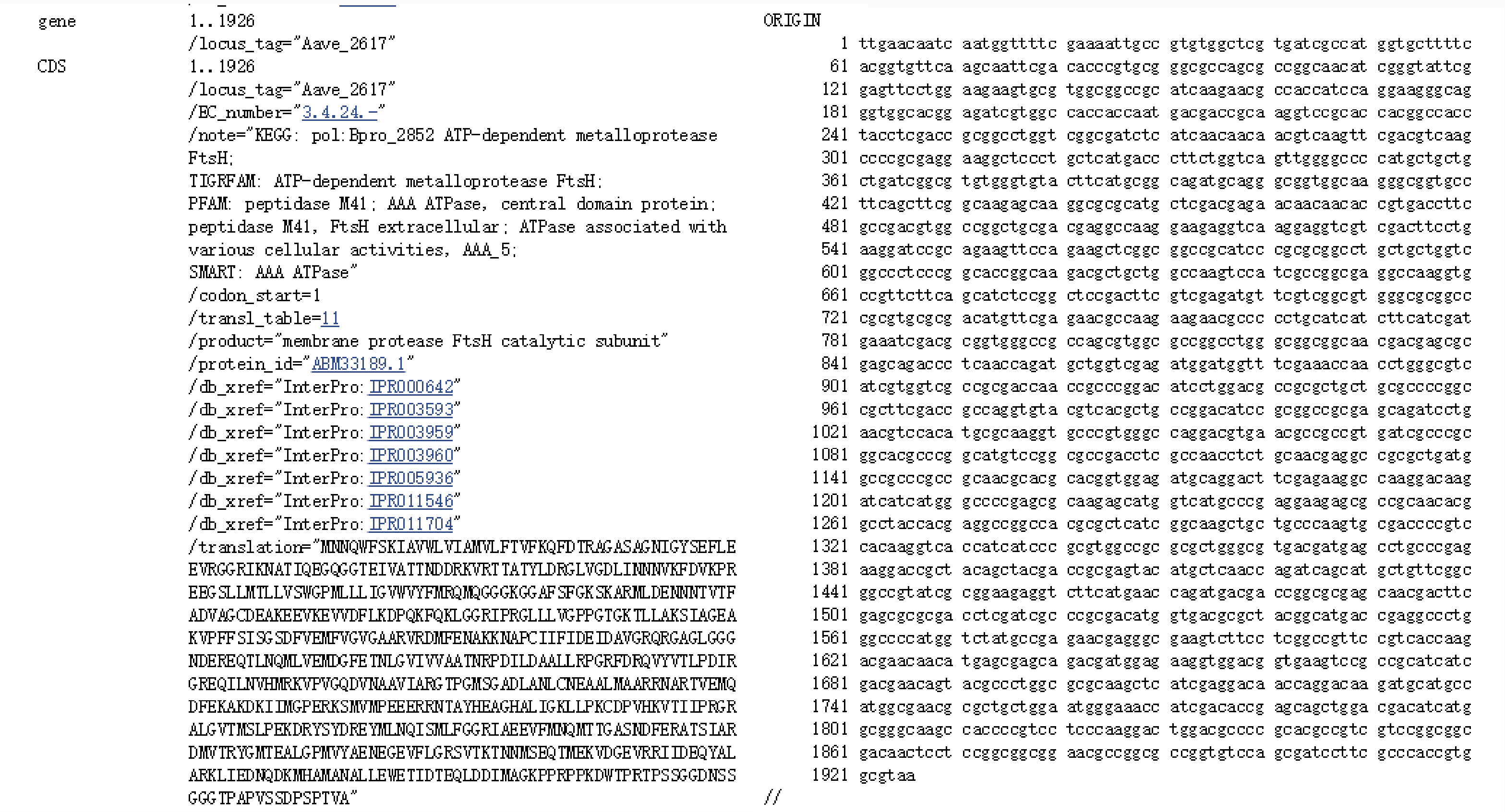
```
LOCUS       CP000512             5352772 bp    DNA     circular BCT 28-JAN-2014
DEFINITION  Acidovorax citrulli AAC00-1, complete genome.
ACCESSION   CP000512 AASX01000000 AASX01000001-AASX01000076
VERSION     CP000512.1
DBLINK      BioProject: PRJNA15708
            BioSample: SAMN02598334
KEYWORDS    .
SOURCE      Acidovorax citrulli AAC00-1
  ORGANISM  Acidovorax citrulli AAC00-1
            Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
            Comamonadaceae; Acidovorax.
REFERENCE   1  (bases 1 to 5352772)
  AUTHORS   Copeland,A., Lucas,S., Lapidus,A., Barry,K., Detter,J.C., Glavina
            del Rio,T., Dalin,E., Tice,H., Pitluck,S., Kiss,H., Brettin,T.,
            Bruce,D., Han,C., Tapia,R., Gilna,P., Schmutz,J., Larimer,F.,
            Land,M., Hauser,L., Kyrpides,N., Kim,E., Stahl,D. and Richardson,P.
  CONSRTM   US DOE Joint Genome Institute
  TITLE     Complete sequence of Acidovorax avenae subsp. citrulli AAC00-1
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 5352772)
  AUTHORS   Copeland,A., Lucas,S., Lapidus,A., Barry,K., Detter,J.C., Glavina
            del Rio,T., Dalin,E., Tice,H., Pitluck,S., Kiss,H., Brettin,T.,
            Bruce,D., Han,C., Tapia,R., Gilna,P., Schmutz,J., Larimer,F.,
            Land,M., Hauser,L., Kyrpides,N., Kim,E., Stahl,D. and Richardson,P.
  CONSRTM   US DOE Joint Genome Institute
  TITLE     Direct Submission
  JOURNAL   Submitted (13-DEC-2006) US DOE Joint Genome Institute, 2800
            Mitchell Drive B100, Walnut Creek, CA 94598-1698, USA
COMMENT     URL -- http://www.jgi.doe.gov
            JGI Project ID: 4001427
            Source DNA and bacteria available from David Stahl
            (dastahl@u.washington.edu)
            Contacts: David Stahl (dastahl@u.washington.edu)
                      Paul Richardson (microbes@cuba.jgi-psf.org)
            Quality assurance done by JGI-Stanford
            Annotation done by JGI-ORNL and JGI-PGF
            Finishing done by JGI-LANL
```

图3.1 西瓜噬酸菌AAC00-1菌株全基因组数据库

基因基本信息
及序列获取

1. *ftsH*基本信息

在库中搜索FtsH，根据注释信息，得到*ftsH*的基因号为*Aave_2617*，碱基序列全长1926bp，氨基酸序列全长641aa，编码FtsH蛋白，属于AAA家族，与多种生命活动相关，此外还有产物蛋白在KEGG、SMART等网站的注释。



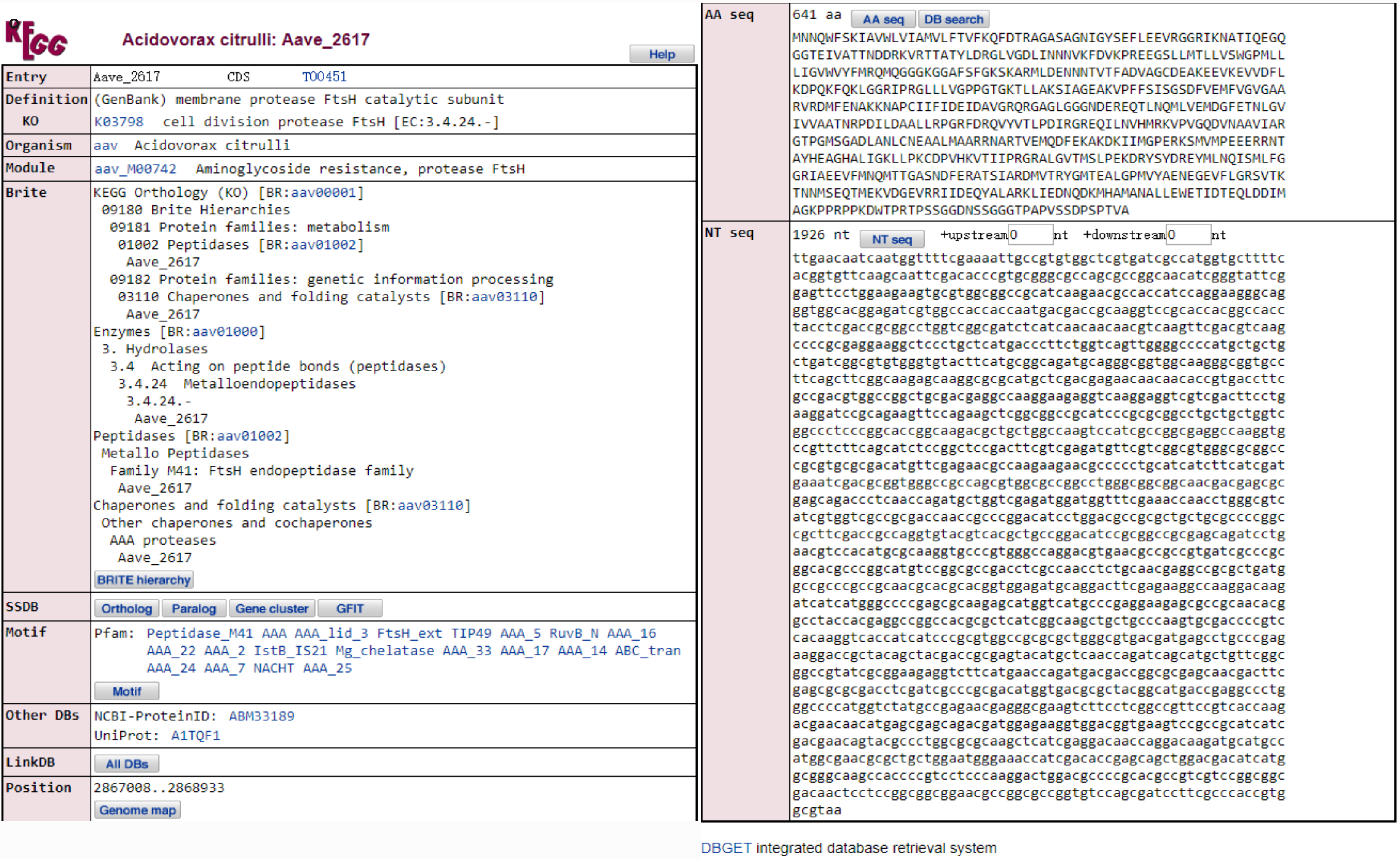
```
     gene            1..1926
                     /locus_tag="Aave_2617"
     CDS             1..1926
                     /locus_tag="Aave_2617"
                     /EC_number="3.4.24.-"
                     /note="KEGG: pol:Bpro_2852 ATP-dependent metalloprotease
                     FtsH;
                     TIGRFAM: ATP-dependent metalloprotease FtsH;
                     PFAM: peptidase M41; AAA ATPase, central domain protein;
                     peptidase M41, FtsH extracellular; ATPase associated with
                     various cellular activities, AAA_5;
                     SMART: AAA ATPase"
                     /codon_start=1
                     /transl_table=11
                     /product="membrane protease FtsH catalytic subunit"
                     /protein_id="ABM33189.1"
                     /db_xref="InterPro:IPR000642"
                     /db_xref="InterPro:IPR003593"
                     /db_xref="InterPro:IPR003959"
                     /db_xref="InterPro:IPR003960"
                     /db_xref="InterPro:IPR005936"
                     /db_xref="InterPro:IPR011546"
                     /db_xref="InterPro:IPR011704"
                     /translation="MNNQWFSKIAVWLVIAMVLFTVFKQFDTRAGASAGNIGYSEFLE
                     EVRGGRIKNATIQEGQGGTEIVATTNDDRKVRTTATYLDRGLVGDLINNNVKFDVKPR
                     EEGSLLMTLLVSWGPMLLLIGVWVYFMRQMQGGGKGGAFSFGKSKARMLDENNNTVTF
                     ADVAGCDEAKEEVKEVVDFLKDPQKFQKLGGRIPRGLLLVGPPGTGKTLLAKSIAGEA
                     KVPFFSISGSDFVEMFVGVGAARVRDMFENAKKNAPCIIFIDEIDAVGRQRGAGLGGG
                     NDEREQTLNQMLVEMDGFETNLGVIVVAATNRPDILDAALLRPGRFDRQVYVTLPDIR
                     GREQILNVHMRKVPVGQDVNAAVIARGTPGMSGADLANLCNEAALMAARRNARTVEMQ
                     DFEKAKDKIIMGPERKSMVMPEEERRNTAYHEAGHALIGKLLPKCDPVHKVTIIPRGR
                     ALGVTMSLPEKDRYSYDREYMLNQISMLFGGRIAEEVFMNQMTTGASNDFERATSIAR
                     DMVTRYGMTEALGPMVYAENEGEVFLGRSVTKTNNMSEQTMEKVDGEVRRIIDEQYAL
                     ARKLIEDNQDKMHAMANALLEWETIDTEQLDDIMAGKPPRPPKDWTPRTPSSGGDNSS
                     GGGTPAPVSSDPSPTVA"
```

```
ORIGIN
        1 ttgaacaatc aatggttttc gaaaattgcc gtgtggctcg tgatcgccat ggtgcttttc
       61 acggtgttca agcaattcga cacccgtgcg ggcgccagcg ccggcaacat cgggtattcg
      121 gagttcctgg aagaagtgcg tggcggccgc atcaagaacg ccaccatcca ggaagggcag
      181 ggtggcacgg agatcgtggc caccaccaat gacgaccgca aggtccgcac cacggccacc
      241 tacctcgacc gcggcctggt cggcgatctc atcaacaaca acgtcaagtt cgacgtcaag
      301 ccccgcgagg aaggctccct gctcatgacc cttctggtca gttggggccc catgctgctg
      361 ctgatcggcg tgtgggtgta cttcatgcgg cagatgcagg gcggtggcaa gggcggtgcc
      421 ttcagcttcg gcaagagcaa ggcgcgcatg ctcgacgaga acaacaacac cgtgaccttc
      481 gccgacgtgg ccggctgcga cgaggccaag gaagaggtca aggaggtcgt cgacttcctg
      541 aaggatccgc agaagttcca gaagctcggc ggccgcatcc cgcgcggcct gctgctggtc
      601 ggccctcccg gcaccggcaa gacgctgctg gccaagtcca tcgccggcga ggccaaggtg
      661 ccgttcttca gcatctccgg ctccgacttc gtcgagatgt tcgtcggcgt gggcgcggcc
      721 cgcgtgcgcg acatgttcga gaacgccaag aagaacgccc cctgcatcat cttcatcgat
      781 gaaatcgacg cggtgggccg ccagcgtggc gccggcctgg gcggcggcaa cgacgagcgc
      841 gagcagaccc tcaaccagat gctggtcgag atggatggtt tcgaaaccaa cctgggcgtc
      901 atcgtggtcg ccgcgaccaa ccgcccggac atcctggacg ccgcgctgct gcgccccggc
      961 cgcttcgacc gccaggtgta cgtcacgctg ccggacatcc gcggccgcga gcagatcctg
     1021 aacgtccaca tgcgcaaggt gcccgtgggc caggacgtga acgccgccgt gatcgcccgc
     1081 ggcacgcccg gcatgtccgg cgccgacctc gccaacctct gcaacgaggc cgcgctgatg
     1141 gccgcccgcc gcaacgcacg cacggtggag atgcaggact tcgagaaggc caaggacaag
     1201 atcatcatgg gccccgagcg caagagcatg gtcatgcccg aggaagagcg ccgcaacacg
     1261 gcctaccacg aggccggcca cgcgctcatc ggcaagctgc tgcccaagtg cgaccccgtc
     1321 cacaaggtca ccatcatccc gcgtggccgc gcgctgggcg tgacgatgag cctgcccgag
     1381 aaggaccgct acagctacga ccgcgagtac atgctcaacc agatcagcat gctgttcggc
     1441 ggccgtatcg cggaagaggt cttcatgaac cagatgacga ccggcgcgag caacgacttc
     1501 gagcgcgcga cctcgatcgc ccgcgacatg gtgacgcgct acggcatgac cgaggccctg
     1561 ggccccatgg tctatgccga gaacgagggc gaagtcttcc tcggccgttc cgtcaccaag
     1621 acgaacaaca tgagcgagca gacgatggag aaggtggacg gtgaagtccg ccgcatcatc
     1681 gacgaacagt acgccctggc gcgcaagctc atcgaggaca accaggacaa gatgcatgcc
     1741 atggcgaacg cgctgctgga atgggaaacc atcgacaccg agcagctgga cgacatcatg
     1801 gcgggcaagc caccccgtcc tcccaaggac tggacgcccc gcacgccgtc gtccggcggc
     1861 gacaactcct ccggcggcgg aacgccggcg ccggtgtcca gcgatccttc gcccaccgtg
     1921 gcgtaa
//
```

图3.2 西瓜噬酸菌AAC00-1菌株全基因组数据库中*ftsH*基本信息

基因基本信息及序列获取

根据基因号，登录KEGG（https://www.kegg.jp/），可以得到*ftsH*的碱基序列和氨基酸序列，同时还可以得到*ftsH*的同源基因、是否处于某条信号通路中、基因簇、在基因图谱中的位置及翻译方向、产物蛋白所属家族及相关注释、在NCBI及UniProt中的登录号等信息。



Acidovorax citrulli: Aave_2617

Entry	Aave_2617 CDS T00451
Definition	(GenBank) membrane protease FtsH catalytic subunit
KO	K03798 cell division protease FtsH [EC:3.4.24.-]
Organism	aav Acidovorax citrulli
Module	aav_M00742 Aminoglycoside resistance, protease FtsH
Brite	KEGG Orthology (KO) [BR:aav00001] 09180 Brite Hierarchies 09181 Protein families: metabolism 01002 Peptidases [BR:aav01002] Aave_2617 09182 Protein families: genetic information processing 03110 Chaperones and folding catalysts [BR:aav03110] Aave_2617 Enzymes [BR:aav01000] 3. Hydrolases 3.4 Acting on peptide bonds (peptidases) 3.4.24 Metalloendopeptidases 3.4.24.- Aave_2617 Peptidases [BR:aav01002] Metallo Peptidases Family M41: FtsH endopeptidase family Aave_2617 Chaperones and folding catalysts [BR:aav03110] Other chaperones and cochaperones AAA proteases Aave_2617
Motif	Pfam: Peptidase_M41 AAA AAA_lid_3 FtsH_ext TIP49 AAA_5 RuvB_N AAA_16 AAA_22 AAA_2 IstB_IS21 Mg_chelatase AAA_33 AAA_17 AAA_14 ABC_tran AAA_24 AAA_7 NACHT AAA_25
Other DBs	NCBI-ProteinID: ABM33189 UniProt: A1TQF1
Position	2867008..2868933

AA seq 641 aa

```
MNNQWFSKIAVWLVIAMVLFTVFKQFDTRAGASAGNIGYSEFLEEVRGGRIKNATIQEGQ
GGTEIVATTNDDRKVRTTATYLDRGLVGDLINNNVKFDVKPREEGSLLMTLLVSWGPMLL
LIGVWVYFMRQMQGGGKGGAFSFGKSKARMLDENNNTVTFADVAGCDEAKEEVKEVVDFL
KDPQKFQKLGGRIPRGLLLVGPPGTGKTLLAKSIAGEAKVPFFSISGSDFVEMFVGVGAA
RVRDMFENAKKNAPCIIFIDEIDAVGRQRGAGLGGGNDEREQTLNQMLVEMDGFETNLGV
IVVAATNRPDILDAALLRPGRFDRQVYVTLPDIRGREQILNVHMRKVPVGQDVNAAVIAR
GTPGMSGADLANLCNEAALMAARRNARTVEMQDFEKAKDKIIMGPERKSMVMPEEERRNT
AYHEAGHALIGKLLPKCDPVHKVTIIPRGRALGVTMSLPEKDRYSYDREYMLNQISMLFG
GRIAEEVFMNQMTTGASNDFERATSIARDMVTRYGMTEALGPMVYAENEGEVFLGRSVTK
TNNMSEQTMEKVDGEVRRIIDEQYALARKLIEDNQDKMHAMANALLEWETIDTEQLDDIM
AGKPPRPPKDWTPRTPSSGGDNSSGGGTPAPVSSDPSPTVA
```

NT seq 1926 nt +upstream 0 nt +downstream 0 nt

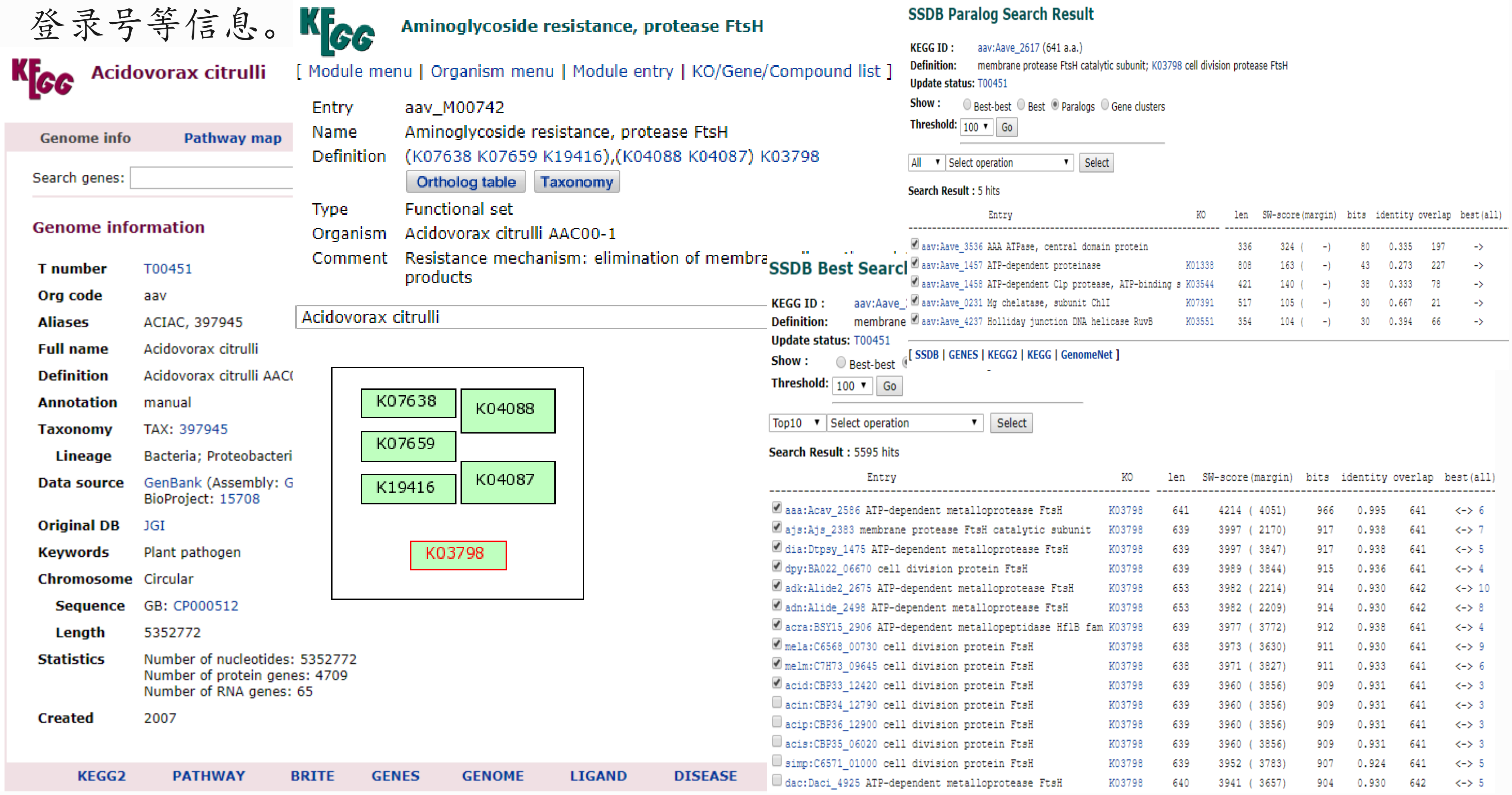
```
ttgaacaatcaatggttttcgaaaattgccgtgtggctcgtgatcgccatggtgcttttc
acggtgttcaagcaattcgacacccgtgcgggcgccagcgccggcaacatcgggtattcg
gagttcctggaagaagtgcgtggcggccgcatcaagaacgccaccatccaggaagggcag
ggtggcacggagatcgtggccaccaccaatgacgaccgcaaggtccgcaccacggccacc
tacctcgaccgcggcctggtcggcgatctcatcaacaacaacgtcaagttcgacgtcaag
ccccgcgaggaaggctccctgctcatgacccttctggtcagttggggccccatgctgctg
ctgatcggcgtgtgggtgtacttcatgcggcagatgcagggcggtggcaagggcggtgcc
ttcagcttcggcaagagcaaggcgcgcatgctcgacgagaacaacaacaccgtgaccttc
gccgacgtggccggctgcgacgaggccaaggaagaggtcaaggaggtcgtcgacttcctg
aaggatccgcagaagttccagaagctcggcggccgcatcccgcgcggcctgctgctggtc
ggccctcccggcaccggcaagacgctgctggccaagtccatcgccggcgaggccaaggtg
ccgttcttcagcatctccggctccgacttcgtcgagatgttcgtcggcgtgggcgcggcc
cgcgtgcgcgacatgttcgagaacgccaagaagaacgcccccctgcatcatcttcatcgat
gaaatcgacgcggtgggccgccagcgtggcgccggcctgggcggcggcaacgacgagcgc
gagcagaccctcaaccagatgctggtcgagatggatggtttcgaaaccaacctgggcgtc
atcgtggtcgccgcgaccaaccgcccggacatcctggacgccgcgctgctgcgcccggc
cgcttcgaccgccaggtgtacgtcacgctgccggacatccgcggccgcgagcagatcctg
aacgtccacatgcgcaaggtgcccgtgggccaggacgtgaacgccgccgtgatcgcccgc
ggcacgcccggcatgtccggcgccgacctcgccaacctctgcaacgaggccgcgctgatg
gccgcccgccgcaacgcacgcacggtggagatgcaggacttcgagaaggccaaggacaag
atcatcatgggccccgagcgcaagagcatggtcatgcccgaggaagagcgccgcaacacg
gcctaccacgaggccggccacgcgctcatcggcaagctgctgcccaagtgcgaccccgtc
cacaaggtcaccatcatcccgcgtggccgcgcgctgggcgtgacgatgagcctgcccgag
aaggaccgctacagctacgaccgcgagtacatgctcaaccagatcagcatgctgttcggc
ggccgtatcgcggaagaggtcttcatgaaccagatgacgaccggcgcgagcaacgacttc
gagcgcgcgacctcgatcgcccgcgacatggtgacgcgctacggcatgaccgaggccctg
ggccccatggtctatgccgagaacgagggcgaagtcttcctcggccgttccgtcaccaag
acgaacaacatgagcgagcagacgatggagaaggtggacggtgaagtccgccgcatcatc
gacgaacagtacgccctggcgcgcaagctcatcgaggacaaccaggacaagatgcatgcc
atggcgaacgcgctgctggaatgggaaaccatcgacaccgagcagctggacgacatcatg
gcgggcaagccaccccgtcctcccaaggactggacgccccgcacgccgtcgtccggcggc
gacaactcctccggcggcggaacgccggcgccggtgtccagcgatccttcgcccaccgtg
gcgtaa
```

DBGET integrated database retrieval system

图3.3 登录KEGG获取*ftsH*基因基本信息

基因基本信息及序列获取

根据基因号，登录KEGG（https://www.kegg.jp/），可以得到*ftsH*的碱基序列和氨基酸序列，同时还可以得到*ftsH* 的同源基因、是否处于某条信号通路中、基因簇、在基因图谱中的位置及翻译方向、产物蛋白所属家族及相关注释信息、在NCBI及UniProt中的登录号等信息。



Acidovorax citrulli

Genome info　Pathway map

Search genes:

Genome information

T number	T00451
Org code	aav
Aliases	ACIAC, 397945
Full name	Acidovorax citrulli
Definition	Acidovorax citrulli AAC
Annotation	manual
Taxonomy	TAX: 397945
Lineage	Bacteria; Proteobacteri
Data source	GenBank (Assembly: G BioProject: 15708
Original DB	JGI
Keywords	Plant pathogen
Chromosome	Circular
Sequence	GB: CP000512
Length	5352772
Statistics	Number of nucleotides: 5352772 Number of protein genes: 4709 Number of RNA genes: 65
Created	2007

KEGG2　PATHWAY　BRITE　GENES　GENOME　LIGAND　DISEASE

Aminoglycoside resistance, protease FtsH

[Module menu | Organism menu | Module entry | KO/Gene/Compound list]

Entry	aav_M00742
Name	Aminoglycoside resistance, protease FtsH
Definition	(K07638 K07659 K19416),(K04088 K04087) K03798 Ortholog table　Taxonomy
Type	Functional set
Organism	Acidovorax citrulli AAC00-1
Comment	Resistance mechanism: elimination of membra products

Acidovorax citrulli

SSDB Paralog Search Result

KEGG ID : aav:Aave_2617 (641 a.a.)
Definition: membrane protease FtsH catalytic subunit; K03798 cell division protease FtsH
Update status: T00451
Show : ○ Best-best ○ Best ● Paralogs ○ Gene clusters
Threshold: 100 Go

All　Select operation　Select

Search Result : 5 hits

	Entry	KO	len	SW-score (margin)	bits	identity	overlap	best (all)
☑	aav:Aave_3536 AAA ATPase, central domain protein		336	324 (　-)	80	0.335	197	->
☑	aav:Aave_1457 ATP-dependent proteinase	K01338	808	163 (　-)	43	0.273	227	->
☑	aav:Aave_1458 ATP-dependent Clp protease, ATP-binding s	K03544	421	140 (　-)	38	0.333	78	->
☑	aav:Aave_0231 Mg chelatase, subunit ChlI	K07391	517	105 (　-)	30	0.667	21	->
☑	aav:Aave_4237 Holliday junction DNA helicase RuvB	K03551	354	104 (　-)	30	0.394	66	->

[SSDB | GENES | KEGG2 | KEGG | GenomeNet]

SSDB Best Searc

KEGG ID : aav:Aave_
Definition: membrane
Update status: T00451
Show : ○ Best-best
Threshold: 100 Go

Top10　Select operation　Select

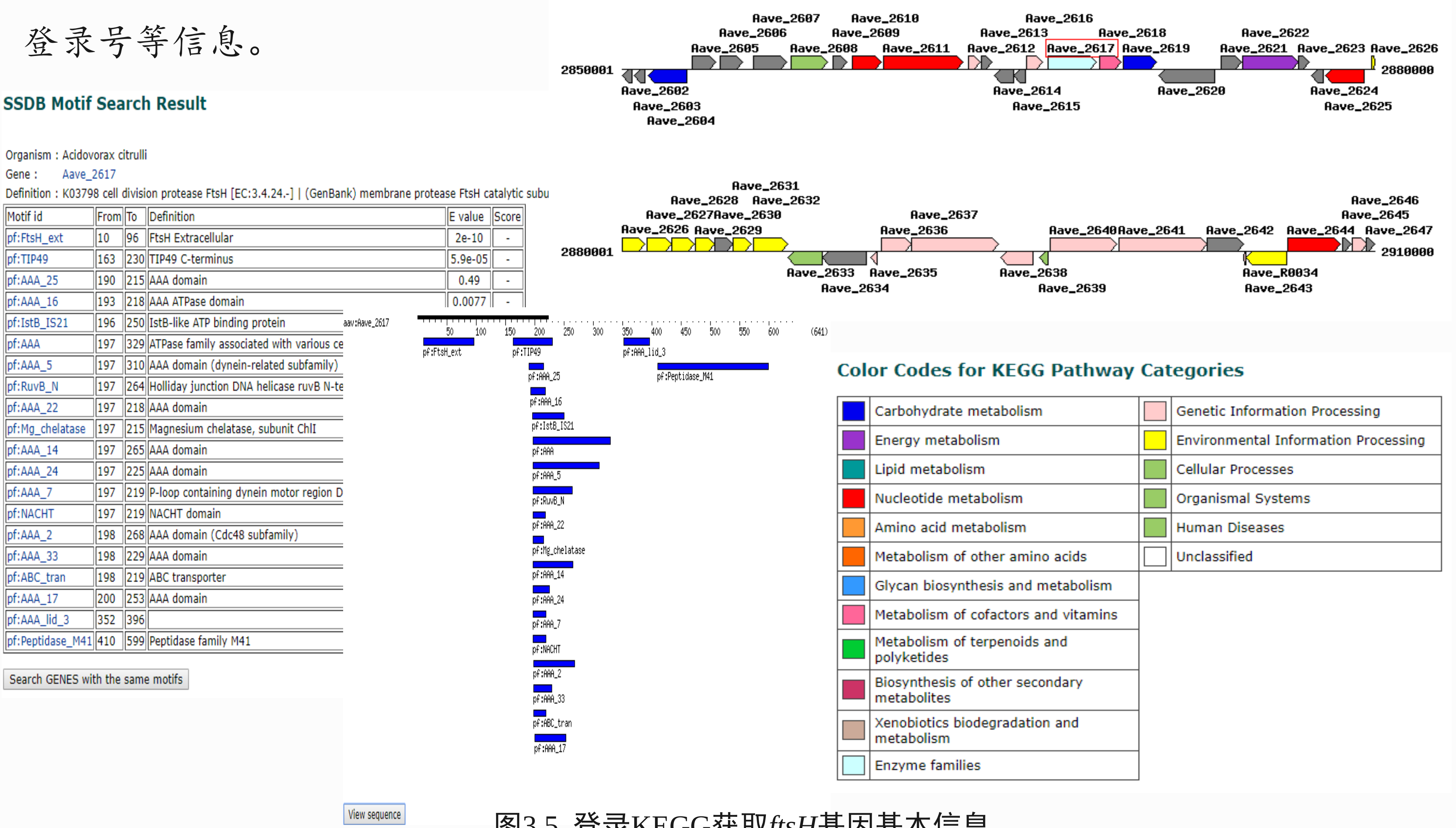
Search Result : 5595 hits

	Entry	KO	len	SW-score (margin)	bits	identity	overlap	best (all)
☑	aaa:Acav_2586 ATP-dependent metalloprotease FtsH	K03798	641	4214 (4051)	966	0.995	641	<-> 6
☑	ajs:Ajs_2383 membrane protease FtsH catalytic subunit	K03798	639	3997 (2170)	917	0.938	641	<-> 7
☑	dia:Dtpsy_1475 ATP-dependent metalloprotease FtsH	K03798	639	3997 (3847)	917	0.938	641	<-> 5
☑	dpy:BA022_06670 cell division protein FtsH	K03798	639	3989 (3844)	915	0.936	641	<-> 4
☑	adk:Alide2_2675 ATP-dependent metalloprotease FtsH	K03798	653	3982 (2214)	914	0.930	642	<-> 10
☑	adn:Alide_2498 ATP-dependent metalloprotease FtsH	K03798	653	3982 (2209)	914	0.930	642	<-> 8
☑	acra:BSY15_2906 ATP-dependent metallopeptidase HflB fam	K03798	639	3977 (3772)	912	0.938	641	<-> 4
☑	mela:C6568_00730 cell division protein FtsH	K03798	638	3973 (3630)	911	0.930	641	<-> 9
☑	melm:C7H73_09645 cell division protein FtsH	K03798	638	3971 (3827)	911	0.933	641	<-> 6
☑	acid:CBP33_12420 cell division protein FtsH	K03798	639	3960 (3856)	909	0.931	641	<-> 3
☐	acin:CBP34_12790 cell division protein FtsH	K03798	639	3960 (3856)	909	0.931	641	<-> 3
☐	acip:CBP36_12900 cell division protein FtsH	K03798	639	3960 (3856)	909	0.931	641	<-> 3
☐	acis:CBP35_06020 cell division protein FtsH	K03798	639	3960 (3856)	909	0.931	641	<-> 3
☐	simp:C6571_01000 cell division protein FtsH	K03798	639	3952 (3783)	907	0.924	641	<-> 5
☐	dac:Daci_4925 ATP-dependent metalloprotease FtsH	K03798	640	3941 (3657)	904	0.930	642	<-> 5

图3.4 登录KEGG获取*ftsH*基因基本信息

基因基本信息及序列获取

根据基因号，登录KEGG（https://www.kegg.jp/），可以得到*ftsH*的碱基序列和氨基酸序列，同时还可以得到*ftsH* 的同源基因、是否处于某条信号通路中、基因簇、在基因图谱中的位置及翻译方向、产物蛋白所属家族及相关注释信息、在NCBI及UniProt中的登录号等信息。



SSDB Motif Search Result

Organism : Acidovorax citrulli

Gene : Aave_2617

Definition : K03798 cell division protease FtsH [EC:3.4.24.-] | (GenBank) membrane protease FtsH catalytic subu

Motif id	From	To	Definition	E value	Score
pf:FtsH_ext	10	96	FtsH Extracellular	2e-10	-
pf:TIP49	163	230	TIP49 C-terminus	5.9e-05	-
pf:AAA_25	190	215	AAA domain	0.49	-
pf:AAA_16	193	218	AAA ATPase domain	0.0077	-
pf:IstB_IS21	196	250	IstB-like ATP binding protein		
pf:AAA	197	329	ATPase family associated with various ce		
pf:AAA_5	197	310	AAA domain (dynein-related subfamily)		
pf:RuvB_N	197	264	Holliday junction DNA helicase ruvB N-te		
pf:AAA_22	197	218	AAA domain		
pf:Mg_chelatase	197	215	Magnesium chelatase, subunit ChlI		
pf:AAA_14	197	265	AAA domain		
pf:AAA_24	197	225	AAA domain		
pf:AAA_7	197	219	P-loop containing dynein motor region D		
pf:NACHT	197	219	NACHT domain		
pf:AAA_2	198	268	AAA domain (Cdc48 subfamily)		
pf:AAA_33	198	229	AAA domain		
pf:ABC_tran	198	219	ABC transporter		
pf:AAA_17	200	253	AAA domain		
pf:AAA_lid_3	352	396			
pf:Peptidase_M41	410	599	Peptidase family M41		

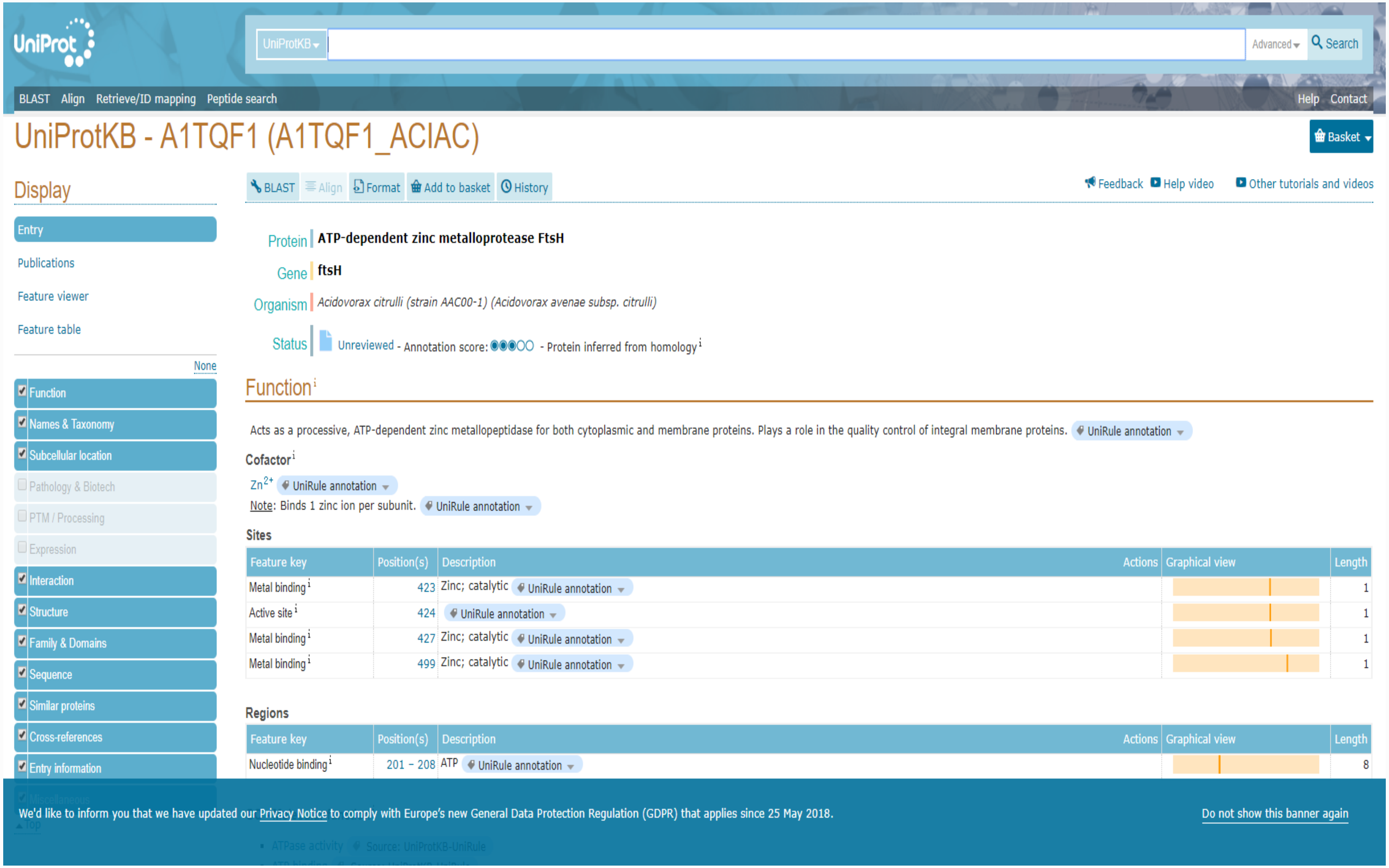
Color Codes for KEGG Pathway Categories

🟦	Carbohydrate metabolism	🟥	Genetic Information Processing
🟪	Energy metabolism	🟨	Environmental Information Processing
🟦	Lipid metabolism	🟩	Cellular Processes
🟥	Nucleotide metabolism	🟩	Organismal Systems
🟧	Amino acid metabolism	🟩	Human Diseases
🟧	Metabolism of other amino acids	⬜	Unclassified
🟦	Glycan biosynthesis and metabolism		
🟥	Metabolism of cofactors and vitamins		
🟩	Metabolism of terpenoids and polyketides		
🟥	Biosynthesis of other secondary metabolites		
🟫	Xenobiotics biodegradation and metabolism		
⬜	Enzyme families		

图3.5 登录KEGG获取*ftsH*基因基本信息

基因基本信息及序列获取

根据基因号，登录UniProt（https://www.uniprot.org/），同样可以得到*ftsH*的碱基序列和氨基酸序列。



UniProtKB - A1TQF1 (A1TQF1_ACIAC)

Protein | ATP-dependent zinc metalloprotease FtsH

Gene | ftsH

Organism | *Acidovorax citrulli (strain AAC00-1) (Acidovorax avenae subsp. citrulli)*

Status | Unreviewed - Annotation score: ●●●○○ - Protein inferred from homology i

Function i

Acts as a processive, ATP-dependent zinc metallopeptidase for both cytoplasmic and membrane proteins. Plays a role in the quality control of integral membrane proteins. UniRule annotation ▾

Cofactor i

Zn^{2+} UniRule annotation ▾

Note: Binds 1 zinc ion per subunit. UniRule annotation ▾

Sites

Feature key	Position(s)	Description	Actions	Graphical view	Length
Metal binding i	423	Zinc; catalytic UniRule annotation ▾			1
Active site i	424	UniRule annotation ▾			1
Metal binding i	427	Zinc; catalytic UniRule annotation ▾			1
Metal binding i	499	Zinc; catalytic UniRule annotation ▾			1

Regions

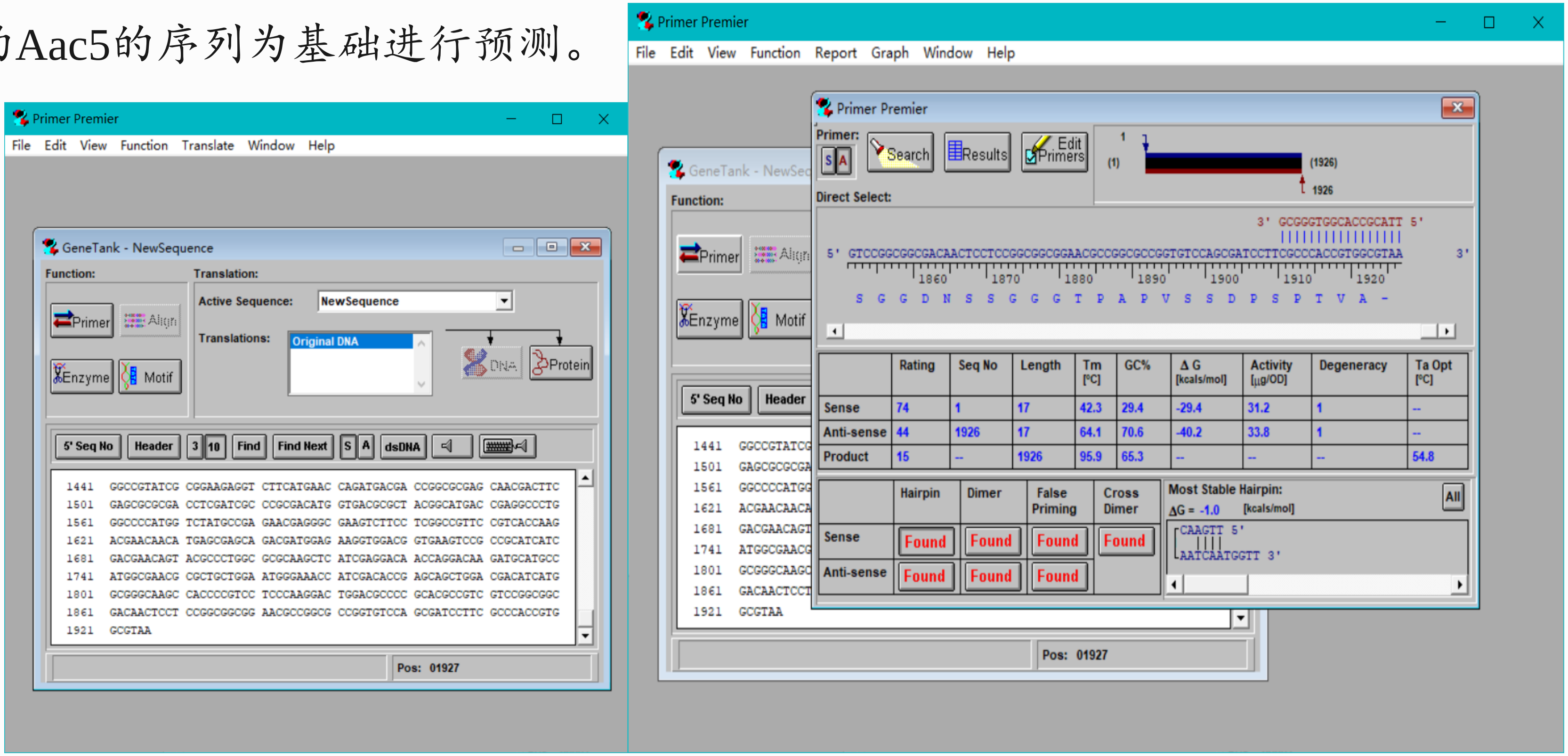
Feature key	Position(s)	Description	Actions	Graphical view	Length
Nucleotide binding i	201 - 208	ATP UniRule annotation ▾			8

We'd like to inform you that we have updated our Privacy Notice to comply with Europe's new General Data Protection Regulation (GDPR) that applies since 25 May 2018.

图3.6 登录UniProt获取基因相关信息

2. *ftsH*碱基序列获取和验证

由于实验室所用菌株为Aac5，但西瓜噬酸菌中进行全基因组测序的菌株为AAC00-1，基于两个菌株同属于II组菌株，序列存在高相似性，因此可以AAC00-1的序列为基础，以Aac5菌株DNA为模板，设计简单引物，克隆目的基因ftsH，然后将克隆结果进行测序，并与AAC00-1的序列进行比对，如比对结果完全一致，则后续预测可基于AAC00-1序列进行，如比对后发现两菌株序列存在差异，则应以测序所得的Aac5的序列为基础进行预测。



```
1441  GGCCGTATCG CGGAAGAGGT CTTCATGAAC CAGATGACGA CCGGCGCGAG CAACGACTTC
1501  GAGCGCGCGA CCTCGATCGC CCGCGACATG GTGACGCGCT ACGGCATGAC CGAGGCCCTG
1561  GGCCCCATGG TCTATGCCGA GAACGAGGGC GAAGTCTTCC TCGGCCGTTC CGTCACCAAG
1621  ACGAACAACA TGAGCGAGCA GACGATGGAG AAGGTGGACG GTGAAGTCCG CCGCATCATC
1681  GACGAACAGT ACGCCCTGGC GCGCAAGCTC ATCGAGGACA ACCAGGACAA GATGCATGCC
1741  ATGGCGAACG CGCTGCTGGA ATGGGAAACC ATCGACACCG AGCAGCTGGA CGACATCATG
1801  GCGGGCAAGC CACCCCGTCC TCCCAAGGAC TGGACGCCCC GCACGCCGTC GTCCGGCGGC
1861  GACAACTCCT CCGGCGGCGG AACGCCGGCG CCGGTGTCCA GCGATCCTTC GCCCACCGTG
1921  GCGTAA
```

Pos: 01927

```
                                                         3' GCGGGTGGCACCGCATT 5'
                                                            |||||||||||||||||
5' GTCCGGCGGCGACAACTCCTCCGGCGGCGGAACGCCGGCGCCGGTGTCCAGCGATCCTTCGCCCACCGTGGCGTAA    3'
          1860      1870      1880      1890      1900      1910      1920
   S  G  G  D  N  S  S  G  G  G  T  P  A  P  V  S  S  D  P  S  P  T  V  A  -
```

	Rating	Seq No	Length	Tm [°C]	GC%	Δ G [kcals/mol]	Activity [μg/OD]	Degeneracy	Ta Opt [°C]
Sense	74	1	17	42.3	29.4	-29.4	31.2	1	--
Anti-sense	44	1926	17	64.1	70.6	-40.2	33.8	1	--
Product	15	--	1926	95.9	65.3	--	--	--	54.8

	Hairpin	Dimer	False Priming	Cross Dimer
Sense	Found	Found	Found	Found
Anti-sense	Found	Found	Found	

Most Stable Hairpin:
ΔG = -1.0 [kcals/mol]

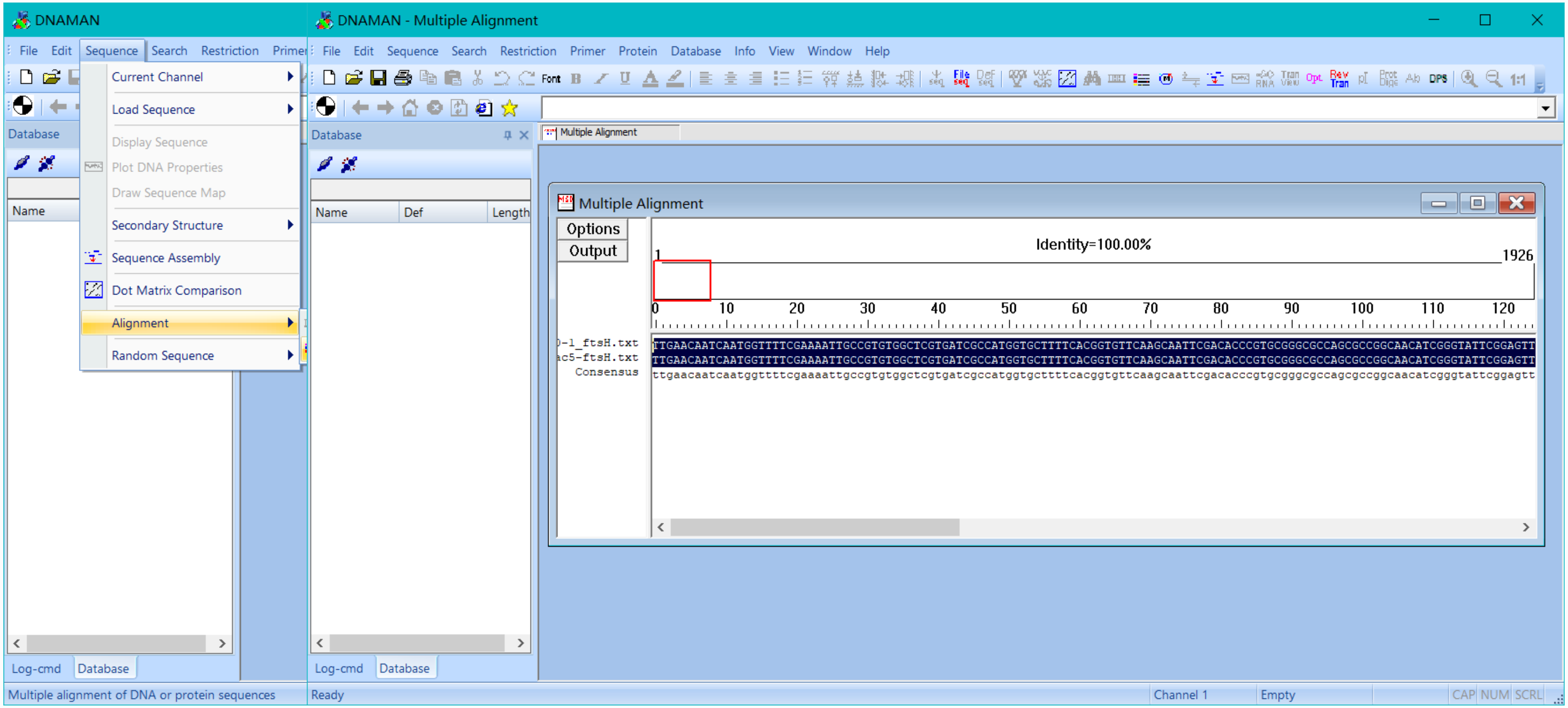
```
┌CAAGTT 5'
   ||||
└AATCAATGGTT 3'
```

Pos: 01927

图3.7 使用Primer Premier 5设计引物

2. *ftsH*碱基序列获取和验证

得到基因*ftsH*在AAC00-1菌株中的碱基序列后，以Aac5菌株DNA为模板，克隆目的基因*ftsH*，将克隆结果测序后，使用DNAMAN对测序结果和AAC00-1的序列进行比对，发现两菌株的*ftsH*序列完全一致（图11），则使用AAC00-1的序列进行后续预测可行。



```
Identity=100.00%
1                                                                                                              1926
0        10        20        30        40        50        60        70        80        90        100       110       120
)-1_ftsH.txt TTGAACAATCAATGGTTTTCGAAAATTGCCGTGTGGCTCGTGATCGCCATGGTGCTTTTCACGGTGTTCAAGCAATTCGACACCCGTGCGGGCGCCAGCGCCGGCAACATCGGGTATTCGGAGTT
ac5-ftsH.txt TTGAACAATCAATGGTTTTCGAAAATTGCCGTGTGGCTCGTGATCGCCATGGTGCTTTTCACGGTGTTCAAGCAATTCGACACCCGTGCGGGCGCCAGCGCCGGCAACATCGGGTATTCGGAGTT
   Consensus ttgaacaatcaatggttttcgaaaattgccgtgtggctcgtgatcgccatggtgcttttcacggtgttcaagcaattcgacacccgtgcgggcgccagcgccggcaacatcgggtattcggagtt
```

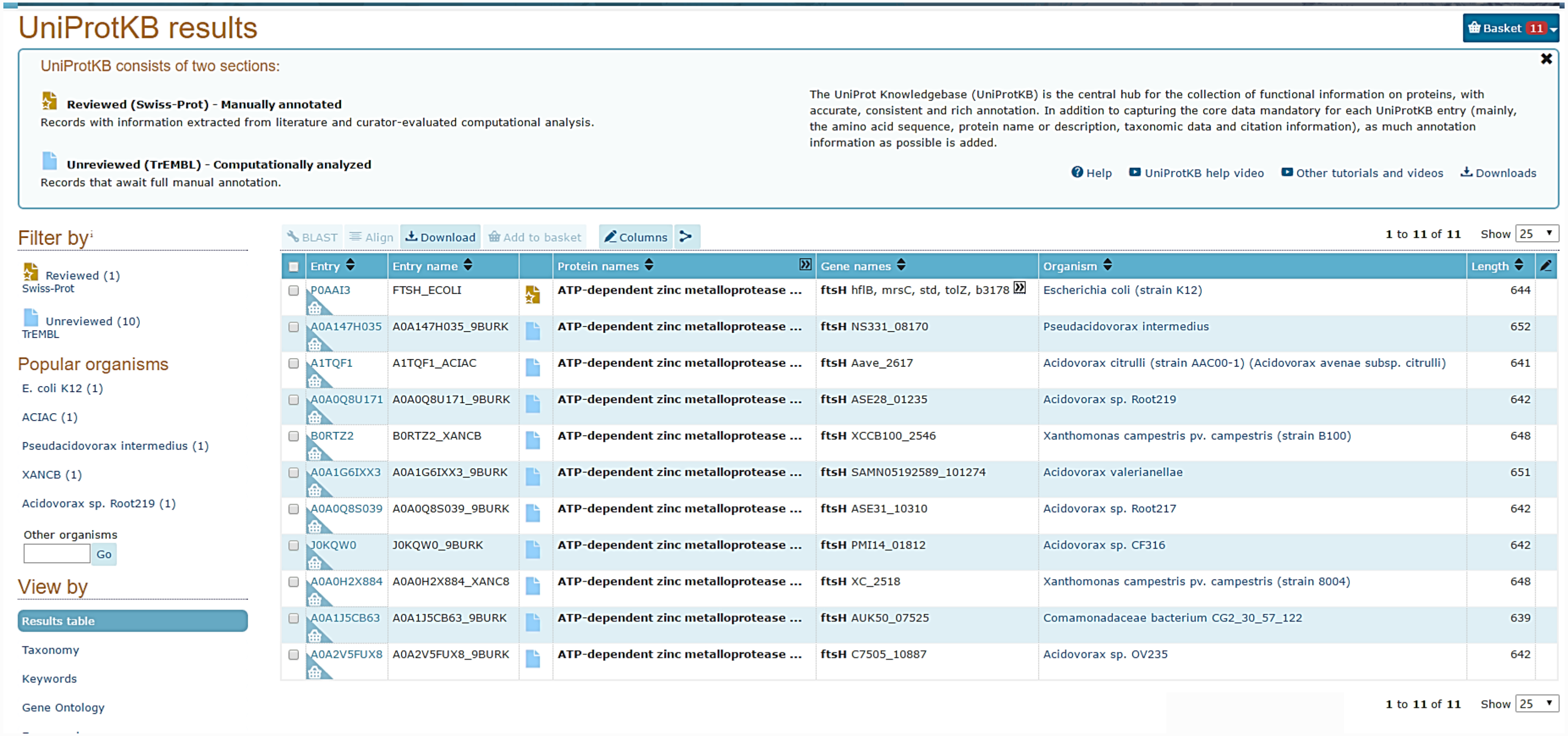
图3.8 使用DNAMAN比对AAC00-1和Aac5的*ftsH*序列

蛋白结构及功能预测

1. 氨基酸序列同源性分析、保守位点分析及系统发育树构建
2. 蛋白理化性质分析、跨膜结构预测和亚细胞定位预测
3. 蛋白二级结构、三级结构预测、结构域预测
4. 基因互作关系预测及蛋白功能预测分析

4.1 FtsH氨基酸序列同源性分析、保守位点分析及系统发育树构建

登录UniProt数据库（https://www.uniprot.org/），获取大肠杆菌及分类学上与西瓜噬酸菌分类地位相近的细菌的FtsH蛋白氨基酸序列。



UniProtKB results

UniProtKB consists of two sections:

Reviewed (Swiss-Prot) - Manually annotated
Records with information extracted from literature and curator-evaluated computational analysis.

Unreviewed (TrEMBL) - Computationally analyzed
Records that await full manual annotation.

The UniProt Knowledgebase (UniProtKB) is the central hub for the collection of functional information on proteins, with accurate, consistent and rich annotation. In addition to capturing the core data mandatory for each UniProtKB entry (mainly, the amino acid sequence, protein name or description, taxonomic data and citation information), as much annotation information as possible is added.

Filter by

Reviewed (1)
Swiss-Prot

Unreviewed (10)
TrEMBL

Popular organisms

- E. coli K12 (1)
- ACIAC (1)
- Pseudacidovorax intermedius (1)
- XANCB (1)
- Acidovorax sp. Root219 (1)

Other organisms

View by

1 to 11 of 11 Show 25

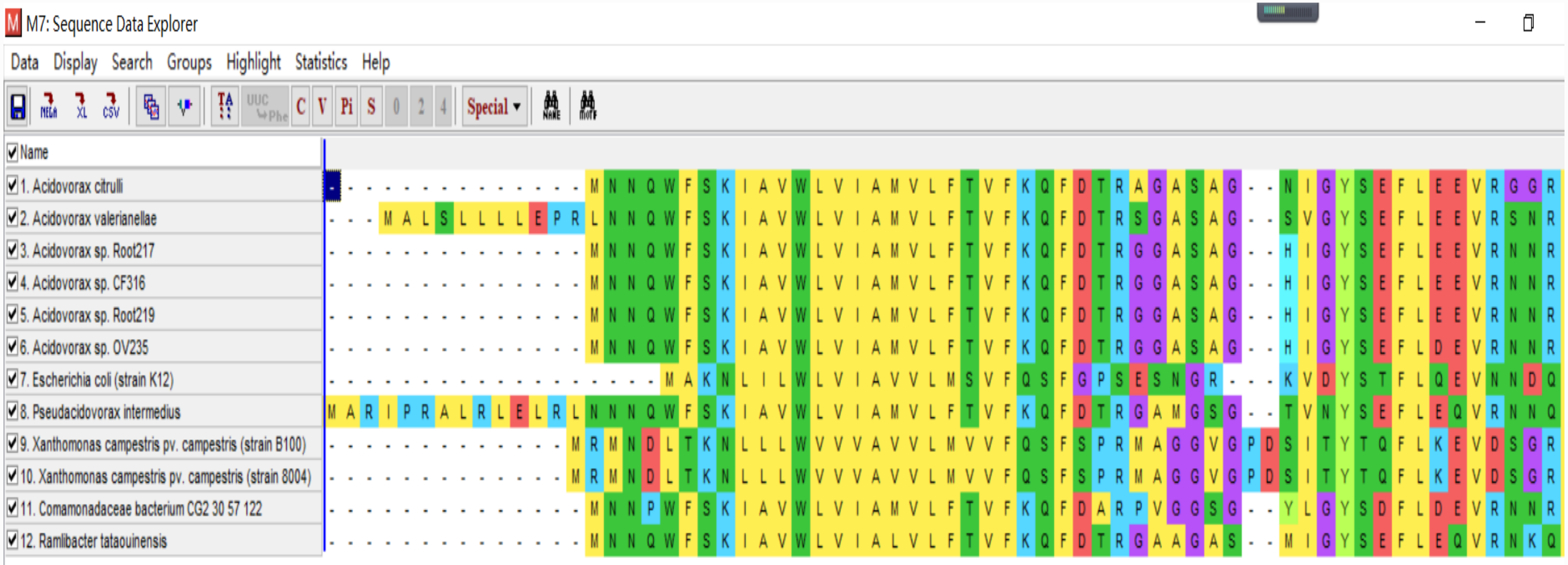
Entry	Entry name		Protein names	Gene names	Organism	Length
P0AAI3	FTSH_ECOLI		**ATP-dependent zinc metalloprotease ...**	**ftsH** hflB, mrsC, std, tolZ, b3178	Escherichia coli (strain K12)	644
A0A147H035	A0A147H035_9BURK		**ATP-dependent zinc metalloprotease ...**	**ftsH** NS331_08170	Pseudacidovorax intermedius	652
A1TQF1	A1TQF1_ACIAC		**ATP-dependent zinc metalloprotease ...**	**ftsH** Aave_2617	Acidovorax citrulli (strain AAC00-1) (Acidovorax avenae subsp. citrulli)	641
A0A0Q8U171	A0A0Q8U171_9BURK		**ATP-dependent zinc metalloprotease ...**	**ftsH** ASE28_01235	Acidovorax sp. Root219	642
B0RTZ2	B0RTZ2_XANCB		**ATP-dependent zinc metalloprotease ...**	**ftsH** XCCB100_2546	Xanthomonas campestris pv. campestris (strain B100)	648
A0A1G6IXX3	A0A1G6IXX3_9BURK		**ATP-dependent zinc metalloprotease ...**	**ftsH** SAMN05192589_101274	Acidovorax valerianellae	651
A0A0Q8S039	A0A0Q8S039_9BURK		**ATP-dependent zinc metalloprotease ...**	**ftsH** ASE31_10310	Acidovorax sp. Root217	642
J0KQW0	J0KQW0_9BURK		**ATP-dependent zinc metalloprotease ...**	**ftsH** PMI14_01812	Acidovorax sp. CF316	642
A0A0H2X884	A0A0H2X884_XANC8		**ATP-dependent zinc metalloprotease ...**	**ftsH** XC_2518	Xanthomonas campestris pv. campestris (strain 8004)	648
A0A1J5CB63	A0A1J5CB63_9BURK		**ATP-dependent zinc metalloprotease ...**	**ftsH** AUK50_07525	Comamonadaceae bacterium CG2_30_57_122	639
A0A2V5FUX8	A0A2V5FUX8_9BURK		**ATP-dependent zinc metalloprotease ...**	**ftsH** C7505_10887	Acidovorax sp. OV235	642

1 to 11 of 11 Show 25

图4.1 使用UniProt下载同源氨基酸序列

蛋白结构及功能预测

使用MEGA7软件对西瓜噬酸菌、大肠杆菌及其他分类地位相近的细菌的FtsH蛋白进行氨基酸序列比对。

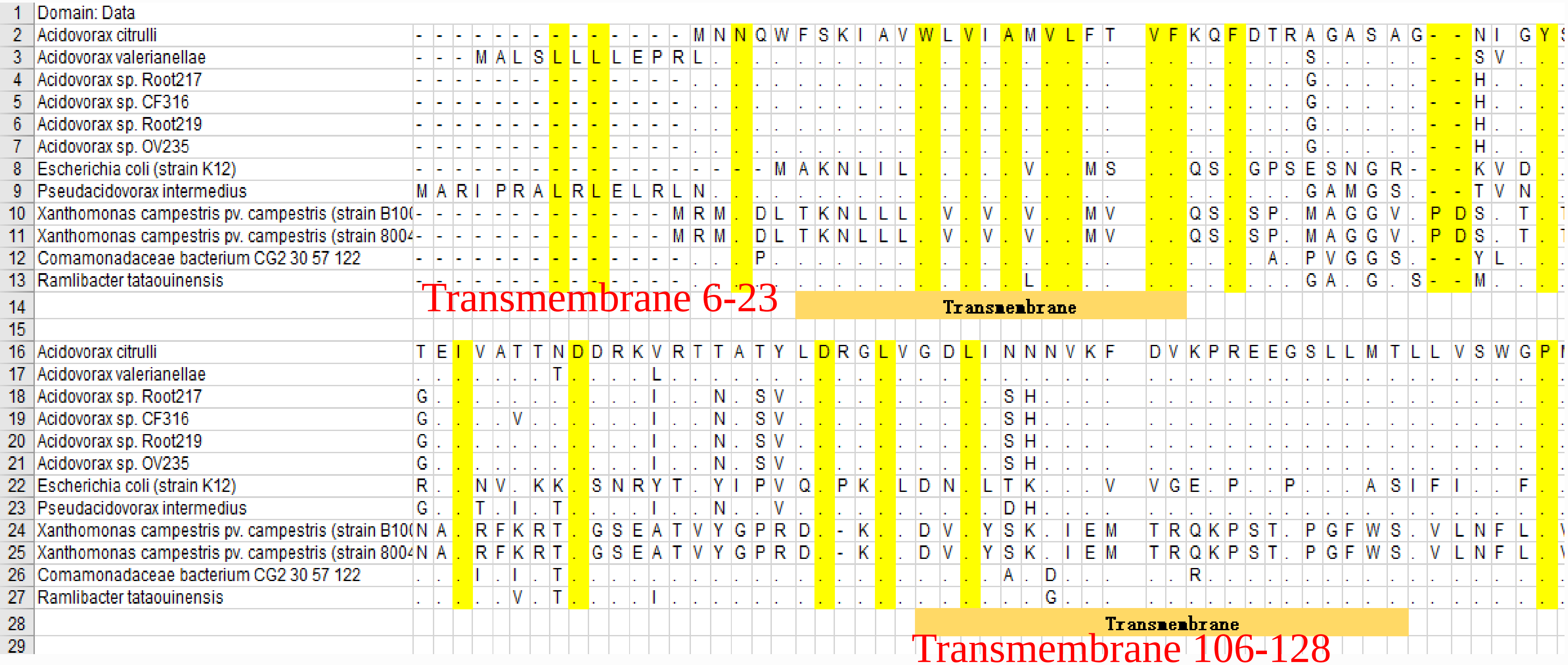


```
☑Name
☑1. Acidovorax citrulli                                  --------------MNNQWFSKIAVWLVIAMVLFTVFKQFDTRAGASAG--NIGYSEFLEEVRGGR
☑2. Acidovorax valerianellae                             ---MALSLLLLEPRLNNQWFSKIAVWLVIAMVLFTVFKQFDTRSGASAG--SVGYSEFLEEVRSNR
☑3. Acidovorax sp. Root217                               --------------MNNQWFSKIAVWLVIAMVLFTVFKQFDTRGGASAG--HIGYSEFLEEVRNNR
☑4. Acidovorax sp. CF316                                 --------------MNNQWFSKIAVWLVIAMVLFTVFKQFDTRGGASAG--HIGYSEFLEEVRNNR
☑5. Acidovorax sp. Root219                               --------------MNNQWFSKIAVWLVIAMVLFTVFKQFDTRGGASAG--HIGYSEFLEEVRNNR
☑6. Acidovorax sp. OV235                                 --------------MNNQWFSKIAVWLVIAMVLFTVFKQFDTRGGASAG--HIGYSEFLDEVRNNR
☑7. Escherichia coli (strain K12)                        ------------------MAKNLILWLVIAVVLMSVFQSFGPSESNGR---KVDYSTFLQEVNNDQ
☑8. Pseudacidovorax intermedius                          MARIPRALRLELRLNNNQWFSKIAVWLVIAMVLFTVFKQFDTRGAMGSG--TVNYSEFLEQVRNNQ
☑9. Xanthomonas campestris pv. campestris (strain B100)   -------------MRMNDLTKNLLLWVVVAVVLMVVFQSFSPRMAGGVGPDSITYTQFLKEVDSGR
☑10. Xanthomonas campestris pv. campestris (strain 8004)  -------------MRMNDLTKNLLLWVVVAVVLMVVFQSFSPRMAGGVGPDSITYTQFLKEVDSGR
☑11. Comamonadaceae bacterium CG2 30 57 122              --------------MNNPWFSKIAVWLVIAMVLFTVFKQFDARPVGGSG--YLGYSDFLDEVRNNR
☑12. Ramlibacter tataouinensis                           --------------MNNQWFSKIAVWLVIALVLFTVFKQFDTRGAAGAS--MIGYSEFLEQVRNKQ
```

图4.2 使用MEGA7对各细菌中的FtsH氨基酸序列进行比对分析

蛋白结构及功能预测

通过氨基酸序列比对，我们发现FtsH蛋白在不同菌种中的氨基酸序列相似性很高，说明FtsH蛋白在原核生物中可能是一个保守蛋白。我们进一步对相对保守的氨基酸序列进行分析发现，这些保守的氨基酸序列多为一些具有特殊功能的结构域。



```
1  Domain: Data
2  Acidovorax citrulli                                  - - - - - - - - - - - - - - M N N Q W F S K I A V W L V I A M V L F T V F K Q F D T R A G A S A G - - N I G Y
3  Acidovorax valerianellae                             - - - M A L S L L L L E P R L . . . . . . . . . . . . . . . . . . . . . . . . . . . . S . . . . . - - S V . .
4  Acidovorax sp. Root217                               - - - - - - - - - - - - - - . . . . . . . . . . . . . . . . . . . . . . . . . . . . . G . . . . . - - H . . .
5  Acidovorax sp. CF316                                 - - - - - - - - - - - - - - . . . . . . . . . . . . . . . . . . . . . . . . . . . . . G . . . . . - - H . . .
6  Acidovorax sp. Root219                               - - - - - - - - - - - - - - . . . . . . . . . . . . . . . . . . . . . . . . . . . . . G . . . . . - - H . . .
7  Acidovorax sp. OV235                                 - - - - - - - - - - - - - - . . . . . . . . . . . . . . . . . . . . . . . . . . . . . G . . . . . - - H . . .
8  Escherichia coli (strain K12)                        - - - - - - - - - - - - - - - - - - M A K N L I L . . . . . V . . M S . . Q S . G P S E S N G R - - - K V D .
9  Pseudacidovorax intermedius                          M A R I P R A L R L E L R L N . . . . . . . . . . . . . . . . . . . . . . . . . . . . G A M G S . - - T V N .
10 Xanthomonas campestris pv. campestris (strain B10    - - - - - - - - - - - - - M R M . D L T K N L L L . V . V . V . . M V . . Q S . S P . M A G G V . P D S . T .
11 Xanthomonas campestris pv. campestris (strain 800    - - - - - - - - - - - - - M R M . D L T K N L L L . V . V . V . . M V . . Q S . S P . M A G G V . P D S . T .
12 Comamonadaceae bacterium CG2 30 57 122               - - - - - - - - - - - - - - . . . P . . . . . . . . . . . . . . . . . . . . . . . A . P V G G S . - - Y L . .
13 Ramlibacter tataouinensis                            - - - - - - - - - - - - - - . . . . . . . . . . . . . . . . L . . . . . . . . . . . . G A . G . S - - M . . .
14                                                                                            Transmembrane
15
16 Acidovorax citrulli                                  T E I V A T T N D D R K V R T T A T Y L D R G L V G D L I N N N V K F D V K P R E E G S L L M T L L V S W G P
17 Acidovorax valerianellae                             . . . . . . . T . . . . L . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
18 Acidovorax sp. Root217                               G . . . . . . . . . . . I . . N . S V . . . . . . . . . . S H . . . . . . . . . . . . . . . . . . . . . . . .
19 Acidovorax sp. CF316                                 G . . . . V . . . . . . I . . N . S V . . . . . . . . . . S H . . . . . . . . . . . . . . . . . . . . . . . .
20 Acidovorax sp. Root219                               G . . . . . . . . . . . I . . N . S V . . . . . . . . . . S H . . . . . . . . . . . . . . . . . . . . . . . .
21 Acidovorax sp. OV235                                 G . . . . . . . . . . . I . . N . S V . . . . . . . . . . S H . . . . . . . . . . . . . . . . . . . . . . . .
22 Escherichia coli (strain K12)                        R . . N V . K K . S N R Y T . Y I P V Q . P K . L D N . L T K . . . V V G E . P . . P . . . A S I F I . . F .
23 Pseudacidovorax intermedius                          G . . T . I . T . . . . I . . N . . V . . . . . . . . . . D H . . . . . . . . . . . . . . . . . . . . . . . .
24 Xanthomonas campestris pv. campestris (strain B10    N A . R F K R T . G S E A T V Y G P R D . - K . . D V . Y S K . I E M T R Q K P S T . P G F W S . V L N F L .
25 Xanthomonas campestris pv. campestris (strain 800    N A . R F K R T . G S E A T V Y G P R D . - K . . D V . Y S K . I E M T R Q K P S T . P G F W S . V L N F L .
26 Comamonadaceae bacterium CG2 30 57 122               . . . I . I . T . . . . . . . . . . . . . . . . . . . . . A . D . . . . . R . . . . . . . . . . . . . . . . .
27 Ramlibacter tataouinensis                            . . . . . V . T . . . . I . . . . . . . . . . . . . . . . . . G . . . . . . . . . . . . . . . . . . . . . . .
28                                                                                                              Transmembrane
29
```

Transmembrane 6-23

Transmembrane 106-128

图4.3 使用MEGA7分析各细菌FtsH氨基酸序列保守区域

蛋白结构及功能预测

通过氨基酸序列比对，我们发现FtsH蛋白在不同菌种中的氨基酸序列相似性很高，说明FtsH蛋白在原核生物中可能是一个保守蛋白。我们进一步对相对保守的氨基酸序列进行分析发现，这些保守的氨基酸序列多为一些具有特殊功能的结构域。

```
30 Acidovorax citrulli                                FSFGKSKARMLDENNNTVTFADVAGCDEAKEEVKEVVDFLKDPQKFQKLGGRIPRGLLLVGPPGTGKTLLAKSIAGEAKV [240]
31 Acidovorax valerianellae                           ..............S.Q............................................................... [240]
32 Acidovorax sp. Root217                             ................Q............................................................... [240]
33 Acidovorax sp. CF316                               ................Q............................................................... [240]
34 Acidovorax sp. Root219                             ................Q............................................................... [240]
35 Acidovorax sp. OV235                               ................Q............................................................... [240]
36 Escherichia coli (strain K12)                      M..........T.DQIKT.................A.L.EY.RE.SR......K..K.V.M.............A....... [240]
37 Pseudacidovorax intermedius                        ................................................................................ [240]
38 Xanthomonas campestris pv. campestris (strain B10( M.....R.KLQG.DQVK................G.L....R..T..T....K....V.M.............A....... [240]
39 Xanthomonas campestris pv. campestris (strain 8004 M.....R.KLQG.DQVK................G.L....R..T..T....K....V.M.............A....... [240]
40 Comamonadaceae bacterium CG2 30 57 122             .........L....T............................A.................................... [240]
41 Ramlibacter tataouinensis                          ................QI......................................................G....... [240]
42                                                    AAA
43                                                    Nucleotide binding
44
45 Acidovorax citrulli                                PFFSISGSDFVEMFVGVGAARVRDMFENAKKNAPCIIFIDEIDAVGRQRGAGLGGGNDEREQTLNQMLVEMDGFETNLGV [320]
46 Acidovorax valerianellae                           ................................................................................ [320]
47 Acidovorax sp. Root217                             ..........................D..................................................... [320]
48 Acidovorax sp. CF316                               ..........................D..................................................... [320]
49 Acidovorax sp. Root219                             ..........................D..................................................... [320]
50 Acidovorax sp. OV235                               ..........................D..................................................... [320]
51 Escherichia coli (strain K12)                      ...T...............S.......Q...A........................H..................G.E.I [320]
52 Pseudacidovorax intermedius                        ................................................................................ [320]
53 Xanthomonas campestris pv. campestris (strain B10( ...................S.......Q...H...............H........H.........L........GGE.. [320]
54 Xanthomonas campestris pv. campestris (strain 8004 ...................S.......Q...H...............H........H.........L........GGE.. [320]
55 Comamonadaceae bacterium CG2 30 57 122             ...................S.........................................................V.. [320]
56 Ramlibacter tataouinensis                          ................................................................................ [320]
57                                                    AAA
58
59 Acidovorax citrulli                                IVVAATNRPDILDAALLRPGRFDRQVYVTLPDIRGREQILNVHMRKVPVGQDVNAAVIARGTPGMSGADLANLCNEAALM [400]
60 Acidovorax valerianellae                           ......................................................PS........................ [400]
61 Acidovorax sp. Root217                             ..............................................I.I....APNI....................... [400]
62 Acidovorax sp. CF316                               ..............................................I.I....APGI....................... [400]
63 Acidovorax sp. Root219                             ..............................................I.I....APNI....................... [400]
64 Acidovorax sp. OV235                               ..............................................I.I....APNI....................... [400]
65 Escherichia coli (strain K12)                      ..I.......V..P............V.G...V.......K....R..LAP.ID..I.......F........V.....F [400]
66 Pseudacidovorax intermedius                        ........................................S.....I.......PSI....................... [400]
67 Xanthomonas campestris pv. campestris (strain B10( ..I.......V..P............V.G...VK......R.....L.LAD..VPL........F........A.....F [400]
68 Xanthomonas campestris pv. campestris (strain 8004 ..I.......V..P............V.G...VK......R.....L.LAD..VPL........F........A.....F [400]
69 Comamonadaceae bacterium CG2 30 57 122             ................................................LSP............................. [400]
70 Ramlibacter tataouinensis                          ............................Q.................I.I.T....SI....................... [400]
71                                                    AAA
72
```

AAA Domain 193-332

图4.4 使用MEGA7分析各细菌FtsH氨基酸序列保守区域

蛋白结构及功能预测

根据UniProt数据库所得FtsH蛋白氨基酸序列，分别对上述细菌的FtsH蛋白通过MEGA7软件进行比对，构建系统发育树，使用NJ邻位法对进化距离进行分析，自展检测重复100次。

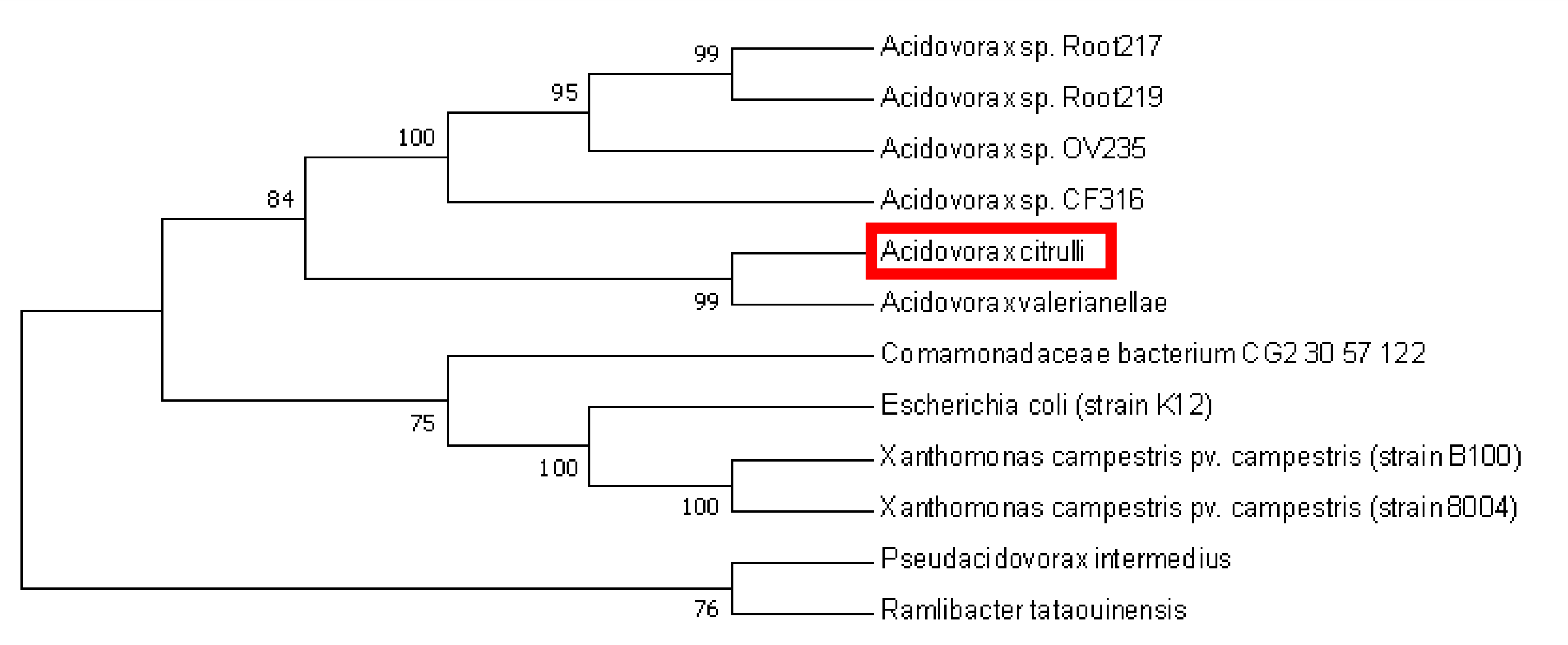
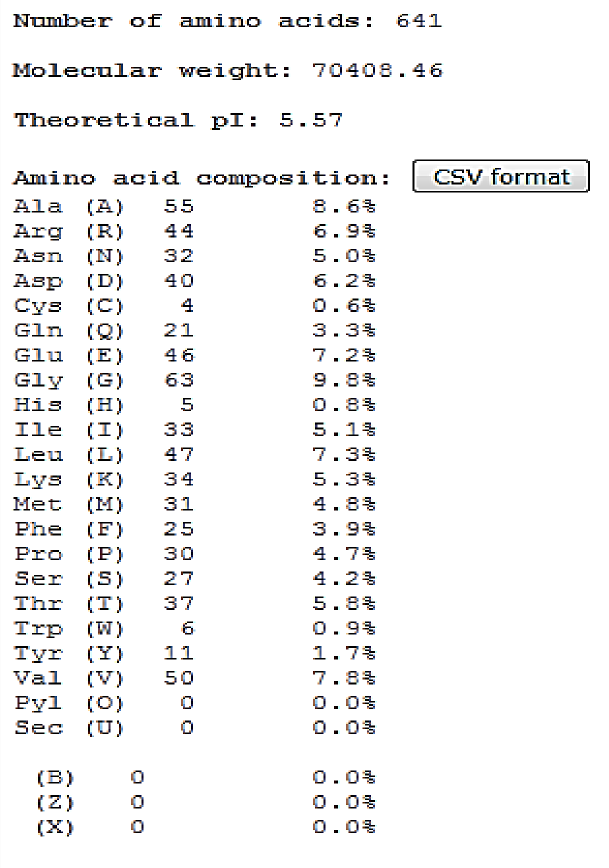


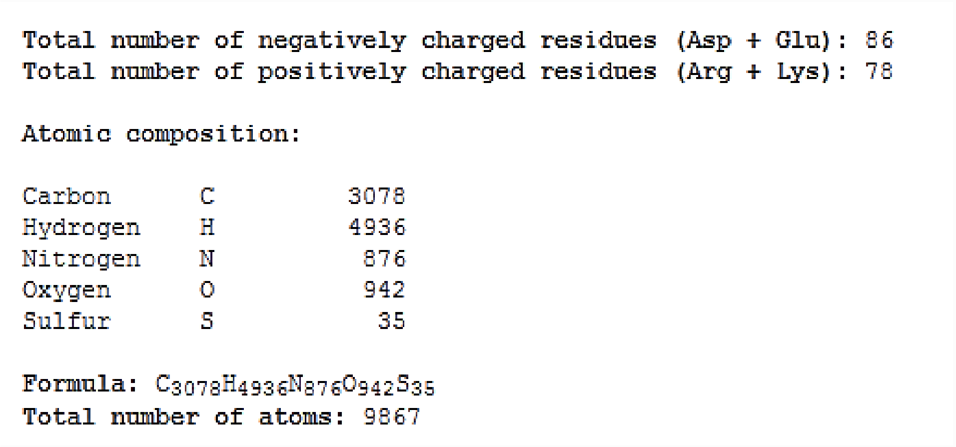
图4.5 FtsH蛋白系统发育树

4.2 FtsH蛋白理化性质分析、跨膜结构预测和亚细胞定位预测

根据FtsH氨基酸序列， 使用ExPASy（https://www.expasy.org/）中的Protpara（https://web.expasy.org/protparam/） 分析FtsH基本理化性质，得到序列所含各种氨基酸的种类和数目、分子量、等电点、酸性氨基酸数目、碱性氨基酸及C，H，O，N，S等原子的个数。



```
Number of amino acids: 641

Molecular weight: 70408.46

Theoretical pI: 5.57

Amino acid composition:
Ala (A)  55        8.6%
Arg (R)  44        6.9%
Asn (N)  32        5.0%
Asp (D)  40        6.2%
Cys (C)   4        0.6%
Gln (Q)  21        3.3%
Glu (E)  46        7.2%
Gly (G)  63        9.8%
His (H)   5        0.8%
Ile (I)  33        5.1%
Leu (L)  47        7.3%
Lys (K)  34        5.3%
Met (M)  31        4.8%
Phe (F)  25        3.9%
Pro (P)  30        4.7%
Ser (S)  27        4.2%
Thr (T)  37        5.8%
Trp (W)   6        0.9%
Tyr (Y)  11        1.7%
Val (V)  50        7.8%
Pyl (O)   0        0.0%
Sec (U)   0        0.0%

 (B)   0           0.0%
 (Z)   0           0.0%
 (X)   0           0.0%
```

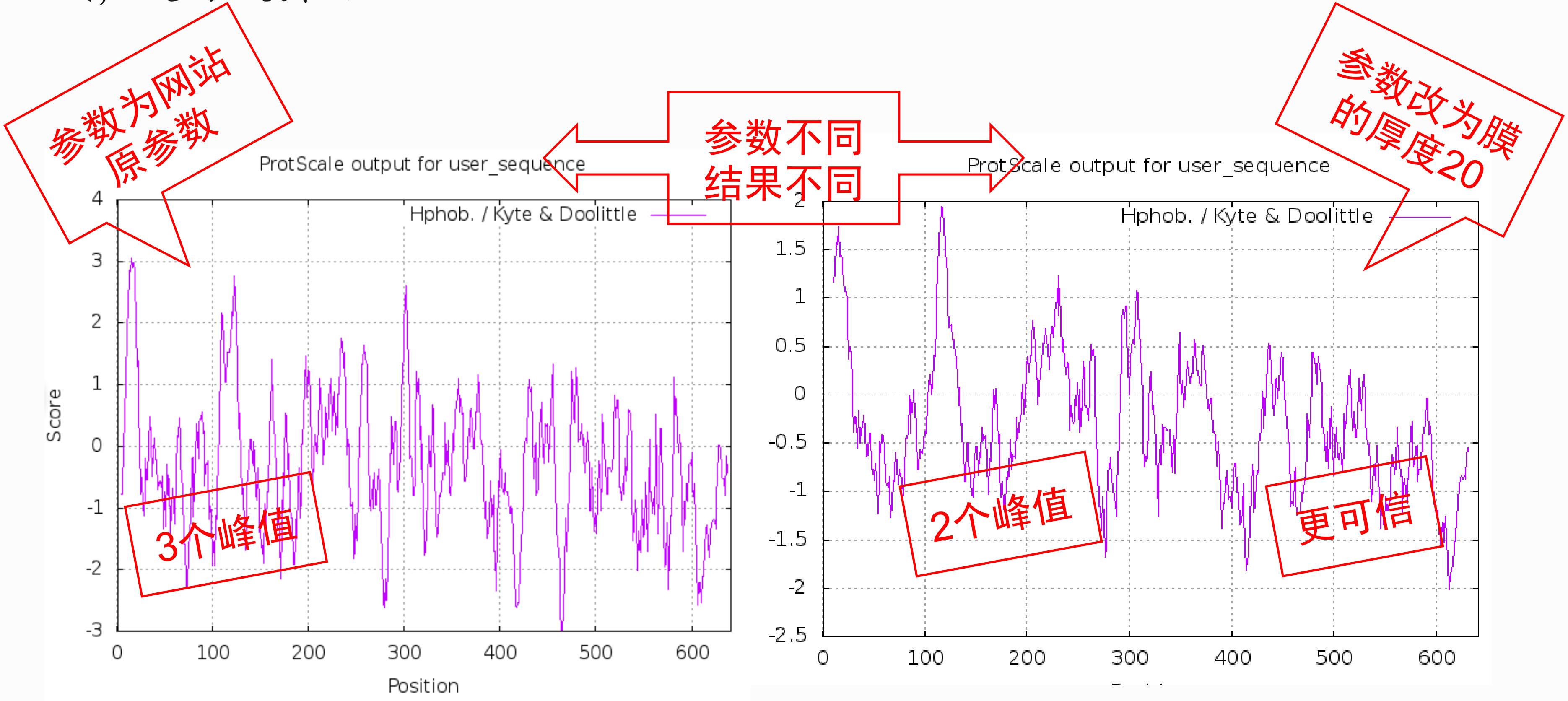


```
Total number of negatively charged residues (Asp + Glu): 86
Total number of positively charged residues (Arg + Lys): 78

Atomic composition:

Carbon      C        3078
Hydrogen    H        4936
Nitrogen    N         876
Oxygen      O         942
Sulfur      S          35

Formula: C3078H4936N876O942S35
Total number of atoms: 9867
```

图4.6 FtsH蛋白理化性质分析

蛋白结构及功能预测

根据氨基酸序列，使用ExPASy（https://www.expasy.org/）中的ProtScale（https://web.expasy.org/protscale/）预测FtsH蛋白亲疏水性和跨膜结构，根据生物膜的普遍厚度设定参数为20aa，得，FtsH蛋白可能含有2个跨膜结构域，是跨膜蛋白。



参数为网站原参数

参数不同

结果不同

参数改为膜的厚度20

3个峰值

2个峰值

更可信

图4.7 使用ProtScale预测分析FtsH蛋白亲疏水性和跨膜结构

蛋白结构及功能预测

根据氨基酸序列，使用TMHMM Server v. 2.0（http://www.cbs.dtu.dk/services/TMHMM/）跨膜结构域预测分析，FtsH蛋白含有2个跨膜结构域，是跨膜蛋白，这一结果与ProtScale的预测结果相一致 。

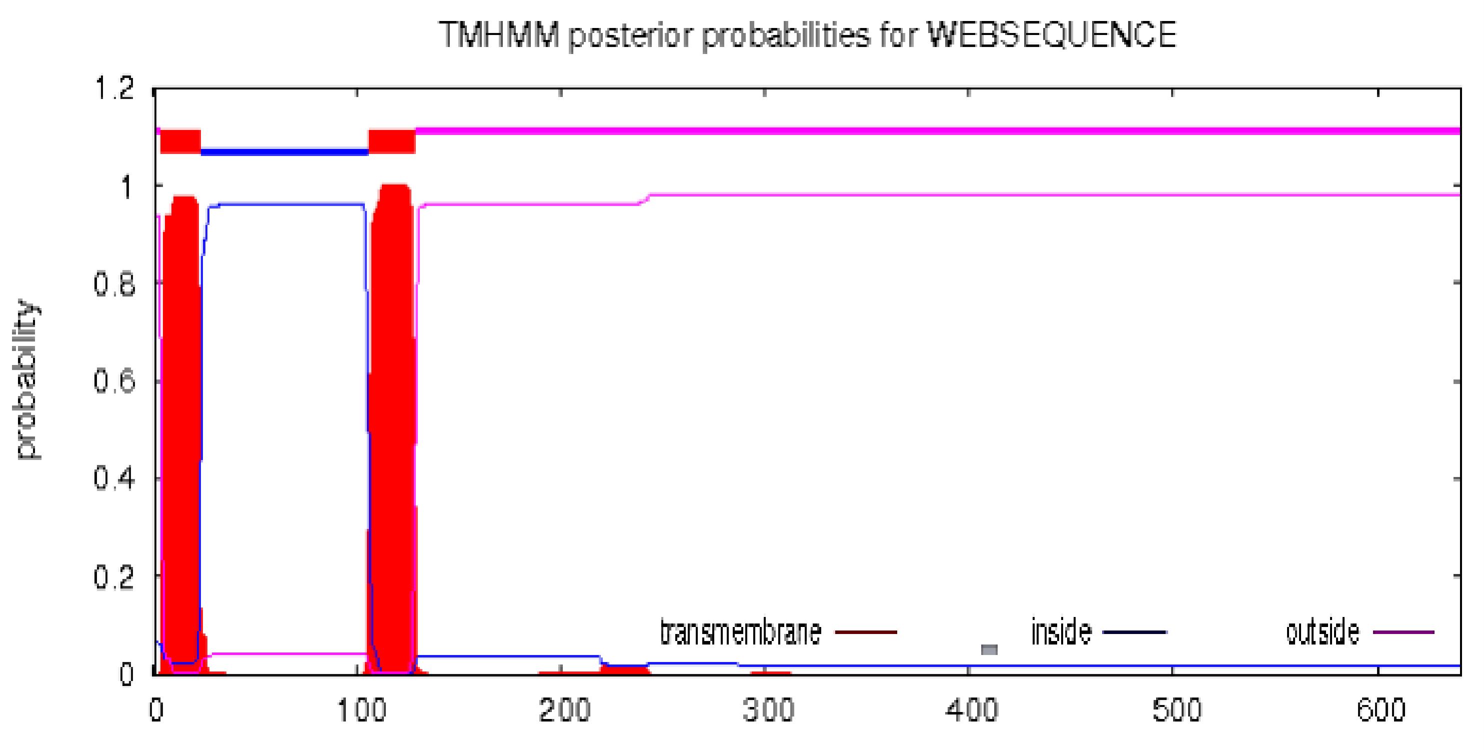


图4.8 使用TMHMM预测分析FtsH蛋白跨膜结构

蛋白结构及功能预测

在UniProt中，我们可以看到，FtsH蛋白的亚细胞定位为膜内、跨膜、膜外三个部分，存在2个跨膜结构域，是跨膜蛋白。这一结果与之前用ProtScalc蛋白亲疏水性预测和TMHMM Server v. 2.0预测跨膜结构域所得结果一致。在UniProt中，我们还可以看到它对氨基酸序列中跨膜位置的预测。

Subcellular location[i]

- Cell inner membrane ℹ UniRule annotation ▾; Multi-pass membrane protein ℹ UniRule annotation ▾; Cytoplasmic side ℹ UniRule annotation ▾

Topology

Feature key	Position(s)	Description	Actions	Graphical view	Length
Transmembrane[i]	6 – 23	Helical UniRule annotation ▾	Add BLAST		18
Transmembrane[i]	106 – 128	Helical UniRule annotation ▾	Add BLAST		23

GO - Cellular component[i]

- integral component of membrane Source: UniProtKB-KW
- plasma membrane Source: UniProtKB-SubCell

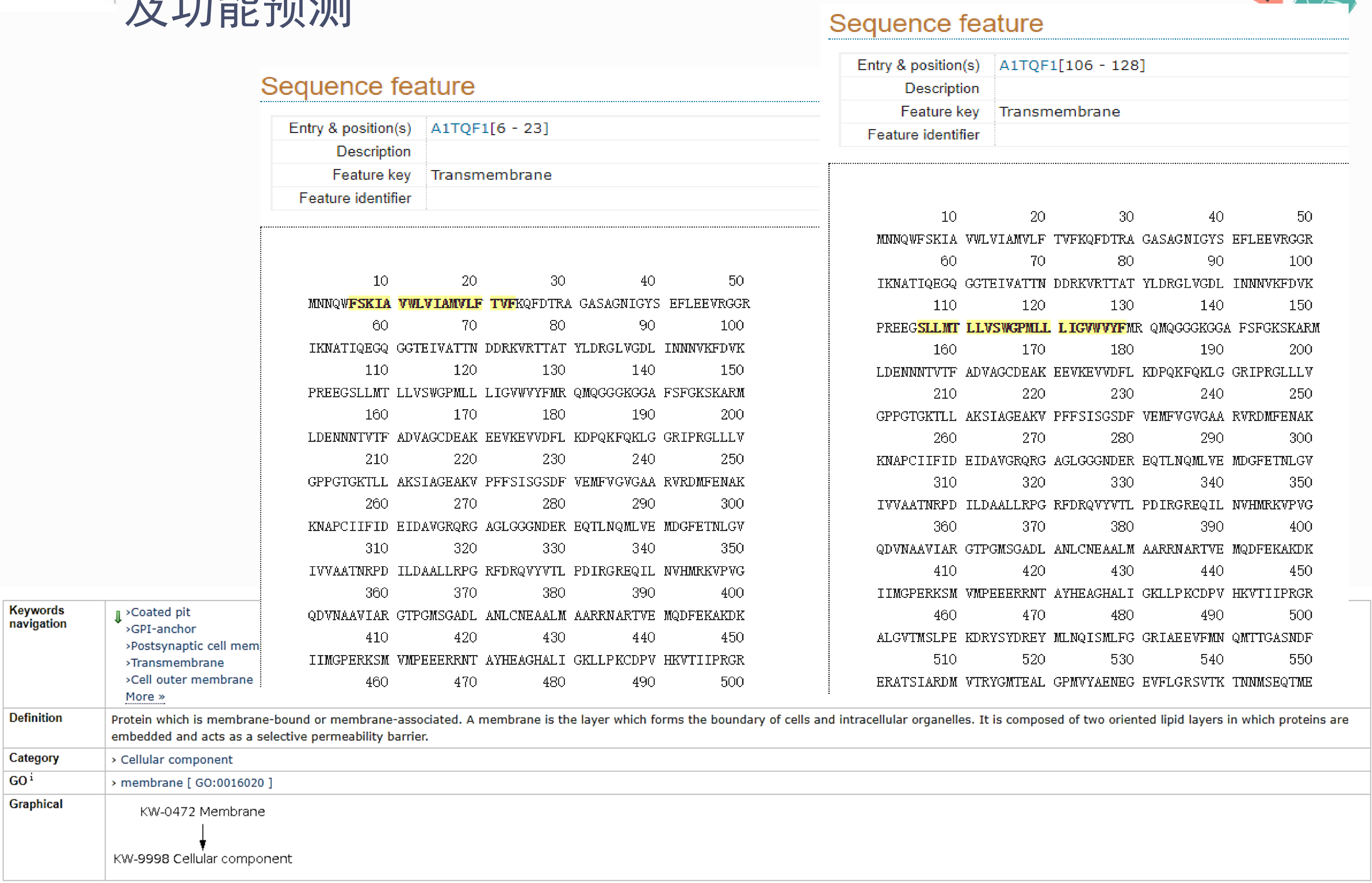
View the complete GO annotation on QuickGO ...

Keywords - Cellular component[i]

Cell inner membrane UniRule annotation ▾, Cell membrane, Membrane

图4.9 使用UniProt预测分析FtsH蛋白跨膜结构

蛋白结构及功能预测



Sequence feature

Entry & position(s)	A1TQF1[6 - 23]
Description	
Feature key	Transmembrane
Feature identifier	

10 20 30 40 50
MNNQW**FSKIA VWLVIAMVLF TVF**KQFDTRA GASAGNIGYS EFLEEVRGGR
60 70 80 90 100
IKNATIQEGQ GGTEIVATTN DDRKVRTTAT YLDRGLVGDL INNNVKFDVK
110 120 130 140 150
PREEGSLLMT LLVSWGPMLL LIGVWVYFMR QMQGGGKGGA FSFGKSKARM
160 170 180 190 200
LDENNNTVTF ADVAGCDEAK EEVKEVVDFL KDPQKFQKLG GRIPRGLLLV
210 220 230 240 250
GPPGTGKTLL AKSIAGEAKV PFFSISGSDF VEMFVGVGAA RVRDMFENAK
260 270 280 290 300
KNAPCIIFID EIDAVGRQRG AGLGGGNDER EQTLNQMLVE MDGFETNLGV
310 320 330 340 350
IVVAATNRPD ILDAALLRPG RFDRQVYVTL PDIRGREQIL NVHMRKVPVG
360 370 380 390 400
QDVNAAVIAR GTPGMSGADL ANLCNEAALM AARRNARTVE MQDFEKAKDK
410 420 430 440 450
IIMGPERKSM VMPEEERRNT AYHEAGHALI GKLLPKCDPV HKVTIIPRGR
460 470 480 490 500

Sequence feature

Entry & position(s)	A1TQF1[106 - 128]
Description	
Feature key	Transmembrane
Feature identifier	

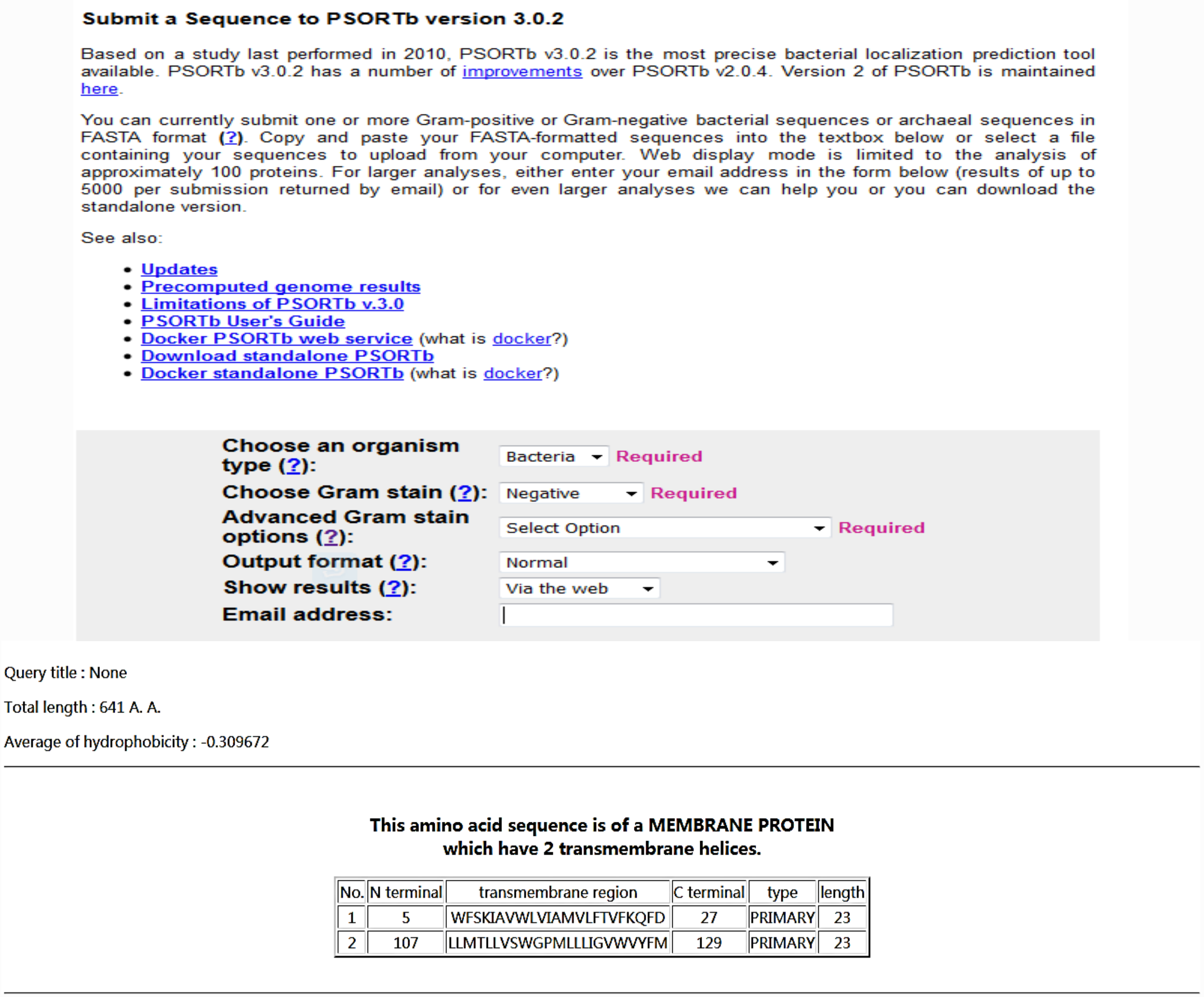
10 20 30 40 50
MNNQWFSKIA VWLVIAMVLF TVFKQFDTRA GASAGNIGYS EFLEEVRGGR
60 70 80 90 100
IKNATIQEGQ GGTEIVATTN DDRKVRTTAT YLDRGLVGDL INNNVKFDVK
110 120 130 140 150
PREEG**SLLMT LLVSWGPMLL LIGVWVYF**MR QMQGGGKGGA FSFGKSKARM
160 170 180 190 200
LDENNNTVTF ADVAGCDEAK EEVKEVVDFL KDPQKFQKLG GRIPRGLLLV
210 220 230 240 250
GPPGTGKTLL AKSIAGEAKV PFFSISGSDF VEMFVGVGAA RVRDMFENAK
260 270 280 290 300
KNAPCIIFID EIDAVGRQRG AGLGGGNDER EQTLNQMLVE MDGFETNLGV
310 320 330 340 350
IVVAATNRPD ILDAALLRPG RFDRQVYVTL PDIRGREQIL NVHMRKVPVG
360 370 380 390 400
QDVNAAVIAR GTPGMSGADL ANLCNEAALM AARRNARTVE MQDFEKAKDK
410 420 430 440 450
IIMGPERKSM VMPEEERRNT AYHEAGHALI GKLLPKCDPV HKVTIIPRGR
460 470 480 490 500
ALGVTMSLPE KDRYSYDREY MLNQISMLFG GRIAEEVFMN QMTTGASNDF
510 520 530 540 550
ERATSIARDM VTRYGMTEAL GPMVYAENEG EVFLGRSVTK TNNMSEQTME

Keywords navigation	↓ ›Coated pit ›GPI-anchor ›Postsynaptic cell mem ›Transmembrane ›Cell outer membrane More »
Definition	Protein which is membrane-bound or membrane-associated. A membrane is the layer which forms the boundary of cells and intracellular organelles. It is composed of two oriented lipid layers in which proteins are embedded and acts as a selective permeability barrier.
Category	› Cellular component
GO ⁱ	› membrane [GO:0016020]
Graphical	KW-0472 Membrane ↓ KW-9998 Cellular component

图4.10 使用UniProt预测分析FtsH蛋白跨膜结构

蛋白结构及功能预测

使用PSORT（http://www.psort.org/psortb/）预测蛋白跨膜结构，也能得到相同的结果。



Submit a Sequence to PSORTb version 3.0.2

Based on a study last performed in 2010, PSORTb v3.0.2 is the most precise bacterial localization prediction tool available. PSORTb v3.0.2 has a number of improvements over PSORTb v2.0.4. Version 2 of PSORTb is maintained here.

You can currently submit one or more Gram-positive or Gram-negative bacterial sequences or archaeal sequences in FASTA format (?). Copy and paste your FASTA-formatted sequences into the textbox below or select a file containing your sequences to upload from your computer. Web display mode is limited to the analysis of approximately 100 proteins. For larger analyses, either enter your email address in the form below (results of up to 5000 per submission returned by email) or for even larger analyses we can help you or you can download the standalone version.

See also:

- Updates
- Precomputed genome results
- Limitations of PSORTb v.3.0
- PSORTb User's Guide
- Docker PSORTb web service (what is docker?)
- Download standalone PSORTb
- Docker standalone PSORTb (what is docker?)

Choose an organism type (?):	Bacteria	Required
Choose Gram stain (?):	Negative	Required
Advanced Gram stain options (?):	Select Option	Required
Output format (?):	Normal	
Show results (?):	Via the web	
Email address:		

Query title : None

Total length : 641 A. A.

Average of hydrophobicity : -0.309672

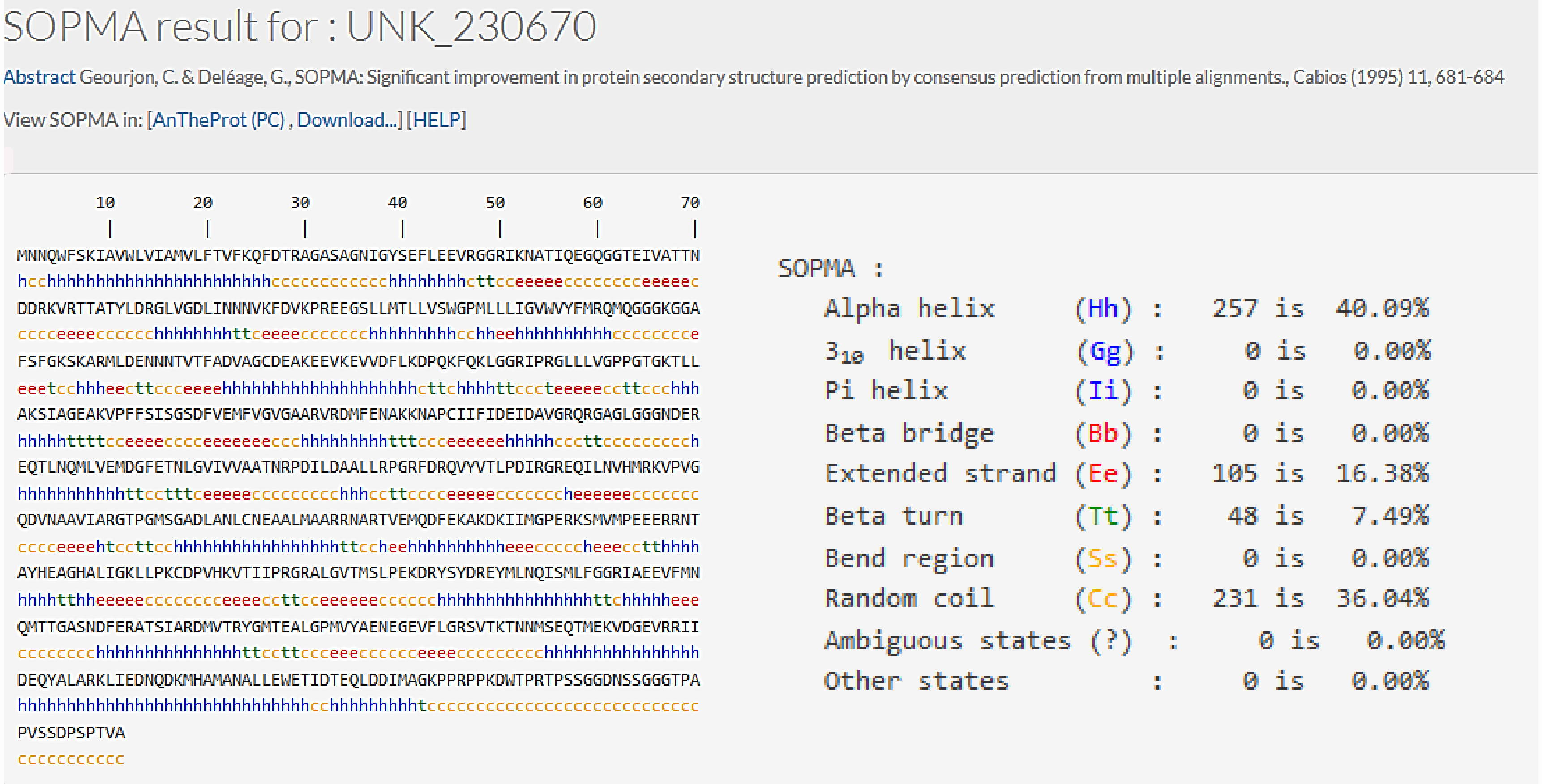
This amino acid sequence is of a MEMBRANE PROTEIN which have 2 transmembrane helices.

No.	N terminal	transmembrane region	C terminal	type	length
1	5	WFSKIAVWLVIAMVLFTVFKQFD	27	PRIMARY	23
2	107	LLMTLLVSWGPMLLLIGVWVYFM	129	PRIMARY	23

图4.11 FtsH蛋白亚细胞定位以及跨膜结构预测

4.3.1 FtsH蛋白二级结构预测

使用NPS@：SOPMA secondary structure prediction（https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html）分析预测FtsH蛋白二级结构，所得结果如下图。



SOPMA result for : UNK_230670

Abstract Geourjon, C. & Deléage, G., SOPMA: Significant improvement in protein secondary structure prediction by consensus prediction from multiple alignments., Cabios (1995) 11, 681-684

View SOPMA in: [AnTheProt (PC) , Download...] [HELP]

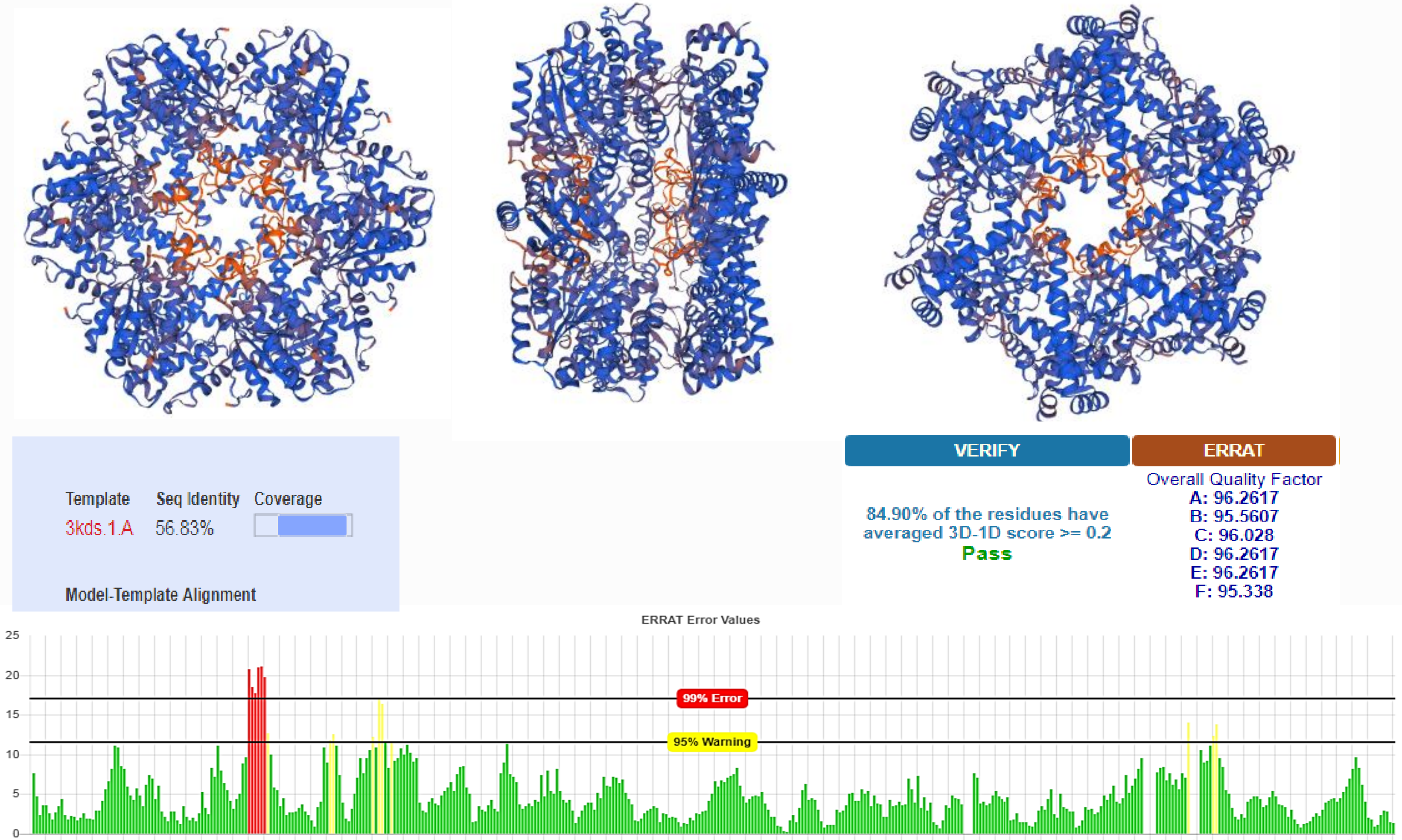
```
        10        20        30        40        50        60        70
         |         |         |         |         |         |         |
MNNQWFSKIAVWLVIAMVLFTVFKQFDTRAGASAGNIGYSEFLEEVRGGRIKNATIQEGQGGTEIVATTN
hcchhhhhhhhhhhhhhhhhhhhhhhcccccccccccchhhhhhhhcttcceeeeecccccccceeeeec
DDRKVRTTATYLDRGLVGDLINNNVKFDVKPREEGSLLMTLLVSWGPMLLLIGVWVYFMRQMQGGGKGGA
cccceeeecccccchhhhhhhhttceeeecccccccchhhhhhhhhcchheehhhhhhhhhhcccccccce
FSFGKSKARMLDENNNTVTFADVAGCDEAKEEVKEVVDFLKDPQKFQKLGGRIPRGLLLVGPPGTGKTLL
eeetcchhheecttccceeeehhhhhhhhhhhhhhhhhhhhcttchhhhttccctteeeeeccttccchhh
AKSIAGEAKVPFFSISGSDFVEMFVGVGAARVRDMFENAKKNAPCIIFIDEIDAVGRQRGAGLGGGNDER
hhhhhttttcceeeecccceeeeeeeccchhhhhhhhhtttccceeeeeehhhhhccctccccccccch
EQTLNQMLVEMDGFETNLGVIVVAATNRPDILDAALLRPGRFDRQVYVTLPDIRGREQILNVHMRKVPVG
hhhhhhhhhhhttcctttceeeeeccccccccchhhccttcccceeeeecccccccheeeeeecccccccc
QDVNAAVIARGTPGMSGADLANLCNEAALMAARRNARTVEMQDFEKAKDKIIMGPERKSMVMPEEERRNT
cccceeeehtccttcchhhhhhhhhhhhhhhhhttcchheehhhhhhhhhheeecccccheeeccttthhhh
AYHEAGHALIGKLLPKCDPVHKVTIIPRGRALGVTMSLPEKDRYSYDREYMLNQISMLFGGRIAEEVFMN
hhhhtthheeeeecccccccceeeeccttcceeeeeecccccchhhhhhhhhhhhhhhhttchhhhheee
QMTTGASNDFERATSIARDMVTRYGMTEALGPMVYAENEGEVFLGRSVTKTNNMSEQTMEKVDGEVRRII
cccccccchhhhhhhhhhhhhhhttccttccceeecccccceeeecccccccccchhhhhhhhhhhhhhhh
DEQYALARKLIEDNQDKMHAMANALLEWETIDTEQLDDIMAGKPPRPPKDWTPRTPSSGGDNSSGGGTPA
hhhhhhhhhhhhhhhhhhhhhhhhhhhhhhcchhhhhhhhhhtcccccccccccccccccccccccccccc
PVSSDPSPTVA
ccccccccccc
```

```
SOPMA :
   Alpha helix      (Hh) :   257 is  40.09%
   3$_{10}$ helix         (Gg) :     0 is   0.00%
   Pi helix         (Ii) :     0 is   0.00%
   Beta bridge      (Bb) :     0 is   0.00%
   Extended strand  (Ee) :   105 is  16.38%
   Beta turn        (Tt) :    48 is   7.49%
   Bend region      (Ss) :     0 is   0.00%
   Random coil      (Cc) :   231 is  36.04%
   Ambiguous states (?)  :     0 is   0.00%
   Other states          :     0 is   0.00%
```

图4.12 FtsH蛋白二级结构分析

蛋白结构及功能预测

4.3.2 FtsH蛋白三级结构预测（Swiss-model（https://www.swissmodel.expasy.org/））



Template Seq Identity Coverage

3kds.1.A 56.83%

Model-Template Alignment

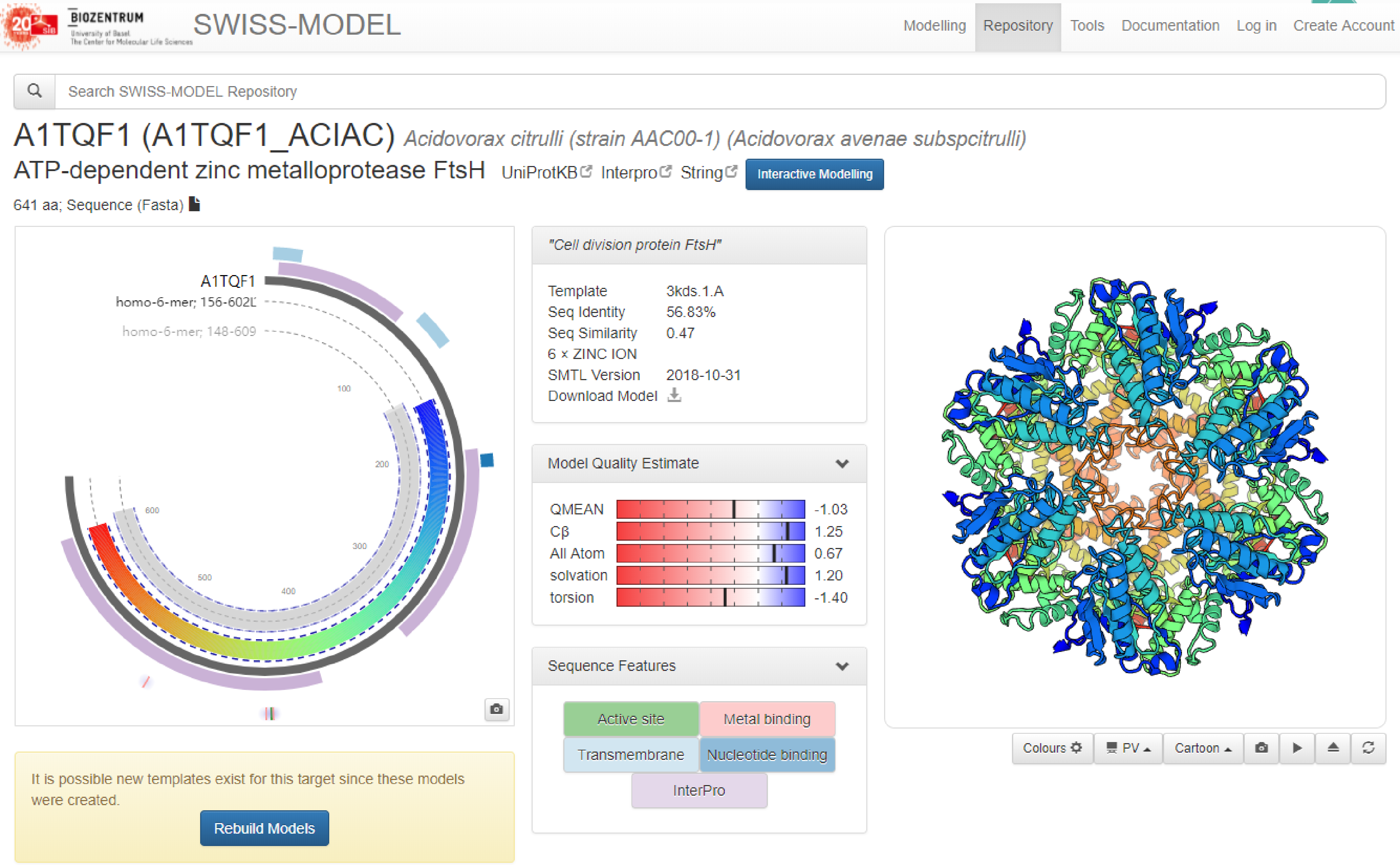
VERIFY	ERRAT
84.90% of the residues have averaged 3D-1D score >= 0.2 Pass	Overall Quality Factor A: 96.2617 B: 95.5607 C: 96.028 D: 96.2617 E: 96.2617 F: 95.338

图4.13 FtsH蛋白三级结构分析

通过VERIFY3D评估所得3D模型，所得评分为84.90%；评估结果显示，模型中Errat部分（图中红色部分）较少，因此所建模型较为可信的。

蛋白结构及功能预测

https://swissmodel.expasy.org/repository/md5/069064be16671fca430dbe10d0372cc4



SWISS-MODEL

Modelling | Repository | Tools | Documentation | Log in | Create Account

Search SWISS-MODEL Repository

A1TQF1 (A1TQF1_ACIAC) *Acidovorax citrulli (strain AAC00-1) (Acidovorax avenae subspcitrulli)*

ATP-dependent zinc metalloprotease FtsH UniProtKB Interpro String Interactive Modelling

641 aa; Sequence (Fasta)

"Cell division protein FtsH"

Template	3kds.1.A
Seq Identity	56.83%
Seq Similarity	0.47
6 × ZINC ION	
SMTL Version	2018-10-31
Download Model	

Model Quality Estimate

QMEAN	-1.03
Cβ	1.25
All Atom	0.67
solvation	1.20
torsion	-1.40

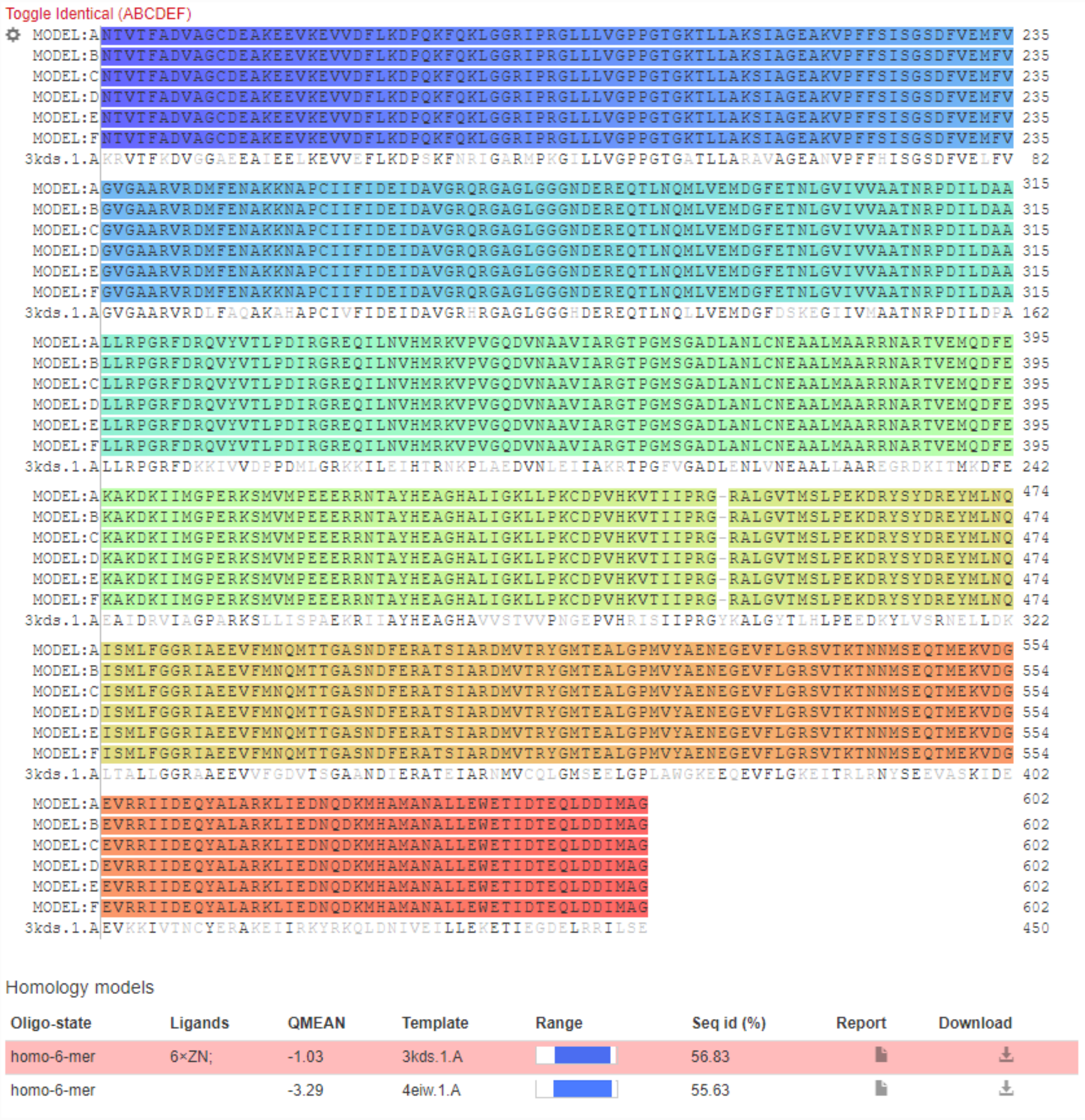
Sequence Features

It is possible new templates exist for this target since these models were created.

图4.14 FtsH蛋白三级结构分析

蛋白结构及功能预测

https://swissmodel.expasy.org/repository/md5/069064be16671fca430dbe10d0372cc4

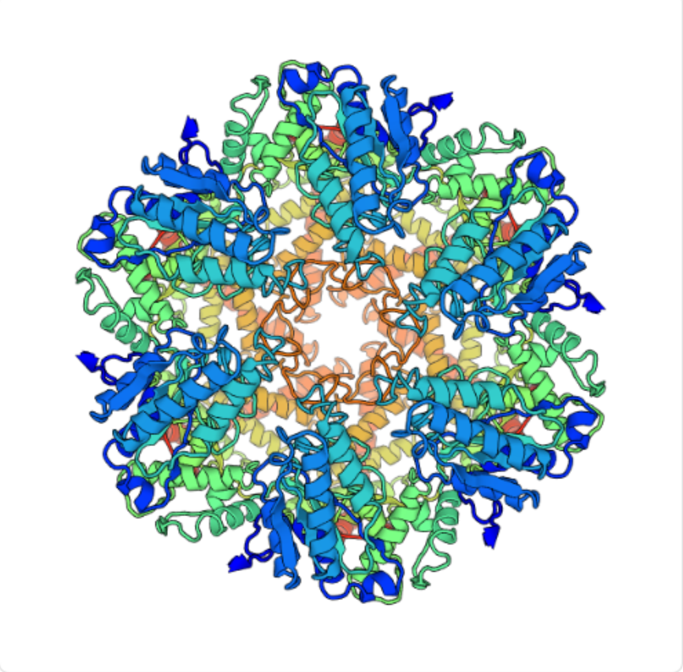


Toggle Identical (ABCDEF)

```
MODEL:A NTVTFADVAGCDEAKEEVKEVVDFLKDPQKFQKLGGRIPRGLLLVGPPGTGKTLLAKSIAGEAKVPFFSISGSDFVEMFV 235
MODEL:B NTVTFADVAGCDEAKEEVKEVVDFLKDPQKFQKLGGRIPRGLLLVGPPGTGKTLLAKSIAGEAKVPFFSISGSDFVEMFV 235
MODEL:C NTVTFADVAGCDEAKEEVKEVVDFLKDPQKFQKLGGRIPRGLLLVGPPGTGKTLLAKSIAGEAKVPFFSISGSDFVEMFV 235
MODEL:D NTVTFADVAGCDEAKEEVKEVVDFLKDPQKFQKLGGRIPRGLLLVGPPGTGKTLLAKSIAGEAKVPFFSISGSDFVEMFV 235
MODEL:E NTVTFADVAGCDEAKEEVKEVVDFLKDPQKFQKLGGRIPRGLLLVGPPGTGKTLLAKSIAGEAKVPFFSISGSDFVEMFV 235
MODEL:F NTVTFADVAGCDEAKEEVKEVVDFLKDPQKFQKLGGRIPRGLLLVGPPGTGKTLLAKSIAGEAKVPFFSISGSDFVEMFV 235
3kds.1.A KRVTFKDVGGAEEAIEELKEVVEFLKDPSKFNRIGARMPKGILLVGPPGTGATLLARAVAGEANVPFFHISGSDFVELFV  82

MODEL:A GVGAARVRDMFENAKKNAPCIIFIDEIDAVGRQRGAGLGGGNDEREQTLNQMLVEMDGFETNLGVIVVAATNRPDILDAA 315
MODEL:B GVGAARVRDMFENAKKNAPCIIFIDEIDAVGRQRGAGLGGGNDEREQTLNQMLVEMDGFETNLGVIVVAATNRPDILDAA 315
MODEL:C GVGAARVRDMFENAKKNAPCIIFIDEIDAVGRQRGAGLGGGNDEREQTLNQMLVEMDGFETNLGVIVVAATNRPDILDAA 315
MODEL:D GVGAARVRDMFENAKKNAPCIIFIDEIDAVGRQRGAGLGGGNDEREQTLNQMLVEMDGFETNLGVIVVAATNRPDILDAA 315
MODEL:E GVGAARVRDMFENAKKNAPCIIFIDEIDAVGRQRGAGLGGGNDEREQTLNQMLVEMDGFETNLGVIVVAATNRPDILDAA 315
MODEL:F GVGAARVRDMFENAKKNAPCIIFIDEIDAVGRQRGAGLGGGNDEREQTLNQMLVEMDGFETNLGVIVVAATNRPDILDAA 315
3kds.1.A GVGAARVRDLFAQAKAHAPCIVFIDEIDAVGRHRGAGLGGGHDEREQTLNQLLVEMDGFDSKEGIIVMAATNRPDILDPA 162

MODEL:A LLRPGRFDRQVYVTLPDIRGREQILNVHMRKVPVGQDVNAAVIARGTPGMSGADLANLCNEAALMAARRNARTVEMQDFE 395
MODEL:B LLRPGRFDRQVYVTLPDIRGREQILNVHMRKVPVGQDVNAAVIARGTPGMSGADLANLCNEAALMAARRNARTVEMQDFE 395
MODEL:C LLRPGRFDRQVYVTLPDIRGREQILNVHMRKVPVGQDVNAAVIARGTPGMSGADLANLCNEAALMAARRNARTVEMQDFE 395
MODEL:D LLRPGRFDRQVYVTLPDIRGREQILNVHMRKVPVGQDVNAAVIARGTPGMSGADLANLCNEAALMAARRNARTVEMQDFE 395
MODEL:E LLRPGRFDRQVYVTLPDIRGREQILNVHMRKVPVGQDVNAAVIARGTPGMSGADLANLCNEAALMAARRNARTVEMQDFE 395
MODEL:F LLRPGRFDRQVYVTLPDIRGREQILNVHMRKVPVGQDVNAAVIARGTPGMSGADLANLCNEAALMAARRNARTVEMQDFE 395
3kds.1.A LLRPGRFDKKIVVDPPDMLGRKKILEIHTRNKPLAEDVNLEIIAKRTPGFVGADLENLVNEAALLAAREGRDKITMKDFE 242

MODEL:A KAKDKIIMGPERKSMVMPEEERRNTAYHEAGHALIGKLLPKCDPVHKVTIIPRG-RALGVTMSLPEKDRYSYDREYMLNQ 474
MODEL:B KAKDKIIMGPERKSMVMPEEERRNTAYHEAGHALIGKLLPKCDPVHKVTIIPRG-RALGVTMSLPEKDRYSYDREYMLNQ 474
MODEL:C KAKDKIIMGPERKSMVMPEEERRNTAYHEAGHALIGKLLPKCDPVHKVTIIPRG-RALGVTMSLPEKDRYSYDREYMLNQ 474
MODEL:D KAKDKIIMGPERKSMVMPEEERRNTAYHEAGHALIGKLLPKCDPVHKVTIIPRG-RALGVTMSLPEKDRYSYDREYMLNQ 474
MODEL:E KAKDKIIMGPERKSMVMPEEERRNTAYHEAGHALIGKLLPKCDPVHKVTIIPRG-RALGVTMSLPEKDRYSYDREYMLNQ 474
MODEL:F KAKDKIIMGPERKSMVMPEEERRNTAYHEAGHALIGKLLPKCDPVHKVTIIPRG-RALGVTMSLPEKDRYSYDREYMLNQ 474
3kds.1.A EAIDRVIAGPARKSLLISPAEKRIIAYHEAGHAVVSTVVPNGEPVHRISIIPRGYKALGYTLHLPEEDKYLVSRNELLDK 322

MODEL:A ISMLFGGRIAEEVFMNQMTTGASNDFERATSIARDMVTRYGMTEALGPMVYAENEGEVFLGRSVTKTNNMSEQTMEKVDG 554
MODEL:B ISMLFGGRIAEEVFMNQMTTGASNDFERATSIARDMVTRYGMTEALGPMVYAENEGEVFLGRSVTKTNNMSEQTMEKVDG 554
MODEL:C ISMLFGGRIAEEVFMNQMTTGASNDFERATSIARDMVTRYGMTEALGPMVYAENEGEVFLGRSVTKTNNMSEQTMEKVDG 554
MODEL:D ISMLFGGRIAEEVFMNQMTTGASNDFERATSIARDMVTRYGMTEALGPMVYAENEGEVFLGRSVTKTNNMSEQTMEKVDG 554
MODEL:E ISMLFGGRIAEEVFMNQMTTGASNDFERATSIARDMVTRYGMTEALGPMVYAENEGEVFLGRSVTKTNNMSEQTMEKVDG 554
MODEL:F ISMLFGGRIAEEVFMNQMTTGASNDFERATSIARDMVTRYGMTEALGPMVYAENEGEVFLGRSVTKTNNMSEQTMEKVDG 554
3kds.1.A LTALLGGRAAEEVVFGDVTSGAANDIERATEIARNMVCQLGMSEELGPLAWGKEEEQEVFLGKEITRLRNYSEEVASKIDE 402

MODEL:A EVRRIIDEQYALARKLIEDNQDKMHAMANALLEWETIDTEQLDDIMAG                                 602
MODEL:B EVRRIIDEQYALARKLIEDNQDKMHAMANALLEWETIDTEQLDDIMAG                                 602
MODEL:C EVRRIIDEQYALARKLIEDNQDKMHAMANALLEWETIDTEQLDDIMAG                                 602
MODEL:D EVRRIIDEQYALARKLIEDNQDKMHAMANALLEWETIDTEQLDDIMAG                                 602
MODEL:E EVRRIIDEQYALARKLIEDNQDKMHAMANALLEWETIDTEQLDDIMAG                                 602
MODEL:F EVRRIIDEQYALARKLIEDNQDKMHAMANALLEWETIDTEQLDDIMAG                                 602
3kds.1.A EVKKIVTNCYERAKEIIRKYRKQLDNIVEILLEKETIEGDELRRILSE                                 450
```

Homology models

Oligo-state	Ligands	QMEAN	Template	Range	Seq id (%)	Report	Download
homo-6-mer	6×ZN;	-1.03	3kds.1.A		56.83		
homo-6-mer		-3.29	4eiw.1.A		55.63		



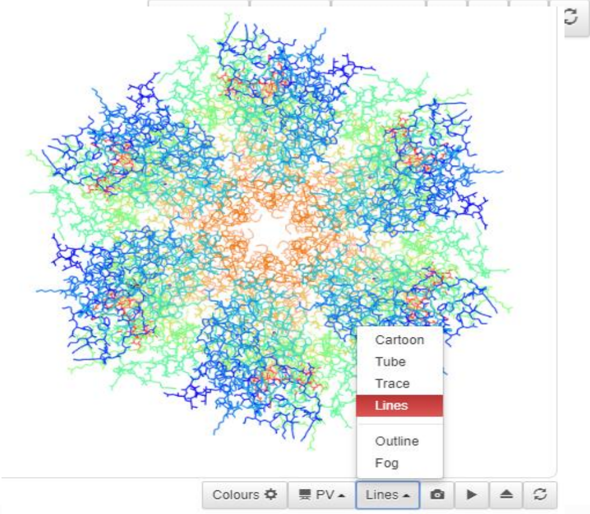


图4.15 FtsH蛋白三级结构分析

蛋白结构及功能预测

4.3.3 FtsH蛋白结构域预测

经SMART（http://smart.embl-heidelberg.de/smart/set_mode.cgi?GENOMIC=1）蛋白结构预测分析，FtsH蛋白由FtsH_ext域、AAA域、肽酶M41域和跨膜结构构成。

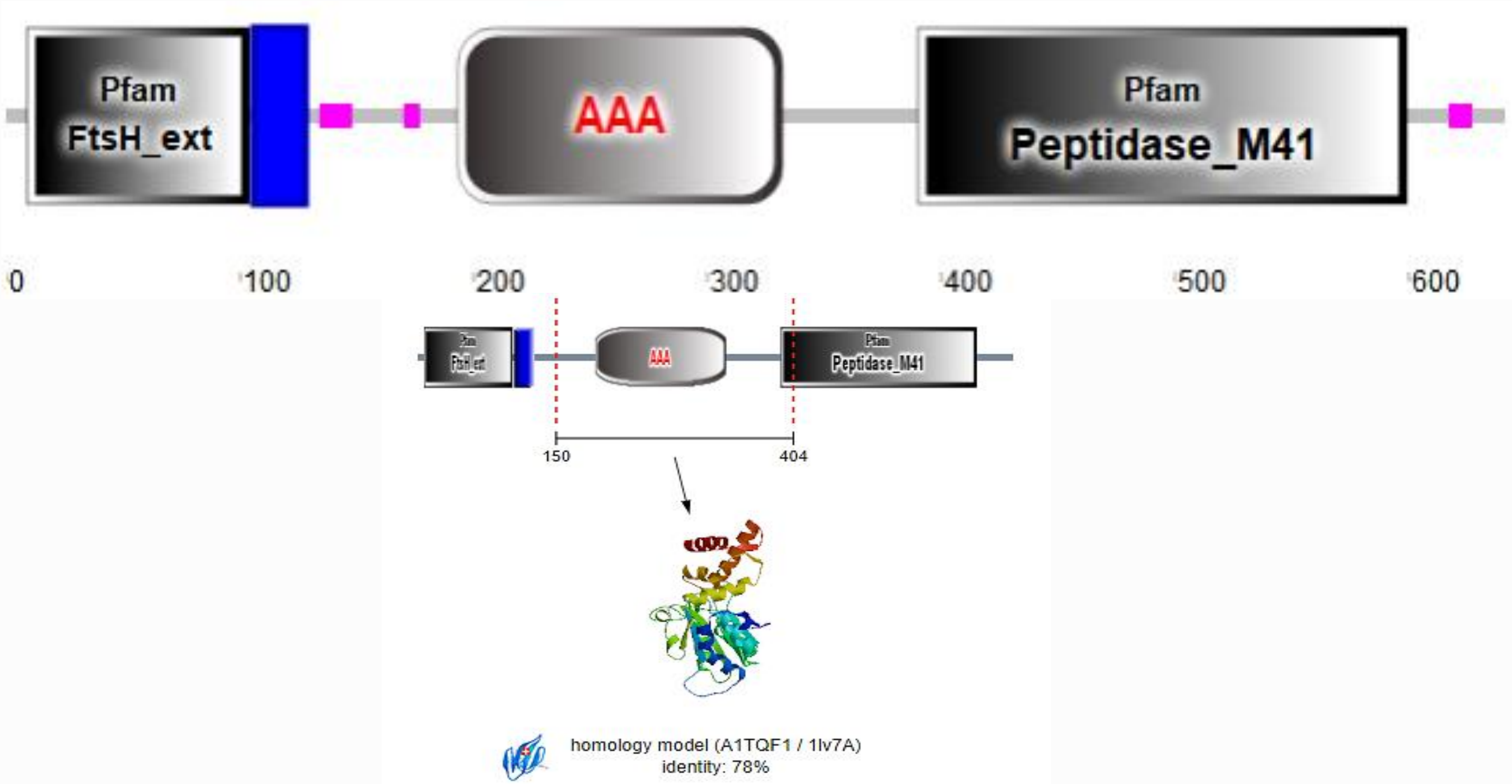
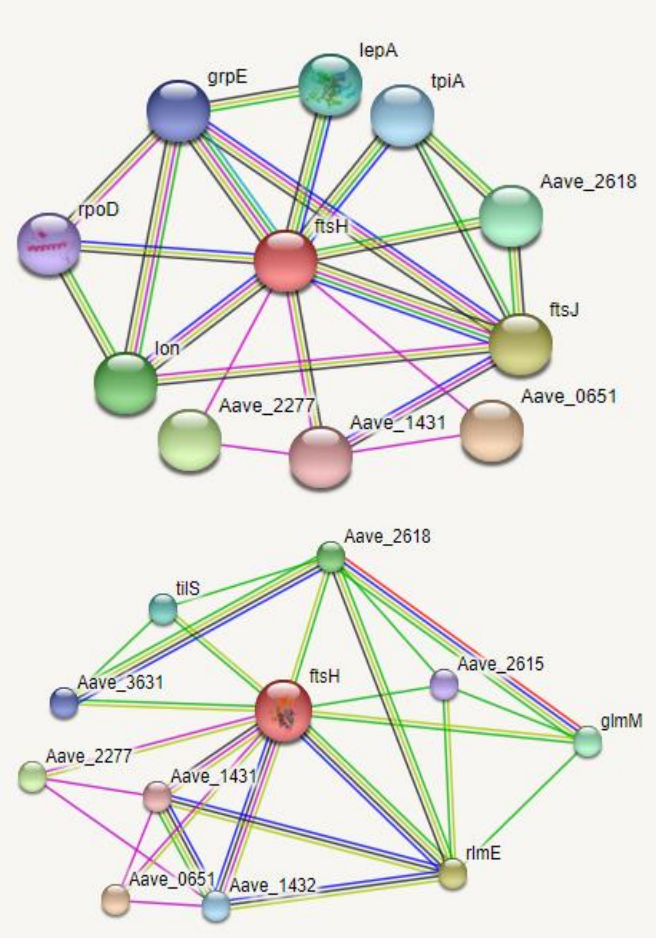


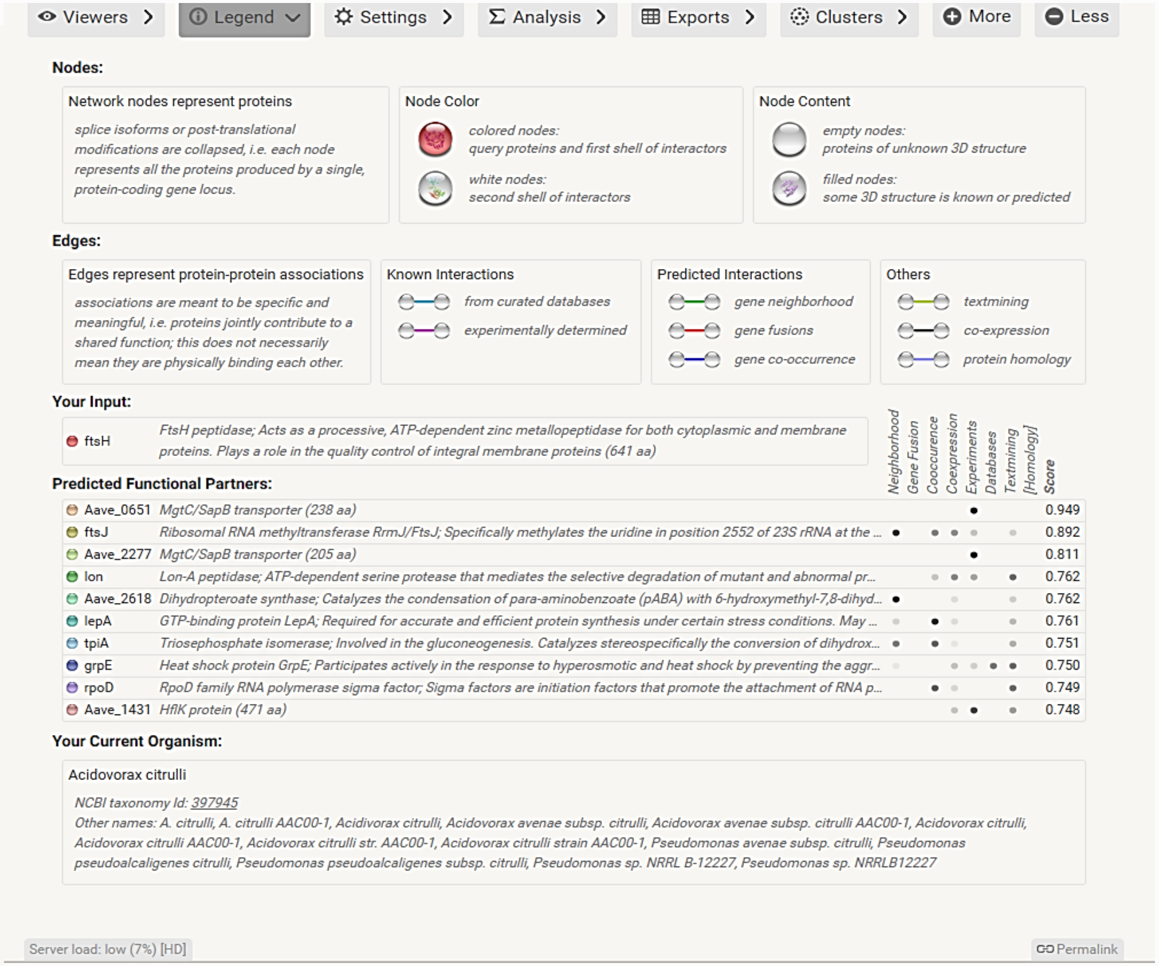
图4.16 FtsH蛋白结构域分析

蛋白结构及功能预测

4.4 基因互作关系预测及功能预测

在STRING分析预测中，与*ftsH*基因调控关系密切的基因有*hflK(Aave_1431)*、*hflC(Aave_1432)*、*Aave_2615*、*Aave_0651*、*Aave_2618*、*Aave_3631*、*Aave_2277*等。





Nodes:

Network nodes represent proteins

splice isoforms or post-translational modifications are collapsed, i.e. each node represents all the proteins produced by a single, protein-coding gene locus.

Node Color

colored nodes: query proteins and first shell of interactors

white nodes: second shell of interactors

Node Content

empty nodes: proteins of unknown 3D structure

filled nodes: some 3D structure is known or predicted

Edges:

Edges represent protein-protein associations

associations are meant to be specific and meaningful, i.e. proteins jointly contribute to a shared function; this does not necessarily mean they are physically binding each other.

Known Interactions

- *from curated databases*
- *experimentally determined*

Predicted Interactions

- *gene neighborhood*
- *gene fusions*
- *gene co-occurrence*

Others

- *textmining*
- *co-expression*
- *protein homology*

Your Input:

ftsH	*FtsH peptidase; Acts as a processive, ATP-dependent zinc metallopeptidase for both cytoplasmic and membrane proteins. Plays a role in the quality control of integral membrane proteins (641 aa)*

Predicted Functional Partners:

		Neighborhood	Gene Fusion	Cooccurence	Coexpression	Experiments	Databases	Textmining	[Homology]	**Score**
Aave_0651	*MgtC/SapB transporter (238 aa)*					●				0.949
ftsJ	*Ribosomal RNA methyltransferase RrmJ/FtsJ; Specifically methylates the uridine in position 2552 of 23S rRNA at the ...*	●		●	●	●		●		0.892
Aave_2277	*MgtC/SapB transporter (205 aa)*					●				0.811
lon	*Lon-A peptidase; ATP-dependent serine protease that mediates the selective degradation of mutant and abnormal pr...*			●	●	●		●		0.762
Aave_2618	*Dihydropteroate synthase; Catalyzes the condensation of para-aminobenzoate (pABA) with 6-hydroxymethyl-7,8-dihyd...*	●			●			●		0.762
lepA	*GTP-binding protein LepA; Required for accurate and efficient protein synthesis under certain stress conditions. May ...*	●		●	●			●		0.761
tpiA	*Triosephosphate isomerase; Involved in the gluconeogenesis. Catalyzes stereospecifically the conversion of dihydrox...*	●		●	●			●		0.751
grpE	*Heat shock protein GrpE; Participates actively in the response to hyperosmotic and heat shock by preventing the aggr...*	●			●	●	●	●		0.750
rpoD	*RpoD family RNA polymerase sigma factor; Sigma factors are initiation factors that promote the attachment of RNA p...*			●	●			●		0.749
Aave_1431	*HflK protein (471 aa)*				●	●		●		0.748

Your Current Organism:

Acidovorax citrulli

NCBI taxonomy Id: 397945

Other names: A. citrulli, A. citrulli AAC00-1, Acidivorax citrulli, Acidovorax avenae subsp. citrulli, Acidovorax avenae subsp. citrulli AAC00-1, Acidovorax citrulli, Acidovorax citrulli AAC00-1, Acidovorax citrulli str. AAC00-1, Acidovorax citrulli strain AAC00-1, Pseudomonas avenae subsp. citrulli, Pseudomonas pseudoalcaligenes citrulli, Pseudomonas pseudoalcaligenes subsp. citrulli, Pseudomonas sp. NRRL B-12227, Pseudomonas sp. NRRLB12227

Server load: low (7%) [HD]

图4.17 *ftsH*互作关系分析

小 结

通过对西瓜噬酸菌中*ftsH*基因的产物蛋白FtsH进行一系列的分析发现：

1. 西瓜噬酸菌FtsH蛋白与同噬酸菌属的细菌的FtsH蛋白亲缘关系较近，与大肠杆菌和青枯菌中FtsH蛋白的亲缘关系较远。西瓜噬酸菌FtsH蛋白含有2个跨膜结构，极有可能是跨膜蛋白，蛋白带负电，弱酸性。
2. 西瓜噬酸菌中的FtsH蛋白的二级结构由α螺旋、β折叠、无规卷曲和延伸链构成，三级结构是同源六聚体，内部含有环状结构，在空间上，包含FtsH_ext域、AAA域、肽酶M41域和跨膜结构，基于AAA结构域的保守性（AAA+蛋白酶均含有ATP酶域和蛋白酶域，ATP酶域催化底物的展开，并将其转移给蛋白酶体，其中包含一个酪氨酸与组氨酸的氨基酸交换过程。蛋白酶体将展开的底物分解成5-25个氨基酸的小肽段。），三级结构中的环状结构可能是底物到达蛋白酶域进行水解的中心孔。
3. 由于含有AAA结构，FtsH蛋白可能通过行使ATP酶和水解酶功能参与细菌细胞的生长发育、膜内在蛋白的质量平衡等过程。肽酶M41域的存在可能使FtsH蛋白与氨基糖苷类抗生素抗性相关。
4. 通过互作分析，西瓜噬酸菌中*ftsH*基因可能与西瓜噬酸菌的盐离子转运、热激反应以及胞外多糖的形成有关，可能影响菌株的生长、致病和生物膜形成等能力。

致 谢

感谢G06组内每一位同学的努力和付出。

很高兴能有机会与大家共同学习。

感谢罗老师，您的每一堂精彩课程都让我们受益匪浅，您的严谨认真也感染和激励着我们，让我们更严格的要求自己，请允许我代表我的组员向您表示真挚的敬意和感谢！

THANKS
敬请老师和同学们
批评指正