

UBIQUITIN AND UBIQUITIN-LIKE PROTEINS

By G16

Ubiquitin of Human

Human中ubiquitin 的四种编码基因

- **UBA52:** 编码结合在核糖体蛋白L40上的单个泛素。
- **RPS27A:** 编码结合在核糖体蛋白S27a上的单个泛素。
- **UBB:** 编码多泛素的先导蛋白。
- **UBC:** 编码多泛素的先导蛋白。

UBA52

- Protein name: Ubiquitin-60S ribosomal protein L40.
- Sequence length: 128AA
- **Cleaved into the following 2 chains:**
 - 1-76:** Ubiquitin
 - 77-128:** 60S ribosomal protein L40

```
    10      20      30      40      50      60
MQIFVKLTG KTITLEVEPS DTIENVKAKI QDKEGIPPDQ QRLIFAGKQL EDGRTLSDYN
    70      80      90     100     110     120
IQKESTLHLV LRLRGGIIEP SLRQLAQKYN CDKMICRKCY ARLHPRAVNC RKKKCGHTNN

LRPKKKVK
```

RPS27A

- Protein name: Ubiquitin-40S ribosomal protein S27a
- Sequence length: 156AA
- Cleaved into the following 2 chains:
 - 1-76: Ubiquitin
 - 77-156:** 40S ribosomal protein S27a

```
      10      20      30      40      50      60
MQIFVKTLTG KTITLEVEPS DTIENVKAKI QDKEGIPPDQ QRLIFAGKQL EDGRTLSDYN
      70      80      90     100     110     120
IQKESTLHLV LRLRGGAKKR KKKSYTTPKK NKHKRKKVKL AVLKYYKVDE NGKISRLRRE
     130     140     150
CPSDECGAGV FMASHFDRHY CGKCCLTYCF NKPEDK
```

UBB

- Protein name: Polyubiquitin-B
- Sequence length: 229AA
- Cleaved into the following chain:
 - 1-76: Ubiquitin
 - 77-152: Ubiquitin
 - 153-228: Ubiquitin
 - 229: Propeptide

```
      10      20      30      40      50      60
MQIFVKTLTG KTITLEVEPS DTIENVKAKI QDKEGIPPDQ QRLIFAGKQL EDGRTLSDYN

      70      80      90     100     110     120
IQKESTLHLV LRLRGGMQIF VKTLTGKTIT LEVEPSDTIE NVKAKIQDKE GIPPDQQRLI

     130     140     150     160     170     180
FAGKQLEDGR TLDYNIQKE STLHLVLR LR GGMQIFVKTL TGKTITLEVE PSDTIENVKA

     190     200     210     220
KIQDKEGIPP DQQRLIFAGK QLEDGRTLSD YNIQKESTLH LVLRLRGGC
```

UBC

- Protein name: Polyubiquitin-C
- Sequence length: 685AA
- Cleaved into the following chain:

1-76, 77-152, 153-228, 229-304, 305-380,
381-456, 457-532, 533-608, 609-684: Ubiquitin
685: Propeptide

```
>sp|P0CG48|UBC_HUMAN Polyubiquitin-C OS=Homo sapiens GN=UBC PE=1 SV=3
MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYN
IQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLI
FAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKA
KIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKT
ITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLR
LRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTL
SDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQ
QRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTLTGKTITLEVEPSDTIE
NVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLHLVLRLRGGMQIFVKTL
TGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDYNIQKESTLH
LVLRLRGGMQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLED
GRTLSDYNIQKESTLHLVLRLRGGV
```

Alignment

```
1  MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDY 60  P62987  RL40_HUMAN
1  MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDY 60  P62979  RS27A_HUMAN
1  MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDY 60  P0CG47  UBB_HUMAN
1  MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDY 60  P0CG47  UBB_HUMAN
1  MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDY 60  P0CG47  UBB_HUMAN
1  MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDY 60  P0CG48  UBC_HUMAN
1  MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDY 60  P0CG48  UBC_HUMAN
1  MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDY 60  P0CG48  UBC_HUMAN
1  MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDY 60  P0CG48  UBC_HUMAN
1  MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDY 60  P0CG48  UBC_HUMAN
1  MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDY 60  P0CG48  UBC_HUMAN
1  MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDY 60  P0CG48  UBC_HUMAN
1  MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDY 60  P0CG48  UBC_HUMAN
1  MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDY 60  P0CG48  UBC_HUMAN
1  MQIFVKTLTGKTITLEVEPSDTIENVKAKIQDKEGIPPDQQRLIFAGKQLEDGRTLSDY 60  P0CG48  UBC_HUMAN
```

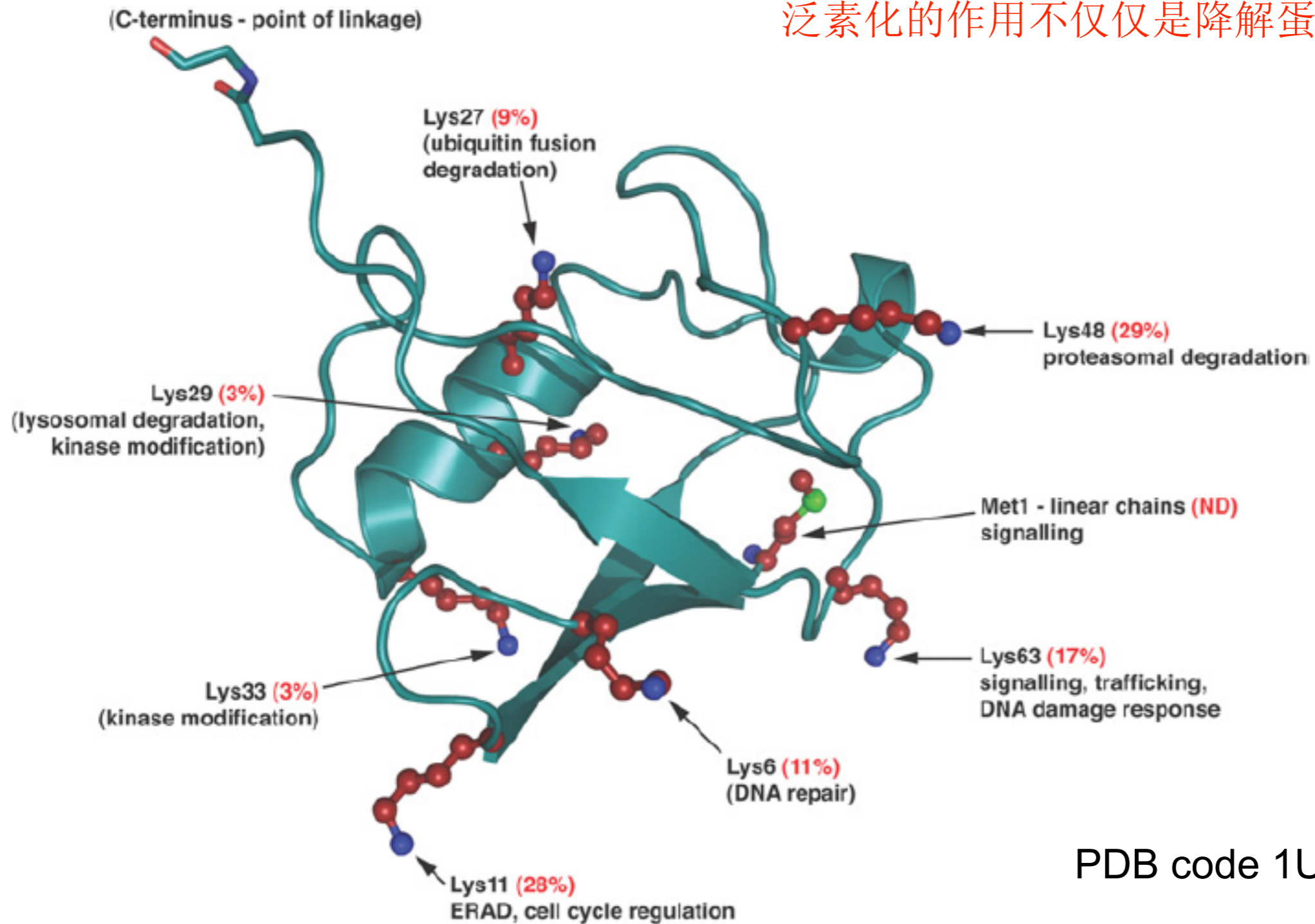
```
61  IQKESTLHLVLRLRGG 76  P62987  RL40_HUMAN
61  IQKESTLHLVLRLRGG 76  P62979  RS27A_HUMAN
61  IQKESTLHLVLRLRGG 76  P0CG47  UBB_HUMAN
61  IQKESTLHLVLRLRGG 76  P0CG47  UBB_HUMAN
61  IQKESTLHLVLRLRGG 76  P0CG47  UBB_HUMAN
61  IQKESTLHLVLRLRGG 76  P0CG48  UBC_HUMAN
61  IQKESTLHLVLRLRGG 76  P0CG48  UBC_HUMAN
61  IQKESTLHLVLRLRGG 76  P0CG48  UBC_HUMAN
61  IQKESTLHLVLRLRGG 76  P0CG48  UBC_HUMAN
61  IQKESTLHLVLRLRGG 76  P0CG48  UBC_HUMAN
61  IQKESTLHLVLRLRGG 76  P0CG48  UBC_HUMAN
61  IQKESTLHLVLRLRGG 76  P0CG48  UBC_HUMAN
61  IQKESTLHLVLRLRGG 76  P0CG48  UBC_HUMAN
61  IQKESTLHLVLRLRGG 76  P0CG48  UBC_HUMAN
61  IQKESTLHLVLRLRGG 76  P0CG48  UBC_HUMAN
```

Similarity = 100%

Ubiquitin of Eukaryotes

泛素及其7个赖氨酸残基

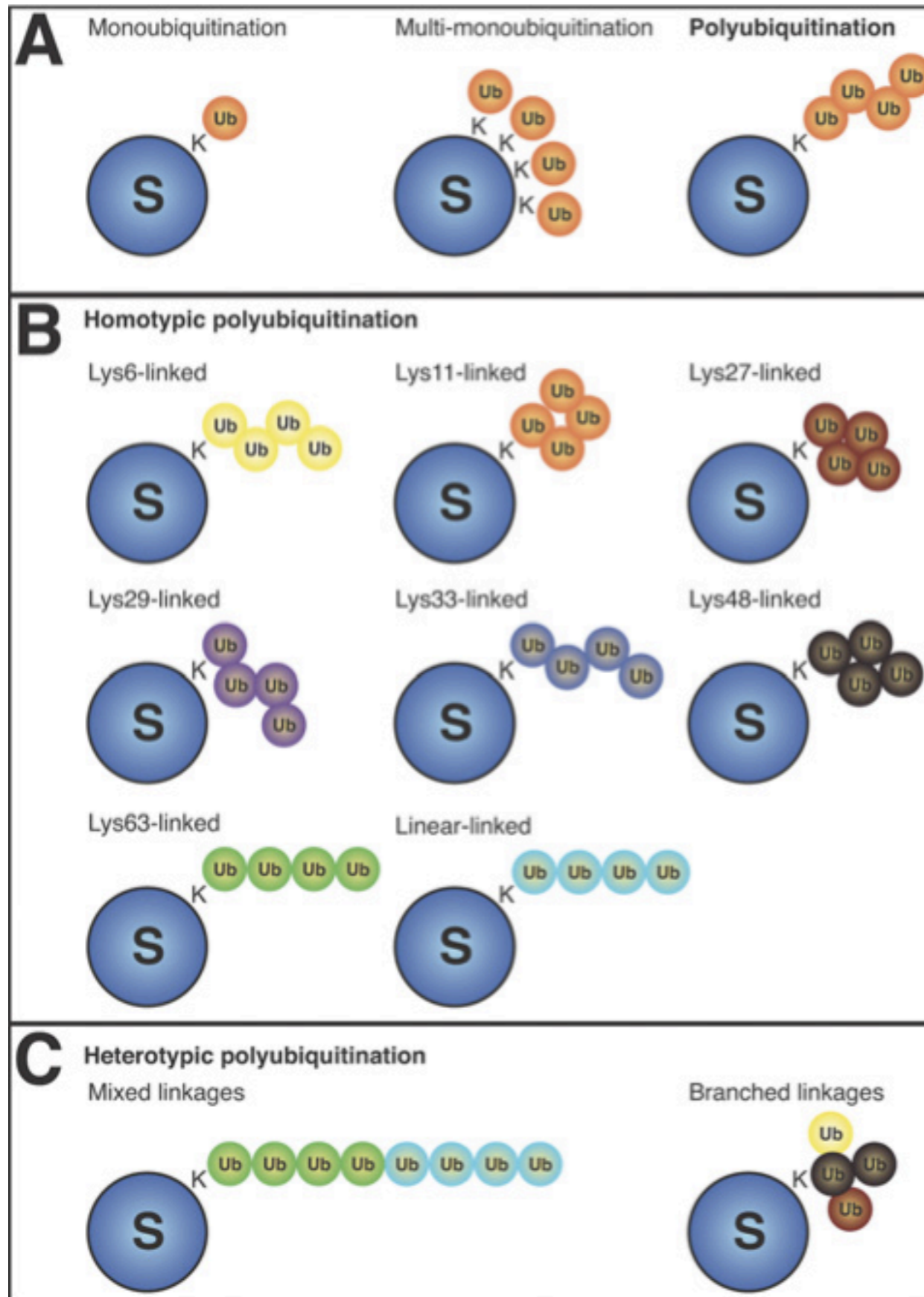
泛素化的作用不仅仅是降解蛋白！



PDB code 1UBQ

| 泛素分子的结合位点 | 功能 |
|-----------|--|
| Lys48 | 通过蛋白酶体降解蛋白 |
| Lys11 | <p>1、与内质网相关的方式降解蛋白。如在人类细胞中，通过Lys11连接的泛素化链会与许多UBA/UBX家族的蛋白相互作用，而UBX结构域会与内质网相关降解途径的另一个调控子AAA蛋白Cdc48/p97相互作用。</p> <p>2、在哺乳动物的细胞周期中发挥作用。这个过程的关键E3连接酶APC/C（anaphase-promoting complex/cyclosome）蛋白复合物，将Lys11连接的泛素化链连接在细胞周期中需要被降解的底物上。</p> |
| Lys63 | <p>1、与内吞作用相关。如调节内吞作用的去泛素化酶AMSH，具有精确的识别Lys63连接的泛素链的特异性</p> <p>2、与DNA损伤应答有关。早期在酵母中证明将内源的泛素63位赖氨酸突变成精氨酸会使细胞对DNA损伤更敏感。</p> <p>3、在信号传导过程中起作用，激活转录因子NF-κB</p> |
| Lys6 | 参与DNA修复 |
| Lys29 | 通过溶酶体降解蛋白。比如Notch信号通路的调控蛋白Deltex就是被Lys29连接的泛素链结合后，通过溶酶体降解 |
| Lys33 | 激酶的修饰。如一些AMPK相关的激酶家族蛋白被Lys33连接的泛素链结合后，会抑制这些激酶的活性。 |
| Lys27 | 目前生物学功能还不是很清楚 |

泛素化的形式



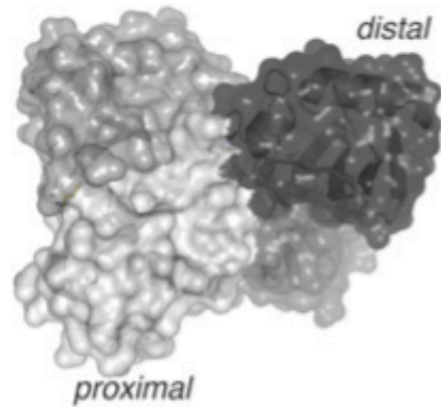
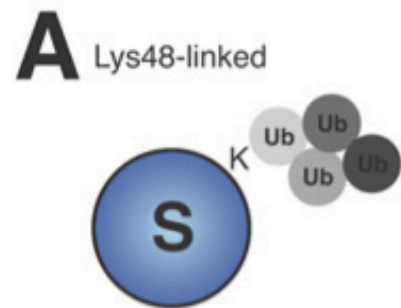
单泛素化、多位点单泛素化、多聚泛素化

同类型的多聚泛素化

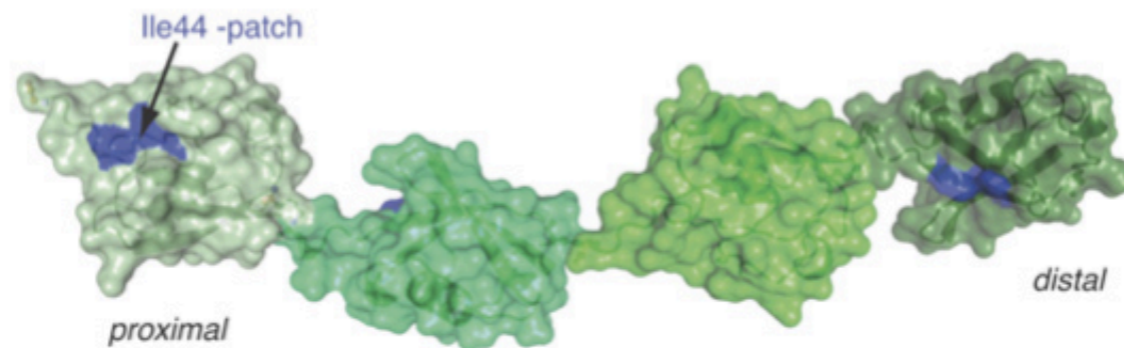
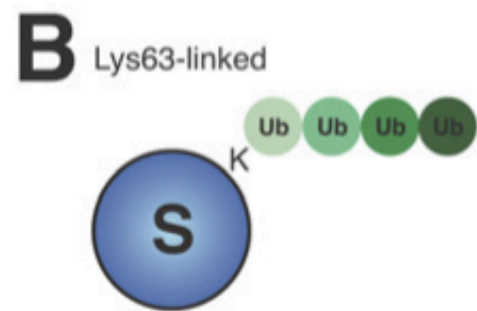
以Lys6-linked为例，后一个泛素分子羧基端甘氨酸都是与前一个泛素分子第6位赖氨酸连接，形成多聚泛素链

不同类型的多聚泛素化

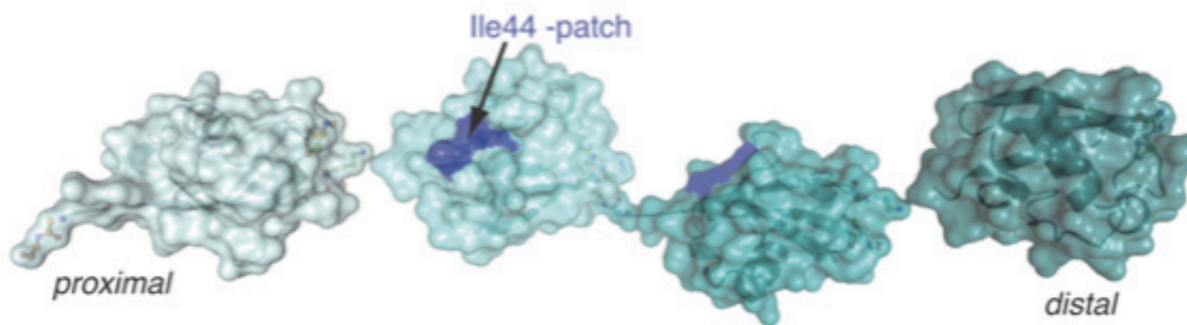
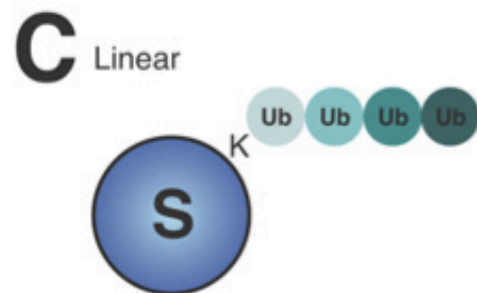
不同位点赖氨酸连接的泛素链结构



当泛素分子羧基端甘氨酸与前一个泛素分子48位赖氨酸连接时，泛素分子会相互作用，形成**紧密结构**。PDB 代码2O6V



63位赖氨酸连接时，形成开放的结构。左图中蓝色区域为疏水的**异亮氨酸**区域，它与泛素结合结构域及去泛素酶相互作用。PDB代码 2JF5

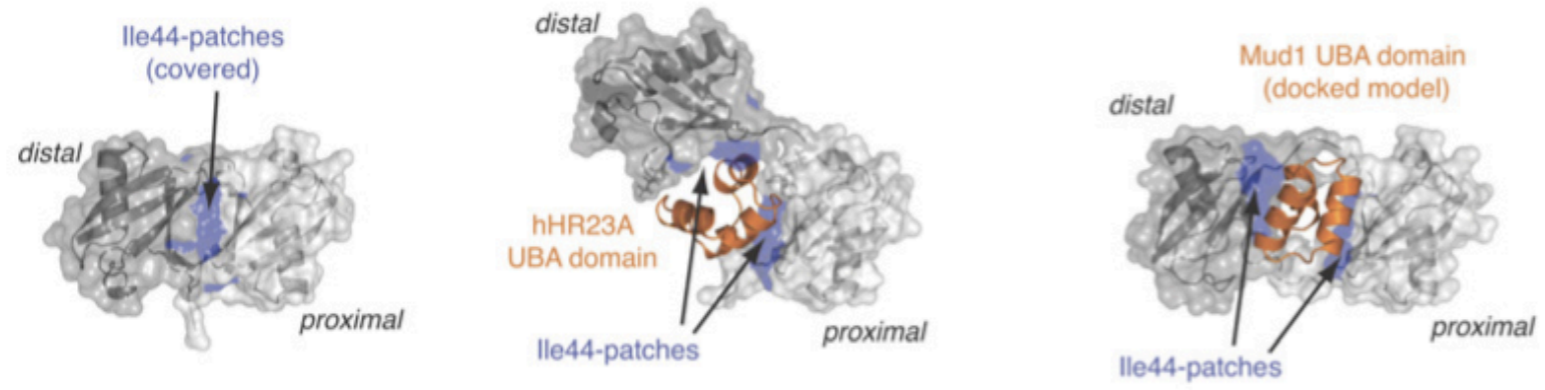


与**起始甲硫氨酸**连接成线状多泛素链，其结构也是开放的，与连在63位赖氨酸时的结构相似。PDB 代码 2W9N

正是由于不同位点赖氨酸连接的泛素链具有不同结构，才会与不同的特异蛋白相互作用，也就发挥不同的生物学功能！

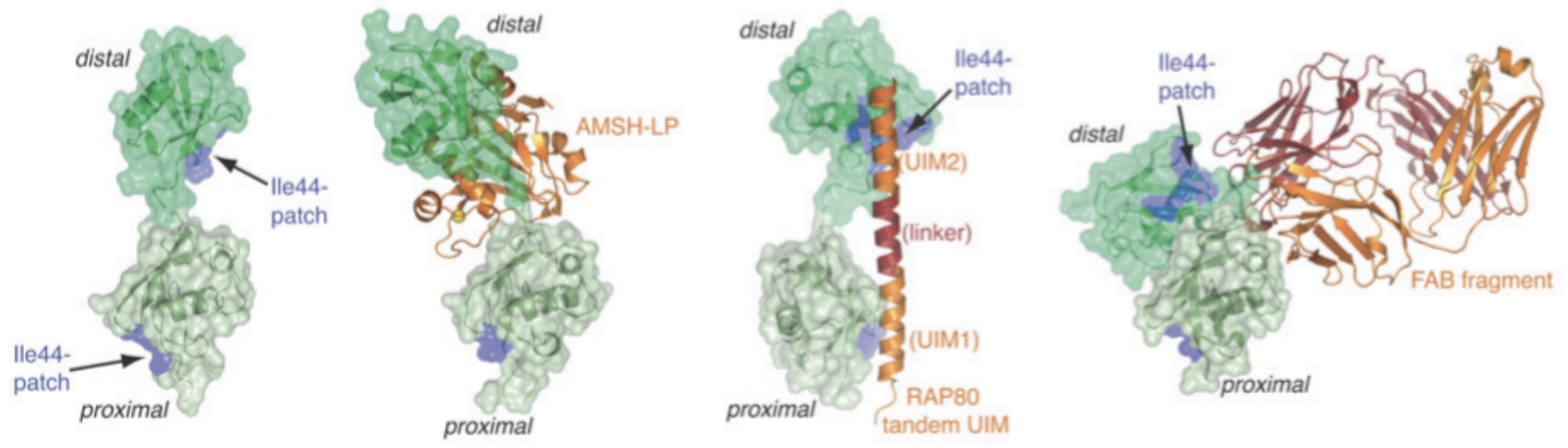
3种泛素化链与蛋白结合的结构

A Lys48 diubiquitin complex



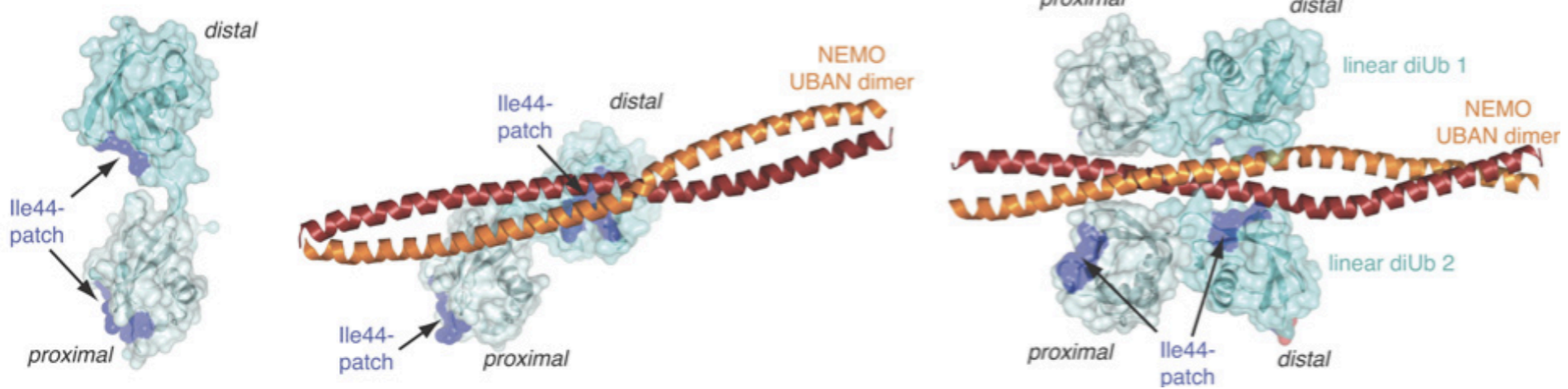
蛋白的泛素结合区域
(UBA) 插入到48位赖氨酸连接的泛素二聚体中，与泛素的异亮氨酸疏水区域相互作用

B Lys63 diubiquitin complexes



63位赖氨酸连接的泛素二聚体与去泛素化酶AMSH-LP或抗体作用的结构

C Linear diubiquitin complex

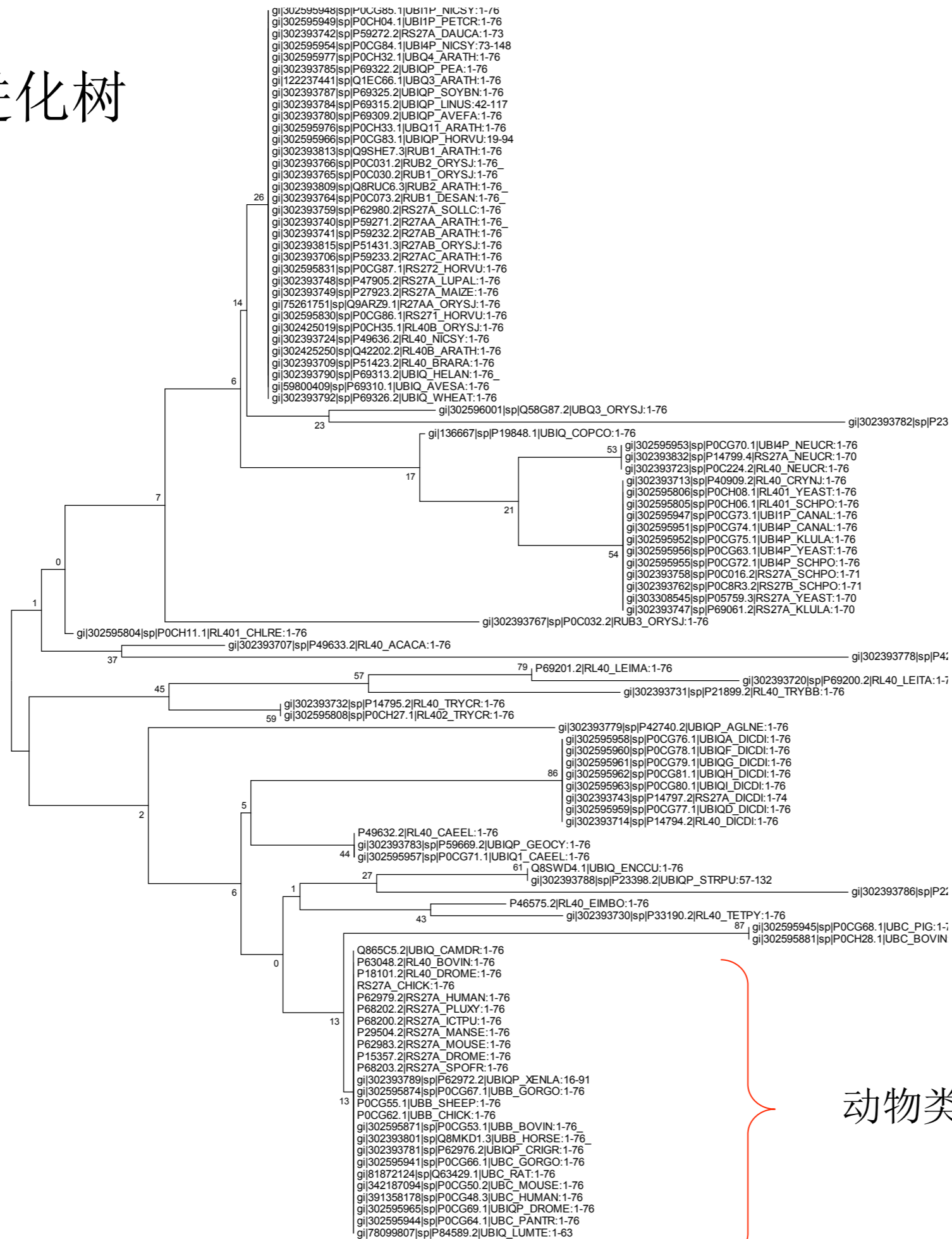


从头连接的线状泛素二聚体与NEMO蛋白的UBAN结构域作用的结构

泛素序列BLAST

- 为了找出泛素在真核生物中的进化关系，将人的泛素序列在NCBI中进行BLAST,尽管搜索条件很严格，但仍找出了很多序列，其中有一些是类泛素。发现在真核生物中查找出来的泛素序列与人的泛素序列相似性都在95%以上，说明泛素在真核生物中非常保守。
- 每个物种保留一条泛素序列，然后将BLAST出来的真核生物的泛素序列用MEGA构建系统发育树，看它们的进化关系。
- 从进化树中可以看到人类的泛素与其他动物的泛素基本在一条分支，而植物类、酵母类等分别在另外的一条分支上。

泛素进化树




动物类

0.5

Ubiquitin-like Proteins of Human

人中的类泛素分子

- UniProt中搜索  18条
- 发现基本上都有类泛素结构域
- 参与各种生理过程，不仅仅是蛋白质降解
 - 可以稳定某些蛋白
 - 底物蛋白的亚细胞定位
 - 细胞周期控制、DNA的复制和修复
 - 免疫系统的激活、细胞自噬
 - 抗病毒

类泛素结构域的相似性

- 提取以上18条类泛素分子中的类泛素结构域，进行序列比对 (tcoffee)
- 总体在序列上保守性不强，其C端较N端保守
- 在结构上可能有相似的相似性——“类泛素折叠”

| BAD AVG GOOD | | | |
|--------------|--------|-------|------|
| sp | A6NDN8 | 23-99 | : 78 |
| sp | Q15843 | 1-76 | : 83 |
| sp | P63165 | 20-97 | : 80 |
| sp | P61956 | 16-95 | : 78 |
| sp | P55854 | 15-92 | : 80 |
| sp | Q6EEV6 | 17-95 | : 76 |
| sp | Q9BVT8 | 103-1 | : 71 |
| sp | Q9NRR5 | 13-87 | : 77 |
| sp | O15205 | 6-81 | : 79 |
| sp | O15205 | 90-16 | : 77 |
| sp | Q8WUN7 | 152-2 | : 82 |
| sp | O95164 | 10-88 | : 73 |
| sp | P11441 | 1-76 | : 83 |
| sp | Q8N7F7 | 1-76 | : 82 |
| sp | Q9BZL1 | 1-73 | : 76 |
| sp | Q96S82 | 18-98 | : 61 |
| sp | P05161 | 2-78 | : 77 |
| sp | P05161 | 79-15 | : 81 |
| cons | | | : 77 |

| | | | | | | | | |
|-------------------|---|-------|----------------|------------|-----------------------------------|-------------------------------|------|----|
| sp A6NDN8 23-99 | 1 | --- | LC | -- | PQVAYVRARELHTLEVTGLETVAQSKAHVASLE | ---- | 35 | |
| sp Q15843 1-76 | 1 | --- | ML | -- | IKVKTTLTGKE | -IEIDIEPTDKVERIKERVEEKEG | ---- | 35 |
| sp P63165 20-97 | 1 | --- | EYIKLKVIGQDSSE | - | IHFVKMTTHLKKLKE | SYCQRQ | ---- | 36 |
| sp P61956 16-95 | 1 | --- | DHINLKVAGQDGSV | - | VQFKIKRHTPLSKLMKAYCERQ | ---- | 36 | |
| sp P55854 15-92 | 1 | --- | DHINLKVAGQDGSV | - | VQFKIKRHTPLSKLMKAYCERQ | ---- | 36 | |
| sp Q6EEV6 17-95 | 1 | --- | H | - | INLKVAGQDGSV | -VQFKIKRQTPLSKLMKAYCEPR | ---- | 35 |
| sp Q9BVT8 103-1 | 1 | --- | LV | -- | LRLKFLNDSE | -QVARAWPHDTIGSLKRTQ | ---- | 32 |
| sp Q9NRR5 13-87 | 1 | --- | IRVTV | -- | KTPKDKE | -EIVICDRASVKEFKEEISRRF | ---- | 33 |
| sp O15205 6-81 | 1 | --- | SCLCVHVRSEEWDL | - | MTFDANPYDSVKKIKEHVRSKT | ---- | 36 | |
| sp O15205 90-16 | 1 | --- | LF | -- | LVE SGDEAKR | -HLLQVRRSSSVAQVKAMIEKTK | ---- | 34 |
| sp Q8WUN7 152-2 | 1 | --- | CQLRL | -- | RLSTGKD | -LKLVVRSTDTVFHMKRRLHAAE | ---- | 34 |
| sp O95164 10-88 | 1 | --- | IN | -- | LRLILVSGKT | -KEFLFSPNDSASDI AKHVYDNWPMDWE | ---- | 39 |
| sp P11441 1-76 | 1 | --- | MQ | -- | LTVKALQGREC | -SLQVPEDELVSTLKQLVSEKL | ---- | 34 |
| sp Q8N7F7 1-76 | 1 | --- | MF | -- | LTVKLLLGQR | -CSLKVSGQESVATLKRLVSRRLK | ---- | 35 |
| sp Q9BZL1 1-73 | 1 | --- | MI | - | EVVCNDRLGKK | -VRVKCNTDDTIGDLKKLIAAQTG | ---- | 36 |
| sp Q96S82 18-98 | 1 | --- | TP | -- | KSILRLPETELGEYSL | -GGYSISFLKQLIAGKQLQ | ---- | 35 |
| sp P05161 2-78 | 1 | G | WD | -- | LTVKMLAGNE | -FQVSLSSSMSVSELKAQITQKI | ---- | 35 |
| sp P05161 79-15 | 1 | DEPLS | -- | ILVRNNKGRS | -STYEVRLTQTVAHLKQQVSGLE | ---- | 37 | |

| | | | |
|------|---|--|----|
| cons | 1 |  | 45 |
|------|---|--|----|

| | | | | | | | | | | |
|----|--------|--------------|---------|-----------|----------|----------------------|----------------------|-----------------------|----------------------|----|
| 36 | --- | GLI | - | PEDKVVL | -- | LAGSPLQN | - | EATLGQCGVEALTTLLEVGR | 73 | |
| 36 | --- | - | IP | - | PQQQLI | -- | YSGKQMN | - | EKTAADYKILGGSVLHLVLA | 72 |
| 37 | --- | GVP | - | MNSLRFL | -- | FEGQRIAD | - | NHTPKELGMEEDVIEVYQE | 74 | |
| 37 | --- | GLS | - | MRQIRFR | -- | FDGQPINE | - | TDTPAQLEMEDEDIDVFQQ | 74 | |
| 37 | --- | GLS | - | MRQIRFR | -- | FDGQPINE | - | TDTPAQLEMEDEDIDVFQQ | 74 | |
| 36 | --- | GLS | - | MKQIRFR | -- | FGGQPISG | - | TDKPAQLEMEDEDIDVFQQ | 73 | |
| 33 | --- | - | GR | - | EQQVRLI | -- | YQGQLLGD | - | DTQTLGSLHLPNCVLHCHVS | 70 |
| 34 | --- | KAQ | - | QDQLVLI | -- | FAGKILKD | - | GDTLNQHG IKDGLTVHLVIK | 71 | |
| 37 | --- | KVP | - | VQDQVLL | -- | LGSKILKP | - | RRSLSSYGIDKEKTIHLL | 73 | |
| 35 | --- | GII | - | PETQIVT | -- | CNGKRLED | - | GKMMADYGIRKGNLLFLAC | 71 | |
| 35 | --- | GVE | - | PGSQRWF | -- | FSGRPLTD | - | KMKFEELKIPKDYVVQVIVS | 72 | |
| 40 | EEQVSS | - | PNILRLI | -- | YQGRFLHG | - | NVTLGALKLPFGKTTV | ---- | 76 | |
| 35 | --- | NVP | - | VRQQRLI | -- | FKGKALAD | - | GKRLSDYSIGPNSKLNLVVK | 72 | |
| 36 | --- | - | VP | - | EEQQHLL | -- | FRGQLLED | - | DKHLSDYCIGPNASINVIMQ | 72 |
| 37 | --- | - | TR | - | WNKIVLK | -- | KWYTIKFD | - | HVSLGDYEIHDGMNLELY | 71 |
| 36 | -- | ESVPDPELIDLI | -- | YCGRKLKD | - | DQTLDFYGIQPGSTVHVLRK | ---- | 75 | | |
| 36 | --- | GVH | - | AFQQLAVHP | SGVALQD | - | RVP LASQGLGPGSTVLLVV | ---- | 74 | |
| 38 | --- | GVQ | - | DDLFWLT | -- | FEGKPLED | - | QLPLGEYGLKPLSTVFMNLR | 75 | |

| | | | | |
|----|---|---|---|----|
| 46 |  | : | : | 90 |
|----|---|---|---|----|

阿尔茨海默症 & 泛素化异常

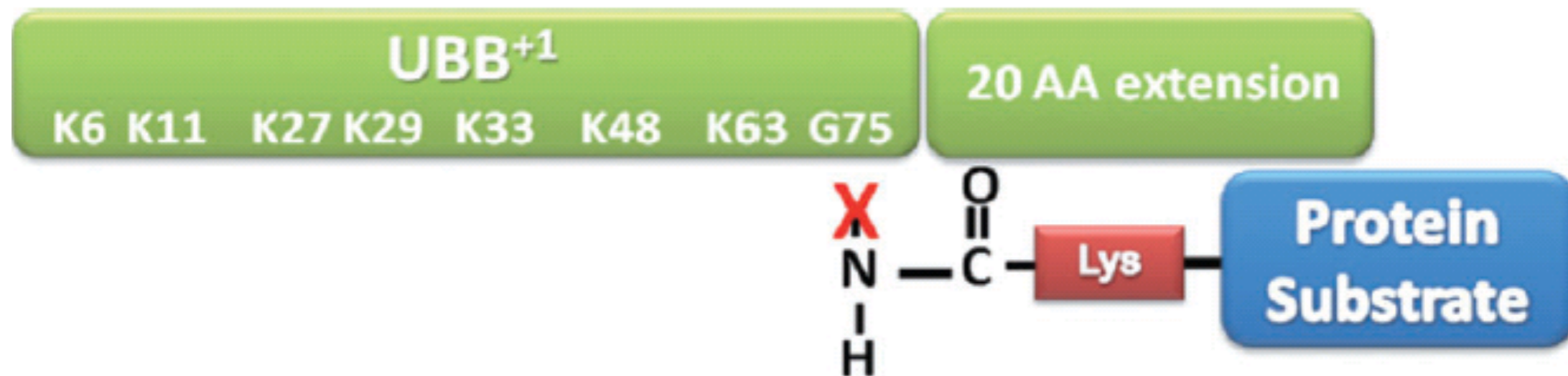
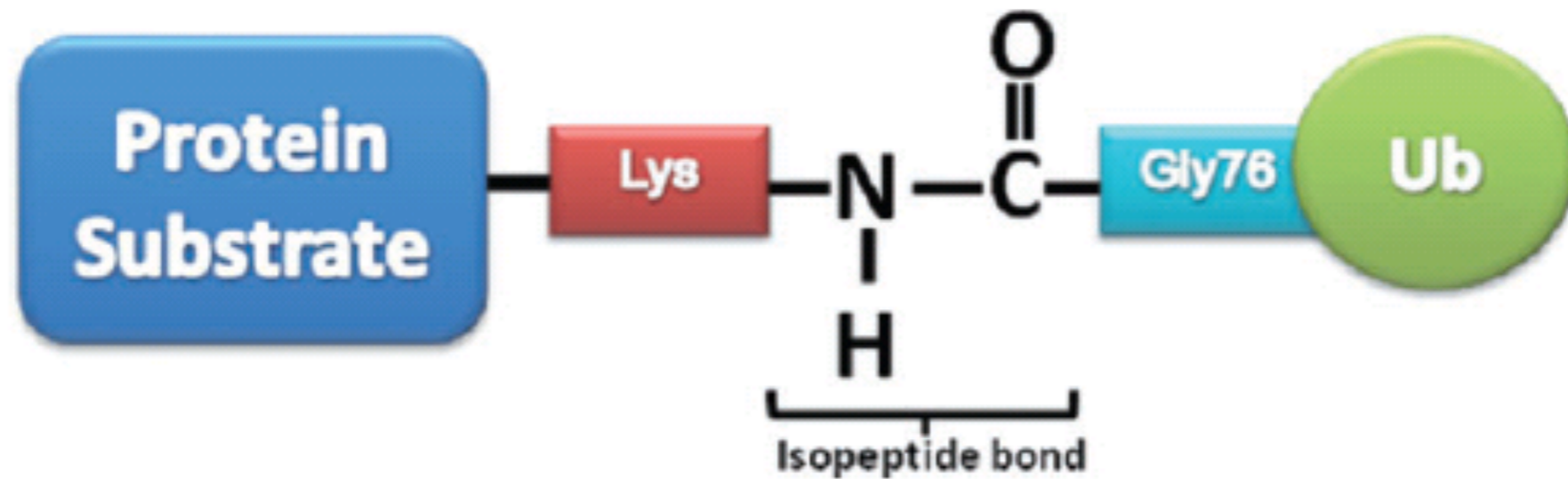
- 阿尔茨海默症

以脑内广泛存在胞外类淀粉样蛋白 ($A\beta$) 沉淀的斑块以及神经元内纤维状缠绕的高度磷酸化微管相关蛋白 (p-tau蛋白) 为重要的特征

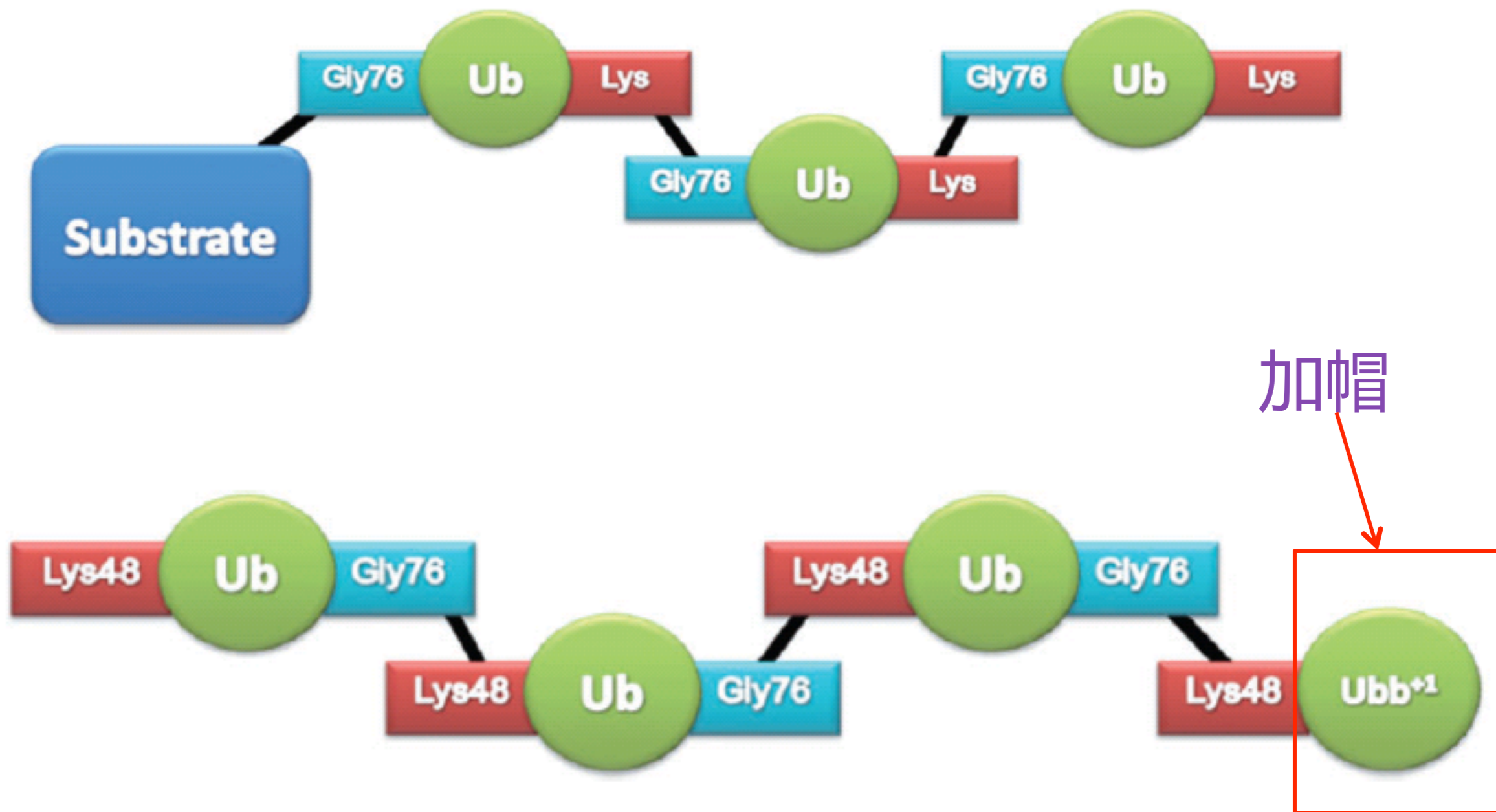
- 泛素化异常

UBB基因mRNA非编码序列的两个核苷酸删除造成的突变 Ubb+1, 其76位甘氨酸缺失, 使得多泛素链在76位连接下一个泛素分子的过程终止, 由此形成很多体内游离的多泛素链

正常的泛素 Vs. 异常泛素



多泛素链的异常

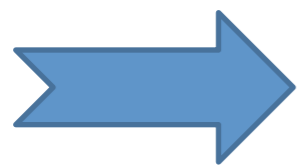


“加帽”多泛素链的影响

- 在泛素-蛋白酶体系统中，被泛素分子标记的蛋白，进入26S蛋白酶体，在降解之前或降解时，多泛素链会与靶蛋白分离，其C端游离-COOH被异肽键酶而降解
- 体内加帽的多泛素链，C端不能被识别，因此体内产生游离的多泛素链
- 这些游离的多泛素链，会竞争性结合蛋白酶体的泛素识别位点而抑制其对靶蛋白的降解

多泛素链的活性

- 除了以上提到的，阿尔茨海默症患者体内游离的多泛素链可能抑制了蛋白酶体活性，从而使异常蛋白降解受到阻碍，与脑内淀粉样沉积有关
- 最近发现，游离的多泛素链具有激活激酶的活性



是否可能与tau蛋白的
过度磷酸化有关？

可能的tau蛋白激酶

- UniProt中输入(tau phosphorylation AND organism:"Human [9606]") AND reviewed:yes
检索出70条相关序列
- 用SMART利用结构域查找
如果是有多泛素链激活而介导的蛋白激酶活性，必定同时含有蛋白激酶结构域和泛素分子结合结构域（UBA）

- Homo sapiens (13)
 - D3DTB1_HUMAN *MAP/microtubule affinity-regulating kinase 1, isoform CRA_b*
 - O75271_HUMAN *R31237_1, (Fragment)*
 - P27448-2 *Isoform 2 of MAP/microtubule affinity-regulating kinase 3*
 - P27448-3 *Isoform 3 of MAP/microtubule affinity-regulating kinase 3*
 - P27448-5 *Isoform 5 of MAP/microtubule affinity-regulating kinase 3*
 - MARK3_HUMAN *MAP/microtubule affinity-regulating kinase 3*
 - SIK1_HUMAN *Serine/threonine-protein kinase SIK1*
 - Q7KZI7-9 *Isoform 9 of Serine/threonine-protein kinase MARK2*
 - MARK2_HUMAN *Serine/threonine-protein kinase MARK2*
 - Q96L34-2 *Isoform 2 of MAP/microtubule affinity-regulating kinase 4*
 - MARK4_HUMAN *MAP/microtubule affinity-regulating kinase 4*
 - MARK1_HUMAN *Serine/threonine-protein kinase MARK1*
 - UPI0001639AC6 *MAP/microtubule affinity-regulating kinase 2 isoform f*

激酶和UBA结构域相似性

- 基于以上五种蛋白，从UniProt中调取序列的激酶结构域和UBA结构域，分别进行序列比对（tcoffee）
- 发现均具有极高的相似性
- 激酶结构域：UBA结构域：

*
BAD AVG GOOD
*

| | |
|-----------------|-------|
| sp Q9P0L2 MARK1 | : 100 |
| sp Q7KZI7 MARK2 | : 100 |
| sp P27448 MARK3 | : 100 |
| sp Q96L34 MARK4 | : 100 |
| sp P57059 SIK1 | : 100 |
| cons | : 100 |

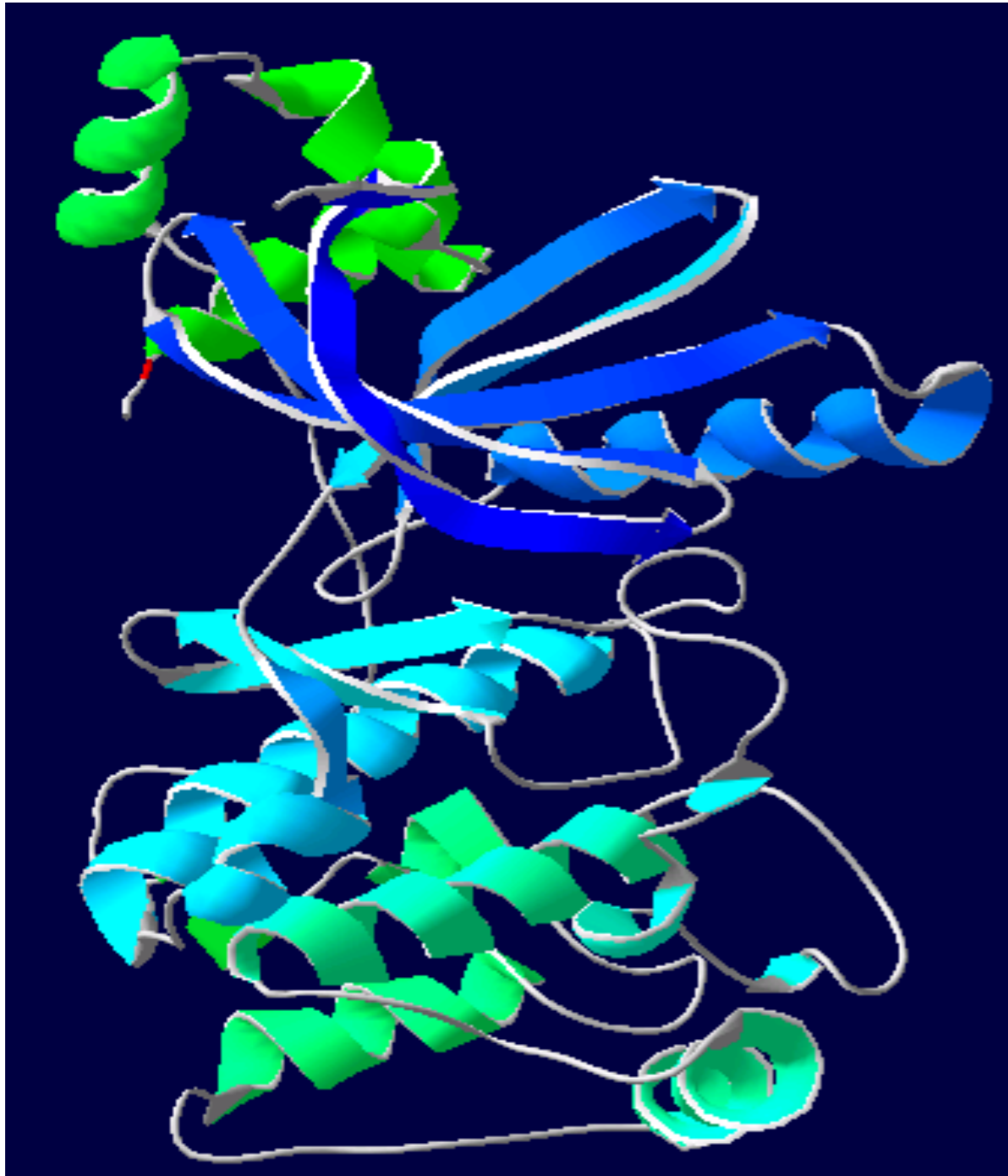
*
BAD AVG GOOD
*

| | |
|-----------------|------|
| sp Q9P0L2 MARK1 | : 94 |
| sp Q7KZI7 MARK2 | : 97 |
| sp P27448 MARK3 | : 96 |
| sp Q96L34 MARK4 | : 96 |
| sp P57059 SIK1 | : 93 |
| cons | : 95 |

结构分析

- 基于以上序列比对结果，发现其有高度一致的序列，猜测其结构也相似，因此选用其中的MARK3为例进行结构分析，以探究UBA结构域与激酶结构域是否有相互作用，进而影响了激酶的活性
- 通过文献检索，找到相关X-ray晶体学文献，找到了MARK3的PDB号，3fe3

3fe3

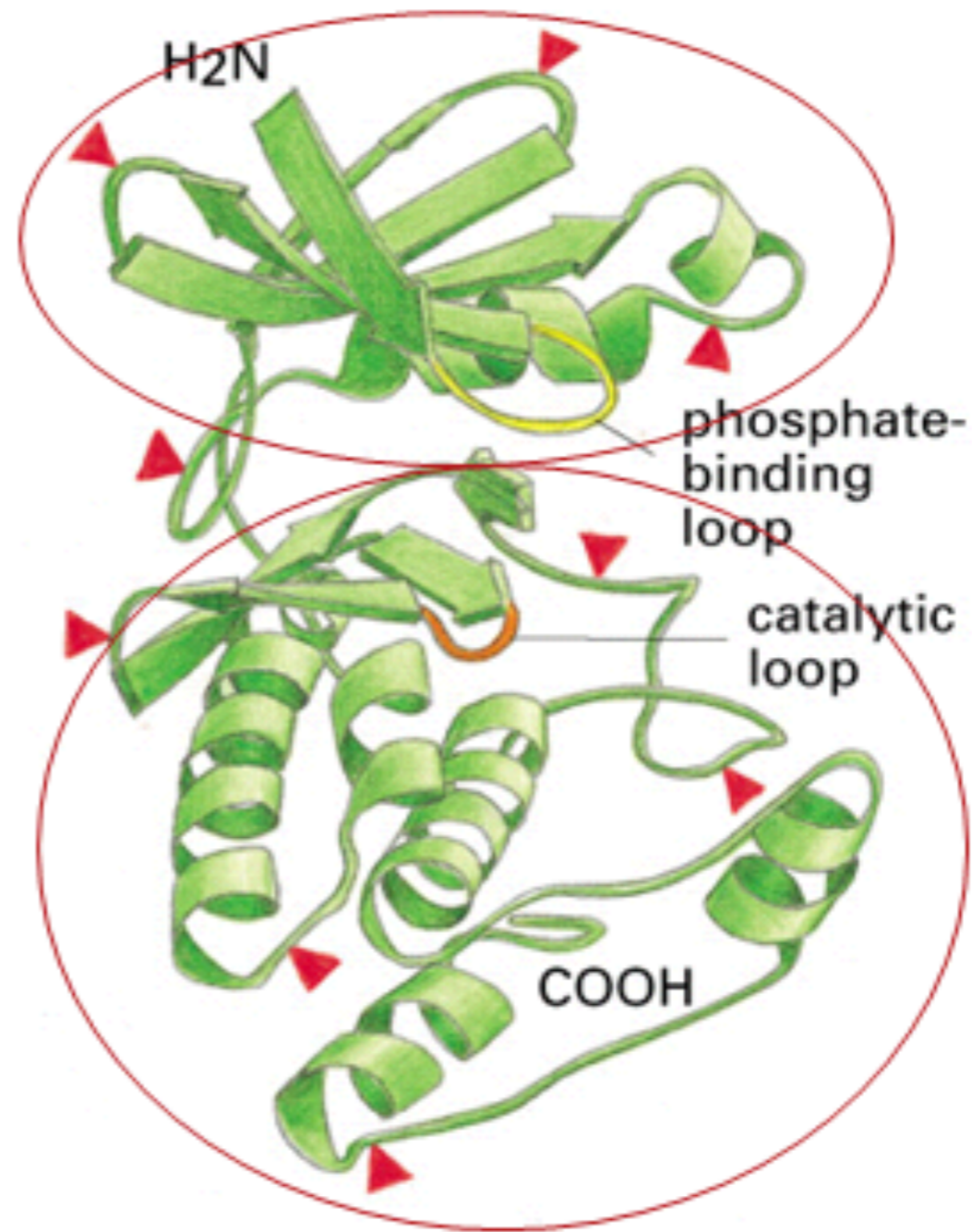


这个蛋白晶体结构只包含了从50到366位氨基酸，且是T211A和S215A双突变的

整体结构是同源二聚体，左图只显示其中一个单体：（红色为C端）

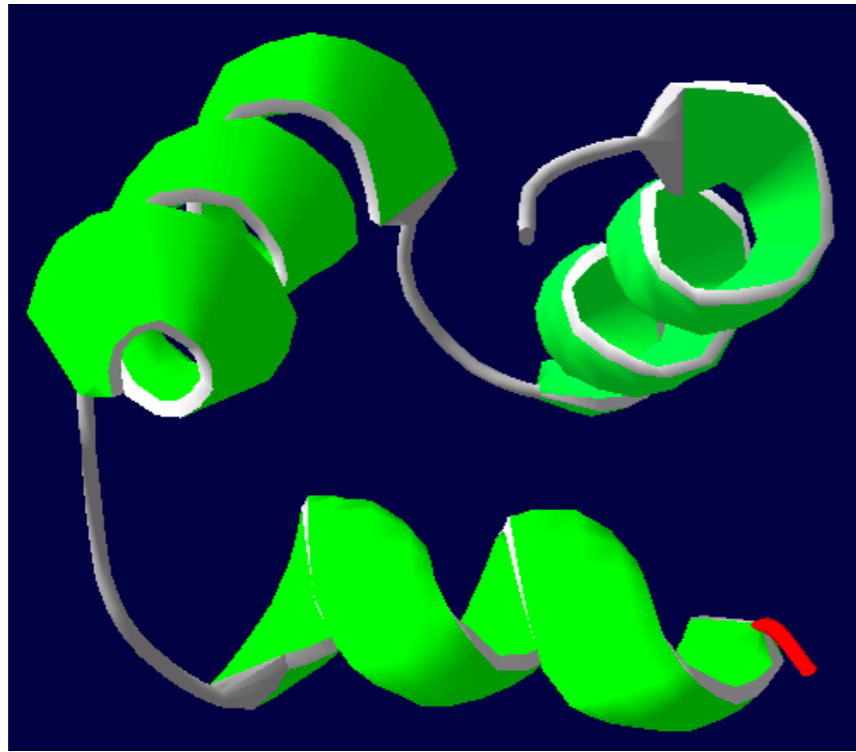
MARK3中56-307位是激酶结构域，326-365是UBA结构域

激酶结构域



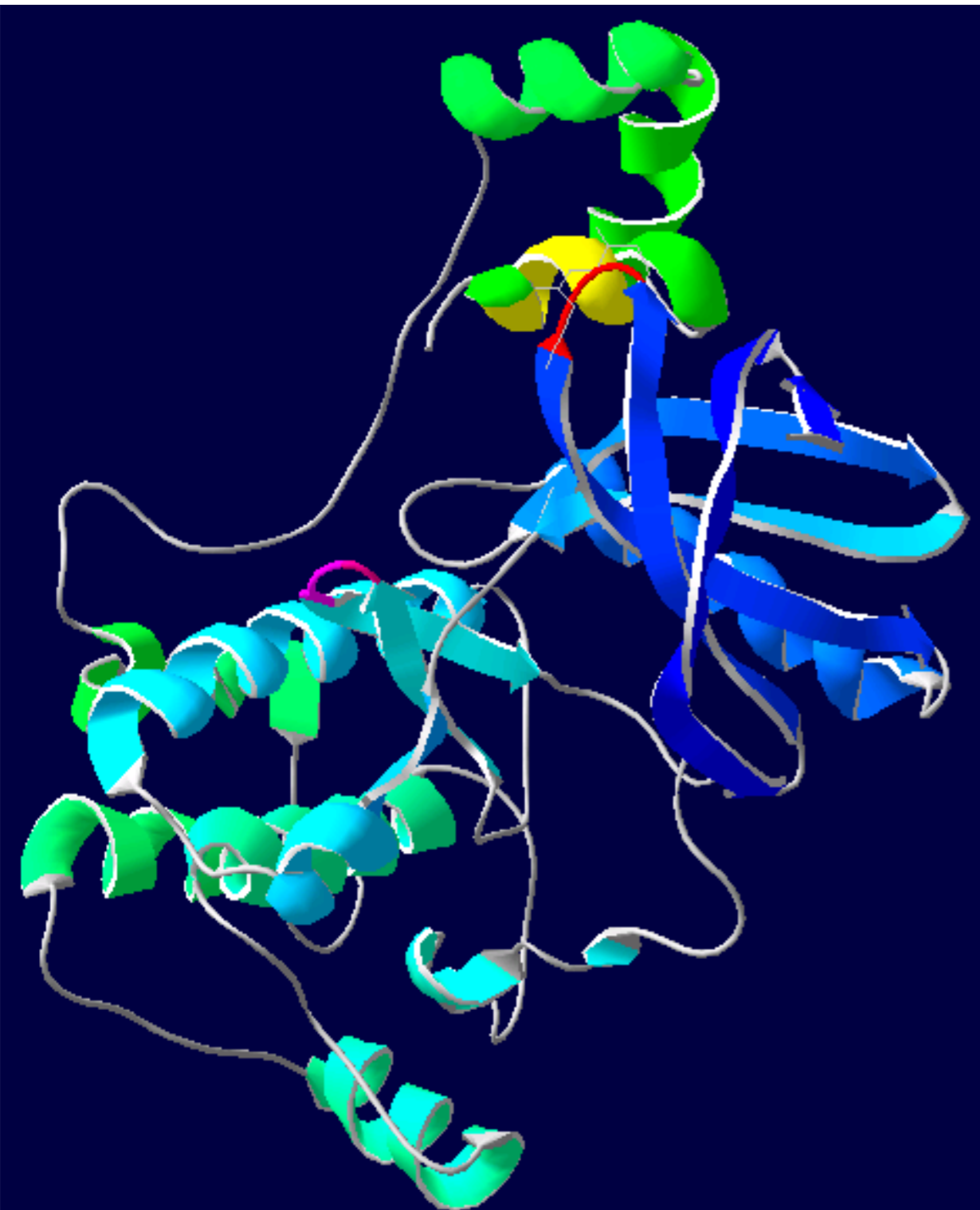
王映真老师课件

UBA结构域



有3个 α -helixs组成，靠近C端（红色标记处）的部分与激酶中的ATP结合位点距离很近

可能是通过C端的氨基酸残基侧链与激酶结构域中的ATP结合loop相互作用，抑制了激酶的活性



分析UBA结构域调节激酶结构域的可能的部位：
（只能是空间上比较近的区域）

红色标示为ATP结合loop，粉红色标示为催化loop，黄色标出的UBA结构域中C端的 α -helix上的一段，从空间位置上分析是最有可能作用于ATP结合位点

各段氨基酸残基分别是：

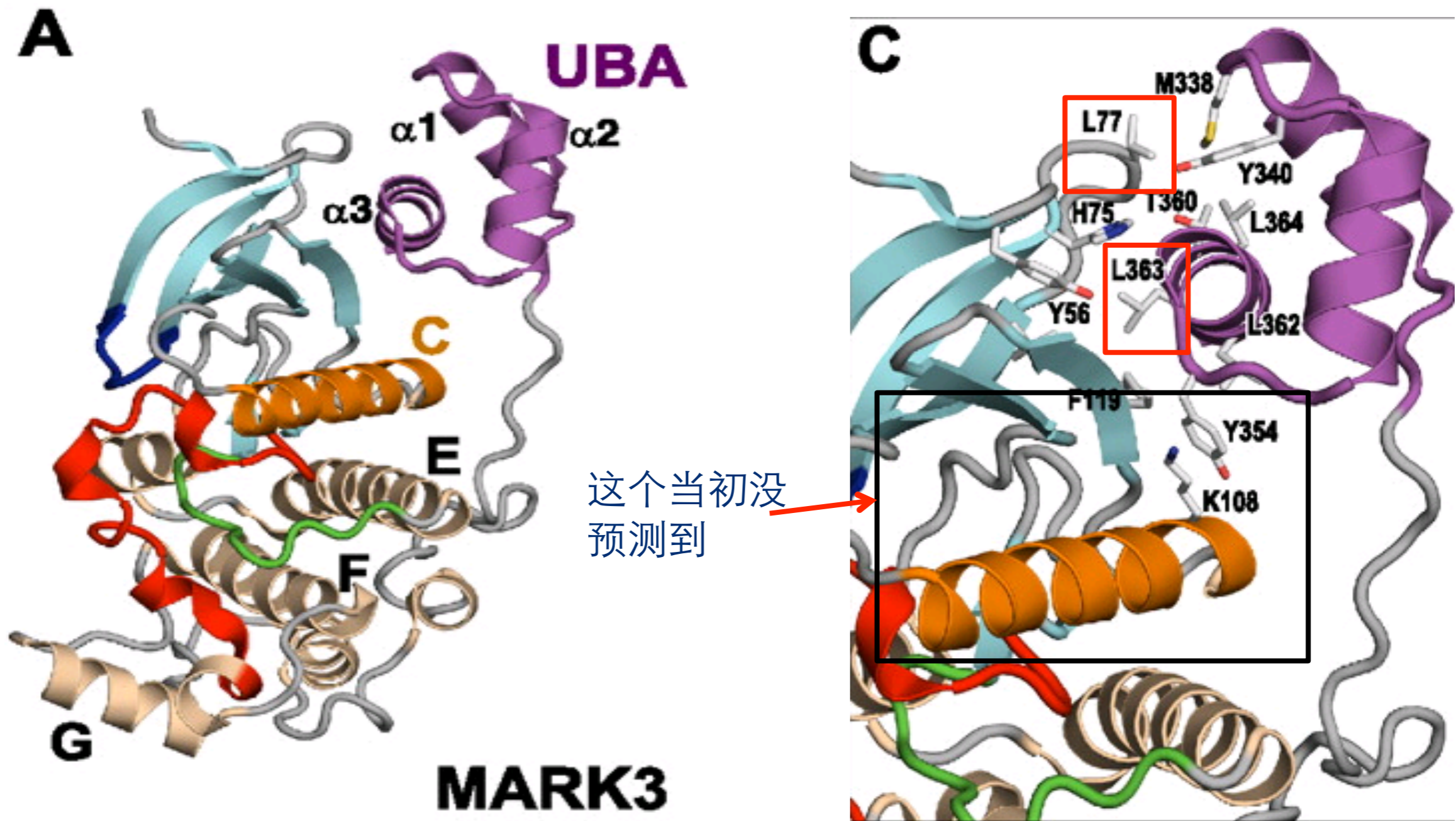
大红色：（ATP结合区）77Leu **78Thr** 79Gly

粉红色：（催化中心）188Ala 189Asp 190Met

黄色：（UBA结构域上作用位点）359Ala 360Thr 361Tyr 362Leu **363Leu**

作用位点

- 从空间位置上看，只能是UBA的C末端
（ 362Leu 363Leu ） 与ATP结合loop （ 77Leu
78Thr 79Gly ） 互作
- 但这只是从空间位置上的分析，不够准确也不精确，因此进行了文献检索，验证相关猜想



UBA与激酶结构域的相互作用

- 文献中指出，由激酶结构域的loop和strands形成了一个groove，而UBA结构域中靠近C端的helix就结合在这个groove中
- 363位leu在UBA结构域与激酶结构域相互作用中起了重要作用，与56位Y、75位H、78位T、80位R、82位V和119位F之间有作用力
- 另外，UBA结构域通过354位Y与108位K相互作用，这是当时没有预测到的

结论

- 基于以上的分析以及文献阅读，认为激酶结构域中ATP结合部位与UBA结构域C端的螺旋互作而抑制了激酶的活性
- 当有游离的多泛素链结合UBA结构域时，将会减弱UBA对ATP结合部位的抑制，使得激酶去抑制，磷酸化下游的tau
- 另外，根据文献，体内体外实验均证明，异常的泛素-蛋白酶体系统中，由于“Ubb+1帽”多泛素链对蛋白酶体有抑制作用，将导致磷酸化的tau蛋白积累
- 可见，多聚泛素链可通过有UBA结构域介导对tau蛋白激酶活性的调节，从而在阿尔茨海默症的病程中起了一定作用

Prokaryotic Ubiquitin- like Protein

“Alignment and Analysis”

- blast the SwissProt for the well-annotated prokaryotic ubiquitin-like proteins, sequence analysis and explore the phylogenetic relationships between the organisms
- blast the RefSeq mainly for the hypothetical prokaryotic ubiquitin-like proteins to predict which protein is a possible “Pup” for further investigation
- multiple global sequence alignment between the pups and the eukaryotic ubiquitin-like protein for further information.

Blast for the Well-Annotated Pups

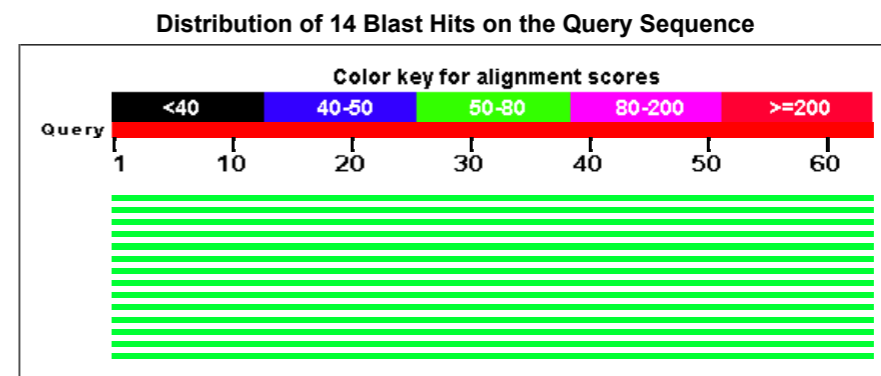
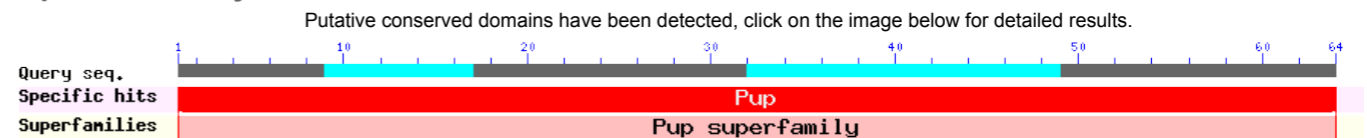
- Job Title: I_BLAST_SwissProt_PSI_minus_6_2;
- Database: UniProtKB/SwissProt
- Program: PSI-BLAST
- Expect Threshold: 0.000001
- Word Size: 2
- Matrix: PAM70
- Gap Costs: Existence 9 Extension 1
- Filter: Low Complexity Regions

1_BLAST_SwissProt_PSI_minus_6_2

Query ID |cl|74564
Description None
Molecule type amino acid
Query Length 64

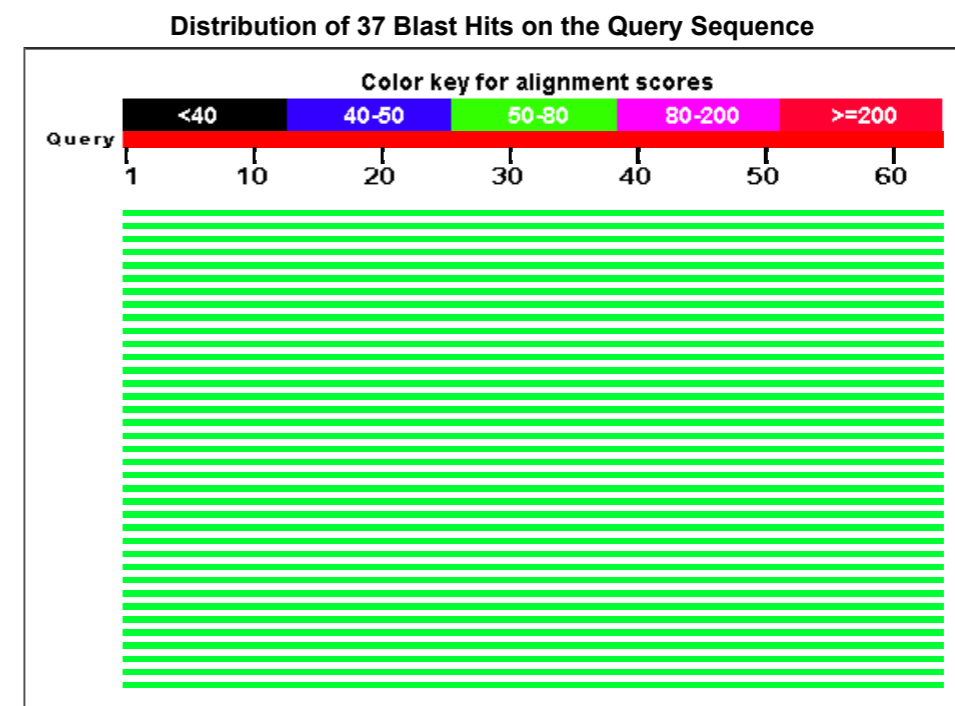
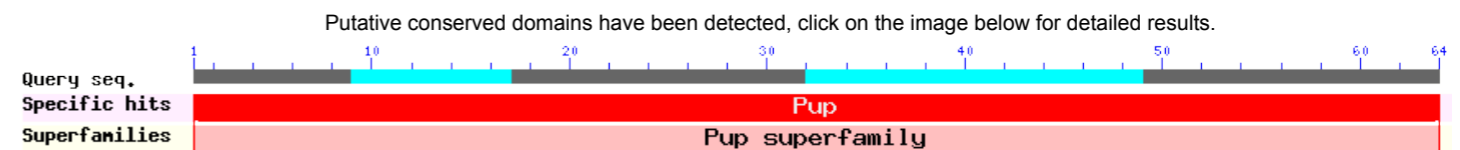
Database Name swissprot
Description Non-redundant UniProtKB/SwissProt sequences
Program BLASTP 2.2.27+

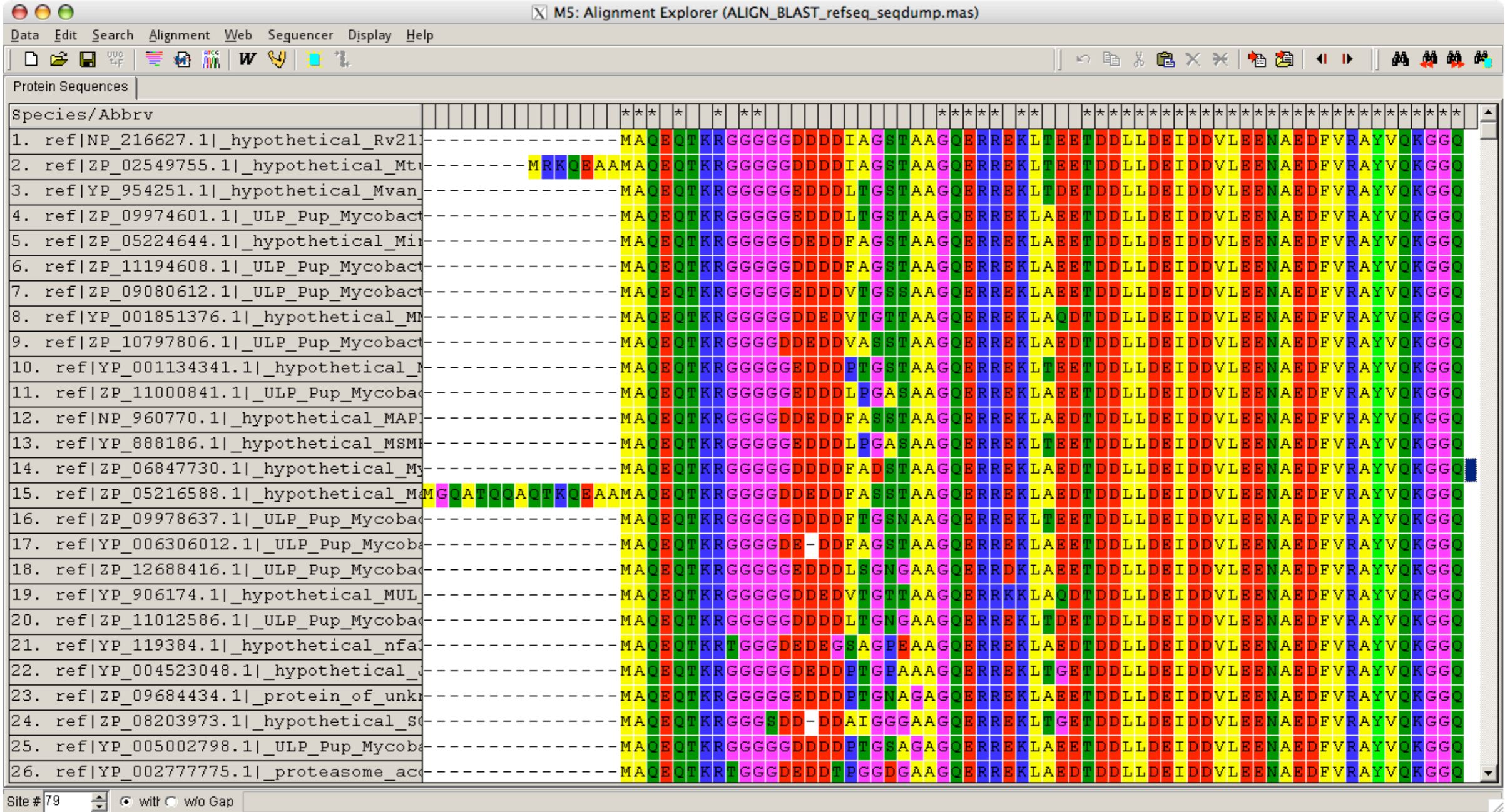
Graphic Summary



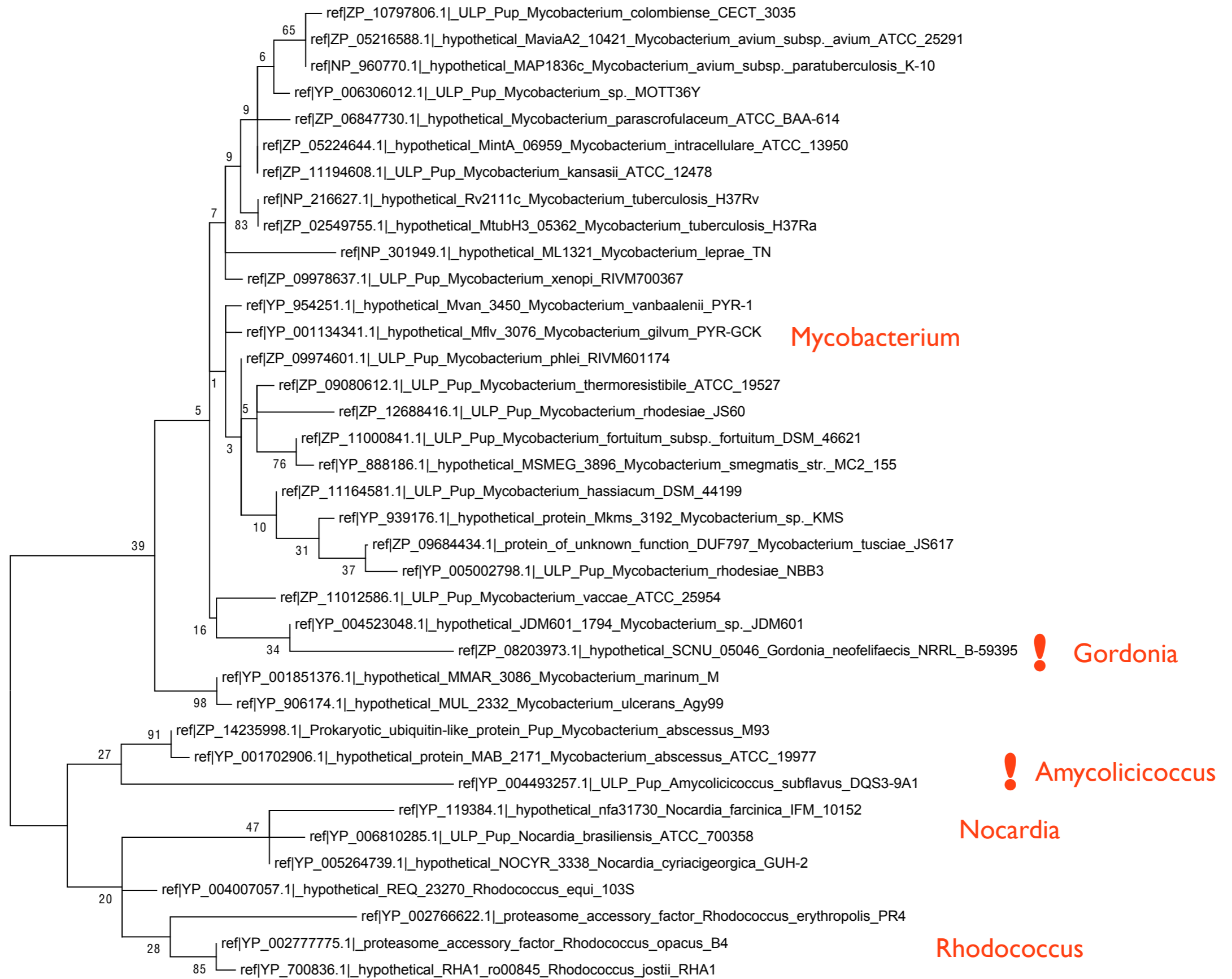
Blast for the Hypothetical Pups

- Job Title: I_BLAST_RefSeq_PSI_minus_6_2;
- Database: RefSeq
- Program: PSI-BLAST
- Expect Threshold: 0.000001
- Word Size: 2
- Matrix: PAM70
- Gap Costs: Existence 9 Extension 1
- Filter: Low Complexity Regions





Multiple Global Alignment



Phylogenetic Tree

Comparison Between PUP and UBL_(Eukaryotes)

- Sequences include: Eukaryotic ubiquitin-like proteins(SUMO2_HUMAN, SUMO3_HUMAN, SUMO4_HUMAN, ISG15_HUMAN, NEDD8_HUMAN, NEDD8_MOUSE, NEDD8_RAT); Prokaryotic ubiquitin-like proteins(THIS_ECOLI, THIS_BACSU, MOAD_BACSU, MOAD_STAAN); PUP_MYCTU.
- Note: only the ubiquitin-like domain!

```
>sp|O33246|PUP_MYCTU
MAQEQTKRGGGGDDDDIAGSTAAGQERREKLTEETDDLLDEIDDDVLEENAEDFVRAYVQKGGQ
>sp|O32583|THIS_ECOLI
MQILFNDQAMQCAAGQTVHELLEQLDQRQAGAALAINQOIVPREQWAQHIVQDGDQILLFQVIAGG
>sp|O31617|THIS_BACSU
MLQLNGKDVKKKDTGTIQDLLASYQLENKIVIVERNKEIIGKERYHEVELCDRDVIEIVHFVGGG
>sp|O31706|MOAD_BACSU
MIKILLFAGLAEQAGTQAEIDMEQATTDDEIKASLKEQYGLSIDTAMIAVNESYVKENTSVSSGDTVAIIPPVSGG
>sp|Q7A441|MOAD_STAAN
MKVLYFAEIKDILQKAQEDIVLEQALTVOQFEDLLFERYPQINNKKFQAVNEEFVQKSDFIQPNDTVALIPPVSGG
>sp|P61956|SUMO2_HUMAN(16-95)
DHINLKVAGQDGSVVQFKIKRHTPLSKLMKAYCERQGLSMRQIRFRFDGQPINETDTPAQLEMEDEDTIDVFQQQTGGVY
>sp|P55854|SUMO3_HUMAN 15-92
DHINLKVAGQDGSVVQFKIKRHTPLSKLMKAYCERQGLSMRQIRFRFDGQPINETDTPAQLEMEDEDTIDVFQQQTGG
>sp|Q6EEV6|SUMO4_HUMAN 17-95
HINLKVAGQDGSVVQFKIKRHTPLSKLMKAYCEPRGLSMKQIRFRFDGQPISGTDKPAQLEMEDEDTIDVFQQQTGGVY
>sp|P05161|ISG15_HUMAN 79-157
DEPLSILVRNNGRSTYEVRLTQTV AHLKQOVSGLEGVQDDLFWLTFEGKPLEDQLPLGEYGLKPLSTVFMNLRRLGG
>sp|Q64339|ISG15_MOUSE 77-155
SEPLSILVRNERGHSNIYEVFLTQTVDTLKKKVSQREQVHEDQFWLSFEGRPMEDEKELLGEYGLKPQCTVIKHLRLGG
>sp|Q15843|NEDD8_HUMAN 1-76
MLIKVKTTLTGKEIEIDIEPTDKVERIKERVEEKEGIPPQQORLIYSGQMNDKTAADYKILGGSVLHLVLALRGG
>sp|P29595|NEDD8_MOUSE 1-76
MLIKVKTTLTGKEIEIDIEPTDKVERIKERVEEKEGIPPQQORLIYSGQMNDKTAADYKILGGSVLHLVLALRGG
>sp|Q71UE8|NEDD8_RAT 1-76
MLIKVKTTLTGKEIEIDIEPTDKVERIKERVEEKEGIPPQQORLIYSGQMNDKTAADYKILGGSVLHLVLALRGG
>sp|O15205|UBD_HUMAN 90-(163+2)
LFLVESGDEAKRHLLQVRRSSVAQVKAMIEKTGTIIPETQIVTCNGKRLEDGKMMADYGIRKGNLLFLACYCIGG
>sp|P63072|UBD_MOUSE 87-(160+2)
LFLVESKNEGQRHLLRVRSSVAQVKEMIESVTSVIPKKQVNCNGKLEDGKIMADYNIKSGSLLFLTHCTGG
>sp|Q921A3|UBD_RAT 86-(159+2)
LSLVESGDEGQRHLLRVRSSVAQVKEMIENVTAVPPKKQIVNCNGKRLEDGKIMADYNIKSGSLLFLTAHCIGG
```

```

sp|O31706|MOAD_BACSU      MIKILLFAGLAEQAGTQAIEIDMEQA-TTDEIKASLKEQYGLESIDTAMI 49
sp|Q7A441|MOAD_STAAN     -MKVLYFAEIKDILQKAQEDIVLEQALTVQQFEDLLFERYPQINNKKFQV 49
sp|P61956|SUMO2_HUMAN_16-95_ --DHINLKVAGQDGSVVQFKIKRHTP--LSKLMKAYCERQGLSM-RQIRF 45
sp|P55854|SUMO3_HUMAN    --DHINLKVAGQDGSVVQFKIKRHTP--LSKLMKAYCERQGLSM-RQIRF 45
sp|Q6EEV6|SUMO4_HUMAN    ---HINLKVAGQDGSVVQFKIKRQTP--LSKLMKAYCEPRGLSM-KQIRF 44
sp|P05161|ISG15_HUMAN    -DEPLSILVRNNGRSSTYEVRLTQT--VAHLKQQVSGLEGVQD-DLFWL 46
sp|Q64339|ISG15_MOUSE   -SEPLSILVRNERGHSNIYEVFLTQT--VDTLKKKVSQREQVHE-DQFWL 46
sp|Q15843|NEDD8_HUMAN   -----MLIKVKTLTGKEIEIDIEPTDK--VERIKERVEEKEGIPP-QQORL 43
sp|P29595|NEDD8_MOUSE   -----MLIKVKTLTGKEIEIDIEPTDK--VERIKERVEEKEGIPP-QQORL 43
sp|Q71UE8|NEDD8_RAT     -----MLIKVKTLTGKEIEIDIEPTDK--VERIKERVEEKEGIPP-QQORL 43
sp|P63072|UBD_MOUSE     -----LFLVESKNEGQRHLLRVRSSS--VAQVKEMIESVTSVIP-KKQVV 43
sp|Q921A3|UBD_RAT       -----LSLVESGDEGQRHLLRVRSSS--VAQVKEMIENVTAAPP-KKQIV 43
sp|O15205|UBD_HUMAN     -----LFLVESGDEAKRHLLQVRRSSS--VAQVKAMIETKTGIIP-ETQIV 43
sp|O32583|THIS_ECOLI    -----MQILFN-----DQAMQCAAGQT--VHELLEQLDQRQAGAA-----L 34
sp|O31617|THIS_BACSU    -----MLQLNG-----KDVKWKKDTG----TIQDLLAS-YQLEN-KIVIV 34
sp|O33246|PUP_MYCTU     -----MAQEQTKRGGG--GGDDDDIAGSTAAGQERREKL 32

```

```

sp|O31706|MOAD_BACSU      AVNESYVKENT-----SVSSGDTVAIIPPVSGG-- 77
sp|Q7A441|MOAD_STAAN     AVNEEFVQKSD-----FIQPNDTVALIPPVSGG-- 77
sp|P61956|SUMO2_HUMAN_16-95_ RFDGQPINETDTPAQLEMEDEDTIDVFQQQTGGVY 80
sp|P55854|SUMO3_HUMAN    RFDGQPINETDTPAQLEMEDEDTIDVFQQQTGG-- 78
sp|Q6EEV6|SUMO4_HUMAN    RFGGQPISGTDKPAQLEMEDEDTIDVFQQPTGGVY 79
sp|P05161|ISG15_HUMAN    TFEGKPLEDQLPLGEYGLKPLSTVFMNLRRLRGG-- 79
sp|Q64339|ISG15_MOUSE   SFEGRPMEDEKELLGEYGLKPQCTVIKHLRLRGG-- 79
sp|Q15843|NEDD8_HUMAN   IYSGKQMNDEKTAADYKILGGSVLHLVLRGG-- 76
sp|P29595|NEDD8_MOUSE   IYSGKQMNDEKTAADYKILGGSVLHLVLRGG-- 76
sp|Q71UE8|NEDD8_RAT     IYSGKQMNDEKTAADYKILGGSVLHLVLRGG-- 76
sp|P63072|UBD_MOUSE     NCNGKKLEDGKIMADYNIKSGSLLFLTTHCTGG-- 76
sp|Q921A3|UBD_RAT       NCNGKRLEDGKIMADYNIKSGSLLFLTAHCIGG-- 76
sp|O15205|UBD_HUMAN     TCNGKRLEDGKMMADYGIRKGNLLFLACYCIGG-- 76
sp|O32583|THIS_ECOLI    AINQQIVPREQ-WAQHIVQDGDQILLFQVIAGG-- 66
sp|O31617|THIS_BACSU    ERNKEIIGKER-YHEVELCDRDVIEIVHFVGGG-- 66
sp|O33246|PUP_MYCTU     TEETDILLDEI--DDVLEENAEDFVRAYVQKGGQ- 64

```

Sequence Alignment

Conclusion of the Comparison

- Significant similarity between the same superfamily, while significant differences between different superfamilies
- the highly conservative sites !