

# 抗原制备方案设计

## 原核重组表达抗原方案设计

分析蛋白质性质和特点，根据实际应用目的制定抗原制备方案。

**1.基本性质分析:**访问专业数据库查看蛋白质修饰情况（糖基化、羟基化、磷酸化、乙酰化等）、结构域特点、剪切加工等情况，如果蛋白质有后加工，应该尽可能选择一个完整“成熟体”区段进行分析。

**2.跨膜区分析:**跨膜区分析的数据库很多，进行跨膜分析时只需要将蛋白整个序列直接拷贝进文本框，即可直接查看结果。

**3.同源性分析:**通过NCBI网站或者专业物种数据库分析其在同物种中与其他蛋白同源性，这就是所谓的特异性原则；同时要检索其与被免疫宿主的同源性，避免宿主“免疫耐受”。

**4.亲水性/免疫原性分析:**用DNASTAR软件进行分析预测。

**5.稀有密码子分析:**将整个基因序列输入原核表达宿主密码子分析数据库即可查看结果；连续出现两个或者两

个以上的稀有密码子就有可能使表达变得极为困难。

作为抗原使用的重组蛋白如果获得全长有困难，可以截取其中抗原性比较好的区段部分重组表达。

经以上几步分析之后，如果决定要截取其中片段表达，需要遵循以下原则：

1.选取某个目标蛋白“成熟体”区段序列作为表达序列。

2.避开跨膜蛋白的“膜内区”。

3.避开与本体其他蛋白和免疫宿主“同源区”。

4.选取亲水性和免疫原性都较好的区段。

5.选取的区段尽可能长（100氨基酸以上）。

符合以上原则的区段如果氨基酸稀有密码子太多或比较集中，需要对其密码子进行优化，优化到大肠杆菌比较偏爱的密码子；如果符合以上原则的区段比较短或者不集中，这时可考虑通过合成多肽制备抗原。

### 实战演示1

以人神经生长因子受体16[ P08138 (NGFR\_HUMAN) ]为例

读者如果需要自行实践操作练习，可以联系我们技术[tech@friendbio.com](mailto:tech@friendbio.com)索取资料信息和软件

人神经生长因子受体16（NGFR）氨基酸序列：

```
MGAGATGRAMDGPRLLLLLLLGVSLGGAKEACPTGLYTHSGECCACNLGEGVAQPCGANQTVCEPCLDSVTFS
DVVSATEPCKPCTECVGLQSMSAPCVEADDAVCRCA YGGYQDETTGRCEACRVCEAGSGLVFSCQDKQNTVCEE
CPDGTYSDEANHVDPCLPCTVCEDETERQLRECTRWADAEEIIPGRWITRSTPPEGS DSTAPSTQEPEAPPEQDLIA
STVAGVVTTVMGSSQPVVTRGTTDNLIPVYCSILAAVVVGLVAYIAFKRWNSCKQNKQGANSRPVNQTPPEGEK
LHSDSGISVDSQSLHDQPHTQTASGQALKGDGGLYSSLPKREEVEKLLNGSAGDTWRHLA GELGYQPEHIDSF
THEACPVRALLASWATQDSATLDALLAALRRIQRADLVESLCSESTATSPV
```

### 基本性质分析

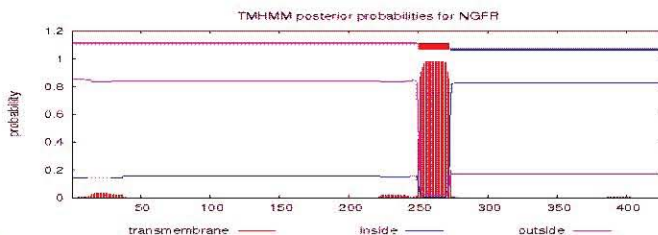
#### Molecule processing

Feature key	Position(s)	Description	Graphical view
Signal peptide	1 - 28		
Chain (PRO_000034591)	29 - 427	Tumor necrosis factor receptor superfamily member 16	

#### Topology

Feature key	Position(s)	Description	Graphical view
Topological domain	29 - 250	Extracellular Evidence: Sequence analysis	
Transmembrane	251 - 272	Helical Evidence: Sequence analysis	
Topological domain	273 - 427	Cytoplasmic Evidence: Sequence analysis	

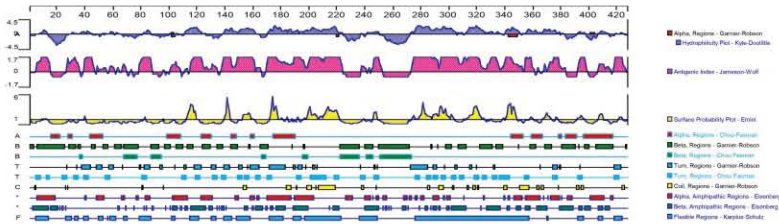
### 跨膜分析



从图中可以清楚看出该蛋白的信号肽、跨膜区、胞内区和胞外区。真核蛋白信号肽在原核中不能加工造成蛋白不表达，所以进行原核表达时要去掉；对于制备抗体一般优先选择胞外区段；跨膜区不但对抗体制备没有贡献，而且还难以在原核表达，因此原核表达制备抗原一定要去除信号肽和跨膜区。

蛋白跨膜区的预测，可以使用在线工具TMHMM Server 2.0进行。对于研究较多的蛋白在Uniprot网站数据库已经给予详细注释说明。

### 亲水性/免疫原性分析



上文已经阐述了作为抗原必须要有三性（亲水性、免疫原性和表面暴露性），通过软件分析可以分析拟选区段三性，从图中可以看出从120-380aa三性都比较好，可以作为优选区域。

### 同源性分析

与本物种内其它蛋白同源性比较

Description	Max score	Total score	Query cover	E value	Ident	Accession
tumor necrosis factor receptor superfamily member 16 precursor [Homo sapiens]	868	868	100%	0.0	100%	<a href="#">NP_002498.1</a>
unnamed protein product [Homo sapiens]	682	725	92%	0.0	100%	<a href="#">BAG64358.1</a>
Chain A, P75ntr Dd:rip2 Card	201	201	22%	4e-63	100%	<a href="#">ZNR3_A</a>
Chain A, P75ntr Dd:rhogdi	192	192	22%	8e-60	100%	<a href="#">ZNR0_A</a>
tumor necrosis factor receptor superfamily, member 11b (osteoprotegerin) [Homo sapiens]	66.6	66.6	38%	8e-11	30%	<a href="#">EAW91978.1</a>
osteoclastogenesis inhibitory factor (OCIF) [Homo sapiens]	65.5	65.5	35%	2e-10	31%	<a href="#">BAA25910.1</a>
osteoprotegerin [Homo sapiens]	65.5	65.5	35%	2e-10	31%	<a href="#">AAB53709.1</a>

上图从NCBI人的蛋白数据库分析比对对蛋白与其它蛋白的同源率和覆盖率。分析结果表明：除了NGFR本身和未命名蛋白外，与其他蛋白同源性低【特异性原则】。

与被免疫动物同源性比较

Description	Max score	Total score	Query cover	E value	Ident	Accession
nerve growth factor receptor [Mus musculus]	133	133	15%	1e-37	96%	<a href="#">ACR50956.1</a>
nerve growth factor receptor [Mus musculus]	91.7	91.7	13%	3e-22	95%	<a href="#">AHL58823.1</a>
unnamed protein product [Mus musculus]	748	748	95%	0.0	93%	<a href="#">BAG29775.1</a>
RecName: Full=Tumor necrosis factor receptor superfamily member 16; AltName: Full=Low affinity neurotrophin receptor p75NTR; AltName: Full=Low-affinity nerve growth factor receptor; Short=NGF receptor; AltName: CD_antigen=CD271; Flags: Precursor	746	746	95%	0.0	93%	<a href="#">Q9Z0W1.1</a>
tumor necrosis factor receptor superfamily member 16 precursor [Mus musculus]	749	749	99%	0.0	92%	<a href="#">NP_150086.2</a>

本蛋白拟制备单克隆抗体，免疫宿主为小鼠，下图从NCBI小鼠的蛋白数据库分析比对表明：与被免疫宿主同源性很高，但是别无选择。差异较多区段主要集中在氨基酸1-330aa【外源性原则】

### 方案确定

综上所述NGFR去掉信号肽，选择29-250aa区段进行原核表达比较利于小鼠单克隆抗体制备。将该区段翻译成大肠杆菌偏爱的密码子基因(序列如下)，全基因合成后插入pGEX-4T-1表达载体，转化大肠杆菌系列表达工程菌诱导表达。

```
AAGGAAGCATGCCCGACCGGCCTGTATACCCATAGTGGTGAATGCTGTAAAGCATGTAATCTGGGTGAAGGT
GTGGCCAGCCGTGTGGTGCCAATCAGACCGTTTGC GAACCGTGCCTGGATAGCGTGACCTTTAGTGATGTGG
TTAGTGCAACCGAACCGTGCAAACCGTGCACCGAATGCGTTGGCCTGCAGAGCATGAGTGCCCCGTGTGTTGA
AGCAGATGATGCAGTGTGCCGTTGCGCATATGGTTATTATCAGGATGAAACCACCGGTCGCTGCGAAGCCTGT
CGCGTTTGC GAAGCAGGTAGCGGCCTGGTGTAGCTGTCAGGATAAACAGAATACCGTGTGTGAAGAATGCC
CGGATGGCACCTATAGTGATGAAGCCAATCATGTGGACCCTTGCCTGCCGTGTACCGTTTGC GAGGATACCGA
ACGTCAGCTGCGCGAATGCACCCGCTGGGCAGATGCAGAATGTGAAGAAATTCGGGGTCGTTGGATTACCCGT
AGCACCCCGCCGGAAGGTAGCGATAGTACCGCACCGAGCACCCAGGAACCGGAAGCCCGCCTGAACAGGAT
CTGATTGCCAGCACCGTGGCAGGCGTTGTTACCACCGTTATGGGCAGCAGCCAGCCGGTTGTTACCCGTGGCA
CCACCGATAATTA
```

本蛋白去除信号肽之后仅有222个氨基酸，分子量偏小，一方面不利于纯化，另一方面，显然不符合抗原分子量/结构复杂原则，所以我们尽可能与大的融合标签（GST,MBP）一起表达。含有GST原核载体主要有pGEX-4T-1, pET-41a等，含有MBP标签的载体主要是pMAL系列载体（NEB公司）。选择此类标签还可促进目的蛋白/肽段的可溶性。

### 多肽抗原方案设计

难于获得困难、分子量小、同源性较高不易与其他区分的抗原，我们可以考虑采用多肽合成方法制备抗原；多肽抗原为半抗原，必须与载体偶联形成完全抗原才可以免疫动物。多肽抗原方案制备抗体核心是多肽选取与抗原表位的设计。多肽设计前首先对抗原蛋白进行系统分析，分析方法与原核重组表达蛋白相同，在此不做赘述。



## 多肽设计通用原则

### 外源性原则

设计多肽序列与被免疫宿主蛋白序列不同源（Blast检索）。

### 特异性原则

设计多肽序列与自身物种其他蛋白序列不同源（Blast检索）。

### 亲水性原则

蛋白质都是将亲水部分暴露在外，只有抗原更好的溶于水相才能更好地激发免疫反应，因此设计的多肽一定要亲水性好（通过DNassist，GCG，ProtScale等软件预测亲水性）。

### C末端性原则

末端肽段比内部肽段暴露性更好，更容易制备抗体，但膜蛋白C端疏水性比较强，一般选择N端多肽。

### 结构原则

序列不形成 $\alpha$ -helix、内部重复或接近重复段的序列。

### 氨基酸原则

序列内Pro(脯氨酸)一般不超过2个；同时尽量包括MHC II类分子偏好的Asp, Tyr, Phe等氨基酸，避免容易发生糖基化的氨基酸（如Asn、Ser、Thr和Hyp）和磷酸化的氨基酸（Ser、Thr和Tyr）。

### 连接臂原则

一般在肽链上加一个联接臂便于多肽与载体偶联，最为常用的是加Cys的-SH，加在对抗体产生影响最小的那一端；选择N末端多肽序列的加在多肽C端；选择C末端序列-SH加在N端，选择中间肽段加在对抗体制备影响最小的那一端。

### 长度原则

抗原表位一般是6~8个氨基酸组成，多肽做抗原的时候一般设计15~20个氨基酸，一方面为防止蛋白质水解，另一方面为了保证表位覆盖率。链太长则易造成链内二级结构，同时增加合成的成本。

### 纯度原则

用于制备抗体的多肽，85%以上的纯度就足够了，纯度越高合成的成本越高。

目前生物信息学在线软件比较多，作为多肽预测的辅助网站主要有：

<http://www.cbs.dtu.dk/services/BepiPred/>

<http://www.epitope-informatics.com/Links.htm>

<http://bio.dfci.harvard.edu/Tools/index.html>

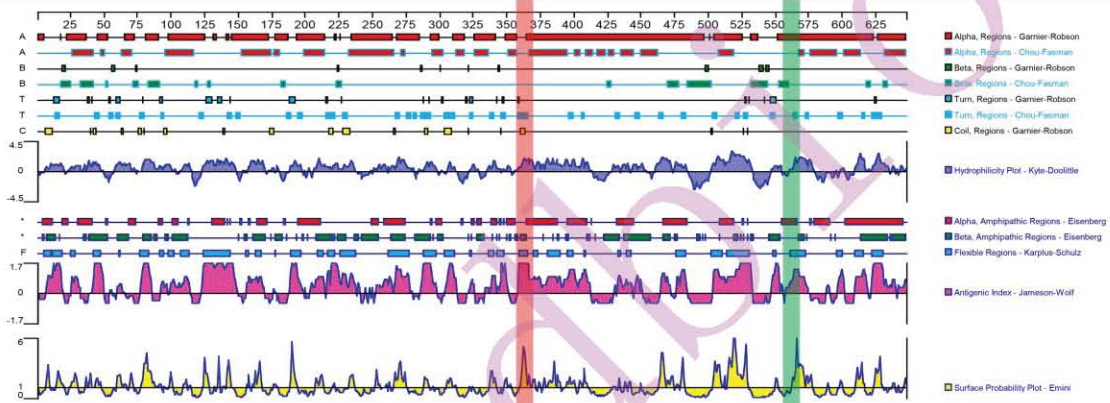
以上数据库只能作为一个辅助参考，最终还是要以实验验证数据为主，商业化的多肽合成公司一般都有其专用数据库和多肽设计经验，设计成功率更高。科研工作者可以请求商业公司协助设计，福因德生物团队接到委托一般会提供三组多肽设计，成功几率基本可以达到98%以上。

## 实战演示2

以下两个项目实例，根据多肽设计原则，结合福因德生物技术团队多年多肽设计经验，每组设计两条肽，其中都有一条多肽成功制备出抗体（绿色为成功肽段，红色为失败肽段）。读者可以根据以上原则认真仔细揣摩。

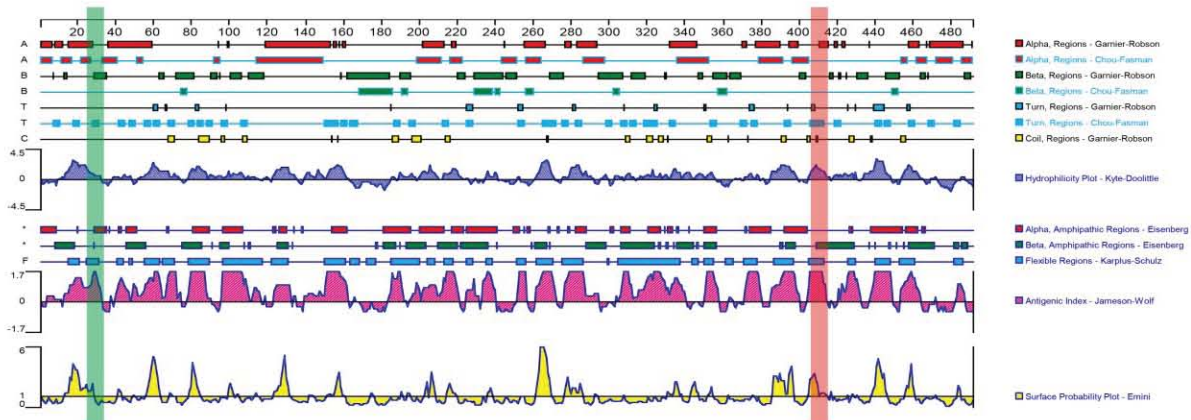
蛋白序列1

MAFFQQPARPIAEGQYTQTIYTLIKEQKFAEAIQHLQYQLQNVPE\$RAALSLLGYCYYYTGQYDMASQMYEQLVTL  
LYPSNEDYKLYYAQSLYKGGMYPEASKAVVKVEGHQKAVTLLVACSYEQDDLTCRRQLDKCAPEDPDTMVN  
TGCIMFKEGKFEAARQKFNDAVQALGYQPELLYNIALCYKTKQFGPALKHLAEIIEKAVREHPELSVGSNTDGME  
VRSVGN\$SQT\$K\$ETALIEAFNLKAAIEYTMKNVEAAKEALTDMP\$PRAEELDPVTLHNSALINMDS\$DPTGGFKLNF  
LLQSPFPFETFANLLLLYCKPSHGFDLAADVLAENPQYAGKLLSPDLYDYLQAAI**GRYKSP\$EAFRRFDEL**ATRH  
VEQLRRLTKIQDARIARDNDAIKRAINEYDEALEAYIPGLMAMASIYWDMELYSNVEKIFRQSAEFCSEHEVWKL  
NVAHTFFMQDNHYKEAIRYYEPVVKKNADNLLGVT\$AIVLANL\$CVSYIMTSQNEEAEELMRKVEKEEERS\$SMQDP  
DKPCFHLCIINLVIGTLYCAKGN\$YEFVSR**IKSLEPYDKKLETDTWY**YAKRCFLALIENLAKHMIVLKDSSFTEIMA  
FLNEAEKHGKDIRVVFNEGKHQ\$RTIASEARMLK\$KMFLLKLRD



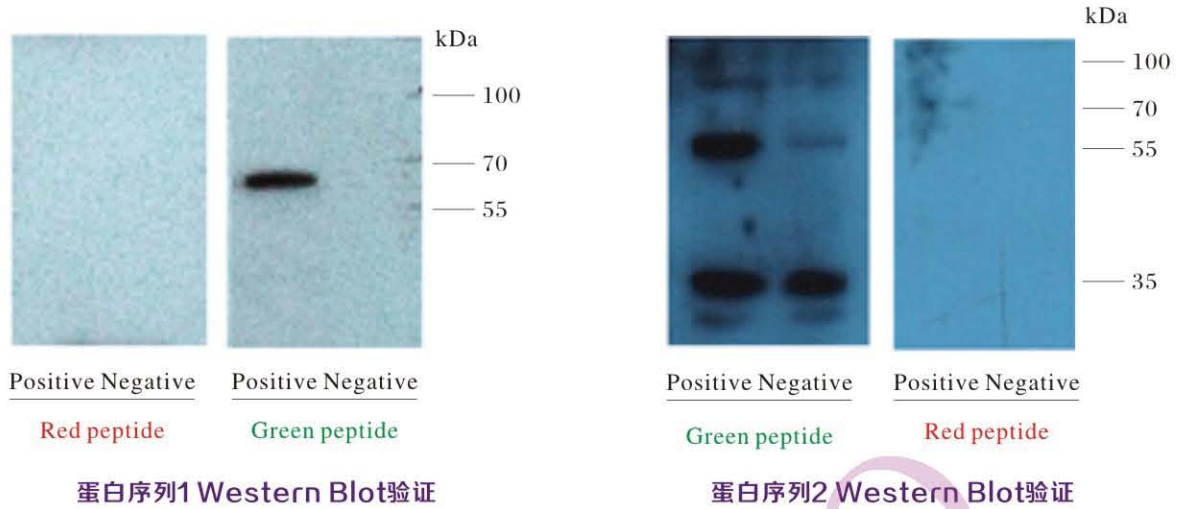
蛋白序列2

MAKLIK\$NVGASLR**ARTHDEDDTMMKQKGAT**GVFRNLAFADADDNLVSTSARAMATSESTKKNFFGG\$SQDNIAS  
IDVTPRSRDAGNGASSWAHADLPTSASKRVG\$TGSASTPVKSATFARTASAQKRAKNATAIQEISAFEHEHAVMDE  
MSGSEDGERPAGLVSGGSAIGATTSTTVIAVRSVARGPSITQQVSTSGSVRAWEEEVKRLIASGRHEDAVRWVAPS  
DGIIRCTVRRVKNFLGHTLAYQLFLD\$GDTFVLAARKRKK\$KASNFVLSTSQEDLGKDS\$DHCIAKLRANFVGTEYG  
LV\$RTGGHISG\$M\$DIDGGAQ\$GGKLAPPAEPFSREEIAVHYKQTALTA\$KGGPRTMLVATPLPEVSWAPSAADGSDS  
LANCLEAARRRELSRMERQL**CMLATRPPEWDP\$SLKAY**TLDFHGRIRASSVKNFQLVHWDHNTDRKGS\$DLVLQF  
GKIDENTDDFALDFTYPLSLQKAF\$AIALASTDTKLCYAL





实验验证结果 (Western Blot)



多肽设计没有可生搬硬套的公式，所有预测性软件也仅仅只能做参考。福因德生物技术团队经过数千例的项目经验积累，与同行业专家不断学习交流，结合最前沿的生物信息学分析技术、虚拟筛选技术和多肽指纹技术等，为国内外医药研发企业与院校提供千例多肽设计和合成服务，每个项目设计三组多肽基本上可以保证项目98%成功。

多肽抗原还是重组蛋白抗原选择问题

初次接触抗体制备的科研人员往往为选择多肽还是重组蛋白作为抗原犹豫不决。一般来说，能用天然蛋白的就不用重组蛋白，能用重组蛋白的就不用多肽，但是多肽在解决特异性和抗原难以获得等方面优势凸显，在明确获知（已发表权威文献或经实验验证）抗原表位的情况下可以优先选择使用多肽抗原。多肽抗原产生的抗体基本上都是针对单一表位的，重组蛋白产生的抗体一般都是针对多个表位的，所以选择多肽抗原检测内源性标本要么不成功，一旦成功其背景较为干净；而重组蛋白产生的抗体往往就没有这么干脆利索。重组蛋白抗原制备抗体的主要优势是其整条肽链上包含着多个抗原表位，所有表位均能激发抗体产生，但正是由于其表位众多，其中难免有一些表位序列与内源性标本中其他蛋白同源或相近似，带来诸多背景问题，有时候甚至“喧宾夺主”，其背景信号干扰甚至超越目标信号。多肽必须与载体蛋白偶联才能成为完全抗原，偶联也是抗体制备至关重要的环节，一般正常偶联方法每个载体蛋白上偶联5~10个多肽，福因德技术团队通过支架偶联技术可以在一个载体蛋白上偶联多达100个多肽，在激发抗体方面更有优势；但并不是每个多肽都适合超高偶联比，由于每个肽理化性质不同，偶联太多分子表面电荷改变会造成蛋白质聚集产生沉淀。除此以外，多肽作为抗原要想产生高品质抗体还必须在佐剂类型和免疫方法上做出更多优化，为此福因德生物研发出多肽抗原专用佐剂，与普通佐剂相比，至少可以提高抗体滴度10倍。

