

单个蛋白质分析案例

Version:1.0

Date: 2012-6-30

注：因技术路线保密的需要，采用的软件用 toolA, toolB,toolC,...替代，正式报告将注明采用的软件。****表示数值不显示。

目录

蛋白质基因信息获取.....	2
蛋白质基本性质分析.....	2
蛋白质二级结构及结构域分析.....	2
信号肽及酶切位点预测.....	3
跨膜区预测.....	3
蛋白质亚细胞定位.....	4
蛋白质酶切位点预测.....	4
蛋白质相互作用分析.....	4
蛋白质翻译后修饰.....	5
乙酰化.....	5
磷酸化.....	6
糖基化.....	6
蛋白质功能分析.....	7
关联分析.....	7
突变对蛋白功能的影响.....	7
结果讨论.....	7
参考文献.....	8

蛋白质基因信息获取

根据给定的[demo]蛋白的 access number 到对应的数据库中查询基因信息和下载序列信息。

蛋白质基本性质分析

该蛋白质为酸性蛋白质，序列由**个氨基酸组成，分子量(molecular weight): ****，等电点(pI): ****，负电荷残基总数(Asp+Glu): **，正电荷残基总数(Arg+Lys): **，分子式: ****，原子总数: ****，半衰期: ****，不稳定系数(instability index): ****，为稳定蛋白，脂肪系数(Aliphatic index): ****，Grand average of hydropathicity(GRAVY)亲水性评估: ****。

蛋白质二级结构及结构域分析

该蛋白质包含多种二级结构，其中以 Helix 居多。Helix 区域总长**** aa，占原始序列总长度的****%。蛋白质二级结构分布图形如图 2.1 所示。同时通过生物信息学网站 toolA 预测得该蛋白质为非球状蛋白质。该蛋白质包括两个 Domain 及两个 Calcium binding 区域，如图 2.2 所示。

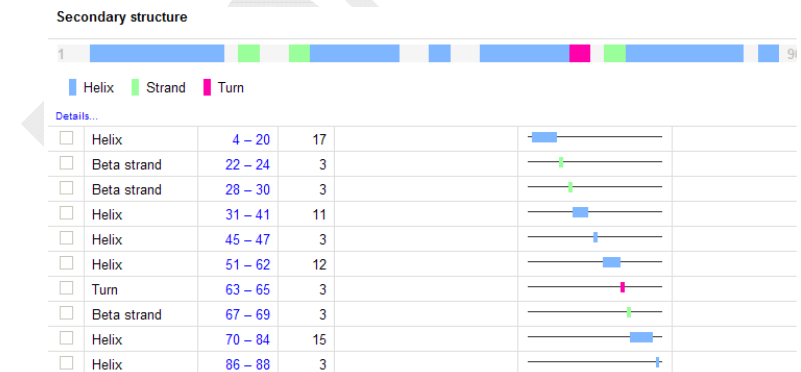


图 2.1 [DEMO]蛋白质二级结构分布

Molecule processing

<input type="checkbox"/> Chain	1 - 90	90	Protein S100-A6		PRO_0000143984
--------------------------------	--------	----	-----------------	--	----------------

Regions

<input type="checkbox"/> Domain	12 - 47	36	EF-hand 1		
<input type="checkbox"/> Domain	48 - 83	36	EF-hand 2		
<input type="checkbox"/> Calcium binding	20 - 33	14	1 [Ref.17]		
<input type="checkbox"/> Calcium binding	61 - 72	12	2 [Ref.17]		

图 2.2 [DEMO]蛋白质结构域分布

信号肽及酶切位点预测

应用在线工具 [toolB](#) 预测蛋白信号肽及其剪切位点。结果表明该蛋白质为非信号肽类型的分泌蛋白。采用在线工具 [toolC](#) 预测该蛋白质可能为非典型分泌蛋白(非信号肽引发的)。采用在线工具 [toolD](#) 预测该蛋白质在 pos. ****氨基酸区域可能为亮氨酸富集的核输出信号****。

跨膜区预测

使用在线工具 [toolE](#) 预测该蛋白质的跨膜区域。预测结果中该蛋白质并不存在跨膜区域，预测结果如图 4.1 所示。

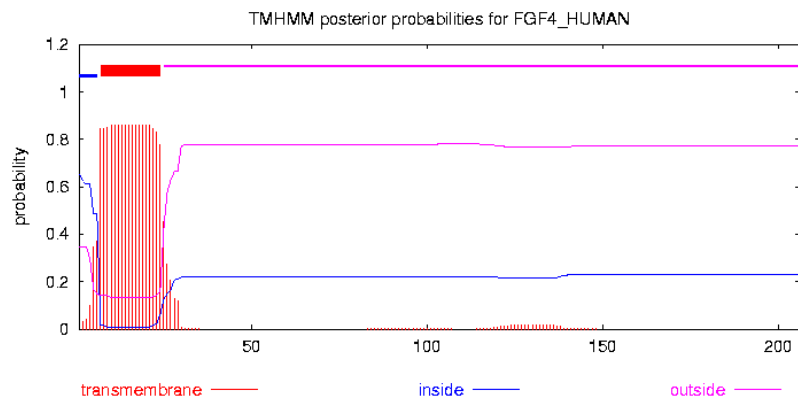


图 4.1 [DEMO]蛋白质跨膜区预测结果

蛋白质亚细胞定位

根据文献可得该蛋白质亚细胞定位如下：Nucleus envelope, Cytoplasm, cell membrane。为了更近一步确定该蛋白质亚细胞定位主要集中在哪里，这里使用工具 [ToolF](#) 进行亚细胞定位预测，预测结果如下：

Cytoplasmic	Nuclear	Cytoskeletal	Peroxisomal	Mitochondrial
****%	****%	****%	****%	****%

蛋白质酶切位点预测

通过工具 [ToolG](#) 进行蛋白质酶切位点预测，结果发现如下酶对该蛋白有酶切作用。

Arg-C proteinase	Trypsin	Asp-N endopeptidase + N-terminal Glu
CNBr	NTCB	chymotrypsin-low specificity
Clostripain	Formic acid	Glutamyl endopeptidase
LysC	LysN	Chymotrypsin-high specificity
Pepsin(pH1.3)	Pepsin(pH>2)	Proteinase K
Staphylococcal peptidase I	Thermolysin	Asp-N endopeptidase

蛋白质相互作用分析

通过 [WebA](#) 网站来挖掘该蛋白质的相互作用信息。结果发现有****个蛋白质与该蛋白质存在直接的互作关系，其中有**个互作关系通过实验及文本挖掘证实，剩余**个通过文本挖掘证实；**个互作关系为 Binding 关系，剩余一个仅仅通过文本挖掘说明存在互作，但并未说明具体的互作关系。[WebA](#) 数据库中的互作关系如图 7.1 所示。

蛋白质行使功能不仅仅与自身特性有关，同时还与相互作用的蛋白质及互作关系相关。这里应用 [DAVID](#) 分析工具分析[DEMO]蛋白及其互作蛋白的功能特性，给出 P value

<= 0.01 显著的功能注释。其中在 acetylation 注释中存在 8 个蛋白质，在 cytoplasm 注释中存在 7 个蛋白质，[DEMO]蛋白被注释到 GO:0048667: cell morphogenesis involved in neuron differentiation.

WebA 数据库中仅仅给出直接互作的关系，并没有给出[DEMO]蛋白上下游的互作关系。7]。

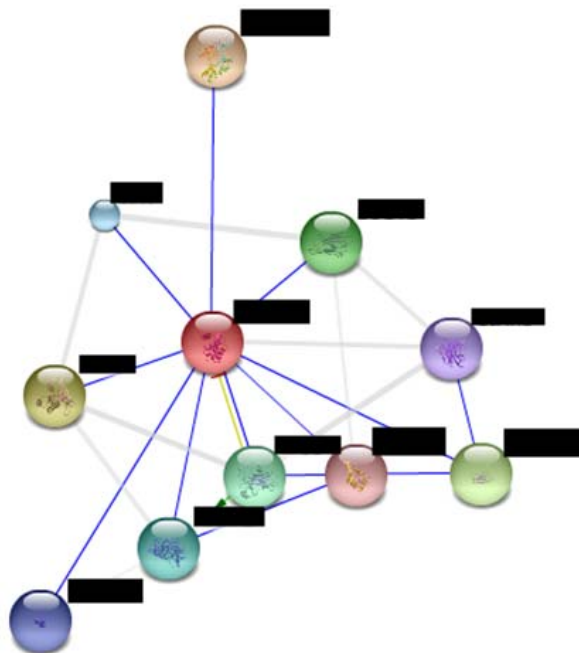


图 7.1 STRING 数据库[DEMO]蛋白互作关系

(注：黑框是后面处理将名字挡了)

蛋白质翻译后修饰

乙酰化

由已经发布的信息可得，该蛋白可能存在乙酰化(Acetylation, pos. 40)。

磷酸化

通过工具 ToolH 进行激酶特异的磷酸化位点预测[8]，结果发现在 pos. 30 位点存在 PKC 磷酸化位点。使用 NetPhos 工具进行磷酸化位点预测分析，结果共得到三个磷酸化位点：Serine (pos. 20, pos. 30), Tyrosine (pos. 73)。磷酸化位点预测结果如图 8.2.1 所示。

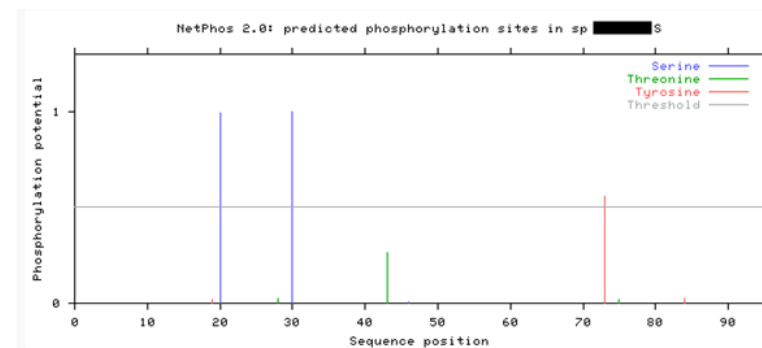


图 8.2.1 [DEMO]磷酸化位点预测结果

糖基化

通过工具 ToolH 进行糖基化位点预测，结果并没有发现糖基化位点；通过工具 NetGlycate 进行赖氨酸糖化作用预测[9]，在 pos. 32, pos. 64 位点可能存在赖氨酸糖化作用。赖氨酸糖化作用预测结果如图 8.3.1 所示。

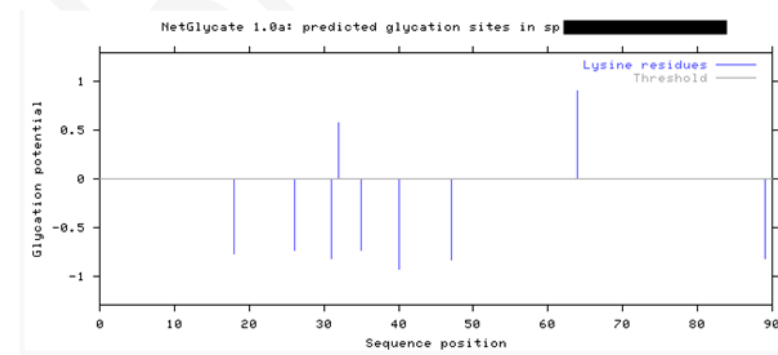


图 8.3.1 [DEMO]赖氨酸糖化作用预测结果

蛋白质功能分析

根据文献可得该蛋白质可能的功能包括：钙传感器、有助于细胞内钙信号、与其它蛋白互作并且间接的重组肌动蛋白细胞骨架及在细胞能动性上发挥作用、钙离子的 binding。该蛋白质属于**** family, ****蛋白可能调控细胞周期及细胞分化，同时可能刺激 Ca²⁺ 依赖的胰岛素释放，刺激催乳素分泌及胞外分泌。该蛋白质在 Gene Ontology(GO)数据库中的注释信息如下：

Biological process	GO:0007409 axonogenesis
	GO:0048146 positive regulation of fibroblast proliferation
	GO:0007165 signal transduction
Molecular function	GO:0048154 S100 beta binding
	GO:0005509 calcium ion binding

关联分析

根据 Genetic Association Database 中的信息可得，DEMO 基因可能与 Dermatitis, Atopic 及 Alzheimer Disease 等疾病相关。

突变对蛋白功能的影响

根据已发表的信息可得在编码该蛋白质的区域存在 4 个 natural variations, dbSNP ID 号为：rs11974, rs1802581, rs1802582, rs2228293。用 SIFT 工具预测突变导致的氨基酸改变是否会影响蛋白质的功能，在结果中 rs1802581, rs1802582, rs2228293 这三个位点的变异可能会影响到蛋白质的功能。

结果讨论

蛋白质疏水性通常根据蛋白质的 GRAVY 值来评价。GRAVY 值的范围在 2 到-2 之间，正值表示此蛋白质疏水，负值表示此蛋白质亲水。.....以下省略

参考文献

- Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A.; **PROTEIN IDENTIFICATION AND ANALYSIS TOOLS ON THE EXPASY SERVER;** (In) John M. Walker (ed): **The Proteomics Protocols Handbook, Humana Press (2005).** pp. 571-607
- <http://www.uniprot.org/uniprot/P06703>
- Jannick Dyrlov Bendtsen, Henrik Nielsen, Gunnar von Heijne and Søren Brunak. **Improved prediction of signal peptides: SignalP 3.0.** J. Mol. Biol., 340:783-795, 2004.
- J. Dyrlov Bendtsen, L. Juhl Jensen, N. Blom, G. von Heijne and S. Brunak **Feature based prediction of non-classical and leaderless protein secretion.** Protein Eng. Des. Sel., 17(4):349-356, 2004
- Tanja la Cour, Lars Kiemer, Anne M. gaard, Ramneek Gupta, Karen Skriver and Sen Brunak **Analysis and prediction of leucine-rich nuclear export signals.** Protein Eng. Des. Sel., 17(6):527-36, 2004.
- Mandinova, A., Atar, D., Schafer, B. W., Spiess, M., Aebi, U., Heizmann, C. W. **Distinct subcellular localization of calcium binding S100 proteins in human smooth muscle cells and their relocation in response to rises in intracellular calcium.** J. Cell Sci. 111: 2043-2054, 1998. [PubMed: [9645951](https://pubmed.ncbi.nlm.nih.gov/9645951/), [related citations](#)]
- Takata M, Shimamoto S, Yamaguchi F, Tokuda M, Tokumitsu H, Kobayashi R. **Regulation of nuclear localization signal-importin α interaction by Ca²⁺/S100A6.** FEBS letters 2010 Nov 19;584(22):4517-23 PMID: 20965181
- Blom N, Sicheritz-Ponten T, Gupta R, Gammeltoft S, Brunak S. **Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence Proteomics.** 2004 Jun;4(6):1633-49. Review.
- Morten Bo Johansen, Lars Kiemer and Søren Brunak **Analysis and prediction of mammalian protein glycation.** Glycobiology, 16:844-853, 2006.