

# NIF3 类超家族蛋白

盖其静 张耀洲\*

(浙江理工大学生命科学学院生物化学研究所, 杭州 310018)

**摘要** NIF3 (NGG1p interacting factor 3)类超家族蛋白是一个含有 NIF3 结构域的超家族, 从细菌到哺乳动物存在着较高的保守性。编码该家族蛋白质的 mRNA 广泛存在于生物的各种组织中。该家族蛋白属于  $\alpha/\beta$  蛋白, 该类蛋白质通过自身形成二聚体或与其他蛋白质结合形成复合物发挥生物学功能。现已证明该家族蛋白多与基因的转录有关, 在一些哺乳动物中还在神经分化中起关键作用。最近研究表明它还和一些人类疾病有关。虽然该家族蛋白数量众多, 广泛存在, 但多数为假想蛋白, 对其相关功能的研究较少, 被归属为“未知功能蛋白质”。也正是因为该家族蛋白数量众多, 广泛存在, 功能未知, 它们已经引起了人们越来越大的研究兴趣。

**关键词** NIF3; 神经分化; 转录; 未知功能

NIF3 (NGG1p interacting factor 3)类超家族蛋白多数属于保守的假想蛋白, 是根据它们的 DNA 序列推测出来的。最初, 由于 NIF3 类超家族蛋白的保守结构域的功能不清楚, 只是发现了一些结构相似非常保守的蛋白质, 所以把它们称为 DUF34(domain unknown function 34)<sup>[1]</sup>。“NIF3 类超家族蛋白”的命名起源于酵母中的 NIF3。酵母的 NIF3 最初是通过酵母双杂交系统得到的, 由于它是与 NGG1p 相互作用而被筛选出来的, 所以被定义为 NGG1p 作用因子 3<sup>[2]</sup>。

目前的研究表明 NIF3 类超家族蛋白是一种胞质蛋白, 存在于线粒体和胞液中<sup>[3]</sup>, 并且广泛存在于各种生物中, 大肠杆菌中的 ybgI、枯草芽孢杆菌中的 yqfo、酵母中的 NIF3 以及小鼠和人中的 NIF3L1 等都属于这个家族<sup>[4]</sup>。而且这个家族的蛋白质在进化中从细菌到哺乳动物有着很高的保守性, 特别是序列的两端保守性更高。例如: 人中的 NIF3L1 的两侧与果蝇中的 anon-35F/36A 同源率为 41%, 与假想蛋白 YGL221c 和酵母的 NIF3 同源率为 37%, 与低等生物尿素分解尿素原体保守的假想蛋白和天蓝色链霉菌种的假想蛋白 SCC30.09c 在两翼表现出了大约 25%~30% 的同源性<sup>[5]</sup>。目前该超家族蛋白被研究和出现的频率越来越高, 并且被证明该类蛋白质和人类的 Williams-Beuren 综合征(Williams-Beuren syndrome, WBS)、青少年肌萎缩侧索硬化硬化症(juvenile amyotrophic lateral sclerosis, ALS2)等疾病有关, 它们已经引起了人们的极大兴趣, 受关注程度在保守的未

知功能蛋白质中位居前 10 位<sup>[4]</sup>。

## 1 NIF3 类超家族蛋白的结构

NIF3 类超家族蛋白是  $\alpha/\beta$  蛋白, 主要是平行的  $\beta$  片层结构。它们一般含有 N 端和 C 端两个作用的结构域, 可以和二价离子结合<sup>[6]</sup>。这个家族在进化上具有较高的保守性, 我们从一些该家族蛋白的多序列比对可以总结出它们的一级结构的保守性主要集中在氨基酸序列的两侧, 有的 NIF3 类超家族蛋白成员的 NIF3 结构域中间又插入了几十个氨基酸残基, 由此我们也可以把该家族蛋白分为两类: 一类为只含有 NIF3 结构域的蛋白质, 大肠杆菌的 ybgI、酵母的 NIF3、果蝇的 anon-35F/36A 等都属于这一类; 另一类是在 NIF3 结构域中间插入了其他结构(一般为 PII 类结构)的蛋白质, 葡萄球菌的 SA1388、人的 NIF3L1 和小鼠的 NIF3L1 等都属于这一类。下面我们就以大肠杆菌的 ybgI(登录号为 P0AFP6)和葡萄球菌的 SA1388(登录号为 2NYD\_B)这两个蛋白质为例介绍一下这两类该家族蛋白质的结构。

大肠杆菌的 ybgI 只含有一个 NIF3 结构域, 它形成两个相连接的相似的  $\alpha/\beta$  结构, 而且都是三层的三明治结构(alpha-beta-alpha)。第一个结构域(N 端)含有 5 个  $\beta$  折叠组成的片层, 在它们的一侧有 2 个  $\alpha$  螺

收稿日期: 2007-05-28 接受日期: 2007-07-02

国家高技术研究发展计划(863 计划)(No.2005AA206120)和国家重点基础研究发展规划(973 计划)(No.2005CB121006)资助

\* 通讯作者。Tel/Fax: 0571-86843198, E-mail: yaozhou@chinagene.com

旋, 另一侧有 3 个  $\alpha$  螺旋。其中, 3 个  $\alpha$  螺旋中的 2 个几乎与  $\beta$  片层平行, 第 3 个  $\alpha$  螺旋较短几乎与  $\beta$  片层垂直并指向第 2 个结构域。第 2 个结构域(C 端)含有 6 个  $\beta$  折叠,  $\beta$  片层的两侧各有 2 个  $\alpha$  螺旋结构, 并有 1 个短的  $\alpha$  螺旋回到了第 1 个结构域<sup>[7]</sup>(图 1A)。而葡萄球菌的 SA1388 含有 2 个 NIF3 结构域, 中间插入一段 PII 类结构。2 个 NIF3 结构域分别具有 5 个  $\beta$  折叠, 并且在  $\beta$  片层两侧各有 2 个  $\alpha$  螺旋结构, 也像其他规则的 NIF3 类蛋白一样具有典型三明治型 NIF3 结构的特点<sup>[8]</sup>(图 1B)。我们将大肠杆菌的 ybgI 和葡萄球菌的 SA1388 这两个蛋白质的氨基酸序列提交到在线软件 SWISS-MODEL(<http://swissmodel.expasy.org/SWISS-MODEL.html>)中对其高级结构进行预测模拟做图, 并用 CE 软件将它们的高级结构进行比对<sup>[9-11]</sup>, 发现它们的三级结构非常相似, 同源性

更高(图 1C)。

在对 ybgI 和 SA1388 这两种蛋白质的四级结构分析中发现, 它们都能自身形成不对称的同源二聚体, 再由 3 个不对称的二聚体形成 1 个环状的四级结构。不同的是, ybgI 形成的环状结构是管状的<sup>[7]</sup>, 而 SA1388 形成的圆形环结构的上下还有多出来的 PII 类结构形成的盖子将管状的开口部分封住, 使其他大分子物质不能够结合到环的内部<sup>[8]</sup>(图 2)。另外, 它们的单体都存在可以和 2 个二价离子结合的活性位点, 并且都位于单体的 2 个 NIF3 结构域之间和环状六聚体的内侧。不同的是 ybgI 的活性位点, 可以和铁、锌等二价离子结合, 而 SA1388 只能和锌离子结合。这个可以和二价金属离子结合的活性位点在 NIF3 类超家族蛋白中非常保守, 主要含有 5 个组氨酸、1 个谷氨酸和 1 个天冬氨酸。由于许多蛋白质

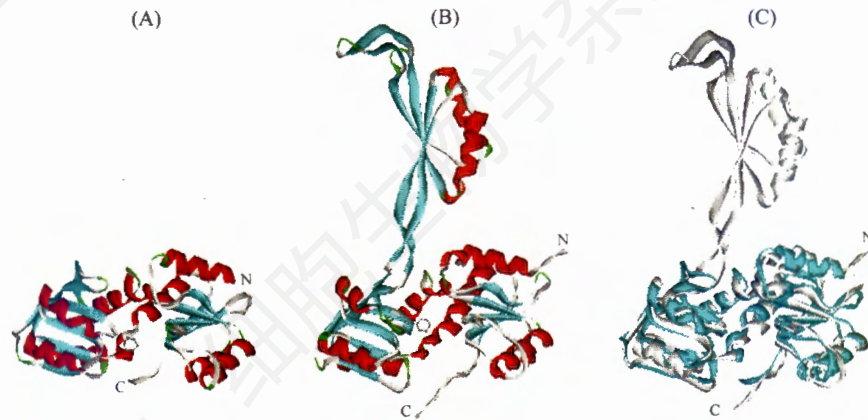


图1 大肠杆菌的 ybgI 和葡萄球菌的 SA1388 三级结构

图中蛋白质的高级结构均由在线软件 SWISS-MODEL 预测得出。A: ybgI 的三级结构; B: SA1388 的三级结构; C: 将 A、B 两种蛋白质用 CE 软件进行高级结构比对的模拟图。

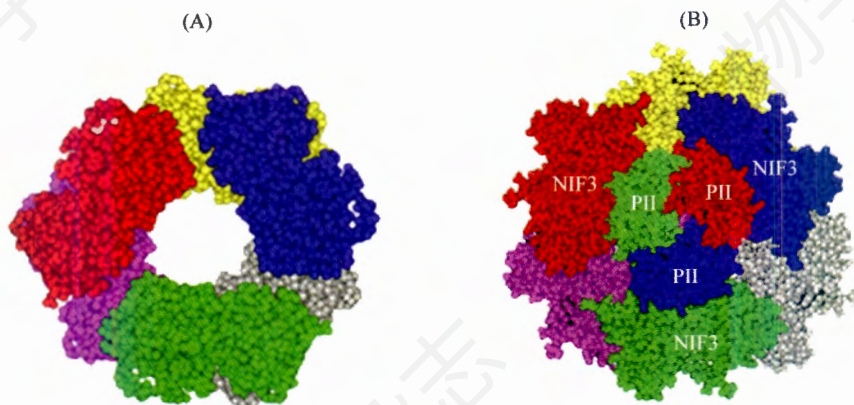


图2 大肠杆菌 ybgI 和葡萄球菌 SA1388 的六聚体比较

图中展示了 ybgI 和 SA1388 两个蛋白质的六聚体图, 每一个单体都由不同颜色标出, 单体的高级结构均用在线软件 SWISS-MODEL 预测。A: ybgI 的六聚体示意图; B: SA1388 的六聚体示意图。

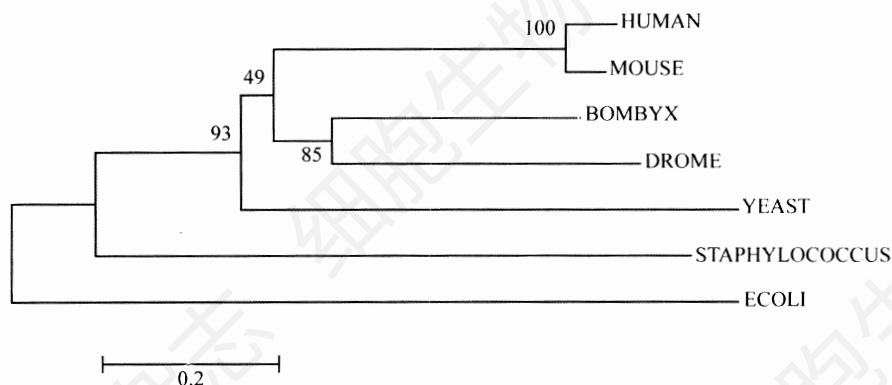


图3 不同物种间的NIF3类超家族蛋白氨基酸序列的分子进化图

HUMAN: *Homo sapien* NIF3-like protein 1(登录号: Q9GZT8); MOUSE: *Mus musculus* NIF3-like protein 1(登录号: Q9EQ80); BOMBYX: *Bombyx mori* NIF3-like protein (登录号: ABI73988.1); DROME: *Drosophila melanogaster* (登录号: AAL39701.2); YEAST: *Saccharomyces cerevisiae* Nif3p(登录号: NP\_011294); ECOLI: *Escherichia coli* UPF0135 protein ybgI (登录号: P0AFP6); STAPHYLOCOCCUS: *Staphylococcus Aureus* Hypothetical Protein Sa1388(登录号: 2NYD\_B)。

的催化位点都是以谷氨酸作为桥连接合锌或锌和铁、锰或镁离子<sup>[12]</sup>的,所以推测NIF3类超家族蛋白的二价离子结合的活性位点是1个催化位点。另外在SA1388的中发现了1个未知的内源性配基,这可能与PII类结构的功能有关。

将人、小鼠、家蚕、果蝇、酵母、葡萄球菌、大肠杆菌的NIF3类蛋白用MEGA3软件构建了氨基酸序列的分子进化树(图3)<sup>[13]</sup>,发现虽然葡萄球菌的SA1388和哺乳动物(人、小鼠)的NIF3类蛋白都存在一个PII类结构,但是由于物种间的差异比较大,在进化上它们的距离比较远,也可以看出相近物种之间的进化距离比较近,符合物种的进化规律。

## 2 NIF3类超家族蛋白的功能

由于NIF3类超家族蛋白是近年来新发现的一种具有保守结构域的蛋白质,而且大多数是根据生物信息学推测出来的,并没有真正得到蛋白质,所以目前对NIF3类超家族蛋白的功能研究很少而且不够深入,作用机制都还不是很清楚。

### 2.1 原核生物中NIF3类蛋白的功能

目前的研究表明在低等生物中大肠杆菌的NIF3类超家族蛋白成员ybgI可结合到大肠杆菌的一种操作子上,但该操作子中的ybgI在其他细菌中并不保守<sup>[14]</sup>。在丝裂霉素C的DNA损伤引起的遗传毒性压力下,大肠杆菌ybgI的表达水平明显增加<sup>[15]</sup>,再结合该蛋白质的四级结构推测它可能参与到DNA代谢的过程,因为参与DNA代谢过程的很多蛋白质都能形成与大肠杆菌ybgI相似的环状结构<sup>[14,16,17]</sup>。而葡

萄球菌中的SA1388的四级结构不是管状的,但是在SA1388中却把NIF3与PII类结构整合到了一起,PII类家族蛋白是一种信号转导蛋白,在微生物中参与氮的调节,它们通过形成三聚体与配基结合发挥生物学功能<sup>[18-20]</sup>。虽然其功能还不了解,但可以根据结构推测在葡萄球菌中可能是通过PII类结构来调节NIF3类结构的功能。

### 2.2 真核生物中NIF3类蛋白的功能

**2.2.1 酵母的NIF3的功能** 在真核单细胞生物中我们以酵母为例介绍一下NIF3的功能。酵母中的NIF3能够和NGG1p的氨基酸末端相互作用,这种作用可以降低GAL4p在葡萄糖富集培养基中的转录活性。NGG1p含有373个氨基酸,是组氨酸乙酰转移酶复合物组成之一ADA的一个亚基。对酵母中的转录作用物的基因研究证明NGG1p/ADA3p复合物作用于转录的基因,可以阻止由于葡萄糖效应而引发的转录快速增加,而这一效应受到NIF3的抑制<sup>[2,21]</sup>。所以说酵母的NIF3与转录有关。在酿酒酵母中鉴定完全蛋白质-蛋白质间相互作用的高产量的酵母双杂交筛选系统揭示NIF3可以与一种ras类GTPase TEM1(YML064C)和核输入/输出蛋白SRP1(YNL189W)相互作用,而这两种蛋白质都是细胞彻底结束有丝分裂所必需的,由此可以总结出NIF3可能与细胞周期的调控有关<sup>[22]</sup>。目前对酵母中的NIF3的研究大体上就这些,它的具体功能作用机制还有待继续研究。

**2.2.2 哺乳动物中NIF3L1的功能** 本文以人和小鼠的NIF3L1为例介绍一下哺乳动物中NIF3类超家族蛋白的功能。人和小鼠的NIF3L1有极高的同源

性,研究表明编码 NIF3L1 的 mRNA 在小鼠的各个组织中广泛存在。通过对未出生小鼠和其胚胎干细胞的 RNA 的 RT-PCR 分析表明 *nif3l1* 基因在整个胚胎发育期都有表达。小鼠的 *nif3l1* 基因定位在染色体 1, C 区。人的 *nif3l1* 基因含有 7 个外显子形成 1.75 kb 的基因组 DNA 被定位于染色体 2q33 上。一个含有绿色荧光蛋白和人的 *nif3l1* 开放阅读框的融合蛋白的定位表明,它多存在于细胞质中<sup>[5]</sup>。

对小鼠细胞中的 NIF3L1 的研究表明它的 C 端可以与 Trip15/CSN2 的 N 端相互作用,作为转录共抑制因子在 P19 胚胎癌细胞的神经分化中起到作用<sup>[23]</sup>。Trip15/CSN2 是一种转录共阻遏子,在神经分化过程中起到关键的调控作用<sup>[24]</sup>。虽然 NIF3 类超家族蛋白被归属为胞质蛋白,但在细胞中过表达 Trip15/CSN2,会发现 NIF3L1 在细胞核中也大量存在<sup>[23]</sup>。进一步研究表明增加有义 NIF3L1 mRNA 的表达会增强维甲酸诱导的 P19 细胞的神经分化,这个过程还伴随着 Oct-3/4 mRNA 的大量减少和 Mash-1 mRNA 的快速增加,这两种 mRNA 编码的转录因子都是神经生成的调节子<sup>[23]</sup>。综上所述,在维甲酸诱导的 P19 细胞的神经分化中 NIF3L1 和 Trip15/CSN2 形成一个新的共抑制复合物,抑制了一些抑制神经生成基因的转录,从而促进了神经分化。

最近的研究表明人的 NIF3L1 和 WBS、ALS2 等一些人类疾病有关。WBS 是在染色体带 7q11.23 上发生染色体重组引起的遗传病。WBS 患者的关键区域基因 *wbscr14* 是半合子,它被与信号转导相关的 14-3-3 所调节。WBSCR14 是一种与碳水化合物(糖)结合的亮氨酸拉链家族的转录因子,它的过表达可以使碳水化合物的效应因子直接结合到 L-型丙酮酸激酶(LPK)的启动子上,从而引起 L-型丙酮酸激酶的增加。通过酵母双杂交系统分析发现 WBSCR14 可以和 14-3-3 与 NIF3L1 共同作用。WBSCR14、14-3-3 和 NIF3L1 在哺乳动物细胞中可以形成一个复合物,使 WBSCR14 留在细胞质当中并且增强 *wbscr14* 基因的转录,通过调节转录起到一定的生理作用<sup>[25]</sup>。ALS2 是一种递进的神经变性疾病,青少年肌萎缩侧索硬化的萎缩型常染色体位于人的染色体 2q33-q34,而 *nif3l1* 基因就位于染色体 2q33 上,所以说虽然人的 NIF3L1 的确切功能还未知,但不难看出 *nif3l1* 基因与 ALS2 有密切关系<sup>[26]</sup>。

综上所述,我们可以总结出,在真核生物中, NIF3 类超家族蛋白主要通过与转录因子相互作用而发挥

生物学功能。它可以和与之相互作用的转录因子一起进入细胞核中共同抑制靶基因的转录,也可以使与它相互作用的转录因子留在细胞质中而引起生理反应,所以在真核生物中该家族蛋白与基因转录密切相关。

### 3 NIF3 类超家族蛋白的研究现状

目前在国内还没有发现 NIF3 类超家族蛋白的相关报道,世界上对这个家族蛋白的研究也相对较少。截止到 2003 年 NIF3 类超家族蛋白仅有 67 个成员<sup>[7]</sup>,随着越来越多生物的基因组的测定,以及结构预测、同源性比对等一系列生物软件和生物信息学的发展,该家族的成员也越来越多。现在已经可以在 NCBI 上检索到一千多条 NIF3 类超家族的相关蛋白质。最近又在许多生物中发现了该家族蛋白的存在。

直到 2003 年才有该家族蛋白结构的相关研究,在功能上也仅对它们的个别成员有初步的了解。在结构上,目前只有 3 种该家族蛋白的结构有所报道<sup>[8]</sup>。在功能上,虽然酵母中的 NIF3 发现的比较早,但是也只是 Martens 等<sup>[2]</sup>通过酵母双杂交的方法在寻找与 NGG1p 相互作用蛋白时发现的,并没有对其功能进行针对性的深入研究,所以现在只知道它和转录有关,它的作用机制和具体的生物学功能还不清楚。目前对该家族蛋白的研究主要集中在哺乳动物细胞中,尤其是人和小鼠的 NIF3L1 研究相对多一些。由于该家族蛋白的作用机制生物化学性质还需要进一步的研究,所以目前它们仍被划分为未知功能蛋白质<sup>[4]</sup>。

本实验室在自己构建的家蚕 cDNA 文库中发现了一个编码含有 NIF3 结构域蛋白的 EST 序列,其编码蛋白的氨基酸序列两侧与其他 NIF3 类超家族蛋白有着很高的同源性,与人和小鼠等哺乳动物两端的同源性可以达到 63%。昆虫中该家族蛋白的功能目前还没有人报道,但以往的研究表明, NIF3 类超家族蛋白可调节神经的分化,而家蚕是完全变态生物,神经分化必定存在于变态的全过程,我们推测家蚕中的 NIF3 类蛋白在家蚕的发育过程中发挥着重要的作用。我们现在正对家蚕中的 NIF3 类蛋白进行功能和结构的研究,目前的工作主要集中在该蛋白质在细胞水平发挥的生物学功能和蛋白质的天然结构上。目的在于对该家族蛋白的功能有更深入的认识,填补 NIF3 类超家族蛋白在昆虫中研究的空白。

### 4 NIF3 类超家族蛋白研究前景

NIF3 类超家族蛋白在生物界中(从细菌到高等



的哺乳动物)广泛存在,并且在各种组织中也分布较广,推测它在生物体中发挥重要而不可缺的作用。但是NIF3类超家族蛋白是一类新的蛋白质研究的并不多,功能还不是很清楚,特别是作用机制有待于进一步的研究,这也为我们进行更深入的研究留有广泛的空间。相信随着研究的深入,该家族蛋白会从未知变为已知蛋白,鉴于NIF3L1和一些疾病相关,对这类家族的深入研究很可能会在一些疑难病症的治疗方面有所突破。

### 参考文献(References)

- [1] Pfam: Protein families database of alignments and HMMs [http://www.sanger.ac.uk/Software/Pfam/]
- [2] Martens JA *et al. J Biol Chem*, 1996, **271**: 15884
- [3] Karniely S *et al. Exp Cell Res*, 2006, **312**: 3835
- [4] Galperin MY *et al. Nucleic Acids Res*, 2004, **32**: 5452
- [5] Tascou S *et al. Cytogenet Cell Genet*, 2000, **90**: 330
- [6] <http://scop.berkeley.edu/data/scop.b.d.bdd.A.A.A.html>
- [7] Ladner JE *et al. BMC Struct Biol*, 2003, **3**: 7
- [8] Saikatendu KS *et al. BMC Struct Biol*, 2006, **6**: 27
- [9] Schwede T *et al. Nucleic Acids Res*, 2003, **31**: 3381
- [10] Guex N *et al. Electrophoresis*, 1997, **18**: 2714
- [11] Shindyalov IN *et al. Protein Eng*, 1998, **11**: 739
- [12] Auld DS. *Biometals*, 2001, **14**: 271
- [13] Kumar S *et al. Brief Bioinform*, 2004, **5**: 150
- [14] Gifford CM *et al. Nucleic Acids Res*, 2000, **28**: 762
- [15] Khil PP *et al. Mol Microbiol*, 2002, **44**: 89
- [16] Kovall R *et al. Science*, 1997, **277**: 1824
- [17] Hingorani MM *et al. Nat Rev Mol Cell Biol*, 2000, **1**: 22
- [18] Arnesano F *et al. J Biol Chem*, 2003, **278**: 45999
- [19] Cheah E *et al. Structure*, 1994, **2**: 981
- [20] Xu Y *et al. J Mol Biol*, 1998, **282**: 149
- [21] Wu M *et al. J Bacteriol*, 1999, **181**: 4755
- [22] Uetz P *et al. Nature*, 2000, **403**: 623
- [23] Akiyama H *et al. J Biol Chem*, 2003, **278**: 10752
- [24] Akiyama H *et al. Brain Res Dev Brain Res*, 2003, **140**: 45
- [25] Merla G *et al. Hum Mol Genet*, 2004, **13**: 1505
- [26] Hadano S *et al. Genomics*, 2001, **71**: 200

## NIF3-like Superfamily Proteins

Qi-Jing Gai, Yao-Zhou Zhang\*

(Institute of Biochemistry, School of Life Science, Zhejiang Sci-Tech University, Hangzhou 310018, China)

**Abstract** NGG1p interacting factor 3 (NIF3)-like protein belongs to a protein superfamily which contains a NIF3 domain that is highly conserved from bacteria to mammals. The mRNAs which encode these proteins are ubiquitous in various kinds of species and tissues. However, due to that most of these proteins are presumed, there are few studies on their functions. So, they are called as “unknown proteins”. The proteins of this super family are  $\alpha/\beta$  proteins, and execute functions by forming dimers or complexes with other proteins. It has been demonstrated that this kind of proteins is involved in gene transcription and particularly plays an essential role in the mammal neurocyte differentiation. Recent studies indicated that NIF3-like proteins also related to some human diseases. Just because these NIF3-like proteins are abundant, ubiquitous and functionally unknown, more and more attentions were attracted to them.

**Key words** NIF3; neural differentiation; transcription; unknown function

Received: May 28, 2007 Accepted: July 2, 2007

This work was supported by the National High Technology Research and Development Program of China (863 Program) (No. 2005AA206120), and the National Basic Research Program of China (973 Program) (No.2005CB121006)

\*Corresponding author. Tel/Fax: 86-571-86843198, E-mail: yaozhou@chinagene.com