

重组斑马鱼 Gli2 蛋白 N 端肽段的 表达及纯化^{*}

谭训刚^{1 2} 张培军¹

(¹中国科学院海洋研究所, 青岛 266071)

(²中国科学院研究生院, 北京 100039)

杜少军³

(³ Center of Marine Biotechnology, University of Maryland Biotechnology Institute

Baltimore MD USA 21202)

Hedgehog (Hh) 信号及其信号传递系统在生物机体发育中起着非常重要的作用。从 Hhs 分泌到其靶基因收到来自 Hhs 的信号的过程中, 包括很多步骤, 而每一步骤都有不同的因子参与作用。在其传递的过程中, 参与受体细胞将胞外的信号传到胞内调节其目标基因的一个重要因子是果蝇分节-极性基因 *cubitus interruptus* (Ci) 及其脊椎动物同源基因-Gli 编码的蛋白。

遗传学家分析认为 Ci 是依赖 Hh 的转录活化的中介物质 (Forbes *et al.*, 1993), 但是对这一复杂的蛋白的功能还远没有搞清楚。Ci 属于 Gli 型锌指蛋白 (ZFP) 家族 (Ruiz 1999)。在脊椎动物中已经鉴定了 3 个 Gli 基因 (*Gli1*, *Gli2*, *Gli3*) (Koebernick *et al.*, 2002), 线虫中有一个 Gli 型的锌指基因-*tra-1* (Zarkower *et al.*, 1992)。Gli 型的锌指蛋白共包括 5 个高度保守的 C2H2-型的锌指结构域。果蝇的实验证实了 Ci 有双重功能, 可作为 Hh 目标基因的转录活化剂或抑制剂。其中 Ci 的 C 端截短形式—Ci-75 (Aza-Blanco *et al.*, 1997), 仍保留了锌指结合结构域和一个 N 端的抑制结构域; 当没有 Hh 信号时, Ci-75 位于细胞核中并抑制 Hh 目标基因的转录 (Koebernick *et al.*, 2002), 并存在于所有的 Hh 应答细胞中。但是 Ci 的全长形式 Ci-155 则分布于那些分泌 Hh 的细胞邻近的细胞中并累积 (Koebernick *et al.*, 2002)。当 Hh 信号抑制了 Ci 的切割时, 导致了有转录激活作用的全长 Ci-155 的积累 (Aza-Blanco *et al.*, 1997), 从而激活 Hh 目标基因的转录, 与 Ci-75 不同的是 Ci-155 主要位于细胞质中。

实验证明, Gli2 在组织培养细胞 (Ruiz 1999) 和转基因果蝇中表达时被降解。与 Ci

^{*} 国家自然科学基金项目 (30210403202) 资助。

通讯作者: 张培军, 男, 研究员, 博士生导师, 现从事海洋生物学等方面的研究。

E-mail: pjzhang@ms.qdio.ac.cn tanxgj@yahoo.com

收稿日期: 2005 年 1 月 4 日。

©1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. <http://www.cnki.net>

不同, 这种降解作用似乎对 Hh 信号不敏感, 而且降解后的蛋白是整个蛋白的一个部分。虽然在爪蟾的胚胎中描述了标记的 G li1 和 G li2 蛋白的截短形式 (Ruiz 1999), 而且观察到 G li2 能被 PKA 磷酸化 (Wang *et al.*, 2000), 但在建立的细胞系中却没有检测到内源性 G li1 和 G li2 蛋白的降解过程 (Ruiz 1999, Dai *et al.*, 1999, Wang *et al.*, 2000)。在果蝇中的研究发现 G li2 和 G li3 是蛋白水解后的产物, 但是 G li2 并不存在依赖 Hh 信号的降解抑制作用 (Aza Blanc *et al.*, 2000)。

专家在对斑马鱼的试验中发现: 在其肌肉发育过程中, G li2 对于肌肉发育尤其是慢肌的发育非常重要 (Lewis *et al.*, 1999, Du *et al.*, 2001)。虽然在爪蟾中通过重组 G li2 标记蛋白已证明 G li2 蛋白存在蛋白降解途径。目前尚没有证据可以证明内源的 G li2 蛋白也存在该机制, 这主要是因为内源 G li2 蛋白表达量低或者性质不稳定造成的。在果蝇中, 用识别蛋白不同部分的抗体证明了 CiG li2 的同源基因的蛋白存在蛋白降解。用同样的方法也可得到斑马鱼体内 G li2 蛋白的表达情况。

本文作者根据 G li2 蛋白的氨基酸序列, 利用计算机程序分析蛋白的亲水 疏水区域, 设计并构建了表达 G li2 蛋白 N 端肽段的基因重组体, 经转化工程菌, 优化表达条件, 获得了表达的可溶性蛋白, 并对该蛋白进行了纯化, 为进一步用于斑马鱼的发育分析打下基础。

一、材料和方法

1. 材料

pfu 多聚酶 (Stratagene USA), 胶回收试剂盒 (Qiagen USA), B-PERTM 细菌蛋白提取液 (Pierce USA), Ni-NTA 蛋白提取介质 (Clontech USA), 15% SDS-PAGE, SDS-PAGE 电泳缓冲液, SDS 甘油, β -巯基乙醇, LB 培养基, IPTG, 氨苄青霉素, Mini 8*10 垂直板电泳系统 (Life Technologies USA), 多种内切酶赞许等材料均由 Promega 公司提供。

质粒 pBluescript II SK (+) (Stratagene USA), 带有 His 标签的表达载体 A, B, C (His A, B, C) (Clontech USA)。

菌株 BL-21。

2 PCR 扩增所需的 DNA 片段

通过 PCR 扩增所需要的 DNA 片段, 为保证序列的正确性, 在 PCR 过程中采用 *pfu* 多聚酶, 所用的引物是 G li2N2-5 (5'-GGATCCC GCAGCCACTGAGAAATCCCCCTTCAAC-CCT-3') 和 G li2N2-3 (5'-GGATCCCCTACGAGTCTCCATAAGGGCTACGGTG GCT-3'), 在每一个引物的 5 端引入 *Bam*H I 切点。PCR 扩增的程序是 94℃ 变性 3 min, 然后是 35 个循环扩增。每个循环扩增中含有 94℃ 变性 30 s, 55℃ 复性 30 s, 72℃ 延伸 60 s, 最后是 72℃ 延伸 10 min。

3 质粒的构建

PCR 产物经胶回收试剂盒回收后, 被克隆入 pBluescript SK (+) 载体的 *Sma* I 位

点。经 *Bam*H I酶切判断含有所需片段的质粒小提后进行测序（ABI373测序仪）。然后含有正确序列的片段被 *Bam*H I酶切出来连接到 *Bam*H I酶切过并去磷酸化的带有 His 标签的表达载体，得到 G li2N2 His B。通过 *Eco*R I和 *Sca* I酶切判断插入方向的正确性，含有正确方向的质粒被测序。含有正确序列以及阅读框架的质粒被进一步大量提取。

4 计算机程序

疏水 亲水分析软件为 http://bioinformatics.weizmann.ac.il/hydrophobicity_h.html;

序列比较 (Clustalw)采用 <http://www.ebi.ac.uk/clustalw>。

5 其他方法

根据《Q IAGEN® Plasmid Purification Handbook (Qiagen, USA)》进行质粒的提取；根据《GST Gene Recombinant System Third Edition Revision 2.18-1123-20(Amersham Pharmacia Biotech USA)》制备感受态细胞 BL-21并转入表达质粒；根据《Qiagen Q IAxpressionist™ fifth edition—a handbook for high-level expression and purification of 6×His tagged proteins (Qiagen, 2001)》进行重组蛋白的诱导表达并对诱导物的浓度以及诱导的时间进行优化；根据《Qiagen Q IAxpressionist™ fifth edition (Qiagen)》以及《Instruction B-PER® Bacterial Protein Extraction Reagent (Pierce, USA)》利用 NiNTA 进行重组的蛋白可溶性分析以及提取进行纯化。

二、结果与讨论

1. 比较斑马鱼的 G li2与果蝇的 Ci

由于斑马鱼的 G li2和果蝇的 Ci是不同物种之间的同源基因，其在蛋白水平有着较高的保守性，在果蝇中 Ci的蛋白的水解位点以及锌指保守区已经通过实验推测出来 (Aza Blanc *et al.*, 1997)。因此通过比较就可以将锌指保守区以及可能的蛋白水解位点推测出来。这个信息决定产生不同抗体的重组蛋白在全序列中的位置。本文作者运用 Clustalw 程序比较了斑马鱼的 G li2以及果蝇的 Ci推测锌指结构区和蛋白水解位点。

根据果蝇 Ci的锌指结构区以及蛋白水解位点，确定出斑马鱼中的锌指结构区和蛋白水解位点 (图 1)。锌指结构区在第 402和 532个氨基酸之间，G li2N 端的肽段应在第 402个氨基酸之前；蛋白水解的位点在第 604和 655个氨基酸之间。

2 亲水 疏水区的决定

亲水 疏水性对蛋白的正确折叠非常重要。正确的折叠形成正确的结构，该结构可以影响蛋白的可溶性以及抗体的识别。对于抗体的产生，可溶的自然结构的蛋白是最好的抗原。因此对于用于识别 G li2的抗体，其抗原的亲水 疏水区必须确定，本文作者通过计算机程序对蛋白的亲水 疏水区进行了确定。

.....		
<i>D anio</i>	N - - - - - A RGH TPP - - - - - LIQPSPALSSRQQTLVAAAAL	320
<i>D rosoph ila</i>	LNFMTPQVAA SGFSIGHMPTSA SLRVNDVHPNLSDSH IQ ITSP T V TK D V S Q V P A A A F S	349
<i>D anio</i>	NN - - - - - -TSSSTSSSTA DTSTESSQNAGGD PAV SSTVN PM IFRSKV	363
<i>D rosoph ila</i>	LKN IDDA REK KGP FK DVV PEQ PSSTSGG VAQV EAD SA SSQ LSD RCYN NV V N N IK G IPG D V	409
<i>D anio</i>	KTEAEGPHPBFGSQ - - - DHLELREELDKDECKQEPEAVYE TNCHWEGCAKEYDTQDQLV	420
<i>D rosoph ila</i>	KVNSRLDEY NCGSISIPSENYDCA NADTTDKDEFDFE TNCHWR SC R IE F IT Q D EL V	469
<i>D anio</i>	HHISNDHIHGEKKEFVCRWVECSREQKPKAQYMLVVHMRRTGKEKPHKCTFEGCSKAYS	480
<i>D rosoph ila</i>	KHNNDH IQTNKKA FVCRWEDCTRGEKPFKAQYMLVVHMRRTGKEKPHKCTFEGCFKAYS	529
<i>D anio</i>	RLENLK THLRSHGTGEKPYVCEHEGCKNKA FSNA SDRAKHQNRTH SNEKPYVCK PGCTKRY	540
<i>D rosoph ila</i>	RLENLK THLRSHGTGEKPYTCEYPGCSKAFSNASDRAKHQNRTH SNEKPYICK APGCTKRY	589
<i>D anio</i>	TDPSSLRKHVKTVHGP EAHVTKK QRGDA PPKHPKNGENEAH TKHV RG RTDG SG E A N S	600
<i>D rosoph ila</i>	TDPSSLRKHVKTVHGA EFYA NKKHKG - - LPLNDANSRLQ QNNSRHNIQ EHN IDSSPCSED	647
<i>D anio</i>	TTR GVEDCQHVK TIK TEN TVM YQSSFGGQSSCSSESP IGSATNND SGV EM AMH S GGSLG	660
<i>D rosoph ila</i>	SHL GKMLGTSSPSKSES DISSNHH LVNG VRASDS - - LLTSPDDIA ENLNIDD GW NCD	705
<i>D anio</i>	DLSALDDTPVVDSTGSPGTSAGVGLQLRKN RAGLLQLEH KKEKIK TV RD SC SW A N A P P Q	720
<i>D rosoph ila</i>	DDVDVADLPN LRAM VN IENGNASASTEGSV LARQRFRGRLQTK G IN S ST IM LCN IPE S	765
<i>D anio</i>	VRNTK LPPIPS IDSLLD APNMGTQM SV PPTQH IGD LSSYEM TM LNQ IHE RRD SSTSTM S	779
<i>D rosoph ila</i>	NRTFG ISELNQR ITE LKMEPGTDA EIK IPK LPNT TIGGYTED PLQNQTSFRNTV SNKQG	824
.....		

图 1 比较果蝇的 C 和斑马鱼的 G l2 蛋白序列
黑体示指保守区，下划线部分示可能的水解位点。

从图 2 中可以看出 G l2 蛋白中氨基酸亲水 疏水区的分布情况。在蛋白的折叠过程中，通常是亲水区在蛋白外侧，而疏水在蛋白内部。如果 G l2N 端肽段的疏水区序列在中间而亲水区在两侧，该蛋白就会比较容易折叠形成可溶型的蛋白。

将 G l2 蛋白和 C 蛋白序列的比较结果与亲水 疏水区结合在一起进行分析，确定出需要克隆和表达的 G l2 蛋白 N 末端肽段。该肽段 (G l2 N 端肽段，G l2N2) 的氨基酸序列为第 86 到第 180 共 95 个氨基酸残基。

3 蛋白产生条件的优化

本文作者采用能够高产率表达外源蛋白的大肠杆菌 BL-21 菌系作菌种。表达质粒被转入 BL-21 细菌中进行扩菌培养，在 IPTG 诱导条件下进行重组蛋白的表达实验。由于重组蛋白是在细菌经 IPTG 诱导到一定时间产生的，因此诱导物 IPTG 的浓度以及诱导时间的长短对重组蛋白的产率具有一定的影响。

(1) 诱导时间对重组蛋白产生量的影响

诱导时间的长短是一个影响重组蛋白量的因素。诱导物加入的时间是 OD₆₀₀ 值在 0.7 到 0.8 由于重组蛋白在细菌中产生，细菌有一个生长曲线，在生长曲线的不同的部分细菌的生长完全不同从而影响重组蛋白的量，也就是诱导时间的长短影响蛋白的产生量。通过在 15% SDS-PAGE 中电泳蛋白，比较不同诱导时间下的重组蛋白的量 (图 3A)。在所有的重组蛋白的表达时间中，诱导 3 h 的蛋白量处于诱导曲线的较高值，从 4 h 以后蛋白的表达量开始降低 (图 3A)，因此本实验采用诱导 3 h

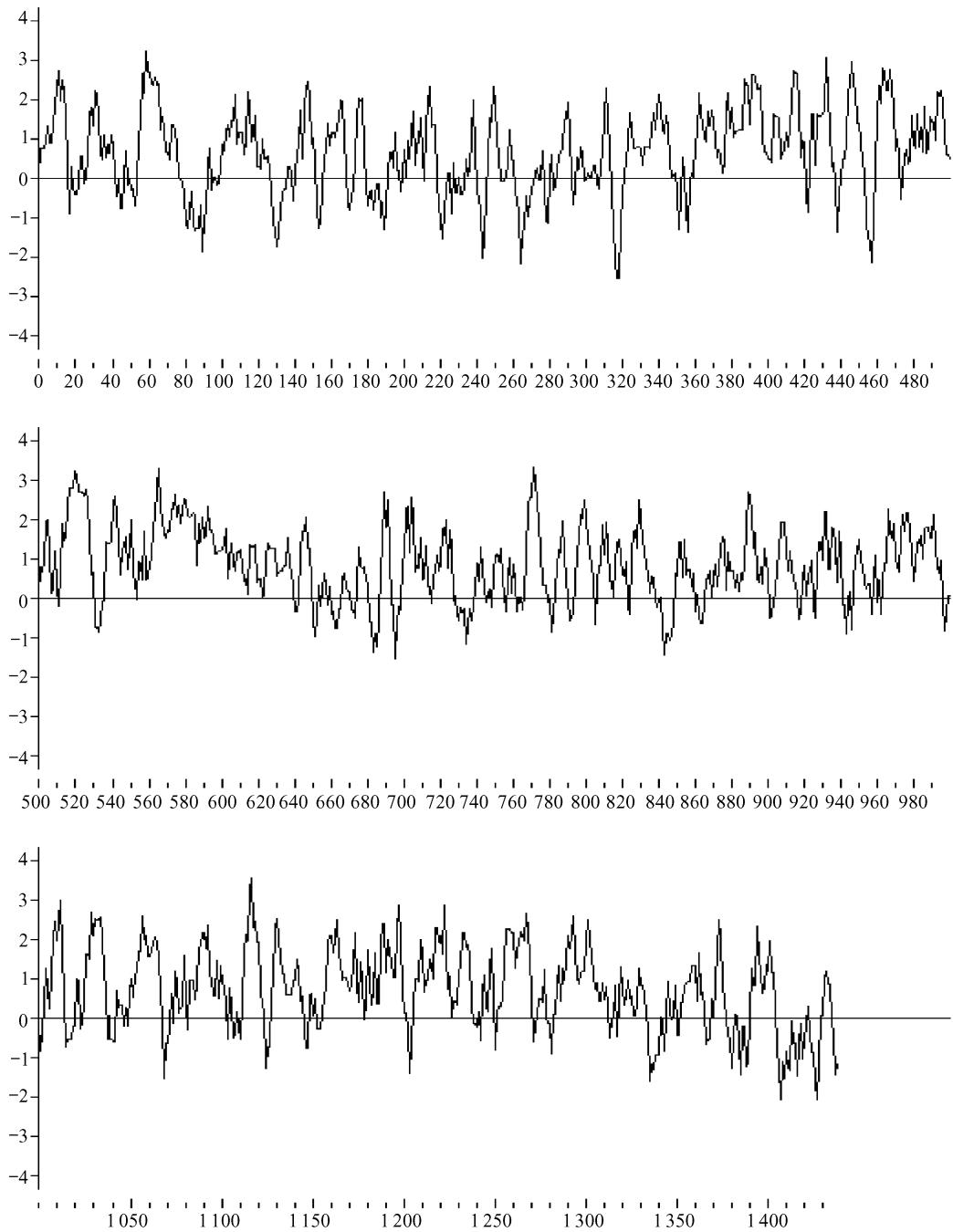


图 2 G12 蛋白的亲水 疏水示意图

X轴以上是亲水区，以下是疏水区。横线所跨区域是 G12 的 N 末端第 86~180 氨基酸。

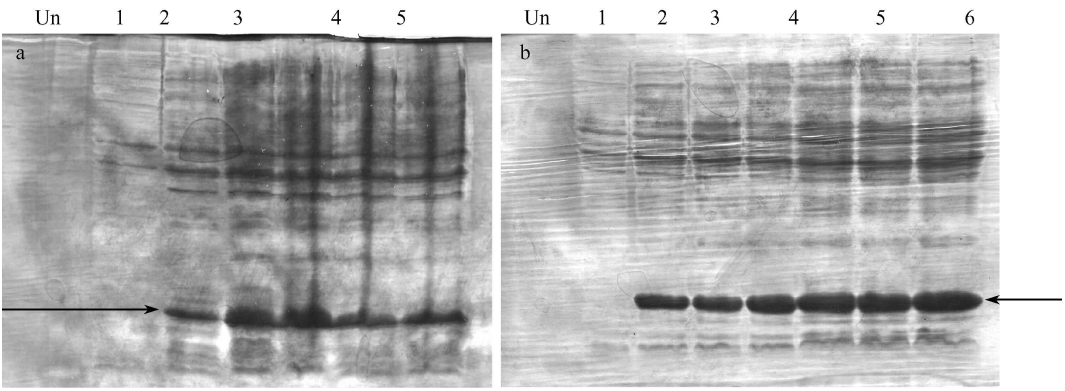


图 3 IPTG 的浓度对蛋白产生的影响以及诱导时间对蛋白产生的影响

a. IPTG 浓度为 0.6 mM 时, 诱导时间对 G li2N2 蛋白的影响。

Un 没有诱导; 诱导时间: 1: 1 h 2: 2 h 3: 3 h 4: 4 h 5: 5 h

b. 诱导 3 h IPTG 浓度对 G li2N2 蛋白的影响。

Un 没有诱导: 1: 0.2 mmol 2: 0.4 mmol 3: 0.6 mmol 4: 1.0 mmol 5: 1.5 mmol 6: 3.0 mmol

(箭所标明的是目的蛋白)

(2) 诱导物 IPTG 浓度对重组蛋白量的影响

实验证明, 诱导物 IPTG 的浓度对产生量有很大的影响。IPTG 的浓度根据产生的重组蛋白的量以及 IPTG 的用量决定。根据这两个标准, IPTG 的浓度设定在 0.2 mmol 到 3 mmol 之间。通过在 15% SDS-PAGE 中电泳蛋白, 比较不同浓度下的重组蛋白的量 (图 3B)。从图中可以看出, 0.6 ~ 1.0 mmol 的 IPTG 诱导产生的蛋白处于曲线最高值。

4. 重组蛋白的可溶性和与亲和柱结合效率

在细菌产生的重组蛋白中, 并不是所有的蛋白均为可溶性的, 通常是包涵体。如果是可溶的, 这意味着该蛋白可以按照自然状态折叠, 这对抗体的产生十分有利。而在大多数情况下, 为了制备抗体, 一个重要的方法是将蛋白变性, 然后复性, 最后纯化复性后的蛋白。如果重组蛋白可溶, 这将节省很多时间。因此, 必须决定蛋白在提取缓冲液中的可溶性。在该实验中, 提取的缓冲液是 B-Per 溶液。

对于可溶性蛋白的纯化, 蛋白和分离介质的结合效率非常重要。因为有些蛋白可能把标签 (6XHis) 折叠在蛋白内部, 这样就不可能纯化该蛋白, 因此需要进行初步实验检测重组蛋白能否与介质 (NiNTA) 结合; 另外用于提取可溶性蛋白的溶液含有多种化学物质, 这些物质对于分离介质可能会有一些伤害或者影响蛋白和介质的结合率造成纯化效率低下, 为避免伤害介质或者提高纯化效率, 就必须进行一些稀释以检测其结合效率。同时也要考虑到稀释后的体积, 避免造成庞大体积, 而不利于纯化。参照 B-Per 溶液说明用洗涤溶液将提取上清液稀释 1~3 倍。

结果表明带标签的 G li2N2 的蛋白是可溶性的 (见图 4), 并且可以和 NiNTA 介质结合; 即使在没有稀释的情况下, B-Per 提取溶液中的蛋白也可和 NiNTA 柱结合

且结合率很高。但是考虑 B-Per溶液中含有多种成分对 NiNTA 介质可能造成损害, 作者采用 1:1 的稀释度来提纯该蛋白。

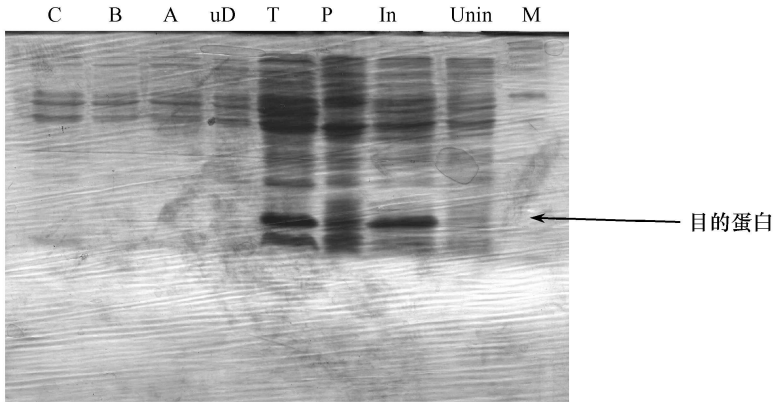


图 4 蛋白在 B-per 溶液中的溶解性以及和 NiNTA 柱结合的效率

A: 1:1 稀释样品; B: 1:2 稀释样品; C: 1:3 稀释样品; uD: 没有稀释的样品; T: 上清提取液; P: 提取后得到的沉淀; In: 诱导产生的蛋白; Unin: 没有诱导产生的蛋白; M: 蛋白标准

5 重组蛋白的纯化

通过上述实验得到可溶性的蛋白后需要将其纯化才能用于抗体的产生, 并且上述实验已经证实将重组蛋白稀释后再可以结合到 NiNTA 柱上, 因此将重组蛋白稀释后将其纯化。本实验采取分批纯化的方式, 将每次纯化过程中的不同的组分 (1 mL) 收集, 进行 SDS-PAGE 电泳纯度分析 (图 5), 然后将其电泳纯的组分收集起来, 最后就得到电泳纯的蛋白。在纯化的过程中 (图 6), 洗脱的第一组分中仍然含有一些杂质, 而在第二组分以后杂质消失, 蛋白的纯度较高, 达到电泳纯, 并且在第二到第四组分中蛋白的浓度比较稳定。因此收集第二到第五组分, 就得到高纯度的蛋白。

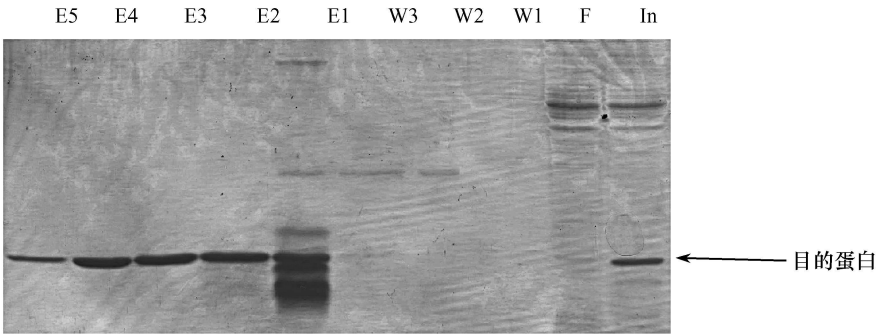


图 5 纯化 G12 蛋白

In: 诱导的纯化前的 G12 蛋白; F: 过柱后滤液; W1, W2: 洗涤组分; E1, E2, E3, E4, E5: 洗脱收集的组分

体内具有活性的蛋白是正确折叠具有一定构型的, 要识别这些蛋白, 产生的抗体的抗原是正确折叠具有相同构型的蛋白否则产生的可能只识别线性非正确折叠变性的蛋白而非具有活性保持正常构型的蛋白。因此在很多抗体制备的过程中, 由于重组蛋白的不可溶性以及非正确折叠, 造成产生的抗体只识别变性下蛋白而非有活性的蛋白。比如 G li2 从抗体方面, 由于本身不能在体内被检测到是由于没有得到好的抗体, 从抗原方面, 由于没有获得可溶性以及正确折叠的蛋白用于抗体的产生, 造成产生的抗体为识别性蛋白而不是有活性的蛋白。本文作者在试验中获得了一个可溶性的 G li2 N 端的肽段。这是第一次报道细菌产生自然状态的一段 G li2 肽段, 并且可以在高浓度下产生, 这对于抗体的生产是一个良好的开端。

在重组蛋白中, 大多数重组蛋白是包涵体, 作者在实验中获得了一个可溶性的 G li2 部分蛋白。在该蛋白的纯化过程中, 一种特殊的溶液 B-per 细菌蛋白提取液被采用, 并且被证明非常有效。经纯化后的蛋白纯度及浓度极高, 对抗体的制备较为有利。

利用该蛋白产生的抗体将有利于检测内源性蛋白, 分析出体内 G li2 蛋白的分布情况以及所处位置, 并分析 G li2 蛋白是否水解产生 C 末端截断的蛋白。为今后更进一步开展蛋白的功能研究奠定了基础。

参 考 文 献

- Aza Blane P., Lin H., Ruiz Alcala A., *et al* 2000. Expression of the vertebrate Gli proteins in *Drosophila* reveals a distribution of activator and repressor activities *Development* **127**: 4293 ~ 4301
- Aza Blane P., Ramirezveber F., Laget M., *et al* 1997. Proteolysis that is inhibited by hedgehog targets cubitus interruptus protein to the nucleus and converts it to a repressor *Cell* **89**: 1043 ~ 1053
- Dai P., Akimaru H., Tanaka Y., *et al* 1999. Sonic hedgehog induced activation of the Gli1 promoter is mediated by GLI3. *J. Biol. Chem.* **274**: 8143 ~ 8152
- Du S. J., Dierhart M. 2001. Gli2 mediation of hedgehog signals in slow muscle induction in zebrafish *Differentiation* **2001**, **67**(3): 84 ~ 91
- Forbes A., Nakano Y., Taylor A., *et al* 1993. Genetic analysis of hedgehog signaling in the *Drosophila* embryo *Development Suppl.* **115** ~ 124
- Koebernick K., Pieler T. 2002. Gli type zinc finger proteins as bipotential transducers of Hedgehog signaling *Differentiation* **70**: 69 ~ 76
- Lewis K. E., Currie P. D., Roy S., *et al* 1999. Control of muscle cell type specification in the zebrafish embryo by Hedgehog signaling *Dev Biol* **216**: 469 ~ 480
- Ruiz Alcala A. 1999. Gli proteins encode context-dependent positive and negative functions: Implications for development and disease *Development* **126**: 3205 ~ 3216
- von Mering C., Basler K. 1999. Distinct and regulated activities of human Gli proteins in *Drosophila* *Curr. Biol.* **9**: 1319 ~ 1322
- Wang B., Fallon J., and Beahy P. 2000. Hedgehog regulated processing of Gli3 produces an anterior/posterior repressor gradient in the developing vertebrate limb *Cell* **100**: 423 ~ 434
- Zarkower D., Hodgkin J. 1992. Molecular analysis of the *C. elegans* sex determining gene tra-1: a gene encoding two zinc finger proteins *Cell* **70**(2): 237 ~ 249

Expression and Purification of N-Term inal Partial Protein of Recombination Z ebrafish G li2

TAN Xungang^{1,2} ZHANG Peijun¹ DU Shaojun³

(^{1,2} *Institute of Oceanology Chinese Academy of Sciences Qingdao 266071*)

(² *Graduate School of Chinese Academy of Sciences Beijing 100039*)

(³ *Center of Marine Biotechnology University of Maryland Biotechnology Institute Baltimore MD USA 21202*)

ABSTRACT

Based on the protein comparison between zebrafish G li2 and fruit fly C₁ and the analysis of hydrophilicity hydrophobicity of G li2 protein by computer programmer the recombination N-terminal partial protein of G li2 protein was designed. After the DNA fragment encoding the N-terminal part was cloned out by PCR, it was ligated with an expression vector containing His tag and transformed into BL-21, a bacterial expression strain to express this protein. After the expression condition was optimized, the protein can be expressed at high level after having been induced for 3 hours when the inducer IPTG was at 0.6 mmol/L. The protein is soluble in solution, a bacterial protein extraction solution. After the protein solution was diluted in B-per solution, the recombination protein can bind to the NiNTA column and be purified highly.