

[19]中华人民共和国国家知识产权局

[51]Int. Cl⁷

[12] 发明专利申请公开说明书

C12N 15/12
C07K 14/47 C12N 15/63
C12N 1/21 C12N 5/10
C07K 16/44

[21] 申请号 98111045.2

[43]公开日 2000年3月8日

[11]公开号 CN 1246532A

[22]申请日 1998.8.31 [21]申请号 98111045.2

[71]申请人 复旦大学

地址 上海市邯郸路220号

[72]发明人 余龙 傅强 刘擎
屠强 赵寿元

[74]专利代理机构 上海专利商标事务所
代理人 徐迅

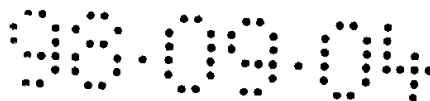
权利要求书1页 说明书14页 附图页数0页

[54]发明名称 人热休克关联蛋白编码序列、其编码的多肽及制备方法

[57]摘要

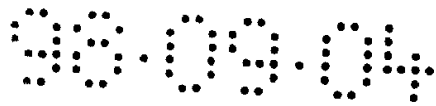
本发明涉及一种新的人基因核苷酸序列。更具体地说,本发明涉及热休克关联蛋白家族的新成员 HSC3 的 cDNA 编码序列。本发明还涉及该 cDNA 序列 编码的多肽、所述多核苷酸序列和所述多肽的生产方法、以及这些多核苷酸和 多肽的应用。

ISSN 1008-4274



权 利 要 求 书

1. 一种分离出的DNA分子, 其特征在于, 它包括: 编码具有人HSC3蛋白活性的多肽的核苷酸序列, 所述的核苷酸序列与SEQ ID NO.3中从核苷酸108-1037位的核苷酸序列有至少70%的同源性; 或者所述的核苷酸序列能在中度严紧条件下与SEQ ID NO.3中从核苷酸108-1037位的核苷酸序列杂交。
5
2. 如权利要求1所述的DNA分子, 其特征在于, 所述的序列编码一多肽, 该多肽具有SEQ ID NO.4所示的序列。
3. 如权利要求1所述的DNA分子, 其特征在于, 该序列具有SEQ ID NO.3中核苷酸108-1037位的序列。
10
4. 一种分离的HSC3蛋白多肽, 其特征在于, 它包括: 具有SEQ ID NO.4氨基酸序列的多肽、或其活性片段, 或其活性衍生物。
5. 如权利要求4所述的多肽, 其特征在于, 该多肽是具有SEQ ID NO.4序列的多肽。
6. 一种载体, 其特征在于, 它含有权利要求1所述的DNA。
15
7. 一种用权利要求6所述载体转化的宿主细胞。
8. 如权利要求7所述的宿主细胞, 其特征在于, 该细胞是大肠杆菌。
9. 如权利要求7所述的宿主细胞, 其特征在于, 该细胞是真核细胞。
10. 一种产生具有HSC3蛋白活性的多肽的方法, 其特征在于, 该方法包括:
20 (a) 将编码具有HSC3蛋白活性的多肽的核苷酸序列可操作地连于表达调控序列, 形成HSC3蛋白表达载体, 所述的核苷酸序列与SEQ ID NO.3中从核苷酸108-1037位的核苷酸序列有至少70%的同源性;
(b) 将步骤(a)中的表达载体转入宿主细胞, 形成HSC3蛋白的重组细胞;
(c) 在适合表达HSC3蛋白多肽的条件下, 培养步骤(b)中的重组细胞;
25 (d) 分离出具有HSC3蛋白活性的多肽。
11. 如权利要求10所述的方法, 其特征在于, 该核苷酸序列为SEQ ID NO.3中从核苷酸108-1037位。
12. 一种能与权利要求4所述的HSC3蛋白多肽特异性结合的抗体。
13. 一种核苷酸分子, 其特征在于, 它是权利要求1所述DNA分子的反义序列。
30
14. 一种探针分子, 其特征在于, 它含有权利要求1所述的DNA分子中约8-100个连续核苷酸。



说明书

人热休克关联蛋白编码序列、其编码的多肽及制备方法

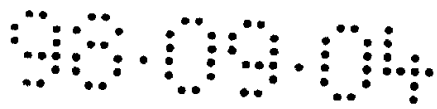
5 本发明涉及基因工程领域，具体地，本发明涉及一种新的人基因核苷酸序列。更具体地说，本发明涉及热休克关联蛋白家族的新成员的cDNA编码序列。本发明还涉及该cDNA序列编码的多肽、所述多核苷酸序列和所述多肽的生产方法、以及这些多核苷酸和多肽的应用。

10 热休克蛋白(hsp)是细胞或生物体在诸如热休克、化学物质等环境压力下，合成受诱导或增加的一类蛋白。值得注意的是，压力的类型和严重程度决定着hsp的应答情况，结果导致细胞存活或死亡。轻度的压力可使细胞作好准备对付以后严重的压力，而没有作好准备的细胞通常不能存活。根据分子量和氨基酸顺序同源度，hsp可分为不同的家族，hsp70家族就是其中之一。

15 这个家族的蛋白不仅在压力下诱导，而且一些被称为热休克关联蛋白(hsc)成员还组成型表达，这表明它们不仅在关键时刻起作用，而且在正常条件下也起作用。最近的研究发现此类蛋白广泛参与蛋白质的代谢活动，包括蛋白质的折叠、跨细胞内膜运输、蛋白水解、蛋白装配，以及蛋白集合体和多蛋白结构的解装配。大肠杆菌中的DnaJ、DnaK就属于这类蛋白，DnaJ作为调控因子调节DnaK
20 的活性，主要是刺激DnaK的ATP酶活，使产生能量用于蛋白折叠等代谢活动。

近几年的研究表明，真核细胞中也存在与DnaJ、DnaK同源的蛋白家族。1989年Sadler等人(*J. Cell. Biol.* 109(1989), 2665-1675)首次描述了酵母中与DnaJ同源的蛋白Sec63。蛋白Sec63含有与DnaJ的N端相似的一段70个氨基酸的区域，该蛋白质参与将蛋白运输到内质网的过程。此后，从酵母中分离克隆到一些DnaJ同源蛋白。1992年Cheetham等人(*Biochem. J.* 284(1992), 469-476)用双螺旋纤维蛋白的多克隆抗血清扫描人类前脑皮层表达cDNA库，首先克隆人类的DnaJ同源cDNA，
25 HSI1,推导出的氨基酸顺序与其它物种中的蛋白很相似。随后几年又有HDI1, HDI2, HSI1-b等与DnaJ同源的人类蛋白被克隆。但在本发明被公布之前，尚没有任何人公开过本发明涉及的人类HSC家族新成员。

30 本发明的一个目的是提供一种新的人多核苷酸，该多核苷酸编码热休克关联蛋白家族的一个新成员，本发明的热休克关联蛋白同系物命名为HSC3。



本发明的另一个目的是提供一种新的人热休克关联蛋白家族成员，它被命名为HSC3。

本发明的再一个目的是提供一种利用重组技术生产所述的新的热休克关联蛋白家族成员的方法。

5 本发明还涉及HSC3多肽和编码序列的应用。

在本发明的一个方面，提供了一种分离出的DNA分子，它包括：编码具有人HSC3蛋白活性的多肽的核苷酸序列，所述的核苷酸序列与SEQ ID NO.3中从核苷酸108-1037位的核苷酸序列有至少70%的同源性；或者所述的核苷酸序列能在中度严紧条件下与SEQ ID NO.3中从核苷酸108-1037位的核苷酸序列杂交。较佳地，所述的序列编码一多肽，该多肽具有SEQ ID NO.4所示的序列。更佳地，该序列具有SEQ ID NO.3中从核苷酸108-1037位的核苷酸序列。

10 在本发明的另一方面，提供了一种分离的HSC3蛋白多肽，它包括：具有SEQ ID NO.4氨基酸序列的多肽、或其活性片段，或其活性衍生物。较佳地，该多肽是具有SEQ ID NO.4序列的多肽。

在本发明的另一方面，提供了一种载体，它含有上述分离出的DNA。

在本发明的另一方面，提供了一种所述载体转化的宿主细胞。

15 在本发明的另一方面，提供了一种产生具有HSC3蛋白活性的多肽的方法，该方法包括：

20 (a)将编码具有HSC3蛋白活性的多肽的核苷酸序列可操作地连于表达调控序列，形成HSC3蛋白表达载体，所述的核苷酸序列与SEQ ID NO.3中从核苷酸108-1037位的核苷酸序列有至少70%的同源性；

(b)将步骤(a)中的表达载体转入宿主细胞，形成HSC3蛋白的重组细胞；

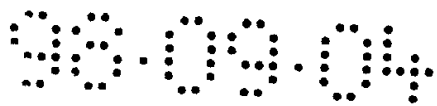
(c)在适合表达HSC3蛋白多肽的条件下，培养步骤(b)中的重组细胞；

25 (d)分离出具有HSC3蛋白活性的多肽。

在本发明的一个具体实施方案中，本发明的分离的多核苷酸全长为1203个核苷酸，其详细序列见SEQ ID NO.3，其中开放读框位于108-1037位核苷酸。

30 在本发明中，“分离的”、“纯化的”或“基本纯的”DNA是指，该DNA或片段已从天然状态下位于其两侧的序列中分离出来，还指该DNA或片段已经与天然状态下伴随核酸的组份分开，而且已经与在细胞中伴随其的蛋白质分开。

在本发明中，术语“HSC3蛋白(或多肽)编码序列”指编码具有HSC3蛋白活性的多肽的核苷酸序列，如SEQ ID NO.3中108-1037位核苷酸序列及其简并序列。



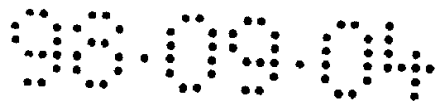
该简并序列是指，位于SEQ ID NO.3序列的编码框108-1037位核苷酸中，有一个或多个密码子被编码相同氨基酸的简并密码子所取代后而产生的序列。由于密码子的简并性，所以与SEQ ID NO.3中108-1037位核苷酸序列同源性低至约70%的简并序列也能编码出SEQ ID NO.4所述的序列。该术语还包括能在中度严紧条件下，更佳地在高度严紧条件下与SEQ ID NO.3中从核苷酸108-1037位的核苷酸序列杂交的核苷酸序列。该术语还包括与SEQ ID NO.3中从核苷酸108-1037位同源性至少70%，更佳地至少80%，更佳地至少90%的核苷酸序列。

该术语还包括能编码具有与人HSC3相同功能的蛋白的、SEQ ID NO.3序列的变异形式。这些变异形式包括(但不限于): 若干个(通常为1-90个，更佳地1-60个，更佳地1-20个，最佳地1-10个)核苷酸的缺失、插入和/或取代，以及在5'和/或3'端添加数个(通常为60个以内，更佳地为30个以内，更佳地为10个以内，最佳地为5个以内)核苷酸。

在本发明中，“基本纯的”蛋白质或多肽是指其至少占样品总物质的至少20%，更佳地至少50%，更佳地至少80%，最佳地至少90%(按干重或湿重计)。纯度可以用任何合适的方法进行测量，如用柱层析、PAGE或HPLC法测量多肽的纯度。基本纯的多肽基本上不含天然状态下的伴随其的组分。

在本发明中，术语“HSC3蛋白多肽”指具有HSC3蛋白活性的SEQ ID NO.4序列的多肽。该术语还包括具有与人热休克关联蛋白相同功能的、SEQ ID NO.4序列的变异形式。这些变异形式包括(但不限于): 若干个(通常为1-50个，更佳地1-30个，更佳地1-20个，最佳地1-10个)氨基酸的缺失、插入和/或取代，以及在C末端和/或N末端添加一个或数个(通常为20个以内，更佳地为10个以内，更佳地为5个以内)氨基酸。例如，在本领域中，用性能相近或相似的氨基酸进行取代时，通常不会改变蛋白质的功能。又比如，在C末端和/或N末端添加一个或数个(通常为20个以内，更佳地为10个以内，更佳地为5个以内)氨基酸通常也不会改变蛋白质的功能。该术语还包括HSC3蛋白的活性片段和活性衍生物。

该多肽的变异形式包括: 同源序列、等位变异体、天然突变体、诱导突变体、在高或低的严谨度条件下能与HSC3 DNA 杂交的DNA所编码的蛋白、以及利用抗HSC3多肽的抗血清获得的多肽或蛋白。本发明还提供了其他多肽，如包含HSC3多肽或其片段的融合蛋白。除了几乎全长的多肽外，本发明还包括了HSC3多肽的可溶性片段。通常，该片段具有HSC3多肽序列的至少约10个连续氨基酸，通常至少约30个连续氨基酸，更佳地至少约50个连续氨基酸，更佳地至少约80个连续氨基酸，最佳地至少约100个连续氨基酸。



发明还提供HSC3蛋白或多肽的类似物。这些类似物与天然HSC3多肽的差别可以是氨基酸序列上的差异，也可以是不影响序列的修饰形式上的差异，或者兼而有之。这些多肽包括天然或诱导的遗传变异体。诱导变异体可以通过各种技术得到，如通过辐射或暴露于诱变剂而产生随机诱变，还可通过定点诱变法或其他已知分子生物学的技术。类似物还包括具有不同于天然L-氨基酸的残基(如D-氨基酸)的类似物，以及具有非天然存在的或合成的氨基酸(如 β 、 γ -氨基酸)的类似物。应理解，本发明的多肽并不限于上述例举的代表性的多肽。

修饰(通常不改变一级结构)形式包括：体内或体外的多肽的化学衍生形式如乙酰化或羧基化。修饰还包括糖基化，如那些在多肽的合成和加工中或进一步加工步骤中进行糖基化修饰而产生的多肽。这种修饰可以通过将多肽暴露于进行糖基化的酶(如哺乳动物的糖基化酶或去糖基化酶)而完成。修饰形式还包括具有磷酸化氨基酸残基(如磷酸酪氨酸，磷酸丝氨酸，磷酸苏氨酸)的序列。还包括被修饰从而提高了其抗蛋白水解性能或优化了溶解性能的多肽。

本发明还包括HSC3多肽编码序列的反义序列。这种反义序列可用于抑制细胞内HSC3的表达。

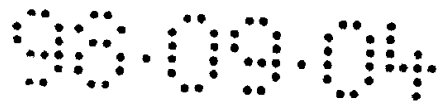
本发明还包括一种探针分子，该分子通常具有HSC3多肽编码序列的8-100个，较佳地15-50个连续核苷酸。该探针可用于检测样品中是否存在编码HSC3的核酸分子。

本发明还包括检测HSC3核苷酸序列的方法，它包括用上述的探针与样品进行杂交，然后检测探针是否发生了结合。较佳地，该样品是PCR扩增后的产物，其中PCR扩增引物对应于HSC3多肽的编码序列，并可位于该编码序列的两侧或中间。引物长度一般为20-50个核苷酸。

在本发明中，可选用本领域已知的各种载体，如市售的各种载体。

在本发明中，术语“宿主细胞”包括原核细胞和真核细胞。常用的原核宿主细胞的例子包括大肠杆菌、枯草杆菌等。常用的真核宿主细胞包括酵母细胞，昆虫细胞、和哺乳动物细胞。较佳地，该宿主细胞是真核细胞，如CHO细胞、COS细胞等。

另一方面，本发明还包括对HSC3 DNA或是其片段编码的多肽具有特异性的多克隆和单克隆抗体，尤其是单克隆抗体。这里，“特异性”是指抗体能结合于HSC3基因产物或片段。较佳地，指那些能与HSC3基因产物或片段结合但不识别和结合于其它非相关抗原分子的抗体。本发明中抗体包括那些能够结合并抑制HSC3蛋白的分子，也包括那些并不影响HSC3蛋白功能的抗体。本发明还包括那



些能与修饰或未经修饰形式的HSC3基因产物结合的抗体。

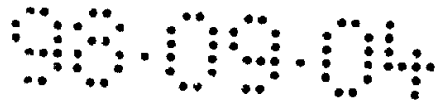
5 本发明不仅包括完整的单克隆或多克隆抗体，而且还包括具有免疫活性的抗体片段，如Fab'或(Fab)₂片段；抗体重链；抗体轻链；遗传工程改造的单链Fv分子(Ladner等人，美国专利No. 4,946,778)；或嵌合抗体，如具有鼠抗体结合特异性但仍保留来自人的抗体部分的抗体。

本发明的抗体可以通过本领域内技术人员已知的各种技术进行制备。例如，纯化的HSC3基因产物或者其具有抗原性的片段，可被施用于动物以诱导多克隆抗体的产生。与之相似的，表达HSC3或其具有抗原性的片段的细胞可用来免疫动物来生产抗体。本发明的抗体也可以是单克隆抗体。此类单克隆抗体可以利用杂交瘤技术来制备(见 Kohler 等人，Nature 256:495, 1975; Kohler 等人，Eur.J.Immunol. 6:511, 1976; Kohler 等人，Eur.J.Immunol. 6:292, 1976; Hammerling 等人，In Monoclonal Antibodies and T Cell Hybridomas, Elsevier, N.Y., 1981)。本发明的抗体包括能阻断HSC3功能的抗体以及不影响HSC3功能的抗体。本发明的各类抗体可以利用HSC3基因产物的片段或功能区，通过常规免疫技术获得。这些片段或功能区可以利用重组方法制备或利用多肽合成仪合成。与HSC3基因产物的未修饰形式结合的抗体，可以用原核细胞(例如*E. Coli*)中生产的基因产物来免疫动物而产生；与翻译后修饰形式结合的抗体(如糖基化或磷酸化的蛋白或多肽)，可以用真核细胞(例如酵母或昆虫细胞)中产生的基因产物来免疫动物而获得。

20 在本发明中，HSC3的cDNA核苷酸序列是如此获得的：以人睾丸λ gt11cDNA文库(购自Clontech公司)为模板，用一对寡核苷酸为引物A1：5'-TACCTATTTGCAGTCACTACCTC -3'(SEQ ID NO.1)为正向引物，寡核苷酸B1：5'- ATCCTTTTCTGACATACCCATTC -3'(SEQ ID NO.2)为反向引物，进行PCR。电泳检测得到约1.2kb的目的片段。经鉴定测序后得到SEQ ID NO.3的全长cDNA序列。

通过同源检索发现，本发明的HSC3的cDNA序列与已被发表并确认为的热休克关联蛋白基因序列高度同源，并且本发明HSC3蛋白也与已知的热休克关联蛋白(如鼠的热休克关联蛋白)高度同源。此外，本发明的HSC3蛋白还具有热休克关联蛋白家族高度保守的氨基酸序列。因此，这表明，与已知鼠热休克关联蛋白等热休克关联蛋白家族成员一样，本发明的HSC3蛋白也属于热休克关联蛋白家族，并且具有相似的功能。

许多慢性炎症是由抗原激活免疫系统引发的，一系列随后的反应会诱导热休



克蛋白的表达，这些蛋白反过来保护细胞和组织免受炎症的有害影响。现有迹象表明，hsp70等热休克蛋白的一些保护功能可能依赖于热休克蛋白的正常细胞功能，即陪伴变性或折叠不正确的蛋白以及伴随它们穿过细胞膜的功能 (Experientia (1994)50: 1031-1038)。有实验表明hsp70家族的成员可改善抗原加工及递呈，从而使免疫应答更有效(Int. Immun. 6(1994)925-930)。这都说明hsp70在炎症中的保护作用，而且hsp70对于真核细胞中受损蛋白的复性也是必需的。DnaJ主要功能是协助hsp70及其同源蛋白的功能，例如蛋白折叠、蛋白装配等。由于与DnaJ蛋白同源，因此本发明的HSC3在蛋白质代谢过程中也具有类似于DnaJ的功能，特别是在炎症反应中可能起到保护作用。

下面结合具体实施例，进一步阐述本发明。应理解，这些实施例仅用于说明本发明而不用来限制本发明的范围。下列实施例中未注明具体条件的实验方法，通常按照常规条件如Sambrook等人，分子克隆：实验室手册(New York: Cold Spring Harbor Laboratory Press, 1989)中所述的条件，或按照制造厂商所建议的条件。

实施例1

HSC3的cDNA的克隆和测序

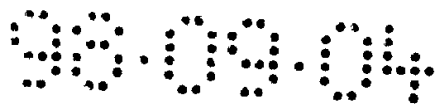
1. PCR扩增

以人睾丸 λ gt11cDNA文库(购自Clontech公司)为模板，用一对寡核苷酸为引物A1: 5'-TACCTATTTGCAGTCACTACCTC-3'(SEQ ID NO.1)为正向引物，寡核苷酸B1: 5'-ATCCTTTTCTGACATACCCATTC-3'(SEQ ID NO.2)为反向引物，进行PCR。A1/B1的PCR条件为93℃ 4分钟，随之以93℃ 1分钟、62℃ 1分钟和72℃ 1分钟进行35个循环，最后72℃ 延伸5分钟。电泳检测得到约1.2kbp的目的片段。

2. PCR产物的测序

将步骤1中获得的PCR扩增产物与pGEM-TTM载体(Promega)连接，转化大肠杆菌JM103，用QIAprep Plasmid试剂盒(QIAGEN)提取质粒，用双链嵌套式缺失试剂盒(Pharmacia)对插入片段进行定向系列缺失，然后用PCR对缺失子进行快速鉴定及排序。用SequiTherm EXCELTM DNA测序试剂盒(Epicentre Technologies)对依次截短的缺失子进行测序，最后用电脑软件拼接顺序，获得全长cDNA序列，共1203bp，详细序列见SEQ ID NO: 3，其中开放读框位于108-1037位核苷酸。

根据得到的全长基因组序列推导出HSC3的氨基酸序列，共309个氨基酸残基，其氨基酸序列详见SEQ ID NO: 4。



实施例2

HSC3同源比较

本发明的全长cDNA顺序进行Basic BLAST同源比较, 采用的数据库是非冗余的GenBank+EMBL+DDBJ+PDB, 结果发现与鼠的热休克关联蛋白 (数据库编号及名称: gb|AF03596|MRJ)同源度最高, 核苷酸顺序与对应的MRJ相似度为90%。

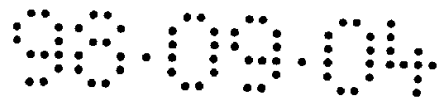
用PCGENE软件将本发明的全长cDNA顺序翻译成蛋白质顺序, ORF位于108-1037位核苷酸, 编码一段长309个氨基酸的多肽。然后将得到的氨基酸顺序进行Basic BLAST同源比较, 采用的数据库是非冗余的GenBank 翻译CDS+PDB+SwissPort+更新SP+PIR, 结果发现与鼠的热休克关联蛋白(数据库编号及名称: gi|3142372|MRJ)同源度最高, 氨基端68个氨基酸残基98%相同, 羧基端的四分之一顺序也有96%相似, 92%相同。

Hsp70是一类热休克蛋白, 它作用的基本机理是在依赖ATP的方式下结合、释放非天然构像的多肽。研究发现hsp70不是单独发挥作用的, 需要另外的热休克蛋白例如DnaJ的参与作用, DnaJ作为调控因子调节hsp70的活性。大肠杆菌DnaJ主要功能特征是直接与hsp70结合, 刺激hsp70的ATP酶活(Trends. Biochem. Sci. 19(1994), 20-25; Trends. Biochem. Sci. 17(1992), 295-299)。在酵母、植物、人类中已鉴定出一些与大肠杆菌DnaJ同源的基因。

遗传和生化实验表明真核生物DnaJ 类似蛋白与大肠杆菌DnaJ功能相关(Cell 71(1992), 1143-1155; J. Biol. Chem. 267(1992), 20927-20931)。比较DnaJ家族成员的氨基酸顺序, 发现所有的DnaJ类似蛋白含有J 结构域, 这是一段对应于大肠杆菌DnaJ氨基端70个氨基酸的一段顺序, 而且DnaJ与hsp70结合并调节其ATP活性很有可能通过该结构域发挥作用(Trends. Biochem. Sci. 19(1994), 176-181)。本发明HSC3氨基端68个氨基酸残基与大肠杆菌的J结构域相似性达81%, 表明HSC3也有J结构域, 属于DnaJ类似蛋白, 是热休克关联蛋白家族的新成员。

在不同的生物中, DnaJ类似蛋白直接或间接地参与胞内活动。它作为调控因子调节hsp70的活性, 正是通过J结构域识别hsp70而发挥功能的(Trends. Biochem. Sci. 19(1994), 176-181)。本发明HSC3作为热休克关联蛋白家族的新成员, 特别是作为DnaJ类似蛋白, 也有调节hsp70的功能, 与蛋白质的代谢活动密切相关。

许多慢性炎症是由抗原激活免疫系统引发的, 其中主要被激活的是单核细胞、巨噬细胞。这些吞噬性的细胞产生大量的活性氧物质(ROS)和细胞因子, 从而调节HSP蛋白的表达, 这些蛋白反过来保护细胞和组织免受炎症的有害影响。保护的机理包括阻碍ROS引发的DNA断裂、脂质过氧化。现有迹象表明, hsp70



等热休克蛋白的一些保护功能可能依赖于热休克蛋白的正常细胞功能，即陪伴变性或折叠不正确的蛋白以及伴随它们穿过细胞膜(Experientia, 1038)。有报道，hsp70家族的一些成员参与抗原递呈及加工的不同步骤(J. Expl. Med. (1989)170: 1799-1803)。有人提出HSP不仅参与主要组织相容性复合物II的装配，还参与随后的抗原递呈(Critical. Rev. Immun.13(1993),71-81)，已有实验表明hsp70家族的成员可改善抗原加工及递呈，从而使免疫应答更有效(Int. Immun. 6(1994),925-930)。这都说明hsp70在炎症中的保护作用，而且hsp70对于真核细胞中受损蛋白的复性也是必需的。与同源的DnaJ蛋白相似(J. Biol. Chem. 271(1996), 11236-11246)，本发明的HSC3蛋白主要功能可能是协助hsp70或同源蛋白的功能，包括蛋白质的折叠、跨细胞内膜运输、蛋白水解、蛋白装配，以及蛋白集合体和多蛋白结构的解装配。这提示本发明在蛋白质代谢过程中起着类似DnaJ的作用，特别是在炎症反应中可能起到保护作用。

实施例3

15 HSC3在大肠杆菌中的表达

在该实施例中，编码HSC3的DNA序列用对应于该DNA序列的5'和3'端的PCR寡核苷酸引物进行扩增，以获得HSC3插入片段。

PCR反应5'寡核苷酸引物序列为：

5'-GGAAGTCGAC ATGGTGGATT ACTATGAAG-3'(SEQ ID NO.5)

20 该引物含有SalI限制性内切酶的酶切位点，在酶切位点之后是由起始密码子开始的HSC3编码序列的19个核苷酸；

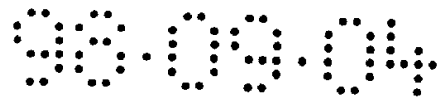
3'端引物序列为：

5'-GTGCTGCAGT TAACAATTCC TTTTGGTAG-3'(SEQ ID NO.6)

25 该引物含有PstI限制性内切酶的酶切位点、一个翻译终止子和HSC3的编码序列。

引物上的限制性内切酶的酶切位点对应于细菌表达载体pQE-9(Qiagen Inc., Chatsworth, CA)上的限制性内切酶酶切位点，该质粒载体编码抗生素抗性(Amp^r)、一个细菌复制起点(ori)、一个IPTG-可调启动子/操纵子(P/O)、一个核糖体结合位点(RBS)、一个6-组氨酸标记物(6-His)以及限制性内切酶克隆位点。

30 用SalI和PstI消化pQE-9载体以及插入片段，随后将插入片段连接到pQE-9载体并保持开放读框在细菌RBS起始。随后用连接混合物转化购自Qiagen，商品名为M15/rep4的E.coli菌株，M15/rep4含有多拷贝的质粒pREP4，其表达lacI阻遏物



并携带卡那霉素抗性(Kan^r)。在含有Amp和Kan的LB培养皿上筛选转化子，抽提质粒，测序验证HSC3的cDNA片段已正确插入了载体。

在补加Amp(100 μg/ml)和Kan(25 μg/ml)的LB液体培养基中过夜培养(O/N)含所需构建物的阳性转化子克隆。过夜(O/N)培养物以1: 100-1: 250的稀释率稀
5 释，然后接种到大体积培养基中，培养细胞至600光密度(OD₆₀₀)为0.4-0.6，随后加入IPTG(“异丙基硫代-β-D-半乳糖苷”)至终浓度为1mM。通过使lacI阻遏物失活，IPTG诱导启动P/O导致基因表达水平提高。继续培养细胞3-4小时，随后离心(6000 × g, 20分钟)。超声裂解包涵体，收集细胞并将细胞沉淀溶于6M的盐酸胍中。澄清后，通过在能使含6-His标记物蛋白紧密结合的条件下，用镍-螯合柱层
10 析从溶液中纯化溶解的HSC3。用6M盐酸胍(pH5.0)从柱中洗脱HSC3。可用几种方法从盐酸胍中变性沉淀蛋白。或者，使用透析步骤除去盐酸胍，或者从镍-螯合柱中分离出的纯化蛋白。纯化后的蛋白可以结合到第二个柱中，该柱中具有递减的线性盐酸胍梯度。在结合到该柱时蛋白质变性，随后用盐酸胍(pH5.0)洗脱。最后，将可溶的蛋白质对含PBS进行透析，然后将蛋白质保存在终浓度为10%(w/v)甘
15 油的贮存液中。

实施例4

HSC3在真核细胞(CHO细胞株)中的表达

参照实施例3，编码HSC3的DNA序列用对应于该DNA序列的5'和3'端的PCR
20 寡核苷酸引物进行扩增，以合成HSC3基因作为插入片段。

PCR反应中使用的5'寡核苷酸引物序列为：

5'-GCAACTGCAG ATGGTGGATT ACTATGAAG -3'(SEQ ID NO.7)

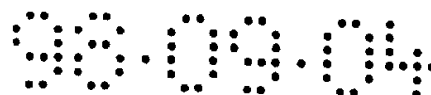
该引物含有PstI限制性限制性内切酶的酶切位点，接之是由起始密码子开始的HSC3编码序列的19个核苷酸；

25 3'端引物序列为：

5'-TCGGATATCT TAACAATTCC TTTTGGTAG -3'(SEQ ID NO.8)

该引物含有EcoRV限制性限制性内切酶的酶切位点、一个翻译终止子和HSC3的编码序列。

引物上的限制性内切酶的酶切位点对应于CHO细胞表达载体
30 pcDNA3(Invitrogen公司)上的限制性内切酶酶切位点，该质粒载体编码抗生素抗性(Amp^r和Neo^r)、一个噬菌体复制起点(f1 ori)、一个病毒复制起点(SV40 ori)、一个病毒启动子(P-CMV)、一个Sp6启动子、一个SV40启动子、一个SV40加尾信号



和相应的polyA顺序、一个BGH加尾信号和相应的polyA顺序。

用PstI和EcoRV消化pcDNA3载体及插入片断，随后将插入片段连接到pcDNA3载体。随后用连接混合物转化E.coli DH5 α 菌株。在含有Amp的LB培养皿上筛选转化子，在补加Amp(100 μ g/ml)的LB液体培养基中过夜培养(O/N)含所需构建物的克隆。抽提质粒，用BamHI酶切鉴定插入片段大小及方向，并测序验证HSC的cDNA已正确插入载体。

质粒转染CHO细胞是采用脂转染法，用Lipofectin试剂盒(GiBcolife)进行的。转染48小时后，经2-3周的持续G418加压筛选，收集细胞及细胞上清测定表达蛋白酶活力。去G418，连续传代培养；对混合克隆细胞极限稀释，选择具有较高蛋白活性的细胞亚克隆。按常规方法大量培养上述阳性亚克隆。48小时后，开始收集细胞及上清，用超声裂解方法破碎细胞。以含0.05%Triton的50mM Tris \cdot HCl(pH7.6)溶液为平衡液及洗脱液，用经预平衡的Superdex G-75柱收集上述蛋白的活性峰。再用50mM Tris \cdot HCl(pH8.0)平衡的DEAE-Sepharose柱，以含1M NaCl的50mM Tris \cdot HCl(pH8.0)溶液为洗脱液进行梯度洗脱，收集上述蛋白的活性峰。然后以PBS(pH7.4)为透析液对表达蛋白溶液进行透析。最后冻干保存。

用12%的SDS-PAGE胶进行电泳，鉴定表达蛋白的分子量大小为35kD。

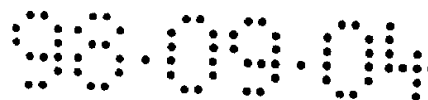
此外，用常规方法对表达蛋白的N端和C端各10个氨基酸长度的氨基酸进行测序，发现与SEQ ID NO: 4的序列一致。

20 实施例5

制备抗体

将实施例3或4中获得的重组蛋白用来免疫动物以产生抗体，具体如下。重组分子用层析法进行分离后备用。也可用SDS-PAGE凝胶电泳法进行分离，将电泳条带从凝胶中切下，并用等体积的完全Freund's佐剂乳化。用50-100 μ g/0.2ml乳化的蛋白，对小鼠进行腹膜内注射。14天后，用非完全Freund's佐剂乳化的同样抗原，对小鼠以50-100 μ g/0.2ml的剂量进行腹膜内注射以加强免疫。每隔14天进行一次加强免疫，至少进行三次。获得的抗血清的特异反应活性用它在体外沉淀HSC3蛋白的能力加以评估。

在本发明提及的所有文献都在本申请中引用作为参考，就如同每一篇文献被单独引用作为参考那样。此外应理解，在阅读了本发明的上述讲授内容之后，本领域技术人员可以对本发明作各种改动或修改，这些等价形式同样落于本申请所附权利要求书所限定的范围。



序列表

(2)SEQ ID NO: 1的信息

(i)序列特征

- 5 (A)长度: 23碱基
(B)类型: 核酸
(C)链性: 单链
(D)拓扑结构: 线性

(ii)分子类型: 寡核苷酸

10 (xi)序列描述: SEQ ID NO: 1

TACCTATTTG CAGTCACTAC CTC 23

(2)SEQ ID NO: 2的信息

(ii)序列特征

- 15 (A)长度: 23碱基
(B)类型: 核酸
(C)链性: 单链
(D)拓扑结构: 线性

(ii)分子类型: 寡核苷酸

20 (xi)序列描述: SEQ ID NO: 2

ATCCTTTTCT GACATACCCA TTC 23

(2)SEQ ID NO: 3的信息:

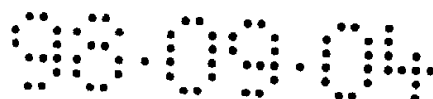
(i)序列特征:

- 25 (A)长度: 1203bp
(B)类型: 核酸
(C)链性: 单链
(D)拓扑结构: 线性

(ii)分子类型: cDNA

30 (xi)序列描述: SEQ ID NO: 3

1 TACCTATTTG CAGTCACTAC CTCTATCACC ACCACCAGCA GAGCCTGAGC TGAGGAAACC
61 ACGGTTCTCA ATACCCAGCA CACCCACTTC CAACTATCTG TTAAAACATG GTGGATTACT



121 ATGAAGTTCT AGGACTGCAA AGATATGCTT CACCTGAGGA CATTAAAAA GCTTATCATA
181 AAGTGGCACT TAAATGGCAC CCTGATAAAA ATCCAGAAAA TAAAGAAGAA GCAGAGAGAA
241 AATTCAAAGA AGTAGCTGAG GCATACGAGG TATTATCAAA TGATGAGAAA CGGGACATTT
301 ATGATAAATA TGGCACAGAA GGATTAACG GAGGTGGAAG TCATTTTGAT GATGAATGTG
5 361 AGTACGGCTT CACATTCCAT AAGCCAGATG ATGTTTTTAA AGAAATTTT CATGAAAGGG
421 ATCCATTTTC TTTTCACTTC TTTGAAGACT CGCTTGAGGA CCTGTAAAT CGTCCAGGAA
481 GCTCCTATGG AAACAGAAAC AGAGATGCAG GATACTTTTT CTCCACTGCC AGTGAATATC
541 CAATTTTTGA GAAATTTTCT TCATATGATA CAGGATATAC ATCACAGGGT TCATTGGGGC
601 ATGAAGGCCT TACTTCTTTC TCTTCCCTGG CTTTTGATAA TAGTGGGATG GACAACTACA
10 661 TATCTGTTAC AACTTCAGAC AAAATCGTTA ATGGCAGAAA TATTAATACA AAGAAAATTA
721 TTGAAAGTGA TCAAGAAAGA GAAGCTGAAG ATAATGGAGA GTTGACATTT TTTCTTGTA
781 ATAGTGTGGC CAATGAAGAG GGCTTTGCAA AAGAATGCAG CTGGAGAACA CAGTCATTCA
841 ACAACTATTC ACCAAATTCT CACAGCTCCA AACATGTATC TCAATATACT TTCGTGGACA
901 ATGATGAGGG AGGTATATCT TGGGTTACCA GCAACAGAGA TCCCCTATT TTCTCAGCAG
15 961 GAGTCAAAGA GGGTGGTAAG AGGAAAAAAA AGAAGCGTAA AGAGGTGCAA AAGAAGTCTA
1021 CCAAAGGAA TTGTAAATT GACTCTTCAA ATATATAACA TTTGAACACA ATTGTGTGTG
1081 TTTTGGTTAA TCACAAATTT TGTAGATAAC ACTTAATACT ATACTAAGAG CTTTTCAACA
1141 CTTTTAGCAG GATTGTGGAC ATTTGGTTAG TAGTTTTTTT GAATGGGTAT GTCAGAAAAG
1201 GAT

20

(2)SEQ ID NO: 4的信息:

(i)序列特征:

(A)长度: 309个氨基酸

(B)类型: 氨基酸

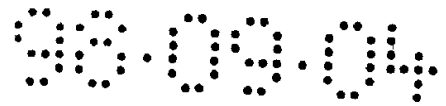
25

(D)拓扑结构: 线性

(ii)分子类型: 多肽

(xi)序列描述: SEQ ID NO: 4

1 Met Val Asp Tyr Tyr Glu Val Leu Gly Leu Gln Arg Tyr Ala Ser
16 Pro Glu Asp Ile Lys Lys Ala Tyr His Lys Val Ala Leu Lys Trp
30 31 His Pro Asp Lys Asn Pro Glu Asn Lys Glu Glu Ala Glu Arg Lys
46 Phe Lys Glu Val Ala Glu Ala Tyr Glu Val Leu Ser Asn Asp Glu
61 Lys Arg Asp Ile Tyr Asp Lys Tyr Gly Thr Glu Gly Leu Asn Gly



76 Gly Gly Ser His Phe Asp Asp Glu Cys Glu Tyr Gly Phe Thr Phe
91 His Lys Pro Asp Asp Val Phe Lys Glu Ile Phe His Glu Arg Asp
106 Pro Phe Ser Phe His Phe Phe Glu Asp Ser Leu Glu Asp Leu Leu
121 Asn Arg Pro Gly Ser Ser Tyr Gly Asn Arg Asn Arg Asp Ala Gly
5 136 Tyr Phe Phe Ser Thr Ala Ser Glu Tyr Pro Ile Phe Glu Lys Phe
151 Ser Ser Tyr Asp Thr Gly Tyr Thr Ser Gln Gly Ser Leu Gly His
166 Glu Gly Leu Thr Ser Phe Ser Ser Leu Ala Phe Asp Asn Ser Gly
181 Met Asp Asn Tyr Ile Ser Val Thr Thr Ser Asp Lys Ile Val Asn
196 Gly Arg Asn Ile Asn Thr Lys Lys Ile Ile Glu Ser Asp Gln Glu
10 211 Arg Glu Ala Glu Asp Asn Gly Glu Leu Thr Phe Phe Leu Val Asn
226 Ser Val Ala Asn Glu Glu Gly Phe Ala Lys Glu Cys Ser Trp Arg
241 Thr Gln Ser Phe Asn Asn Tyr Ser Pro Asn Ser His Ser Ser Lys
256 His Val Ser Gln Tyr Thr Phe Val Asp Asn Asp Glu Gly Gly Ile
271 Ser Trp Val Thr Ser Asn Arg Asp Pro Pro Ile Phe Ser Ala Gly
15 286 Val Lys Glu Gly Gly Lys Arg Lys Lys Lys Lys Arg Lys Glu Val
301 Gln Lys Lys Ser Thr Lys Arg Asn Cys

(2)SEQ ID NO: 5的信息

(i)序列特征

- 20 (A)长度: 29碱基
(B)类型: 核酸
(C)链性: 单链
(D)拓扑结构: 线性
(ii)分子类型: 寡核苷酸

25 (xi)序列描述: SEQ ID NO: 5

GGAAGTCGAC ATGGTGGATT ACTATGAAG 29

(2)SEQ ID NO: 6的信息

(i)序列特征

- 30 (A)长度: 29碱基
(B)类型: 核酸
(C)链性: 单链

(D)拓扑结构: 线性
 (iii)分子类型: 寡核苷酸
 (xi)序列描述: SEQ ID NO: 6
 GTGCTGCAGT TAACAATTCC TTTTGGTAG 29

5

(2)SEQ ID NO: 7的信息

(i)序列特征

(A)长度: 29碱基

(B)类型: 核酸

10

(C)链性: 单链

(D)拓扑结构: 线性

(iv)分子类型: 寡核苷酸

(xi)序列描述: SEQ ID NO: 7

GCAACTGCAG ATGGTGGATT ACTATGAAG 29

15

(2)SEQ ID NO: 8的信息

(i)序列特征

(A)长度: 29碱基

(B)类型: 核酸

20

(C)链性: 单链

(D)拓扑结构: 线性

(v)分子类型: 寡核苷酸

(xi)序列描述: SEQ ID NO: 8

TCGGATATCT TAACAATTCC TTTTGGTAG 29