



(12)发明专利

(10)授权公告号 CN 109293783 B

(45)授权公告日 2019.10.11

(21)申请号 201811254050.7

C12N 15/70(2006.01)

(22)申请日 2018.10.25

C12N 1/21(2006.01)

C12R 1/19(2006.01)

(65)同一申请的已公布的文献号

申请公布号 CN 109293783 A

(43)申请公布日 2019.02.01

(73)专利权人 山西锦波生物医药股份有限公司

地址 030032 山西省太原市经济技术开发区
经北街18号

(72)发明人 杨霞

(74)专利代理机构 北京市柳沈律师事务所

11105

代理人 张文辉

(51)Int.Cl.

C07K 19/00(2006.01)

C12N 15/62(2006.01)

(56)对比文件

吴铭等.重组胶原蛋白多肽在大肠杆菌中的表达优化.《湖北农业科学》.2014,第53卷(第10期),2443-2447.

杨立霞等.重组生产胶原蛋白的研究进展.《河北化工》.2007,第30卷(第7期),43-45.

徐立群.类人胶原蛋白真核表达载体的构建及在毕赤酵母中的分泌表达.《中国优秀硕士学位论文全文数据库 基础科学辑》.2013,(第S2期),A006-189.

审查员 李颖

权利要求书1页 说明书10页

序列表10页 附图3页

(54)发明名称

多肽、其生产方法和用途

(57)摘要

本发明涉及多肽、其生产方法和用途。所述多肽包含N端区域和C端区域。本发明的多肽具有显著的细胞黏附效果。

1. 多肽,所述多肽由以SEQ ID No. 5或SEQ ID No. 6所示的氨基酸序列组成。
2. 多核苷酸,其编码根据权利要求1所述的多肽。
3. 表达载体,其包含根据权利要求2所述的多核苷酸。
4. 宿主细胞,其包含根据权利要求3所述的表达载体。
5. 根据权利要求4所述的宿主细胞,其中所述宿主细胞是大肠杆菌。
6. 根据权利要求1所述的多肽的生产方法,其包括:
 - (1) 在生产培养基中培养根据权利要求4或5所述的宿主细胞并生产多肽;和
 - (2) 收获并纯化多肽。
7. 根据权利要求6所述的生产方法,所述方法包括对多肽进行酶切的步骤。
8. 组合物,其包含根据权利要求1所述的多肽。
9. 根据权利要求8所述的组合物,其中所述组合物是医疗器材、组织工程产品、化妆品或保健品。
10. 根据权利要求1所述的多肽在制备组合物中的用途。
11. 根据权利要求10所述的用途,其中所述组合物是医疗器材、组织工程产品、化妆品或保健品。
12. 根据权利要求1所述的多肽在制备用于促进细胞黏附的药物中的用途。

多肽、其生产方法和用途

技术领域

[0001] 本发明属于基因工程技术领域,涉及多肽、其生产方法和用途。

背景技术:

[0002] 胶原蛋白一般为白色、透明、无分支的原纤维,是皮肤和骨骼的基础支撑物,可以占到蛋白质总量的25%~35%,主要分布于人体的皮肤、血管、骨骼、筋腱、牙齿和软骨等处,是这些组织的主要基质和支架,保护并连结各种组织,在体内发挥着重要的生理功能。因此,胶原蛋白可以广泛的应用在医药和化妆品等行业中。

[0003] 当前市场上销售的胶原蛋白产品都是取自猪、牛、鱼等动物组织中。以胶原蛋白的氨基酸组成而言:哺乳动物猪、牛与人的相似度为95%,鱼与人的相似度65%。虽然哺乳动物猪、牛等与人的胶原蛋白相似度较高,其仍然难以避免病毒感染以及致敏性的危险。而食用鱼类所萃取的胶原蛋白人体的利用率低于65%,无法完全被人体吸收与利用。所以其常用于食品运用,少部分运用于化妆品,但无法被用于医疗器材或较精密的组织工程产品。所以,目前的胶原蛋白只能在化妆品和保健品中使用,根本无法发挥胶原蛋白的原本生物学功能。

[0004] 从结构上来说,人体天然的胶原蛋白的结构非常的复杂,所以才导致人源胶原蛋白极难通过常规手段表达和大量制备。胶原蛋白最普遍的结构特征是由3条肽链形成的三螺旋结构,即由3条A肽链以右手超螺旋方式形成蛋白质,这样的三股螺旋区域被称为胶原区域。每个A肽链在分子结构上都是由重复出现的Gly-X-Y(X、Y代表Gly之外的任何氨基酸残基,X往往是Pro,Y往往是Hyp)肽段构成左手螺旋,3条链在氨基酸残基的相互作用下,以同一轴为中心,以右手超螺旋方式形成稳定的三股螺旋结构。在生物体中,胶原蛋白的合成和修饰从原胶原开始,经历了羟基化、糖基化、相互交联等诸多化学变化,受到了多种生物酶的复杂调控。原胶原除了含有胶原链之外,还含有球状的头部和尾部。没有这些头部和尾部,胶原链就不会折叠成为正确的三螺旋,从而缺乏胶原蛋白的生物学活性。因此,按照原始基因序列制备的胶原蛋白不可能在体外自发的组织形成正确的空间结构。这样的困难严重阻碍了人胶原蛋白的研发和生产。

[0005] 生产胶原蛋白的传统方法是利用酸、碱、酶解法处理动物来源的组织,提取胶原蛋白衍生物。这些方法提取的胶原蛋白本身已经丧失了原本的生物学活性,无法应用于生物医学领域发挥真正的功能。大多数经过制备后的胶原蛋白产品,拉伸强度较弱;纯胶原蛋白在体内降解较快,以及可能存在潜在的抗原性等,同时由于胶原蛋白的来源、加工工艺和原料配比的差异,产品的营养成分及饲用价值也不相同,且皮料在加工过程中不仅要接触许多化学物质,且易受到细菌的感染,严重限制了胶原蛋白的应用。虽然国外研究机构通过培育含人胶原蛋白基因的小鼠,得到了含有人胶原蛋白的乳汁,但是这样生产的成本过高,生产周期过长,无法投入大规模生产。

发明内容

[0006] 针对上述现有技术的缺陷,本发明提供了:

[0007] 1.多肽,所述多肽包含SEQ ID No.1(人I型胶原蛋白COL1A1的全长序列)中的至少60个连续的氨基酸残基的N端区域和包含以SEQ ID No.2所示的序列GAPGPCCGG的C端区域,

[0008] 其中优选地,所述多肽是胶原蛋白多肽,优选是人源胶原蛋白多肽;

[0009] 其中优选地,所述N端区域为重复n次的基本重复单元,n为大于等于1的整数,优选地为1、2、3、4、5、6、7、8、9或10,优选地n为4;

[0010] 其中优选地,所述N端区域与C端区域是连续的或者间隔1个或多个氨基酸残基。

[0011] 2.根据1所述的多肽,其中所述N端区域包含选自SEQ ID No.3和SEQ ID No.4的序列。

[0012] 3.1或2所述的多肽,其包含

[0013] a) SEQ ID No.5或SEQ ID No.6的氨基酸序列

[0014] b) 与SEQ ID No.5或SEQ ID No.6的氨基酸序列具有90%、92%、95%、96%、97%、98%或99%同一性的氨基酸序列,其保留SEQ ID No.5或SEQ ID No.6的氨基酸序列的细胞黏附效果;

[0015] c) SEQ ID No.5或SEQ ID No.6的氨基酸序列中添加、取代、缺失或插入1个或多个氨基酸残基的氨基酸序列,优选为保守氨基酸取代,所述氨基酸序列保留SEQ ID No.5或SEQ ID No.6的氨基酸序列的细胞黏附效果;或

[0016] d) 由核苷酸序列编码的氨基酸序列,所述核苷酸序列与编码SEQ ID No.5或SEQ ID No.6的氨基酸序列的多核苷酸序列在严格条件下杂交,其保留SEQ ID No.5或SEQ ID No.6的氨基酸序列的细胞黏附效果,所述严格条件是中等严格条件,中-高严格条件,高严格条件或非常高严格条件。

[0017] 4.多核苷酸,其编码根据1-3中任一项所述的多肽。

[0018] 5.表达载体,其包含根据4所述的多核苷酸。

[0019] 6.宿主细胞,其包含根据5所述的表达载体,其中所述宿主细胞优选是大肠杆菌。

[0020] 7.根据权利要求1所述的多肽的生产方法,其包括:

[0021] (1) 在生产培养基中培养根据6所述的宿主细胞并生产多肽;

[0022] (2) 收获并纯化多肽;和

[0023] (3) 任选地对多肽进行酶切。

[0024] 8.组合物,优选医疗器材、组织工程产品、化妆品或保健品,其包含根据1所述的多肽。

[0025] 9.根据1所述的多肽在制备组合物,优选医疗器材、组织工程产品、化妆品、保健品中的用途。

[0026] 10.根据1所述的多肽促进细胞黏附的用途。

[0027] 与现有技术相比,本发明具有以下特点:

[0028] (1) 本发明首次选择的I型胶原蛋白序列为长期筛选优化的序列;相比于I型胶原蛋白的其它片段具有显著更好的细胞黏附效果。本发明的重组I型胶原蛋白片段以较短的序列长度便更好地实现了I型胶原蛋白片段的细胞黏附活性。

[0029] (2) 采用大肠杆菌表达系统,适于大规模放大,20小时即可完成一轮发酵,生产成本非常低,由于对基因序列进行了大肠杆菌的密码子优化以及选用2×YT培养基,使得产量非常大;

[0030] (3) 生产的重组人源胶原蛋白具有非常好的亲水性和稳定性,其氨基酸组成与天然胶原蛋白氨基酸序列相应部分100%相同,应用于人体不会产生免疫排斥和过敏反应,可以广泛应用于生物医药和化妆品行业;

[0031] (4) 本发明产品经过活性检测,具有达到甚至超过人体天然蛋白的生物学活性,是目前报道过的最优化的胶原蛋白设计方案,可以在人体中行使天然蛋白的功能,达到真正的产品应用的目的。

附图说明

[0032] 图1为用于本发明载体pET32a-C1S4T (SEQ ID NO.5)、pET32a-C1S5T (SEQ ID NO.6) 构建的质粒pET32a的图谱;

[0033] 图2为本发明C1S4T蛋白纯化酶切后得到的目的蛋白电泳图;C1S4T蛋白的电泳检测分子量约为36kDa,对应于包含SEQ ID NO.5的氨基酸序列的蛋白质。

[0034] 图3为本发明C1S5T蛋白纯化酶切后得到的目的蛋白电泳图;C1S5T蛋白的电泳检测分子量约为36kDa,对应于包含SEQ ID NO.6的氨基酸序列的蛋白质。

[0035] 图4为本发明C1S4T蛋白与人胶原蛋白相比较的生物活性检测结果。

[0036] 图5为本发明C1S5T蛋白与人胶原蛋白相比较的生物活性检测结果。

具体实施方式

[0037] 下文提供进一步的描述以便于理解本发明。

[0038] 如本文中使用的,“医疗器械”是指直接或者间接用于人体的仪器、设备、器具、体外诊断试剂及校准物、材料以及其他类似或者相关的物品。

[0039] 如本文中使用的,“组织工程产品”是指用于组织工程的产品。组织工程是一门以细胞生物学和材料科学相结合,进行体外或体内构建组织或器官的新兴学科。

[0040] 在本发明中,选择I型胶原蛋白序列为筛选优化的序列。所述人胶原I型 COL1A1的序列是NCBI参照序列:NP_000079 (SEQ ID No.1),参见 https://www.ncbi.nlm.nih.gov/protein/NP_000079.2

[0041] MFSFVDLRLLLLLLAATALLTHGQEEGQVEGQDEDIPPITCVQNGRLRYHDR DVWKPEPCRICVCDNGK
VLCDDVICDETKNCPGAEVPEGECCPVC PDGS ESPTDQETTGVGPKGDTGPRGPRGPAGPPGRDGI PGQPGLPG
PPGPPGPP GPPGLGGNFAPQLSYGYDEKSTGGISVPGPMGSPGRGLPGPPGAPGPQG FQPPGEPGEPGASGP
MGPRGPPGPPGKNGDDGEAGKPRPGERGPPGP QGARGLPGTAGLPGMKGHRGFSGLDGA KG DAGPAGPKGEPGS
PGENGA PGQMGRPLPGERGRPGAPGARGNDGATGAAGPPGPTGPAGPPGFP GAVGAKGEAGPQGRGSE
PQGVREPGPPGAGAAGPAGNPGADGQP GAKGANGAPGIAGAPGFPGARGPSGPQGGPPGPKGNSGEPGAPG
SKG DTGAKGEPGPVGVQGGPPAGEEGKRGARGEPTGLPGPPGERGGPGS RGFPGADVAGPKGPAGERGSP
GPAGPKGSPGEAGRPGEAGLP GAKGLT GSPGSPGPDGKTGPPGAGQDGRGPPGPPGARGQAGVMGFPGPKGAAG
EPGKAGERGVPPGAVGPAGKDGEAGA QGPPGAPGAGERGEQGPAGSPGFQGLPGPAGPPGEAGKPEQGVPGD
LGAPGSPGARGERGFGERGVQGGPPGAPRGANGAPGNDGAKGDAGAPGAPGSQGAPGLQMPGE RGAAGLPGP

KGDRGDAGPKGADGSPGKDGVRGLTGPIGPPGAPAGPGDK GESGSPGAPGTGARGAPGDRGEPGPPGAPGAFAGP
 PGADGQPAGKGEPEG DAGAKGDAGPPGAPGAPGPPGPIGNVAGAPGAKGARGGAPGPPGATGFPGA AGRVGPAGPSG
 NAGPPGPPGAPGKEGKGRGETGPAGRPGEVGPAGPPG PAGEKGSPPGADGAPGAPGTPGPQGIAGQRGVVGLPG
 QRGERGFPLPGPSGEPGKQGPSGASGERGPPGPMGPPGLAGPPGESGREGAPGAEGSP GRDGSPGAKGDRGETG
 PAGPPGAPGAPGAPGVPAGKSGDRGETGPAG PAGVPVGVARGPAGPQGRGDKGETGEQGDRIKGRHGFSG
 LQGPAGP PGSPGEQGPSGASGPAGPRGPPGSAGAPGKDLNGLPGPIGPPGPRGRGTG DAGVPGPPGPPGPPGPP
 GPPSAGFDFSLPQPPQEKAHDGGRYRADDAN VVRDRDLEVDTTLSLSQQIENIRSPEGSRKNPARTCRDLKM
 CHSDWKS EYWIDPNQGCNLDAIKVFCNMETGETCVYPTQPSVAQKNWYISKPNPKDK RHVWFGESMTDGFQFEY
 GGQGS DPADVAIQLTFLRLMSTEASQNTYHC KNSVAYMDQQTGNLKKALLLQGSNEIEIRAEGNSRFTYSVTVD
 GCTSHTG AWGKTVIEYKTTKTSRLPIIDVAPLDVGAPDQEFQFDVGPVCF (SEQ ID No.1)

[0042] 上述序列中粗体下划线部分即为本发明选择的氨基酸序列。申请人经过大量的研究发现,选择的上述序列比商品化的人胶原蛋白或SEQ ID No.1中的其他序列实现更好的黏附效果。在本发明中,多肽不是SEQ ID No.1的全长序列。

[0043] 本发明部分基于以下发现:包含SEQ ID No.1中的至少60个连续的氨基酸残基的N端区域和以SEQ ID No.2所示的序列GAPGPCCGG的C端区域的多肽能够比商品化的人胶原蛋白实现更好的黏附效果,如实施例证明。本领域技术人员可以适当选择构成N端区域的连续的氨基酸残基。例如,连续的氨基酸残基的长度可以是60-100、72-84等等。

[0044] 在本发明中,对几种具体的N端区域进行了测试:

[0045] (1) GPAGSPGFQGLPGPAGPPGEAGKPGEQGVPGDLGAPGPSGARGER GFPGERGVQPPGPA (SEQ ID No.3);

[0046] (2) GEKGSPPGADGAPGAPGTPGPQGIAGQRGVVGLPGQRGERGFPLPGPS GEPGKQGPSGAS (SEQ ID No.4);

[0047] 在本发明中,C端氨基酸序列可以为GAPGPCCGG (SEQ ID No.2),该序列增强胶原活性的端序列肽段。

[0048] 多肽在本文中可以是重组人源胶原蛋白C1S4T,为单链结构,包括氨基酸249个,基本重复单元为 GPAGSPGFQGLPGPAGPPGEAGKPGEQGVPGDLGAPGPSGARGERGFPE RGVQPPGPA (SEQ ID No.3),为人胶原蛋白I型肽段,C端氨基酸序列为 GAPGPCCGG (SEQ ID No.2),为增强胶原活性的端序列肽段。C1S4T的氨基酸序列如下: GPAGSPGFQGLPGPAGPPGEAGKPGEQGV PGDLGAPGPSGARGERGFPE RGVQPPGAPGAPGSPGFQGLPGPAGPPGEAGKPGEQGVPGDLGAPGPSG ARG ERGFPERGVQPPGAPGAPGSPGFQGLPGPAGPPGEAGKPGEQGV GDLGAPGPSGARGERGFPERGVQPPG PAGPAGSPGFQGLPGPAGPPGE AGKPGEQGVPGDLGAPGPSGARGERGFPERGVQPPGAPGAPGPCCGG (SEQ ID No.5)。C1S4T的DNA序列如下: GGTCCGGCCGGTAGCCCGGTTTTCAAGGTCTGCCGGTCCCGCTGGT CCTCCGGGTGAGGCTGGTAAACCCGGTGAGCAAGGTGTTCCCGGTGAT CTGGGTGCACCGGTCCGAGTGGTGCA CGTGGTGAGCGTGGCTTTCCG GGTGAGCGTGGCGTTCAAGGTCCCCCGGTCCGGCTGGTCCGGCTGGT AGTCC CGGTTTTCAAGGTCTGCCCGTCCCGCTGGTCTCCGGTGAA GCCGGTAAACCGGCGAGCAAGGTGTTCCGGG TGATTTAGGTGCCCC GGTCCGAGCGGTGCACGTGGTGGAGCGCGGCTTCCCGGTGAACCGG TGTTCAAGGTC CCCCCGTCCGGCTGGTCCCGCTGGTAGTCCGGGTTT CCAAGGTTTACCCGGTCCGGCTGGTCCCCCGGTGAAG CTGGTAAACC GGGTGAACAAGGTGTTCCGGGTGATTTAGGCGCACCCGGTCTAGCGG TGCACGTGGTGAGCGT GGCTTCCCGGTGAACGTGGTGTCAAGGTCC GCCCGTCCCGCTGGTCCGGCTGGTAGCCCGGTTTTCAAGGT

CTGCC GGGTCCGGCTGGTCCTCCGGGCGAAGCTGGTAAGCCGGGTGAGCAAG GTGTGCCGGGTGACTTAGGTGC ACCGGGTCCGAGTGGTGCACGTGGC GAGCGTGGTTTTCCGGGCGAACGTGGTGTCAAGGTCCGCCGGGTCCG GCCGGTGCACCGGGTCCGTGTTGTGGTGGT (SEQ ID No.7)。

[0049] 多肽在本文中可以是人源胶原蛋白C1S5T,为单链结构,包括249个氨基酸,基本重复单元为 GEKSPGADGPAGAPGTPGPQGIAGQRGVVGLPGQRGERGFPGLPGPSG EPGKQGPSGAS (SEQ ID No.4),为人胶原蛋白I型肽段,C端氨基酸序列为GAPGPCCG (SEQ ID No.2),为增强胶原活性的端序列肽段。C1S5T的氨基酸序列如下: GEKSPGADGPAGAPGTPGPQGIAGQRGVVGLP GQRGERGFPGLPGPSG EPGKQGPSGASGEKSPGADGPAGAPGTPGPQGIAGQRGVVGLPGQRGE RGFPLPGP SGEPGKQGPSGASGEKSPGADGPAGAPGTPGPQGIAGQRG VVGLPGQRGERGFPGLPGPSGEPGKQGPSGASGE KGSPGADGPAGAPGTP GPQGIAGQRGVVGLPGQRGERGFPGLPGPSGEPGKQGPSGASGAPGPCC GG (SEQ ID No.6)。C1S5T的DNA序列如下: GGTGAAAAGGCAGCCCGGGTGCCGATGGTCCCGCTGGTGCACCGG TA CACCGGGTCTCAAGGATTGCCGGTCAACGTGGTGTGTGGGTCT GCCGGTCAGCGTGGTGAACCGGTTTTTC CGGGTCTGCCGGTCCGA GTGGTGAACCGGTAACAAGGTCCGAGCGGTGCCAGTGGTAAAAA GGTAGCCCC GGTGCAGACGGTCCCGCTGGTGCCCCGGTACACCGGT CCTCAAGGCATTGCTGGTCAGCGTGGCGTTGTGGGT CTGCCCCGGTCAG CGTGGCGAGCGTGGTTTTCCCGGTTTACCCGGTCCGAGTGGCGAGCCC GGTAAGCAAGGTCC GAGTGGTGCCAGCGGTGAGAAGGGTAGTCCGGG TGCAGACGGTCCCGCTGGTGCCCCGGTACCCCGGGTCCGCA AGGTAT TGCTGGTCAACGTGGTGTGTGGTTTACCGGGTACGCGCGGCAACG TGGTTTCCCGGGTCTGCCCG GTCCGAGTGGCGAGCCGGTAAGCAAG GTCCGAGCGGCGCAAGCGGCGAAAAAGGTAGTCCGGGTGCAGATGGT CCCGCTGGTGCACCGGTACACCGGTCTCAAGGTATCGCTGGTCAG CGCGGTGTTGTTGGTCTGCCGGGTCAA CGCGGTGAACGTGGTTTCCCG GGTCTGCCCGTCCGAGTGGCGAACCGGTAACAAGGTCCGAGCGG TGCCAG CGGTGCACCGGGTCCGTGTTGTGGTGGT (SEQ ID No.8)。

[0050] 在本发明中,重组人源胶原蛋白可以通过本领域中常规的方法进行。例如,可以如下步骤生产:(1)大肠杆菌基因工程菌的构建;(2)大肠杆菌基因工程菌的发酵培养;(3)重组人源胶原蛋白的诱导和表达;以及(4)重组人源胶原蛋白的纯化和任选的酶切。

[0051] 在步骤(1)中,大肠杆菌基因工程菌的构建可以如下进行:(1)利用PCR方法对人源性I型胶原蛋白的基因螺旋区的DNA片段进行密码子优化和拼接重组,最终得到目的基因片段;(2)将得到的目的基因片段插入PET-32a表达载体中得到重组表达质粒;(3)将重组表达质粒转入大肠杆菌感受态细胞 BL21 (DE3)中,筛选得到阳性大肠杆菌基因工程菌。

[0052] 在步骤(2)与(3)中,大肠杆菌基因工程菌的发酵培养和重组人源胶原蛋白的诱导和表达可以如下进行:(1)从LAB平板中挑取优选后的大肠杆菌基因工程菌单菌落,置于10ml的LB培养基中37℃,220rpm培养12-16小时;(2)将菌液按照1:100接种到2×YT培养基中放大培养,37℃培养约3小时,待OD₆₀₀在0.4-0.6时,加入终浓度为0.5mM IPTG进行诱导,16℃继续培养20小时,离心收集菌体。

[0053] 在步骤(4)中,重组人源胶原蛋白多肽的纯化和酶切可以如下进行:(1)用磷酸盐缓冲液(40mM NaH₂PO₃,500mM NaCl,pH 7.8)重悬细菌,超声破碎,离心收集上清液;(2)利用NI-NTA亲和柱结合重组人源胶原蛋白,10mM 咪唑漂洗杂蛋白后,加入Prescission Protease (PPase)蛋白酶4℃,16h柱上酶切,最后获得目的胶原蛋白多肽。

[0054] 大肠杆菌仅仅是例示性的。宿主细胞可以是真核细胞,例如真菌和酵母,原核细胞,例如肠杆菌科细菌。应当理解,本领域技术人员可以通过用其它表达菌株替换上述大

肠杆菌菌株作为宿主细胞。

[0055] 本发明的肽包含以SEQ ID No.5或SEQ ID No.6所示的序列或或以 SEQ ID No.5或SEQ ID No.6所示的序列中取代、缺失、插入和/或添加一个、两个或多个氨基酸的序列，所述肽显示黏附活性。“多个”可以是2、3、4、5、6、7、或8个。

[0056] 氨基酸取代意味着在相同位置，某个氨基酸残基被其他氨基酸残基替代。插入的氨基酸残基可以在任何位置插入，插入的氨基酸残基也可以全部或部分彼此相邻，或插入的氨基酸之间都不彼此相邻。可以从SEQ ID NO:1 的序列中删除1、2或3个氨基酸，只要所述肽显示抑制神经细胞的活性或除皱活性。

[0057] 本领域技术人员已知，本发明所述的肽可以在氨基酸序列之间的一个或多个位置进行翻译后修饰。翻译后修饰的例子可以包括磷酸化作用、乙酰化作用和脱酰氨基作用。

[0058] 本发明还提供SEQ ID No.5或6所示的肽的类似物，只要类似物显示细胞黏附活性。这些类似物与天然的肽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。

[0059] 在本发明中，取代可以是保守氨基酸取代，指与SEQ ID NO:5或6的氨基酸序列相比，有至多3个，更佳地至多2个氨基酸或1个氨基酸被性质相似或相近的氨基酸所替换而形成肽。这些保守性变异肽可以根据表1进行氨基酸替换而产生。

[0060]

最初的残基	代表性的取代	优选的取代
Ala (A)	Val;Leu;Ile	Val
Arg (R)	Lys;Gln;Asn	Lys
Asn (N)	Gln;His;Lys;Arg	Gln
Asp (D)	Glu	Glu
Cys (C)	Ser	Ser
Gln (Q)	Asn	Asn
Glu (E)	Asp	Asp
Gly (G)	Pro;Ala	Ala
His (H)	Asn;Gln;Lys;Arg	Arg
Ile (I)	Leu;Val;Met;Ala;Phe	Leu
Leu (L)	Ile;Val;Met;Ala;Phe	Ile
Lys (K)	Arg;Gln;Asn	Arg
Met (M)	Leu;Phe;Ile	Leu
Phe (F)	Leu;Val;Ile;Ala;Tyr	Leu
Pro (P)	Ala	Ala
Ser (S)	Thr	Thr
Thr (T)	Ser	Ser

Trp (W)	Tyr;Phe	Tyr
Tyr (Y)	Trp;Phe;Thr;Ser	Phe
Val (V)	Ile;Leu;Met;Phe;Ala	Leu

[0061] 序列同一性:参数“序列同一性”描述两个氨基酸序列之间或两个核苷酸序列之间的相关性。

[0062] 就本发明而言,两个氨基酸序列之间的序列同一性程度使用如EMBOSS 软件包(EMBOSS:The European Molecular Biology Open Software Suite,Rice 等,2000,Trends Genet.16:276-277),优选3.0.0版或更高版本的Needle程序中所执行的Needleman-Wunsch算法(Needleman和Wunsch,1970,J.Mol.Biol. 48:443-453)来测定。使用的可选参数为缺口罚分(gap penalty)10,缺口延伸罚分(gap extension penalty)0.5和EBLOSUM62(BLOSUM62的EMBOSS版)取代 矩阵。使用Needle标记为“最高同一性(longest identity)”的输出结果(使用 -nobrief选项获得)作为同一性百分比,并计算如下:

[0063] (同样的残基×100)/(比对长度-比对中缺口的总数)

[0064] 就本发明而言,杂交表示多核苷酸在非常低至非常高的严格条件下与标记的核酸探针杂交,所述核酸探针对应于SEQ ID NO:5或6的多肽编码序列。可使用例如X射线片(X-ray film)检测在这些条件下与核酸探针杂交的分子。

[0065] 对于长度至少100个核苷酸的长探针,将非常低至非常高的严格条件定义为在42℃,在5X SSPE、0.3%SDS、200微克/ml已剪切并且变性的鲑精 DNA中,并且对于非常低和低严格性为25%的甲酰胺、对于中和中-高严格性为35%的甲酰胺、或对于高和非常高严格性为50%的甲酰胺,根据标准的 Southern印迹法进行预杂交和杂交最佳12至24小时。使用2X SSC、0.2%SDS 至少在45℃(非常低严格性),至少在50℃(低严格性),至少在55℃(中严格性),至少在60℃(中-高严格性),至少在65℃(高严格性),至少在70℃(非常高严格性)将载体材料最终洗涤三次,每次15分钟。

[0066] 实施例

[0067] 提供以下实施例来阐述本发明。本领域技术人员应当理解实施例仅仅是 例示性的而非限制性的。本发明仅仅由所附权利要求书的范围限定。

[0068] 实施例1:重组人源胶原蛋白多肽的构建及表达

[0069] C1S4T基因表达载体的构建及表达

[0070] 1.实施例1中使用的人源胶原蛋白C1S4T全长基因序列以SEQ ID No.7 显示。该序列已经针对大肠杆菌的密码子进行了密码子优化。

[0071] 2.C1S4T基因全长747bp,根据优化后的C1S4T密码子基因序列SEQ ID No.7,委托上海华津生物科技有限公司进行基因片段的合成,并将合成后的 C1S4T基因片段通过BamH I(NEB公司货号:R0136L)和Xho I(NEB公司, 货号:R0146L)的酶切位点插入PET32a表达载体。将该构建成功的表达质粒 转化大肠杆菌感受态细胞BL21(DE3)(Merck公司)。具体过程为:1:取1μl的 该质粒于100μl的大肠杆菌感受态细胞BL21(DE3)中,冰上静置30min。2:将该混合物于42℃水浴锅中热激90s,然后迅速置于冰上静置2min。3:向该混合物中加入600μl无抗性的LB,37℃,220rpm条件下培养1h。4:取200μl该 菌液均匀的涂布在含有氨苄青霉素抗性的LB平板上(10g/L蛋白胨,5g/L酵母 提取物,10g/L氯化钠,15g/L琼脂,100μg/ml氨苄青霉素)。5:将平板倒置 培养于37℃温箱中,培养约20h待长出清晰可见的菌落。

[0072] 3.从转化好的LB平板中挑取单克隆菌落于10ml LB(含100 μ g/ml氨苄抗生素)培养基中培养12h-16h后,再按照1:100的比例转接到2 \times YT培养基(16g/L 蛋白胨,10g/L酵母提取物,5g/L氯化钠)中进行扩大培养,37 $^{\circ}$ C,220rpm培养至菌液OD600在0.4-0.6时,加入终浓度为0.5mM IPTG(Sigma公司,货号:I5502-1G)进行诱导表达,诱导条件为18 $^{\circ}$ C、180rpm培养20h。最后离心收集菌体,保存于-20 $^{\circ}$ C或者立即进入下步纯化。

[0073] 4.用磷酸盐缓冲液(pH 7.8)(40mM磷酸二氢钠,500mM氯化钠)约50ml重悬(1L)菌体沉淀,利用高压破菌仪器(新芝生物)进行破菌后,13000rpm离心30min,使可溶性蛋白与包涵体充分分离。

[0074] 5.用5倍柱体积的结合缓冲液(Binding buffer)(40mM NaH₂PO₃,500mM NaCl,pH 7.8)平衡Ni-NTA(Qiagen公司,货号:30210)亲和柱。然后加入蛋白上清于4 $^{\circ}$ C条件下孵育0.5-1h,使目的重组蛋白充分结合到柱材上。再用200ml含有10mM咪唑(Sigma公司)的洗涤缓冲液(washing buffer)(10mM咪唑,40mM NaH₂PO₃,500mM NaCl,pH 7.8)漂洗杂蛋白。最后加入适量具有His标签的Prescission Protease(简称PPase)(Sigma,SAE0045)蛋白酶,于4 $^{\circ}$ C孵育16h后,收集穿流液,即为去除载体蛋白的目的胶原蛋白。所得产物透析过夜,冻干为干粉待用。

[0075] 6.所得C1S4T蛋白利用SDS-PAGE检测纯度。具体过程为:取纯化后的蛋白液40 μ l,加入10 μ l 5 \times 的蛋白上样缓冲液(250mM的Tris-HCl(pH:6.8),10%SDS,0.5%溴酚蓝,50%甘油,5% β -巯基乙醇),置于100 $^{\circ}$ C沸水中煮10min,然后每孔10 μ l加入SDS-PAGE蛋白胶中,电压80V跑2h后,用考马斯亮蓝染色液(0.1%考马斯亮蓝R-250,25%异丙醇,10%冰醋酸)进行蛋白染色20min,再利用蛋白脱色液(10%醋酸,5%乙醇)进行脱色。最后与人天然胶原蛋白相对照测量蛋白活性。

[0076] C1S5T基因表达载体的构建及表达

[0077] 1.人源胶原蛋白C1S5T全长基因序列以SEQ ID No.8显示。该序列已经针对大肠杆菌的密码子进行了密码子优化。

[0078] 2.C1S5T基因全长747bp,根据优化后的C1S5T密码子基因序列,委托上海华津生物科技有限公司进行基因片段的合成,并将合成后的C1S5T基因片段通过BamH I(NEB公司货号:R0136L)和Xho I(NEB公司,货号:R0146L)的酶切位点插入PET32a表达载体。将该构建成功的表达质粒转化大肠杆菌感受态细胞BL21(DE3)(Merck公司)。具体过程为:1:取1 μ l的该质粒于100 μ l的大肠杆菌感受态细胞BL21(DE3)中,冰上静置30min。2:将该混合物于42 $^{\circ}$ C水浴锅中热激90s,然后迅速置于冰上静置2min。3:向该混合物中加入600 μ l无抗性的LB,37 $^{\circ}$ C,220rpm条件下培养1h。4:取200 μ l该菌液均匀的涂布在含有氨苄青霉素抗性的LAB平板上(10g/L蛋白胨,5g/L酵母提取物,10g/L氯化钠,15g/L琼脂,100 μ g/ml氨苄抗生素)。5:将平板倒置培养于37 $^{\circ}$ C温箱中,培养约20h待长出清晰可见的菌落。

[0079] 3.从转化好的LB平板中挑取单克隆菌落于10ml LB(含100 μ g/ml氨苄抗生素)培养基中培养12h-16h后,再按照1:100的比例转接到2 \times YT培养基(16g/L 蛋白胨,10g/L酵母提取物,5g/L氯化钠)中进行扩大培养,37 $^{\circ}$ C,220rpm培养至菌液OD600在0.4-0.6时,加入终浓度为0.5mM IPTG(Sigma公司,货号:I5502-1G)进行诱导表达,诱导条件为18 $^{\circ}$ C、180rpm培养20h。最后离心收集菌体,保存于-20 $^{\circ}$ C或者立即进入下步纯化。

[0080] 4.用磷酸盐缓冲液(pH 7.8)(40mM磷酸二氢钠,500mM氯化钠)约50ml重悬(1L)菌

体沉淀,利用高压破菌仪器(新芝生物)进行破菌后,13000rpm离心30min,使可溶性蛋白与包涵体充分分离。

[0081] 5.用5倍柱体积的结合缓冲液(Binding buffer)(40mM NaH₂PO₃,500mM NaCl,pH 7.8)平衡Ni-NTA亲和柱(Qiagen公司,货号:30210)。然后加入蛋白上清于4℃条件下孵育0.5-1h,使目的重组蛋白充分结合到柱材上。再用200ml含有10mM咪唑(Sigma公司)的洗涤缓冲液(washing buffer)(10mM咪唑,40mM NaH₂PO₃,500mM NaCl,pH 7.8)漂洗杂蛋白。最后加入适量具有His标签的Prescission Protease(简称PPase)蛋白酶(Sigma,SAE0045),于4℃孵育16h后,收集穿流液,即为去除载体蛋白的目的胶原蛋白。所得产物透析过夜,冻干为干粉待用。

[0082] 6.所得C1S5T蛋白利用SDS-PAGE检测纯度。具体过程为:取纯化后的蛋白液40μl,加入10μl 5×的蛋白上样缓冲液(250mM的Tris-HCl(pH:6.8),10%SDS,0.5%溴酚蓝,50%甘油,5%β-巯基乙醇),置于100℃沸水中煮10min,然后每孔10μl加入SDS-PAGE蛋白胶中,电压80V跑2h后,用考马斯亮蓝染色液(0.1%考马斯亮蓝R-250,25%异丙醇,10%冰醋酸)进行蛋白染色20min,再利用蛋白脱色液(10%醋酸,5%乙醇)进行脱色。最后与人天然胶原蛋白相对照测量蛋白活性。

[0083] 结果

[0084] 图2和图3的电泳图表明得到表观分子量为36kDa的C1S4T和C1S5T,分子量分别对应于SEQ ID NO:5和6的氨基酸序列的多肽,表明多肽C1S4T和C1S5T得到正确表达。

[0085] 实施例2C1S4T和C1S5T蛋白的活性检测

[0086] 胶原蛋白的活性检测方法可以参考文献Juming Yao,Satoshi Yanagisawa,Tetsuo Asakura,Design,Expression and Characterization of Collagen-Like Proteins Based on the Cell Adhesive and Crosslinking Sequences Derived from Native Collagens,J Biochem.136,643-649(2004)。具体实施方法如下:

[0087] 1、利用紫外吸收法检测待测蛋白样品的浓度,包括对照人胶原蛋白(Sigma,C7774)、C1S4T和C1S5T蛋白样品。具体为分别测定样品在215nm和225nm下的紫外光吸收,利用经验公式 $C(\mu\text{g}/\text{mL}) = 144X(A_{215} - A_{225})$ 计算蛋白质浓度,注意需在 $A_{215} < 1.5$ 的情况下检测。该方法的原理是测定肽键在远紫外光下的特征吸收,不受生色团含量的影响,干扰物质少,操作简便,适合检测考马斯亮蓝不显色的人胶原蛋白及其类似物。(参考文献为Walker JM.The Protein Protocols Handbook,second edition.Humana Press.43-45.)。检测完蛋白浓度后,用PBS将所有待测蛋白浓度调整到0.5mg/ml。

[0088] 2、向96孔板中加入100μl各种蛋白溶液和空白PBS溶液对照,室温静置60min。

[0089] 3、每孔中加入10⁵个培养状态良好的3T3细胞(来自清华大学童佩老师),37℃孵育60min。

[0090] 4、每孔用PBS清洗4次。

[0091] 5、用LDH检测试剂盒(Roche,04744926001)检测OD492nm的吸光度。根据空白对照的数值,可以计算出细胞的贴壁率。计算公式如下:细胞贴壁率=(测试孔-空白孔)×100%/(阳性孔-空白孔)。细胞的贴壁率即可以反应胶原蛋白的活性。蛋白的活性越高,越能在短时间给细胞提供优质的外环境,帮助细胞贴壁。

[0092] 结果参见图4和图5。

[0093] 图4和图5的结果表明,两种人重组胶原蛋白(即C1S4T和C1S5T)与商品化的人胶原蛋白相比,皆具有相当或更好的黏附活性。本发明的重组I型胶原蛋白C1S4T和C1S5T实现了I型胶原蛋白的细胞黏附活性。

[0001] 序列表
 [0002] <110> 山西锦波生物医药股份有限公司
 [0003] <120> 多肽、其生产方法和用途
 [0004] <130> C18P2983
 [0005] <160> 8
 [0006] <170> PatentIn version 3.5
 [0007] <210> 1
 [0008] <211> 1464
 [0009] <212> PRT
 [0010] <213> 人
 [0011] <400> 1
 [0012] Met Phe Ser Phe Val Asp Leu Arg Leu Leu Leu Leu Leu Ala Ala Thr
 [0013] 1 5 10 15
 [0014] Ala Leu Leu Thr His Gly Gln Glu Glu Gly Gln Val Glu Gly Gln Asp
 [0015] 20 25 30
 [0016] Glu Asp Ile Pro Pro Ile Thr Cys Val Gln Asn Gly Leu Arg Tyr His
 [0017] 35 40 45
 [0018] Asp Arg Asp Val Trp Lys Pro Glu Pro Cys Arg Ile Cys Val Cys Asp
 [0019] 50 55 60
 [0020] Asn Gly Lys Val Leu Cys Asp Asp Val Ile Cys Asp Glu Thr Lys Asn
 [0021] 65 70 75 80
 [0022] Cys Pro Gly Ala Glu Val Pro Glu Gly Glu Cys Cys Pro Val Cys Pro
 [0023] 85 90 95
 [0024] Asp Gly Ser Glu Ser Pro Thr Asp Gln Glu Thr Thr Gly Val Glu Gly
 [0025] 100 105 110
 [0026] Pro Lys Gly Asp Thr Gly Pro Arg Gly Pro Arg Gly Pro Ala Gly Pro
 [0027] 115 120 125
 [0028] Pro Gly Arg Asp Gly Ile Pro Gly Gln Pro Gly Leu Pro Gly Pro Pro
 [0029] 130 135 140
 [0030] Gly Pro Pro Gly Pro Pro Gly Pro Pro Gly Leu Gly Gly Asn Phe Ala
 [0031] 145 150 155 160
 [0032] Pro Gln Leu Ser Tyr Gly Tyr Asp Glu Lys Ser Thr Gly Gly Ile Ser
 [0033] 165 170 175
 [0034] Val Pro Gly Pro Met Gly Pro Ser Gly Pro Arg Gly Leu Pro Gly Pro
 [0035] 180 185 190
 [0036] Pro Gly Ala Pro Gly Pro Gln Gly Phe Gln Gly Pro Pro Gly Glu Pro
 [0037] 195 200 205
 [0038] Gly Glu Pro Gly Ala Ser Gly Pro Met Gly Pro Arg Gly Pro Pro Gly

[0039]	210	215	220
[0040]	Pro Pro Gly Lys Asn Gly Asp Asp Gly Glu Ala Gly Lys Pro Gly Arg		
[0041]	225	230	235 240
[0042]	Pro Gly Glu Arg Gly Pro Pro Gly Pro Gln Gly Ala Arg Gly Leu Pro		
[0043]	245	250	255
[0044]	Gly Thr Ala Gly Leu Pro Gly Met Lys Gly His Arg Gly Phe Ser Gly		
[0045]	260	265	270
[0046]	Leu Asp Gly Ala Lys Gly Asp Ala Gly Pro Ala Gly Pro Lys Gly Glu		
[0047]	275	280	285
[0048]	Pro Gly Ser Pro Gly Glu Asn Gly Ala Pro Gly Gln Met Gly Pro Arg		
[0049]	290	295	300
[0050]	Gly Leu Pro Gly Glu Arg Gly Arg Pro Gly Ala Pro Gly Pro Ala Gly		
[0051]	305	310	315 320
[0052]	Ala Arg Gly Asn Asp Gly Ala Thr Gly Ala Ala Gly Pro Pro Gly Pro		
[0053]	325	330	335
[0054]	Thr Gly Pro Ala Gly Pro Pro Gly Phe Pro Gly Ala Val Gly Ala Lys		
[0055]	340	345	350
[0056]	Gly Glu Ala Gly Pro Gln Gly Pro Arg Gly Ser Glu Gly Pro Gln Gly		
[0057]	355	360	365
[0058]	Val Arg Gly Glu Pro Gly Pro Pro Gly Pro Ala Gly Ala Ala Gly Pro		
[0059]	370	375	380
[0060]	Ala Gly Asn Pro Gly Ala Asp Gly Gln Pro Gly Ala Lys Gly Ala Asn		
[0061]	385	390	395 400
[0062]	Gly Ala Pro Gly Ile Ala Gly Ala Pro Gly Phe Pro Gly Ala Arg Gly		
[0063]	405	410	415
[0064]	Pro Ser Gly Pro Gln Gly Pro Gly Gly Pro Pro Gly Pro Lys Gly Asn		
[0065]	420	425	430
[0066]	Ser Gly Glu Pro Gly Ala Pro Gly Ser Lys Gly Asp Thr Gly Ala Lys		
[0067]	435	440	445
[0068]	Gly Glu Pro Gly Pro Val Gly Val Gln Gly Pro Pro Gly Pro Ala Gly		
[0069]	450	455	460
[0070]	Glu Glu Gly Lys Arg Gly Ala Arg Gly Glu Pro Gly Pro Thr Gly Leu		
[0071]	465	470	475 480
[0072]	Pro Gly Pro Pro Gly Glu Arg Gly Gly Pro Gly Ser Arg Gly Phe Pro		
[0073]	485	490	495
[0074]	Gly Ala Asp Gly Val Ala Gly Pro Lys Gly Pro Ala Gly Glu Arg Gly		
[0075]	500	505	510
[0076]	Ser Pro Gly Pro Ala Gly Pro Lys Gly Ser Pro Gly Glu Ala Gly Arg		
[0077]	515	520	525

[0078]	Pro Gly Glu Ala Gly Leu Pro Gly Ala Lys Gly Leu Thr Gly Ser Pro
[0079]	530 535 540
[0080]	Gly Ser Pro Gly Pro Asp Gly Lys Thr Gly Pro Pro Gly Pro Ala Gly
[0081]	545 550 555 560
[0082]	Gln Asp Gly Arg Pro Gly Pro Pro Gly Pro Pro Gly Ala Arg Gly Gln
[0083]	565 570 575
[0084]	Ala Gly Val Met Gly Phe Pro Gly Pro Lys Gly Ala Ala Gly Glu Pro
[0085]	580 585 590
[0086]	Gly Lys Ala Gly Glu Arg Gly Val Pro Gly Pro Pro Gly Ala Val Gly
[0087]	595 600 605
[0088]	Pro Ala Gly Lys Asp Gly Glu Ala Gly Ala Gln Gly Pro Pro Gly Pro
[0089]	610 615 620
[0090]	Ala Gly Pro Ala Gly Glu Arg Gly Glu Gln Gly Pro Ala Gly Ser Pro
[0091]	625 630 635 640
[0092]	Gly Phe Gln Gly Leu Pro Gly Pro Ala Gly Pro Pro Gly Glu Ala Gly
[0093]	645 650 655
[0094]	Lys Pro Gly Glu Gln Gly Val Pro Gly Asp Leu Gly Ala Pro Gly Pro
[0095]	660 665 670
[0096]	Ser Gly Ala Arg Gly Glu Arg Gly Phe Pro Gly Glu Arg Gly Val Gln
[0097]	675 680 685
[0098]	Gly Pro Pro Gly Pro Ala Gly Pro Arg Gly Ala Asn Gly Ala Pro Gly
[0099]	690 695 700
[0100]	Asn Asp Gly Ala Lys Gly Asp Ala Gly Ala Pro Gly Ala Pro Gly Ser
[0101]	705 710 715 720
[0102]	Gln Gly Ala Pro Gly Leu Gln Gly Met Pro Gly Glu Arg Gly Ala Ala
[0103]	725 730 735
[0104]	Gly Leu Pro Gly Pro Lys Gly Asp Arg Gly Asp Ala Gly Pro Lys Gly
[0105]	740 745 750
[0106]	Ala Asp Gly Ser Pro Gly Lys Asp Gly Val Arg Gly Leu Thr Gly Pro
[0107]	755 760 765
[0108]	Ile Gly Pro Pro Gly Pro Ala Gly Ala Pro Gly Asp Lys Gly Glu Ser
[0109]	770 775 780
[0110]	Gly Pro Ser Gly Pro Ala Gly Pro Thr Gly Ala Arg Gly Ala Pro Gly
[0111]	785 790 795 800
[0112]	Asp Arg Gly Glu Pro Gly Pro Pro Gly Pro Ala Gly Phe Ala Gly Pro
[0113]	805 810 815
[0114]	Pro Gly Ala Asp Gly Gln Pro Gly Ala Lys Gly Glu Pro Gly Asp Ala
[0115]	820 825 830
[0116]	Gly Ala Lys Gly Asp Ala Gly Pro Pro Gly Pro Ala Gly Pro Ala Gly

[0117]	835	840	845
[0118]	Pro Pro Gly Pro Ile Gly Asn Val Gly Ala Pro Gly Ala Lys Gly Ala		
[0119]	850	855	860
[0120]	Arg Gly Ser Ala Gly Pro Pro Gly Ala Thr Gly Phe Pro Gly Ala Ala		
[0121]	865	870	875
[0122]	Gly Arg Val Gly Pro Pro Gly Pro Ser Gly Asn Ala Gly Pro Pro Gly		
[0123]	885	890	895
[0124]	Pro Pro Gly Pro Ala Gly Lys Glu Gly Gly Lys Gly Pro Arg Gly Glu		
[0125]	900	905	910
[0126]	Thr Gly Pro Ala Gly Arg Pro Gly Glu Val Gly Pro Pro Gly Pro Pro		
[0127]	915	920	925
[0128]	Gly Pro Ala Gly Glu Lys Gly Ser Pro Gly Ala Asp Gly Pro Ala Gly		
[0129]	930	935	940
[0130]	Ala Pro Gly Thr Pro Gly Pro Gln Gly Ile Ala Gly Gln Arg Gly Val		
[0131]	945	950	955
[0132]	Val Gly Leu Pro Gly Gln Arg Gly Glu Arg Gly Phe Pro Gly Leu Pro		
[0133]	965	970	975
[0134]	Gly Pro Ser Gly Glu Pro Gly Lys Gln Gly Pro Ser Gly Ala Ser Gly		
[0135]	980	985	990
[0136]	Glu Arg Gly Pro Pro Gly Pro Met Gly Pro Pro Gly Leu Ala Gly Pro		
[0137]	995	1000	1005
[0138]	Pro Gly Glu Ser Gly Arg Glu Gly Ala Pro Gly Ala Glu Gly Ser		
[0139]	1010	1015	1020
[0140]	Pro Gly Arg Asp Gly Ser Pro Gly Ala Lys Gly Asp Arg Gly Glu		
[0141]	1025	1030	1035
[0142]	Thr Gly Pro Ala Gly Pro Pro Gly Ala Pro Gly Ala Pro Gly Ala		
[0143]	1040	1045	1050
[0144]	Pro Gly Pro Val Gly Pro Ala Gly Lys Ser Gly Asp Arg Gly Glu		
[0145]	1055	1060	1065
[0146]	Thr Gly Pro Ala Gly Pro Ala Gly Pro Val Gly Pro Val Gly Ala		
[0147]	1070	1075	1080
[0148]	Arg Gly Pro Ala Gly Pro Gln Gly Pro Arg Gly Asp Lys Gly Glu		
[0149]	1085	1090	1095
[0150]	Thr Gly Glu Gln Gly Asp Arg Gly Ile Lys Gly His Arg Gly Phe		
[0151]	1100	1105	1110
[0152]	Ser Gly Leu Gln Gly Pro Pro Gly Pro Pro Gly Ser Pro Gly Glu		
[0153]	1115	1120	1125
[0154]	Gln Gly Pro Ser Gly Ala Ser Gly Pro Ala Gly Pro Arg Gly Pro		
[0155]	1130	1135	1140

[0156]	Pro Gly Ser Ala Gly Ala Pro Gly Lys Asp Gly Leu Asn Gly Leu
[0157]	1145 1150 1155
[0158]	Pro Gly Pro Ile Gly Pro Pro Gly Pro Arg Gly Arg Thr Gly Asp
[0159]	1160 1165 1170
[0160]	Ala Gly Pro Val Gly Pro Pro Gly Pro Pro Gly Pro Pro Gly Pro
[0161]	1175 1180 1185
[0162]	Pro Gly Pro Pro Ser Ala Gly Phe Asp Phe Ser Phe Leu Pro Gln
[0163]	1190 1195 1200
[0164]	Pro Pro Gln Glu Lys Ala His Asp Gly Gly Arg Tyr Tyr Arg Ala
[0165]	1205 1210 1215
[0166]	Asp Asp Ala Asn Val Val Arg Asp Arg Asp Leu Glu Val Asp Thr
[0167]	1220 1225 1230
[0168]	Thr Leu Lys Ser Leu Ser Gln Gln Ile Glu Asn Ile Arg Ser Pro
[0169]	1235 1240 1245
[0170]	Glu Gly Ser Arg Lys Asn Pro Ala Arg Thr Cys Arg Asp Leu Lys
[0171]	1250 1255 1260
[0172]	Met Cys His Ser Asp Trp Lys Ser Gly Glu Tyr Trp Ile Asp Pro
[0173]	1265 1270 1275
[0174]	Asn Gln Gly Cys Asn Leu Asp Ala Ile Lys Val Phe Cys Asn Met
[0175]	1280 1285 1290
[0176]	Glu Thr Gly Glu Thr Cys Val Tyr Pro Thr Gln Pro Ser Val Ala
[0177]	1295 1300 1305
[0178]	Gln Lys Asn Trp Tyr Ile Ser Lys Asn Pro Lys Asp Lys Arg His
[0179]	1310 1315 1320
[0180]	Val Trp Phe Gly Glu Ser Met Thr Asp Gly Phe Gln Phe Glu Tyr
[0181]	1325 1330 1335
[0182]	Gly Gly Gln Gly Ser Asp Pro Ala Asp Val Ala Ile Gln Leu Thr
[0183]	1340 1345 1350
[0184]	Phe Leu Arg Leu Met Ser Thr Glu Ala Ser Gln Asn Ile Thr Tyr
[0185]	1355 1360 1365
[0186]	His Cys Lys Asn Ser Val Ala Tyr Met Asp Gln Gln Thr Gly Asn
[0187]	1370 1375 1380
[0188]	Leu Lys Lys Ala Leu Leu Leu Gln Gly Ser Asn Glu Ile Glu Ile
[0189]	1385 1390 1395
[0190]	Arg Ala Glu Gly Asn Ser Arg Phe Thr Tyr Ser Val Thr Val Asp
[0191]	1400 1405 1410
[0192]	Gly Cys Thr Ser His Thr Gly Ala Trp Gly Lys Thr Val Ile Glu
[0193]	1415 1420 1425
[0194]	Tyr Lys Thr Thr Lys Thr Ser Arg Leu Pro Ile Ile Asp Val Ala

[0195]	1430	1435	1440
[0196]	Pro Leu Asp Val Gly Ala Pro Asp Gln Glu Phe Gly Phe Asp Val		
[0197]	1445	1450	1455
[0198]	Gly Pro Val Cys Phe Leu		
[0199]	1460		
[0200]	<210>	2	
[0201]	<211>	9	
[0202]	<212>	PRT	
[0203]	<213>	人工序列	
[0204]	<220>		
[0205]	<223>	多肽C端序列	
[0206]	<400>	2	
[0207]	Gly Ala Pro Gly Pro Cys Cys Gly Gly		
[0208]	1	5	
[0209]	<210>	3	
[0210]	<211>	60	
[0211]	<212>	PRT	
[0212]	<213>	人工序列	
[0213]	<220>		
[0214]	<223>	N端重复单元1	
[0215]	<400>	3	
[0216]	Gly Pro Ala Gly Ser Pro Gly Phe Gln Gly Leu Pro Gly Pro Ala Gly		
[0217]	1	5	10 15
[0218]	Pro Pro Gly Glu Ala Gly Lys Pro Gly Glu Gln Gly Val Pro Gly Asp		
[0219]		20	25 30
[0220]	Leu Gly Ala Pro Gly Pro Ser Gly Ala Arg Gly Glu Arg Gly Phe Pro		
[0221]		35	40 45
[0222]	Gly Glu Arg Gly Val Gln Gly Pro Pro Gly Pro Ala		
[0223]		50	55 60
[0224]	<210>	4	
[0225]	<211>	60	
[0226]	<212>	PRT	
[0227]	<213>	人工序列	
[0228]	<220>		
[0229]	<223>	N端重复单元2	
[0230]	<400>	4	
[0231]	Gly Glu Lys Gly Ser Pro Gly Ala Asp Gly Pro Ala Gly Ala Pro Gly		
[0232]	1	5	10 15
[0233]	Thr Pro Gly Pro Gln Gly Ile Ala Gly Gln Arg Gly Val Val Gly Leu		

[0234]	20	25	30
[0235]	Pro Gly Gln Arg Gly Glu Arg Gly Phe Pro Gly Leu Pro Gly Pro Ser		
[0236]	35	40	45
[0237]	Gly Glu Pro Gly Lys Gln Gly Pro Ser Gly Ala Ser		
[0238]	50	55	60
[0239]	<210> 5		
[0240]	<211> 249		
[0241]	<212> PRT		
[0242]	<213> 人工序列		
[0243]	<220>		
[0244]	<223> C1S4T的氨基酸序列		
[0245]	<400> 5		
[0246]	Gly Pro Ala Gly Ser Pro Gly Phe Gln Gly Leu Pro Gly Pro Ala Gly		
[0247]	1	5	10
[0248]	Pro Pro Gly Glu Ala Gly Lys Pro Gly Glu Gln Gly Val Pro Gly Asp		
[0249]	20	25	30
[0250]	Leu Gly Ala Pro Gly Pro Ser Gly Ala Arg Gly Glu Arg Gly Phe Pro		
[0251]	35	40	45
[0252]	Gly Glu Arg Gly Val Gln Gly Pro Pro Gly Pro Ala Gly Pro Ala Gly		
[0253]	50	55	60
[0254]	Ser Pro Gly Phe Gln Gly Leu Pro Gly Pro Ala Gly Pro Pro Gly Glu		
[0255]	65	70	75
[0256]	Ala Gly Lys Pro Gly Glu Gln Gly Val Pro Gly Asp Leu Gly Ala Pro		
[0257]	85	90	95
[0258]	Gly Pro Ser Gly Ala Arg Gly Glu Arg Gly Phe Pro Gly Glu Arg Gly		
[0259]	100	105	110
[0260]	Val Gln Gly Pro Pro Gly Pro Ala Gly Pro Ala Gly Ser Pro Gly Phe		
[0261]	115	120	125
[0262]	Gln Gly Leu Pro Gly Pro Ala Gly Pro Pro Gly Glu Ala Gly Lys Pro		
[0263]	130	135	140
[0264]	Gly Glu Gln Gly Val Pro Gly Asp Leu Gly Ala Pro Gly Pro Ser Gly		
[0265]	145	150	155
[0266]	Ala Arg Gly Glu Arg Gly Phe Pro Gly Glu Arg Gly Val Gln Gly Pro		
[0267]	165	170	175
[0268]	Pro Gly Pro Ala Gly Pro Ala Gly Ser Pro Gly Phe Gln Gly Leu Pro		
[0269]	180	185	190
[0270]	Gly Pro Ala Gly Pro Pro Gly Glu Ala Gly Lys Pro Gly Glu Gln Gly		
[0271]	195	200	205
[0272]	Val Pro Gly Asp Leu Gly Ala Pro Gly Pro Ser Gly Ala Arg Gly Glu		

[0273]	210	215	220
[0274]	Arg Gly Phe Pro Gly Glu Arg Gly Val Gln Gly Pro Pro Gly Pro Ala		
[0275]	225	230	235
[0276]	Gly Ala Pro Gly Pro Cys Cys Gly Gly		
[0277]	245		
[0278]	<210> 6		
[0279]	<211> 249		
[0280]	<212> PRT		
[0281]	<213> 人工序列		
[0282]	<220>		
[0283]	<223> C1S5T的氨基酸序列		
[0284]	<400> 6		
[0285]	Gly Glu Lys Gly Ser Pro Gly Ala Asp Gly Pro Ala Gly Ala Pro Gly		
[0286]	1	5	10
[0287]	Thr Pro Gly Pro Gln Gly Ile Ala Gly Gln Arg Gly Val Val Gly Leu		
[0288]	20	25	30
[0289]	Pro Gly Gln Arg Gly Glu Arg Gly Phe Pro Gly Leu Pro Gly Pro Ser		
[0290]	35	40	45
[0291]	Gly Glu Pro Gly Lys Gln Gly Pro Ser Gly Ala Ser Gly Glu Lys Gly		
[0292]	50	55	60
[0293]	Ser Pro Gly Ala Asp Gly Pro Ala Gly Ala Pro Gly Thr Pro Gly Pro		
[0294]	65	70	75
[0295]	Gln Gly Ile Ala Gly Gln Arg Gly Val Val Gly Leu Pro Gly Gln Arg		
[0296]	85	90	95
[0297]	Gly Glu Arg Gly Phe Pro Gly Leu Pro Gly Pro Ser Gly Glu Pro Gly		
[0298]	100	105	110
[0299]	Lys Gln Gly Pro Ser Gly Ala Ser Gly Glu Lys Gly Ser Pro Gly Ala		
[0300]	115	120	125
[0301]	Asp Gly Pro Ala Gly Ala Pro Gly Thr Pro Gly Pro Gln Gly Ile Ala		
[0302]	130	135	140
[0303]	Gly Gln Arg Gly Val Val Gly Leu Pro Gly Gln Arg Gly Glu Arg Gly		
[0304]	145	150	155
[0305]	Phe Pro Gly Leu Pro Gly Pro Ser Gly Glu Pro Gly Lys Gln Gly Pro		
[0306]	165	170	175
[0307]	Ser Gly Ala Ser Gly Glu Lys Gly Ser Pro Gly Ala Asp Gly Pro Ala		
[0308]	180	185	190
[0309]	Gly Ala Pro Gly Thr Pro Gly Pro Gln Gly Ile Ala Gly Gln Arg Gly		
[0310]	195	200	205
[0311]	Val Val Gly Leu Pro Gly Gln Arg Gly Glu Arg Gly Phe Pro Gly Leu		

[0312]	210	215	220
[0313]	Pro Gly Pro Ser Gly Glu Pro Gly Lys Gln Gly Pro Ser Gly Ala Ser		
[0314]	225	230	235 240
[0315]	Gly Ala Pro Gly Pro Cys Cys Gly Gly		
[0316]	245		
[0317]	<210> 7		
[0318]	<211> 747		
[0319]	<212> DNA		
[0320]	<213> 人工序列		
[0321]	<220>		
[0322]	<223> C1S4T的DNA序列		
[0323]	<400> 7		
[0324]	ggtccggccg gtagcccggg ttttcaaggt ctgccgggtc ccgctggtcc tccgggtgag	60	
[0325]	gctggtaaac ccggtgagca aggtgttccc ggtgatctgg gtgcaccggg tccgagtggg	120	
[0326]	gcacgtggtg agcgtggctt tccgggtgag cgtggcggtc aaggtcccc cgggtccggt	180	
[0327]	ggtccggctg gtagtcccgg tttccaaggt ctgccgggtc ccgctggtcc tccgggtgaa	240	
[0328]	gccggtaaac cgggcgagca aggtgttccg ggtgatttag gtgccccggg tccgagcggg	300	
[0329]	gcacgtggtg agcgcggctt cccgggtgaa cgcggtgttc aaggtcccc cgggtccggt	360	
[0330]	ggtcccgtg gtagtccggg tttccaaggt ttaccgggtc cggctggtcc ccccggtgaa	420	
[0331]	gctggtaaac cgggtgaaca aggtgttccg ggtgatttag gcgcaccggg tccgagcggg	480	
[0332]	gcacgtggtg agcgtggctt cccgggtgaa cgtggtgttc aaggtccgcc cgggtccggt	540	
[0333]	ggtccggctg gtagccccgg tttccaaggt ctgccgggtc cggctggtcc tccgggcgaa	600	
[0334]	gctggtaagc cgggtgagca aggtgtgccg ggtgacttag gtgcaccggg tccgagtggg	660	
[0335]	gcacgtggcg agcgtggttt tccgggcgaa cgtggtgttc aaggtccgcc ggggtccggc	720	
[0336]	ggtgcaccgg gtccgtgttg tggtggt	747	
[0337]	<210> 8		
[0338]	<211> 747		
[0339]	<212> DNA		
[0340]	<213> 人工序列		
[0341]	<220>		
[0342]	<223> C1S5T的DNA序列		
[0343]	<400> 8		
[0344]	ggtgaaaaag gcagcccggg tgccgatggt cccgctggtg caccgggtac accgggtcct	60	
[0345]	caaggtattg ccggtcaacg tggtgttgtg ggtctgccgg gtcagcgtgg tgaacgcggg	120	
[0346]	tttccgggtc tgccgggtcc gtagtggtaaa cccggtaaac aaggtccgag cgggtccagt	180	
[0347]	ggtgaaaaag gtagcccggg tgcagacggt cccgctggtg cccccgtac accgggtcct	240	
[0348]	caaggcattg ctggtcagcg tggcgttgtg ggtctgcccg gtcagcgtgg cgagcgtggg	300	
[0349]	tttcccgttt taccgggtcc gtagtggcgag cccggtaagc aaggtccgag tggtgccagc	360	
[0350]	ggtgagaagg gtagtccggg tgcagacggt cccgctggtg cccccgtac cccgggtccg	420	

[0351] caaggtattg ctggtcaacg tgggtgttggt ggtttaccgg gtcagcgagg cgaacgtggt 480
[0352] ttcccgggtc tgcccgggtcc gagtggcgag ccgggtaagc aaggtccgag cggcgcaagc 540
[0353] ggcgaaaaag gtagtccggg tgcagatggt cccgctggtg caccgggtac accgggtcct 600
[0354] caaggtatcg ctggtcagcg cgggtgttggt ggtctgccgg gtcaacggcg tgaacgtggt 660
[0355] ttcccgggtc tgcccgggtcc gagtggcgaa ccgggtaaac aaggtccgag cggtgccagc 720
[0356] ggtgcaccgg gtccgtgttg tgggtggt 747

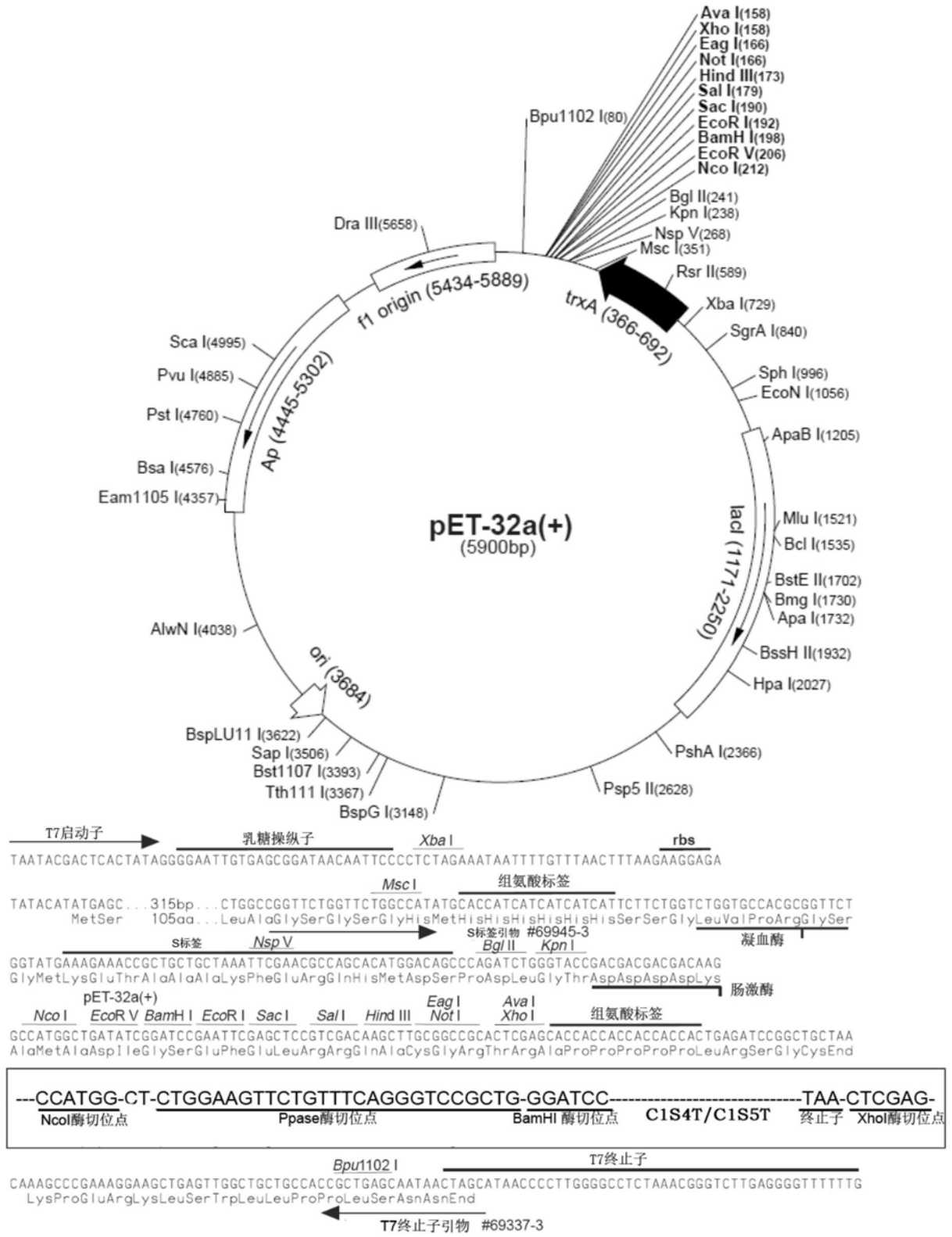


图1

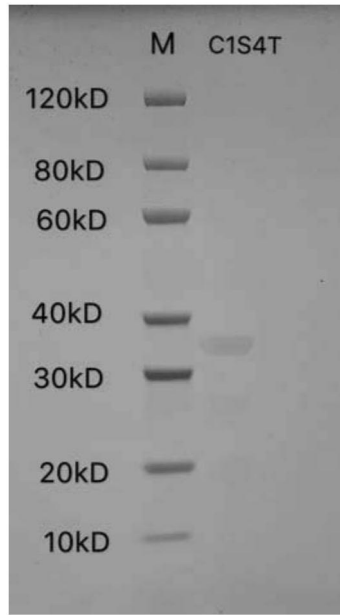


图2

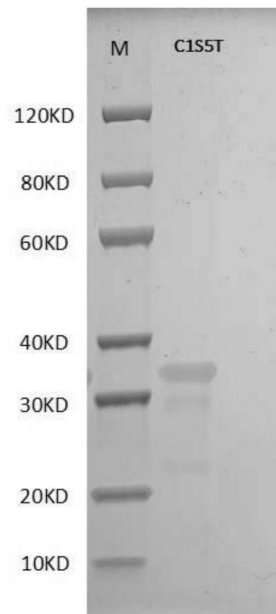


图3

细胞黏附活性检测

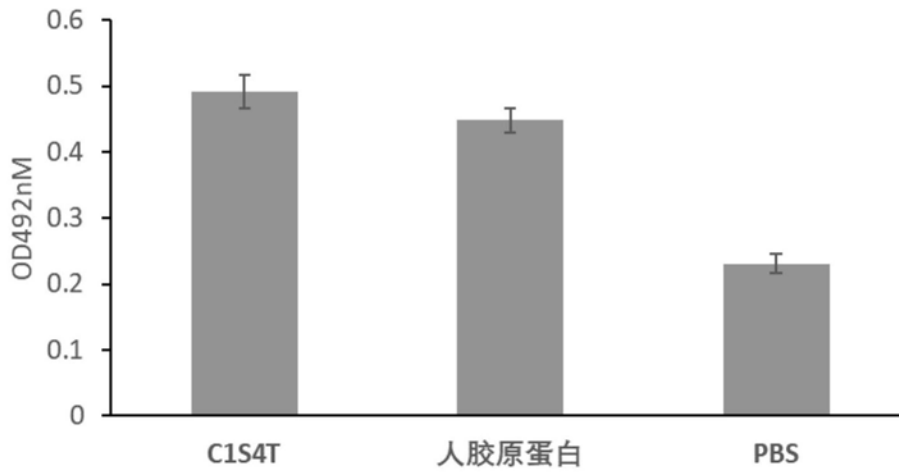


图4

细胞黏附活性检测

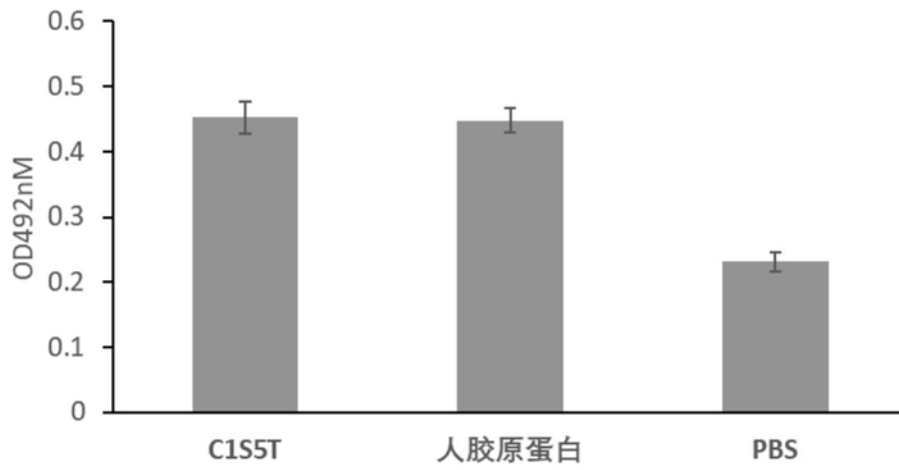


图5