(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 110257347 B
(45)授权公告日 2021.06.22

(21)申请号 201910557170.2
(22)申请日 2019.06.25
(65)同一申请的已公布的文献号
申请公布号 CN 110257347 A
(43)申请公布日 2019.09.20
(73)专利权人 成都英普博集生物科技有限公司
地址 610000 四川省成都市高新区天府大道中段1388号1栋5层522号
(72)发明人 刘懿
(74)专利代理机构 成都顶峰专利事务所(普通合伙) 51224
代理人 王霞
(51)Int.Cl.
C12N 9/02(2006.01)
C12N 15/70(2006.01)
C07K 19/00(2006.01)
审查员 温婧

权利要求书1页 说明书17页
序列表4页 附图3页

(54)发明名称

硫氧还蛋白突变体、其制备方法及其在重组融合蛋白生产中的应用

(57)摘要

本发明公开了一种硫氧还蛋白突变体、其制备方法及其在重组融合蛋白生产中的应用，属于生物技术领域，该硫氧还蛋白突变体是具有高温可溶性的非天然蛋白；本发明的硫氧还蛋白突变体在长时间高温条件下仍旧能够保持可溶性；硫氧还蛋白突变体同时具有热稳定性和在细菌宿主菌体内的可溶表达性，即表达产生的重组融合蛋白能够快速释放，能够快速提取、提取需要的试剂和实验设备价格低廉且能够用于工业上的大规模生产，高温环境下的可溶稳定性好和不易变性。

CN 110257347 B

1.一种硫氧还蛋白突变体,其特征在于,该硫氧还蛋白突变体是具有高温可溶性的非天然蛋白;

该硫氧还蛋白突变体是在硫氧还蛋白模型蛋白的氨基酸序列的基础上人工突变得到;

所述硫氧还蛋白模型蛋白的氨基酸序列如SEQ ID NO.1所示,该硫氧还蛋白突变体是由组1中的氨基酸位置被取代而得到,该硫氧还蛋白突变体的突变位置为:

T15I+D23T+D43E+R44Q+K51 R+S71 M+T74I+D85E+Q91 K+K93P。

2.一种如权利要求1所述硫氧还蛋白突变体在重组融合蛋白生产中的应用,其特征在于,所述重组融合蛋白的结构有两种,分别命名为结构体1和结构体2,其中:结构体1的结构为:硫氧还蛋白突变体——连接肽——切割位点——目标异源多肽或蛋白质,结构体2的结构为:硫氧还蛋白突变体——切割位点——目标异源多肽或蛋白质;

结构体1的制备方法为:目标异源多肽或蛋白质与切割位点结合后得到中间体1,中间体1与连接肽结合后得到中间体2,中间体2与硫氧还蛋白突变体的氨基末端或羧基末端结合得到所述结构体1;

结构体2的制备方法为:目标异源多肽或蛋白质与切割位点结合后得到中间体1,中间体1与硫氧还蛋白突变体的氨基末端或羧基末端结合得到所述结构体2;所述目标异源多肽或蛋白质为由数量不大于60的氨基酸组成的聚合物。

硫氧还蛋白突变体、其制备方法及其在重组融合蛋白生产中的应用

技术领域

[0001] 本发明属于生物技术领域，具体涉及硫氧还蛋白突变体，这些突变导致硫氧还蛋白突变体热稳定性、可溶性和净电荷特性的改变，硫氧还蛋白突变体的制备方法以及利用该硫氧还蛋白突变体在重组融合蛋白生产中的应用。

背景技术

[0002] 许多有着生物活性的多肽和蛋白质在现代农业、生物化工、医药健康、美容护理和食品保健等领域具有潜在的用途和商业价值，但是它们的应用往往受到工业化生产条件的限制，例如在热，有机溶剂和各种化学品操作条件下易变性。然而，在这些条件下稳定的蛋白质已广泛应用于工业生产。

[0003] 蛋白质的三维结构对其功能至关重要。事实上，蛋白质功能和结构的多样性完全取决于不同氨基酸序列所折叠形成的不同三维构象。关于蛋白质三维结构与其生物学和结构性质之间的关系研究是蛋白质化学发展的重要方面。

[0004] 折叠蛋白质的稳定性决定于不同稳定蛋白质结构的相互作用，如疏水和亲水相互作用，氢键，二级结构等之间的平衡，以及由去折叠自由能变化引起的不稳定趋势。因此，通过改变影响蛋白质稳定性相互作用的氨基酸，可以增加或减少蛋白质的稳定性。然而，构象和其他变化可能使氨基酸替代的效果难以预测，实际这也是蛋白质工程研究领域中的重要挑战。

[0005] 许多肽和蛋白质可以通过重组方法在多种表达系统中产生，例如各种细菌，真菌，哺乳动物、植物或昆虫细胞。然而，当细菌被用作异源基因表达宿主时，通常会出现以下一些困难与问题。首先，编码多肽的异源基因通常在细菌中难以成功表达。由于它们氨基酸序列长度太短，大多数多肽异源表达时难以形成稳定和可溶的构象，从而受到宿主细胞中蛋白酶和肽酶的降解。那些在大肠杆菌或其他细菌宿主中能直接表达的多肽通常以不溶性或包涵体形式存在，这种情况使得它们无法直接用于生物活性筛选，同时对这些多肽的提取与纯化带来了困难。其次，即便产生的多肽不以包涵体形式存在，由于它们具有高度的构象自由度，也可能具有大量潜在的不同结构，而这些结构的生物活性会有高低差别。因此，多肽在生产中即使一级结构一致，但活性大小也会因其高级结构的不同而受影响。这也是短肽异源表达时遇到的另一个困难。

[0006] 此外，蛋白质在细菌细胞中异源表达时，也经常会以包涵体形式产生。这些包涵体通常需要通过重溶与复性的过程才能恢复蛋白的生物活性，但复性的过程影响因素复杂，因此，复性的效果具有不确定性。而且，复性的方法在技术上通常具有难度，使用的试剂昂贵，对重组蛋白的生产来说成本过高。

[0007] 为了解决上述问题，本领域已经使用某些肽或蛋白作为目标肽或蛋白质的融合标签，以使目标异源多肽或蛋白质以融合蛋白的形式在细菌表达系统中重组表达。常用的融合标签包括弹性蛋白样多肽(ELPs)、麦芽糖结合蛋白(MBP)、钙调素结合蛋白(CaMBPs)、N-

利用物质A(NusA)、翻译起始因子IF2、谷胱甘肽-S-转移酶(GST)、肽基脯氨酰顺反异构酶(PPIases)、二硫键形成蛋白A(DsbA)和硫氧还蛋白(Trx)等。

[0008] 然而,通过这些标签蛋白融合生产的目标肽或蛋白即便在细菌宿主细胞中以可溶形式表达,在后续下游生产工艺中仍然具有其他潜在缺点。这些胞内成功表达的重组多肽或蛋白质进入下游纯化工艺的最初步骤即破胞释放。通常在工业化生产中,采用的破胞方法主要有三类:一是物理破胞,主要是细胞冻融、冷冻干燥和高压匀浆破胞。物理法除了高压匀浆可进行大规模工业化生产外,其余均难以适应规模化生产所需,然而高压匀浆法所用设备昂贵,处理时间通常在10小时以上,不利于压力敏感或易蛋白酶降解的重组多肽或蛋白的生产;二是化学破胞,主要是用溶媒和表面活性剂破胞提取,该方式不仅有机溶剂使用量大,对重组多肽或蛋白的空间构型有潜在影响,容易导致生物活性丧失,还会造成环境污染。三是生物破胞,主要是利用一些复合溶菌酶进行破胞,然而,酶法破胞时间长且一般不彻底,重组多肽或蛋白收率不高,同时外源加入的溶菌酶使得后续纯化过程复杂化,导致产率降低,此外,对成品质量控制也会造成潜在隐患。

[0009] 因此,本领域现有使用的破胞方法对细菌宿主异源表达的重组多肽或蛋白(简称为重组融合蛋白)释放与提取均不能满足高效、低廉且可规模化生产的工业化要求。

发明内容

[0010] 本发明所要解决的技术问题在于针对上述现有技术中的不足,提供硫氧还蛋白突变体。本发明采用的技术方案是:一种硫氧还蛋白突变体,该硫氧还蛋白突变体是具有高温可溶性的非天然蛋白。

[0011] 本发明的工作原理或有益效果为:硫氧还蛋白突变体在长时间高温条件下仍旧能够保持可溶性;硫氧还蛋白突变体同时具有热稳定性和在细菌宿主菌体内的可溶表达性,即表达产生的重组融合蛋白能够快速释放,能够快速提取、提取需要的试剂和实验设备价格低廉且能够用于工业上的大规模生产,高温环境下的可溶稳定性好和不易变性。

[0012] 进一步限定,该硫氧还蛋白突变体是在硫氧还蛋白模型蛋白的氨基酸序列的基础上人工突变得到。

[0013] 本发明通过对天然硫氧还蛋白的分子改造,在保留天然硫氧还蛋白在细菌宿主体内可溶表达性质的同时,赋予其在对细菌宿主发酵液高温搅拌破胞时,长期保持在水相中稳定可溶的特性,而宿主细胞蛋白、其他杂质蛋白和细胞碎片因不具备高温下在水相中长期稳定的性质,经过一定时间保温处理后,会逐渐聚集沉淀,经固液分离处理,从破胞混悬液中获得分离提取,最后本发明提供的硫氧还蛋白突变体会以较高的纯度保留在清液中。因为热破胞不需要特殊设备,(通常在发酵罐中即可进行),破胞保温时间与发酵液体积无关,因此,通过使用本发明提供的硫氧还蛋白突变体可有效满足重组多肽或蛋白质的高效、低廉和规模化生产的商业化需求。

[0014] 本发明还提供硫氧还蛋白突变体的制备方法,包括以下步骤:

[0015] 1)以硫氧还蛋白模型蛋白的编码基因为模板,对其进行随机突变,从而获得编码硫氧还蛋白突变体的基因突变库;

[0016] 2)将硫氧还蛋白模型蛋白的基因突变库克隆至表达载体,获得重组表达载体,然后将所述表达载体转化至适宜的表达宿主,并筛选出表达量高的硫氧还蛋白突变体表达文

库；

[0017] 3)对硫氧还蛋白突变体表达文库中的硫氧还蛋白突变体编码基因进行测序分析，计算出Δ ΔG值和Δ净电荷，并筛选出Δ ΔG有显著下降且Δ净电荷提升的硫氧还蛋白突变体优选文库；

[0018] 4)将硫氧还蛋白突变体优选文库的转化子置于含有抗生素的液体培养基中培养至对数生长期，加入诱导剂开始诱导，继续培养至终点，然后于培养液中加入等体积的缓冲液混匀，于70-100℃水浴中孵育一段时间，取样离心后将上清样进行SDS-PAGE分析，根据条带光密度值大小进行硫氧还蛋白突变体热稳定性及热可溶性能的复筛；

[0019] 5)使用编码具有硫氧还蛋白模型蛋白的基因或一种或多种同源基因，获得具有耐热可溶性能提升的硫氧还蛋白突变体的突变基因，可选择性地将步骤1)到4)重复一次或若干次；

[0020] 6)对获得的具有耐热可溶性能提升的硫氧还蛋白突变体进行表达。

[0021] 本发明还提供硫氧还蛋白突变体在重组融合蛋白生产中的应用，所述重组融合蛋白的结构有两种，分别命名为结构体1和结构体2，其中：结构体1的结构为：硫氧还蛋白突变体——连接肽——切割位点——目标异源多肽或蛋白质，结构体2的结构为：硫氧还蛋白突变体——切割位点——目标异源多肽或蛋白质；结构体1的制备方法为：目标异源多肽或蛋白质与切割位点结合后得到中间体1，中间体1与连接肽结合后得到中间体2，中间体2与硫氧还蛋白突变体的氨基末端或羧基末端结合得到所述结构体1；结构体2的制备方法为：目标异源多肽或蛋白质与切割位点结合后得到中间体1，中间体1与硫氧还蛋白突变体的氨基末端或羧基末端结合得到所述结构体2，所述目标异源多肽或蛋白质为由数量不大于60的氨基酸组成的聚合物。

附图说明

[0022] 图1是表示部分高表达硫氧还蛋白突变体表达筛选实验的SDS-PAGE检测结果图，图中箭头指示硫氧还蛋白突变体条带，其中：顶部数字为部分硫氧还蛋白突变体筛选用编号，M为蛋白Marker。

[0023] 图2是表示部分硫氧还蛋白突变体优选文库的热稳定性及热可溶性能复筛的SDS-PAGE电泳检测结果图，图中矩形框指示硫氧还蛋白突变体条带，M代表蛋白Marker。

[0024] 图3是表示重组工程菌在30L发酵罐培养过程中，重组融合蛋白mTrx39-GSGSG-DDDDK-Arg^{34}GLP-1(9-37)随时间的表达的SDS-PAGE电泳检测结果图，图中箭头指示重组融合蛋白mTrx39-GSGSG-DDDDK-Arg^{34}GLP-1(9-37)条带。

[0025] 图4是表示在重组肠激酶裂解反应中，通过反相HPLC测定的重组融合蛋白mTrx39-GSGSG-DDDDK-Arg^{34}GLP-1(9-37)裂解为Arg^{34}GLP-1(9-37)随时间变化的色谱分析图，检测波长214nm，其中横坐标为检测时间，纵坐标为紫外检测响应值。

[0026] 图5是表示经等电点沉淀后获得的Arg^{3}4GLP-1(9-37)沉淀重溶后，通过反相HPLC测定色谱分析图，其中横坐标为检测时间，纵坐标为紫外检测响应值。

具体实施方式

[0027] 下面结合实施例对本发明做进一步的详细说明，以令本领域技术人员参照说明书

文字能够据以实施。应当理解，本文所使用的诸如“具有”、“包含”以及“包括”术语并不排除一个或多个其它元件或其组合的存在或添加。

[0028] 在文所使用的描述中，“若干”的含义是两个或两个以上，除非另有明确具体的限定。术语“Δ ΔG”在本文中被定义为蛋白质解折叠自由能差值，计算公式为：Δ ΔG＝ΔGTrxWT-ΔGmTrx，量纲kcal/mol，其值反映了野生型蛋白和突变体结构之间的稳定性差异，该值越小，表明突变体比野生型蛋白结构越稳定。术语“ΔGTrxWT”在本文中被定义为硫氧还蛋白模型蛋白分子的解折叠自由能，量纲kcal/mol。术语“ΔGmTrx”在本文中被定义为硫氧还蛋白突变体分子的解折叠自由能，量纲kcal/mol。术语“Δ净电荷”在本文中被定义为在pH＝8.0环境下蛋白质分子净电荷差异，计算公式为：Δ净电荷＝净电荷TrxWT-净电荷mTrx，其值反映了野生型蛋白和突变体结构之间的带电量差异，该值越小，表明突变体比野生型蛋白分子带电量越高。术语“净电荷TrxWT”在本文中被定义为硫氧还蛋白模型蛋白分子的净电荷数。术语“净电荷mTrx”在本文中被定义为硫氧还蛋白突变体分子的净电荷数。术语“光密度”或“OD”在本文中被定义为利用凝胶成像系统中投射白光拍摄记录的蛋白凝胶电泳图中各蛋白条带的入射光与透射光比值的对数：OD＝lg(入射光/透射光)或OD＝lg(1/透光率)，为无量纲值。术语“连接肽”在本文中被定义为连接硫氧还蛋白突变体与切割位点之间的一段多肽序列，具体为硫氧还蛋白突变体C端与切割位点N端之间的一段氨基酸序列。术语“切割位点”在本文中被定义为蛋白内切酶或化学试剂识别的一段特定氨基酸序列。

[0029] 本发明共包含七个方面，第一方面，本发明提供了用于多肽和蛋白质融合表达的一系列硫氧还蛋白突变体(mTrx)，它们是具有可溶性的模型蛋白的突变体。本发明提供的硫氧还蛋白突变体与具有可溶性的模型蛋白相比，所述硫氧还蛋白突变体在长时间高温条件下仍旧能够保持可溶性。

[0030] 在本发明的上下文中，“硫氧还蛋白突变体(mTrx)”指任何这样的蛋白：其具有高温可溶性，但并非从天然来源获得，其氨基酸序列与Geobacillus stearothermophilus和Bacillaceae和Anoxybacillus tepidamans和Falsibacillus pallidus的天然硫氧还蛋白的氨基酸序列有所不同。

[0031] 本发明提供的硫氧还蛋白突变体，所述硫氧还蛋白突变体是在硫氧还蛋白模型蛋白氨基酸序列的基础上人工突变得到，所述硫氧还蛋白模型蛋白为来源于Geobacillus stearothermophilus的硫氧还蛋白，其氨基酸序列如SEQ ID NO.1所示。所述硫氧还蛋白突变体，其优选地在SEQ ID NO.1以下区域内至少一个氨基酸位置被修饰：由位置8、10、12、13、15、17、23、29、33、36、43、44、51、56、58、71、74、82、85、91、93、94和102组成的组1中的氨基酸序列。更优的选择是硫氧还蛋白突变体在SEQ ID NO.1以下区域内至少一个氨基酸位置被修饰：由位置12、13、15、23、29、33、43、44、51、56、58、71、74、82、85、91、93和102组成的组2中的氨基酸序列。最优选地，硫氧还蛋白突变体在SEQ ID NO.1以下区域内至少一个氨基酸位置被修饰：由位置12、15、23、44、51、56、71、85、93和102组成的组3中的氨基酸序列。

[0032] 在氨基酸位置上的修饰包括被另一氨基酸取代，所述另一氨基酸选自20种L构型天然氨基酸，如表1所示。或者，氨基酸位置上的修饰包括删除所述位置上的氨基酸。另外，氨基酸位置上的修饰可包括取代所述氨基酸C端或N端侧的一个或多个氨基酸。

[0033] 表1 20种L构型天然氨基酸

[0034]

氨基酸名称	三字母缩写	单字母缩写
丙氨酸	Ala	A
精氨酸	Arg	R
天冬氨酸	Asp	D
天冬酰胺	Asn	N
半胱氨酸	Cys	C
谷氨酰胺	Gln	Q
谷氨酸	Glu	E
甘氨酸	Gly	G
组氨酸	His	H
异亮氨酸	Ile	I
亮氨酸	Leu	L
赖氨酸	Lys	K
甲硫氨酸	Met	M
苯丙氨酸	Phe	F
脯氨酸	Pro	P
丝氨酸	Ser	S
苏氨酸	Thr	T
色氨酸	Trp	W
酪氨酸	Tyr	Y
缬氨酸	Val	V

[0035] 本发明中所述的具有硫氧还蛋白结构的模型蛋白选自下列蛋白构成的组:具有硫氧还蛋白结构且优选具有SEQ ID NO.1的氨基酸序列,即Geobacillus stearothermophilus菌的硫氧还蛋白或具有SEQ ID NO.2的氨基酸序列,即Bacillaceae菌种的硫氧还蛋白或具有SEQ ID NO.3的氨基酸序列,即Anoxybacillus tepidamans菌种的硫氧还蛋白或具有SEQ ID NO.4的氨基酸序列,即Falsibacillus pallidus菌种的硫氧还蛋白。

[0036] 所述模型蛋白具有与SEQ ID NO.1有至少70%,优选至少75%,更优选至少80%,更优选至少85%,更优选至少90%,最优选至少95%的百分比同一性的氨基酸序列,或具有与SEQ ID NO.2有至少70%,优选至少75%,更优选至少80%,更优选至少85%,更优选至少90%,最优选至少95%的百分比同一性的氨基酸序列,或具有与SEQ ID NO.3有至少70%,优选至少75%,更优选至少80%,更优选至少85%,更优选至少90%,最优选至少95%的百分比同一性的氨基酸序列,或具有与SEQ ID NO.4有至少70%,优选至少75%,更优选至少80%,更优选至少85%,更优选至少90%,最优选至少95%的百分比同一性的氨基酸序列的蛋白。作为用于本发明的具有耐高温且可溶的模型蛋白而言,更优选的是具有高温可溶性,并且具有SEQ ID NO.1的氨基酸序列或具有SEQ ID NO.2的氨基酸序列或具有SEQ ID NO.3的氨基酸序列或具有SEQ ID NO.4的氨基酸序列的蛋白。作为具有高温可溶性的模型蛋白而言,最优选的是Geobacillus stearothermophilus菌的硫氧还蛋白(SEQ ID NO.1)。

[0037] 本发明优选提供了选自下述系列硫氧还蛋白模型的突变体,所述系列由具有SEQ

ID NO.1的氨基酸序列的TrxA和具有SEQ ID NO.2的氨基酸序列的Trx和具有SEQ ID NO.3的氨基酸序列的Trx和具有SEQ ID NO.4的氨基酸序列的Trx，以及具有与SEQ ID NO.1有至少70％，优选至少75％，更优选至少80％，更优选至少85％，更优选至少90％，最优选至少95％的百分比同一性的氨基酸序列，或具有与SEQ ID NO.2有至少70％，优选至少75％，更优选至少80％，更优选至少85％，更优选至少90％，最优选至少95％的百分比同一性的氨基酸序列且具有高温可溶性的蛋白结构，并且，所述突变体具有至少在以下位点的修饰：

[0038] 选自8、10、12、13、15、17、23、29、33、36、43、44、51、56、58、71、74、82、85、91、93、94和102的一个或多个氨基酸位置。选自12、13、15、23、29、33、43、44、51、56、58、71、74、82、85、91、93和102的一个或多个氨基酸位置。选自12、15、23、44、51、56、71、85、93和102的一个或多个氨基酸位置。选自12、15、23、44、51、56、71、85、93和102的2个或多个氨基酸位置，更优选地在位置15+51或23+51或85+93或15+93或23+93或51+102或15+102或15+23。选自12、13、15、23、29、33、43、44、51、56、58、71、74、82、85、91、93和102的3个或多个氨基酸位置，更优选地在位置23+85+93或15+85+93或51+85+93或44+85+93或85+93+102或23+44+51或44+51+85。选自12、13、15、23、29、33、43、44、51、56、58、71、74、82、85、91、93和102的4个或多个氨基酸位置，更优选地在位置23+51+85+93或15+51+85+93或15+23+85+93或23+44+51+85或44+51+85+93或23+44+85+93。选自12、13、15、23、29、33、43、44、51、56、58、71、74、82、85、91、93和102的5个或多个氨基酸位置，更优选地在位置15+23+51+85+93或23+51+85+93+102或23+44+51+85+93或15+44+51+85+93或12+56+71+93+102或12+15+56+71+102。选自12、13、15、23、29、33、43、44、51、56、58、71、74、82、85、91、93和102的6个或多个氨基酸位置，更优选地在位置15+23+29+44+51+85或15+23+44+51+85+93或12+23+56+71+93+102或12+56+71+85+93+102或15+23+44+51+85+93或23+44+51+85+93+102。选自12、13、15、23、29、33、43、44、51、56、58、71、74、82、85、91、93和102的7个或多个氨基酸位置，更优选地在位置15+23+44+51+85+93+102或15+23+44+51+71+85+93或15+23+44+51+74+85+93或12+15+23+44+51+85+93或15+23+43+44+51+85+93。选自12、13、15、23、29、33、43、44、51、56、58、71、74、82、85、91、93和102的8个或多个氨基酸位置，更优选地在位置12+15+23+44+51+85+93+102或15+23+44+51+56+85+93+102或15+23+44+51+71+85+93+102或15+23+29+44+51+85+93+102或12+15+23+44+51+56+85+93或15+23+29+44+51+56+71+85。选自12、13、15、23、29、33、43、44、51、56、58、71、74、82、85、91、93和102的9个或多个氨基酸位置，更优选地在位置12+15+23+44+51+56+71+85+93或15+23+43+44+51+74+85+91+93或15+23+29+44+51+71+85+93+102或15+23+44+51+56+71+85+93+102或12+15+23+44+51+71+85+93+102。选自12、13、15、23、29、33、43、44、51、56、58、71、74、82、85、91、93和102的10个或多个氨基酸位置，更优选地在位置12+15+23+29+44+51+56+71+85+93或15+23+43+44+51+71+74+85+91+93或15+23+43+44+51+74+85+91+93+102或12+15+23+44+51+56+71+85+93+102。选自12、13、15、23、29、33、43、44、51、56、58、71、74、82、85、91、93和102的11个氨基酸位置，更优选地在位置12+15+23+44+51+56+71+85+91+93+102或12+15+23+29+44+51+56+71+85+93+102或12+15+23+44+51+56+71+74+85+93+102或12+15+23+43+44+51+56+71+85+93+102。选自12、13、15、23、29、33、43、44、51、56、58、71、74、82、85、91、93和102的12个氨基酸位置，更优选地在位置12+15+23+29+44+51+56+71+85+91+93+102或12+15+23+44+51+56+71+74+85+91+93+102或12+15+23+43+44+51+56+71+85+91+93+102或12+15+23+43+44+51+56+71+74+85+93+102。选自12、13、15、23、29、33、43、

44、51、56、58、71、74、82、85、91、93和102的13个氨基酸位置，更优选地在位置12+15+23+29+44+51+56+71+74+85+91+93+102或12+15+23+43+44+51+56+71+74+85+91+93+102或12+15+23+44+51+56+71+74+82+85+91+93+102或12+15+23+44+51+56+58+71+74+85+91+93+102。选自12、13、15、23、29、33、43、44、51、56、58、71、74、82、85、91、93和102的14个氨基酸位置，更优选地在位置12+15+23+43+44+51+56+71+74+82+85+91+93+102或12+15+23+29+43+44+51+56+71+74+85+91+93+102或12+13+15+23+43+44+51+56+71+74+85+91+93+102或12+15+23+33+43+44+51+56+71+74+85+91+93+102或12+15+23+43+44+51+56+58+71+74+85+91+93+102。选自12、13、15、23、29、33、43、44、51、56、58、71、74、82、85、91、93和102的15个氨基酸位置，更优选地在位置12+13+15+23+29+43+44+51+56+71+74+85+91+93+102或12+13+15+23+33+43+44+51+56+71+74+85+91+93+102或12+15+23+43+44+51+56+58+71+74+82+85+91+93+102。选自12、13、15、23、29、33、43、44、51、56、58、71、74、82、85、91、93和102的16个氨基酸位置，更优选地在位置12+13+15+23+29+43+44+51+56+58+71+74+85+91+93+102或12+13+15+23+33+43+44+51+56+58+71+74+85+91+93+102或12+15+23+33+43+44+51+56+58+71+74+82+85+91+93+102。选自12、13、15、23、29、33、43、44、51、56、58、71、74、82、85、91、93和102的17个氨基酸位置，更优选地在位置12+13+15+23+33+43+44+51+56+58+71+74+82+85+91+93+102或12+13+15+23+29+33+43+44+51+56+58+71+74+85+91+93+102或12+13+15+23+29+43+44+51+56+58+71+74+82+85+91+93+102。选自12、13、15、23、29、33、43、44、51、56、58、71、74、82、85、91、93和102的18个氨基酸位置。

[0039] 本发明还提供前述定义的硫氧还蛋白突变体，其为Geobacillus stearothermophilus菌的硫氧还蛋白的突变体，所述硫氧还蛋白突变体在选自以下氨基酸位置上被修饰：由位置D8、T10、A12、A13、T15、D17、D23、C29、R33、A36、D43、R44、K51、N56、D58、S71、T74、Q82、D85、Q91、K93、E94和K102组成的组，其中数字前的字母表示各自的氨基酸(单字母缩写)，后面的数字是SEQ ID NO.1中所述Geobacillus stearothermophilus的硫氧还蛋白的氨基酸序列中的氨基酸位置。

[0040] 优选的硫氧还蛋白突变体是至少在下述位点具有修饰的Geobacillus stearothermophilus硫氧还蛋白突变体：

[0041] 选自D8、T10、A12、A13、T15、D17、D23、C29、R33、A36、D43、R44、K51、N56、D58、S71、T74、Q82、D85、Q91、K93、E94和K102的一个或多个氨基酸位置。选自A12、A13、T15、D23、C29、R33、D43、R44、K51、N56、D58、S71、T74、Q82、D85、Q91、K93和K102的一个或多个氨基酸位置。选自A12、T15、D23、R44、K51、N56、S71、D85、K93和K102的一个或多个氨基酸位置。选自A12、T15、D23、R44、K51、N56、S71、D85、K93和K102的2个或多个氨基酸位置。选自A12、A13、T15、D23、C29、R33、D43、R44、K51、N56、D58、S71、T74、Q82、D85、Q91、K93和K102的3个或多个氨基酸位置。选自A12、A13、T15、D23、C29、R33、D43、R44、K51、N56、D58、S71、T74、Q82、D85、Q91、K93和K102的4个或多个氨基酸位置。选自A12、A13、T15、D23、C29、R33、D43、R44、K51、N56、D58、S71、T74、Q82、D85、Q91、K93和K102的5个或多个氨基酸位置。选自A12、A13、T15、D23、C29、R33、D43、R44、K51、N56、D58、S71、T74、Q82、D85、Q91、K93和K102的6个或多个氨基酸位置。选自A12、A13、T15、D23、C29、R33、D43、R44、K51、N56、D58、S71、T74、Q82、D85、Q91、K93和K102的7个或多个氨基酸位置。选自A12、A13、T15、D23、C29、R33、D43、R44、K51、N56、D58、S71、T74、Q82、D85、Q91、K93和K102的8个或多个氨基酸位置。选自A12、A13、T15、D23、C29、

R33、D43、R44、K51、N56、D58、S71、T74、Q82、D85、Q91、K93和K102的9个或多个氨基酸位置。选自A12、A13、T15、D23、C29、R33、D43、R44、K51、N56、D58、S71、T74、Q82、D85、Q91、K93和K102的10个或多个氨基酸位置。选自A12、A13、T15、D23、C29、R33、D43、R44、K51、N56、D58、S71、T74、Q82、D85、Q91、K93和K102的11个氨基酸位置。选自A12、A13、T15、D23、C29、R33、D43、R44、K51、N56、D58、S71、T74、Q82、D85、Q91、K93和K102的12个氨基酸位置。选自A12、A13、T15、D23、C29、R33、D43、R44、K51、N56、D58、S71、T74、Q82、D85、Q91、K93和K102的13个氨基酸位置。选自A12、A13、T15、D23、C29、R33、D43、R44、K51、N56、D58、S71、T74、Q82、D85、Q91、K93和K102的14个氨基酸位置。选自A12、A13、T15、D23、C29、R33、D43、R44、K51、N56、D58、S71、T74、Q82、D85、Q91、K93和K102的15个氨基酸位置。选自A12、A13、T15、D23、C29、R33、D43、R44、K51、N56、D58、S71、T74、Q82、D85、Q91、K93和K102的16个氨基酸位置。选自A12、A13、T15、D23、C29、R33、D43、R44、K51、N56、D58、S71、T74、Q82、D85、Q91、K93和K102的17个氨基酸位置。选自A12、A13、T15、D23、C29、R33、D43、R44、K51、N56、D58、S71、T74、Q82、D85、Q91、K93和K102的18个氨基酸位置。

[0042] Geobacillus stearothermophilus的硫氧还蛋白中各个位置的优选修饰如下(使用氨基酸单字母缩写代码)：

[0043] SEQ ID NO.1位置8的D被Q、E、A、K、R、S或T取代，更优选地被Q、E、A、或K取代，最优选被A取代。SEQ ID NO.1位置10的T被D、E、Q、K、R、S或I取代，更优选地被D、I或Q取代，最优选被I取代。SEQ ID NO.1位置12的A被G、E、Q、K、R、S或I取代，更优选地被G或I取代，最优选被I取代。SEQ ID NO.1位置13的A被D、E、Q、K、R、S或I取代，更优选地被I或Q取代，最优选被I取代。SEQ ID NO.1位置15的T被I、V、A、T取代或被删除，更优选地被I或Q取代，最优选I取代。SEQ ID NO.1位置17的D被N、E、Q、K、R、S、G或A取代，更优选地被E或Q取代，最优选E取代。SEQ ID NO.1位置23的D被T、Q、K、R、S、G或N取代，更优选地被S或T取代，最优选T取代。SEQ ID NO.1位置29的C被R、D、S、T、N、K、W取代或被删除，更优选地被R、S、K、W取代或被删除，最优选被S取代和被删除。SEQ ID NO.1位置33的R被D、E、Q、K、N或R取代，更优选地被K、Q或N取代，最优选被K取代。SEQ ID NO.1位置36的A被G、S、T、K、R、H、L取代或被删除，更优选地被L、S、K取代或被删除，最优选被删除。SEQ ID NO.1位置43的D被Q、Y、W、P、I或E取代，更优选地被Q或E取代，最优选被E取代。SEQ ID NO.1位置44的R被L、A、Q、S、T、K或H取代，更优选地被L、A或Q取代，最优选被Q取代。SEQ ID NO.1位置51的K被Q、D、R、E或I取代，更优选地被Q或R取代，最优选K51R取代。SEQ ID NO.1位置56的N被E、Q、R、K、W或Y取代，更优选地被K、R或Q取代，最优选被K取代。SEQ ID NO.1位置58的D被W、R、T、D、Q、H或E取代，更优选地被R、Q或E取代，最优选被E取代。SEQ ID NO.1位置71的S被I、L、K、D、M、E或N取代，更优选地被I、L或M取代，最优选被M取代。SEQ ID NO.1位置74的T被I、L、D、S、M或V取代，更优选地被I或M取代，最优选被I取代。SEQ ID NO.1位置82的E被Q、N、R或K取代，更优选地被K、N或R取代，最优选被K取代。SEQ ID NO.1位置85的D被R、Q、H、N、E或K取代，更优选地被E、Q或K取代，最优选被E取代。SEQ ID NO.1位置91的Q被E、Y、I、N或R取代，更优选地被R或K取代，最优选被K取代。SEQ ID NO.1位置93的K被P、M、S、I、Q或N取代，更优选地被P或M取代，最优选被P取代。SEQ ID NO.1位置94的E被D、S、E、Q、N、K或R取代，更优选地被D、S、E或Q取代，最优选E94Q取代。SEQ ID NO.1位置102的K被D、S、E、H、Y、K或R取代，更优选地被H、Y或S取代，最优选被H取代。

[0044] 高度优选的硫氧还蛋白突变体如表2所示，每个突变体由其编号标识：突变被引入

来自Geobacillus stearothermophilus的硫氧还蛋白(SEQ ID NO.1)。根据实施例3中的ΔΔG和Δ净电荷计算方法,以来自Geobacillus stearothermophilus的硫氧还蛋白(SEQ ID NO.1)的ΔG和净电荷为模型对照。

[0045] 本发明提供的硫氧还蛋白突变体与具有耐热性的硫氧还蛋白模型蛋白相比,ΔΔG更小或Δ净电荷更大,更利于后期基于高温和电荷性质开发的下游纯化工艺。硫氧还蛋白突变体ΔΔG降低了至少-1kcal/mol、更优选地至少-2kcal/mol、更优选地至少-3kcal/mol、更优选地至少-4kcal/mol、更优选地至少-5kcal/mol、更优选地至少-6kcal/mol、更优选地至少-7kcal/mol、更优选地至少-8kcal/mol、更优选地至少-9kcal/mol、更优选地至少-10kcal/mol、更优选地至少-11kcal/mol、更优选地至少-12kcal/mol、更优选地至少-13kcal/mol。硫氧还蛋白突变体Δ净电荷提高了至少-0.8、更优选地至少-1、更优选地至少-2、更优选地至少-3。对高温环境更稳定且净电荷更大的硫氧还蛋白突变体更有利于应用在目标异源多肽或蛋白的生产方法中,详细方法在下文会进一步描述。最有价值地是,本发明提供的硫氧还蛋白突变体同时具有热稳定性和在细菌宿主菌体内的可溶表达性,以及高温环境下的可溶稳定性。

[0046] 第二方面,本发明提供编码本发明硫氧还蛋白突变体的多核苷酸。编码本发明硫氧还蛋白突变体的多核苷酸可以是编码本发明合适氨基酸序列的任何多核苷酸。或者,本发明的多核苷酸可包括含编码序列,其中多种氨基酸的密码子选择不同于来自Geobacillus stearothermophilus中的密码子选择。例如所述密码子可以被优化来适合特定的宿主细胞的密码子偏好需求,并用于编码硫氧还蛋白突变体的DNA片段转化。

[0047] 第三方面,本发明提供包含前文所述多核苷酸的表达载体。

[0048] 第四方面,本发明提供经转化的宿主细胞,其利用本发明的多核苷酸或本发明的表达载体转化。经转化的宿主细胞可用于生产本发明的硫氧还蛋白突变体,或所述宿主细胞可用于生产目的异源多肽或蛋白。

[0049] 用于生产本发明硫氧还蛋白突变体的宿主细胞优选地是下述宿主细胞,所述宿主细胞因其在细胞外或细胞内有效生产多肽或蛋白质而被本领域所共知,例如微生物,如真菌、酵母菌和细菌。优选的宿主细胞的实例包括,但不仅限于以下属:Saccharomyces(例如S.cerevisiae)、Pichia(例如P.pastoris)、Bacillus(例如B.subtilis、B.licheniformis)、Escherichia(例如E.coli)。

[0050] 用于生产目的异源多肽或蛋白的宿主细胞优选地是下述宿主细胞,所述宿主细胞因其在细胞外或细胞内有效生产多肽或蛋白质而被本领域所共知。优选的宿主细胞的实例包括,但不仅限于以下属:Saccharomyces(例如S.cerevisiae)、Pichia(例如P.pastoris)、Bacillus(例如B.subtilis、B.licheniformis)、Escherichia(例如E.coli)。

[0051] 第五方面,本发明提供了利用本发明前文所述硫氧还蛋白突变体与目标异源多肽或蛋白质的重组融合蛋白,该重组融合蛋白的结构为:(1)硫氧还蛋白突变体——连接肽——切割位点——目标异源多肽或蛋白质或(2)硫氧还蛋白突变体——切割位点——目标异源多肽或蛋白质,目标异源多肽或蛋白质与连接肽和/或切割位点结合后融合到本发明提供的硫氧还蛋白突变体的氨基末端或羧基末端。所述连接肽包含柔性连接肽或刚性连接肽或无连接肽,所述切割位点包含但不限于蛋白内切酶切割位点或化学切割位点。

[0052] 本发明连接肽优选地是下列柔性连接肽或刚性连接肽,所述的连接肽在重组融合

蛋白生产中的功能被本领域所共知，优选的柔性连接肽实例包括，但不仅限于以下类别(氨基酸单字母缩写)：(GGGGS)n(n≤6)，GSAGSAAGSGEF，GSGSG，GSG和(G)n(n≤8)；优选的刚性连接肽实例包括，但不仅限于以下类别(氨基酸单字母缩写)：(EAAAK)n(n≤6)，(XP)n(X可以设计为任意一种氨基酸，n≤18)，(PT)nP(n≤18)和A(EAAAK)4ALEA(EAAAK)4A。

[0053] 本发明切割位点包含的位点优选地是下列蛋白内切酶切割位点或化学切割位点，所述的蛋白内切酶切割位点或化学切割位点能有效地被相应的蛋白内切酶或化学试剂识别并特异性地切割而被本领域所共知，优选的蛋白内切酶和化学试剂的实例包括，但不仅限于以下类别：凝血酶，烟草蚀纹病毒蛋白酶，3C蛋白酶，V8蛋白酶，肠激酶，胰蛋白酶和赖氨酰肽链内切酶，所述化学试剂包括溴化氰，羟胺和甲酸。

[0054] 第六方面，本发明提供了利用本发明前文所述结构为：(1)硫氧还蛋白突变体——连接肽——切割位点——目标异源多肽或蛋白质；(2)硫氧还蛋白突变体——切割位点——目标异源多肽或蛋白质的重组融合蛋白，以及生产目标异源多肽和蛋白质的方法。该方法包括在有利于生产上述重组融合蛋白的条件下培养本发明的经转化的宿主细胞，以及培养结束后破胞回收上述重组融合蛋白。优选的上述重组融合蛋白中目标异源多肽或蛋白质是指由60个或更少氨基酸组成的聚合物，实例包括，但不仅限于以下类型：胰高血糖素样肽-1(Glucagon-like peptide-1，GLP-1)及其类似聚合物(例如GLP-1(7-37)、GLP-1(9-37)、GLP-1(7-36)、GLP-1(11-37)、Arg34GLP-1(7-37)、Arg34GLP-1(9-37))；甲状旁腺激素(Parathyroid hormone，PTH)及其类似聚合物(例如hPTH(1-34))；胰高血糖素样肽-2(Glucagon-like peptide-2，GLP-2)及其类似聚合物(例如GLP-2(1-33)、Lys17Arg30GLP-2(1-33)、)；促肾上腺皮质激素(Adrenocorticotropic hormone，ACTH)及其类似聚合物；黑色素细胞刺激素(Melanocyte-stimulating hormones，MSH)及其类似聚合物；肾上腺髓质素(Adrenomedullin ADM)及其类似聚合物；心房钠尿肽(Atrial natriuretic peptide，ANP)及其类似聚合物；催产素(Oxytocin，Oxt)及其类似聚合物；生长激素释放激素(Growth hormone-releasing hormone，GHRH)及其类似聚合物；血管活性肠肽(Vasoactive intestinal peptide，VIP)及其类似聚合物；尿鸟苷蛋白(Uroguanylin)及其类似聚合物。在优选的实施方案中，本发明提供通过培养Escherichia coli宿主菌株来生产目标异源多肽或蛋白Arg34GLP-1(9-37)的方法，所述菌株已用编码本发明硫氧还蛋白突变体的选定的本发明多核苷酸转化。高度优选的硫氧还蛋白突变体选自表2组成的组。

[0055] 第七方面，本发明提供获得本发明硫氧还蛋白突变体的方法，其中所述方法包括以下步骤：

[0056] 以硫氧还蛋白模型蛋白的编码基因为模板，对其进行随机突变，从而获得编码硫氧还蛋白突变体的基因突变库；

[0057] 将硫氧还蛋白模型蛋白的编码基因突变库克隆至表达载体，获得重组表达载体，然后将所述表达载体转化至适宜的表达宿主，并筛选出表达量高的硫氧还蛋白突变体表达文库；

[0058] 对上述文库中的硫氧还蛋白突变体编码基因进行测序分析，采用本领域所共知的算法软件，例如Foldx、Rosetta Δ ΔG、BindProfX、PROTEIN CALCULATOR v3.4等，计算出Δ ΔG值和Δ净电荷，并筛选出Δ ΔG有显著下降且Δ净电荷提升的硫氧还蛋白突变体优选文库；

[0059] 将上述筛选出的硫氧还蛋白突变体优选文库的转化子于含有抗生素的液体培养基中培养至对数生长期，加入诱导剂开始诱导，继续培养至终点，然后于培养液中加入等体积的缓冲液混匀，于70-100℃水浴中孵育一段时间，取样离心后将上清样进行SDS-PAGE分析，根据条带光密度值大小进行硫氧还蛋白突变体热稳定性及热可溶性能的复筛；

[0060] 使用编码具有硫氧还蛋白模型蛋白的基因或一种或多种同源基因，获得具有耐热可溶性能提升的硫氧还蛋白突变体的突变基因，可选择性地将步骤1到4重复一次或若干次；

[0061] 可根据本领域已知的方法，对获得的具有耐热可溶性能提升的硫氧还蛋白突变体进行表达。

[0062] 优选的硫氧还蛋白模型蛋白选自由下述蛋白组成的组：来自Geobacillus stearothermophilu的硫氧还蛋白，其优选的具有SEQ ID NO.1的氨基酸序列，更优选地至少70％、更优选地至少75％、更优选地至少80％、更优选地至少85％、更优选地至少90％、最优选地至少95％的同一性，例如表2中概括的硫氧还蛋白。

[0063] 可采用分子生物学中所熟知的手段来进行上述步骤1-6的实施。可采用分子生物学中的任何基因突变手段进行编码基因突变，较为适宜的手段有易错PCR(Error-prone PCR)和基因改组(DNA Shuffling)。

[0064] 实施例1

[0065] 步骤1)硫氧还蛋白突变基因库的建立

[0066] 将来自Geobacillus stearothermophilu的硫氧还蛋白氨基酸序列(SEQ ID NO.1)按照E.coli密码子偏好性优化后由通用生物系统(安徽)有限公司合成，将此基因命名为TrxWT，并插入载体pET28a的多克隆位点，得到重组载体pET28a-TrxWT。设计引物并以硫氧还蛋白TrxWT基因为模板，进行易错PCR扩增，并将获得的随机突变片段混合物插入载体pET28a，得到含有突变基因的重组载体pET28a-TrxWT_epmut mix，将重组载体混合物转入感受态细胞E.coli BL21(DE3)，获得库容量约11000-12000的表达文库。随机选择了50个单菌落进行双酶切验证，结果显示大约25％为假阳性转化子，随后，随机选择了15个阳性转化子的菌株进行测序分析，以测定突变率。

[0067] 步骤2)高表达硫氧还蛋白突变体表达文库的建立

[0068] 将获得的硫氧还蛋白突变体大肠杆菌表达文库中阳性转化子接种于含有50μg/mL硫酸卡那霉素的LB液体培养基的96孔深孔板中，37℃，700rpm培养过夜。将所得培养液作为种子，以1％接种量接种于新鲜的含50μg/mL硫酸卡那霉素的LB液体培养基中，于96孔深孔板中37℃，700rpm培养4-6h，将温度调至30℃并加入IPTG至终浓度0.4mM，继续振荡培养20-25h。所得培养液于室温4000rpm离心10min，收集菌体，保存于-20℃。所得菌体样品加入一定体积的上样缓冲液，100℃煮沸5min制样。SDS-PAGE分析目的蛋白条带的光密度值，其中部分筛选结果如图1中的(a)和(b)所示，由图中结果可知硫氧还蛋白突变体获得了正常表达。

[0069] 步骤3)硫氧还蛋白突变体优选文库的建立

[0070] 从高表达硫氧还蛋白突变体表达文库获得的阳性转化子中随机挑选若干进行序列测定分析，根据测序结果利用本领域所共知的算法软件，例如Foldx、RosettaΔΔG、BindProfX、PROTEIN CALCULATOR v3.4等，以来自Geobacillus stearothermophilus的模

板硫氧还蛋白(SEQ ID NO.1)的ΔG和净电荷为对照,按照下述公式计算出每个硫氧还蛋白突变体与模板硫氧还蛋白之间的Δ ΔG值和Δ净电荷,其中部分计算结果如表2所示。从表中结果可知筛选获得的硫氧还蛋白突变体无论是Δ ΔG还是Δ净电荷都较天然来源模板硫氧还蛋白的性质得到提升,其中Δ ΔG值越小,表明硫氧还蛋白突变体稳定性越强;Δ净电荷值越小,表明硫氧还蛋白突变体极性越大。

[0071] Δ ΔG＝ΔG TrxWT-ΔGmTrx;

[0072] Δ净电荷＝净电荷TrxWT-净电荷mTrx;

[0073] 表2硫氧还蛋白突变体及其Δ ΔG和Δ净电荷

[0074]

突变体编号	突变位置	稳定及净电荷性能	
		ΔΔG (kcal/mol)	Δ净电荷
mTrx 1	D23T+K93P	-5.34	-1.0
mTrx 2	T15I+D23T+C29S+R44Q+K51R+D85E+K93P	-7.87	-1.8
mTrx 3	K93P	-2.87	-2.0
mTrx 4	D23T+R44Q+K51R+D85E+K93P	-8.24	-2.0
mTrx 5	T15I+D23T+R44Q	-4.22	-1.0
mTrx 6	T15I+K93P	-4.57	-2.0
mTrx 7	T15I+D23T+R44Q+K51R+D85E+K93P	-8.87	-2.0
mTrx 8	T15I+D23T+R44Q+K51R+D85E+Q91K+K93P	-9.52	-1.0
mTrx 9	R44Q+K51R+D85E+K93P	-6.47	-3.0
mTrx 10	D85E+K93P	-5.16	-2.0

[0075]

mTrx 11	D23T+K51R+K93P	-5.37	-1.0
mTrx 12	C29S+K93P	-1.87	-1.8
mTrx 13	D23T+R44Q+K51R	-5.85	-1.0
mTrx 14	T15I+D23T+R44Q+K51R+D85E+K93P+K102H	-9.56	-3.0
mTrx 15	T15I+D23T+R44Q+K51R+T74I+D85E+K93P	-9.56	-2.0
mTrx 16	T15I+D23T+D43E+R44Q+K51R+D85E+K93P	-9.43	-2.0
mTrx 17	T15I+D23T+R44Q+K51R+N56K+D85E+K93P	-9.56	-1.0
mTrx 18	T15I+D23T+C29S+R44Q+K51R+D85E	-9.58	-0.8
mTrx 19	T15I+D23T+R44Q+K51R	-6.06	-1.0
mTrx 20	D23T+R44Q+K51R+D85E	-6.49	-1.0
mTrx 21	T15I+D23T+R44Q+K51R+N56K+D85E+K93P+K102H	-10.34	-2.0
mTrx 22	D23T+K51R+D85E+K93P+K102H	-9.17	-2.0
mTrx 23	T15I+D23T+D43E+R44Q+K51R+T74I+D85E+Q91K+K93P	-11.26	-1.0
mTrx 24	K51R+D85E+K93P	-6.82	-2.0
mTrx 25	T15I+D23T+R44Q+K51R+D85E	-7.19	-1.0
mTrx 26	R44Q+K51R+D85E	-4.62	-2.0
mTrx 27	T15I+D23T+R44Q+K51R+D85E+K93M	-7.16	-2.0
mTrx 28	T15I+D23T+D43E+R44Q+K51R+T74I+D85E+Q91K+K93P+K102H	-11.72	-2.0
mTrx 29	T15I+D23T+C29S+R44Q+K51R+S71M+D85E+K93P+K102H	-9.61	-2.7
mTrx 30	T15I+K51R+D85E+K93P	-7.62	-2.0
mTrx 31	A12I+T15I+D23T+R44Q+K51R+D85E+K93P	-9.62	-2.0
mTrx 32	T15I+D23T+R44Q+K51R+S71M+D85E+K93P	-9.92	-2.0
mTrx 33	D43E+T74I+Q91K	-2.67	-0.1
mTrx 34	A12I+T15I+D23T+R44Q+K51R+D85E+K93P+K102H	-10.28	-3.0
mTrx 35	T15I+D23T+R44Q+K51R+S71M+D85E+K93P+K102H	-10.48	-3.0
mTrx 36	A12I+T15I+D23T+R44Q+K51R+N56K+D85E+K93P	-10.27	-1.0
mTrx 37	T15I+D23T+C29S+R44Q+K51R+N56K+S71M+D85E	-10.61	-1.0
mTrx 38	T15I+D23T+R44Q+K51R+N56K+S71M+D85E+K93P+K102H	-11.02	-2.0
mTrx 39	T15I+D23T+D43E+R44Q+K51R+S71M+T74I+D85E+Q91K+K93P	-12.40	-1.0
mTrx 40	A12I+D23T+N56K+S71M+K93P+K102H	-8.24	-1.0
mTrx 41	A12I+T15I+D23T+C29S+R44Q+K51R+N56K+S71M+D85E+K93P	-10.23	-0.8
mTrx 42	A12I+T15I+D23T+C29S+R44Q+K51R+N56K+S71M+D85E+K93P+K102H	-7.74	-1.8

[0076]

mTrx 43	A12I+T15I+D23T+C29S+R44Q+K51R+N56K+S71M+D85E+Q91K+K93P+K102H	-8.00	-0.8
mTrx 44	A12I+T15I+D23T+R44Q+K51R+S71M+D85E+K93P+K102H	-11.23	-3.0
mTrx 45	A12I+N56K+S71M+K93P+K102H	-6.07	-2.0
mTrx 46	A12I+T15I+D23T+R44Q+K51R+N56K+S71M+D85E+K93P	-11.13	-1.0
mTrx 47	A12I+T15I+D23T+C29S+R44Q+K51R+N56K+S71M+T74I+D85E+Q91K+K93P+K102H	-8.41	-0.8
mTrx 48	A12I+T15I+D23T+R44Q+K51R+N56K+S71M+D85E+K93P+K102H	-11.85	-2.0
mTrx 49	A12I+T15I+D23T+R44Q+K51R+N56K+S71M+D85E+Q91K+K93P+K102H	-8.40	-1.0
mTrx 50	A12I+N56K+S71M+D85E+K93P+K102H	-7.55	-2.0
mTrx 51	A12I+C29S+N56K+S71M+K102H	-2.22	-0.8
mTrx 52	A12I+T15I+D23T+C29S+D43E+R44Q+K51R+N56K+S71M+T74I+D85E+Q91K+K93P+K102H	-8.95	-0.8
mTrx 53	A12I+T15I+D23T+R44Q+K51R+N56K+S71M+T74I+D85E+K93P+K102H	-8.16	-2.0
mTrx 54	A12I+T15I+D23T+R44Q+K51R+N56K+S71M+T74I+D85E+Q91K+K93P+K102H	-8.79	-1.0
mTrx 55	A12I+T15I+D23T+D43E+R44Q+K51R+N56K+S71M+D85E+K93P+K102H	-8.29	-2.0
mTrx 56	A12I+T15I+D23T+D43E+R44Q+K51R+N56K+S71M+D85E+Q91K+K93P+K102H	-8.94	-1.0
mTrx 57	A12I+T15I+D23T+D43E+R44Q+K51R+N56K+S71M+T74I+Q82K+D85E+Q91K+K93P+K102H	-9.84	1.0
mTrx 58	A12I+T15I+D23T+D43E+R44Q+K51R+N56K+S71M+T74I+D85E+K93P+K102H	-8.70	-2.0
mTrx 59	A12I+T15I+D23T+D43E+R44Q+K51R+N56K+S71M+T74I+D85E+Q91K+K93P+K102H	-9.34	-1.0
mTrx 60	A12I+T15I+N56K+S71M+K102H	-4.77	-1.0
mTrx 61	A12I+A13I+T15I+D23T+C29S+R33K+D43E+R44Q+K51R+N56K+D58E+S71M+T74I+Q82K+D85E+Q91K+K93P+K102H	-10.14	1.2
mTrx 62	A12I+A13I+T15I+D23T+C29S+R33K+D43E+R44Q+K51R+N56K+D58E+S71M+T74I+D85E+Q91K+K93P+K102H	-9.73	-0.8
mTrx 63	A12I+A13I+T15I+D23T+R33K+D43E+R44Q+K51R+N56K+S71M+T74I+D85E+Q91K+K93P+K102H	-9.53	-1.0
mTrx 64	A12I+A13I+T15I+D23T+C29S+D43E+R44Q+K51R+N56K+S71M+T74I+D85E+Q91K+K93P+K102H	-9.14	-0.8

[0077]

mTrx 65	A12I+A13I+T15I+D23T+R33K+D43E+R44Q+K51R+N56K+D58E+S71M+T74I+Q82K+D85E+Q91K+K93P+K102H	-10.46	1.0
mTrx 66	A12I+A13I+T15I+D23T+R33K+D43E+R44Q+K51R+N56K+D58E+S71M+T74I+D85E+Q91K+K93P+K102H	-10.09	-1.0
mTrx 67	A12I+A13I+T15I+D23T+C29S+D43E+R44Q+K51R+N56K+D58E+S71M+T74I+D85E+Q91K+K93P+K102H	-9.70	-0.8
mTrx 68	A12I+A13I+T15I+D23T+D43E+R44Q+K51R+N56K+S71M+T74I+D85E+Q91K+K93P+K102H	-9.53	-1.0
mTrx 69	K51R+K102H	-2.71	-2.0
mTrx70	R44Q+K102H	-2.09	-3.0
mTrx71	D23T+K102H	-2.77	-1.0
mTrx72	T15I+K102H	-2.41	-2.0
mTrx73	D23T+K51R	-4.23	0.0
mTrx74	T15I+D23T	-3.74	0.0
mTrx75	T15I+K51R	-3.61	-1.0
mTrx76	T15I+D85E+K93P	-6.96	-2.0
mTrx77	D23T+D85E+K93P	-7.75	-1.0
mTrx78	R44Q+D85E+K93P	-5.22	-3.0
mTrx79	D85E+K93P+K102H	-5.09	-3.0
mTrx80	T15I+D23T+K51R	-5.94	0.0
mTrx81	T15I+D23T+D85E+K93P	-6.60	-1.0
mTrx82	D23T+R44Q+D85E+K93P	-5.31	-2.0
mTrx83	D23T+K51R+D85E+K93P	-8.48	-1.0
mTrx84	T15I+D23T+K51R+D85E+K93P	-9.17	-1.0
mTrx85	T15I+R44Q+K51R+D85E+K93P	-7.04	-3.0
mTrx86	D23T+R44Q+K51R+D85E+K93P+K102H	-7.00	-3.0

[0078] 由表2可知,已获得的86个硫氧还蛋白突变体,其具有ΔΔG从-2.22kcal/mol到-12.40kcal/mol范围内的降低和Δ净电荷从0.0到-3.0范围内的提高。

[0079] 步骤4)硫氧还蛋白突变体优选文库的热稳定性及热可溶性能复筛

[0080] 将获得的硫氧还蛋白突变体优选文库获得的阳性转化子接种于含有50μg/mL硫酸卡那霉素的LB液体培养基的96孔深孔板中,37℃,700rpm培养过夜。将所得培养液作为种子,以1%接种量接种于新鲜的含50μg/mL硫酸卡那霉素的LB液体培养基的24孔板中,37℃,500rpm培养4-6h,将温度调至30℃并加入乳糖至终浓度2%,继续振荡培养20-25h。所得培养液加入等体积的缓冲液混匀后,于90℃水浴和100℃沸水浴中分别孵育30min和60min,取样离心12000rpm,3min,上清及沉淀分别制样进行SDS-PAGE。根据SDS-PAGE光密度值结果分析硫氧还蛋白突变体的热稳定性及热可溶性能的情况,其中部分筛选结果如图2所示,表明这些硫氧还蛋白突变体较之培养液中菌体蛋白、细胞碎片和其他杂质蛋白而言具有在高温条件下可溶且稳定,不易分解的特性,这也与表2中显示的硫氧还蛋白突变体ΔΔG和Δ净电荷性能改变相呼应;同时也表明,可以通过一步离心的简单步骤或工艺将本专利提供的硫氧还蛋白突变体与培养液中菌体蛋白、细胞碎片和其他杂质蛋白进行有效分离。根据结果剔除上清中目标蛋白量显著较少的阳性转化子,其余加甘油保存于-80℃超低温冰箱。

[0081] 步骤5)硫氧还蛋白突变体/Arg34GLP-1(9-37)重组融合蛋白工程菌的构建

[0082] 选择硫氧还蛋白突变体优选文库复筛获得的硫氧还蛋白突变体mTrx39(氨基酸序列为SEQ ID NO.5)和连接肽GSGSG(氨基酸单字母缩写),肠激酶切割位点DDDDK(氨基酸单字母缩写)构建结构为:mTrx39-GSGSG-DDDDK-Arg34GLP-1(9-37)的重组融合蛋白(氨基酸序列为SEQ ID NO.6),质粒构建方法参阅例如中国专利201810858509.8,具体为将上述测序正确的单个阳性转化子即工程菌接种于含有硫酸卡那霉素(50g/mL)的LB液体培养基,37℃,220rpm培养过夜。将所得培养液作为种子,1%接种量接种于新鲜的含硫酸卡那霉素(50ug/mL)的LB液体培养基中,即1mL种子加入100mL新鲜培养基中,37℃培养4h,将温度调至30℃并加入IPTG(异丙基硫代半乳糖苷)至IPTG终浓度0.4mM/L,继续振荡培养20h。将所得培养液于室温4000rpm离心10min,收集菌体,多余菌体保存于-20℃。取一定体积的菌体,向其中加入一定体积的Sample loadingbuffer,100℃煮沸5min制样,获得重组表达载体pET28a-mTrx39/(R34)9-37。将上述构建正确的表达载体热激转化大肠杆菌BL21(DE3)中,涂布于含有50μg/mL硫酸卡那霉素的LB平板培养基。大肠杆菌感受态制备以及热激转化的方法参见《分子克隆实验指南》。待转化子长出,筛选若干送出测序,测序正确的转化子保存备用,测序正确的转化子即为重组工程菌。

[0083] 步骤6)硫氧还蛋白突变体/Arg34GLP-1(9-37)重组融合蛋白的生产

[0084] 将构建好的重组工程菌从平皿三接种于含有50μg/mL硫酸卡那霉素的LB液体摇瓶培养基中,37℃,220rpm培养约8-12h。将所得培养液作为1级种子,以1%接种量接种于新鲜的含50μg/mL硫酸卡那霉素的2级种子液体培养基中:磷酸氢二胺0.2%、磷酸二氢钾0.675%、一水柠檬酸0.093%、七水硫酸镁0.07%、葡萄糖2%、微量盐5ml/L(微量盐组成:七水硫酸亚铁1%、五水硫酸锰0.05%、七水硫酸锌0.225%、五水硫酸铜0.1%、十水四硼酸钠0.023%、二水氯化钙0.2%、七钼酸铵0.01%、盐酸5mol/L),34℃,220rpm培养约10-12h,将所得2级种子培养液,以2-5%接种量接种于30L发酵罐中,发酵培养基与2级种子培养基一致,发酵过程中根据菌体生长情况补加葡萄糖和七水硫酸镁,发酵pH通过添加浓氨水控制在pH 6.8-7.0。发酵前期温度37℃,发酵至6-8h后开始补料,发酵约10-14h时将温度调至30℃并加入乳糖至终浓度2%,继续发酵至22-26h。图3为SDS-PAGE检测的发酵过程中不同发酵取样点,重组融合蛋白(mTrx39-GSGSG-DDDDK-Arg34GLP-1(9-37))的表达结果。从该结果可知重组融合蛋白(mTrx39-GSGSG-DDDDK-Arg34GLP-1(9-37))获得了正常表达,且随着发酵时间延长至25h时均无分解现象发生。

[0085] 所得培养液加入等体积的缓冲液混匀后,于80℃水浴中孵育约15min,放出料液,采用碟片离心机～12000rpm离心收集上清,经澄清膜过滤后,用5-10kD超滤膜浓缩8-10倍,再用反应缓冲液置换后,将料液转入玻璃反应釜,加入重组肠激酶25℃酶切6-8h,调节pH至目标异源多肽等电点附近沉淀Arg34GLP-1(9-37),板框过滤后获得重组多肽Arg34GLP-1(9-37)粗品。图4是通过反相HPLC分析从重组融合蛋白mTrx39-GSGSG-DDDDK-Arg34GLP-1(9-37)酶切获得Arg34GLP-1(9-37)不同时间取样检测的结果,随着反应进程目标异源多肽Arg34GLP-1(9-37)不断生成,重组融合蛋白mTrx39-GSGSG-DDDDK-Arg34GLP-1(9-37)不断裂解成重组融合蛋白mTrx39-GSGSG-DDDDK。通过酶切反应,85%的重组融合蛋白mTrx39-GSGSG-DDDDK-Arg34GLP-1(9-37)被裂解,用等电点沉淀后获得纯度75%的Arg34GLP-1(9-37)粗品,如图5所示,通过一步沉淀的简单步骤或工艺即可获得较高纯度的目标异源多肽Arg34GLP-1(9-37)。分析条件:YMC-Pack PROTEIN-RP色谱柱,流动相A为20%乙腈+0.1%三

氟乙酸，流动相B为100％乙腈+0.1％三氟乙酸，流速为1ml/min，按如下梯度洗脱：0-8min流动相B由10％升至20％，8-25min流动相B由20％升至55％，检测波长214nm，柱温35℃。

[0086] 本发明不局限于上述可选实施方式，任何人在本发明的启示下都可得出其他各种形式的产品，但不论在其形状或结构上作任何变化，凡是落入本发明权利要求界定范围内的技术方案，均落在本发明的保护范围之内。

[0001] 序列表

[0002] <110> 成都英普博集生物科技有限公司

[0003] <120> 硫氧还蛋白突变体、其制备方法及其在重组融合蛋白生产中的应用

[0004] <160> 6

[0005] <170> SIPOSequenceListing 1.0

[0006] <210> 1

[0007] <211> 105

[0008] <212> PRT

[0009] <213> 嗜热脂肪地芽胞杆菌(Geobacillus stearothermophilus)

[0010] <400> 1

[0011] Met Ala Ile Val Asn Ala Thr Asp Gln Thr Phe Ala Ala Glu Thr Lys

[0012] 1 5 10 15

[0013] Asp Gly Leu Thr Leu Val Asp Phe Trp Ala Pro Trp Cys Gly Pro Cys

[0014] 20 25 30

[0015] Arg Met Ile Ala Pro Val Leu Glu Glu Leu Asp Arg Glu Met Gly Asp

[0016] 35 40 45

[0017] Lys Val Lys Ile Val Lys Val Asn Val Asp Glu Asn Gln Glu Thr Ala

[0018] 50 55 60

[0019] Ser Lys Phe Gly Val Met Ser Ile Pro Thr Leu Leu Val Phe Lys Asn

[0020] 65 70 75 80

[0021] Gly Glu Leu Val Asp Lys Ala Val Gly Tyr Gln Pro Lys Glu Ala Leu

[0022] 85 90 95

[0023] Val Gln Leu Val Gly Lys His Val Ser

[0024] 100 105

[0025] <210> 2

[0026] <211> 105

[0027] <212> PRT

[0028] <213> 芽孢杆菌(Bacillaceae)

[0029] <400> 2

[0030] Met Ala Ile Val Asn Ala Thr Asp Gln Thr Phe Ala Ala Glu Thr Lys

[0031] 1 5 10 15

[0032] Asp Gly Leu Thr Leu Val Asp Phe Trp Ala Pro Trp Cys Gly Pro Cys

[0033] 20 25 30

[0034] Arg Met Ile Ala Pro Val Leu Glu Glu Leu Asp Arg Glu Met Gly Asp

[0035] 35 40 45

[0036] Lys Val Lys Ile Val Lys Val Asn Val Asp Glu Asn Gln Glu Thr Ala

[0037] 50 55 60

[0038] Ser Lys Phe Gly Val Met Ser Ile Pro Thr Leu Leu Val Phe Lys Asn

[0039] 65 70 75 80

[0040] Gly Glu Leu Val Asp Lys Ala Ile Gly Tyr Gln Pro Lys Glu Ala Leu

[0041] 85 90 95

[0042] Val Gln Leu Val Gly Lys His Val Ser

[0043] 100 105

[0044] <210> 3

[0045] <211> 104

[0046] <212> PRT

[0047] <213> 喜温无氧芽孢杆菌(Anoxybacillus tepidamans)

[0048] <400> 3

[0049] Met Ala Ile Val Asn Ala Thr Asp Gln Thr Phe Val Thr Glu Thr Ser

[0050] 1 5 10 15

[0051] Thr Gly Val Thr Leu Val Asp Phe Trp Ala Pro Trp Cys Gly Pro Cys

[0052] 20 25 30

[0053] Arg Met Ile Ala Pro Val Leu Glu Glu Val Asp Gln Glu Met Gly Asp

[0054] 35 40 45

[0055] Lys Val Lys Ile Val Lys Val Asn Val Asp Glu Asn Gln Glu Thr Ala

[0056] 50 55 60

[0057] Ser Lys Tyr Gly Val Met Ser Ile Pro Thr Leu Leu Val Phe Lys Asp

[0058] 65 70 75 80

[0059] Gly Asn Val Val Asp Lys Thr Val Gly Phe Gln Pro Lys Glu Ala Leu

[0060] 85 90 95

[0061] Val Gln Leu Leu Gln Gln His Val

[0062] 100

[0063] <210> 4

[0064] <211> 104

[0065] <212> PRT

[0066] <213> 白色假芽孢杆菌(Falsibacillus pallidus)

[0067] <400> 4

[0068] Met Ala Ile Thr Asn Ala Thr Asp Gln Asn Phe Gly Ala Glu Thr Ser

[0069] 1 5 10 15

[0070] Gln Gly Leu Val Leu Ala Asp Phe Trp Ala Pro Trp Cys Gly Pro Cys

[0071] 20 25 30

[0072] Lys Met Ile Ala Pro Val Leu Glu Glu Leu Asp Ser Glu Met Gly Asp

[0073] 35 40 45

[0074] Lys Val Lys Ile Val Lys Val Asp Val Asp Glu Asn Gln Glu Thr Ala

[0075] 50 55 60

[0076] Ser Ser Phe Gly Val Met Ser Ile Pro Thr Leu Ile Val Leu Lys Asp

[0077] 65 70 75 80

[0078] Gly Glu Val Val Asp Lys Val Ile Gly Phe Gln Pro Lys Glu Ala Leu
[0079] 85 90 95
[0080] Ala Glu Leu Leu Asn Lys His Ala
[0081] 100
[0082] <210> 5
[0083] <211> 104
[0084] <212> PRT
[0085] <213> 人工序列(Artificial Sequence)
[0086] <400> 5
[0087] Met Ala Ile Val Asn Ala Thr Asp Gln Thr Phe Ile Ala Glu Thr Lys
[0088] 1 5 10 15
[0089] Asp Gly Leu Thr Leu Val Asp Phe Trp Ala Pro Trp Ser Gly Pro Cys
[0090] 20 25 30
[0091] Arg Met Ile Ala Pro Val Leu Glu Glu Leu Asp Arg Glu Met Gly Asp
[0092] 35 40 45
[0093] Lys Val Lys Ile Val Lys Val Lys Val Asp Glu Asn Gln Glu Thr Ala
[0094] 50 55 60
[0095] Ser Lys Phe Gly Val Met Met Ile Pro Thr Leu Leu Val Phe Lys Asn
[0096] 65 70 75 80
[0097] Gly Glu Leu Val Asp Lys Ala Val Gly Tyr Gln Pro Lys Glu Ala Leu
[0098] 85 90 95
[0099] Val Gln Leu Val Gly His His Val
[0100] 100
[0101] <210> 6
[0102] <211> 143
[0103] <212> PRT
[0104] <213> 人工序列(Artificial Sequence)
[0105] <400> 6
[0106] Met Ala Ile Val Asn Ala Thr Asp Gln Thr Phe Ile Ala Glu Thr Lys
[0107] 1 5 10 15
[0108] Asp Gly Leu Thr Leu Val Asp Phe Trp Ala Pro Trp Ser Gly Pro Cys
[0109] 20 25 30
[0110] Arg Met Ile Ala Pro Val Leu Glu Glu Leu Asp Arg Glu Met Gly Asp
[0111] 35 40 45
[0112] Lys Val Lys Ile Val Lys Val Lys Val Asp Glu Asn Gln Glu Thr Ala
[0113] 50 55 60
[0114] Ser Lys Phe Gly Val Met Met Ile Pro Thr Leu Leu Val Phe Lys Asn
[0115] 65 70 75 80
[0116] Gly Glu Leu Val Asp Lys Ala Val Gly Tyr Gln Pro Lys Glu Ala Leu

[0117] 85 90 95

[0118] Val Gln Leu Val Gly His His Val Gly Ser Gly Ser Gly Asp Asp Asp

[0119] 100 105 110

[0120] Asp Lys Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly

[0121] 115 120 125

[0122] Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Arg Gly Arg Gly

[0123] 130 135 140

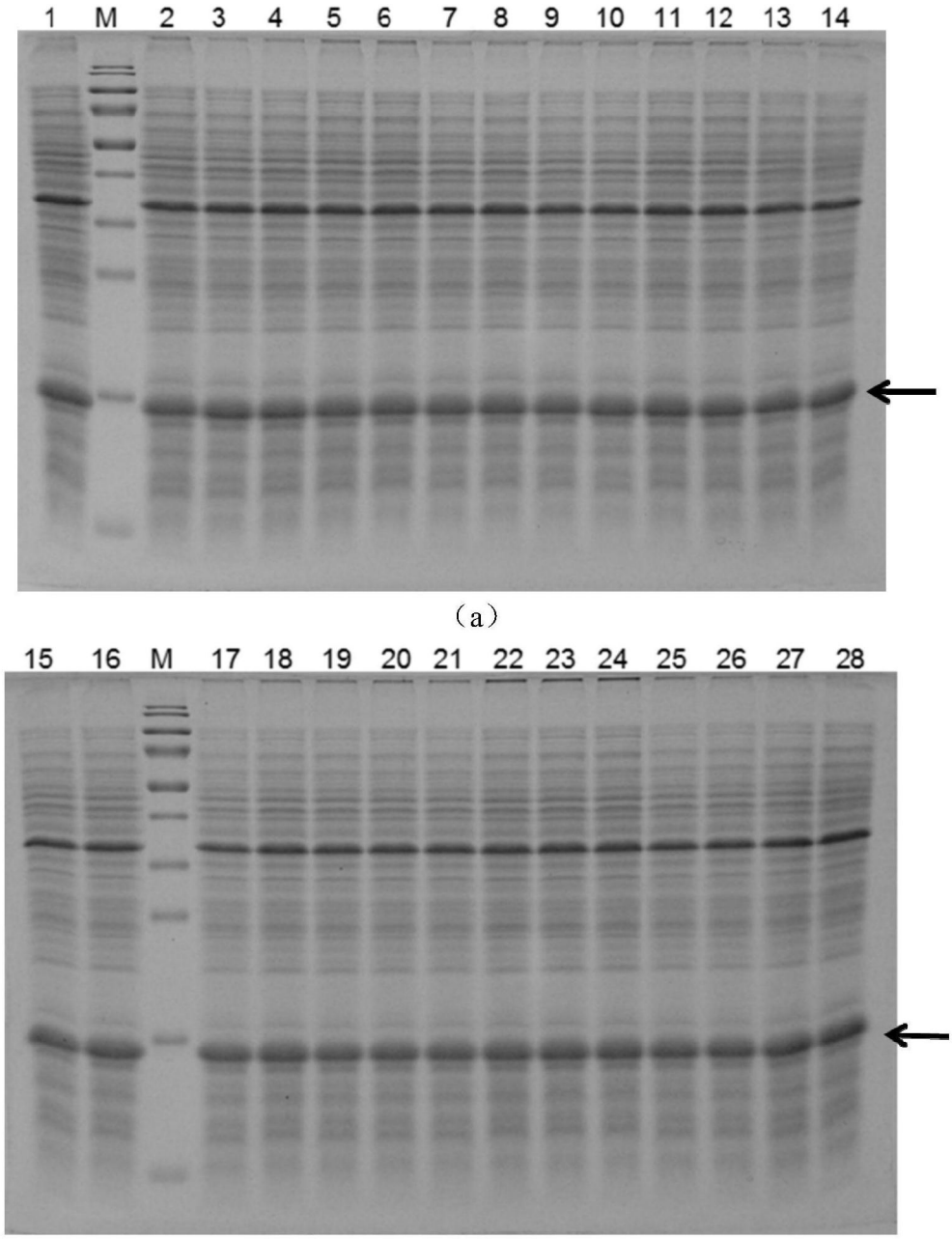

1 M 2 3 4 5 6 7 8 9 10 11 12 13 14
(a)
15 16 M 17 18 19 20 21 22 23 24 25 26 27 28
(b)

图1

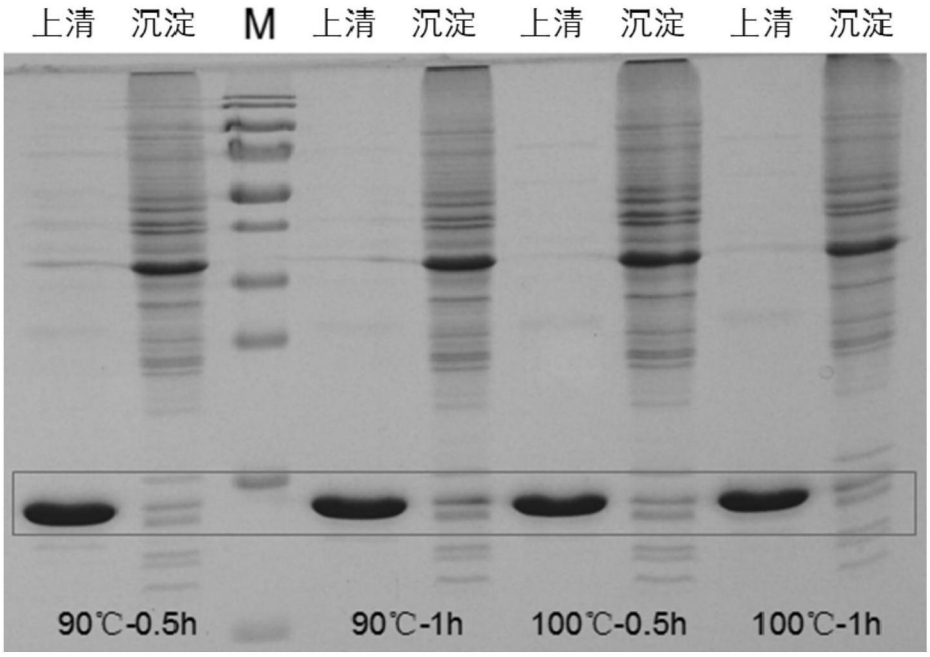

上清
沉淀
M
上清
沉淀
上清
沉淀
上清
沉淀
90℃-0.5h
90℃-1h
100℃-0.5h
100℃-1h

图2

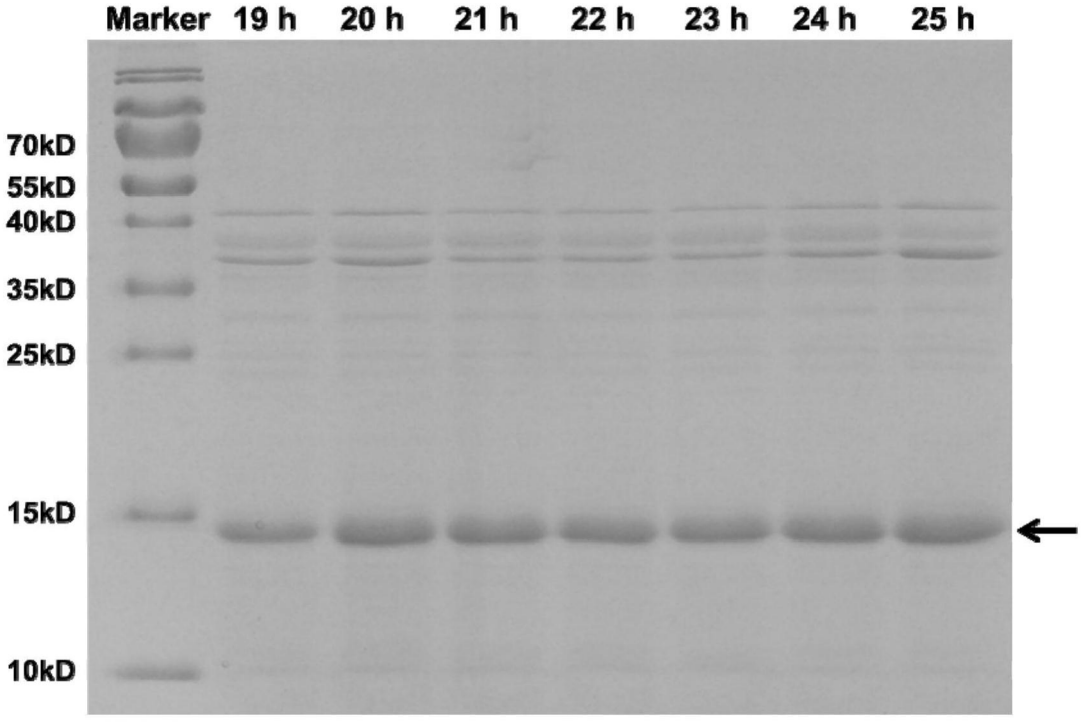

Marker
19 h
20 h
21 h
22 h
23 h
24 h
25 h
70kD
55kD
40kD
35kD
25kD
15kD
10kD

图3

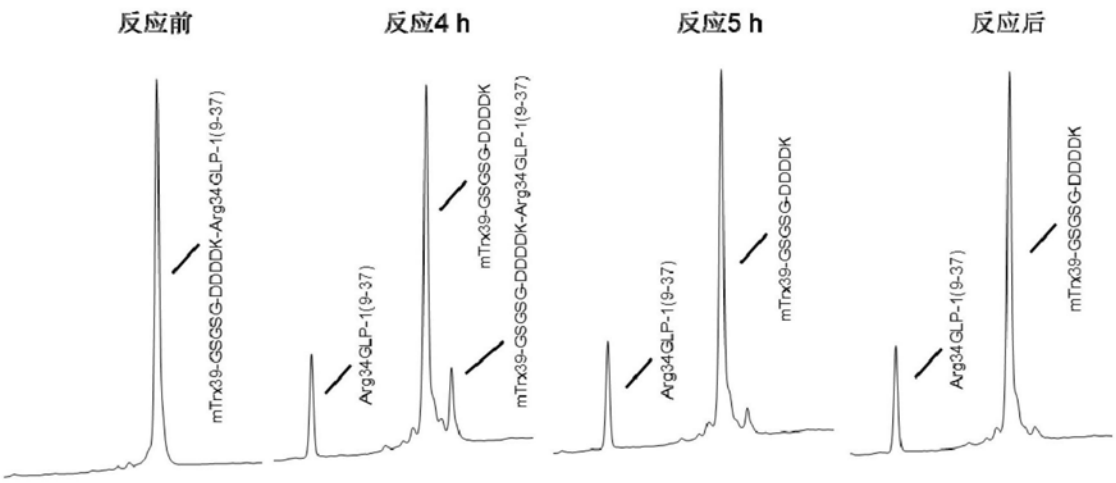

反应前
反应4 h
反应5 h
反应后
mTrx39-GSGSG-DDDDK-Arg34GLP-1(9-37)
Arg34GLP-1(9-37)
mTrx39-GSGSG-DDDDK
mTrx39-GSGSG-DDDDK-Arg34GLP-1(9-37)
Arg34GLP-1(9-37)
mTrx39-GSGSG-DDDDK
Arg34GLP-1(9-37)
mTrx39-GSGSG-DDDDK

图4

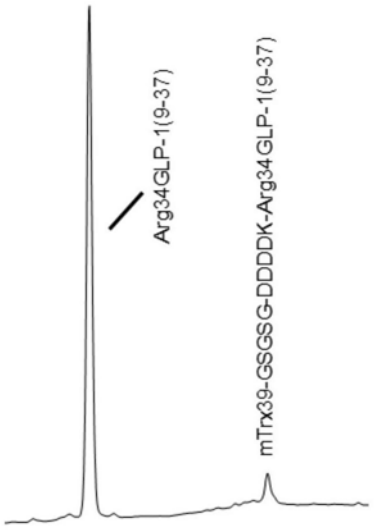

Arg34GLP-1(9-37)
mTrx39-GSGSG-DDDDK-Arg34GLP-1(9-37)

图5