



(12) 发明专利申请

(10) 申请公布号 CN 113980141 A

(43) 申请公布日 2022. 01. 28

(21) 申请号 202111256498.4

G12P 19/02 (2006.01)

(22) 申请日 2021.10.27

G12P 19/14 (2006.01)

(71) 申请人 湖北大学

地址 430062 湖北省武汉市武昌区友谊大道368号

(72) 发明人 马立新 翟超 杨军 王飞

(74) 专利代理机构 武汉智嘉联合知识产权代理
事务所(普通合伙) 42231

代理人 姜婷

(51) Int. Cl.

G07K 19/00 (2006.01)

G12N 9/22 (2006.01)

G12N 15/70 (2006.01)

G12N 9/96 (2006.01)

G12N 9/42 (2006.01)

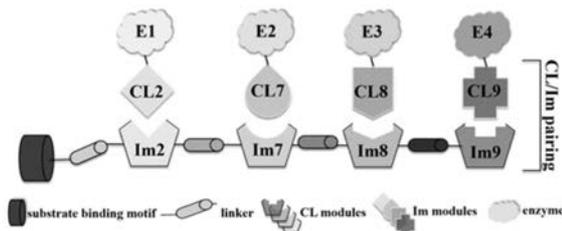
权利要求书2页 说明书8页
序列表7页 附图4页

(54) 发明名称

基于大肠杆菌素E家族DNA酶的蛋白质复合物及其在人工蛋白支架中的应用

(57) 摘要

本发明公开了基于大肠杆菌素E家族DNA酶的蛋白质复合物及其在人工蛋白支架中的应用,所述蛋白质复合物中包括:CL2蛋白与Im2蛋白形成的相互作用对,或CL7蛋白与Im7蛋白形成的相互作用对,或CL8蛋白与Im8蛋白形成的相互作用对,或CL9蛋白与Im9蛋白形成的相互作用对中的任意一种或多种。本发明通过对CE家族的DNA酶CE2、CE7、CE8和CE9的羧基端DNase结构域进行蛋白质工程改造,获得丧失DNA酶活性,但仍保留有与对应的免疫蛋白Im具有超高亲和力的突变体,并构建得到了CL2/Im2,CL7/Im7,CL8/Im8和CL9/Im9蛋白质相互作用对。研究发现这些蛋白质相互作用对具有耐热性,高亲和力,高特异性,分子量小,组装速度快等性质,基于此构建了人工蛋白质支架,为人工多酶复合物的构建搭建了新型平台。



1. 一种基于大肠杆菌素E家族DNA酶的蛋白质复合物,其特征在于,所述蛋白质复合物中包括:CL2蛋白与Im2蛋白形成的相互作用对,或CL7蛋白与Im7蛋白形成的相互作用对,或CL8蛋白与Im8蛋白形成的相互作用对,或CL9蛋白与Im9蛋白形成的相互作用对中的任意一种或多种;其中,

所述CL2蛋白的氨基酸序列如SEQ ID NO.2所示;

所述CL7蛋白的氨基酸序列如SEQ ID NO.5所示;

所述CL8蛋白的氨基酸序列如SEQ ID NO.8所示的;

所述CL9蛋白的氨基酸序列如SEQ ID NO.11所示。

2. 根据权利要求1所述的蛋白质复合物,其特征在于,

所述Im2蛋白的氨基酸序列如SEQ ID NO.3所示;

所述Im7蛋白的氨基酸序列如SEQ ID NO.6所示;

所述Im8蛋白的氨基酸序列如SEQ ID NO.9所示;

所述Im9蛋白的氨基酸序列如SEQ ID NO.12所示。

3. 根据权利要求1所述的蛋白质复合物,其特征在于,

所述CL2蛋白是在如SEQ ID NO.1所示的CE2蛋白基础上突变得到的;

所述CL7蛋白是在如SEQ ID NO.4所示的CE7蛋白基础上突变得到的;

所述CL8蛋白是在如SEQ ID NO.7所示的CE7蛋白基础上突变得到的;

所述CL9蛋白是在如SEQ ID NO.10所示的CE7蛋白基础上突变得到的。

4. 如权利要求1-3任一项所述的蛋白质复合物在制备人工蛋白质支架中的应用。

5. 一种基于权利要求1-3任一项所述的蛋白质复合物的支架蛋白,其特征在于,所述支架蛋白中包括:通过linker依次串联的纤维素结合模块CBM3a,Im2蛋白,Im7蛋白,Im8蛋白以及Im9蛋白;

所述支架蛋白上的各Im蛋白模块可与相应的CL蛋白连接形成相互作用对,通过将目的蛋白与CL蛋白连接,实现目的蛋白在支架蛋白上的有序装配。

6. 根据权利要求5所述的支架蛋白,其特征在于,所述纤维素结合模块CBM3a来源于热纤梭菌*Clostridium thermocellum*。

7. 如权利要求5所述的支架蛋白的构建方法,其特征在于,所述方法包括:

步骤一、采用PCR扩增的方法分别得到CBM3a,Im2蛋白,Im7蛋白,Im8蛋白以及Im9蛋白的基因片段;

步骤二、采用Overlapping PCR技术将步骤一得到的5段基因片段依次连接;

步骤三、将步骤二连接好的基因进行重组载体构建,并导入受体细胞,诱导表达,得到所述支架蛋白。

8. 如权利要求5所述的支架蛋白在制备多酶复合物体系中的应用。

9. 一种基于权利要求5所述的支架蛋白的多酶复合物体系,其特征在于,所述多酶复合物体系中包括:4个CL-tagged纤维素酶以及如权利要求5所述的支架蛋白,4个CL-tagged纤维素酶与支架蛋白上对应的Im蛋白模块连接形成相互作用对,即得到所述多酶复合物;

其中所述4个CL-tagged纤维素酶包括:与CL2蛋白连接的外切葡聚糖酶,与CL7蛋白连接的内切葡聚糖酶突变体,与CL8蛋白连接的内切葡聚糖酶,以及与CL9蛋白连接的 β -葡萄糖苷酶。

10. 如权利要求9所述的多酶复合物体系在水解纤维素底物中的应用。

基于大肠杆菌素E家族DNA酶的蛋白质复合物及其在人工蛋白 支架中的应用

技术领域

[0001] 本发明属于合成生物学及纳米生物技术领域,具体涉及基于大肠杆菌素E家族DNA酶的蛋白质复合物及其在人工蛋白支架中的应用。

背景技术

[0002] 高亲和力的蛋白质相互作用对在许多领域都有广泛的应用,如免疫共沉淀、蛋白质pull-down、酵母双杂交和酶的固定化等,其中利用蛋白质相互作用对自组装形成的多酶复合物体系因具备出色的协同催化能力而备受关注。对于“一锅法”合成反应而言,利用多个有着正交性的蛋白质相互作用对实现多个酶分子按照特定的顺序进行组装,以人工多酶复合物的形式进行级联催化反应,有利于底物通道的形成,可以防止中间产物的扩散,促进中间产物及时转化,从而避免某些有毒中间体对酶活性的干扰,提高反应效率。目前,蛋白质-蛋白质的相互作用主要通过来源于纤维小体的基本组装元件Cohesin-Dockerin以及人工设计的SpyCatcher/SpyTag、SnoopCatcher/SnoopTag和RIAD/RIDD等蛋白质相互作用对实现。然而现有的蛋白质相互作用对在实际应用时存在诸多问题。首先,强相互作用对的数量有限,目前常用的相互作用对的亲和力都不强,在构建较为复杂的多酶复合物时,结构不稳定,容易坍塌。其次,对于某些放热反应,如纤维素的降解,需要热稳定的蛋白质相互作用对结合热稳定酶来构建多酶复合物,目前此类蛋白质相互作用对更是少之又少。此外,大多数蛋白质相互作用对都需要在高盐离子条件下发生相互作用,如Cohesin-Dockerin的作用受钙离子调控,限制了它们在体内的应用。因此,针对以上问题挖掘具有优良性质的新蛋白质相互作用对具有重要意义。

[0003] 大肠杆菌产生的大肠杆菌素是一种细胞毒素,可以杀死亲缘关系比较近的杆菌。大肠杆菌素E家族的DNA酶包括CE2、CE7、CE8和CE9,是一类结构高度相似的非专一性核酸内切酶。在宿主细胞中,CE蛋白与各自对应的免疫蛋白Im2、Im7、Im8和Im9结合,抑制CE蛋白与DNA的结合,从而消除它们对于宿主的毒性。CE蛋白和Im蛋白的结合是目前自然界已知最强的蛋白质相互作用之一, K_d 达到 $10^{-14} \sim 10^{-17}$ 。Im与CE的结合为“双重识别机制”,即Im蛋白质首先通过保守元件(螺旋III)与CE核酸酶形成序列无关的复合物;随后Im蛋白质上含有特异性残基的螺旋II与对应的CE核酸酶结合,稳定正交复合物(Li W., et al, Biochemistry.1998,37,11771-11779;Keeble AH,Kleanthous C.,J Mol Biol.2005,352,656-671)。

发明内容

[0004] 本发明的目的在于提供基于大肠杆菌素E家族DNA酶的蛋白质复合物及其在人工蛋白支架中的应用,本发明通过对CE家族的DNA酶CE2、CE7、CE8和CE9的羧基端DNase结构域进行蛋白质工程改造,获得丧失DNA酶活性,但仍保留与对应Im蛋白的超高亲和力的突变体,并构建得到了CL2/Im2,CL7/Im7,CL8/Im8和CL9/Im9蛋白质相互作用对。研究发现这些

蛋白质相互作用对具有耐热性,高亲和力,高特异性,分子量小,组装速度快等性质,基于此构建了人工蛋白质支架,为人工多酶复合物的构建搭建了新型平台。

[0005] 本发明的目的之一在于提供一种基于大肠杆菌素E家族DNA酶的蛋白质复合物,所述蛋白质复合物中包括:CL2蛋白与Im2蛋白形成的相互作用对,或CL7蛋白与Im7蛋白形成的相互作用对,或CL8蛋白与Im8蛋白形成的相互作用对,或CL9蛋白与Im9蛋白形成的相互作用对中的任意一种或多种;其中,

[0006] 所述CL2蛋白的氨基酸序列如SEQ ID NO.2所示;

[0007] 所述CL7蛋白的氨基酸序列如SEQ ID NO.5所示;

[0008] 所述CL8蛋白的氨基酸序列如SEQ ID NO.8所示的;

[0009] 所述CL9蛋白的氨基酸序列如SEQ ID NO.11所示。

[0010] 进一步地,

[0011] 所述Im2蛋白的氨基酸序列如SEQ ID NO.3所示;

[0012] 所述Im7蛋白的氨基酸序列如SEQ ID NO.6所示;

[0013] 所述Im8蛋白的氨基酸序列如SEQ ID NO.9所示;

[0014] 所述Im9蛋白的氨基酸序列如SEQ ID NO.12所示。

[0015] 进一步地,

[0016] 所述CL2蛋白是在如SEQ ID NO.1所示的CE2蛋白基础上突变得到的;

[0017] 所述CL7蛋白是在如SEQ ID NO.4所示的CE7蛋白基础上突变得到的;

[0018] 所述CL8蛋白是在如SEQ ID NO.7所示的CE7蛋白基础上突变得到的;

[0019] 所述CL9蛋白是在如SEQ ID NO.10所示的CE7蛋白基础上突变得到的。

[0020] 本发明的目的之二在于提供了所述蛋白质复合物在制备人工蛋白质支架中的应用。

[0021] 本发明的目的之三在于提供了一种基于所述蛋白质复合物的支架蛋白,所述支架蛋白中包括:通过linker依次串联的纤维素结合模块CBM3a,Im2蛋白,Im7蛋白,Im8蛋白以及Im9蛋白;

[0022] 所述支架蛋白上的各Im蛋白模块可与相应的CL蛋白连接形成相互作用对,通过将目的蛋白与CL蛋白连接,实现目的蛋白在支架蛋白上的有序装配。

[0023] 进一步地,所述纤维素结合模块CBM3a来源于热纤梭菌*Clostridium thermocellum*。

[0024] 进一步地,所述linker包括:CBM3a与Im2蛋白之间通过Linker1(TTTTTPTTASSTTSTTSATSG)连接,Im2蛋白与Im7蛋白之间通过Linker2(NTLTIGGGNPGGGNPGGGTNP GT)连接,Im7蛋白和Im8蛋白之间通过Linker3(GASGSGTTTTTTTTSTTTGGTDP)连接,Im8蛋白和Im9蛋白之间通过Linker4(SSWEGISTDDGIVHSVRISGNTLS)连接。

[0025] 本发明的目的之四在于提供了所述支架蛋白的构建方法,包括:

[0026] 步骤一、采用PCR扩增的方法分别得到CBM3a,Im2蛋白,Im7蛋白,Im8蛋白以及Im9蛋白的基因片段;

[0027] 步骤二、采用Overlapping PCR技术将步骤一得到的5段基因片段依次连接;

[0028] 步骤三、将步骤二连接好的基因进行重组载体构建,并导入受体细胞,诱导表达,

得到所述支架蛋白。

[0029] 本发明的目的之五在于提供了所述支架蛋白在制备多酶复合物体系中的应用。

[0030] 本发明的目的之六在于提供了一种基于所述支架蛋白的多酶复合物体系,所述多酶复合物体系中包括:4个CL-tagged纤维素酶以及如权利要求5所述的支架蛋白,4个CL-tagged纤维素酶与支架蛋白上对应的Im蛋白模块连接形成相互作用对,即得到所述多酶复合物;

[0031] 其中所述4个CL-tagged纤维素酶包括:与CL2蛋白连接的外切葡聚糖酶,与CL7蛋白连接的内切葡聚糖酶突变体,与CL8蛋白连接的内切葡聚糖酶,以及与CL9蛋白连接的 β -葡萄糖苷酶。

[0032] 进一步地,所述外切葡聚糖酶为通过DNA shuffling得到的外切葡聚糖酶Cel48S_{m3},所述内切葡聚糖酶突变体为来源于Clostridium thermocellum的内切葡聚糖酶突变体Cel18A*(K276R,G283P,S329G,S375T),所述与CL8蛋白连接的内切葡聚糖酶为来源于Caldicellulosiruptor bescii的内切葡聚糖酶GH5D,所述 β -葡萄糖苷酶为来源于Caldicellulosiruptor owensensis的 β -葡萄糖苷酶CoGH1A

[0033] 本发明的目的之七在于提供了所述多酶复合物体系在水解纤维素底物中的应用。

[0034] 与现有技术相比,本发明的有益效果是:本发明提供了一种基于大肠杆菌素E家族DNA酶的蛋白质复合物及其在人工蛋白支架中的应用,本发明通过对CE家族的DNA酶CE2、CE7、CE8和CE9的羧基端DNase结构域进行蛋白质工程改造,获得丧失DNA酶活性,但仍保留有与对应的免疫蛋白Im蛋白具有超高亲和力的突变体CL2、CL7、CL8和CL9,并构建得到了CL2/Im2,CL7/Im7,CL8/Im8和CL9/Im9蛋白质相互作用对。研究发现上述CL/Im蛋白质相互作用对工具箱具有高特异性、超高亲和力和高热稳定性、结构统一等特性,因此基于该CL/Im工具箱可构建稳定、耐热蛋白质支架,并能够按照设定的顺序进行多酶组装的平台,为构建具备高鲁棒性的多酶分子机器提供了基础。本发明中将4种纤维素水解酶组装到该人工蛋白质支架上形成多酶复合物,在高温环境下更高效水解纤维素底物。同时,在任何涉及到蛋白质-蛋白质相互作用的领域都能够使用以上蛋白质相互作用对。

附图说明

[0035] 图1为本发明实施例1中4种CL与Im蛋白纯化后的SDS-PAGE检测图;

[0036] 图2为本发明实施例2中采用Native-PAGE检测CL和Im蛋白质的相互作用(2-B)以及CL和Im蛋白质之间相互作用的耐热性(2-C);

[0037] 图3为本发明实施例3中支架蛋白组装多酶复合物的原理示意图;

[0038] 图4为本发明实施例3中含有1-4个Im模块的支架蛋白纯化后的SDS-PAGE检测结果图;

[0039] 图5为本发明实施例3中支架蛋白Scaf-CIQ的热稳定性检测结果图;

[0040] 图6为本发明实施例4中装配有4种纤维素水解酶的多酶复合物MEC-CIQ的组装情况检测结果图;

[0041] 图7为本发明实施例4中人工多酶复合物MEC-CIQ中4种纤维素水解酶的协同作用检测结果图。

具体实施方式

[0042] 下面将结合本发明中的实施例,对本发明的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本发明一部分实施例,而不是全部的实施例。基于本发明中的实施例,本领域普通技术人员在没有做出创造性劳动条件下所获得的所有其它实施例,都属于本发明保护的范围。

[0043] 实施例1 CL蛋白与Im蛋白的表达与纯化

[0044] 1、通过与CL7蛋白质进行氨基酸序列比对,基于4种蛋白质序列的高度相似性,将CE家族的DNA酶CE2、CE7、CE8和CE9的羧基端DNase结构域进行蛋白质工程改造,获得丧失DNA酶活性,但仍保留有与对应的免疫蛋白Im蛋白具有超高亲和力的突变体CL2、CL7、CL8和CL9,其中四组CL-Im及4种野生型CE蛋白质序列为:

[0045] (1) CE2蛋白:如SEQ ID NO.1所示;CL2蛋白:如SEQ ID NO.2所示;Im2蛋白:如SEQ ID NO.3所示;

[0046] (2) CE7蛋白:如SEQ ID NO.4所示;CL7蛋白:如SEQ ID NO.5所示;Im7蛋白:如SEQ ID NO.6所示;

[0047] (3) CE8蛋白:如SEQ ID NO.7所示;CL8蛋白:如SEQ ID NO.8所示;Im8蛋白:如SEQ ID NO.9所示;

[0048] (4) CE9蛋白:如SEQ ID NO.10所示;CL9蛋白:如SEQ ID NO.11所示;Im9蛋白:如SEQ ID NO.12所示。

[0049] 2、各突变体及Im蛋白基因交由武汉金开瑞公司进行基因合成,将4种CL和Im蛋白的编码基因分别采用本领域常规方法构建至pET23a载体中。重组质粒转入大肠杆菌BL21(DE3),诱导表达并纯化,具体为:表达菌株接种到添加100 μ g/mL氨苄青霉素的LB液体培养基中,37 $^{\circ}$ C培养过夜;其后将种子液按1:100的比例接种到TB培养基中,待OD₆₀₀值达到0.8~1.0时,加入终浓度为1mM的IPTG,18 $^{\circ}$ C,220rpm诱导16~18h;收集菌体,用TBS buffer (pH 7.4)充分重悬菌体,高压细胞破碎仪重复破碎5遍,随后用10000rpm,30min离心,将上清转移到预处理好的重力柱中。

[0050] 预处理方法如下:吸取4mL镍珠至60mL蛋白纯化重力柱,用3倍体积破菌buffer平衡镍珠。将上清转移至重力柱后,封口膜封口,于静音混合器上4 $^{\circ}$ C孵育1h,然后用3倍柱体积的TBS(含10mM咪唑)洗涤,重复3次。随后依次用含30mM、50mM、200mM的TBS洗涤各3次,每次3倍柱体积,分别收集流出液。将各梯度咪唑洗脱后的收集液用SDS-PAGE检测。结果如图1所示,结果显示,所有CL与Im蛋白质均能成功表达并纯化。重组蛋白样品经超滤换液并浓缩后分装,于-80 $^{\circ}$ C冷冻保存。

[0051] 实施例2 Native-PAGE

[0052] 1、Native-PAGE非变性聚丙烯酰胺凝胶制备

[0053] 对于CL-Im蛋白质相互作用对的结合特异性验证,采用30%聚丙烯酰胺(37.5:1)、甘油电泳体系,具体配方如表1所示。

[0054] 表1 Native-PAGE非变性聚丙烯酰胺凝胶配方

	分离胶		浓缩胶	
	组分	体积 (mL)	组分	体积 (mL)
	37.5:1 丙烯酰胺	1.6	37.5:1 丙烯酰胺	0.83
	1.5 M Tris-HCl pH8.8	1	1.M Tris-HCl pH6.8	0.63
[0055]	60% 甘油	0.587	60% 甘油	-
	10% APS	0.015	10% APS	0.050
	TEMED	0.004	TEMED	0.005
	ddH ₂ O	0.794	ddH ₂ O	3.45
	总体积	4	总体积	5

[0056] 2、Native-PAGE

[0057] 为验证CL蛋白与Im蛋白结合的特异性,本实施例将4种CL蛋白质和4种Im蛋白质分别混合,得到16种组合,室温孵育30min后进行Native-PAGE,上层胶时电压为120V,下层胶电压为180V,当溴酚蓝指示带泳动到胶底部时结束电泳,然后进行染色。电泳检测结果如图2-B所示,电泳结果表明CL与Im之间的相互作用具有良好的特异性。

[0058] 进一步验证CL与Im蛋白质结合的热稳定性,具体为:将16种组合的CL-Im蛋白质混合体系,室温孵育30min后,将其分别在100℃水浴加热1h,对高温处理后的蛋白质混合样品进行Native-PAGE,检测结果如图2-C所示。结果表明Im和CL蛋白在高温处理后发生微弱结构变化,导致活性胶蛋白质条带模糊,但彼此正交的4组Im-CL形成的复合物在高温环境下均没有发生明显变化,由此可知Im与CL的相互作用具有良好的热稳定性。

[0059] 实施例3支架蛋白Scaf-CIQ的构建及热稳定性验证

[0060] 1、本实施例基于上述CL与Im蛋白质结合可形成相互作用对,构建相应的支架蛋白,具体如下:

[0061] CIQ系统中的支架蛋白Scaf-CIQ为来源于Clostridium thermocellum(生存环境为60~90℃)的纤维素结合模块CBM3a,通过linker与Im2、Im7、Im8和Im9串联表达出的融合蛋白,即为支架蛋白Scaf-CIQ。CBM3a的存在赋予支架蛋白Scaf-CIQ与微晶纤维素(Avicel)、磷酸膨胀纤维素(PASC)等纤维素基质结合的能力,达到靶向纤维素基质和固定化酶的目的。所述支架蛋白Scaf-CIQ的原理如图3所示,支架蛋白Scaf-CIQ上的各Im模块只能与包含相应CL-tag的目的蛋白如酶进行结合,从而实现4种目的蛋白在支架蛋白Scaf-CIQ上的有序装配,组装而成的复合物被命名为MEC-CIQ(multienzyme complexes-CL-Im quartet)。

[0062] 所述支架蛋白构建过程如下:

[0063] 通过PCR扩增出四种Im蛋白及CBM3a的对应基因片段,采用overlapping PCR技术将其整合,得到含有1-4个Im模块的基因序列,并进行重组载体构建,具体为:所述支架蛋白选用的载体骨架为经过NdeI和XhoI双酶切的pET23a载体,设计引物时,CBM3a基因的正向引物含有15个与pET23a载体NdeI位点侧翼序列同源的核苷酸,根据Im模块数量的不同,Im2、Im7、Im8和Im9编码序列分别带有与pET23a载体XhoI位点侧翼15个碱基同源的重组序列。

[0064] 其中CBM3a与Im2蛋白之间通过Linker1(氨基酸序列:TTTTTPTASSTTSTTSATSG,

Inoue H, et al. *Biotechnol Biofuels*, 2015, 8, 77.) 连接, Im2蛋白与Im7蛋白之间通过 Linker2 (氨基酸序列: NTLTIGGGNPGGGNPGGGTNP GT, Liu Y, et al. *Ind Microbiol Biotechnol*, 2015, 42, 1591-1599) 连接, Im7蛋白和Im8蛋白之间通过 Linker3 (氨基酸序列: GASGSGTTTTTTTTSTTTGGTDP, Miao Y, et al. *AMB Express*, 2018, 8, 44) 连接, Im8蛋白和Im9蛋白之间通过 Linker4 (氨基酸序列: SSWEGISTDDGIVHSVRISGNK TLS, Sakuraba H, et al. *Biochim Biophys Acta*, 2008, 1784, 563-571) 连接。

[0065] PCR扩增出含不同Linker序列的四种Im蛋白质及CBM3a的基因片段, 将CBM3a、Im2 (Im2+Im7、Im2+Im7+Im8、Im2+Im7+Im8+Im9) 和线性化的pET23a线性载体按摩尔比3:1混合, 5 μ L体系中加入0.5U的T5核酸外切酶及对应的NEB Buffer4, 加入大肠杆菌DH5 α 感受态后, 冰水混合浴上静置5min, 基因的同源臂区域将产生3'端突出的黏性末端, 基因片段转入感受态细胞后, 在大肠杆菌体内经过同源重组, 将得到含有1-4个Im模块与CBM3a完整基因序列的环状载体, 其中仅包含一个Im2模块的支架蛋白命名为Scaf-CIM, 同时包含Im2和Im7两个模块的支架蛋白命名为Scaf-CIB, 同时包含Im2、Im7和Im8三个模块的支架蛋白命名为Scaf-CIT, 同时包含Im2、Im7、Im8和Im9四个模块的支架蛋白命名为Scaf-CIQ。

[0066] 重组载体转入大肠杆菌BL21 (DE3) 进行诱导表达, 继而通过Ni-NTA方法进行纯化, 并对纯化蛋白进行检测, 检测结果如图4所示。SDS-PAGE结果表明, 含CBM3a和1-4个Im模块的支架蛋白均能通过Ni-NTA纯化方法得到较高纯度的蛋白质。

[0067] 2、为探究所述支架蛋白Scaf-CIQ的热稳定性, 所采用的方法如下:

[0068] 所述支架蛋白Scaf-CIQ的羧基端为Im9, 氨基端为CBM3a, 已知Im9与CL9特异性结合, CBM3a可与磷酸膨胀纤维素结合。基于此, 探究Scaf-CIQ热稳定性时, 采用磷酸膨胀纤维素为固定化基质, 通过氨基端融合有CL9-tag的sfGFP (CL9-sfGFP) 与Im9的结合, 可在磷酸膨胀纤维素上呈现绿色荧光。根据绿色荧光的强弱, 可判断出Scaf-CIQ蛋白质的完整性, 即高温处理后的Scaf-CIQ是否保留有氨基端的CBM3a与羧基端的Im9功能。

[0069] 本实施例中将支架蛋白Scaf-CIQ于70 $^{\circ}$ C、75 $^{\circ}$ C、80 $^{\circ}$ C和85 $^{\circ}$ C条件下分别孵育3h和6h, 12000rpm离心5min, 取上清与磷酸膨胀纤维素混合并加入过量CL9-sfGFP, 28 $^{\circ}$ C摇床220rpm孵育1h, 然后用2M NaCl洗涤磷酸膨胀纤维素3次, 去除游离蛋白质, 用100 μ L TBS buffer重悬样品, 采用荧光分光光度计对各样品中的CL9-sfGFP进行荧光强度定量。以未处理的Scaf-CIQ与CL9-sfGFP、磷酸膨胀纤维素结合测得的荧光强度为100%, 以未添加Scaf-CIQ的样品作为负对照。相对荧光强度的检测结果如图5所示, 结果显示, 在85 $^{\circ}$ C下培养3h后, 所述支架蛋白Scaf-CIQ对纤维素和CL9-sfGFP的结合能力约为50%, 在6h后仍有30%以上的结合能力, 即说明本发明所述的支架蛋白Scaf-CIQ具有较高的热稳定性, 因此可应用于高温环境下的多酶复合物组装和应用。

[0070] 实施例4多酶复合物MEC-CIQ的组装及应用

[0071] 1、含4种纤维素水解酶的多酶复合物MEC-CIQ的组装情况

[0072] 本实施例将多酶复合物组装系统MEC-CIQ应用于人工纤维小体领域。具体为: 选用的嗜热纤维素酶基因包括: (1) 通过DNA shuffling得到的外切葡聚糖Cel148S_{m3}、(2) 来源于*Clostridium thermocellum*的内切葡聚糖酶突变体Cel18A* (K276R, G283P, S329G, S375T)、(3) 来源于*Caldicellulosiruptor bescii*的内切葡聚糖酶GH5D; (4) 来源于*Caldicellulosiruptor owensensis*的 β -葡萄糖苷酶CoGH1A (各纤维素酶分别参考文献

Smith M.A., et al. *Febs Journal*. 2012, 279, 4453-65; Anbar M., et al. *Appl Environ Microbiol*. 2012, 78, 3458; Dvortsov IA, et al., *Int J Biol Macromol*, 2018, 107, 305-11; Peng X., et al. *Biotechnol Biofuels*, 2016, 9, 98.)。

[0073] 采用将上述四种嗜热纤维素酶基因分别与四种CL蛋白基因连接,形成CL-tagged嗜热纤维素酶,具体操作过程为:

[0074] 以经过NcoI和XhoI双酶切并琼脂糖凝胶回收后的pET28a为载体骨架,将上述四种嗜热纤维素酶基因视为片段A(A1、A2、A3和A4),四种CL蛋白基因视为片段B(B1、B2、B3和B4),A和B之间通过富含Pro/Thr的linker序列(TTVTTPQTS)连接。设计引物时,CL-tag在纤维素酶的N端情况下,B的反向引物和A的正向引物含有15个碱基长度的重叠区,B的正向引物和线性载体NcoI酶切位点一端含有15个相同的碱基,A的反向引物和线性载体XhoI酶切位点一端含有15个相同的碱基连接。CL-tag在纤维素酶的C端情况下,A的反向引物和B的正向引物含有15个碱基长度的重叠区,A的正向引物和线性载体NcoI酶切位点一端含有15个相同的碱基,B的反向引物和线性载体XhoI酶切位点一端含有15个相同的碱基连接。如此得到的PCR产物和线性载体的末端有15个相同的碱基。将片段顺序确定后的PCR产物A和B(如A1和B1组成基因A1-B1或B1-A1)与线性化载体在T5核酸外切酶的介导下进行三片段转化,在大肠杆菌体内经过同源重组后,基因A、B和线性载体组装成完整的环状质粒,含有CL-tagged嗜热纤维素酶基因。

[0075] 所有纤维素酶的重组载体均通过大肠杆菌BL21(DE3)表达,并通过Ni-NTA方式纯化,纯化后的蛋白质与Scaf-CIQ搭配组装成热稳定人工纤维小体,从而达到在高温环境下更好水解纤维素底物的目的,具体操作如下:

[0076] 本实施例中将体系定为500 μ L,其中4个CL-tagged纤维素酶终浓度为3 μ M,室温条件下与等摩尔Scaf-CIQ孵育30min,通过AKTA蛋白纯化仪进行分子筛层析,选用的分子筛柱型号为Superose 6Increase 10/300GL。以TBS buffer为纯化buffer,流速为0.5mL/min。根据峰值和SDS-PAGE检测多酶复合物的组装,检测结果如图6所示。结果显示,分子筛结果为单一峰,即在支架蛋白Scaf-CIQ的参与下,CL2-Ce148S_{m3}、Ce18A*-CL7、GH5D-CL8、CoGH1A-CL9成功组装在一起;并且SDS-PAGE结果表明该收集峰中含有预期的5种蛋白质,且分子量与预期一致。

[0077] 2、多酶复合物MEC-CIQ中4种纤维素水解酶的协同作用

[0078] 为验证组装成功的人工纤维小体,即包含4种纤维水解酶的多酶复合物MEC-CIQ的性能,采用PASC为纤维素底物,进行纤维素酶活性测定。

[0079] 采用的测定温度为65 $^{\circ}$ C、70 $^{\circ}$ C、75 $^{\circ}$ C、80 $^{\circ}$ C,反应时间为3h。在含有10mM CaCl₂的400 μ L醋酸缓冲液体系中各种纤维素酶的终浓度为1 μ M,PASC的量为2.5%。将反应完毕的体系12000rpm离心1min,取120 μ L上清与180 μ L DNS溶液混合,沸水浴处理10min后立即置于冰水混合物中冷却。6000rpm离心1min,取250 μ L反应液在OD₅₄₀处测定吸光度。将测得的数值代入至测定的葡萄糖标准曲线中,计算出每组反应生成的葡萄糖含量,每组实验均设置三组独立平行,并以未进行多酶复合物组装的游离的4种纤维素水解酶作为对照。

[0080] 检测结果如图7所示,结果显示以PASC为底物时,每组实验中均能体现出多酶协同作用优势。其中,MEC-CIQ体系在70 $^{\circ}$ C反应3h后的还原糖产量提升幅度最大,为游离酶体系的160%。基于此,本发明构建的MEC-CIQ系统可成功应用于多酶复合物组装,且在高温环境

下表现出卓越的性能。

[0081] 以上所述,仅为本发明较佳的具体实施方式,但本发明的保护范围并不局限于此,任何熟悉本技术领域的技术人员在本发明揭露的技术范围内,可轻易想到的变化或替换,都应涵盖在本发明的保护范围之内。

35 40 45
 Phe Asp Asp Phe Arg Glu Thr Phe Trp Glu Glu Val Ser Lys Asp Pro
 50 55 60
 Asp Leu Ser Lys Gln Phe Lys Gly Ser Asn Lys Thr Asn Ile Gln Lys
 65 70 75 80
 Gly Lys Ala Pro Phe Ala Arg Lys Lys Asp Gln Val Gly Gly Arg Glu
 85 90 95
 Arg Phe Glu Leu Asn His Gln Lys Pro Ile Glu Gln Asp Gly Gly Val
 100 105 110
 Tyr Asp Met Asn Asn Ile Arg Val Thr Thr Pro Lys Arg Asn Ile Asp
 115 120 125
 Ile Glu Gly
 130
 <210> 3
 <211> 88
 <212> PRT
 <213> 大肠杆菌 (Escherichia coli)
 <400> 3
 Met Glu Leu Lys His Ser Ile Ser Asp Tyr Thr Glu Ala Glu Phe Leu
 1 5 10 15
 Glu Phe Val Lys Lys Ile Cys Arg Ala Glu Gly Ala Thr Glu Glu Asp
 20 25 30
 Asp Asn Lys Leu Val Arg Glu Phe Glu Arg Leu Thr Glu His Pro Asp
 35 40 45
 Gly Ser Asp Leu Ile Tyr Tyr Pro Arg Asp Asp Arg Glu Asp Ser Pro
 50 55 60
 Glu Gly Ile Val Lys Glu Ile Lys Glu Trp Arg Ala Ala Asn Gly Lys
 65 70 75 80
 Ser Gly Phe Lys Gln Gly Leu Glu
 85
 <210> 4
 <211> 132
 <212> PRT
 <213> 大肠杆菌 (Escherichia coli)
 <400> 4
 Ser Lys Arg Asn Lys Pro Gly Lys Ala Thr Gly Lys Gly Lys Pro Val
 1 5 10 15
 Asn Asn Lys Trp Leu Asn Asn Ala Gly Lys Asp Leu Gly Ser Pro Val
 20 25 30

Pro Asp Arg Ile Ala Asn Lys Leu Arg Asp Lys Glu Phe Lys Ser Phe
 35 40 45
 Asp Asp Phe Arg Lys Lys Phe Trp Glu Glu Val Ser Lys Asp Pro Glu
 50 55 60
 Leu Ser Lys Gln Phe Ser Arg Asn Asn Asn Asp Arg Met Lys Val Gly
 65 70 75 80
 Lys Ala Pro Lys Thr Arg Thr Gln Asp Val Ser Gly Lys Arg Thr Ser
 85 90 95
 Phe Glu Leu His His Glu Lys Pro Ile Ser Gln Asn Gly Gly Val Tyr
 100 105 110
 Asp Met Asp Asn Ile Ser Val Val Thr Pro Lys Arg His Ile Asp Ile
 115 120 125
 His Arg Gly Lys
 130
 <210> 5
 <211> 131
 <212> PRT
 <213> 人工序列(Artificial Sequence)
 <400> 5
 Met Ser Lys Ser Asn Glu Pro Gly Lys Ala Thr Gly Glu Gly Lys Pro
 1 5 10 15
 Val Asn Asn Lys Trp Leu Asn Asn Ala Gly Lys Asp Leu Gly Ser Pro
 20 25 30
 Val Pro Asp Arg Ile Ala Asn Lys Leu Arg Asp Lys Glu Phe Glu Ser
 35 40 45
 Phe Asp Asp Phe Arg Glu Thr Phe Trp Glu Glu Val Ser Lys Asp Pro
 50 55 60
 Glu Leu Ser Lys Gln Phe Ser Arg Asn Asn Asn Asp Arg Met Lys Val
 65 70 75 80
 Gly Lys Ala Pro Lys Thr Arg Thr Gln Asp Val Ser Gly Lys Arg Thr
 85 90 95
 Ser Phe Glu Leu Asn His Gln Lys Pro Ile Glu Gln Asn Gly Gly Val
 100 105 110
 Tyr Asp Met Asp Asn Ile Ser Val Val Thr Pro Lys Arg Asn Ile Asp
 115 120 125
 Ile Glu Gly
 130
 <210> 6
 <211> 87

<212> PRT

<213> 大肠杆菌 (Escherichia coli)

<400> 6

```

Met Glu Leu Lys Asn Ser Ile Ser Asp Tyr Thr Glu Ala Glu Phe Val
1           5           10           15
Gln Leu Leu Lys Glu Ile Glu Lys Glu Asn Val Ala Ala Thr Asp Asp
           20           25           30
Val Leu Asp Val Leu Leu Glu His Phe Val Lys Ile Thr Glu His Pro
           35           40           45
Asp Gly Thr Asp Leu Ile Tyr Tyr Pro Ser Asp Asn Arg Asp Asp Ser
           50           55           60
Pro Glu Gly Ile Val Lys Glu Ile Lys Glu Trp Arg Ala Ala Asn Gly
65           70           75           80
Lys Pro Gly Phe Lys Gln Gly
           85

```

<210> 7

<211> 132

<212> PRT

<213> 大肠杆菌 (Escherichia coli)

<400> 7

```

Ser Lys Arg Asn Lys Pro Gly Lys Ala Thr Gly Lys Gly Lys Pro Val
1           5           10           15
Gly Asp Lys Trp Leu Asp Asp Ala Gly Lys Asp Ser Gly Ala Pro Ile
           20           25           30
Pro Asp Arg Ile Ala Asp Lys Leu Arg Asp Lys Glu Phe Lys Asn Phe
           35           40           45
Asp Asp Phe Arg Arg Lys Phe Trp Glu Glu Val Ser Lys Asp Pro Glu
           50           55           60
Leu Ser Lys Gln Phe Asn Pro Gly Asn Lys Lys Arg Leu Ser Gln Gly
65           70           75           80
Leu Ala Pro Arg Ala Arg Asn Lys Asp Thr Val Gly Gly Arg Arg Ser
           85           90           95
Phe Glu Leu His His Asp Lys Pro Ile Ser Gln Asp Gly Gly Val Tyr
           100           105           110
Asp Met Asp Asn Leu Arg Ile Thr Thr Pro Lys Arg His Ile Asp Ile
           115           120           125
His Arg Gly Gln
           130
<210> 8

```

<211> 131

<212> PRT

<213> 人工序列(Artificial Sequence)

<400> 8

Met	Ser	Lys	Ser	Asn	Glu	Pro	Gly	Lys	Ala	Thr	Gly	Glu	Gly	Lys	Pro
1				5					10					15	
Val	Gly	Asp	Lys	Trp	Leu	Asp	Asp	Ala	Gly	Lys	Asp	Ser	Gly	Ala	Pro
			20					25					30		
Ile	Pro	Asp	Arg	Ile	Ala	Asp	Lys	Leu	Arg	Asp	Lys	Glu	Phe	Glu	Asn
			35				40						45		
Phe	Asp	Asp	Phe	Arg	Glu	Thr	Phe	Trp	Glu	Glu	Val	Ser	Lys	Asp	Pro
			50				55				60				
Glu	Leu	Ser	Lys	Gln	Phe	Asn	Pro	Gly	Asn	Lys	Lys	Arg	Leu	Ser	Gln
65				70						75					80
Gly	Leu	Ala	Pro	Arg	Ala	Arg	Asn	Lys	Asp	Thr	Val	Gly	Gly	Arg	Arg
				85					90					95	
Ser	Phe	Glu	Leu	Asn	His	Gln	Lys	Pro	Ile	Glu	Gln	Asp	Gly	Gly	Val
			100					105						110	
Tyr	Asp	Met	Asp	Asn	Leu	Arg	Ile	Thr	Thr	Pro	Lys	Arg	Asn	Ile	Asp
			115					120						125	
Ile	Glu	Gly													
			130												

<210> 9

<211> 85

<212> PRT

<213> 大肠杆菌(Escherichia coli)

<400> 9

Met	Glu	Leu	Lys	Asn	Ser	Ile	Ser	Asp	Tyr	Thr	Glu	Thr	Glu	Phe	Lys
1				5					10					15	
Lys	Ile	Ile	Glu	Asp	Ile	Ile	Asn	Cys	Glu	Gly	Asp	Glu	Lys	Lys	Gln
			20					25					30		
Asp	Asp	Asn	Leu	Glu	His	Phe	Ile	Ser	Val	Thr	Glu	His	Pro	Ser	Gly
			35				40						45		
Ser	Asp	Leu	Ile	Tyr	Tyr	Pro	Glu	Gly	Asn	Asn	Asp	Gly	Ser	Pro	Glu
			50				55					60			
Ala	Val	Ile	Lys	Glu	Ile	Lys	Glu	Trp	Arg	Ala	Ala	Asn	Gly	Lys	Ser
65				70						75					80
Gly	Phe	Lys	Gln	Gly											
			85												

<210> 10

<211> 132

<212> PRT

<213> 大肠杆菌 (Escherichia coli)

<400> 10

Ser Lys Arg Asn Lys Pro Cys Lys Ala Thr Gly Lys Gly Lys Pro Val
 1 5 10 15
 Gly Asp Lys Trp Leu Asp Asp Ala Gly Lys Asp Ser Gly Ala Pro Ile
 20 25 30
 Pro Asp Arg Ile Ala Asp Lys Leu Arg Asp Lys Glu Phe Lys Ser Phe
 35 40 45
 Asp Asp Phe Arg Lys Ala Val Trp Glu Glu Val Ser Lys Asp Pro Glu
 50 55 60
 Leu Ser Lys Asn Leu Asn Pro Ser Asn Lys Ser Ser Val Ser Lys Gly
 65 70 75 80
 Tyr Ser Pro Phe Thr Pro Lys Asn Gln Gln Val Gly Gly Arg Lys Val
 85 90 95
 Tyr Glu Leu His His Asp Lys Pro Ile Ser Gln Gly Gly Glu Val Tyr
 100 105 110
 Asp Met Asp Asn Ile Arg Val Thr Thr Pro Lys Arg His Ile Asp Ile
 115 120 125
 His Arg Gly Lys
 130

<210> 11

<211> 131

<212> PRT

<213> 人工序列 (Artificial Sequence)

<400> 11

Met Ser Lys Ser Asn Glu Pro Gly Lys Ala Thr Gly Glu Gly Lys Pro
 1 5 10 15
 Val Gly Asp Lys Trp Leu Asp Asp Ala Gly Lys Asp Ser Gly Ala Pro
 20 25 30
 Ile Pro Asp Arg Ile Ala Asp Lys Leu Arg Asp Lys Glu Phe Glu Ser
 35 40 45
 Phe Asp Asp Phe Arg Glu Thr Val Trp Glu Glu Val Ser Lys Asp Pro
 50 55 60
 Glu Leu Ser Lys Asn Leu Asn Pro Ser Asn Lys Ser Ser Val Ser Lys
 65 70 75 80
 Gly Tyr Ser Pro Phe Thr Pro Lys Asn Gln Gln Val Gly Gly Arg Lys

				85					90					95		
Val	Tyr	Glu	Leu	Asn	His	Gln	Lys	Pro	Ile	Glu	Gln	Gly	Gly	Glu	Val	
				100				105						110		
Tyr	Asp	Met	Asp	Asn	Ile	Arg	Val	Thr	Thr	Pro	Lys	Arg	Asn	Ile	Asp	
				115				120						125		
Ile	Glu	Gly														
				130												
<210>	12															
<211>	86															
<212>	PRT															
<213>	大肠杆菌 (Escherichia coli)															
<400>	12															
Met	Glu	Leu	Lys	His	Ser	Ile	Ser	Asp	Tyr	Thr	Glu	Ala	Glu	Phe	Leu	
1				5				10						15		
Gln	Leu	Val	Thr	Thr	Ile	Cys	Asn	Ala	Asp	Thr	Ser	Ser	Glu	Glu	Glu	
				20				25						30		
Leu	Val	Lys	Leu	Val	Thr	His	Phe	Glu	Glu	Met	Thr	Glu	His	Pro	Ser	
				35				40						45		
Gly	Ser	Asp	Leu	Ile	Tyr	Tyr	Pro	Lys	Glu	Gly	Asp	Asp	Asp	Ser	Pro	
				50				55						60		
Ser	Gly	Ile	Val	Asn	Thr	Val	Lys	Gln	Trp	Arg	Ala	Ala	Asn	Gly	Lys	
65				70						75					80	
Ser	Gly	Phe	Lys	Gln	Gly											
				85												

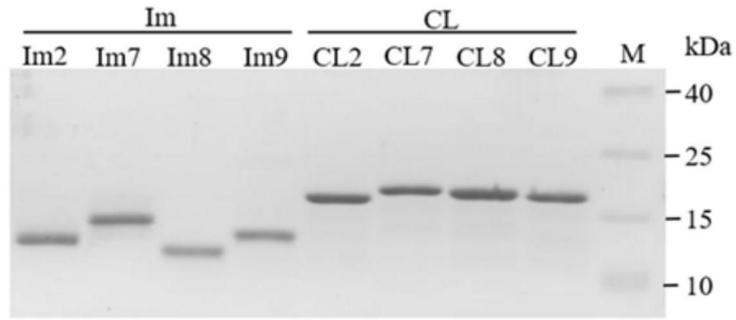


图1

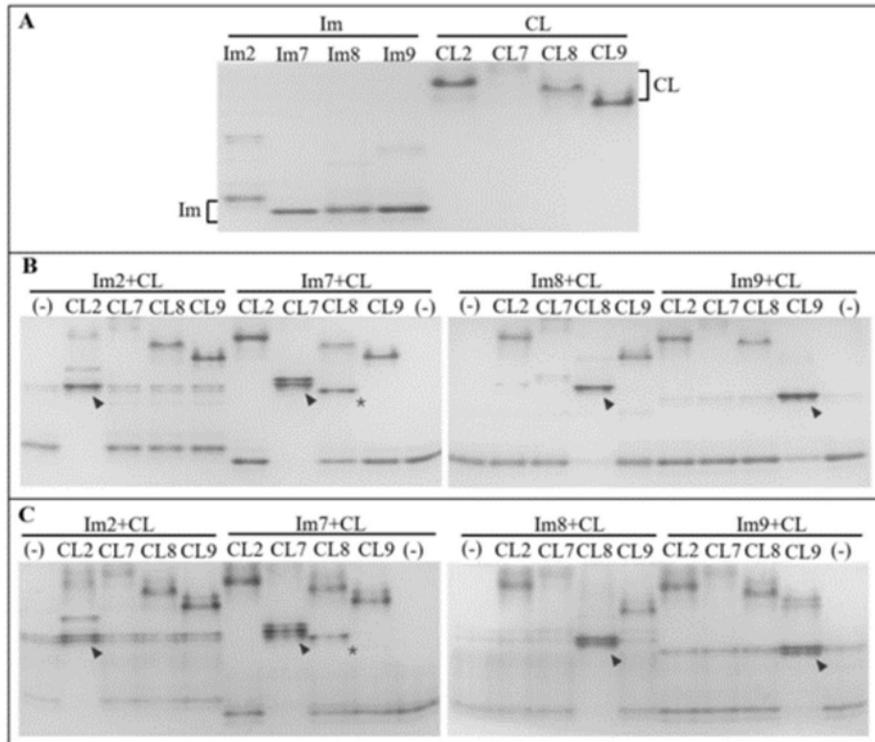


图2

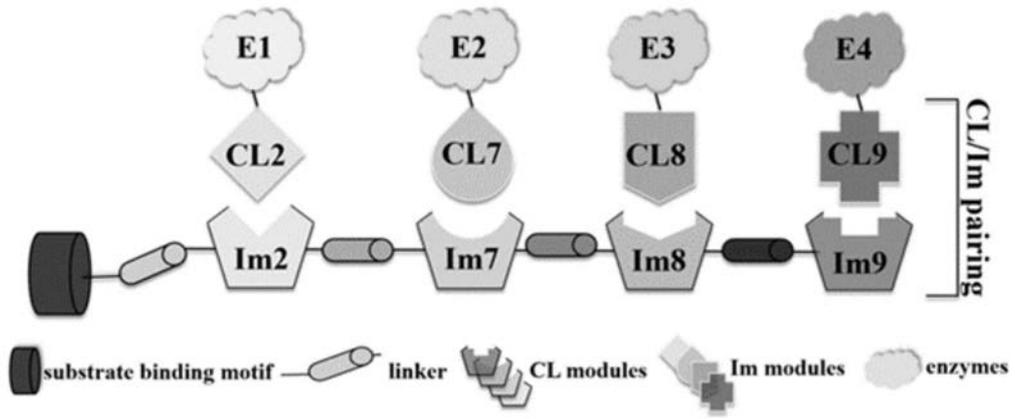


图3

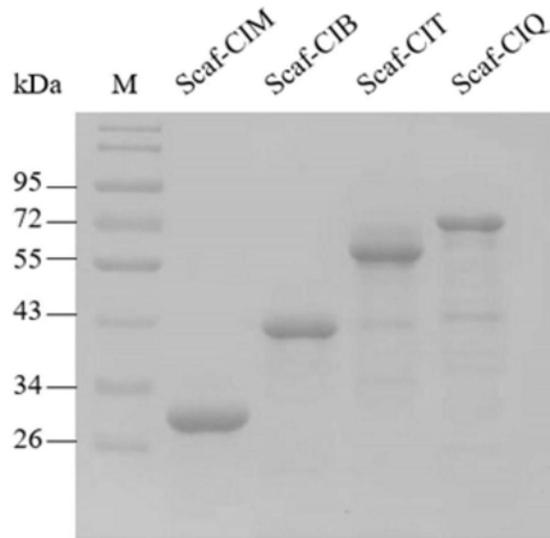


图4

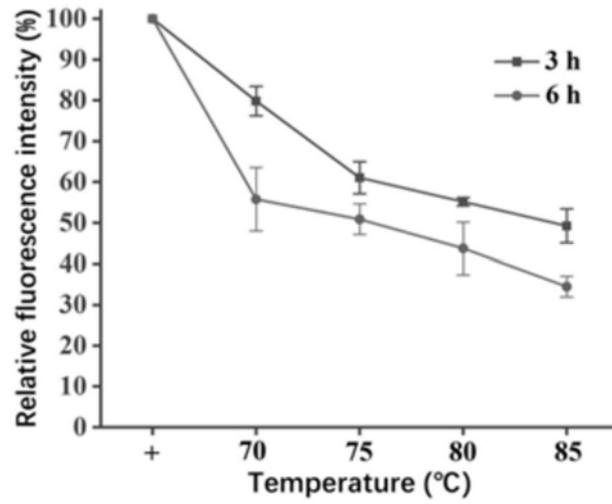


图5

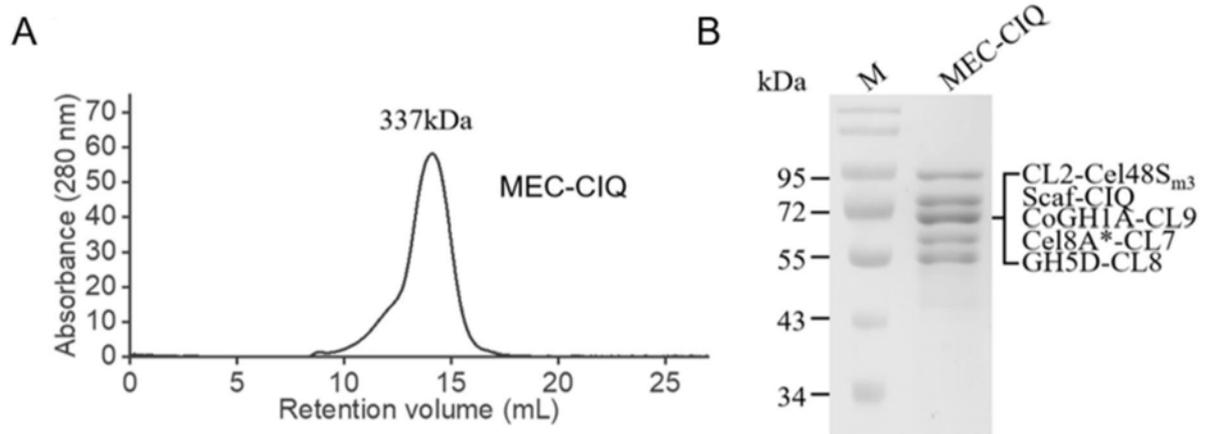


图6

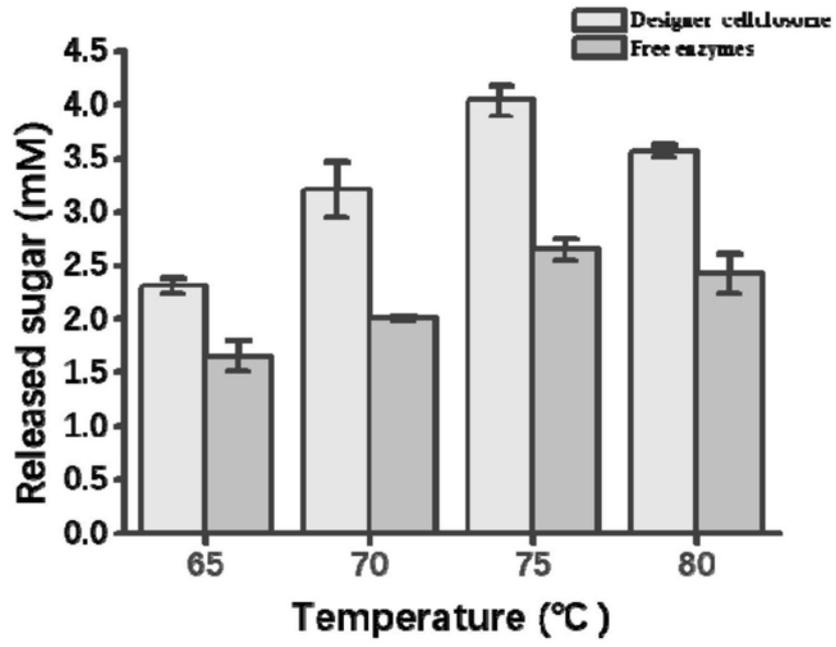


图7