

[19] 中华人民共和国国家知识产权局



[12] 发明专利申请公布说明书

[21] 申请号 200780048162.0

[51] Int. Cl.

C07K 14/435 (2006.01)

C12N 9/42 (2006.01)

C12N 9/24 (2006.01)

C12N 15/09 (2006.01)

[43] 公开日 2009 年 10 月 28 日

[11] 公开号 CN 101568551A

[22] 申请日 2007.11.6

[21] 申请号 200780048162.0

[30] 优先权

[32] 2006.11.6 [33] IN [31] 2411/DEL/2006

[86] 国际申请 PCT/IN2007/000521 2007.11.6

[87] 国际公布 WO2008/056376 英 2008.5.15

[85] 进入国家阶段日期 2009.6.25

[71] 申请人 科学与工业研究委员会

地址 印度新德里

[72] 发明人 迪维亚·卡普尔

桑吉夫·库马·钱德拉扬

舒伯尔·阿赫迈德 斯瓦蒂·夏尔马

马尼施·达塔 巴尔温德尔·辛格

卡蒂肯扬·苏布拉马尼亚恩

普马南达·古普塔萨玛

[74] 专利代理机构 中科专利商标代理有限责任公
司

代理人 程金山

权利要求书 4 页 说明书 40 页 序列表 3 页
附图 19 页

[54] 发明名称

重组中温活性热稳定蛋白及其设计和生物合成方法

[57] 摘要

本发明涉及检验基于 β 折叠的蛋白结构的一部分或完整表面的可变性，特别集中于酶。所述改变是通过将一种蛋白的一部分或完整表面取代/移植到可叠合多肽主链的同源蛋白表面上，通过使用 β 折叠的结构特征仅改变参与底物/配体结合和催化作用的表面区域进行的。所述移植涉及将选择的构成一种酶/蛋白中待改变的表面区域的不连续残基组替换为位于另一种酶/蛋白中类似位置处的不连续残基组，这是以促进通过组合来自这两种酶/蛋白的残基形成的新蛋白的折叠和功能的方式进行的。本发明还涉及利用该表面改造方法来选择性组合来自不同生命领域的酶/蛋白特征，所述特征一般不通过天然进化组合，诸如构建这样的新蛋白，所述新蛋白保留在其上移植了嗜中温同系物的活性表面的嗜热性

酶的大部分热稳定支架，以构建具有最佳功能 pH 和温度的中温活性功能特征的热稳定蛋白。

1. 重组中温活性热稳定蛋白，其包括嗜热性蛋白的结构稳定性特征和嗜中温蛋白的活性特征。

2. 按照权利要求 1 的重组中温活性热稳定蛋白，包括具有单结构域的蛋白质，其中主要不连续的氨基酸组被结构-同源蛋白中结构等价位置处存在的不同的不连续氨基酸组替换。

3. 按照权利要求 1 和 2 的重组中温活性热稳定蛋白，其中“客体”或“供体”祖蛋白的一部分或完整表面有效地移植到“主体”或“接纳体”祖蛋白的结构核心上。

4. 按照权利要求 1-3 的重组中温活性热稳定蛋白，其中所述客体和主体祖蛋白具有同源的结构和相同的功能。

5. 按照权利要求 1-4 的重组中温活性热稳定蛋白，其中所述客体祖蛋白是嗜中温蛋白，且所述主体祖蛋白是嗜热性蛋白。

6. 按照权利要求 1-5 的重组中温活性热稳定蛋白，其中所述同源的客体和主体祖蛋白的功能-活性表面区域由基于 β 折叠的二级结构组成。

7. 按照权利要求 1-6 的重组中温活性热稳定蛋白，其中使用的嗜热性主体祖蛋白和嗜中温客体先祖具有分别由 SEQ ID NO:1 和 SEQ ID NO:2 表示的氨基酸序列。

8. 按照权利要求 1-7 的重组中温活性热稳定蛋白，其包括由 SEQ ID NO:3 表示的氨基酸序列，其具有由 SEQ ID NO:1 表示的主体蛋白的结构稳定性特征（热-稳定性）和由 SEQ ID NO:2 表示的客体蛋白的活性特征（中温-活性）。

9. 按照权利要求 1-8 的重组中温活性热稳定蛋白，其具有下列特征：

- i. 分子质量：20-30 千道尔顿；
- ii. 残基数目：约 200-300 个；
- iii. 等电点：在 4-8 范围内；
- iv. 最佳活性 pH：在 4-8 范围内；
- v. 由解链温度 (T_m) 定义的热稳定性：80-95°C，和

vi. 由最佳活性温度 (T_{OA}) 定义的中温-活性: 在 30-60°C 范围内。

10. 按照权利要求 1-9 的重组中温活性热稳定蛋白, 其中所述蛋白及其先祖是属于水解酶组的单-结构域 β 折叠蛋白结构类的酶, 并且选自由纤维素酶、木聚糖酶、淀粉酶和蛋白酶组成的组, 优选地为纤维素酶。

11. 按照权利要求 1 的重组中温活性热稳定蛋白, 其还可以起生物催化剂的作用。

12. 用于设计包括主体祖蛋白的结构稳定性特征和不同的且结构同源的客体祖蛋白的活性特征的重组中温活性热稳定蛋白的方法。

13. 按照权利要求 12 的用于设计重组中温活性热稳定蛋白的方法, 其中所述结构-同源的主体和客体祖蛋白包括这样的结构, 在所述结构中, 由位于基于 β 折叠的二级结构中的残基提供底物-结合和催化功能。

14. 用于设计重组中温活性热稳定蛋白的方法, 所述方法包括下列步骤:

a. 叠合所述客体 and 主体祖蛋白的结构, 并在多肽主链水平上, 优选地, 在包括底物-结合和催化功能的结构区域中, 评估它们结构的叠合性程度,

b. 通过使用步骤 (a) 的信息, 生成所述主体和客体祖蛋白的蛋白序列以及编码 DNA 序列的基于结构类似的序列对比,

c. 鉴定在主体祖蛋白序列中构成基于 β 折叠的‘活性表面’的特定氨基酸残基, 并且使用获自步骤 (b) 的对比序列以在所述客体祖蛋白中鉴定相应类似的不连续残基和残基组,

d. 确定在来自所述主体和客体祖蛋白构成所述两种蛋白的活性表面的两个相应氨基酸残基组中相同的特定残基, 并且消除所述相同的残基, 从而获得在所述两种蛋白中的结构-等价位置处存在的一组相应的不同残基,

e. 将所述主体祖蛋白的‘活性表面’中与所述客体祖蛋白中相应的残基不相同的残基替换为后者蛋白中的残基, 由此构成新型中温活性热稳定蛋白,

f. 通过已知的重组 DNA 方法, 生物合成所述新型中温活性热稳定蛋白, 随后通过已知方法分离和纯化蛋白, 和

g. 通过活性测量, 证实在理想的中温活性热稳定蛋白中的所述主体祖蛋白的结构稳定性特征和所述客体祖蛋白的物理和化学活性特征。

15. 按照权利要求 12-14 的方法, 其中构成所述产物中温活性热稳定蛋白的结构核心的氨基酸残基源自所述两种祖蛋白中的一种, 而构成所述产物蛋白的一部分或完整表面的残基源自另一种祖蛋白。

16. 按照权利要求 12-15 的方法, 其中所述两种祖蛋白是结构同源的, 它们的主链原子的坐标以 0.5-2.5 埃的均方根偏差 (RMSD) 叠合。

17. 按照权利要求 12-16 的方法, 其中所述主体和客体祖蛋白二者中的最佳功能温度和结构解链温度相差 5°C 以内。

18. 按照权利要求 12-17 的方法, 其中所述两种祖蛋白的结构-同源‘活性表面’区域主要由基于 β 折叠的二级结构组成。

19. 按照权利要求 12-18 的方法, 其中使用的客体先祖是 SEQ ID NO:1 的嗜热性蛋白 RM Cel12A。

20. 按照权利要求 12-19 的方法, 其中使用的主体先祖是 SEQ ID NO:2 的嗜中温蛋白 TR Cel12A。

21. 按照权利要求 12-20 的方法, 其中构成所述产物蛋白的结构核心的氨基酸序列源自 RM Cel12A, 而构成所述产物蛋白一部分表面的残基源自 TR Cel12A。

22. 按照权利要求 12-21 的方法, 其中所述 RM Cel12A 和 TR Cel12A 蛋白是结构同源的, 它们的主链原子的坐标以 0.5-2.5 埃的均方根偏差 (RMSD) 叠合。

23. 按照权利要求 12-22 的方法, 其中 RM Cel12A 和 TR Cel12A 二者中的最佳功能温度和结构解链温度相差 5°C 以内。

24. 按照权利要求 12-23 的方法, 其中 RM Cel12A 和 TR Cel12A 的结构-同源‘活性表面’区域主要由基于 β 折叠的二级结构组成。

25. 按照权利要求 12-24 的方法, 其中所述 SEQ ID NO:3 的产物 MT Cel12A 是通过将所述主体先祖活性表面包括的残基替换为所述客体先祖活性表面包括的结构-相似的残基, 而来源于所述主体先祖。

26. 按照权利要求 12-25 的方法, 其中获得的新型中温活性热稳定(MT Cel2A)产物具有所述嗜中温客体先祖 TR Cel12A 的最佳活性温度和所述同

源嗜热性主体先祖 RM Cel12A 的结构稳定性。

27. 按照权利要求 12-26 的方法，其中所述中温活性热稳定蛋白(MT Cel2A)具有下列特征：

- i. 分子质量： 20-30 千道尔顿；
- ii. 残基数目： 约 200-300 个；
- iii. 等电点： 在 4-8 范围内；
- iv. 最佳活性 pH： 在 4-8 范围内；
- v. 由解链温度 (T_m) 定义的热稳定性： 80-95°C，和
- vi. 由最佳活性温度 (T_{OA}) 定义的中温-活性： 在 30-60°C 范围内。

28. 按照权利要求 12-18 的方法，其有效用于在不改变所述酶活性的化学特征或底物的化学限定的条件下调整酶的物理功能特征。

29. 按照权利要求 12-19 的方法，其有效用于重组来自两种非常不同的生命领域的酶结构稳定性与蛋白稳定性和活性特征。

30. 按照权利要求 12-19 的方法，其中所述中温活性热稳定蛋白有效用于纺织工业的应用中，诸如牛仔布的石洗。

重组中温活性热稳定蛋白及其设计和生物合成方法

发明领域

本发明提供包括嗜热性蛋白的结构稳定性特征和嗜中温蛋白的活性特征的重组中温活性热稳定蛋白（meso-active thermo-stable protein）及其设计和生物合成方法，其中包括嗜热性蛋白的底物-结合和催化活性表面的主要不连续氨基酸组被结构-同源的嗜中温蛋白中的结构等价位置处存在的不同的不连续氨基酸组所替换。

更具体地，本发明描述通过基于理性的残基替换由一组残基改变，或“理论上移植”，其表面的一个部分（或多个部分）而谨慎改变蛋白分子功能行为的一般适用方法，所述一组残基包括不同的、独立演化的、结构同源的蛋白质中的一个功能等价表面（或多个表面），所述蛋白在结构稳定性、和物理活性方面具有明显不同的特征，但是具有相同的化学特征。变化的蛋白中的许多不同的不连续放置的残基被同源蛋白等价位置处使用的残基合理替换意欲引起将所述变化的蛋白折叠为三维结构，其充满相同蛋白的未改变形式的原始结构稳定性特征，和另一种（同源）蛋白的活性特征（例如，最佳活性的温度）。这促进产生组合来源于不同生命形式和不同生命领域的酶的理想特征的酶。

现有技术和发明背景

*蛋白质序列-结构关系：常识：*蛋白质是通过肽键连接在一起的氨基酸的天然聚合物，其在合成时折叠成具有某些生物活性的三维结构。蛋白质的三维结构受到其氨基酸序列的控制（Anfinsen, 1973. 科学（*Science*）181, 223）。氨基酸序列包括指示由蛋白链形成三维结构所需的全部相关信息。在进化过程中通过天然引入突变产生的或通过遗传工程技术引起的蛋白质序列的改变导致蛋白质结构的改变。所述改变可以是细微的或深切的。如果所述改变是细微的，则它们通常仅涉及蛋白质特定区域显微结构的微小变化，表现为变化的残基附近的局部残基群（所述蛋白中包埋的，或位

于其表面上的)的形状改变,但对该蛋白质总体形状或功能,或其肽主链穿过其三维结构所采用的轨道不具有任何深切或明显的影响。然而,当所述改变是深切的时候,它们改变蛋白质的整体形状以及主链穿过蛋白结构所采用的轨道。有时,由突变产生的深切变化甚至可以引起所述链失去稳定折叠为特定三维结构的能力(导致聚集和沉淀)。不能总使突变对蛋白结构的影响与其序列中发生的变化相联系或针对其进行校准。尽管现在所述非常有限改变的作用可以通过计算机建模,但是实验研究序列改变的作用在所有情形中(无论这些改变是细微的或是深切的)成为基本的,因为还不能对确定这些改变的作用所涉及的所有参数、以及物理力进行建模。现今充分已知的是深切和细微的序列改变均可以以挑战可预期性的方式导致结构深切或细微的改变。

因此,有时可以看到来自两种在进化上不相关的生物体的两种蛋白质具有在它们的总形状和折叠结构上几乎相同的多肽主链(尽管不是氨基酸残基侧链),即使这两种蛋白质具有无相似性的完全不同的氨基酸序列。然而,这作为超越这样的规则的例外,所述规则是,通常相似尺寸的蛋白质仅倾向于采用基本相似的主链结构,如果它们具有稍微相似的氨基酸序列(包括所有残基的至少 20%的一致性)。然而,所述蛋白质的外部形状特征非常不同,且这是由于所述蛋白质特有的特异性相互作用残基(侧链)组以由其特异性氨基酸序列决定的方式对各种蛋白质的主链进行特异性“修饰”。因此,主链结构的保守性与广泛的功能保守性相关;功能性的精确热力学和动力学参数几乎完全受到由蛋白质表面上存在的侧链所决定的蛋白质外部形状特征的影响。

为了总结以上讨论,氨基酸序列和蛋白质结构之间的精确关系是非常微妙的;尚不理解或领会这种关系的所有方面,且尚不可能在不进行必要实验或不参考特定结构情形的条件下预期在序列中进行的特定变化对蛋白质结构的影响。在这点上,和在与通过序列改变再改造任何蛋白质表面的特定关系中,折叠和稳定性的考虑因素起作用,因为改造完整或部分的蛋白质表面影响其结构-形成能力,及其结构稳定性。通常,在改造蛋白质的所有尝试中,无论关于它们的表面还是它们的内部,可以采用两种不同的方法,这些是:(i)基于结构-功能分析的理性改造方法,和谨慎引入特

定的突变, 和 (ii) 更依赖于随机过程诸如基因改组, 或筛选表现出随机产生的变体群的噬菌体的非理性的(基于组合学的和基于定向进化的)方法, 随后基于结合特征进行选择。在蛋白质工程领域中, 理性方法是最初采用的方法。然而, 因为做出的改变的作用的不可预期性, 证明它不够令人满意。随后, 最近可用的重组 DNA 技术使得基于组合学的(组合的)方法也可行。所述理性方法的不够令人满意的结果最初导致转换为组合方法。然而, 逐渐地, 研究可以通过完全随机方法(对于 n 个残基的链有 20^n 种变化)令人信服地产生的序列变化的甚至非常小的级分的不可行性导致采用试图组合这两种方法的最佳部分的杂交方法。这些杂交方法包括理性选择待进行改变的蛋白质中的残基或结构位点, 和非-理性(基于组合学研究)研究在所述位点处进行突变的影响。以下, 我们简要评述先前的工作者如何采用理性、非-理性或杂交方法, 从而帮助随后对这些和我们在本发明中采用的方法进行区分。

纯理性蛋白质改造的实例: Rutter 和同事在胰蛋白酶的活性位点的两个位置引入位点-特异性突变(在对其活性位点进行结构分析后), 从而降低催化速度, 但增强酶针对其天然底物的底物特异性(Craik 等, 1985. 科学 (*Science*) 228, 291-297)。许多其他小组随后基于蛋白质结构的理性分析引入有限的突变, 从而产生具有改变的速度和/或亲和性特征的变体。Estell 和同事在关于酶活性位点附近的静电学方面, 改造了另一种蛋白酶, 即枯草杆菌蛋白酶, 从而改变由于它们静电学特征而不同的底物结合的优选性(Wells 等, 1987. 美国国家科学院学报 (*Proc. Natl. Acad. Sci. USA*) 84, 1219-1223)。Perham 和同事在谷胱甘肽还原酶中进行了理性的突变, 以在将其辅酶特异性由 NADP^+ 改变为 NAD^+ 的同时保持其底物特异性不变(Scrutton 等, 1990. 自然 (*Nature*) 343, 38-43)。在九十年代中进行了所述工作的若干实例, 在所有这些中在活性位点附近引入有限的理性-选择的突变, 以在活性位点附近改变关于蛋白质-配体相互作用方面的酶特征。随后, 进行了下述更大胆的改造尝试。Benkovic 和同事成功地利用细胞核转运因子 2 的结构同源蛋白支架设计了 scytalone dehydratase-样酶, 从而证明所述理性改造方法在通过重新设计支架蛋白的主要部分开发新实体中的功效(Nixon 等, 1999. 国家科学院学报 (*Proct Natl Acad Sci.*) 96, 3568-3571)。

类似地, Hellinga 和同事分析了核糖-结合蛋白的结构, 针对预期可能赋予该蛋白丙糖磷酸异构酶 (TIM) 活性的 18-22 位点-定向的突变合理地选择了位点, 并证明这是实验的情形(Dwyer 等, 2004. 科学 (*Science*) 304, 1967-1971)。随后, 组合的方法也获得了成功。

纯非-理性 (组合) 蛋白质改造的实例: 定向的进化由将随机分布的突变低频率引入到目的基因中, 随后选择具有理想性质的突变的 (变体) 蛋白质组成(Roberto 等, 2005. 现代生物技术观点 (*Current Opinion in biotechnology*) . 16, 378-384)。证明了定向进化是修饰蛋白质的有力工具, 并且现在已经成为广泛使用的方法。然而, 其主要用于利用错误-倾向的 PCR 技术将突变随机引入蛋白质序列来寻找温度-敏感的以及诸如此类的突变体, 和用于发展新型结合试剂 (通过噬菌体-展示组合方法, 其包括通过将简并寡核苷酸结合到编码 DNA 中随机化蛋白质的部分)。存在非常少量的用于改变酶活性的纯非-理性方法的实例, 这大概是因为用于随机引入突变的机制在不进行一些位点的理性选择的条件下, 通常不能被控制和限制在蛋白质表面特定的区域内, 这是因为组合空间的完全研究是不可能的 (存在太多可以产生的变体)。利用人雌激素受体 α 配体结合结构域, Zhao 和同事使用随机诱变和体外定向进化, 从而发展新型皮质酮活性(Chen 等, 2005. 分子生物学杂志 (*J. Mol. Biol.*) 348, 1273-1282)。

纯粹的随机改造方法应用的另一个著名实例 (然而, 基于随机化位点的半-理性选择) 是 Bryan 和同事的实例, 他们利用定向共同进化, 来改变无钙的枯草杆菌蛋白酶的稳定性和催化活性(Strausberg 等, 2005. 生物化学 (*Biochemistry*) 44, 3272-3279)。定向进化深切改变热稳定性而非活性的应用的一个实例是 Rao 和同事的实例, 他们使用该方法开发高度稳定的脂肪酶(Acharya 等, 2004. 分子生物学杂志 (*J. Mol. Biol.*) 341, 1271-1281)。基本上, 现今的趋势是混合所述理性和非-理性的方法, 一些这样的实例引用如下。

杂交 (理性-组合的) 蛋白质改造实例: 如通过由 Kim 和同事在乙二醛酶 II 的 $\alpha\beta/\beta\alpha$ 金属水解酶支架上发展新的催化活性 (β -内酰胺酶活性) 所例证地, 一些小组使用了杂交方法, 其组合理性和组合的要素以获得成功的结果(Park 等, 2006. 科学 (*Science*) . 311, 535-538)。该方法成功的第

二个实例是 Peimbert 和 Segovia 的实例，他们在 D-Ala D-Ala 转氨酶折叠上发展了 β 内酰胺酶活性(Peimbert 和 Segovia, 2003. *蛋白工程 (Protein Engg.)* 16, 27-35)。另外一个实例是该方法改变 NHR 人雌激素受体特异性，相对于天然配体，即 17β -雌二醇，所述受体更支持合成的配体，即 4,4'-二羟基偶苯酰(Chockalingam 等, 2005. 美国国家科学院学报 (Proc. Natl. Acad. Sci. USA) 102, 5691-5696)。

通过盐桥或二硫化物修饰进行蛋白质结构稳定性改造的实例：另外，尽管尚未详细研究这个方面，但是可以注意到也已经尝试了包括理性方法的蛋白质改造，以通过引入如特异性静电相互作用或其他另外的键诸如二硫键获得特定蛋白质的结构稳定性。所述尝试基于这样的知识，即表面盐桥(Anderson 等, 1990. *生物化学 (Biochemistry)* 29, 2403-2408)以及二硫键(Creighton, 1986. *酶学方法 (Methods. Enzymol.)* 131, 83-106)为蛋白质提供额外的稳定性。然而，所述理性尝试没有获得什么成功，正如通过 Perham 和同事的工作所例证地(Scrutton 等, 1998. *FEBS 通信 (FEBS Letters)* 241, 46-50)，他们通过设计将二硫键引入到谷胱甘肽还原酶中，以试图提高其稳定性，并且生成形成预期的二硫键但不显示额外的结构稳定性的活性酶。

我们能够发现的试图以某种方式组合一种蛋白质的结构稳定性和另一种（相关的）蛋白质活性的温度方案的蛋白质构建体的仅有的先前实例包括反复试验（trial-and-error）方法，其中重组了由源自两种不同的同源蛋白质的连续的残基序列段组成的结构域，以产生嵌合蛋白。我们已经能够发现四个制备所述嵌合蛋白的实例，两个涉及 β 葡萄糖苷酶，来自 Hayashi 和同事的工作(Singh 和 Hayashi, 1995. *生物化学杂志 (J. Biol. Chem.)* 270, 21928-21933; ; Goyal 等, 2001. *分子催化作用 B: 酶杂志 (J. Mol. Catalysis. B: Enzymatic)* 16, 43-51)，一个涉及柠檬酸合酶，来自 Danson 和同事的工作 (Arnott 等, 2000. *分子生物学杂志 (J. Mol. Biol.)* 304, 657-668)，和一个涉及抗生物素蛋白(Hytonen 等, 2007. 美国专利 7,268,216)。在第一个实例中，制备来自根癌农杆菌 (*Agrobacterium tumefaciens*) 和 *gilvus* 纤维弧菌 (*Cellvibrio gilvus*) 的同源- β 葡萄糖苷酶 (~37%的序列同一性；40%的序列相似性) 的嵌合体(Singh 等, 1995. 同前引用)。在第二个实例中，制

备来自根癌农杆菌和海栖热袍菌 (*Thermotoga maritima*) 的同源- β 葡萄糖苷酶的嵌合体(Goyal 等, 2001. 同前引用)。在第三个实例中, 制备来自嗜酸热原体 (*Thermoplasma acidophilum*) 和激烈火球菌 (*Pyrococcus furiosus*) 的同源柠檬酸合酶(Arnott 等, 2000. 同前引用)。在所有三个实例中, 目的是为了获得具有提高的酶稳定性和改变的最佳功能温度和 pH 的酶性质的嵌合体。在第四个实例中, 它是在专利文献中找到的(Hytonen 等, 2007. 同前引用), 鸡抗生物素蛋白的热稳定性是通过其称为 $\beta 4$ 的结构域被不同抗生物素蛋白-相关 (AVR) 蛋白的完整 $\beta 4$ 结构域替换得到提高的。

本发明的公开内容、本发明的新颖性、和‘嵌合体’方法与本发明提议的方法之间的区别:

在本发明中, 我们已经研究了序列变化对蛋白质表面的微观-结构和宏观-结构的影响 (来自蛋白质的进化比较), 其具体涉及使用所述变化, 通过设计 (而非通过非-理性的组合方法) 产生组合一种蛋白质的结构特征和源自不同生物体的另一种 (同源) 蛋白质的功能特征的新型蛋白。因此, 在本发明中, 我们的重点是对蛋白质表面的重新改造, 其具体涉及改造限定蛋白质酶活性和/或其他功能 (例如, 蛋白质-蛋白质相互作用) 的物理参数。

发明内容

我们叠合两种基于 β -折叠的 β 薄卷饼型 (jelly-roll) 折叠蛋白 RM Cell12A (SEQ ID NO:1) 和 TR Cell12A (SEQ ID NO:2) 的结构, 它们是纤维素-降解酶, 称为纤维素酶。我们证实这两种酶的多肽主链原子可与 1.1 埃的均方根偏差重叠。我们然后在这两种酶中构建约 60 个类似残基的对应组, 其包括它们的活性表面。在此, 活性表面定义为这两种酶中的扭曲的/弯曲的、结构-同源的 β 折叠的完整的溶剂-暴露表面, 其包含底物 (纤维素)-结合沟。认为 SEQ ID NO:1 是主体酶 (host enzyme), 且保留其大部分氨基酸序列, 我们然后在位于折叠的酶结构活性表面上的位置处将突变引入该序列中, 从而使所有所述残基被另一种 (客体) 酶结构中的结构-类似位置处使用的残基替换, 所述另一种 (客体) 酶结构具有由 SEQ ID

NO:2 表示的氨基酸序列。该替换的目的是为了将组成其活性表面的 60 个主要不连续残基的完整组（完全来自主体酶的多肽链）替换为组成客体酶的活性表面的结构-等价残基。该替换导致形成由 SEQ ID NO:3 表示的宿主酶突变形式。这实现在主体酶突变形式的折叠结构上对客体酶的活性表面的模仿，以使 SEQ ID NO:3 折叠，从而表现出 RM Cel12A 的结构稳定性特征和 TR Cel12A 的活性特征。由于 RM Cel12A 是嗜热性纤维素酶，且 TR Cel12A 是嗜中温纤维素酶，我们将 RM Cel12A 的该种突变形式（引入 TR Cel12A 的活性表面）称为中温活性热稳定纤维素酶（meso-active thermo-stable cellulase），即 MT Cel12A。利用 X-射线晶体学、圆二色谱学、质谱学、凝胶过滤色谱法和凝胶电泳、和纤维素酶活性测定法，在结构上和功能上表征 MT Cel12A。证明了 MT Cel12A 的确是一种中温活性热稳定酶，其组合 TR Cel12A 的活性表面和 RM Cel12A 的大部分结构稳定性特征。

本发明中提出的方法的新颖性：

总之，在现有技术和背景中，上述改造方法的所有成功应用涉及试图使用蛋白质改造（利用理性或组合的方法）来改变：（a）蛋白/酶活性位点关于对结合底物或其他配体（例如，辅酶）的优选性的化学特异性，或（b）结合底物或配体的热力学和动力学参数。相反，在本发明中，我们提出针对酶/蛋白质功能的物理特征（即，最佳结合的温度，或活性），而非与例如优选结合特定化学结构的底物或配体相关的功能的化学特征，而改变蛋白质/酶表面。我们提出使用理性改造方法而非组合的或杂交的方法，但是其是足够系统的，以至于具有适用于所有这样的蛋白质的高度可能性，所述蛋白质的功能部分（即其活性表面、或活性结构域、或亚-结构域）主要为特定类型的二级结构，即 β -折叠结构。在不干扰所述折叠的条件下将突变引入 β -折叠（在所述折叠完整的朝向溶剂的一侧）比将突变引入位于蛋白质表面上的 β -螺旋中更容易进行，因为在 β -折叠中交替的残基以相反的方向背离朝向，且“紧邻的”的残基彼此不相互作用或彼此相互干扰。在另一方面，在 β -螺旋中，“紧邻的”残基处于螺旋的同一侧或同一面，彼此非常大地互相影响，由于该原因，在不在直接相邻的残基中生成伴随（补

偿) 突变条件下进行的甚至单个突变可以极大地扰乱结构-形成和稳定性。我们的情形是突变作用的不可预知性对螺旋结构比对折叠结构更适用得多。我们提出利用关于所有- β 结构的结构/功能关系的详细理解, 在理性设计的 β 折叠表面上产生突变, 从而探究理性方法可以进行到什么程度。可以注意到, 不存在关于这种性质的实例, 并且因此没有在此引用任何内容。

‘嵌合体’方法和本发明之间的区别

如已经阐述地, 我们的目的之一是构建具有嗜热性(热稳定的)类似物的结构稳定性特征和嗜中温类似物作用于相同底物的功能特征的新型蛋白质/酶。而且, 如已经阐述地(在现有技术和背景中), 先前在该方向(即具有相似的总体目标)的唯一类型的尝试涉及交换包括连续残基序列段的完整结构域, 从而产生嵌合的蛋白。在嵌合蛋白中, 不存在在任何单独的二级结构或超二级结构元件(诸如 β 折叠)水平上重建蛋白表面, 而涉及来自两种执行相似功能的多-结构域蛋白的完全自主-折叠的结构域的混合。在蛋白质中, 通常, 公知结构域之间的相互作用水平总是远低于(就涉及的残基-残基接触的数目而言)结构域中结构元件之间的相互作用水平、或单一二级或超二级结构元件中的残基之间的相互作用水平。实际上, 从结构观点来看, 两种自主折叠的结构域通常甚至不直接相互作用, 而结构域中二级和超二级结构元件中的残基参与广泛且亲密的相互作用。因此, 将来自不同蛋白质的结构域组合到同一多肽链中并设法保持这两种结构域中的功能一般是简单且微不足道的工作, 因为关于结构域-结构域相互作用的关注少得多。

所述的连续序列段的简单(嵌合体-型)组合不能成功地进行重建单-结构域蛋白质, 因为在所述嵌合体中, 来自两种蛋白质的序列将位于表面上和内部中。不能迫使来自在两种不同蛋白质中预进化的二级结构元件的残基在任何蛋白质内部的两种结构的交界面处彼此接触和结合, 因为交界面的形成需要高度的形状互补性和化学相容性。简单的嵌合体不能避免在它们内部的结构-结构交界面处的形状互补性和化学不相容性的这种冲突。大概作为这样的结果, 不存在构建包括单结构域蛋白的任何嵌合体的已知

成功实例，在所述单结构域蛋白中，连续的氨基酸序列段被来自另一种蛋白质的等价序列段所替代。

我们的发明通过仅改造表面残基和仅在基于 β 折叠的结构内进行改造，避免了上述问题，并成功地组合了两种单结构域蛋白的特征。我们的方法成功的原因如下：(i) 彼此邻近的表面残基仅需要在蛋白质表面的一个限定的区域内与溶剂且彼此间相互作用，其中所有残基来自于同一蛋白质，并且因此具有参与相互作用的预-进化方案，从而构建所述表面。与在嵌合体中不同，这些残基不必须与另一组来自具有包埋的内部位置的不同蛋白质的残基接合，由此消除形状互补性和化学不相容性的问题。(ii) 我们使用结构元件（二级或超二级结构元件）的完整表面，诸如例如， β 折叠的完整的溶剂-暴露表面，而不是任何结构元件的一部分表面。这消除了结构元件表面内的形状互补性和化学相容性问题，并将该问题仅限制到所述表面的边缘，在所述边缘处其与不同结构元件表面的另一部分接触，在此处，在任何情形中，由于溶剂暴露，存在更高程度的灵活性和适应性。

因此，总之，现有技术中引用的嵌合体方法在以下方面与本发明提出的方法不同：

- 他们对相似酶的选择基于已知序列的酶之间的序列相似性。相反，我们的选择基于已知结构的酶之间的结构（主链）的叠合性。
- 他们对提高稳定性、最佳温度和最佳 pH 的尝试基于这样的假设，即这些特征确定地处于由总序列中的连续序列段限定的各个结构域中，没有提供支持该假设的证明，即所述特征确实是由单独的结构域-样结构限定的。相反，我们的方法不将特性归咎于连续的残基序列段。而是，我们认为最佳 pH 和温度特征是单独构成活性表面的溶剂-暴露残基（通过链折叠三维集合在一起的明确的不连续残基组）的相互作用和灵活性的函数。我们认为结构稳定性在很大程度上是构成蛋白质的疏水性中心的隐蔽残基的函数。因此，我们的新型蛋白质的构建涉及一起位于结构（而非位于序列！）中的不连续残基的突变。我们基于结构原则和在文献中首次说明的理解，制造这样的突变，即通过同

源嗜中温酶的活性表面仅移植热稳定酶的活性表面。

- 作为上述方法中基本区别的结果，Hayashi 和同事(Singh 和 Hayashi, 1995. 生物化学杂志 (*J. Biol. Chem.*) 270, 21928-21933; Goyal 等, 2001. 分子催化作用 B: 酶杂志 (*J. Mol. Catalysis. B: Enzymatic*) 16, 43-51)以及 Danson 和同事(Arnott 等, 2000. 分子生物学杂志 (*J. Mol. Biol.*) 304, 657-668)设法仅构建具有这样的稳定性和功能特征的嵌合体，所述嵌合体是制备所述嵌合体所来源的两种酶的序列的中间产物，或以完全不可预期的方式与两种亲本序列完全不同。相反，我们的方法提出理性和可预期地混合一种亲本的功能特征（即精确的最佳 pH 和温度）和另一种亲本的稳定性特征，从而衍生在很大程度上保持这些特征的蛋白。
- 在所述嵌合体中，不基于在蛋白质的内部或蛋白质的表面上的残基位置选择来自祖先酶的残基。这使得该方法仅对简单混合来自多-结构域蛋白质的编码结构域的序列有效。相反，我们的方法集中于基于结构选择溶剂-暴露的残基。

发明目的

本发明的目的是提供包括嗜热性蛋白的结构稳定性特征和嗜中温蛋白的活性特征的重组中温活性热稳定蛋白及其合成方法，其中主要不连续的氨基酸组被结构-同源蛋白的结构等价部分中存在的不同的不连续氨基酸组代替。

- 而且，本发明提供用于系统改变任何主要由 β 结构构成的蛋白质的功能表现的方法学。
- 本发明的另一个目的是将一种蛋白质的完整表面或所述表面的一部分取代/移植到具有高度叠合性的主链的同源蛋白上。
- 另一个目的是将来自嗜中温蛋白的活性表面移植到嗜热性蛋白上，从而生成重组的热稳定蛋白。
- 另一个目的是生成表现出嗜热性蛋白的稳定性特征和嗜中温蛋白的活性特征的重组热稳定蛋白，由此提供用于组合经天然进

化一般不组合的结构和功能属性的方法学。

发明概述

本发明涉及基于 β 折叠的蛋白质结构，并特别集中于检验蛋白质表面可变性的酶。这是通过将一种蛋白的完整表面取代/移植到重叠主链轨迹的同源蛋白表面上（即，完整表面移植）或仅改变参与催化和/或底物/配体结合的表面区域（即，活性表面移植）而进行的。本发明还涉及选择性混合和匹配不连续的和/或连续的残基，从而完成移植包括 β 链和/或来自 β 折叠结构的中间环的表面。本发明还涉及进行表面改造，其涉及将嗜热性功能特征的蛋白转化为嗜中温功能特征的蛋白，同时保持嗜热性蛋白的结构特征。

因此，本发明提供包括嗜热性蛋白结构稳定性特征和嗜中温蛋白活性特征的重组的中温活性热稳定蛋白。

在本发明的一个实施方案中，所述重组中温活性热稳定蛋白包括具有单结构域结构的蛋白，其中主要不连续的氨基酸组被在结构-同源蛋白的结构等价部分处存在的不同的不连续氨基酸组代替。

在本发明的一个实施方案中，将“客体”或“供体”祖蛋白的一部分或整体理论上移植到“主体”或“受体”祖蛋白的结构核心上。

然而，在本发明的另一个实施方案中，所述客体和主体祖蛋白具有同源结构和相同的功能。

然而，在本发明的另一个实施方案中，所述客体祖蛋白是嗜中温蛋白，和所述主体祖蛋白是嗜热性蛋白。

在本发明的一个实施方案中，所述同源客体和主体祖蛋白的功能-活性表面区域由基于 β 折叠的二级结构组成。

在本发明的一个实施方案中，使用的嗜热性主体祖蛋白和嗜中温客体先祖分别具有由 SEQ ID NO:1 和 SEQ ID NO:2 表示的氨基酸序列，并且重组的中温活性热稳定蛋白包括由 SEQ ID NO:3 表示的氨基酸序列。

在本发明的另一个实施方案中，所述重组中温活性热稳定蛋白具有下列特征：

1. 分子质量： 20-30 千道尔顿；

2. 残基数目： 约 200-300 个；
3. 等电点： 在 4-8 范围内；
4. 最佳活性 pH： 在 4-8 范围内；
5. 解链温度 (T_m)： 80-95℃，和
6. 最佳活性温度 (T_{OA})： 在 30-60℃ 范围内。

在本发明的另一个实施方案中，所述蛋白或其先祖是属于水解酶组的单-结构域 β 折叠蛋白的结构类的酶，并选自由纤维素酶、木聚糖酶、淀粉酶和蛋白酶组成的组，优选地，是纤维素酶。

在本发明的另一个实施方案中，所述重组中温活性热稳定蛋白还可以起生物催化剂的作用。

在本发明的另一个实施方案中，用于设计重组中温活性热稳定蛋白的方法包括主体祖蛋白的结构稳定性特征和不同的且结构同源的客体祖蛋白的活性特征。

在本发明的一个实施方案中，所述结构-同源的主体和客体祖蛋白包括这样的结构，在所述结构中由位于基于 β 折叠的二级结构上的残基提供底物-结合和催化功能。

在本发明的一个实施方案中，用于涉及重组中温活性热稳定蛋白的方法包括下列步骤：

a. 叠合所述客体和主体祖蛋白的结构，并在多肽主链水平上，优选地，在包括所述底物-结合和催化功能的结构区域中，评估它们结构的叠合性程度，

b. 通过使用步骤 (a) 的信息，生成所述主体和客体祖蛋白的蛋白序列以及编码 DNA 序列的基于结构类似的序列对比，

c. 鉴定在主体祖蛋白序列中构成基于 β 折叠的‘活性表面’的特定氨基酸残基，并且使用获自步骤 (b) 的对比序列以在所述客体祖蛋白中鉴定相应类似的不连续残基和残基组，

d. 确定在来自所述主体和客体祖蛋白构成所述两种蛋白的活性表面的两个相应氨基酸残基组中相同的特定残基，并且消除所述相同的残基，从而获得在所述两种蛋白中的结构-等价位置处存在的一组相应的不同残基，

- e. 在基于 β 折叠的‘活性表面’中的位置处突变所述主体祖蛋白，所述位置与所述客体祖蛋白的相应残基不同，
- f. 将在所述主体祖蛋白中存在的残基替换为所述客体祖蛋白中存在的残基，导致构成新型中温活性热稳定蛋白，
- g. 通过已知的重组 DNA 方法生物合成该新型中温活性热稳定蛋白，随后通过已知方法分离和纯化，和
- h. 通过活性测量证实在该理想的中温活性热稳定蛋白中的主体祖蛋白的结构稳定特征和客体祖蛋白的物理和化学活性特征。

在本发明的一个实施方案中，构成产物中温活性热稳定蛋白的结构核心的氨基酸残基源自所述两种祖蛋白中的一种，而构成该产物蛋白的一部分或完整表面的残基源自另一种祖蛋白。

在本发明的一个实施方案中，所述两种祖蛋白是结构同源的，它们的主链原子的坐标以 0.5-2.5 埃的均方根偏差 (RMSD) 叠合。

在本发明的一个实施方案中，使用的主体和客体祖蛋白中的最佳功能和结构解链温度相差 5°C 以内。

然而，在本发明的另一个实施方案中，所述两种祖蛋白的结构-同源‘活性表面’区域主要由基于 β 折叠的二级结构组成。

然而，在本发明的另一个实施方案中，使用的客体先祖是 SEQ ID NO:1 的嗜热性蛋白 RM Cel12A，且使用的主体先祖是 SEQ ID NO:2 的嗜中温蛋白 TR Cel12A。

在本发明的一个实施方案中，构成产物蛋白的结构核心的氨基酸残基源自 RM Cel12A，而构成产物蛋白一部分表面的残基源自 TR Cel12A。

在本发明的一个实施方案中，RM Cel12A 和 TR Cel12A 蛋白是结构同源的，它们的主链原子的坐标以 0.5-2.5 埃的均方根偏差 (RMSD) 叠合。

在本发明的一个实施方案中，RM Cel12A 和 TR Cel12A 蛋白中的最佳功能和结构解链温度相差 5°C 以内。

在本发明的一个实施方案中，RM Cel12A 和 TR Cel12A 的结构-同源性‘活性表面’区域主要由基于 β 折叠的二级结构组成。

在本发明的一个实施方案中，SEQ ID NO:3 的产物 MT Cel12A 是通过将主体先祖活性表面包括的残基替换为客体先祖活性表面包括的结构-相

似残基，而衍生自主体先祖的。

在本发明的另一个实施方案中，获得的新型中温活性热稳定(MT Cel2A)产物具有嗜中温客体先祖 TR Cel12A 的最佳活性温度和同源嗜热性主体先祖 RM Cel12A 的结构稳定性。

在本发明的另一个实施方案中，所述中温活性热稳定蛋白(MT Cel2A)具有下列特征：

1. 分子质量： 20-30 千道尔顿；
2. 残基数目： 约 200-300 个；
3. 等电点： 在 4-8 范围内；
4. 最佳活性 pH： 在 4-8 范围内；
5. 解链温度 (T_m)： 80-95°C，和
6. 最佳活性温度 (T_{OA})： 在 30-60°C 范围内。

然而，在本发明的另一个实施方案中，所述方法在不改变所述酶活性的化学特征或所述底物的化学限定的条件下，有效调整酶的物理功能特征。

然而，在本发明的另一个实施方案中，所述方法有效用于重组来自两种非常不同的生命领域的酶结构稳定性与蛋白质稳定性和活性特征。

在本发明的一个实施方案中，所述中温活性热稳定蛋白有效用于纺织工业的应用中，诸如牛仔布 (denim fabrics) 的石洗 (stone-washing)。

附图和表格简述

图 1: RM Cel12A (红色),和 TR Cel12A (蓝色)之间的结构相似性和相异性。

图 A 显示以 1.1Å 的 RMSD 叠合的多肽主链。注意夹心的上和下折叠，以及由顶部折叠形成的凹沟构成纤维素-结合 ‘活性表面’。图 B 和 C，分别地，显示 TR Cel12A 和 RM Cel12A 表面的顶视图，强调 19-21 Å 长的纤维素-结合沟。注意这两个沟的显微结构特征的相异性。

图 2: RM (红色)、TR (蓝色)、和建模的 MT (红-蓝) Cel12 的基于-结构的序列对比和表面描绘。

图 A. 绿色框突出由 SEQ ID NO:1 表示的 RM Cel12A (1H0B)和由 SEQ

ID NO:2 表示的 TR Cel12A (10A2)中的‘活性表面’残基,其为保守的(三行)或不保守的(两行)。在最上一行的单个残基表示在预测的和获得的序列之间的差异,在文献中报告的关于 1HOB 的所述残基显示在我们通过对我们的克隆 (SEQ ID NO:1) 测序获得所述残基的上方。在中间一行中提到的序列是 MT Cel12A (SEQ ID NO:3)的序列。注意来自 β 链的交替‘活性表面’残基簇。图 B. RM (红)、TR (蓝)、和建模的 MT (红-蓝) Cel12A 的表面的侧视图。对 MT Cel12A 模型进行彩色-编码,从而指示来自 RM 和 TR 先祖的残基的不同来源。

图 3: RM Cel12A 和 MT Cel12A 在 pH 8.0 和 pH 5.0 时的四级和二级结构。

在缺乏(黑线)和存在(红线) 100 mM NaCl 条件下的洗脱液 Superdex-75 层析图,将关于 MT Cel12A 的显示在图 A (pH 8.0)和图 B (pH 5.0)中,并且将关于 RM Cel12A 的显示在图 C (pH 8.0)和图 D (pH 5.0)中。注意在缺乏盐的条件下在 1.23 ml 处(单体)的 MT 和 RM Cel12A 的洗脱,和后来在存在盐和低 pH 的条件下的体积。在存在 100 mM NaCl 条件下在 pH 8.0 (黑线)和 pH 5.0 (红线)下的酶的远-UV CD 光谱显示在图 E (MT Cel12A)和图 F (RM Cel12A)中。在缺乏盐的条件下获得相同的光谱(未显示)。

图 4: RM Cel12A 和 MT Cel12A 在 pH 8.0 和 pH 5.0 时的热稳定性。

218nm 处的平均残基椭圆率的温度-依赖性变化显示在图 A (RM Cel12A, pH 5.0)、图 B (MT Cel12A, pH 5.0)、图 C (RM Cel12A, pH 8.0)和图 D (MT Cel12A, pH 8.0)中。注意在所有情况下,结构改变仅发生在 85 °C 以上,且 MT Cel12A 在 pH 8.0 和 pH 5.0 时的 RM Cel12A-样热稳定性。

图 5: MT Cel12A 和 RM Cel12A 的温度-和 pH-依赖性活性。

图 A: 作为温度的函数,在 pH 5.0 测量的 MT Cel12A (空心三角形)和 RM Cel12A (实心方块)的活性变化。图 B: 作为 pH 的函数,在 50 °C 测量的 MT Cel12A (空心三角形)和 RM Cel12A (实心方块)的活性变化。平均值和标准误差条各基于 5 次实验。注意 MT Cel12A 具有与 TR Cel12A 接近的最佳温度和 pH (本图),即使它具有与 RM Cel12A 相当的结构稳定性(图 4)。

图 6: 在不同温度和 pH 值处比较 RM Cel12A 和 MT Cel12A 活性。

纵坐标表示关于 DNSA 反应的还原糖在 550nm 处的光学密度。为了允许更好的视觉比较, 关于 RM Cel12A 数据的刻度 (0-8) 与关于 MT Cel12A 的那些 (0-1.5) 不同地显示。

图 7: 比较 MT Cel12A(3B7M)测定的结构和 RM Cel12A(1H0B)与 TR Cel12A (1OA2)的已知结构。

图 A: 叠合 MT Cel12A (蓝)、RM Cel12A (红) 和 TR Cel12A (绿)的主链。图 B: 由氨基酸极性彩色-编码的 RM Cel12A (左上)、TR Cel12A (右上) 和 MT Cel12A (中下)的活性表面沟的视图。图 C: 针对它们的主链 (如细带表示地) 叠合的 MT Cel12A (绿) 和 TR Cel12A (蓝)的活性表面侧链。图 D: 与图 C 中显示的相同的活性表面侧链, 不存在主链 (为了清楚)。图 E: MT Cel12A (绿)、RM Cel12A (红)和 TR Cel12A (蓝)侧视图的所有-原子表面表示法, 显示 MT Cel12A 与两种亲本酶的的相似性。

图 8: 参与 β 折叠形成的链中的残基的交替几何排列, 和由来自折叠中相邻链的侧链形成表面。

利用软件 PYMOL 生成来自 RM Cel12A 的三条链的图示。图 A: 作为 β 折叠一部分的链中的交替残基以相反方向背离朝向该折叠的平面。图 B: 来自 β 折叠中的相邻链的、以相同方向背离朝向该折叠的侧链彼此相邻。图 C: 图 B 中显示的侧链的这些原子彼此相互作用。图 D: 图 C 中显示的侧链的原子充分相互作用, 以至于足以形成表面。

图 9: RM Cel12A 和 TR Cel12A 之间结构同源性和构成这两种酶的活性表面的 β 折叠的视觉描述。

图 A: 主链的重叠。尽管在 RM Cel12A 和 TR Cel12A 之间仅存在~28% 的序列同一性, 但是其结构高度同源, 并可以以 1.1Å 的 RMSD 叠合。由两种弯曲的 β 折叠组成的蛋白质夹入到 β 薄卷饼型折叠中, 纤维素-结合沟位于上方折叠的凹陷表面中。RM Cel12A 主链显示为黄色。其活性表面显示为青色。TR Cel12A 主链显示为红色。图 B: RM Cel12A 的活性表面。显示在所述两种蛋白的叠合主链上的是 RM Cel12A 的活性表面, 其来自构成纤维素-结合沟的上方 β 折叠的残基组成。RM Cel12 主链显示为红色, 且 TR Cel12A 为蓝色。图 C: TR Cel12A 的活性表面。显示在所述两

种蛋白的叠合主链上的是 TR Cel12A, 其由来自构成纤维素-结合沟的上方 β 折叠的残基组成。RM Cel12 主链显示为红色, 且 TR Cel12A 为蓝色。该嗜中温活性表面和嗜热性同系物的活性表面的视觉比较显示本工作中试图进行的表面-改造的深度的理论“切除”视图。

图 10: TR Cel12A 的“温度对比活性”、“pH 对比活性”和“温度对比结构含量”特征。

图 A 和 B: 如 Karlsson, J., Siika-aho, M., Tenkanen, M. & Tjerneld, F. (2002). 生物技术杂志 (*J. Biotechnol.*) **99**, 63-78 报告地, TR Cel12A 活性随温度和 pH 的变化。图 C: 如 Sandgren, M., Gualfetti, P.J., Shaw, A., Gross, L.S., Saldajeno, M., Day, A.G., Jones, T.A. & Mitchinson, C. (2003). 蛋白质科学 (*Protein Science*) **12**, 848-860 报告地, TR Cel12A CD 信号 (结构含量) 随温度的变化。本附图中显示的图来自于以上引用的论文。

图 11: RM Cel12A 和 MT Cel12A 的 DNA 和蛋白质序列。

最上一行——M Cel12A 蛋白序列。第二行——RM Cel 12A DNA 序列。第三行——MT Cel12A DNA 序列。第四行——MT Cel12A 蛋白序列。注意: 仅显示残基编号。未显示碱基编号。显示的残基编号是来自原始参考文献中真实序列的那些, 忽视缺口。在标记物后, RM 序列以 Met (残基 1) 起始, 其未显示。

图 12: 合成编码 RM Cel112A (通过来自基因组 DNA 的 PCR) 和 MT Cel12A (通过利用诱变引物的 SOE-PCR) 的基因的示意图。编号是分配的引物编号 (见表 4)。星号标记表示引物是诱变的。诱变的引物结合 DNA 碱基变化, 所述变化设计为 (i) 改变 RM 残基为结构类似的 TR 残基, (ii) 包括与 RM 中无类似性的来自 TR 的环, (iii) 通过破坏备选 (不理想的) 引物-模板缔合的可能性进行沉默突变, 从而最优化 PCR 的可行性, 和 (iv) 进行沉默突变, 从而破坏二级结构, 认为所述二级结构可能在基因表达过程中在 mRNA 水平影响翻译效率。注意: 利用引物 1N 和 12, 从海洋红嗜热盐菌 (*R.marinus*) 基因组 DNA 中扩增关于 RM Cel 12A 的基因。利用诱变引物 1-11 和不包括突变的引物 12, 通过剪接编码 RM Cel12A 的基因的合适的突变部分构建关于 MT Cel12A 的基因。

图 13: 显示来自 PCR 和 SOE-PCR 反应的扩增的凝胶。凝胶显示 a) 通过 PCR 从来自海洋红嗜热盐菌的基因组 DNA(反应号 1)生成编码 RM Cel12A 的基因, 和 b) 编码 MT Cel12A 的基因的合成中涉及的步骤, 其涉及来自编码 RM Cel12A 的基因的区域诱变 PCR 扩增和通过剪接重叠延伸 (splicing-by-overlap-extension, SOE) PCR (反应号 2-11) 装配所述区域。

图 14: DNA 测序电泳图验证编码 MT Cel12A 的基因的正确构建。

图 15: 显示 MT Cel12A (图 A&B) 和 RM Cel12A (图 C&D) 的非-变性纯化的凝胶。

图 A: 泳道 1—沉淀。泳道 2—溶胞产物 (在 10mM 咪唑中)。泳道 3—流出液。泳道 4—洗涤液 (在 20mM 咪唑中)。泳道 5—分子量标记 (从上到下: 分别为 116, 66, 45, 35, 25, 18.4, 14.4 kDa)。泳道 6-9—洗脱在 1M 咪唑中的 MT Cel12A 级分。图 B: 泳道 10-15—洗脱在 1M 咪唑中的 MT Cel12A 级分。泳道 16—分子量标记 (从上到下: 分别为 116, 66, 45, 35, 25, 18.4 kDa)。图 C: 泳道 1—沉淀。泳道 2—溶胞产物 (在 10mM 咪唑中)。泳道 3—流出液。泳道 4—洗涤液 (在 20mM 咪唑中)。泳道 5—分子量标记 (从上到下: 分别为 116, 66, 45, 35, 25, 18.4, 14.4 kDa)。泳道 6-9—洗脱在 1M 咪唑中的 MT Cel12A 级分。图 D: 泳道 10-15—洗脱在 1M 咪唑中的 MT Cel12A 馏分。泳道 16—分子量标记 (从上到下: 分别为 116, 66, 45, 35, 25, 18.4 kDa)。

图 16: RM Cel12A 和 MT Cel12A 的 MS 特征。

利用 IgG 校准、以具有 >100 ppm 的精确性的线性模式在 ABI Voyager DE-STR MALDI-TOF 质谱仪上收集数据。RM Cel12A 的预测质量是 26215.93 Da。MTC Cel12A 的预测质量是 25037.01 Da。注意 1: 观察到的质量在对这些质量预测的误差范围内。注意 2: 对于 m 值 $m=1, z=2$ & $m=1, z=1$ & $m=2, z=1$ & $m=3, z=1$ 观察到 m/z 峰。

图 17: 凝胶过滤校准和以不同浓度运行 MTC Cel12A。该附图显示 MT Cel12A 样品的堆叠的凝胶过滤洗脱层析图, 每份样品为 50 μ l 体积, 所述样品是在将蛋白质浓度稀释为插入数据中显示的数值后上样的。四种样品的吸光度值范围明显不同, 并且因此在标准化后显示它们堆叠, 从而

促进观察洗脱体积的变化，如果存在变化的话。利用 50 mM Tris pH 8.0 平衡样品和 SMART Superdex-75 柱。观察到的结果是 MT Cel12A 的稀释不引起凝胶过滤洗脱体积（已知与分子流体力学体积有关）的任何变化。在稀释时，任何缔合、或解离的缺乏，以及在以上校准图中插入其洗脱体积(1.23 ml)，表明 MT Cel12A 是~25 kDa 的稳定单体。

表 1: 通过 LSQMAN 叠合程序生成的相似残基对(具有 RMSD 数据)。该表仅显示存在结构叠合的残基对。以上未显示未叠合的环。

表 2: 活性表面移植的详细信息。该表显示为 RM Cel12A 和 TR Cel12A 各自的活性表面沟贡献侧链的所有结构-类似的残基对。构成 β 链的残基显示为绿色突出标记的格中。在经过移植保守的残基显示为黄色。其他残基来自于连接链的环，其侧链朝向溶剂。重要的注意：上表仅列出上方折叠和指向溶剂的表面环上存在的残基，而没有列出在上方折叠上向下指向下方折叠的残基。

表 3: 基因片段合成和装配的序列（PCR 和 SOE-PCR 条件）。

表 4: 用于基因合成的引物序列。

表 5: 序列信息差异表。

表 6: 蛋白质的结构-生物化学和物理-化学性质。对中温活性热稳定酶 (MT Cel12A)测量的性质与由文献已知的嗜中温(TR Cel12A)和嗜热性(RM Cel12A)先祖的性质一起显示在比较表格中。

发明详述

本发明提供包括嗜热性蛋白结构稳定性特征和嗜中温蛋白活性特征的重组中温活性热稳定蛋白及其合成方法，其中包括嗜热性蛋白表面的底物-结合和催化-活性区域的主要不连续的氨基酸组被结构-同源的嗜中温蛋白中的结构等价位置处存在的不同的不连续氨基酸组代替。

我们的发明提出重建酶“活性表面”。认为酶“活性表面”在酶反应的任何阶段，可以包括所有包括与催化作用直接有关的活性位点的残基以及结合底物或接触底物原子有关的残基。值得注意的是，即使在与底物-类似物结合的酶的结构在原子水平详细已知的情形中，例如通过 X-射线晶体学技术已知，人们却从来不能确信哪些其他残基（除在晶体结构中实

际看到与底物接触的那些外)也在酶处理过程中与底物接触,因为许多酶被认为在它们结合和催化循环过程中改变构象。因此,假定仍然不存在确定知晓任何酶完整活性表面的全部范围的方法,实际上,我们应该将“活性表面”定义为酶中参与底物-结合或催化作用的特定的二级结构(或超二级结构)元件的与溶剂接触的全部残基组。因此,例如,如果底物-结合和催化残基处于基于 β 折叠的结构的溶剂-暴露面,则我们应该将指定所述特定 β 折叠的完整的溶剂-暴露面为“活性表面”。

本发明提出并且示范用于酶活性表面的“表面重建”或“表面重新改造”的理性方法。所述重建涉及对两种具有广泛叠合的主链(“广泛叠合的”定义为具有 0.5-2.5 埃均方根偏差的主链原子叠合性)的结构-同源酶的残基使用的比较。所述比较一旦完成,随后用另一种酶的完整活性表面对一种酶中包括一个完整活性表面的所有残基进行基于残基-对-残基(residue-by-residue)突变的替换,——我们称为“表面移植”的方法,因为其导致活性表面从一种酶有效移植到另一种酶。如果以这样的方式选择在它们的结构稳定性和功能性最佳条件方面不同的两种酶,则所述“移植”作用导致基于折叠构建新型酶生物催化剂,所述酶生物催化剂结合“主体”酶的结构稳定性特征和“客体”酶的活性特征,所述主体和客体是进行结构同源性比较的两种酶。因此,所述主体和客体酶有效地成为新型酶生物催化剂的祖先酶,因为客体酶促成新型酶结构的活性表面,且主体酶促成所有其他残基和主要部分。

公知所述活性表面的形成是形成酶的总体三维结构的结果,即其依赖于酶的折叠。由于活性位点构成活性表面的子集,所以断定所述活性表面也是由于折叠形成的。然而,重要的是理解通过折叠仅形成几乎-正确的三维结构不暗示任意酶的活性位点或“活性表面”的正确形成和功能执行。这是因为可以在结构上充分干扰所述活性位点引起终止执行功能,而不以任何可辨别的方式干扰酶的总体结构,正如 Tsou 和同事关于通常在整体解折叠前可以观察到失活作用的许多酶的解折叠和折叠行为所示范地,这提示所述活性位点在构象上比其余的分子更灵活(Tsou, 1993. 科学(*Science*) 262, 380-381; Tsou, 1995. *Biochim. Biophys. Acta* 1253, 151-162; Yang 和 Tsou, 1995. 生物化学杂志(*Biochem J.*) 305, 379-384)。因而断定

这对活性表面也是正确的，即活性表面的形成和功能执行也可以表现出脱离展示所述表面的酶的折叠的一些自主性。本发明提出并示范用于验证和利用所述活性表面脱离酶的总体三维结构的支持支架的自主性的理性方法，这是通过显示可以怎样在酶中通过由多（主要）不连续残基替换实现的“移植”容易地模拟活性表面而进行的。

本发明示范利用具有 β 薄卷饼型折叠结构的酶进行活性表面移植的可行性；然而，本发明提出本发明中开发和公开的特殊的理性方法可以用于利用任何两种结构-同源酶进行所述活性表面移植，所述酶使用基于 β 折叠的活性表面，所述基于 β 折叠的活性表面使用以 0.5-2.5 埃的 RMSD 叠合的主链原子。

重建 β 折叠结构的溶剂-暴露面

在基于 β 折叠的蛋白，诸如 β 薄卷饼型折叠酶中，所述活性表面由 β -折叠形成。我们的发明的一个延伸是如果考虑到这样的事实，即 β -链上交替的残基以相反方向背离朝向，则可能突变酶的完整的基于 β 折叠的表面。那些从折叠内指向溶剂的残基通常形成所述活性表面，而指向内部的那些与另一个折叠的残基疏水性相互作用，这两组残基之间不存在相互作用。因此，我们的情况是从 β 折叠中去除朝向溶剂的残基组并用来自结构-相似（同源）酶的等价的朝向溶剂的残基组替换它们是可行的，即换言之，将一种酶的完整活性表面“理论上移植”到另一种酶上，如我们在此提出的在 β 薄卷饼型折叠酶中，或实际上在任何基于 β 折叠的表面中，所述表面提供分别处理隐蔽残基的机会，所述隐蔽残基主要参与由在蛋白表面上暴露和参与其功能的那些确定酶的总折叠和结构稳定性。

移植可行性的结构分析

我们的移植方法仅适用于基于 β 折叠的蛋白结构。由于 β 折叠中处在延伸构象中的肽键的反式构型，链中任意残基的侧链以与其在每一侧的两个紧邻残基完全相反的方向朝向背离折叠的方向。折叠中的交替残基组因此以相反方向背离朝向，并对其侧链发展完全独立的包装流程，这两面彼此之间具有非常小的任意相互影响范围。在优选的平面折叠中，同一面上

的残基有时可以太过远离，以至于不能有效相互作用形成表面；因此，在多-折叠结构中，残基通过堆叠在一起的折叠获得刚性构象，这容许朝向的侧链组之间的相互作用。然而，由长链构成的 β 折叠很少是完全平面的，大多数天然是弯曲的。尽管所述弯曲的折叠也可以堆叠（如，例如发生在 Cel12A 酶的顶部和底部折叠之间），但是一组堆叠的折叠中的顶部折叠的凹入、溶剂-暴露面上的残基潜在地可以相互作用，从而形成自主包装流程（图 1A-1C）和与所述堆叠中其他折叠有效隔离的表面。这造成选择性重建顶部折叠的凹面的可能性（图 1A，2B）。我们继续对形成 RM Cel12A 活性表面的顶部折叠进行重建，如下文详述。

RM Cel12A 和 TR Cel12A 共享约 28%的氨基酸序列同一性。两种结构的详细检验容许鉴定在两种蛋白质中的类似残基位置。关于每种蛋白质中的每对类似残基，我们鉴定具有参与构成或毗邻纤维素-结合沟的链的主链原子的那些，并然后分别将这些指定为具有这样的侧链，所述侧链是：（a）向上朝向所述沟，（b）向下（离开所述沟的方向）朝向下方的 β 折叠，或（c）落入分隔链的环区域内。然后，从所述列表中除去朝下的残基，它们是不具有参与沟形成的侧链的环结构中存在的残基。剩下的残基，以及在每种蛋白中在其他蛋白质中没有负体的某些沟-组成残基显示在表 2 中，其还总结了关于哪些残基位置是保守的，和哪些位置构成结构-类似的非-相同残基对的详细信息。在表 2 中，可以观察到来自交替残基位置的数组侧链存在于单独的“框”中，其中这些源自相同的 β 链，尽管构成所述表面的残基是由整个蛋白质序列获得的（图 2A）。

总之，可以观察到由 RM Cel12A 中的总共 66 个残基，和由 TR Cel12A 中的 57 个位于结构-类似位置处的残基形成所述沟本身和毗邻所述沟的上升区域。仅 29 个参与形成 TR Cel12A 活性表面的残基由 RM Cel12A 中的结构-类似负体的不-相同的组表示。此外，RM Cel12A 中的两个环，即分别地，12 个残基 (Cys66-Leu77)和 4 个残基 (Ser115-Gly118)，在 TR Cel12A 中被各自 3 个残基的非-结构-同源环[(Ile67-Gln69)和 (His107-Thr110)]替换。而且，TR Cel12A 中与 RM Cel12A 中的 Asp160 和 Trp161 对应的类似位置之间存在残基插入(Ala153)。因此（见表 2），RM Cel12A 和 TR Cel12A

活性表面之间的区别由来自 TR Cel12A 的总共 36 个残基 [29 (类似的) + 3 (来自一个环) + 3 (来自另一个环) + 1 (来自插入)] 限定。

因此, 我们改造的目的是在这 36 个位置处用 TR Cel12A 残基替换 RM Cel12A 残基。这两种酶的序列, 以及结合这 36 个残基变化的重新设计的海洋红嗜热盐菌酶的序列显示在图 2A 中, 所述重新设计的海洋红嗜热盐菌酶命名为“中温活性热稳定” Cel12A, 或 MT Cel12A。编码 MT Cel12A 的基因的合成是通过多突变-引入的聚合酶链反应 (PCR) 步骤, 以及随后通过重叠延伸 PCR(SOE-PCR) 程序由逐步剪接连接所获得的扩增子而完成的。克隆引入 6xHis N-末端亲和标记物的基因构建体, 过表达并且纯化, 以及类似地克隆、过表达和纯化未修饰的 RM Cel12A 作为对照, 其蛋白质质量是通过 MALDI-TOF 质谱法证实的。

合成 MT Cel12A

合成编码 MT Cel12A 的基因是通过多突变-引入的聚合酶链反应 (PCR) 程序, 以及随后通过重叠延伸 PCR(SOE-PCR) 程序由逐步剪接装配所获得的扩增子而实现的[见图 12(SOE PCR 流程)、表 3(SOE 装配详细信息)、表 4 (使用的引物的详细信息)、图 13 (PCR 结果)、图 14 (DNA 序列电泳图)、和表 5 中的序列的一组差异]。克隆、过表达和纯化引入 6xHis 亲和标记物的基因构建体, 以及类似地克隆、过表达和纯化未修饰的海洋红嗜热盐菌 Cel12A (RM Cel12A) 作为对照分子 (图 15)。这两种蛋白的质量是通过 MALDI-TOF 质谱法证实的 (图 16)。

MT Cel12A 如 RM Cel12A 一样地折叠

通过非-变性条件下的相同程序纯化的 MT Cel12A (图 3A) 和 RM Cel12A (图 3C) 在 pH 8.0 的凝胶过滤色谱法过程中表现出相同的洗脱体积, 这提示单体状态。在存在盐 (100 mM NaCl), 和 pH 5.0 下, MT Cel12A 在洗脱中显示出相同的延迟 (图 3A, 3B), 这说明与柱基质相互作用的疏水性残基簇可能的暴露; 在存在盐, 和 pH 5.0 下, 对于 RM Cel12A 也观察到所述延迟的洗脱 (图 3C, 3D)。以四种不同样品浓度、三个十进制量

级范围内层析的 MT Cel12A, 以完全相同的洗脱体积洗脱, 这进一步证实其为单体, 并在浓度升高时不经历缔合。

MT Cel12A (图 3E)和 RM Cel12A (图 3F)的远-UV 圆二色谱彼此类似, 并是典型的 β_1 型(β 折叠)二级结构, 其在 216nm 处具有相似信号强度的特有负波段最大值。在从酸性转移到碱性 pH 时, 酶均不经历任何结构改变 (图 3E、3F)。RM Cel12A 光谱在~225-235 nm 区域内显示出另外的负波段, 该负波段在 MT Cel12A 光谱中较不明显, 这提示可能存在一些微小的结构差别。

MT Cel12A 从 RM Cel12A 继承其 T_m

通过监测作为 25-98°C 温度升高函数的平均残基椭圆率 (MRE) 的减小, 评估 MT Cel12A 和 RM Cel12A 的热稳定性。MT Cel12A 表现出~93°C 的 T_m , 其显示结构在 pH 5.0 完全解链 (图 4B), 和在 pH 8.0 不完全解链 (图 4D), 由此证明其具有与 RM Cel12A 相似的结构 (热) 稳定性水平, 所述 RM Cel12A 表现出 96°C 的 T_m , 其显示结构在 pH 8.0 近似-完全解链 (图 4C), 和在 pH 5.0 不完全解链 (图 4A)。MT Cel12A 的极端热稳定性明显归功于其使其大多数残基, 即其完整的疏水性核心和其大部分表面, 来源于热稳定的 RM Cel12A。两种酶在两种不同 pH 值下加热时完全解折叠的事实可能可以阐释如下。对于 RM Cel12A 在 pH 5.0 可获得的高度稳定性, MT Cel12A 在该 pH 下不能获得, 因为 MT Cel12A 具有在 pH 5.0 时发展为其最高活性 (且大概, 因此, 最高构象灵活的) 状态的活性表面; 该活性表面在 pH 5.0 时的热不稳定性可能起始继续完成的解链过程。相同的论据可能适用于 RM Cel12A, 其具有在 pH7.0 时发展为最佳活性的活性表面, 在 pH 8.0 或 pH 6.0 具有相似的活性, 但在 pH 5.0 具有低得多的活性。因此, 每种酶似乎更易于在其拥有的活性表面的最佳活性 pH (或接近该 pH) 下热解链。

MT Cel12A 从 TR Cel12A 继承其 T_{oa}

如作为恒定 pH5.0 下温度的函数 (图 5A); 作为恒定温度 50°C 下 pH 的函数 (图 5B); 并且针对不同 pH 和温度的组合 (图 6) 测量的由 RM

Cell12A 和 MT Cell12A 对底物羧基-甲基纤维素 (CMC) 所显示的活性, 共同表示许多令人感兴趣的领悟。

第一, (a) 与由 RM Cell12A 显示的 90°C 的 T_{oa} 相比, MT Cell12A 具有显著降低的 55°C 的 T_{oa} , 其接近于已知的由 TR Cell12A 显示的 50°C 的 T_{oa} (图 5A)。这证明 TR Cell12A 的活性表面即使在移植到 RM Cell12A 的结构支架上后, 仍以其原始性质起作用 (在 MT Cell12A 中), 其没有由具有高出 42°C 的 T_m (即 96°C 代替 54°C; 见表 6) 的支架的适用性引起 T_{oa} 上移。

第二, (b) 尽管由 MT Cell12A 显示的 55°C 的 T_{oa} , 其活性对照温度曲线图 (图 5A) 比关于 TR Cell12A 已知的相应曲线图 (图 10) 显著更宽。即使在已知由于其结构在 54°C 解链 TR Cell12A 不显示活性的高温下, 仍存在显著的活性保留 (例如, 在 70°C 保留最大活性的 70%, 和在 90°C 保留 7-8% 的活性)。因此, 明显地, 对在 MT Cell12A 中起作用的输入的 TR Cell12A 活性表面赋予另外的稳定性, 即使该活性表面的 T_{oa} 在很大程度上保持不变。

第三, (c) MT Cell12A 具有比 RM Cell12A 的低 1.0 单位的最佳活性 pH, 即 pH 6.0 代替 7.0 (图 5B), 这进一步确定在 MT Cell12A 中起作用的 TR Cell12A 的自主性并为解释它们移植能力的原理提供支持。

最后, (d) 在 pH 5.0 和 50°C, MT Cell12A 显示出 RM Cell12A 活性的 1/5, (图 6)。该值在其它温度和 pH 值时更低, 这可能是由于移植的活性表面在其新结构背景中的‘脆性’, 这可以进一步通过诱变改进。尽管如此, 在 pH 5.0 和 50°C 时, MT Cell12A 几乎是 RM Cell12A 活性的 1/5 的事实提示从定性和定量的观点来看, 结构成分的折叠和装配在新酶中几乎完全地发生。这些结果还证明, 与引入的改变的深度相比, 改变的支架的任何显微结构 (对活性的) 影响是非常小的。

MT Cell12A 继承 TR Cell12A 活性表面和 RM Cell12A 支架以及 (剩余的) 表面

我们通过 X-射线晶体学确定了 MT Cell12A 的结构。该结构保存在蛋白质数据库 (PDB) 中, 其具有标识码 (PDB ID) 3B7M。该蛋白质结晶

为四聚体。该四聚体的几何排列是这样的，即各个亚基与另外两个亚基相互作用。然而，在溶剂中使用的 pH 和离子强度条件下，作为已经提到地，所述分子是单体，在储存时也观察到二聚体群（数据未显示），这表明结晶条件可以有利于这种二聚体形式的二聚体的沉积。单体结构的详细视图显示在图 7 的图 A 中，其显示与 RM Cel12A (红)和 TR Cel12A (绿) 的已知多肽主链结构叠合的确定的 MT Cel12A (蓝)主链结构的带状结构草图。该图揭示 MT Cel12A 的大多数主链结构与 RM Cel12A 和 TR Cel12A 的相似。在图 7 的图 B 中，以残基极性编码显示所有三种蛋白质的沟的表面视图，从而证明 MT Cel12A 的沟与 TR Cel12A 的沟共享特征。在图 7 的图 C 中，观察到构成与 MT Cel12A (绿) 和 TR Cel12 A (蓝)叠合的活性表面的大多数侧链，图 D 以稍微放大的形式显示相同的侧链，且没有带状结构(为了清楚)。直接证明 MT Cel12A 的活性表面与 TR Cel12A 的类似，其移植的残基在这两种情形中采用相同的几何排列。在图 7 的图 E 中，MT Cel12A (绿)、RM Cel12A (红) 和 TR Cel12A (蓝)的表面的侧视图均从相同的角度显示。可以看出，在很大程度上，MT Cel12A 的表面与 RM Cel12A 共享特征，因为其大部分残基和其全部表面（不包括活性表面沟）源自 RM Cel12A。总之，结构的详细信息证明尽管引入 36 个不连续突变，MT Cel12A 采用与 RM Cel12A 非常类似的折叠结构，其具有与 TR Cel12A 非常类似的表面（除了环区域中的一些微小偏差外）。这为本发明中已经提出的实验观察提供进一步的支持，所述实验观察显示 MT Cel12A 具有 RM Cel12A 的结构稳定性和 TR Cel12A 的物理活性特征。

推论讨论

MT Cel12A 是混合嗜中温亲本的功能性特征和嗜热性亲本的结构特征的新型酶，所述混合通过从所述嗜中温获得不连续残基组（构成负责底物结合的弯曲 β 折叠的溶剂-暴露面）并使用其代替所述嗜热性中的结构-类似残基组实现的。当然，提供稳定的结构支架的确在高温下将移植的活性表面的功能性明显地增强到显著的程度，尽管 TR Cel12A 的原始 T_{oa} 延续到 MT Cel12A 中。尝试的“活性表面移植”的成功证明酶中 T_{oa} 和 T_m 的关联不是必须的，并且显示如何利用蛋白质工程来分开并独立分类酶中

的热稳定性和功能性。采用易于结晶和结构确定的所述彻底重新-改造的酶的显著折叠和功能不仅证明酶活性表面在其较宽的总体三维结构背景中操作的自主性，还为新型理性方法提供“概念-的-证据”示范，所述新型理性方法能够在较宽的酶范围内用于缩短生物体中酶功能性和稳定性的天然共同进化，其目的是产生具有自然界通常不产生的特征的分子。

本发明涉及所有 β 蛋白的功能行为的改变，其通过下列步骤进行：

1) 鉴定具有不同物理功能特征（例如，酶功能的最佳温度）和相同化学功能特征（即，作用在相同底物上）的两种蛋白质，称为 A 和 B。

2) 确定所需改造转化的性质，并且还确定 A 和 B 中哪种蛋白质被认为是“客体”或供体结构，和哪种是“主体”或接纳体结构[例如，所述主体蛋白应该是需要保持其结构稳定性的蛋白，例如，嗜热性蛋白，且所述客体蛋白应该是需要将其功能行为经过涉及特定残基改变（如以下步骤 3-6 中给出地确定）的表面移植引入到所述主体中的蛋白，例如，嗜中温蛋白]。

3) 叠合 A 和 B 的多肽主链。

4) 鉴定位于 A 和 B 中类似表面位置处的残基对。

5) 鉴定位于 A 和 B 中与底物结合有关的类似位置处的残基对子集。

6) 确定将要产生的残基变化的表（包括从链和环区域中的连续或不连续的取代、插入和缺失）。

7) 构建靶多肽氨基酸序列，其是利用合适的诱变引物和来自主体基因的野生型模板区域，通过重叠序列延伸 (SOE) 聚合酶链反应 (PCR) 引起的剪接合成的（引入未改变和改变的残基），从而生成突变的蛋白。

8) DNA 测序以验证正确的基因合成。

9) 利用亲和标记物（例如，用于固定金属亲和层析的 6xHis 标记物）克隆、过表达和纯化突变的和其他形式的酶。

10) 针对生物化学和生物物理特征纯化酶。

11) 利用标准测定法绘制活性对比温度曲线图。

12) 利用标准测定法绘制活性对比 pH 曲线图。

13) 远-UV CD 光谱法确定二级结构。

14) 通过在远-UV 范围内监测圆二色 (CD) 信号获得热解链曲线图。

15) 质谱法确定完整酶的分子量。

16) 凝胶过滤色谱法确定所述蛋白质的四级结构状态（即是单体还是多聚体）。

17) X-射线结晶学确定所述酶的三维结构，并验证是否发生了预期的结构变化。

下列实施例通过例证本发明的方式给出，并且因此不应该将其解释为对本发明范围的限制。

实施例：

实施例 1：设计中温活性热稳定(MT) Cel12A

我们利用软件 LSQMAN 检查了 RM Cel12A 和 TR Cel12A 之间的结构同源性，从而计算主链的叠合。我们使用相同的软件量化主链叠合性水平（图 1，图 A），并确定所述主链能够以 1.1 Å 的 RMSD 叠合，尽管这两种蛋白的表面特征是显著不同的（图 1，图 B）。叠合的主链可以用于确定处于这两种酶中的类似位置处的残基号（表 1）。使用该信息产生这两种酶基于结构-类似的蛋白质（图 2）和 DNA（图 11）序列对比。然后进行视觉检验 RM Cel12A 酶的结构，从而鉴定构成纤维素-结合沟和所述沟周围相关区域的残基。确定这些位于构成所述蛋白的弯曲顶部 β 折叠的 β 链和中间环中。形成弯曲的 β 折叠的链区域由分别属于 RM Cel12A 链的位置 8-29, 53-77, 100-140, 159-167, 200-210 的残基序列段限定。对于上述 RM Cel12A 序列段中的每个残基，我们确定侧链是向上指向生物素-结合沟中，还是沿着沟的侧面（即，具有与纤维素底物相互作用的可能性），还是位于沿着沟排列的壁外侧（不具有与底物相互作用的可能性）。这组成构成 RM Cel12A 活性表面的残基子集，其显示在表 2 的第一栏中。我们然后检查了 TR Cel12A 序列/结构中相应的类似残基，从而确定是否对于 RM Cel12A 中每个残基，TR Cel12A 中的类似残基也参与形成纤维素-结合表面。我们发现对于位于链中的所有残基和对于位于环中的大部分残基，每对结构-类似残基的两个组成部分均具有完全相同的结构排列，即二者均向上指向纤维素-结合表面（活性表面）或向下指向下方的 β 折叠，或指向侧面。因此，鉴定为构成 RM Cel12A 活性表面的残基子集在 TR Cel12A 中

具有类似物，这显示在表 2 的第二栏中。然而，可以注意到，类似残基对不在所有残基位置存在。在某些区域中，环可能不叠合。因此，残基取代和插入的最终定义还包括非-类似的残基环，如表 2 所示。可以注意到，在两种酶中，一些残基是保守的。在表面移植过程中不需要对这些进行突变。关于 RM Cel12A、MT Cel12A、和 TR Cel12A 蛋白质序列的基于结构的多序列对比显示在图 2，图 A 中。

表 1

	10A2(TR Ce112A)		1H08(RM Ce112A)	
片段	GLU-A 4	THR-A 2	0.86 A	
	LEU-A 5	SER-A 3	1.16 A	
	CYS-A 6	CYS-A 4	1.06 A	*
	GLY-A 7	ASP-A 5	1.06 A	
	ARG-A 8	GLN-A 6	0.76 A	
	TRP-A 9	TRP-A 7	0.95 A	*
	ASP-A 10	ALA-A 8	0.74 A	
	ALA-A 11	THR-A 9	0.19 A	
	ARG-A 12	PHE-A 10	0.36 A	
	ASP-A 13	THR-A 11	0.16 A	
片段	VAL-A 14	GLY-A 12	2.73 A	
	GLY-A 17	ASN-A 13	2.65 A	
	ARG-A 18	GLY-A 14	1.47 A	
	TYR-A 19	TYR-A 15	0.66 A	*
	ARG-A 20	THR-A 16	0.32 A	
	VAL-A 21	VAL-A 17	0.37 A	*
	ILE-A 22	SER-A 18	0.35 A	
	ASN-A 23	ASN-A 19	0.27 A	*
	ASN-A 24	ASN-A 20	0.58 A	*
	VAL-A 25	LEU-A 21	0.26 A	*
	TRP-A 26	TRP-A 22	0.38 A	*
	GLY-A 27	GLY-A 23	0.66 A	*
片段	ALA-A 28	ALA-A 24	1.22 A	*
	THR-A 30	GLY-A 29	3.23 A	
	ALA-A 31	PHE-A 30	1.77 A	*
	GLN-A 32	GLY-A 31	1.38 A	
	CYS-A 33	CYS-A 32	1.16 A	*
	ILE-A 34	VAL-A 33	1.58 A	
	GLU-A 35	THR-A 34	1.55 A	*
	VAL-A 36	VAL-A 35	1.25 A	*
	GLY-A 37	VAL-A 36	1.49 A	*
片段	GLY-A 41	ALA-A 42	2.90 A	
	ASN-A 42	SER-A 43	1.57 A	*
	PHE-A 43	TRP-A 44	1.60 A	*
	THR-A 44	HIS-A 45	1.84 A	*
	ILE-A 45	ALA-A 46	2.41 A	*
片段	THR-A 46	ASP-A 47	3.46 A	*
	ALA-A 48	TRP-A 48	1.68 A	*
	ASP-A 49	GLN-A 49	2.61 A	*
	HIS-A 50	TRP-A 50	1.21 A	*
片段	ASP-A 51	SER-A 51	1.83 A	*
	ASN-A 52	GLY-A 53	2.89 A	*
	GLY-A 53	GLN-A 54	0.96 A	*
	ASN-A 54	ASN-A 55	0.47 A	*
	ASN-A 55	ASN-A 56	0.75 A	*
	VAL-A 56	VAL-A 57	0.27 A	*
	ALA-A 57	LYS-A 58	0.36 A	*
	ALA-A 58	SER-A 59	0.61 A	*
	TYR-A 59	TYR-A 60	0.75 A	*
	PRO-A 60	GLN-A 61	0.44 A	*
	ALA-A 61	ASN-A 62	0.24 A	*
	ILE-A 62	SER-A 63	0.17 A	*
	TYR-A 63	GLN-A 64	0.38 A	*
片段	PHE-A 64	ILE-A 65	1.18 A	*
	GLY-A 65	ALA-A 66	2.94 A	*
	PRO-A 78	LYS-A 70	1.48 A	*
	ARG-A 79	ARG-A 71	0.84 A	*
	ARG-A 80	THR-A 72	0.69 A	*
	VAL-A 81	VAL-A 73	1.02 A	*
	GLN-A 82	ASN-A 74	1.84 A	*
	GLU-A 83	SER-A 75	0.57 A	*
	LEU-A 84	ILE-A 76	1.03 A	*
	SER-A 85	SER-A 77	1.18 A	*
	ASP-A 86	SER-A 78	1.24 A	*
	VAL-A 87	MET-A 79	1.21 A	*
	ARG-A 88	PRO-A 80	0.63 A	*
	THR-A 89	THR-A 81	0.41 A	*
	SER-A 90	THR-A 82	0.63 A	*
	TRP-A 91	ALA-A 83	0.77 A	*
	THR-A 92	SER-A 84	0.69 A	*
	LEU-A 93	TRP-A 85	0.84 A	*
	THR-A 94	SER-A 86	0.89 A	*
	PRO-A 95	TYR-A 87	1.65 A	*
	ILE-A 96	SER-A 88	2.05 A	*
片段	GLY-A 99	ILE-A 92	1.38 A	*
	ARG-A 100	ARG-A 93	0.41 A	*
	TRP-A 101	ALA-A 94	0.49 A	*
	ASN-A 102	ASN-A 95	0.61 A	*
	ALA-A 103	VAL-A 96	0.66 A	*
	ALA-A 104	ALA-A 97	0.63 A	*
	TYR-A 105	TYR-A 98	0.49 A	*
	ASP-A 106	ASP-A 99	0.49 A	*
	ILE-A 107	LEU-A 100	1.19 A	*
	TRP-A 108	PHE-A 101	1.26 A	*
	PHE-A 109	THR-A 102	0.59 A	*
	SER-A 110	ALA-A 103	0.97 A	*
	PRO-A 111	ALA-A 104	1.23 A	*
	VAL-A 112	ASN-A 105	1.36 A	*
	THR-A 113	PRO-A 106	2.01 A	*
	ASN-A 114	ASN-A 107	2.50 A	*
片段	GLY-A 121	GLY-A 113	5.34 A	*
	GLY-A 122	ASP-A 114	1.08 A	*

ALA-A 123	TYR-A 115	0.69 A	
GLU-A 124	GLU-A 116	0.58 A	*
LEU-A 125	LEU-A 117	0.47 A	*
MET-A 126	MET-A 118	0.43 A	*
ILE-A 127	ILE-A 119	0.47 A	*
TRP-A 128	TRP-A 120	0.71 A	*
LEU-A 129	LEU-A 121	0.81 A	*
ASN-A 130	GLY-A 122	1.08 A	*
TRP-A 131	LYS-A 123	1.71 A	*
ASN-A 132	TYR-A 124	1.09 A	*
GLY-A 133	GLY-A 125	1.31 A	*
GLY-A 134	ASP-A 126	1.62 A	*
VAL-A 135	ILE-A 127	1.32 A	*
MET-A 136	GLY-A 128	1.19 A	*
PRO-A 137	PRO-A 129	1.11 A	*
GLY-A 138	ILE-A 130	1.31 A	*
GLY-A 139	GLY-A 131	0.53 A	*
SER-A 140	SER-A 132	0.23 A	*
ARG-A 141	SER-A 133	0.20 A	*
VAL-A 142	GLN-A 134	1.20 A	*
ALA-A 143	GLY-A 135	2.23 A	*
THR-A 144	THR-A 136	1.19 A	*
VAL-A 145	VAL-A 137	0.84 A	*
GLU-A 146	ASN-A 138	1.04 A	*
LEU-A 147	VAL-A 139	1.08 A	*
ALA-A 148	GLY-A 140	0.48 A	*
GLY-A 149	GLY-A 141	1.59 A	*
ALA-A 150	GLN-A 142	0.35 A	*
THR-A 151	SER-A 143	0.36 A	*
TRP-A 152	TRP-A 144	0.66 A	*
GLU-A 153	THR-A 145	0.63 A	*
VAL-A 154	LEU-A 146	0.64 A	*
TRP-A 155	TYR-A 147	0.05 A	*
TYR-A 156	TYR-A 148	0.25 A	*
ALA-A 157	GLY-A 149	0.47 A	*
ASP-A 158	TYR-A 150	2.37 A	*
TRP-A 159	ASN-A 151	3.08 A	*
ASP-A 160	GLY-A 152	3.40 A	*
TRP-A 161	MET-A 154	1.09 A	*
ASN-A 162	GLN-A 155	0.94 A	*
TYR-A 163	VAL-A 156	0.36 A	*
ILE-A 164	TYR-A 157	0.38 A	*
ALA-A 165	SER-A 158	0.22 A	*
TYR-A 166	PHE-A 159	0.19 A	*
ARG-A 167	VAL-A 160	0.56 A	*
ARG-A 168	ALA-A 161	1.13 A	*
THR-A 169	GLN-A 162	1.58 A	*
THR-A 170	THR-A 163	1.80 A	*
PRO-A 171	ASN-A 164	1.79 A	*
THR-A 172	THR-A 165	1.26 A	*
THR-A 173	THR-A 166	0.41 A	*
SER-A 174	ASN-A 167	1.10 A	*
VAL-A 175	TYR-A 168	0.85 A	*
SER-A 176	SER-A 169	2.29 A	*
LEU-A 178	GLY-A 170	0.30 A	*
ASP-A 179	ASP-A 171	1.14 A	*
LEU-A 180	VAL-A 172	0.89 A	*
LYS-A 181	LYS-A 173	0.94 A	*
ALA-A 182	ASN-A 174	0.41 A	*
PHE-A 183	PHE-A 175	0.43 A	*
ILE-A 184	PHE-A 176	0.66 A	*
ASP-A 185	ASN-A 177	1.02 A	*
ASP-A 186	TYR-A 178	1.41 A	*
ALA-A 187	LEU-A 179	1.78 A	*
ALA-A 188	ARG-A 180	2.59 A	*
ALA-A 189	ASP-A 181	3.19 A	*
ILE-A 193	TYR-A 185	1.56 A	*
ARG-A 194	ASN-A 186	1.95 A	*
PRO-A 195	ALA-A 187	0.61 A	*
GLU-A 196	GLY-A 189	0.91 A	*
TRP-A 197	GLN-A 190	0.24 A	*
TYR-A 198	TYR-A 191	0.63 A	*
LEU-A 199	VAL-A 192	1.25 A	*
HIS-A 200	LEU-A 193	1.33 A	*
ALA-A 201	SER-A 194	1.02 A	*
VAL-A 202	TYR-A 195	0.49 A	*
GLU-A 203	GLN-A 196	0.43 A	*
THR-A 204	PHE-A 197	0.41 A	*
GLY-A 205	GLY-A 198	0.62 A	*
PHE-A 206	THR-A 199	0.53 A	*
GLU-A 207	GLU-A 200	0.55 A	*
LEU-A 208	PRO-A 201	0.37 A	*
TRP-A 209	PHE-A 202	1.06 A	*
GLU-A 210	THR-A 203	2.91 A	*
GLY-A 211	GLY-A 204	2.82 A	*
ALA-A 213	GLY-A 206	2.09 A	*
GLY-A 214	THR-A 207	1.58 A	*
LEU-A 215	LEU-A 208	0.77 A	*
ARG-A 216	ASN-A 209	1.35 A	*
SER-A 217	VAL-A 210	2.55 A	*
ALA-A 218	ALA-A 211	1.31 A	*
ASP-A 219	SER-A 212	1.05 A	*
PHE-A 220	TRP-A 213	0.88 A	*
SER-A 221	THR-A 214	1.04 A	*
VAL-A 222	ALA-A 215	2.12 A	*
THR-A 223	SER-A 216	1.67 A	*
VAL-A 224	ILE-A 217	1.62 A	*
GLN-A 225	ASN-A 218		

表 2

海洋红嗜热盐菌链中的原始残基	里氏木霉 (<i>T. reesei</i>) 链中的相应 (类似) 残基	结构起源	采取的作用
Arg8	Gln6	环	突变的
Trp9	Trp7	链	保守的
Ala11	Thr9	链	突变的
Asp13	Thr11	链	突变的
Arg20	Thr16	链	突变的
Ile22	Ser18	链	突变的
Asn24	Asn20	链	保守的
Trp26	Trp22	链	保守的
Asn54	Asn55	环	保守的
Asp55 (报告序列中的 Asn 55)	Asn56	环	突变的
Val56	Val57	环	保守的
Tyr59	Tyr60	链	保守的
Ala61	Asn62	链	突变的
Tyr63	Gln64	链	突变的
Gly65	Ala66	链	突变的
Cys66-Leu77 即, Cys66, His67, Trp68, Gly69, Ala70, Cys71, Thr72, Ser73, Asn74, Ser75, Gly76, Leu77	Ile67-Gln69 即, Ile67, Pro68, Gln69	环, 在 RM 和 TR 序列之间不具有 结构一致性	突变的
Arg100	Arg93	链	保守的
Asn102	Asn95	链	保守的
Ala104	Ala97	链	保守的
Asp106	Asp99	链	保守的
Trp108	Phe101	链	突变的
Ser110	Ala103	链	突变的
Pro111	Ala104	环	突变的
Gly112 (报告序列中的 Val112)	Asn105	环	突变的
Thr113	Pro106	环	突变的
Asn114	Asn107	环	保守的
Ser115-Gly121 即, Ser115, Ser116, Asn117, Gly118, Tyr119, Ser120, Gly121 (报告序列中的 Gly116)	His108-Asp113 即, His108, Val109, Thr110, Tyr111, Ser112, Gly113	环, 但在 RM 和 TR 序列之间不具 有结构一致性	Tyr111, Ser112 和 Gly113 保守的; His108, Val109, 和 Thr110 突变的
Gly122	Asp114	链	突变的
Glu124	Glu116	链	保守的
Met126	Met118	链	保守的
Trp128	Trp120	链	保守的

Trp131	Lys123	环	突变的
Gly134	Asp126	环	突变的
Gly138	Ile130	环	突变的
Ser140	Ser132	环	保守的
Trp159	Asn151	环	突变的
Asp160	Gly152	环	突变的
--	Ala153	环	插入的
Trp161	Met154	链	突变的
Tyr163	Val156	链	突变的
Ala165	Ser158	链	突变的
Arg167	Val160	链	突变的
His200	Leu193	环	突变的
Ala201	Ser194	链	突变的
Glu203	Gln196	链	突变的
Gly205	Gly198	链	保守的
Glu207	Glu200	链	保守的
Trp209	Phe202	链	突变的
Glu210	Thr203	环	突变的

实施例 2: MT Cell12A 结构的建模

通过用 TR Cell12A 中结构-类似残基的坐标在计算机芯片上代替 RM Cell12A 残基的坐标首先生成 MT Cell12A 的坐标文件来生成移植的蛋白 (MT Cell12A) 的模型。利用 AMBER 8.0 通过能量最小化去除原子撞击和键角异常 (DRMS 0.01; 使用急剧下降和共轭梯度方法)。该模型以全-原子表面表示, 以及 RM Cell12A 和 TR Cell12A 的类似表示显示在图 2B 中。

实施例 3: 基因合成、克隆和 DNA 测序

我们构建两种编码天然-存在 RM Cell12A 酶重组形式和重组非-天然-存在中温活性热稳定酶 MT Cell12A 的基因。用于基因合成的方案, 用于 PCR 和剪接-重叠序列延伸 (SOE) PCR 的引物和条件, 以及合成的基因的全序列和编码的氨基酸序列的详细信息分别在图 11、表 3 和 4 和图 12、和图 13 中给出。可以注意到基因合成引入 (i) Bam HI 限制位点 (其 5' 末端侧连 12 个碱基对的突出端, 以促进消化), 该位点在编码 RM Cell12A 和 MT Cell12A 的基因中均紧接在编码起始 N-端残基 (苏氨酸) 的密码子之前, (ii) 终止密码子, 其在编码最后的 C-端残基 (谷氨酰胺) 之后, 和 (iii) Hind III 限制位点 (其 3' 末端侧连 12 个碱基对的突出端, 以促进消

化), 其紧接在终止密码子之后。消化 Bam HI 和 Hind III 限制位点, 从而容许将基因插入和连接进入 pQE-30 (Qiagen) 载体。除选择标记 (氨苄青霉素抗性) 外, 所述载体提供可诱导的启动子、转录起始位点、翻译起始位点、和 N-端亲合标记物(N-MRGSHHHHGS-C)。载体中用于编码所述标记物最后两个残基, 即 G 和 S 的碱基一起构成 Bam HI 位点, 这容许插入合成的基因。所述载体还在 Hind III 位点后提供终止密码子, 但是我们优选使用在 Hind III 位点前、紧随 C-端谷氨酰胺之后的终止密码子, 如已经提及地。将连接的载体转化进入携带基因型 *hsdR17 recA1 lacF'* [*proAB+lacI^q lacZ Δ15 Tn10 (tet^r)*] 的感受态大肠杆菌 (*E.coli*) XL-1 blue 细胞中。在含有四环素和氨苄青霉素的 LB 平板上进行克隆的选择。使用应用生物系统 DNA 测序仪 (3130 XL 测序仪) 进行克隆的自动化 DNA 测序。利用 ABI MiniPrep 试剂盒, 从 XL-1 blue 细胞中纯化质粒。使用具有序列 5'-CGGATAACAATTCACACAG-3' 的载体-特异性正向引物, 或具有序列 5'-GTTCTGAGGTCATTACTGG-3' 的载体-特异性反向引物进行热- (循环) -测序反应, 所述正向引物用于通读编码所述蛋白 N-末端的基因区域, 所述反向引物用于通读编码所述蛋白 C-末端的区域。反应使用 ABI 现成反应混合物 (大染料终止子 v3.1 循环测序 RR-100 (Big Dye Terminator v3.1 cycle sequencing RR-100)), 所述混合物在 96°C 变性 (5 分钟), 随后进行 30 个循环的 96°C 变性 (1 分钟), 50°C 退火 (1 分钟), 和 60°C 延伸 (2 分钟)。在上样到分析器上之前, 通过标准程序结合使用 EDTA 和乙醇的洗涤进行反应-后的清理, 显示编码 MT Cell12A 的基因的序列的电泳图提供在图 14 中。

表 3

反应号	正向引物	反向引物	模板 DNA	退化温度/时间	延伸温度/时间	产物 (bp)	酶	Mg ²⁺ 浓度 (mM)
1	引物 1N	引物 12	RM 基因组	60°C; 2 分钟	75°C; 3 分钟	714	Triple Master	2.5
2	引物 1	引物 2	反应 1 的产物	55°C; 2 分钟	75°C; 3 分钟	219	Triple Master	2.5
3	引物 3	引物 4	反应 1 的产物	55°C; 2 分钟	75°C; 3 分钟	174	Deep Vent	4.0

4	引物 5	引物 6	反应 1 的产物	60℃; 2 分钟	75℃; 3 分钟	114	Triple Master	2.5
5	引物 7	引物 8	反应 1 的产物	60℃; 2 分钟	75℃; 3 分钟	117	Triple Master	2.5
6	引物 9	引物 10	反应 1 的产物	60℃; 2 分钟	75℃; 3 分钟	156	Triple Master	2.5
7	引物 11	引物 12	反应 1 的产物	60℃; 2 分钟	75℃; 3 分钟	99	Triple Master	2.5
8	引物 1	引物 4	反应 2 和 3 的 产物	55℃; 2 分钟	75°; 3 分钟	360	Triple Master	2.5
9	引物 5	引物 8	反应 4 和 5 的 产物	55℃; 2 分钟	75℃; 3 分钟	207	Triple Master	2.5
10	引物 9	引物 12	反应 6 和 7 的 产物	60℃; 2 分钟	75℃; 3 分钟	234	Triple Master	2.5
11	引物 1	引物 12	反应 8、 9 和 10 的产物	60℃; 2 分钟	75°; 3 分钟	681	Deep Vent	2.0

表 4

引物号	引物序列
引物 1N	5'-ACTTATACTATAGGATCCACTGTGAGCTGTGTTGTAGATGGGACGCGCGC-3'
引物 1	5'-ACTTATACTATCGGATCCACGGTCGAGCTGTGCGGACAGTGGGACAGAGAACGGTGGCT GGGGGGCGCTACACGGTGAGCAACAACGTATGGGGC-3'
引物 2	5'-CTGCGGGATGGCAAACCTGGATGTTCCGATAGGCGGCCACGTT-3'
引物 3	5'-GCCTATCCGAACATCCAGTTTGCCATCCGACGCCGCGCGTGTGCAAGAACTGTCCGAC GTG-3'
引物 4	5'-CAGCTCGGCGTCTCCAGAATACGTCACGTGGTTTGGGTTGGCGGCGAAGAAGATGTCGTA GCGGC-3'
引物 5	5'-CGCTGGAATGCCGCCTACGACATCTTCTTCGCCGCCAACCCAAACCACGTGACGTATTCTG GAGACGCCGAGCTG-3'
引物 6	5'-GATAGGCATCACATCGCCGTTTTTGTTCAGCC-3'
引物 7	5'-AAAAACGGCGATGTGATGCCTATCGGCAGCCGCGTGGCCACCG-3'
引物 8	5'-CCTCACGTAGCTGATCACATTATCGCACCATTGTGAGCATAACAGACTTCCC-3'
引物 9	5'-TGGTATGCTGACAATGGTGCGATGAATGTGATCAGCTACGTGAGGACGACGCCC-3'
引物 10	5'-GCCCGTTTGACCGACAGCAGATACCACTCCGG-3'
引物 11	5'-CTGCTGTGCGTGCAAACGGGCTTTGAACTCTTACTGTGGTGCCGGTCTGCGAAG CGCC-3'
引物 12	5'-ACTTATACTATCAAGCTTCTACTGCACAGTTACGGAAAAATCGGC-3'

表 5:

RM Cel12A 中的残基 (PubMed 和 PDB 中报告的)	RM Cel12A 中的残基 (我们发现的)	MT Cel12A 的残基 (我们使用的)	TR Cel12A 中的残基 (报告的)	位置	注释
Asp49 (D)	Glu	Glu	Gln49	不是活性表面的一部分	因为 Asp 和 Glu 均带负电荷, 我们决定保留在我们克隆的基因中发现的 Glu
Asn55 (N)	Asp	Asn	Asn56	活性表面的一部分	由于该残基是活性表面的一部分, 且 TR Cel12A 中的类似残基是 Asn, 所以我们确定在 MT Cel12A 中应该是 Asn
Val112 (V)	Gly	Asn	Asn105	- Do -	- Do
Gly116 (G)	Ser	His	His108	- Do -	- Do
Ser176 (S)	Thr	Thr	Ser169	不是活性表面的一部分	在 Wicher 等, 应用微生物学和生物技术 (Appl. Microb. Biotechnol.) 55, 578-584 (2001) 中, 报告了 RM Cel12A 原始序列的作者将 Ser176 修正为 Thr176。由于我们在该位置发现为 Thr, 所以我们保留它。不将其改变为 Ser, 因为它不在活性表面中。

实施例 4: 蛋白质表达和纯化

小规模纯化: 将设置在摇动烧瓶中 2 升总体积中的 XL-1 blue 细胞的二级培养物在 37°C 生长直到至 O.D₆₀₀ ~0.6, 并用 0.4 mM IPTG 进行诱导。为了生成 RM Cel12A, 诱导后 4 小时收获细胞, 而对于 MT Cel12A, 我们发现在 12 小时后收获对获得高产率更好。裂解收获的细胞, 并在非-变性条件下对其进行标准的基于 Ni-NTA 的亲合纯化(Qiagen), 区别是利用 1M 咪唑进行洗脱。随后将咪唑透析掉, 以获得纯的折叠蛋白。MT Cel12A 的非-变性纯化 PAGE 模式图显示在图 15A 和 15B 中, RM Cel12A 的非-变性纯化 PAGE 模式图显示在图 15C 和 15D 中。

大规模纯化: 对诱导后 12 小时从培养物中收获的细菌细胞使用基于超声波-或基于 dynamill 的裂解作用, 并重悬在 50 mM NaH₂PO₄, 300 mM NaCl, pH 8.0 中。基于离心去除细胞碎片, 然后通过向上清液中加入硫酸铵达到 80%的饱和水平, 以及随后基于离心收集沉淀的蛋白。将沉淀的蛋白溶解在 pH 5.0 的 10 mM 柠檬酸盐缓冲液中, 并进行透析以去除痕量的

硫酸铵。在 65°C 加热重新溶解在柠檬酸盐缓冲液中的蛋白质 15-30 分钟，从而加热-变性和加热-沉淀裂解物中除存在的酶，即 RM Cel12A 或 MT Cel12A（依赖于纯化这些蛋白的哪一种）之外的所有细菌细胞质蛋白。随后进行离心，去除热-沉淀的蛋白，将纯 RM Cel12A 或 MT Cel12A 保留在上清液中，其可作为考马斯-染色的 SDS-PAGE 中单一条带辨别。

实施例 5：质谱法

在 Voyager DE-STR MALDI-TO 质谱仪上对纯化的 RM Cel12A 和 MT Cel12A 进行质谱鉴定。获得的质量值充分落入确定这些质量范围的质量精确性中预期的误差范围内。图 16 的图例中提供了进一步的详细信息。

实施例 6：远-UV CD 光谱法

为了波长扫描和温度（结构解链）扫描，使用 0.1-0.3 mg/ml 范围内的蛋白质浓度，使用径长为 0.1 或 0.2 cm 的杯子，和 250-195 nm 范围内的扫描原始椭圆率（ θ ）值，以 6-9 升/分钟的 N₂ 气流速，在 Jasco J-810 分光偏振计上，进行 CD 光谱测量。对于在固定波长(218 nm)处的 pH-依赖性温度扫描，考虑到 218 nm 处 25°C 的 θ 值对应于完全折叠的状态，且为零的 θ 值对应于完全未折叠状态，将原始观察到的椭圆率值转换为部分解折叠值。对于标准波长扫描，利用公式 $[\theta] = \{\theta_{\text{obs}} \times 100 \times \text{平均残基重量}\} / \{\text{浓度 (mg/ml)} \times \text{径长 (cm)}\}$ ，将原始 θ 转换为平均残基椭圆率值。CD 结果提供在图 3 和 4 中。

实施例 7：凝胶过滤色谱法

使用与 SMART 色谱工作站（法玛西亚(Pharmacia)）连接的、预校准和适当预平衡的微-分析 Superdex-75 凝胶过滤柱（床体积 2.4 ml），在存在不同浓度的盐，和使用不同蛋白浓度条件下，检查处于 pH 5.0 或 pH 8.0 缓冲液中正常条件下的 RM Cel12A 和 MT Cel12A 的流体力学体积和洗脱行为。校准数据显示在图 17 图 A 中，且显示 MT Cel12A 缺乏任何解离的对照（进一步支持单体状态）显示在图 17 图 B 中。

实施例 8：内切葡聚糖酶活性测定

通过标准的基于 DNS-终止的方法(Miller 等, 1960. 分析生物化学 (*Anal. Biochem.*) 2, 127-132)测定 MT Cel12 A 和 RM Cel12A 的酶活性。对于在多种温度 (具有固定的 pH), 或在多种 pH 值 (具有固定的温度) 下进行总酶活性测量, 使用以下详述的条件。*温度扫描* (图 5A) —为了在各种温度下温育, 我们使用处在 pH 5.0 的柠檬酸盐缓冲液中的最终 MT Cel12A 浓度 0.1 mg/ml 和 NaCl 浓度 100 mM (终浓度 35 mM)。对每次实验, 用水使反应体积达到 1ml, 并向其中加入 1ml 1.8% 的 CMC (羧甲基纤维素), 并温育该溶液 1 小时。对每种温度进行总共 5 次这样的实验。关于 RM Cel12A, 我们使用处在 pH 5.0 的柠檬酸盐缓冲液中的最终酶浓度 0.04 mg/ml 和 NaCl 浓度 100 mM (终浓度 35 mM)。类似地使反应体积达到 1ml, 并向其中加入 1ml 1.8% CMC, 并温育该溶液 20 分钟。对每种温度进行总共 5 次这样的实验。在下列温度下温育各容纳 2ml 的不同管: 10, 20, 30, 40, 50, 60, 70, 80, 和 90°C。还在 100°C 对 RM Cel12A 进行测量。温育后, 将 3ml DNS (二硝基水杨酸) 试剂加入到每个管中, 并将所述管煮沸 15 分钟, 从而通过 DNS 与由纤维素酶对 CMC 底物的作用释放的还原糖反应而产生颜色。也在没有酶的条件下温育对照反应。通过 550nm 的吸光度测量评估颜色的形成, 并减去关于对照的值 (~0.03)。还使用葡萄糖进行标准校准图, 以评估还原糖——但没有报告这些, 因为受监测的相关参数具有相对活性。认为获得的最高活性是 100%, 且将剩余的活性转换为观察到的最大活性的百分比数值。*pH 扫描* (图 5B) —为了在各种 pH 下温育, 我们使用处在适当缓冲液中的最终 MT Cel12A 浓度 0.1 mg/ml 和 NaCl 浓度 100 mM (最终浓度 35mM)。使反应体积达到 1ml, 并向其中加入 1ml 1.8% CMC, 并在 50°C 温育该溶液 1 小时。对每种 pH 进行总共 5 次这样的实验。关于 RM Cel12A, 我们使用处在适当缓冲液中的最终酶浓度 0.01 mg/ml 和 NaCl 浓度 100 mM (终浓度 35 mM), 使其达到 1ml, 并向其中加入 1ml 1.8% CMC, 并在 50°C 培养该溶液 20 分钟。对每种 pH 进行总共 5 次这样的实验。不同的管容纳 2ml 处于下列 pH 值范围内的各种以上混合物: 3, 4, 5, 6, 7, 8, 9, 和 10。对于 3.0-6.0 范围内的 pH, 使用柠檬酸盐缓冲液。对于 pH 7.0-9.0, 使用 tris 缓冲液。对于 pH 10.0, 使用碳酸盐/碳酸氢盐缓冲液。

严格按照关于以上的温度扫描, 进行利用 DNS 的颜色形成反应。认为获得的最高活性是 100%, 且将剩余的活性转换为观察到的最大活性的百分比数值。在相同条件下比较 *RM Cell12A* 和 *MT Cell12A* 的活性(图 6)——为了比较, 在以下绝对相同的条件下对 *RM Cell12A* 和 *MT Cell12A* 进行活性测定: 在 45 mM 柠檬酸盐缓冲液(用于在 pH 5.0 比较)或 45 mM tris 缓冲液(用于在 pH 8.0 比较)中, 在存在 100mM NaCl 的条件下, 使用相同最终浓度(0.1 mg/ml)的两种酶, 使其达到 1ml, 并加入到 1.0 ml 1.8 % CMC 中, 以形成总体积 2ml。在两种不同温度, 即 50℃和 90℃, 在这两种 pH 下, 使用这两种酶一式三份地进行温育 1 小时。严格按照关于以上的温度和 pH 扫描, 进行利用 DNS 的颜色形成反应。

实施例 9: 蛋白质参数

RM Cell12A。长度: 236 个残基(包括具有序列 N-MRGSHHHHHHGS-C 的 12 个残基长的 N-端亲合标记物)。分子量: 26215.93 Da。消光系数 (280nm): 280 nm 处的 92940.1 O.D 等价于 0.28 mg/ml。等电点 (pI): 5.53。**MT Cell12A**。长度: 227 个残基(包括具有序列 N-MRGSHHHHHHGS-C 的 12 个残基长的 N-端亲合标记物)。分子量: 25037.01 Da。消光系数 (280nm): 280 nm 处的 56,000.1 O.D 等价于 0.45 mg/ml。等电点 (pI): 5.39。

表 6

	TR Cell12A [@]	RM Cell12A [@]	MT Cell12A
残基数 [*]	218	225 ⁺	215
分子质量 [#]	23512 Da	26216 Da ⁺	25037 Da
等电点	5.56	5.53	5.59
最佳活性 pH	5.0	7.0	6.0
T _m	54 °C	96 °C	93 °C
T _{oa}	50 °C	90 °C	55 °C
T _{og}	28 °C	65 °C	不适用

* 排除在 *RM Cell12A* 和 *MT Cell12A* 上存在的 N-MRGSHHHHHHGS-C 标记物的长度。

包括在RM Cel12A 和MT Cel12A 上存在的N- MRGSHHHHHHGS-C 标记物的质量。

+ 我们的RM Cel12A 克隆在所述标记物后仅具有 224 个残基, 因为我们去除了 N-端的甲硫氨酸。

@关于 TR Cel12A 和 RM Cel12A 的详细信息来自公开的文献, 或我们的数据。TR Cel12A 的详细信息来自 Sandgren 等, 2003. *蛋白质科学 (Protein Science)* 12, 848-860; Simmons, 1977, 第二届国际真菌大会 (*Second Intl. Mycol. Congress*), 坦帕, 美国佛罗里达, 第 618 页; Karlsson 等, 2002. 生物技术杂志 (*J. Biotechnol.*) 99, 63-78。RM Cel12A 的详细信息来自 Crennel 等, 2002. 分子生物学杂志 (*J. Mol. Biol.*) 320, 883-897; Bjornsdottir 等, 2006. 嗜极端生物 (*Extremophiles*) 10, 1-16; Hallorsdottir 等, 1998. 应用微生物学和生物技术 (*Appl. Microbiol. Biotechnol.*) 49, 277-284。

实施例 10: 结构鉴定

利用悬滴蒸汽扩散法, 在下列条件下, 使用约 10 mg/ml 的起始蛋白浓度使蛋白结晶: 0.2 M $\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$, 20 % PEG 3350, pH 4.5。晶体在 20°C, 2-3 天后, 生长为尺寸 0.5 x 0.4 x 0.8 mm。利用旋转式阳极 X-射线发生器 (Rigaku UltraX, 日本) 和图像板检测器 (MAR 研究 (MARresearch), 德国) 收集数据。晶体衍射到 2.3 Å 分辨率。利用程序的 HKL 组简化和测量数据 (Denzo 和 Scalepack)。利用 RM Cel12A 作为模型, 和如在 CCP4 中执行的 MOLREP 程序, 通过分子替换法确定结构。利用 CNS 和 CCP4 进行细化。通过利用 Coot 程序计算傅里叶和差别傅里叶图谱, 在绘图工作站上检查所述模型。利用 PROCHECK 和 WHATCHECK 程序验证所述模型, 从而检查任何误差。将模型和结构因子存放在蛋白质数据库中 (PDB ID No. 3B7M)。

优点

- 本发明阐明的设计原理促进将基于 β 折叠结构的任何活性表面从任何蛋白质/酶“移植”结构-同源的蛋白质/酶中, 所述移植不管序列同一性水平, 但以严格依赖于供体 (客体) 和接纳体 (主体) 酶主链、特别是在

移植涉及的两种酶类似区域内的叠合能力水平的方式进行。

- 酶活性表面的成功移植，如通过本发明所述示范地，揭示酶活性表面是能够主要不依赖其主体酶的支持结构支架的结构和稳定性起作用的酶显微结构和功能的自主单位，这是因为在所有支持活性表面功能的条件下保持所述主体结构；同样地，该发现（和相关方法）容许研究者重组来自两种非常不同的生命领域的酶结构稳定性和蛋白质活性特征，如本发明中所示。

- 本发明示范调整酶物理功能特征，诸如例如其最佳活性温度的严格理性方法的第一个用途。

- 本发明证明酶活性位点功能特征的差别显著归功于与底物分子结合的表面的特征（结构/稳定性/灵活性/化学）的不同，而非仅归功于直接参与催化作用的残基。

- 尽管，本发明示范了使用 β 折叠结构的原理，但是通过仔细的结构分析可以将该概念和方法外推至移植还包括螺旋和其他结构的表面。

- 尽管，本发明示范包括对小分子底物结合和催化作用的活性表面的移植，但是可将该概念和方法外推至不包括任何催化化学活性的蛋白-蛋白和蛋白-小分子相互作用。

- 尽管，本发明示范仅移植酶表面的一部分，但是，可以将该概念和方法外推至酶之间、或非-酶蛋白质之间的完整-表面移植，从而以自然界中一般不采用的方式，组合一种酶的核心结构稳定性特征和另一种同源酶的表面特征和功能性。

<110> 科学与工业研究委员会

<120> 重组中温活性热稳定蛋白及其设计和生物合成方法

<130> 0215NF2006

<160> SEQ ID NO:1

<170> PatentIn version 3.4

<210> SEQ ID NO:1

<211> 224

<212> PRT

<213> 海洋红嗜热盐菌 (Rhodothermus marinus)

<400> SEQ ID NO:1

```

Thr Val Glu Leu Cys Gly Arg Trp Asp Ala Arg Asp Val Ala
 2           5           10          15

Gly Gly Arg Tyr Arg Val Ile Asn Asn Val Trp Gly Ala Glu Thr
 20          25          30

Ala Gln Cys Ile Glu Val Gly Leu Glu Thr Gly Asn Phe Thr Ile
 35          40          45

Thr Arg Ala Glu His Asp Asn Gly Asn Asp Val Ala Ala Tyr Pro
 50          55          60

Ala Ile Tyr Phe Gly Cys His Trp Gly Ala Cys Thr Asn Ser Ser
 65          70          75

Gly Leu Pro Arg Arg Val Gln Glu Leu Ser Asp Val Arg Thr Ser
 80          85          90

Trp Thr Leu Thr Pro Ile Thr Thr Gly Arg Trp Asn Ala Ala Tyr
 95          100         105

Asp Ile Trp Phe Ser Pro Gly Thr Asn Ser Ser Asn Gly Tyr Ser
 110         115         120

Gly Gly Ala Glu Leu Met Ile Trp Leu Asn Trp Asn Gly Gly Val
 125         130         135

Met Pro Gly Gly Ser Arg Val Ala Thr Val Val Leu Ala Gly Ala
 140         145         150

Thr Trp Glu Val Trp Tyr Ala Asp Trp Asp Trp Asn Tyr Ile Ala
 155         160         165

Tyr Arg Arg Thr Thr Pro Thr Thr Ser Val Thr Glu Leu Asp Leu
 170         175         180

Lys Ala Phe Ile Asp Asp Ala Val Ala Arg Gly Tyr Ile Arg Pro
 185         190         195

Glu Trp Tyr Leu His Ala Val Glu Thr Gly Phe Glu Leu Trp Glu
 200         205         210

Gly Gly Ala Gly Leu Arg Ser Ala Asp Phe Ser Val Thr Val Gln
 215         220         225

```

<110> 科学与工业研究委员会

<120> 组中温活性热稳定蛋白及其设计和生物合成方法

<130> 0215NF2006

<160> SEQ ID NO:2

<170> PatentIn version 3.4

<210> SEQ ID NO:2

<211> 218

<212> PRT

<213> 里氏木霉 (Trichoderma reesei)

<400> SEQ ID NO:2

```

Glu Thr Ser Cys Asp Gln Trp Ala Thr Phe Thr Gly Asn Gly Tyr
1      5      10      15
Thr Val Ser Asn Asn Leu Trp Gly Ala Ser Ala Gly Ser Gly Phe
20     25     30
Gly Cys Val Thr Val Val Ser Leu Ser Gly Gly Ala Ser Trp His
35     40     45
Ala Asp Trp Gln Trp Ser Gly Gly Gln Asn Asn Val Lys Ser Tyr
50     55     60
Gln Asn Ser Gln Ile Ala Ile Pro Gln Lys Arg Thr Val Asn Ser
65     70     75
Ile Ser Ser Met Pro Thr Thr Ala Ser Trp Ser Tyr Ser Gly Ser
80     85     90
Asn Ile Arg Ala Asn Val Ala Tyr Asp Leu Phe Thr Ala Ala Asn
95     100    105
Pro Asn His Val Thr Tyr Ser Gly Asp Tyr Glu Leu Met Ile Trp
110    115    120
Leu Gly Lys Tyr Gly Asp Ile Gly Pro Ile Gly Ser Ser Gln Gly
125    130    135
Thr Val Asn Val Gly Gly Gln Ser Trp Thr Leu Tyr Tyr Gly Tyr
140    145    150
Asn Gly Ala Met Gln Val Tyr Ser Phe Val Ala Gln Thr Asn Thr
155    160    165
Thr Asn Tyr Ser Gly Asp Val Lys Asn Phe Phe Asn Tyr Leu Arg
170    175    180
Asp Asn Lys Gly Tyr Asn Ala Ala Gly Gln Tyr Val Leu Ser Tyr
185    190    195
Gln Phe Gly Thr Glu Pro Phe Thr Gly Ser Gly Thr Leu Asn Val
200    205    210
Ala Ser Trp Thr Ala Ser Ile Asn
215

```

<110> 科学与工业研究委员会

<120> 重组中温活性热稳定蛋白及其设计和生物合成方法

<130> 0215NF2006

<160> SEQ ID NO:3

<170> PatentIn version 3.4

<210> SEQ ID NO:3

<211> 215

<212> PRT

<213> 海洋红嗜热盐菌 (Rhodothermus marinus)

<400> SEQ ID NO:3

```

Thr Val Glu Leu Cys Gly Gln Trp Asp Thr Arg Thr Val Ala Gly
 2             5             10             15

Gly Arg Tyr Thr Val Ser Asn Asn Val Trp Gly Ala Glu Thr Ala
          20             25             30

Gln Cys Ile Glu Val Gly Leu Glu Thr Gly Asn Phe Thr Ile Thr
          35             40             45

Arg Ala Glu His Asp Asn Gly Asn Asn Val Ala Ala Tyr Pro Asn
          50             55             60

Ile Gln Phe Ala Ile Pro Gln Pro Arg Arg Val Gln Glu Leu Ser
          65             70             75

Asp Val Arg Thr Ser Trp Thr Leu Thr Pro Ile Thr Thr Gly Arg
          80             85             90

Trp Asn Ala Ala Tyr Asp Ile Phe Phe Ala Ala Asn Pro Asn His
          95             100            105

Val Thr Tyr Ser Gly Asp Ala Glu Leu Met Ile Trp Leu Asn Lys
          110            115            120

Asn Gly Asp Val Met Pro Ile Gly Ser Arg Val Ala Thr Val Glu
          125            130            135

Leu Ala Gly Ala Thr Trp Glu Val Trp Tyr Ala Asp Asn Gly Ala
          140            145            150

Met Asn Val Ile Ser Tyr Val Arg Thr Thr Pro Thr Thr Ser Val
          155            160            165

Thr Glu Leu Asp Leu Lys Ala Phe Ile Asp Asp Ala Val Ala Arg
          170            175            180

Gly Tyr Ile Arg Pro Glu Trp Tyr Leu Leu Ser Val Gln Thr Gly
          185            190            195

Phe Glu Leu Phe Thr Gly Gly Ala Gly Leu Arg Ser Ala Asp Phe
          200            205            210

Ser Val Thr Val Gln
          215

```

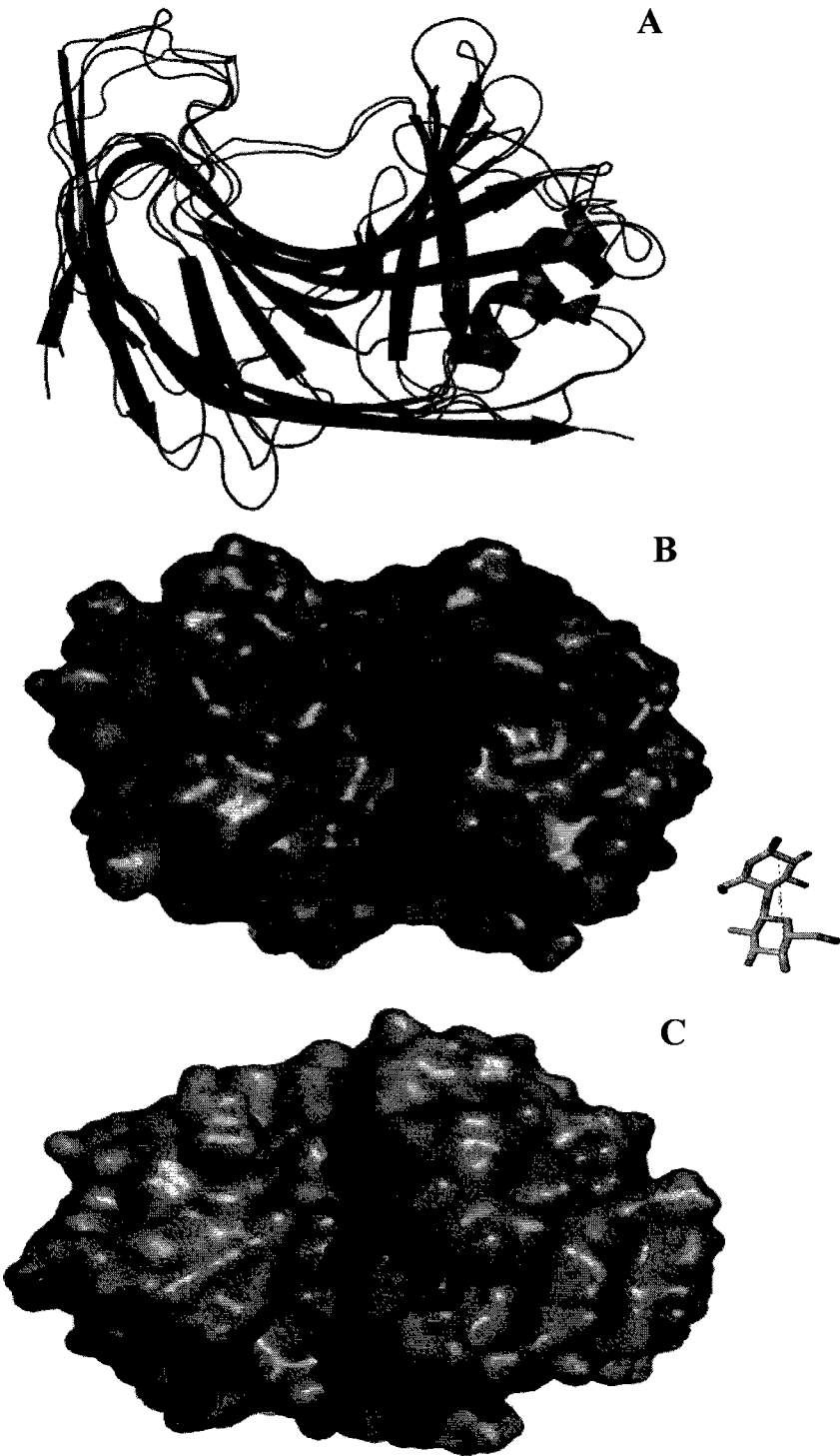



图 1

A

1H0B TVELCGRWDARDVAGGRYRVINNVGAET---AQCI EVGL-ET-GNFTIT 46
MT TVELCGQWDRTIVAGGRYTVSNNVGAET---AQCI EVGL-ET-GNFTIT
10A2 -ETSCDQWATFTIGN--GYTVSNNLWASAGSGFGCVTVVSLSGGASWHAD 47

D N SN
1H0B RAEHDN-GNDVAAYPAIYF---GCHWGACTNSSGLPRRVQELSDVRTSW 91
MT RAEHDN-GNNVAAYPNIQFAIPQ-----PRRVQELSDVRTSW
10A2 -WQWSGGQNNVKSQNSQIAIPQ-----KRTVNSISSMPTTA 83

V G
1H0B TLTPIT-TGRWNAAYDIWFSPGTNSSNG---YSGGAEELMIWLNWNGGVMP 137
MT TLTPIT-TGRWNAAYDIIFAANPN----FVTYSGDAELMIWLNKNGGVMP
10A2 SWSYSGSNIRANVAYDLETAANPN----HVTYSGDYELMIWLKGYGDIGP 129

S
1H0B GGSRVATVELAGATWEVWYADWD-WNYIAYRRTTPTTSVTELDLKAFIDD 186
MT IGSRVATVELAGATWEVWYADNGAMNVI SYVRTTPTTSVTELDLKAFIDD
10A2 IGSQGTVNVGGSWTLYYGYNGAMQVYSFVAQTNTTNYS-GDVKNFFNY 178

1H0B AVARGYIRP-EWYLHAVETGFELWEGGAGLRSADFSVTVQ 225
MT AVARGYIRP-EWYLLSMQTGFELFTGGAGLRSADFSVTVQ
10A2 LRDNKGYNAAAGQYVLSYQFGTEPFTGSGTLNVASWTASIN 218

B

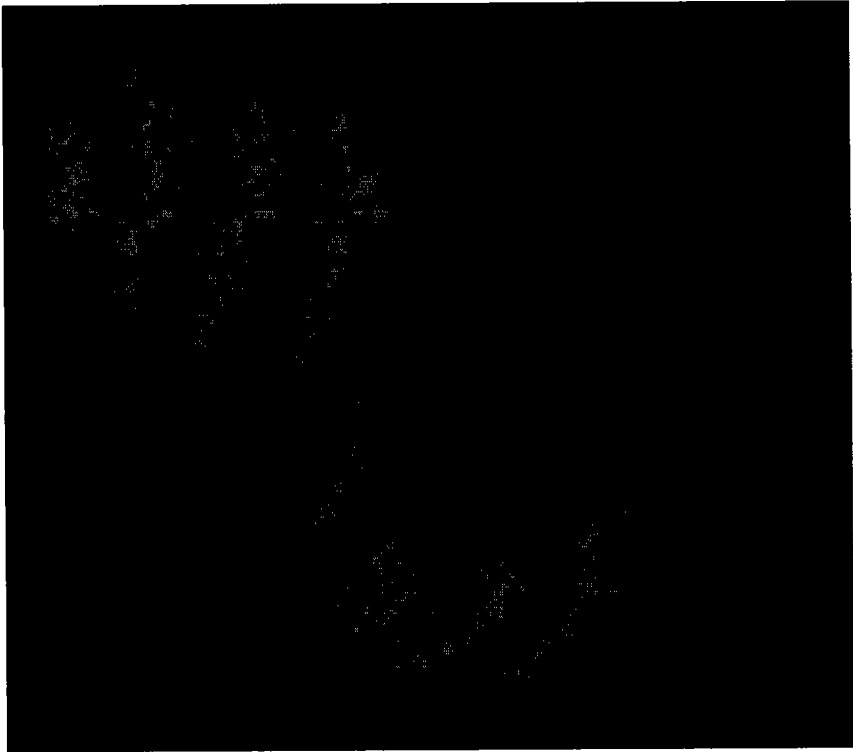


图 2

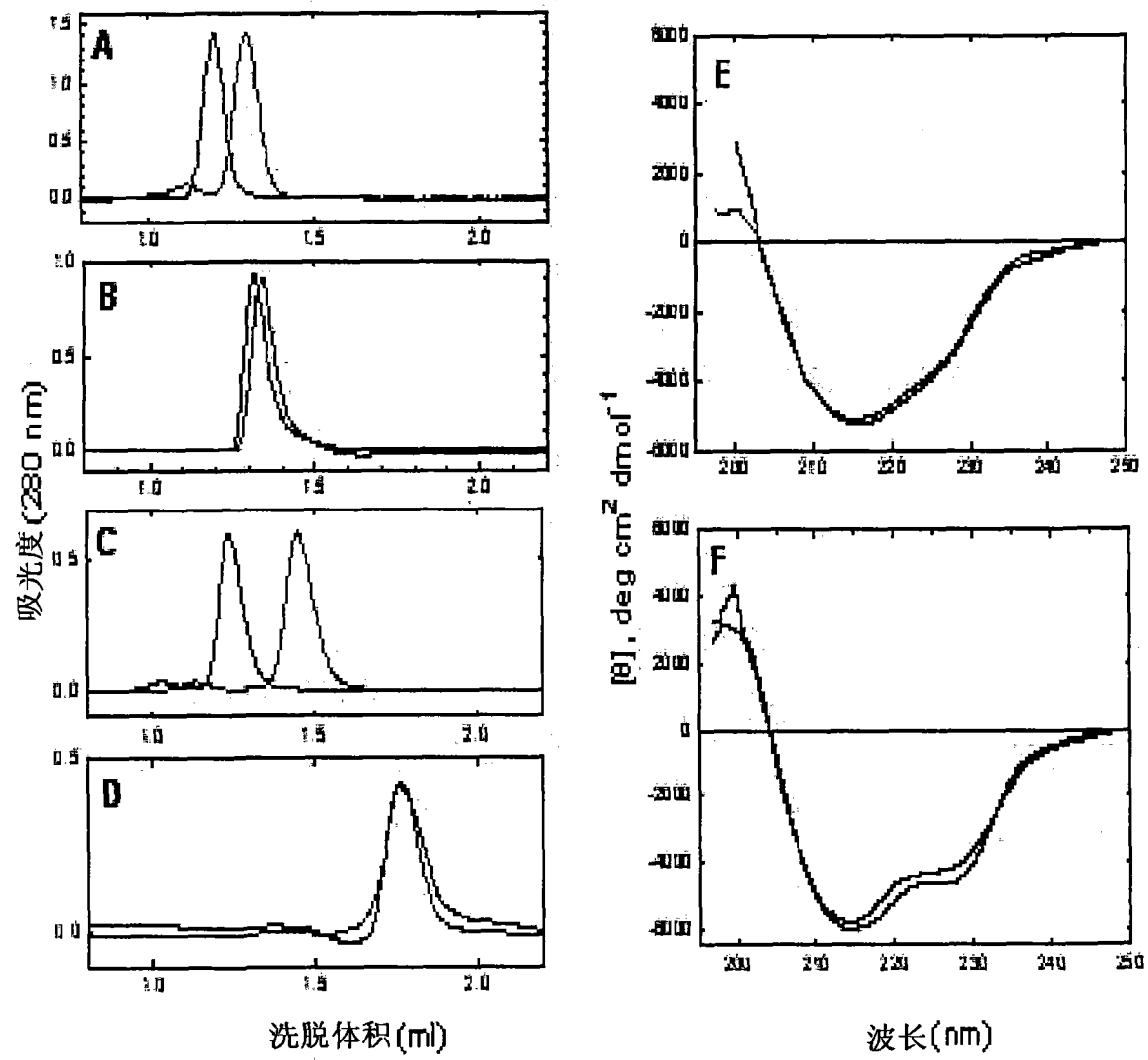


图 3

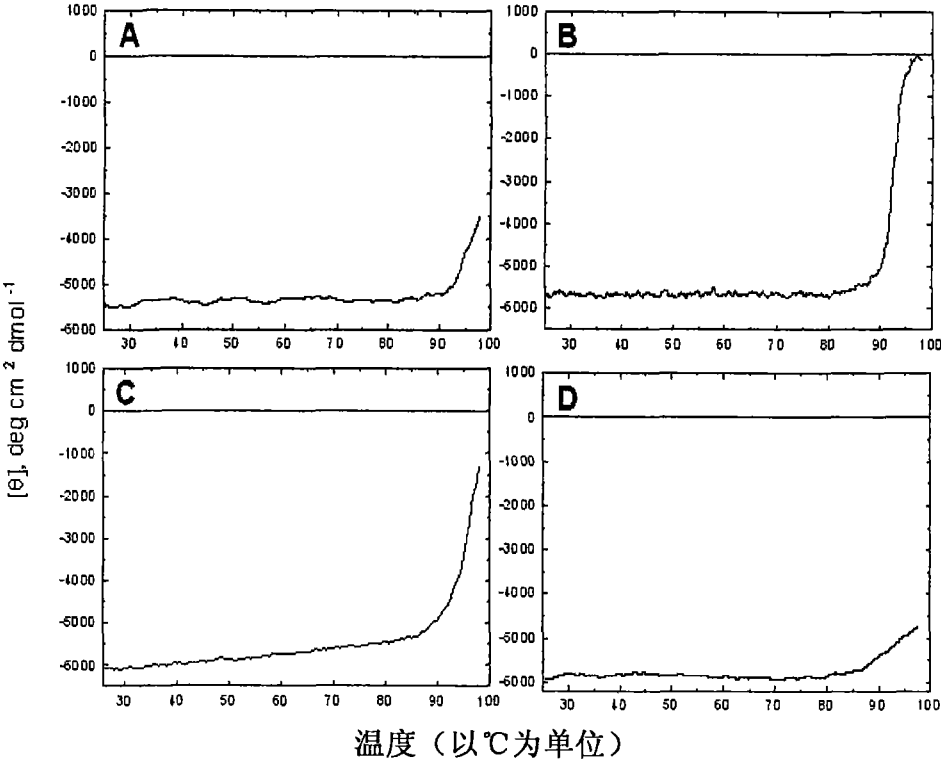


图 4

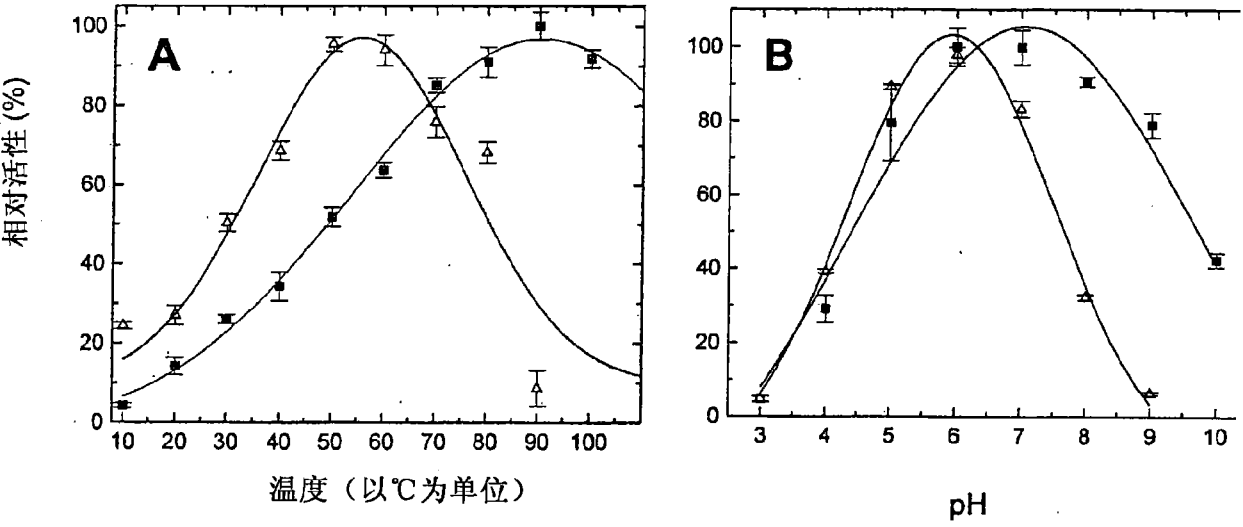


图 5

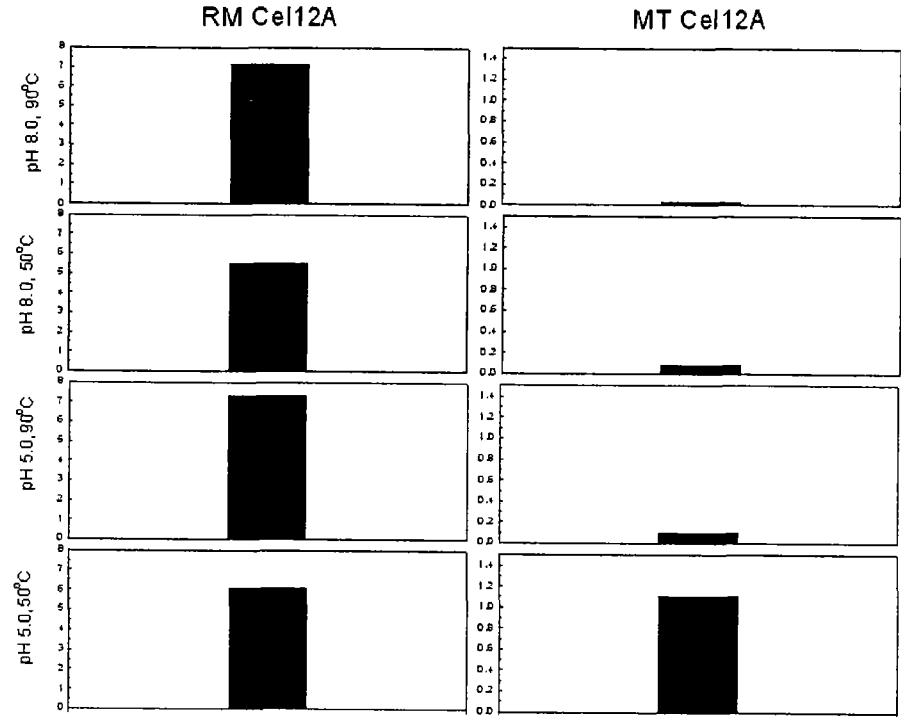


图 6

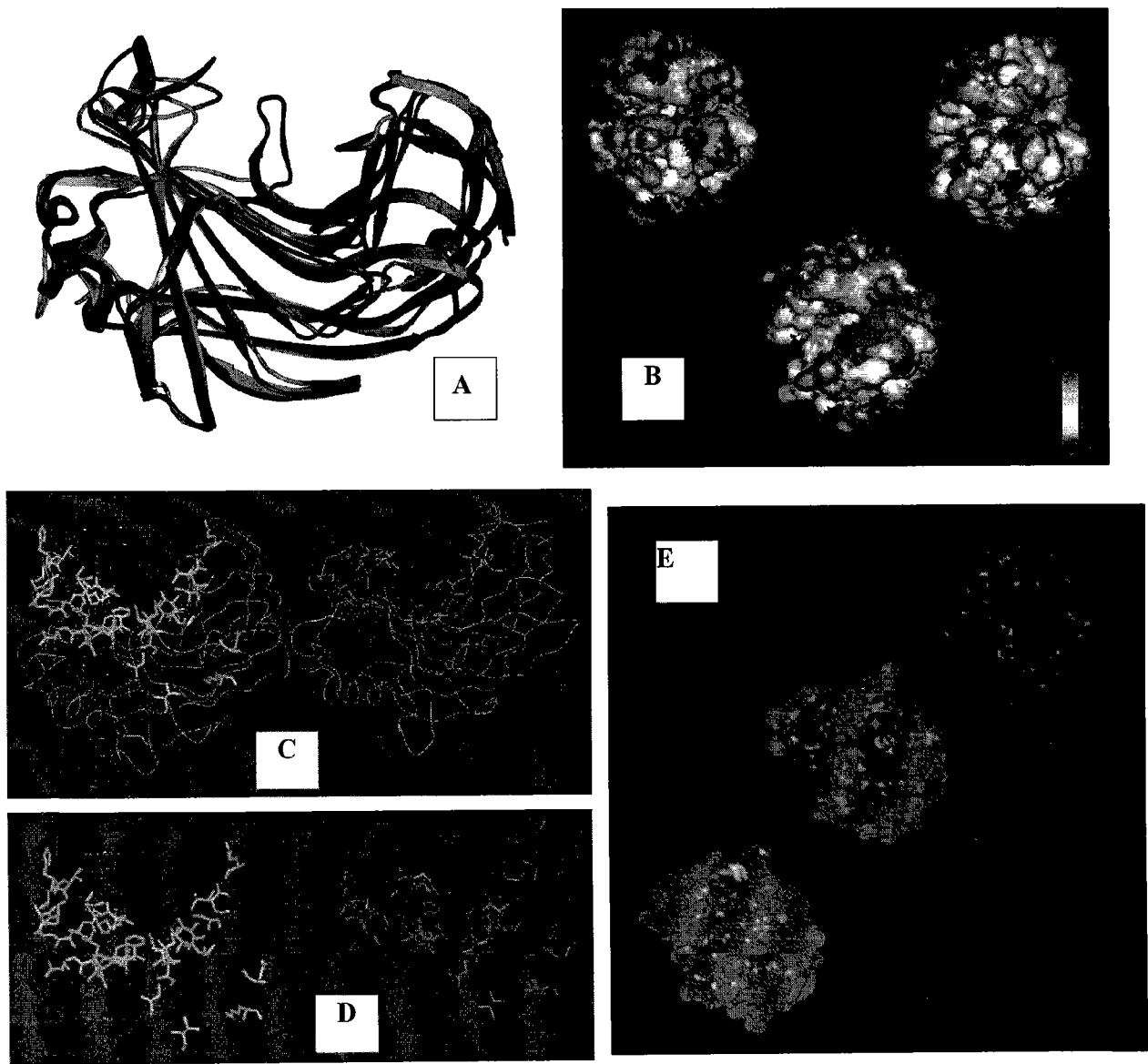


图 7

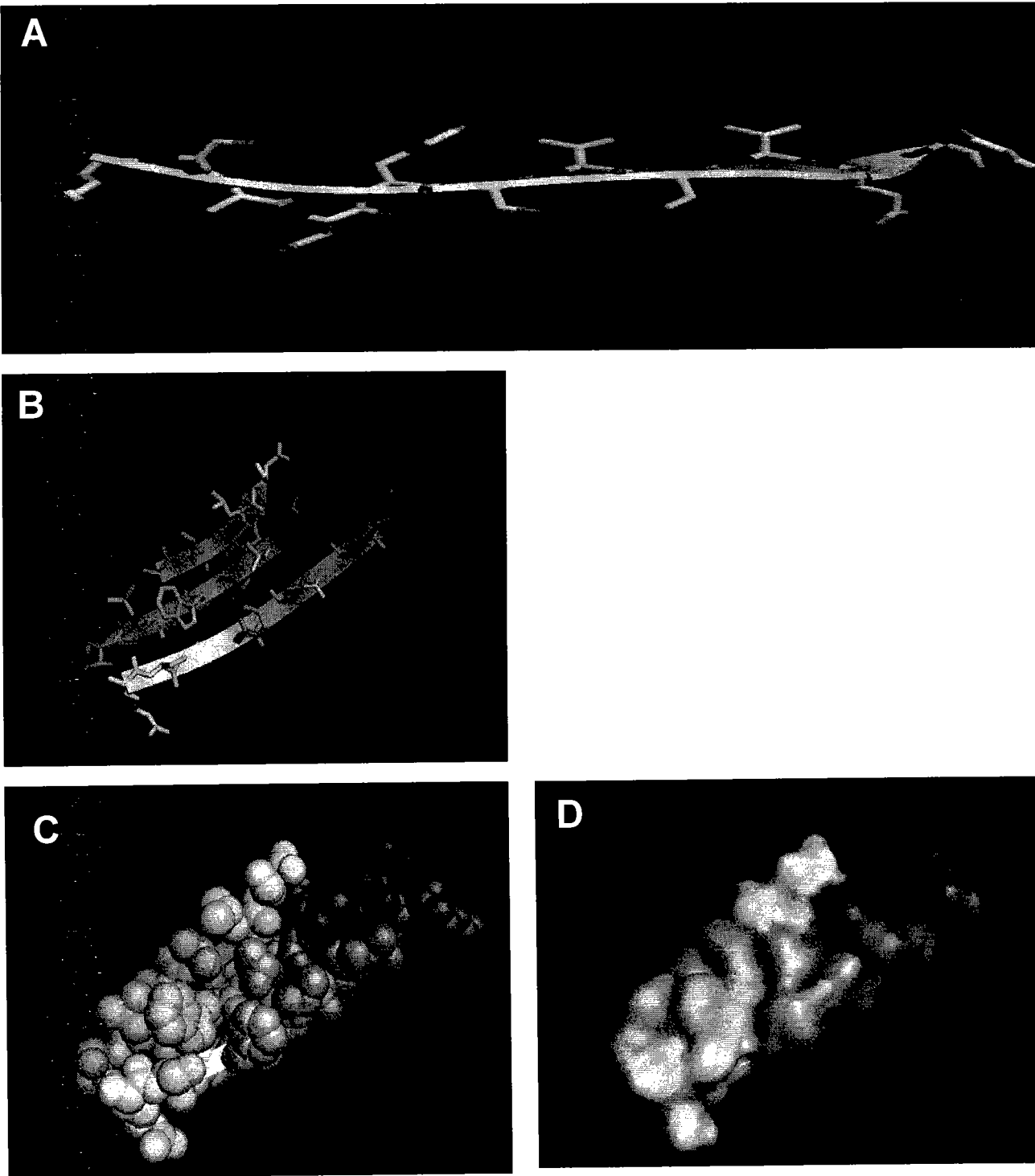


图 8

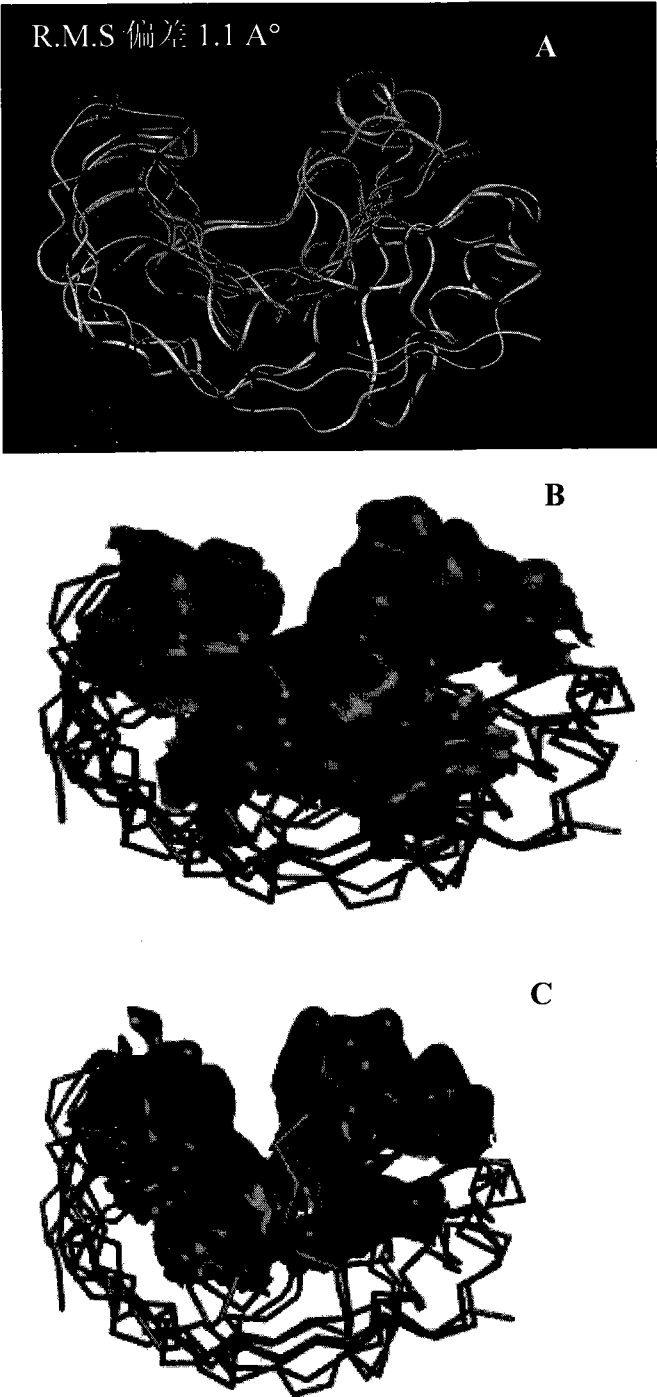


图 9

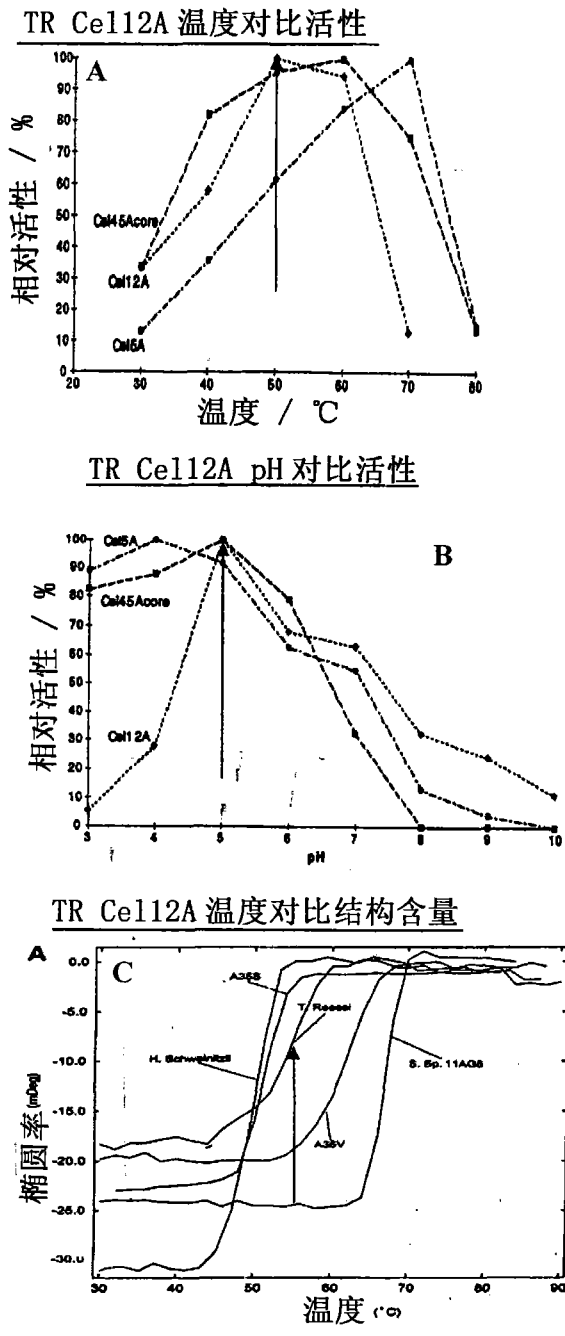


图 10

N- 端标记		M	R	G	S	H	H	H	H	H	H	G	S	
-12														
1		ATG	AGA	GGA	TCG	CAT	CAC	CAT	CAC	CAT	CAC	GGA	TCC	
1		ATG	AGA	GGA	TCG	CAT	CAC	CAT	CAC	CAT	CAC	GGA	TCC	
-12		M	R	G	S	H	H	H	H	H	H	G	S	

2	T	V	E	L	C	G	R	W	D	A	R	D	V	A	G	G	R	Y	R	V	21
	ACT	GTC	GAG	CTG	TGT	GGT	AGA	TGG	GAC	GCG	CGC	GAT	GTG	GCT	GGT	GGG	CGC	TAC	CCG	GTG	
	ACG	GTC	GAG	CTG	TGC	GGG	CAG	TGG	GAC	ACG	AGA	ACG	GTG	GCT	GGG	GGG	CGC	TAC	ACG	GTG	
1	T	V	E	L	C	G	Q	W	D	T	R	T	V	A	G	G	R	Y	T	V	20

22	I	N	N	V	W	G	A	E	T	A	Q	C	I	E	V	G	L	E	T	G	41
	ATC	AAC	AAC	GTA	TGG	GGT	GCG	GAG	ACC	GCA	CAG	TGC	ATC	GAG	GTC	GGG	CTG	GAG	ACG	GGC	
	AGC	AAC	AAC	GTA	TGG	GGG	GCG	GAG	ACC	GCA	CAG	TGC	ATC	GAG	GTC	GGG	CTG	GAG	ACG	GGC	
21	S	N	N	V	W	G	A	E	T	A	Q	C	I	E	V	G	L	E	T	G	40

42	N	F	T	I	T	R	A	E	H	D	N	G	N	D	V	A	A	Y	P	A	61
	AAT	TTT	ACG	ATC	ACG	CGG	GCC	GAG	CAC	GAC	AAC	GGC	AAC	GAC	GTG	GCC	GCC	TAT	CCG	GCC	
	AAT	TTT	ACG	ATC	ACG	CGG	GCC	GAG	CAC	GAC	AAC	GGC	AAC	AAC	GTG	GCC	GCC	TAT	CCG	AAC	
41	N	F	T	I	T	R	A	E	H	D	N	G	N	N	V	A	A	Y	P	N	60

62	I	Y	F	-	-	-	-	G	C	H	W	G	A	C	T	N	S	S	G	L	77
	ATC	TAT	TTT	-	-	-	-	GGG	TGC	CAC	TGG	GGC	GCC	TGC	ACA	AAT	AGT	TCG	GGA	TTG	
	ATC	CAG	TTT	GCC	ATC	CCG	CAG	-	-	-	-	-	-	-	-	-	-	-	-	-	
61	I	Q	F	A	I	P	Q	-	-	-	-	-	-	-	-	-	-	-	-	-	67

78	P	R	R	V	Q	E	L	S	D	V	R	T	S	W	T	L	T	P	I	T	97
	CCG	CGG	CGC	GTG	CAG	GAA	CTG	TCC	GAC	GTG	CGC	ACG	AGC	TGG	ACG	CTC	ACG	CCG	ATC	ACG	
	CCG	CGC	CGI	GTG	CAA	GAA	CTG	TCC	GAC	GTG	CGC	ACG	AGC	TGG	ACG	CTC	ACG	CCG	ATC	ACG	
68	P	R	R	V	Q	E	L	S	D	V	R	T	S	W	T	L	T	P	I	T	87

98	T	G	R	W	N	A	A	Y	D	I	W	F	S	P	G	T	N	-	-	-	114
	ACA	GGC	CGC	TGG	AAT	GCC	GCC	TAC	GAC	ATC	TGG	TTT	AGT	CCC	GCC	ACG	AAT	-	-	-	
	ACA	GGG	CGC	TGG	AAT	GCC	GCC	TAC	GAC	ATC	TTT	TTT	GCC	GCC	AAC	CCA	AGC	CAC	GTG	ACG	
88	T	G	R	W	N	A	A	Y	D	I	F	F	A	A	N	P	N	H	V	T	107

115	S	S	N	G	Y	S	G	G	A	E	L	M	I	W	L	N	W	N	G	G	134
	TCC	AGC	AAC	GGC	TAC	AGC	GGA	GGC	GCC	GAG	CTG	ATG	ATC	TGG	CTG	AAC	TGG	AAC	GGC	GGC	
	-	-	-	-	TAT	TCT	GGA	GAC	GCC	GAG	CTG	ATG	ATC	TGG	CTG	AAC	AAA	AAC	GGC	GAT	
108	-	-	-	-	Y	S	G	D	A	E	L	M	I	W	L	N	K	N	G	D	123

135	V	M	P	G	G	S	R	V	A	T	V	E	L	A	G	A	T	W	E	V	154
	GTG	ATG	CCG	GGT	GGC	AGC	CGC	GTG	GCC	ACC	GTG	GAA	CTG	GCC	GGG	GCC	ACC	TGG	GAA	GTC	
	GTG	ATG	CCT	ATC	GGC	AGC	CGC	GTG	GCC	ACC	GTG	GAA	CTG	GCC	GGG	GCC	ACC	TGG	GAA	GTC	
124	V	M	P	I	G	S	R	V	A	T	V	E	L	A	G	A	T	W	E	V	143

155	W	Y	A	D	W	D	-	W	N	Y	I	A	Y	R	R	T	T	P	T	T	173
	TGG	TAT	GCC	GAC	TGG	GAC	-	TGG	AAC	TAC	ATC	GCC	TAC	CGG	CGT	ACG	ACG	CCC	ACC	ACT	
	TGG	TAT	GCT	GAC	AAT	GGT	GCG	ATG	AAT	GTG	ATC	AGC	TAC	GTG	AGG	ACG	ACG	CCC	ACC	ACT	
144	W	Y	A	D	N	G	A	M	N	V	I	S	Y	V	R	T	T	P	T	T	163

174	S	V	T	E	L	D	L	K	A	F	I	D	D	A	V	A	R	G	Y	I	193
	TCG	GTG	ACC	GAA	CTG	GAC	CTG	AAG	GCC	TTT	ATC	GAC	GAC	GCG	GTC	GCC	CGC	GGC	TAC	ATC	
	TCG	GTG	ACC	GAA	CTG	GAC	CTG	AAG	GCC	TTT	ATC	GAC	GAC	GCG	GTC	GCC	CGC	GGC	TAC	ATC	
164	S	V	T	E	L	D	L	K	A	F	I	D	D	A	V	A	R	G	Y	I	183

194	R	P	E	W	Y	L	H	A	V	E	T	G	F	E	L	W	E	G	G	A	213
	CCG	CCG	GAG	TGG	TAC	CTG	CAC	GCG	GTA	GAG	ACG	GGC	TTT	GAA	CTC	TGG	GAG	GGC	GGG	GCC	
	CCG	CCG	GAG	TGG	TAC	CTG	CTG	TCG	GTA	CAA	ACG	GGC	TTT	GAA	CTC	TTT	ACT	GGT	GGT	GCC	
184	R	P	E	W	Y	L	L	S	V	Q	T	G	F	E	L	F	T	G	G	A	203

214	G	L	R	S	A	D	F	S	V	T	V	Q									225
	GGT	CTG	CGA	AGC	GCC	GAT	TTT	TCC	GTA	ACG	GTG	CAG									
	GGT	CTG	CGA	AGC	GCC	GAT	TTT	TCC	GTA	ACT	GTG	CAG									
204	G	L	R	S	A	D	F	S	V	T	V	Q									215

图 11

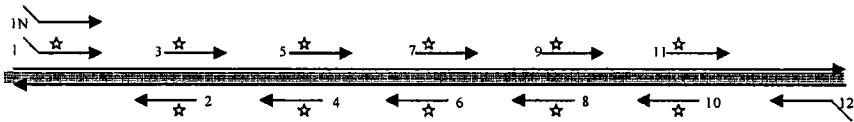


图 12

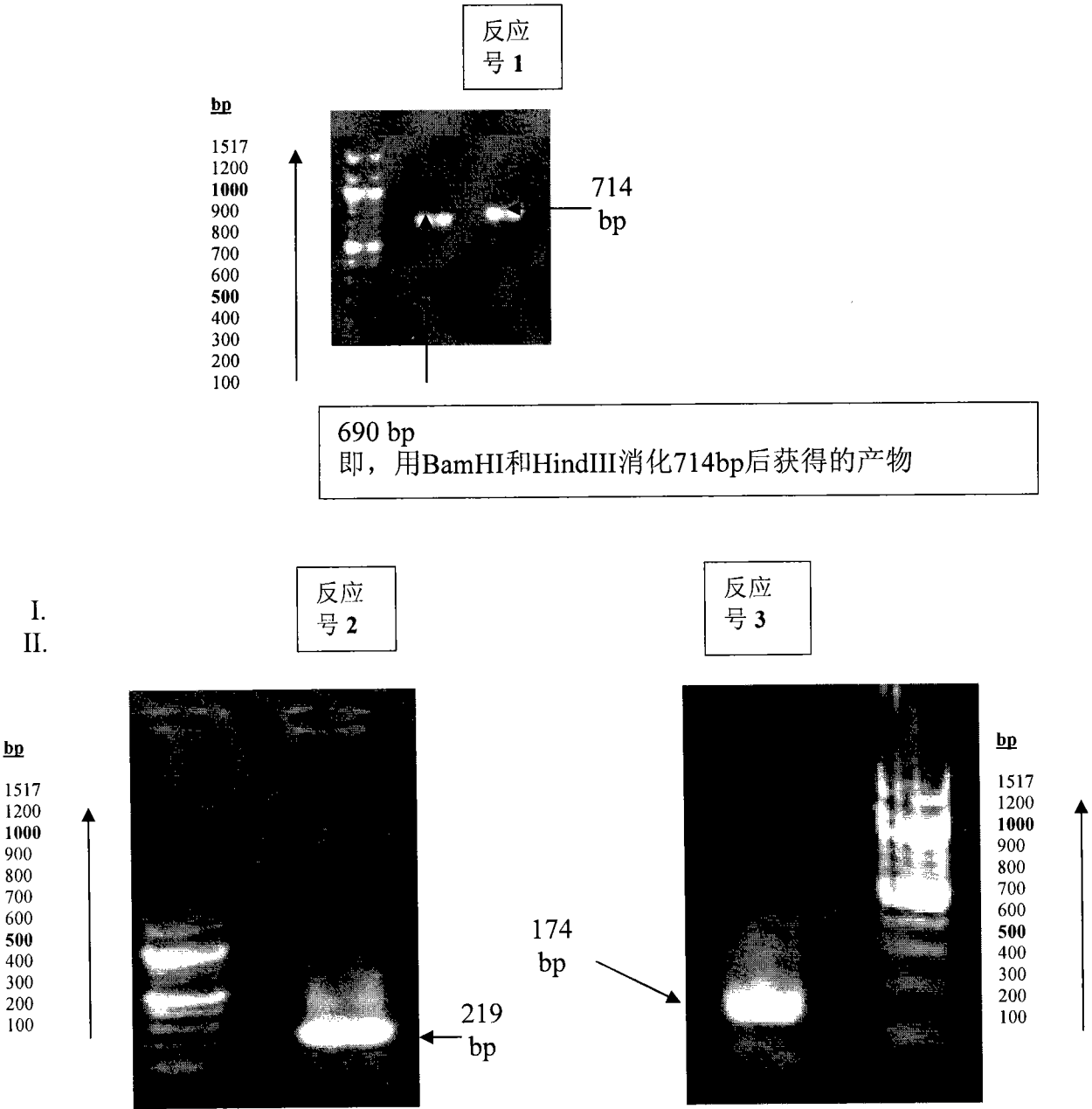


图 13

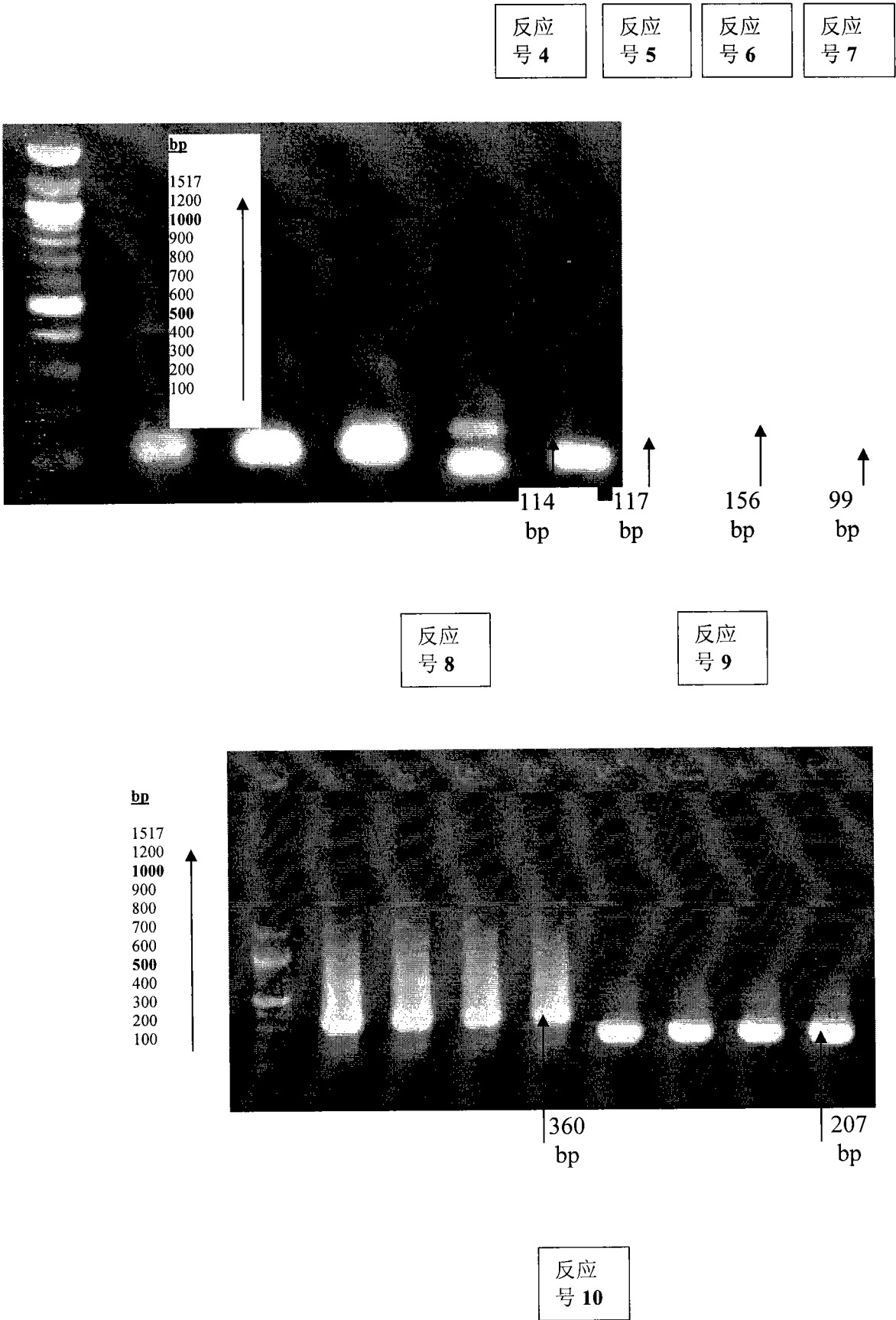


图 13（续）

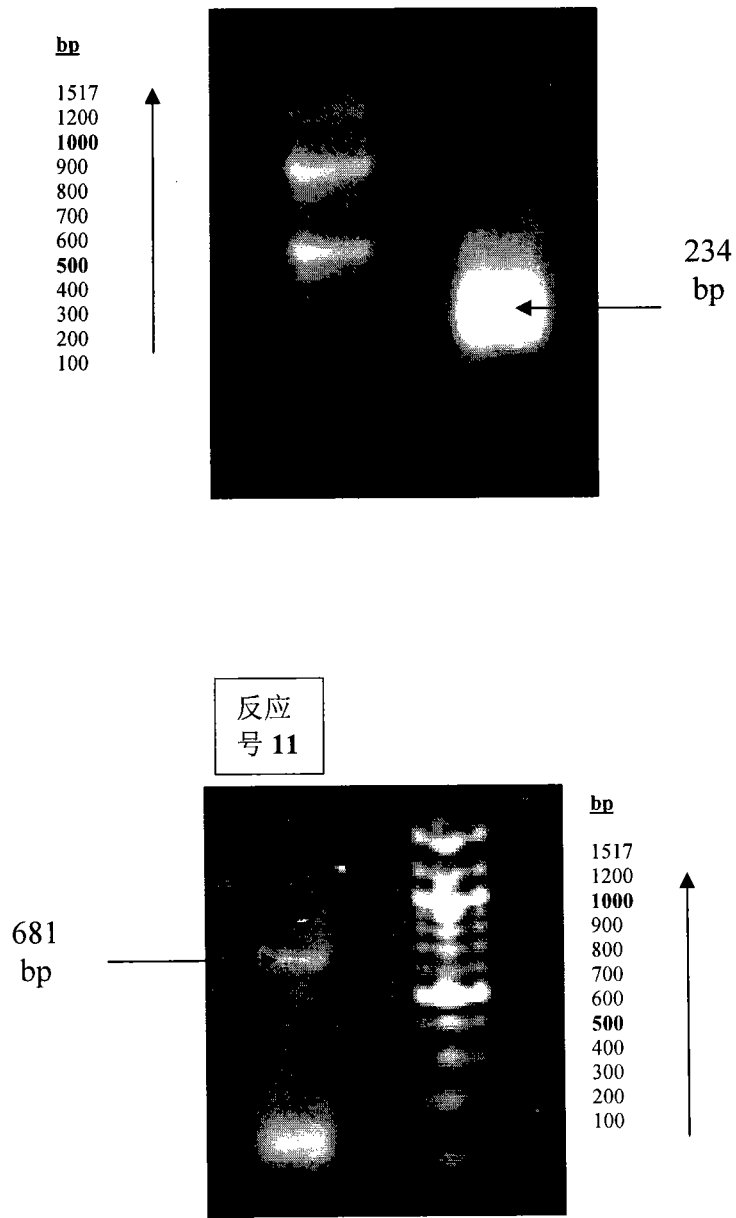


图 13 (续)

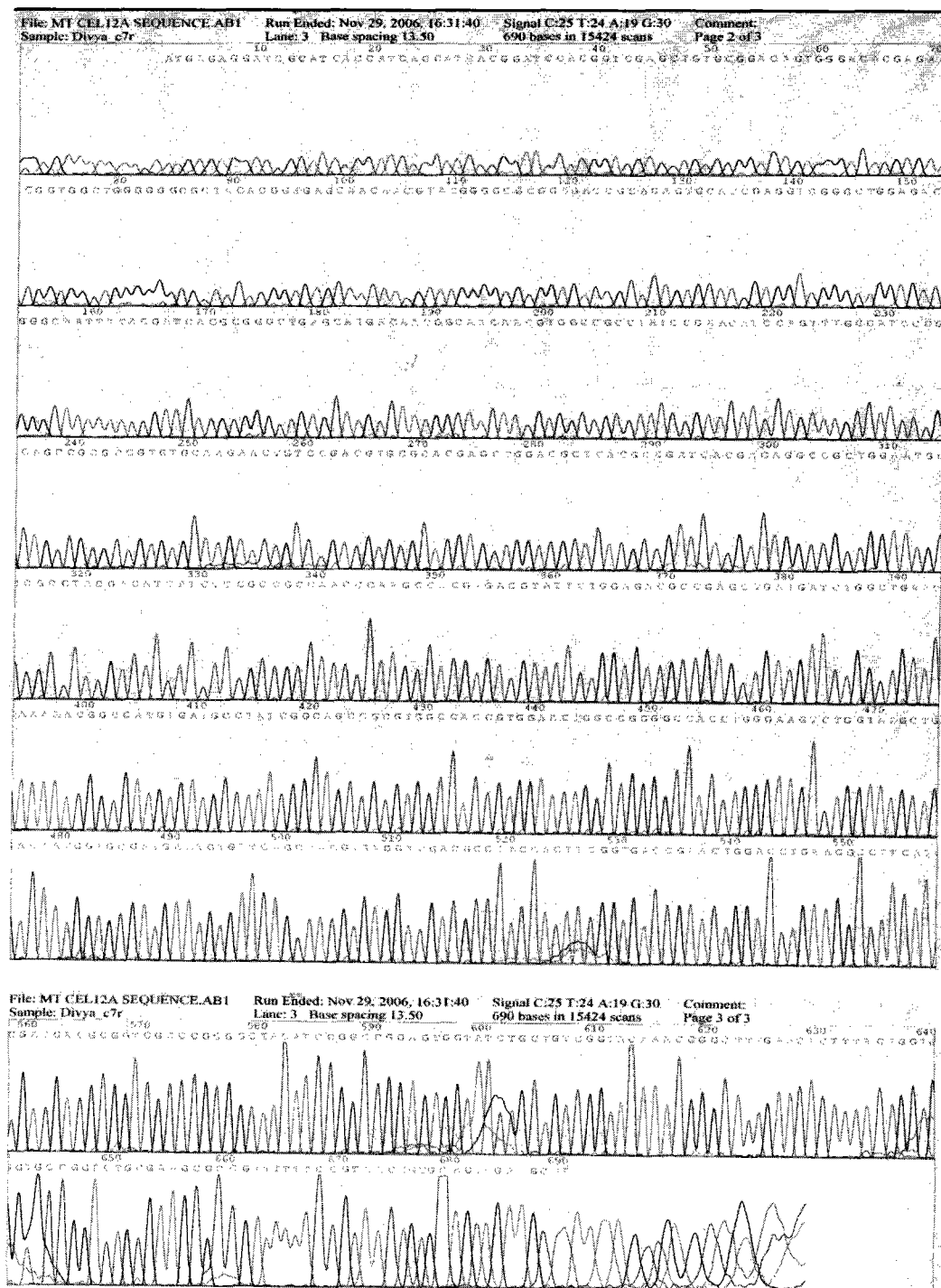


图 14

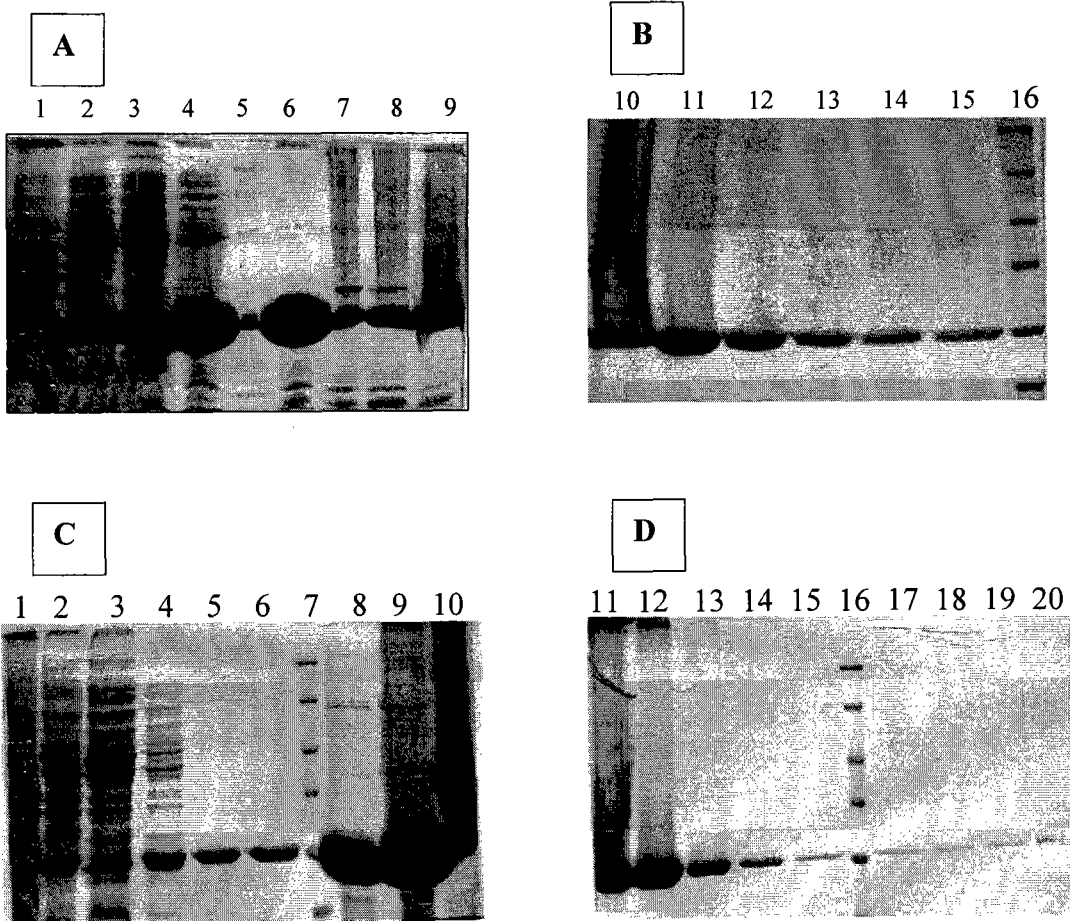
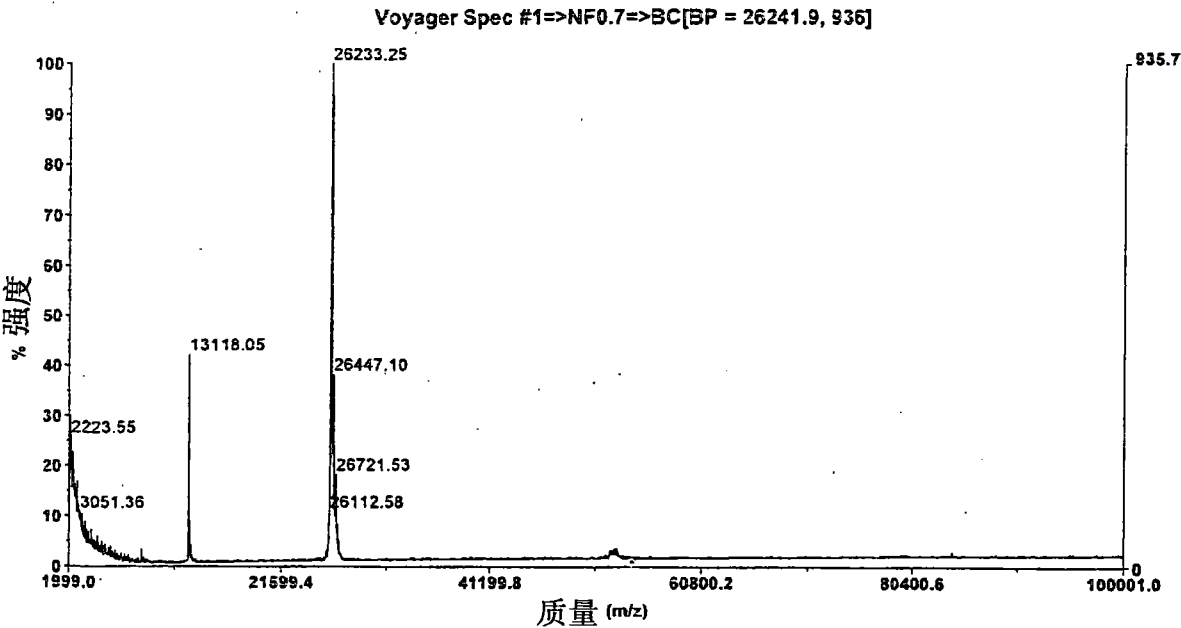


图 15

关于 RM Cell12A 的 MALDI-TOF MS 数据



关于 MT Cell12A 的 MALDI-TOF MS 数据

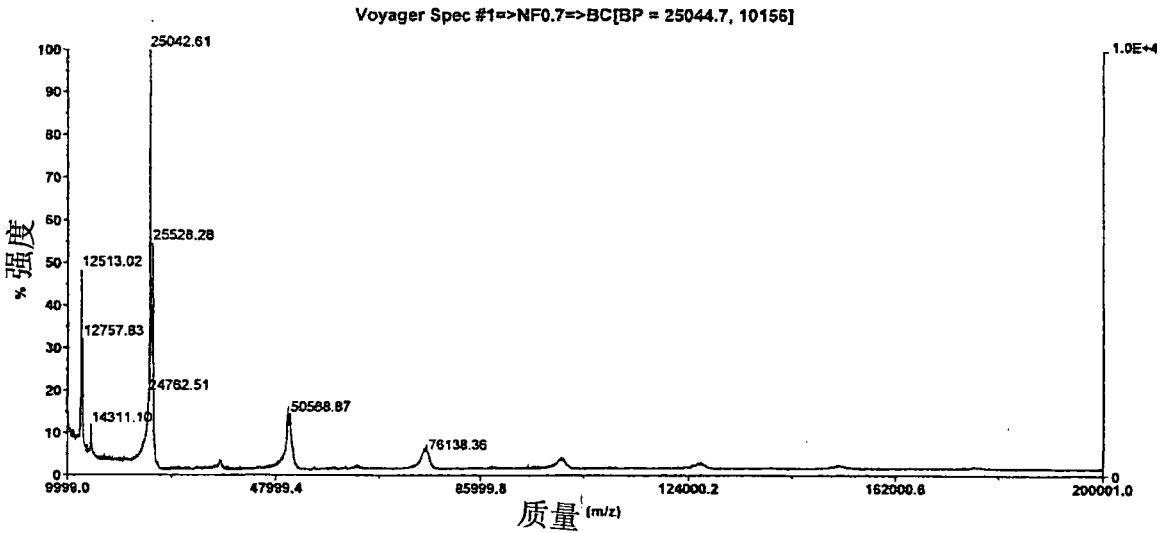


图 16

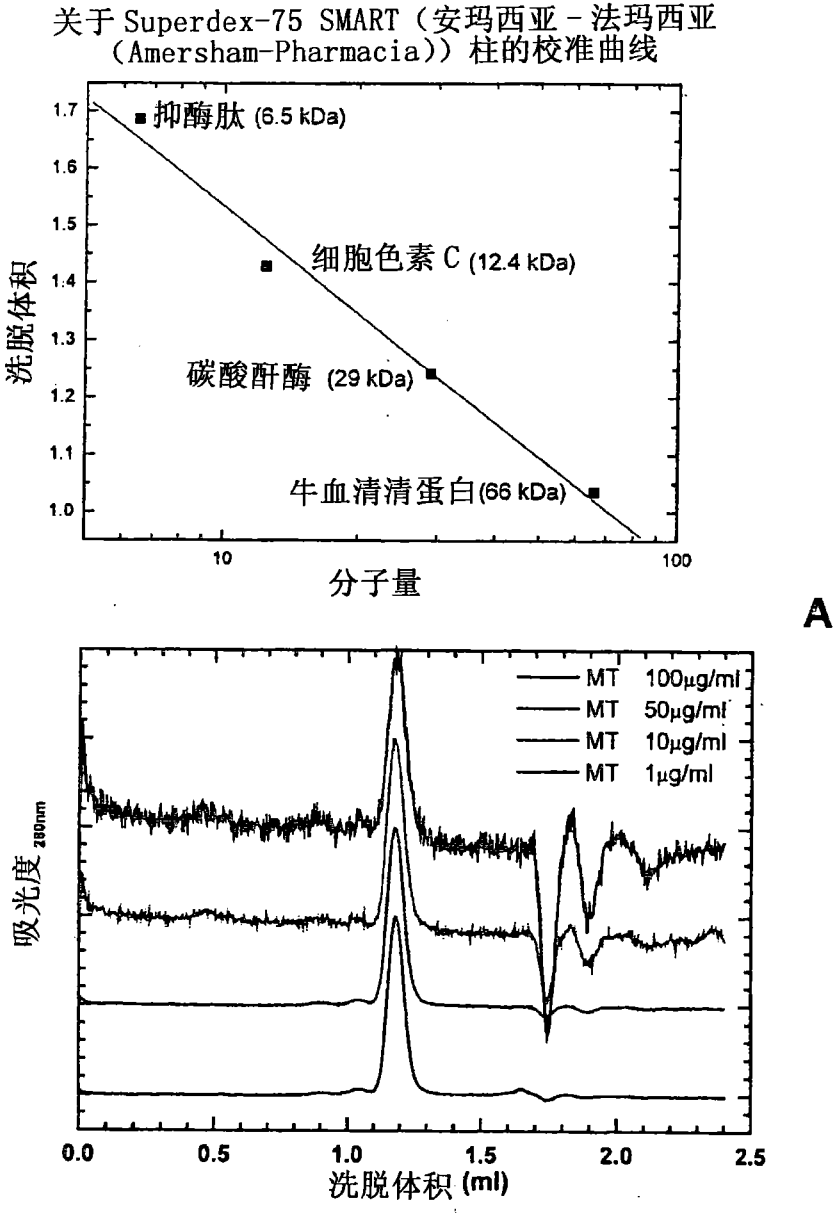


图 17