



(12)发明专利



(10)授权公告号 CN 104053667 B

(45)授权公告日 2017.09.15

(21)申请号 201280040585.9

(22)申请日 2012.06.21

(65)同一申请的已公布的文献号

申请公布号 CN 104053667 A

(43)申请公布日 2014.09.17

(30)优先权数据

11181357.2 2011.09.15 EP

61/571,159 2011.06.21 US

(85)PCT国际申请进入国家阶段日

2014.02.20

(86)PCT国际申请的申请数据

PCT/EP2012/062036 2012.06.21

(87)PCT国际申请的公布数据

W02012/175643 EN 2012.12.27

(73)专利权人 非营利性组织佛兰芒综合大学生
物技术研究所

地址 比利时根特

专利权人 布鲁塞尔自由大学

利兰斯坦福青年大学托管委员会

密执安大学董事会

(72)发明人 J·斯泰亚特 E·帕顿

T·拉尔曼斯 B·科比尔卡

S·拉斯穆森 S·格拉涅尔

R·K·苏娜哈拉

(74)专利代理机构 北京市中咨律师事务所

11247

代理人 史文静 黄革生

(51)Int.Cl.

C07K 14/435(2006.01)

C07K 16/28(2006.01)

审查员 马静

权利要求书2页 说明书67页

序列表34页 附图23页

(54)发明名称

针对GPCR:G蛋白复合物的结合结构域及来自其的用途

(57)摘要

本发明涉及G蛋白偶联受体(GPCR)的结构生物学和信号传递的领域。本发明特别涉及针对和/或特异性结合GPCR:G蛋白复合物的结合结构域。还提供了编码这类结合结构域的核酸序列,和表达或能够表达这类结合结构域的细胞。本发明的结合结构域可用作G蛋白偶联受体的结构和功能表征的通用工具,所述G蛋白偶联受体与下游的异源三聚体G蛋白复合,并与多种天然或合成配体结合,用于研究G蛋白活化的动力学特征,以及用于利用GPCR:G蛋白复合物的筛选和药物发现工作。

1. 结合结构域,其针对和/或特异性结合包含GPCR和G蛋白的复合物,其中所述结合结构域特异性结合G蛋白,而不结合GPCR,其中所述结合结构域是包含这样的氨基酸序列的免疫球蛋白单可变结构域,所述氨基酸序列包含根据下列通式(1)的4个框架区(FR1至FR4)和3个互补决定区(CDR1至CDR3): FR1-CDR1-FR2-CDR2-FR3-CDR3-FR4 (1),其中CDR1是SEQ ID NO:13,CDR2是SEQ ID NO:25,且CDR3是SEQ ID NO:37;或者其中CDR1是SEQ ID NO:14,CDR2是SEQ ID NO:26,且CDR3是SEQ ID NO:38;或者其中CDR1是SEQ ID NO:15,CDR2是SEQ ID NO:27,且CDR3是SEQ ID NO:39;或者其中CDR1是SEQ ID NO:16,CDR2是SEQ ID NO:28,且CDR3是SEQ ID NO:40;或者其中CDR1是SEQ ID NO:17,CDR2是SEQ ID NO:29,且CDR3是SEQ ID NO:41;或者其中CDR1是SEQ ID NO:18,CDR2是SEQ ID NO:30,且CDR3是SEQ ID NO:42。
2. 根据权利要求1的结合结构域,其中,相比与异源三聚体G蛋白单独结合,所述结合结构域与复合物结合的亲和力更高。
3. 根据权利要求1的结合结构域,其中复合物还包含受体配体。
4. 根据权利要求3的结合结构域,其中复合物由GPCR、G蛋白和受体配体组成。
5. 根据权利要求3或4的任一项的结合结构域,其中所述受体配体是激动剂。
6. 根据权利要求1的结合结构域,其中G蛋白处于不含核苷酸的形态中。
7. 根据权利要求1的结合结构域,其特异性结合位于所述G蛋白的 α 和 β 亚基之间的界面上的构象表位。
8. 根据权利要求1的结合结构域,其
 - a) 在存在核苷酸条件下阻止或抑制复合物解离,或
 - b) 阻止或抑制核苷酸与G蛋白结合,或
 - c) 能从G蛋白上移走核苷酸。
9. 根据权利要求6或8的结合结构域,其中所述核苷酸是鸟嘌呤核苷酸或其类似物。
10. 根据权利要求1至4和6至8的任一项的结合结构域,其中GPCR处于活性构象中。
11. 根据权利要求1至4和6至8的任一项的结合结构域,其中G蛋白选自Gs、Gi、Go、Gt、Ggust、Gz、Golf、Gq、G12、G13。
12. 根据权利要求1至4和6至8的任一项的结合结构域,其中GPCR选自Gs偶联受体、Gi偶联受体、Go偶联受体、Gt偶联受体、Ggust偶联受体、Golf偶联受体、Gq偶联受体、G12偶联受体、G13偶联受体。
13. 根据权利要求1至4和6至8的任一项的结合结构域,其中所述GPCR是人蛋白质。
14. 根据权利要求1至4和6至8的任一项的结合结构域,其中免疫球蛋白单可变结构域是纳米抗体(V_{HH})。
15. 根据权利要求14的结合结构域,其中纳米抗体由选自SEQ ID NO:1至6的氨基酸序列组成。
16. 根据权利要求1至4和6至8的任一项的结合结构域,其中所述结合结构域包含在多肽中。
17. 根据权利要求1至4和6至8的任一项的结合结构域,其中所述结合结构域被固定在固体支持物上。

18. 包含根据权利要求1至17的任一结合结构域的复合物。
19. 根据权利要求18的复合物,其还包含GPCR、G蛋白和任选的受体配体。
20. 根据权利要求18至19的任一结合结构域的复合物,其是结晶的。
21. 核酸序列,其编码根据权利要求1至17的任一结合结构域的氨基酸序列。
22. 重组载体,其包含根据权利要求21的核酸序列。
23. 细胞,其包含根据权利要求22的载体或根据权利要求21的核酸序列。
24. 根据权利要求23的细胞,其表达或能够表达GPCR和/或G蛋白。
25. 根据权利要求23至24的任一结合结构域的细胞的细胞培养物。
26. 根据权利要求23至24的任一结合结构域的细胞或根据权利要求25的细胞培养物的膜制品。
27. 根据权利要求1至17的任一结合结构域的用途,用于使处于功能构象状态的复合物稳定化,所述复合物包含GPCR和G蛋白和任选的受体配体。
28. 根据权利要求27的用途,其用于阻止复合物在存在核苷酸条件下解离,所述核苷酸是鸟嘌呤核苷酸或其类似物。
29. 根据权利要求1至17的任一结合结构域的用途,用于使复合物结晶和/或解析复合物结构,所述复合物包含GPCR和G蛋白和任选的受体配体。
30. 根据权利要求1至17的任一结合结构域或根据权利要求23至24的任一结合结构域的细胞或根据权利要求26的膜制品的用途,用于筛选调控GPCR的信号传递活性的化合物。
31. 根据权利要求27至30的任一结合结构域的用途,其中GPCR处于活性构象中。
32. 根据权利要求1至17的任一结合结构域的用途,用于捕获一个或多个相互作用的蛋白质。
33. 捕获和/或纯化包含GPCR和G蛋白的复合物的方法,方法包括步骤:
 - a) 提供根据权利要求1至17的任一结合结构域,和
 - b) 允许结合结构域与包含GPCR和G蛋白以及任选地受体配体的复合物结合,和
 - c) 任选的分离步骤b)中形成的复合物。
34. 确定包含GPCR和G蛋白的复合物的晶体结构的方法,方法包括步骤:
 - a) 提供根据权利要求1至17的任一结合结构域,和
 - b) 允许结合结构域与包含GPCR和G蛋白和任选地受体配体的复合物结合,和
 - c) 使步骤b)中形成的复合物结晶。
35. 根据权利要求1至17的任一结合结构域的用途,用于调控GPCR受体信号传递。
36. 根据权利要求35的用途,用于调控G蛋白介导的GPCR受体信号传递。
37. 生产针对和/或特异性结合包含GPCR和G蛋白的复合物的结合结构域的方法,方法包括步骤:
 - a) 在合适的细胞表达系统中表达根据权利要求21的核酸,和任选的
 - b) 分离和/或纯化结合结构域。

针对GPCR:G蛋白复合物的结合结构域及来自其的用途

发明领域

[0001] 本发明涉及G蛋白偶联受体 (GPCR) 的结构生物学和信号传递的领域。本发明特别涉及针对和/或特异性结合GPCR:G蛋白复合物的结合结构域。还提供了编码这类结合结构域的核酸序列,和表达或能够表达这类结合结构域的细胞。本发明的结合结构域可用作G蛋白偶联受体的结构和功能表征的通用工具,所述G蛋白偶联受体与下游的异源三聚体G蛋白复合,并与各种天然或合成配体结合,所述工具用于研究G蛋白活化的动力学特征,以及用于利用GPCR:G蛋白复合物的筛选和药物发现工作。

[0002] 背景

[0003] 7次跨膜受体 (7TMR),也被称为G蛋白偶联受体 (GPCR) 是人基因组中的最大类的受体,是用于医学治疗的最常见靶向蛋白质类型。在过去60年间,从药理学至体内功能表征,对不同GPCR的理解发生了实质性进展。目前高分辨率的结构研究为GPCR活化和构成的活性的分子机制提供了深入观察(例如,Rasmussen等人,2011)。然而,尚缺乏GPCR如何与其下游靶标相互作用并调控其活性的分子细节。与其下游蛋白质复合的GPCR的结构令人非常感兴趣,不仅因为这类相互作用是药理学相关的,而且因为对分子间相互作用的原子水平的理解是破解功能选择性的秘密、不同激动剂使一种受体产生不同下游效应的能力的关键。现有的结构数据支持这样的观点,即,即使尺寸小,GPCR也是具有多种信号传递输出的复杂变构组织。

[0004] 一旦激活,GPCR就以GTP依赖性的方式,通过三种蛋白质的复合物传达信号,所述复合物被称为异源三聚体G蛋白或 $G\alpha\beta\gamma$ 。胞外配体与GPCR的结合调控了催化 $G\alpha\beta\gamma$ 中的GDP-GTP交换的能力,从而调控了第二信使的胞内水平。失活的 $G\alpha\beta\gamma$ 异源三聚体包括2个主要元件: $G\alpha \cdot GDP$ 和 $G\beta\gamma$ 异源二聚体。 $G\beta\gamma$ 隔离 $G\alpha$ 上的switch II元件,使其不能与第二信使系统相互作用,所述第二信使系统例如涉及cAMP、二酰基甘油和钙的那些。活化的GPCR催化从 $G\alpha$ 释放GDP,允许GTP结合和释放活化的 $G\alpha$ -GTP亚基。在这一状态下,switch II形成了被GTP的 γ 磷酸稳定的螺旋,使其能够与效应子,如腺苷酸环化酶相互作用。虽然关于 $G\alpha$ 亚基如何与其下游靶标相互作用并调控其活性的理解取得了较大进展,但尚不清楚活化的GPCR如何通过催化 $G\alpha\beta\gamma$ 上的核苷酸交换而启动该过程。

[0005] 药物研发的工作一般集中在与特定催化位点或活性位点竞争性结合的小分子配体上,利用靶的静态模型作为起点。该方法已鉴别和验证了多个目前正在使用的可行的活性位点治疗剂。然而,新药化合物的高失败率(仅约8%的I期临床治疗剂最终获得食品药品监督管理局(Food and Drug Administration)批准,每种药物的保守成本为8亿美元)反映出多数工作是不成功的,通常一旦靶被视为不可用于药物研发(undrugable)就会被放弃(Lee& Craik,2009)。相当一部分的这类失败是由于所讨论的靶的最主要的构象不对应可用于药物研发(drugable)的构象,后者是对治疗适应症而言药物必须结合才有效的构象。例如,已经证实了获得结合激动剂的活性状态的GPCR结构的尝试是非常困难的,因为在缺少G蛋白的条件下,该状态是内在不稳定的。目前,利用构象选择性稳定的纳米抗体或 **Xaperones®**,开始能够获得活性状态的GPCR的结构,所述纳米抗体或 **Xaperones®**

模拟G蛋白并且增加了激动剂在orthosteric位点的亲和力(Rasmussen等人,2011)。这证实了**Xaperones®**使最有挑战性的药物靶的结构锁定在治疗相关构象中的能力(Steyaert&Kobilka,2011),及其用于定向药物研发的用途,所述定向药物研发允许特异性的筛选对特定的靶具有较高灵敏度和选择性的潜在药物(WO2012007593)。该技术方法的一个限制是需要为每个GPCR靶鉴别特定的稳定化纳米抗体,这不仅费时费钱,而且暗示要利用不同的工具,如用于免疫和选择目的的生物材料等。

[0006] 因此,需要研发新的直接的工具,用于GPCR药物靶的结构和药理学分析。

[0007] 发明概述

[0008] 阐明与其下游异源三聚化G蛋白复合且结合至多种天然和合成配体的GPCR的结构和功能特征,对于理解GPCR信号转导的机制以及药物研发工作都是有价值的。例如,迄今为止尚未成功获得活性状态的三元复合物(包含激动剂、GPCR和G蛋白)的结构,由于对其生物学的理解贫乏,获得该结构是信号转导的结构生物学非常需要的。该复合物的晶体生成被证实是非常困难的,因为一个伙伴部分(受体)需要去垢剂使其溶解,而G蛋白在去垢剂中则是不稳定的。此外,复合物形成所需要的核苷酸也在瞬时过程中使复合物解离。

[0009] 因此,同时也是非常令人意外的是,本发明人鉴别了使GPCR和G蛋白的复合物稳定的工具,该工具能够捕获和纯化这类复合物,最终使这类复合物结晶。这有利于通过例如基于病毒筛选或设计的结构、高通量筛选或基于片段的药物发现,鉴别配体或药物化合物(参见例如,实施例10)。更具体而言,本发明人鉴别了适用于活性状态的复合物(包含激动剂、GPCR和G蛋白)的结构和功能分析的结合结构域,特别是免疫球蛋白单可变结构域(参见例如,实施例4-7)。令人感兴趣的是,证实了一些上述GPCR:G蛋白复合物-选择性结合结构域是特异性针对G蛋白的,而不针对GPCR。例如,在不与 β -肾上腺素能受体接触的条件下,鉴别了结合位于的 $G_{\alpha s}$ 和 $G_{\beta \gamma}$ 界面上的 G_s 的结合结构域(参见例如,实施例3),因此所述结合结构域可用于捕获和稳定其他的 G_s 偶联受体,如对精氨酸血管加压素受体2(V2R)证实的(参见例如,实施例9)。因此,特别有利的是,结合结构域针对GPCR:G蛋白复合物中的G蛋白,因为这类结合结构域可用作稳定和捕获与特定G蛋白相互作用的GPCR范围的活性状态复合物的一般工具。

[0010] 因此,根据第一个方面,本发明涉及针对和/或特异性结合包含GPCR和G蛋白的复合物的结合结构域。更具体而言,相比分别单独与G蛋白和/或与GPCR结合,本文所述的结合结构域以更高的亲和力结合GPCR:G蛋白复合物。此外,本文所述的结合结构域增强了G蛋白对GPCR的亲和力。因此,本发明提供了针对和/或特异性结合包含GPCR和G蛋白的复合物的构象表位的结合结构域,所述结合结构域将复合物稳定或锁定在特定的构象状态,更具体而言是活性构象状态。

[0011] 一般而言,本文所述的结合结构域可结合任何由GPCR和G蛋白的复合物可利用或可接近的构象表位。这些构象表位可以由包含在复合物中的单个蛋白质呈递,和/或仅在形成复合物时呈递。此外,这些构象表位可以或可以不由单个蛋白质单独呈递。根据一个特定的实施方案,本文所述的结合结构域特异性结合包含在复合物中的G蛋白,而不结合GPCR。

[0012] 通常,G蛋白本质上是结合核苷酸的形式。更具体而言,G蛋白(或至少 α 亚基)结合GTP或GDP,这取决于特定GPCR的活化状态,如本文中另外描述的。结合GPCR的激动剂促进与结合了GDP的 $G_{\alpha \beta \gamma}$ 异源三聚体相互作用,导致 G_{α} 上的GDP交换为GTP和G蛋白功能性解离为G

α -GTP和G β γ 亚基,这是进一步的胞外信号传递所需的。在特定的实施方案中,本文所述的结合结构域在缺少核苷酸条件下特异性结合GPCR:G蛋白复合物,并使其稳定,更具体的是,结合结构域结合和稳定这样的GPCR:G蛋白复合物,其中G蛋白处于不含核苷酸的形态。在特定的实施方案中,本文所述的结合结构域特异性结合位于所述G蛋白的 α 和 β - γ 亚基之间的界面上的构象表位,并由此阻断GDP/GTP结合位点,和干扰GDP/GTP结合。由此且令人惊讶的,本文所述的结合结构域在存在核苷酸(特别是鸟嘌呤核苷酸或其类似物,例如GTP γ S)的条件下阻止或抑制GPCR:G蛋白复合物解离。本文所述的结合结构域还阻止或抑制核苷酸与G蛋白的结合。

[0013] 优选的,本文所述的结合结构域针对和/或特异性结合包含GPCR、G蛋白和一个或多个受体配体的复合物。通常,受体配体是激动剂、或正变构调节剂、或其组合。

[0014] 根据另一个优选的实施方案,本文所述的结合结构域针对和/或特异性结合包含Gs蛋白偶联受体和Gs蛋白的复合物;或包含Gi蛋白偶联受体和Gi蛋白的复合物;或包含Gt蛋白偶联受体和Gt蛋白的复合物;或包含Ggust蛋白偶联受体和Ggust蛋白的复合物;或包含Gz蛋白偶联受体和Gz蛋白的复合物;或包含Golf蛋白偶联受体和Golf蛋白的复合物;或包含Gq蛋白偶联受体和Gq蛋白的复合物;或包含G12偶联受体和G12蛋白的复合物;或包含G13偶联受体和G13蛋白的复合物。包含在复合物中的GPCR和/或G蛋白可来自相同或不同的物种,特别是来自哺乳动物物种。优选的,GPCR是人蛋白质。

[0015] 一般而言,本发明的结合结构域可以是能够特异性结合GPCR:G蛋白复合物的任何非天然存在的分子或其部分。特别的是,本文所述的结合结构域是包含这样的氨基酸序列的免疫球蛋白单可变结构域,所述氨基酸序列包含根据下列通式(1)的4个框架区(FR1至FR4)和3个互补决定区(CDR1至CDR3):

[0016] FR1-CDR1-FR2-CDR2-FR3-CDR3-FR4 (1)。

[0017] 优选的,本文所述的结合结构域是源自骆驼科物种的免疫球蛋白单可变结构域,特别是纳米抗体或V_HH。

[0018] 根据特定的实施方案,本文所述的结合结构域是包含这样的氨基酸序列的免疫球蛋白单可变结构域,所述氨基酸序列包含根据下列通式(1)的4个框架区(FR1至FR4)和3个互补决定区(CDR1至CDR3):

[0019] FR1-CDR1-FR2-CDR2-FR3-CDR3-FR4 (1),

[0020] 并且其中CDR1选自:

[0021] a) SEQ ID NO:13-18,

[0022] b) 与SEQ ID NO:13-18具有至少80%氨基酸同一性的多肽,

[0023] c) 与SEQ ID NO:13-18具有3、2或1个氨基酸差异的多肽,

[0024] 并且其中CDR2选自:

[0025] a) SEQ ID NO:25-30,

[0026] b) 与SEQ ID NO:25-30具有至少80%氨基酸同一性的多肽,

[0027] c) 与SEQ ID NO:25-30具有3、2或1个氨基酸差异的多肽,

[0028] 并且其中CDR3选自:

[0029] a) SEQ ID NO:37-42,

[0030] b) 与SEQ ID NO:37-42具有至少80%氨基酸同一性的多肽,

[0031] c) 与SEQ ID NO:37-42具有3、2或1个氨基酸差异的多肽。

[0032] 在特别优选的实施方案中,本发明提供了包含这样的氨基酸序列的免疫球蛋白单可变结构域,所述氨基酸序列包含根据下列通式(1)的4个框架区(FR1至FR4)和3个互补决定区(CDR1至CDR3):

[0033] FR1-CDR1-FR2-CDR2-FR3-CDR3-FR4 (1);

[0034] 其中CDR1是SEQ ID NO:13;其中CDR2是SEQ ID NO:25;且其中CDR3是SEQ ID NO:37。

[0035] 根据非常特定的实施方案,结合结构域,特别是免疫球蛋白单可变结构域具有选自SEQ ID NO:1至6的氨基酸序列。

[0036] 此外,本文描述的结合结构域还可以包含在多肽中。此外,结合结构域还可以固定在固体支持物上。

[0037] 在一个特定的方面,本发明涉及针对和/或特异性结合G蛋白的结合结构域。

[0038] 本发明的另一个方面设想了包含本文所述的结合结构域的复合物。复合物特别包含GPCR、G蛋白,和任选的受体配体。在某些应用中,本文所述的复合物是晶体。

[0039] 本发明还提供了核酸序列,特别是编码任何本文所述结合结构域的任何氨基酸序列的核酸序列,以及包含任何本文所述核酸序列的重组载体。本发明的特别优选的方面是包含任何本文所述载体或核酸的细胞,所述细胞可以表达或能够表达GPCR和/或G蛋白。根据本发明的细胞的细胞培养物及其来源的膜制品也落入本发明的范围内。

[0040] 本文描述的结合结构域、复合物和细胞可用于多种上下文和应用中。因此,本发明的一个方面相应的涉及本文所述的结合结构域的用途,用于使包含GPCR和G蛋白,和任选的受体配体的复合物稳定化,所述受体配体处于功能构象状态,更具体的处于活性构象状态。在一个特定的实施方案中,本文所述的结合结构域可用于阻止复合物在存在核苷酸条件下解离,所述核苷酸特别是鸟嘌呤核苷酸或其类似物,如GTP γ S。因此,作为工具稳定GPCR:G蛋白复合物和阻断功能构象状态(优选活性构象状态)的GPCR的结合结构域在一系列应用中是非常有效的,如下文概括的。

[0041] 本发明的目标是使用本文所述的结合结构域使复合物结晶和/或解析复合物结构,所述复合物包含GPCR和G蛋白和任选的受体配体。

[0042] 还认为落入本发明范围内的是使用本文所述的结合结构域,或本文所述的其来源的细胞或膜制品,来筛选调控GPCR的信号传递活性的化合物。

[0043] 此外,本文所述的结合结构域可用于捕获一个或多个相互作用的蛋白质,特别是与G蛋白和/或GPCR相互作用的蛋白质。

[0044] 根据特定的实施方案,本发明提供了用于捕获和/或纯化包含GPCR和G蛋白的复合物的方法,方法包括步骤:

[0045] a) 提供本文所述的结合结构域,和

[0046] b) 允许结合结构域与包含GPCR和G蛋白,和任选受体配体的复合物结合,和

[0047] c) 任选的分离步骤b)中形成的复合物。

[0048] 在另一个特定的实施方案中,本发明涉及确定包含GPCR和G蛋白的复合物的晶体结构的方法,方法包括步骤:

[0049] a) 提供本文所述的结合结构域,和

[0050] b) 允许结合结构域与包含GPCR和G蛋白,和任选受体配体的复合物结合,和

[0051] c) 使步骤b)中形成的复合物结晶。

[0052] 一些本文所述的结合结构域可具有治疗用途。因此,本发明的目标还是使用本文所述的结合结构域调控GPCR受体信号传递,特别是G蛋白介导的GPCR受体信号传递。

[0053] 本发明还涵盖了生产针对和/或特异性结合包含GPCR和G蛋白的复合物的结合结构域的方法,方法包括步骤:

[0054] a) 在合适的细胞表达系统中表达本文所述的核酸,和任选的

[0055] b) 分离和/或纯化结合结构域。

[0056] 本发明的另一个方面涉及筛选针对和/或特异性结合包含GPCR和G蛋白的复合物的结合结构域的方法,方法包括步骤:

[0057] a) 提供多个结合结构域,和

[0058] b) 在所述多个结合结构域中,筛选与包含GPCR和G蛋白的复合物结合的结合结构域,和

[0059] c) 分离与复合物结合的结合结构域。

[0060] 根据本文的其他描述,本发明的其他目标也是显而易见的。

[0061] 附图简述

[0062] 图1: β_2 AR:Gs复合物的G蛋白循环。a,与 β_2 AR结合的胞外激动剂导致跨膜区段的细胞质端的构象重排,所述区段使Gs异源三聚体(α 、 β 和 γ)结合受体(R、R*)。GDP是在形成R:G复合物时从 α 亚基释放的。GTP结合不含核苷酸的 α 亚基,导致 α 和 $\beta\gamma$ 亚从受体解离。亚基调控其相应的效应子蛋白——腺苷酸环化酶(AC)和 Ca^{2+} 通道。在 α 亚基中的GTP水解为GDP后,自 α 和 $\beta\gamma$ 亚基重新装配为Gs异源三聚体。b,维持在去垢剂微团中的纯化的不含核苷酸的 β_2 AR:Gs蛋白质复合物。Gs α 亚基由2个结构域组成:Ras结构域(α Ras)和 α -螺旋状结构域(α AH)。两个结构域都参与核苷酸结合。在不合核苷酸的状态下,相对于 α Ras结构域, α AH结构域具有可变的位置。

[0063] 图2:形成稳定的 β_2 AR:Gs复合物。通过下述的组合效应:1)以非常缓慢的解离速率使高亲和力激动剂与受体结合(如Rasmussen等人,2011所述);2)在存在腺苷三磷酸双磷酸酶的条件下,形成不含核苷酸的复合物,该酶水解释放的GDP,阻止它重新结合和导致较不稳定的R:G相互作用;和3)将DDM去垢剂交换为稳定复合物的MNG-3,实现稳定的 β_2 AR:Gs复合物。

[0064] 图3:核苷酸类似物、pH和纳米抗体对 β_2 AR:Gs复合物稳定性的效应。a)分析凝胶过滤显示了核苷酸GDP和GTP γ S(0.1mM)导致 β_2 AR-365:Gs复合物的解离。b)模拟核苷酸磷酸基的磷酸盐——焦磷酸盐和三钠磷酸甲酸盐(forcarnet)(5mM使用)不导致复合物解构。作为添加剂,它们改善了两种T4L- β_2 AR:Gs复合物(不含纳米抗体),T4L- β_2 AR:Gs:Nb37和T4L- β_2 AR:Gs:Nb35的晶体生长。c)确定pH上下限,指导结晶筛选的制备。出于相同的目的,使用各种浓度的NaCl确定离子强度的效应(数据未显示)。复合物在20、100和500mM中是稳定的,而在2.5M NaCl中解离。d)纳米抗体35(Nb35,红色虚线)结合T4L- β_2 AR:Gs:B1167107三元复合物(蓝色实线),形成R:G:Nb35复合物(红色实线),与处理过的R:G复合物(绿色虚线)相反,R:G:Nb35复合物对GTP γ S处理不敏感(绿色实线)。Nb35和Nb37结合位于Gs异源三聚体上的独立表位,形成R:G:Nb35:Nb37复合物(紫色实线)。e)纳米抗体36(Nb36,红色虚

线)结合R:G复合物(黑色实线),形成R:G:Nb36复合物(红色实线),后者对GTP γ S处理较不敏感(绿色实线)。Nb36和Nb37结合位于Gs异源三聚体上的独立表位,形成R:G:Nb36:Nb37复合物(紫色实线)。f) 纳米抗体37(Nb37,绿线)结合R:G复合物(黑色实线),形成R:G:Nb37复合物(红色实线)。g) R:G:Nb37复合物对GTP γ S处理较不敏感(蓝色实线),与处理过的R:G复合物相反(蓝色虚线)。

[0065] 图4:MNG-3对R:G复合物的稳定效应。a) 在4℃孵育48hr后,对在DDM(黑色)、MNG-3(蓝色)或2种MNG-3类似物(红色和绿色)中纯化的 β_2 AR-365:Gs复合物进行分析凝胶过滤。与DDM相反,R:G复合物在MNG去垢剂中是稳定的。b) 通过3H-二氢阿普洛尔(3H-DHA)饱和结合测定的,稀释未结合配体的纯化 β_2 AR在低于去垢剂的临界微团浓度(CMC)的DDM或MNG-3中的效应。20sec后,稀释维持在低于CMC1000倍的DDM中的 β_2 AR导致丧失3H-DHA结合(黑色数据点)。相反,24hr后,在稀释至低于CMC1000倍的MNG-3中的 β_2 AR仍保持了全部与放射性配体结合的能力。

[0066] 图5:R:G复合物的纯度和均质性。a) 在T4L- β_2 AR:G纯化的各个阶段获得的样品的分析性SDS-PAGE/考马斯蓝染色。在过量的异源三聚体中的B1167107激动剂结合的、去磷酸化的,和去糖基化的受体用于与G-蛋白的功能性部分的最佳偶联效率。通过与固定在M1树脂上的受体的相互作用,实现Gs的功能纯化,同时洗去无功能/未结合的Gs。b) 4次连续预备的凝胶过滤之一的代表性洗脱曲线,红色指出了分级。在所示短划线内合并含有R:G复合物的级分,旋转浓缩,并分析纯度和均质性,通过SDS-PAGE/考马斯蓝(a,倒数第2道至右侧),在c) 凝胶过滤,和通过d) 阴离子交换层析。上图显示了在形成复合物前用 λ PP酶处理过的 β_2 AR-365:Gs复合物的分析性IEC的洗脱曲线,与未脱磷酸化的复合物相比,获得异质的制品(下图)。在存在多种化学品的条件下(图1中的例子),使用b)中的所示虚线外侧的级分中的较不均质的材料进行R:G复合物的分析凝胶过滤实验。

[0067] 图6:制备与Nb35的R:G复合物的纯化过程的流程图。

[0068] 图7:纯化Nb35,和确定R:G:Nb的混合比例。a) 在镍亲和层析纯化纳米抗体35(Nb35)后的预备离子交换层析。将作为小峰和均质大峰洗脱在两组中的纳米抗体(显示为红色)收集、旋转浓缩并且在如(b)所示确定与R:G复合物的适当混合比后,用于晶体学分析。b) 基于蛋白质浓度,将激动剂结合的T4L- β_2 AR:Gs复合物与略微过量的Nb35混合(R:G复合物比Nb35为1-1.2摩尔比),并通过分析性凝胶过滤验证所述混合物。

[0069] 图8:海绵样中间相中的T4L- β_2 AR:Gs:Nb35复合物晶体。

[0070] 图9: β_2 AR:Gs复合物的整体结构。a) 复合物的晶格堆积显示了晶体内受体和G蛋白的交替叠层。在含水层的蛋白之间形成了丰富的接触。b) 不对称单元内容的整体结构显示 β_2 AR(绿色)与激动剂(黄色球)结合,并参与和Gs α (桔色)的广泛的相互作用。Gs α 以及G β (青色)和G γ (紫色)构成了异源三聚体G蛋白Gs。结合Gs的纳米抗体(Nb35,红色)在 α 和 β 亚基之间结合G蛋白。纳米抗体(Nb35)促进结晶,与 β_2 AR的氨基末端融合的T4溶菌酶(紫红色)也促进结晶。c) 省略了结晶帮助的生物学复合物,显示了它在细胞膜中的位置和方向。

[0071] 图10:在结合了BI-167107的T4L- β_2 AR:Gs:Nb35复合物的BI-167107中的Nb35和Gs的相互作用。a) 对Nb35(红色)的CDR1(填充表示)与G β (填充表示,青色)的相互作用的两幅代表性图。b) 对Nb35(红色)的CDR3(填充表示)与Gs α (填充表示,桔色)和G β (填充表示,青色)的相互作用的两幅代表性图。通过与Gs α 和G β 相互作用,Nb35可以降低复合物的构象柔

性。c) 对Nb35 (填充表示, 红色) 与GaS (桔色) 的框架区的相互作用的两幅代表性图。

[0072] 图11: Nb35与相邻复合物的GaS亚基之间的晶体接触。涉及 $-x, y-1/2, -z+1$ 对称的相关复合物 (a) 和 $x, y-1, z$ 对称的相关复合物 (b) 的Nb35 (红色, 填充表示) 和GaS (桔色) 的晶体接触。

[0073] 图12: 比较活性和失活的 β 2AR结构。a) 与失活的结合卡拉洛尔 (carazolol) 的 β 2AR结构 (蓝色; Rosenbaum等人, 2007) 相比, β 2AR:Gs结构 (绿色) 的侧面图和胞质视图。对于TM5和TM6的胞内结构域可见显著的结构改变。如两种结构中的Glu268的 α -碳 (黄色箭头) 测量的, TM5延伸出两条螺旋角, 而TM6则向外移动了**14Å**。b) 与纳米抗体-稳定的活性状态的 β 2AR:Nb80结构 (桔色, Rasmussen等人, 2011) 比较的 β 2AR:Gs。c) 从胞质侧观察到的E/DRY和NPxxY基序中的残基位置, 以及 β 2AR:Gs和 β 2AR:Nb80结构的其他关键残基位置。除了在 β 2AR:Nb80结构中 与纳米抗体相互作用的Arg131外, 所有残基都占据了非常相似的位置。

[0074] 图13: R:G界面中的残基的电子密度图。a) TM3的胞质末端的D/ERY基序。b) 在E/DRY基序的Arg131和C末端GaS的Tyr391之间的堆积相互作用。c) TM7的胞质末端的Np_{xx}Y。d) Thr68和Tyr141与E/DRY基序的Asp130的相互作用。IL2的Phe139包埋在GaS的疏水口袋中。e) GaS的 β 1- α 1环 (P环) 参与核苷酸结合。电子密度图是以1 σ 勾画轮廓的2Fo-Fc图。

[0075] 图14: 受体:G蛋白相互作用。a, b) 通过打开跨膜螺旋5和6, GaS的 α 5-螺旋进入了在受体胞内侧形成的空穴中。a) 在跨膜核心中, 相互作用 主要是非极性的。一个例外涉及 α 5-螺旋的Tyr391倚靠TM3的保守DRY序列的Arg131堆积 (也参见图13)。Arg131还倚靠TM7的保守NPxxY序列的Tyr堆积。b) 由于存在 α 5螺旋, 受体与TM5和TM3形成了极性相互作用的网络。c) 受体残基Thr68和Asp130通过Tyr141与 β 2AR的IL2螺旋相互作用, 摆放螺旋的位置, 使得受体的Phe139进入 (dock) G蛋白表面的疏水口袋中, 从而在结构上关联了受体-G蛋白与 β 2AR的高保守DRY基序相互作用。

[0076] 图15: Gas的构象改变。a) 比较 β 2AR:Gs复合物中的Gas (桔色) 与结合GTP γ S的Gas (灰色) (PDB ID: 1AZT; Sunahara等人, 1997)。GTP γ S显示为球形。相对于在结合GTP γ S的状态下的位置, Gas的螺旋状结构域 (GasAH) 表现出剧烈的移位。b) 旋转Gas的 α 5螺旋, 将其替换为 β 2AR, 干扰 (perturb) β 6- α 5环, 所述环另外形成GTP γ S结合口袋的一部分。c) Gas的 β 1- α 1环 (P环) 和 β 6- α 5环分别与GTP γ S-Gas结构中的GTP γ S的磷酸和嘌呤环相互作用。d) 在不合核苷酸的 β 2AR:Gs结构中, β 1- α 1和 β 6- α 5环被重排。

[0077] 图16: Nb37抑制GTP γ S与Gas结合。孵育Bodipy-GTP γ S (100nM) 与1 μ M纯化的Gas, 在存在递增浓度的Nb37的条件下, 实时测量的荧光增加。

[0078] 图17: Nb35不影响GTP γ S与Gas结合。孵育Bodipy-GTP γ S (100nM) 与1 μ M纯化的Gas, 在存在递增浓度的Nb35的条件下, 实时测量的荧光增加。

[0079] 图18: Nb35抑制GTP γ S与Gs $\alpha\beta\gamma$ 异源三聚体结合。孵育Bodipy-GTP γ S (100nM) 与1 μ M纯化的Gs $\alpha\beta\gamma$ 异源三聚体, 在存在递增浓度的Nb35的条件下, 实时测量的荧光增加。

[0080] 图19: 纯化稳定的AVP:NT4LV2R:Gs复合物。a) 在FLAG标签亲和纯化后, 使用Ni-NTA纯化AVP:NT4LV2R:Gs复合物的示意图。b) 亲和纯化的AVP:NT4LV2R:Gs的SEC色谱图。c) 监控纯化方案的SDS-page。道1: 经过FLAG标签亲和柱的流过物; 道2: 纯化前的AVP、NT4LV2R和Gs的混合物; 道3: 分子标志物; 道4: SEC洗脱的AVP:NT4LV2R:Gs复合物; 道5: FLAG-标签亲和纯化后的Ni-NTA后的AVP:NT4LV2R:Gs复合物。

[0081] 图20:SEC监控的AVP:NT4LV2R:Gs复合物的稳定性。短划线:冰上孵育24hr后,AVP:NT4LV2R:Gs复合物的SEC色谱图。蓝线:冰上孵育48hr后,AVP:NT4LV2R:Gs复合物的SEC色谱图。红线:用10 μ M拮抗剂SR121463孵育复合物后,AVP:NT4LV2R:Gs的SEC色谱图。

[0082] 发明详述

[0083] 定义

[0084] 本发明将就一些特定实施方案并参考附图进行描述,但本发明并不限于此,而仅受权利要求限制。权利要求中的任何提及表示都不应解释为限制范围。本文描述的附图仅是示意性而非限制性的。在附图中,为了举例说明目的,一些元素的尺寸可能被夸大,并非按比例绘制。当术语“包括/包含”在本说明书和权利要求中使用,它不排除其他元素或步骤。在使用不定冠词或定冠词的地方,当提及单数名称例如“一个/一种”、“所述/该”时,除非另有具体说明,否则这涵盖该名词的复数形式。此外,在说明书和权利要求中的术语第一、第二、第三等用于区分相似元素,而不一定用于描述顺序或时间次序。应当理解,如此使用的术语在合适环境下是可互换的,并且本文描述的本发明的实施方案能够以与本文描述或举例说明不同的其他顺序操作。

[0085] 除非本文中另外定义,否则本发明使用的科学和技术术语和词汇都应具有本领域普通技术人员通常理解的含义。一般而言,本文使用的命名法,和本文所述的分子和细胞生物学、遗传学和蛋白质和核酸化学和杂交的技术是本领域普遍已知和常规使用的。除非另外指出,否则一般根据本领域普遍已知的常规方法和本说明书全文引用和讨论的各种常见和更具体的参考文献所述,实施本发明的方法和技术。参见例如,Sambrook等人,Molecular Cloning:A Laboratory Manual,第2版,Cold Spring Harbor Laboratory Press,Cold Spring Harbor,N. Y. (1989);Ausubel等人,Current Protocols in Molecular Biology, Greene Publishing Associates (1992,和2002年增刊)。

[0086] 术语“结合结构域”或“蛋白质结合结构域”一般指任何能够使用特定的分子间相互作用,结合蛋白质或肽的非天然存在的分子,或其部分。多种分子可作为蛋白质结合结构域发挥作用,包括但不限于蛋白质性分子(蛋白质、肽、蛋白质样或含蛋白质的)、核酸分子(核酸、核酸样或含核酸的)和碳水化合物分子(碳水化合物、碳水化合物样、含碳水化合物的)。说明书中还可可见更详细的描述。

[0087] 术语“多肽”、“蛋白质”、“肽”在本文中可互换的使用,指任何长度的氨基酸的聚合形态,可包括编码的和非编码的氨基酸,化学或生物化学修饰的或衍生的氨基酸,和具有相似的肽骨架的多肽。

[0088] 如本文中使用的,术语“蛋白质复合物”或简单的“复合物”指一类两条或多条相关的多肽链。蛋白质复合物中的蛋白质是通过非共价的蛋白质-蛋白质相互作用连接的。“四级结构”是蛋白质复合物中的相关折叠蛋白质的结构排列。可理解,复合物可以是包含2、3、4、5、6或更多个多肽的多聚复合物。复合物还可以额外的包含非蛋白质性分子。

[0089] 如本文中使用的,术语“核酸分子”、“多核苷酸”、“多聚核酸”、“核酸”可互换的使用,指任何长度的核苷酸的聚合形态,不论是脱氧核糖核苷酸或核糖核苷酸,或其类似物。多核苷酸可具有任何三维结构,可实施任何已知或未知的功能。多核苷酸的非限制例子包括基因、基因片段、外显子、内含子、信使RNA (mRNA)、转运RNA、核糖体RNA、核酶、cDNA、重组多核苷酸、分支多核苷酸、质粒、载体、任何序列的分离的DNA、控制区、任何序列的分离的

RNA、核酸探针和引物。核酸分子可以是线性或环状的。

[0090] 如本文中使用的,术语“配体”或“受体配体”意指胞内或胞外特异性结合GPCR的分子。在不是限制性目的的条件下,配体可以是蛋白质、(多)肽、脂类、小分子、蛋白质支架、抗体、抗体片段、核酸、碳水化合物。配体可以是合成的或天然存在的。术语“配体”包括“天然配体”,这是天然GPCR的内源性的、天然的配体。在绝大部分情况下,配体是与细胞表达的GPCR接触(例如结合)时,增加或减少胞内应答的“调节剂”。作为调节剂的配体例子包括激动剂、部分激动剂、逆激动剂和拮抗剂,说明书中还可可见对其更详细的描述。

[0091] 术语蛋白质的“构象”或“构象状态”一般指在任何瞬间,蛋白质可能采用的结构范围。本领域技术人员将认识到,构象或构象状态的決定因素包括反映在蛋白质的氨基酸序列(包括修饰的氨基酸)的蛋白质初级结构,和蛋白质周围的环境。蛋白质的构象或构象状态还涉及结构特征,如蛋白质二级结构(例如, α -螺旋、 β -片层等)、三级结构(例如,多肽链的三维折叠)和四级结构(例如,多肽链与其他蛋白质亚基的相互作用)。对多肽链的翻译后修饰和其他修饰(如配体结合、磷酸化、硫酸化、糖基化或连接疏水基等)可影响蛋白质的构象。此外,环境因素(如周围溶液的pH、盐浓度、离子强度和重量摩尔渗透压浓度、与其他蛋白质和辅助因子的相互作用等)可影响蛋白质构象。可以通过针对活性或与另一种分子的结合的功能测定,或通过物理学方法(如X射线晶体学、NMR或自旋标记等),以及其它方法,确定蛋白质的构象状态。关于蛋白质构象和构象状态的一般讨论,可参考Cantor和Schimmel, *Biophysical Chemistry, Part I: The Conformation of Biological Macromolecules*, W.H. Freeman and Company, 1980, 和 Creighton, *Proteins: Structures and Molecular Properties*, W.H. Freeman and Company, 1993。“特定构象”或“特定构象状态”是蛋白质可采用的一系列构象或构象状态的任何子集。

[0092] 如本文中使用的,“功能性构象”或“功能性构象状态”指这样的事实,即,蛋白质具有不同的构象状态,所述构象状态具有动态的活性范围,特别是从无活性至最大活性的范围。功能性构象状态的例子包括活性构象和失活构象。应明确,“功能性构象状态”意在覆盖GPCR的任何具有任何活性的构象状态,包括无活性;而并非意在覆盖蛋白质的变性状态。本文考虑的一类特殊的功能性构象是“可用于药物研发的构象”,一般指靶蛋白的独特的治疗相关的构象状态。作为示例, β_2 肾上腺素能受体的活性构象对应于该受体用于治疗哮喘的可用于药物研发的构象。因此,可理解,可用于药物研发的能力受到取决于治疗的适应症的特定构象的限制。

[0093] 如本文中使用的,词语“锁定(locking)”或“捕捉(trapping)”或“固定(fixing)”或“冻结(freezing)”当涉及(本文定义的)GPCR的功能性构象状态时,指将GPCR保持或保留在其能够另外呈现的可能的构象子集中,这是由于GPCR:G蛋白复合物与根据本发明的结合结构域的相互作用的效应而产生的。相应的,如本文中使用的,“构象捕捉的”或“构象固定的”或“构象锁定的”或“构象冻结的”蛋白质是由于GPCR:G蛋白复合物与根据本发明的结合结构域的相互作用的效应,保持在其能够另外呈现的可能的构象子集中的蛋白质。在该上下文中,特异性或选择性结合蛋白质的特定构象或构象状态的结合结构域指这样的结合结构域,所述结合结构域与构象或构象状态子集中的蛋白质结合的亲和力比与所述蛋白质可呈现的其他构象或构象状态结合的亲和力更高。本领域技术人员将认识到特异性或选择性结合蛋白质的特定构象或构象状态的结合结构域将使该特定的构象或构象状态稳定。

[0094] 如本文中使用的,在抗体上下文中的术语“互补决定区”或“CDR”指H(重)或L(轻)链的可变区(也分别缩写为 v_H 和 v_L),含有能够特异性结合抗原性靶标的氨基酸序列。这些CDR区负责抗体对特定抗原决定子结构的基础特异性。这类区域也被称为“超变区”。CDR代表了可变区内不连续的氨基酸区段,而不论物种如何,发现重链和轻链可变区内的这些关键氨基酸序列的位置定位在可变链的氨基酸序列中具有相似的位置。所有规则抗体的可变重链和可变轻链都具有3个CDR区,对于各的轻链(L)和重链(H)而言,所述CDR彼此互不连续(被称为L1、L2、L3、H1、H2、H3)。特别的是,免疫球蛋白单可变结构域(如纳米抗体(如本文别处定义的))一般包括单条氨基酸链,其包含4个“框架序列或框架区”或FR(被称为FR1、FR2、FR3、FR4)和3个“互补决定区”或CDR(被称为CDR1、CDR2、CDR3),彼此不连续。CDR序列(以及FR序列)的描述是基于对V-结构域和V样结构域的IMGT特有的编号系统(Lefranc等人,2003)。

[0095] 如本文中使用的,“表位”指多肽的抗原决定子。表位可包括对于表位而言独有的空间构象中的3个氨基酸。一般而言,表位由至少4、5、6、7个这类氨基酸组成,更常见的是由至少8、9、10个这类氨基酸组成。确定氨基酸的空间构象的方法是本领域已知的,包括例如x射线晶体学和多维核磁共振。

[0096] 如本文中使用的,“构象表位”指包含空间构象中的氨基酸的表位,所述空间构象对于多肽的折叠3维构象是独一无二的。一般而言,构象表位由线性序列中不连续的氨基酸组成,这些氨基酸在蛋白质的折叠结构中聚集在一起。然而,构象表位还可以由采用多肽的折叠3维构象所特有(且不存在于变性状态中)的构象的氨基酸线性序列组成。在蛋白质复合物中,构象表位由一条或多条多肽的线性序列中的不连续的氨基酸组成,所述氨基酸在不同的折叠多肽折叠并在特有的四级结构中联结时聚集在一起。相似的,构象表位在本文中还可以由一条或多条多肽的氨基酸线性序列组成,所述氨基酸聚集在一起并采用了四级结构特有的构象。

[0097] 如本文中使用的,术语“特异性”指结合结构域,特别是免疫球蛋白(如抗体)或免疫球蛋白片段(如纳米抗体)相对于不同的抗原,优先结合一种抗原的能力,不必然暗示了高亲和力(如本文中别处定义的)。可以特异性结合特定抗原或抗原决定子(例如,表位)和/或对特定抗原或抗原决定子(例如,表位)具有亲和力的结合结构域,特别是免疫球蛋白(如抗体)或免疫球蛋白片段(如纳米抗体)被称为“抗”或“针对”所述抗原或抗原决定子。如果对两种不同的抗原或抗原决定子都是特异性的,则根据本发明的结合结构域被称为与两种不同的抗原或抗原决定子是“交叉反应性的”。

[0098] 如本文中使用的,术语“亲和力”指结合结构域,特别是免疫球蛋白(如抗体)或免疫球蛋白片段(如纳米抗体)结合抗原的程度,所述程度使抗原和结合结构域的平衡移向存在通过其结合而形成的复合物。因此,例如,当抗原和抗体(片段)以相对相等的浓度组合时,高亲和力的抗体(片段)将结合可利用的抗原,使平衡移向所得到的复合物的高浓度。解离常数(K_d)通常用于描述在蛋白质结合结构域和抗原靶标之间的亲和力。通常,解离常数小于 $10^{-5}M$ 。优选的,解离常数小于 $10^{-6}M$,更优选的小于 $10^{-7}M$ 。最优的,解离常数小于 $10^{-8}M$ 。

[0099] 如本文中使用的,术语“特异性结合”或“特异的结合”一般指结合结构域,特别是免疫球蛋白(如抗体)或免疫球蛋白片段(如纳米抗体)优先结合存在于不同抗原的均质混合物中的特定抗原的能力。在某些实施方案中,特异性结合的相互作用将区分样品中想要

的和不要的抗原,在一些实施方案中大于约10至100倍或更多(例如,大于约1000或10,000倍)。

[0100] 术语“特异性结合”、“选择性结合”、“优先结合”及其语法等价物在本文中可互换的使用。术语“构象特异性”或“构象选择性”在本文中也可互换的使用。

[0101] “缺失”在本文中定义为氨基酸或核苷酸序列的改变,其中相比亲代多肽或核酸的氨基酸序列或核苷酸序列,分别缺少了一个或多个氨基酸或核苷酸残基。在蛋白质或其片段的上下文中,缺失可涉及缺失约2、约5、约10个,多达约20、多达约30或多达约50个或更多个氨基酸。蛋白质或其片段可含有一个以上的缺失。在GPCR的上下文中,缺失可特别是环缺失,或N-和/或C-末端缺失。

[0102] “插入”或“添加”是氨基酸或核苷酸序列的这样的改变,即,相比亲代蛋白质的氨基酸序列或核苷酸序列,分别导致添加了一个或多个氨基酸或核苷酸残基。“插入”一般指在多肽的氨基酸序列中添加一个或多个氨基酸残基,而“添加”可以是插入,或指在N-或C-末端、或两端增加氨基酸残基。在蛋白质或其片段的上下文中,插入或添加通常是约1、约3、约5、约10个,多达约20、多达约30或多达约50个或更多个氨基酸。蛋白质或其片段可含有一个以上的插入。

[0103] 如本文中使用的,“取代”是相比亲代蛋白质或其片段的氨基酸序列或核苷酸序列,一个或多个氨基酸或核苷酸分别被不同的氨基酸或核苷酸替换的结果。可理解,蛋白质或其片段可具有保守性氨基酸取代,其对蛋白质的活性基本没有影响。保守性取代是有意的组合,如gly、ala、val、ile、leu、met;asp、glu;asn、gin;ser、thr;lys、arg;cys、met;和phe、tyr、trp。

[0104] 术语“序列同一性”在本文中用于指在比较窗口中,序列基于核苷酸-核苷酸或氨基酸-氨基酸的相同的程度。因此,“序列同一性百分比”如下计算的:通过在比较窗口比较两条最佳比对的序列,确定两条序列中存在的相同核酸碱基(例如,A、T、C、G、I)或相同氨基酸残基(例如,Ala、Pro、Ser、Thr、Gly、Val、Leu、Ile、Phe、Tyr、Trp、Lys、Arg、His、Asp、Glu、Asn、Gln、Cys和Met)的位置数,获得匹配位置数,用匹配位置数除以比较窗口中的位置总数(即,窗口大小),将结果乘以100,生成序列同一性百分比。可以手工的确定序列同一性百分比,或利用本领域可得到的计算机程序。

[0105] 如本文中使用的,“晶体”或“晶体状结构”指这样的固体材料,其组成原子、分子或离子以在伸展在所有的三维空间中的有序重复的模式排列。从液体或溶解在液体中的物质形成晶体状结构的过程通常被称为“结晶”或“晶体发生”。蛋白质晶体几乎只在溶液中生长。最常见的方法是逐渐降低其组分分子的溶解度。溶液中的晶体生长的特征是2个步骤:微观级微晶(可能仅具有100个分子)的成核,然后是所述微晶的生长,理想的生长为衍射-质量的晶体。

[0106] 如本文中使用的,“X射线晶体学”是确定晶体内的原子排列的方法,其中一束X射线轰击晶体,并向多个特定方向衍射。根据这些衍射束的角度和强度,检晶器(crystallographer)可以产生晶体内的电子密度的三维图像。根据该电子密度,可以确定晶体内的原子的平均位置,及其化学键,其混乱和多种其他信息,如本领域技术人员已知的。

[0107] 如本文中使用的,术语“原子坐标”指分子结构内的原子三维坐标组。在一个实施

方案中,根据生物物理学领域的普遍技术人员普遍已知的方法,使用X射线晶体学获得原子坐标。简要的描述如下,可以通过从晶体衍射X射线,获得X射线衍射模式。衍射数据用于计算包含所述晶体的晶胞的电子密度图;所述图用于建立晶胞内的原子位置(即,原子坐标)。本领域技术人员可理解,通过X射线晶体学确定的结构坐标组合有标准差。在其他实施方案中,可以使用其他的实验生物物理学结构确定方法,获得原子坐标,所述方法可包括电子衍射(也称为电子晶体学)和核磁共振(NMR)方法。在其他实施方案中,可以使用分子建模工具获得原子坐标,所述工具可基于一种或多种的从头开始的蛋白质折叠算法、能量最小化和基于同源性的建模。这些技术是生物物理学和生物信息学领域的普遍技术人员普遍已知的。

[0108] “解析结构”在本文中用于指确定蛋白质的原子排列或原子坐标,通常是通过生物物理学方法进行的,如X射线晶体学。

[0109] 术语“化合物”或“测试化合物”或“候选化合物”或“药物候选化合物”在本文中用于描述在测定中测试的任何天然存在的或合成的分子,所述测定如筛选测定或药物发现测定。同样地,这类化合物包括有机和无机的化合物。化合物包括特征是低分子量的多核苷酸、脂类或激素类似物。其他生物聚合的有机测试化合物包括包含约2至约40个氨基酸的小肽或肽样分子(肽模拟物),和包含约40至约500个氨基酸的较大的多肽,如抗体、抗体片段或抗体缀合物。测试化合物还可以是蛋白质支架。出于高通量的目的,可使用测试化合物文库,如提供充分多样性的组合文库或随机文库。例子包括但不限于天然化合物文库、变构化合物文库、肽文库、抗体片段文库、合成的化合物文库、基于片段的文库、噬菌体展示文库等。更详细的描述可见于说明书的其它部分。

[0110] 如本文中使用的,术语“确定”、“测量”、“评估”、“监控”和“测定”在本文中可互换的使用,包括定量和定性的确定。

[0111] 术语“生物学活性”涉及GPCR时指具有天然存在的GPCR的生物化学功能(例如,结合功能、信号转导功能或由于配体结合的改变构象的能力)的GPCR。

[0112] 如本文中使用的,术语“治疗有效量”、“治疗有效剂量”和“有效量”意指实现理想结果所需的量。

[0113] 如本文中使用的,术语“可药用的”意指不是生物学或其他方式不理想的材料,即所述材料可以与化合物一起施用给个体,而不导致任何不理想的生物学效应或与包含其的药物组合物中的任何其他组分以有害的方式相互作用。

[0114] 发明详述

[0115] 尽管可以激活GPCR的配体极其多样,但它们只与相对少数的胞内蛋白质相互作用,诱导深奥的生理学变化。异源三聚体G蛋白、 β -抑制蛋白和GPCR激酶在其活化状态特异性识别GPCR的能力是普遍已知的,但从结构和功能的观点出发则对其理解有限。因此,令人惊讶且有利的是,鉴别了特异性结合GPCR:G蛋白复合物并能够将复合物稳定或锁定在功能性构象状态(特别是活化构象状态的)结合结构域。此外,结合结构域是用于稳定和捕获处于结合G蛋白状态中的想要的GPCR的一般性工具,这类GPCR一般被认为代表GPCR的活性状态(如本文定义)。

[0116] 因此,本发明的第一个方面涉及针对和/或特异性结合包含GPCR和G蛋白的复合物的结合结构域。

[0117] 本发明的结合结构域可以是任何能够特异性结合包含GPCR和G蛋白的复合物的非天然存在的分子或其部分(如上文定义)。根据优选的实施方案,本文所述的结合结构域是蛋白质支架。术语“蛋白质支架”一般指形成结构的折叠单元,特别是蛋白质或肽结构,其包括用于结合另一种分子,例如蛋白质,的框架(参见例如,Skerra (2000),综述)。结合结构域可源自天然存在的分子,例如源自先天性或适应性免疫系统的组分,或者可以完全是人为设计的。结合结构域可以是基于免疫球蛋白的,或者可以基于蛋白质中存在的结构域,包括但不限于微生物蛋白质、蛋白酶抑制剂、毒素、纤连蛋白、脂笼蛋白、单链反平行卷曲螺旋蛋白或重复基序蛋白。本领域已知的结合结构域的例子包括但不限于:抗体、重链抗体(hcAb)、单结构域抗体(sdAb)、微抗体(minibody)、源自骆驼重链抗体的可变结构域(VHH或纳米抗体)、源自鲨鱼抗体的新抗原受体的可变结构域(VNAR)、alphabody、蛋白A、蛋白G、设计的锚蛋白重复结构域(DARPin)、III型纤连蛋白重复、anticalin、knottin、改造的CH2结构域(纳米抗体)、肽和蛋白质、脂肽(例如,pepducin)、DNA和RNA(参见例如,Gebauer & Skerra, 2009; Skerra, 2000; Starovasnik等人, 1997; Binz等人, 2004; Koide等人, 1998; Dimitrov, 2009; Nygren等人, 2008; W02010066740)。通常,当使用选择方法生成特定类型的结合结构域时,使用包含共同序列或框架序列的组合文库筛选与目标分子(如蛋白质)的结合,所述共同序列或框架序列含有随机的、潜在的相互作用残基。

[0118] 本发明的结合结构域可以针对和/或特异性结合任何选定的GPCR:G蛋白复合物。优选的靶复合物是天然存在的GPCR和G蛋白的复合物,或可选的在例如GPCR和G蛋白的非天然存在的变体(本文另外所述)的情况下,是这样的复合物,其中的GPCR和G蛋白将在恰当的生理条件下关联。本领域技术人员可理解,GPCR和G蛋白之间的结构关系决定了是否可以形成特定的GPCR:G蛋白复合物,这将在下文关于G蛋白家族成员和GPCR家族成员中进一步详述。

[0119] “G蛋白”意指鸟嘌呤核苷酸结合蛋白的家族,该家族参与将化学信号传递到细胞外,并且导致细胞内的改变。G蛋白是配体与GPCR的胞外结构域结合后,胞内信号转导的关键分子组分。它们也被称为“异源三聚体G蛋白”或“大型G蛋白”。G蛋白由3个亚基组成: α 、 β 和 γ ,其分类大部分基于不同 α 亚基的同一性和后续转导事件的性质。G蛋白的其他分类来自cDNA序列同源性分析。G蛋白结合鸟苷二磷酸(GDP)或鸟苷三磷酸(GTP),具有高同源性的鸟嘌呤核苷酸结合结构域和用于与受体和效应子相互作用的不同结构域。不同的G α 蛋白亚类(如G α_s 、G α_i 、G α_q 和G α_{12} 等)通过不同的通路传递信号,所述通路涉及第二信使分子,如cAMP、肌醇三磷酸(IP3)、二酰甘油、胞内Ca²⁺和RhoA GTP酶。为了进一步示例, α 亚基(39-46kDa)含有鸟嘌呤核苷酸结合位点,具有GTP酶活性; β (37kDa)和 γ (8kDa)亚基紧密关联,并作为 $\beta\gamma$ 异二聚体发挥功能。目前描述,有23种类型的 α 亚基(包括一些剪接同种型)、6种 β 亚基和11种 γ 亚基。G蛋白和亚基的类型是下标的:因此,例如,Gs蛋白的 α 亚基(其活化腺苷酸环化酶)是G α_s ;其他G蛋白包括G α_i ,其结构上不同于Gs(不同的 α 亚基类型)并抑制腺苷酸环化酶。表1提供了其他实施例。

[0120] 通常,G蛋白本质上是结合核苷酸的类型。更具体而言,根据特定GPCR的激活状态,G蛋白(至少 α 亚基)结合GTP或GDP。结合GPCR的激动剂促进与GDP结合的G $\alpha\beta\gamma$ 异源三聚体的相互作用,导致G α 上的GDP交换为GTP,和G蛋白功能上解离为G α -GTP和G $\beta\gamma$ 亚基。分离的G α -GTP和G $\beta\gamma$ 亚基可以独立或平行的调节下游的细胞效应子(通道、激酶或其他酶,参见表1)。

G γ 的内在GTP酶活性导致GTP水解为GDP,和G α -GDP与G $\beta\gamma$ 亚基重新关联,和信号传递的终止。因此,G蛋白作为调控分子开关发挥作用,能够通过 α 和 $\beta\gamma$ 亚基效应触发分叉(bifurcating)信号。开关通过受体开启,再在数秒内自身关闭,所述时间足以显著放大信号转导。

[0121] 如本文中使用的,“G蛋白偶联受体”或“GPCR”是享有共同结构基序的多肽,具有7个在22至24个疏水氨基酸之间的区域,所述区域形成7个 α 螺旋,每个螺旋跨越细胞膜。通过编号鉴别每个跨越,即,跨膜-1(TM1)、跨膜-2(TM2)等。还通过在细胞膜外部或“胞外”侧的跨膜-2和跨膜-3、跨膜-4和跨膜-5,以及跨膜-6和跨膜-7之间的氨基酸区域连接跨膜螺旋,所述区域分别被称为“胞外”区1、2和3(EC1、EC2和EC3)。还通过在细胞膜内部或“胞内”侧的跨膜-1和跨膜-2、跨膜-3和跨膜-4、以及跨膜-5和跨膜-6之间的氨基酸区域连接跨膜螺旋,所述区域分别被称为“胞内”区1、2和3(IC1、IC2和IC3)。受体的“羧基”(“C”)末端位于细胞内的胞内空间,受体的“氨基”(“N”)末端位于细胞外的胞外空间。任何上述区域都可以通过分析GPCR的一级氨基酸序列来方便的鉴别。

[0122] GPCR结构和分类是本领域普遍已知的,关于GPCR的其他讨论可见于Probst等人,1992;Marchese等人,1994;**Lagerström&Schiöth**,2008;Rosenbaum等人,2009;和下列书籍:Wiley-Liss出版的Jurgen Wess(编著)Structure-Function Analysis of G Protein-Coupled Receptors(第1版;1999年10月15日);John Wiley&Sons出版的Kevin R.Lynch(编著)Identification and Expression of G Protein-Coupled Receptors(1998年3月);和CRC Press出版的Tatsuya Haga(编著),G Protein-Coupled Receptors(1999年9月24日);和Academic Press出版的Steve Watson(编著)G-Protein Linked Receptor Factsbook(第1版;1994)。可以根据序列同源性,将GPCR分为若干不同家族的类型。虽然所有的GPCR具有相似的7次跨膜 α -螺旋的结构,该受体类型中的不同家族彼此没有表现出任何序列同源性,因此提示其跨膜结构域结构的相似性可定义共同的功能需求。当可获得人基因组的第一份草图时,则可能对GPCR储库进行全面理解。Fredriksson及其同事将802个人GPCR基于系统发生规则分类。这显示绝大部分人GPCR可见于五大主要家族,被称为视紫红质、黏附素(Adhesion)、分泌素(Secretin)、谷氨酸型(Glutamate)、Frizzled/Taste2(Fredriksson等人,2003)。

[0123] 视紫红质家族的成员(对应于A类(Kolakowski,1994)或原分类体系中的1类(Foord等人(2005))仅具有小的胞外环,并且配体与跨膜豁口(cleft)中的残基发生相互作用。这是迄今为止最大的一类(占>90%的GPCR),含有气味物、小分子(如儿茶酚胺和胺)、(神经)肽和糖蛋白激素的受体。视紫红质是该家族的代表物质,是第一个结构被解析的GPCR(Palczewski等人,2000)。第一个结构被解析的与扩散型配体相互作用的受体—— β_2 -AR(Rosenbaum等人,2007)也属于该家族。基于系统发生的分析,B类GPCR或2类(Foord等人,2005)受体最近被细分为2个家族:黏附素和分泌素(Fredriksson等人,2003)。黏附素和分泌素受体的特征是参与配体结合的、相对较长的氨基末端胞外结构域。关于跨膜结构域的方向所知甚少,但该方向很可能与视紫红质的方向迥异。这类GPCR的配体是激素,如胰高血糖素、分泌素、促性腺激素释放激素和甲状旁腺激素。谷氨酸家族受体(C类或3类受体)也具有大的胞外结构域,发挥与“捕蝇草(Venus fly trap)”类似的功能,因为其打开并且将所结合的激动剂关闭在内。家族成员是代谢型谷氨酸、Ca²⁺-敏感型和 γ -氨基丁酸(GABA)-B受

体。

[0124] GPCR包括但不限于血清素嗅觉受体、糖蛋白激素受体、趋化因子受体、腺苷受体、生物胺受体、黑皮质素受体、神经肽受体、趋化性受体、促生长素抑制素受体、阿片样物质受体、褪黑激素受体、降钙素受体、PTH/PTHrP受体、胰高血糖素受体、分泌素受体、黑寡妇蜘蛛毒素受体、代谢型谷氨酸受体、钙受体、GABA-B受体、信息素受体、蛋白酶活化型受体、视紫红质和其他与G蛋白偶联的7次跨膜区段受体。GPCR还包括这类彼此关联作为同源或异源二聚体或作为更高阶的寡聚体的GPCR受体。GPCR的氨基酸序列(及编码其的cDNA的核苷酸序列)是可以方便获得的,例如参考GenBank (<http://www.ncbi.nlm.nih.gov/entrez>)。

[0125] 因此,根据特定的实施方案,本发明提供了针对和/或特异性结合GPCR:G蛋白复合物的结合结构域,其中G蛋白选自Gs、Gi、Go、Gt、Ggust、Gz、Golf、Gq、G12和G13。在一个优选的实施方案中,G蛋白是Gs。在另一个优选的实施方案中,G蛋白是Gi。在仍然另一个优选的实施方案中,G蛋白是Gt,更具体的是转导素(transducin)。与其对应的是,包含在复合物中的GPCR选自Gs偶联受体、Gi偶联受体、Go偶联受体、Gt偶联受体、Ggust偶联受体、Golf偶联受体、Gq偶联受体、G12偶联受体和G13偶联受体。在一个优选的实施方案中,GPCR是Gs偶联受体。在另一个优选的实施方案中,GPCR是Gi偶联受体。在仍然另一个优选的实施方案中,GPCR是Gt偶联受体,更具体的是视紫红质。表1提供了特定的非限制性例子。

[0126] 表1.G蛋白偶联受体和信号传递通路的关系的非限制性例子。

[0127]

G 蛋白 家族	α 亚基	效应子/信号传递通路	用途/受体
Gi 家族			
Gi Go	α_i α_o	抑制腺苷酸环化酶 (cAMP ↓) 关闭 Ca^{2+} 通道 (Ca^{2+} ↓)	乙酰胆碱 M2 & M4 受体 腺苷 A1 & A3 受体 肾上腺素能 α_2A 、 α_2B & α_2C 受体 Apelin 受体 钙敏感型受体 趋化因子 CXCR4 受体 多巴胺 D2、D3、D4 GABAB 受体 谷氨酸 mGluR2、mGluR3、 mGluR4、mGluR6、mGluR7 & mGluR8 受体 组胺 H2 & H3 & H4 受体 褪黑激素 MT1、MT2 & MT3 受体 毒蕈碱 M2 & M4 受体 阿片样物质 δ 、 κ 、 μ & 伤害感受 肽受体 前列腺素 EP1、EP3、FP & TP 受体 血清素 5-HT1 & 5-HT5 受体
Gt	α_t (转导 素)	激活磷酸二酯酶 6 (视觉)	视紫红质
Ggust	α_{gust} (味转 导素)	激活磷酸二酯酶 6 (视觉)	味觉受体
Gz	α_z	抑制腺苷酸环化酶 (cAMP ↓)	未知
Gs 家 族			
Gs	α_s	激活腺苷酸环化酶 (cAMP ↑)	5-HT 受体 5-HT4 和 5-HT7 型 ACTH 受体

[0128]

			腺苷受体 A2a 和 A2b 型 精氨酸血管加压素受体 2 β -肾上腺素能受体 β 1、 β 2 和 β 3 型 降钙素受体 降钙素基因相关肽受体 促肾上腺皮质激素释放激素受体 多巴胺受体 D1 样家族 (D1 和 D5) FSH-受体 胃抑制性多肽受体 胰高血糖素受体 组胺 H2 受体 黄体生成素/绒毛膜促性腺激素受体 黑皮质素受体 甲状旁腺激素受体 1 前列腺素受体 D2 和 I2 型 分泌素受体 促甲状腺素受体
Golf	α olf	激活腺苷酸环化酶 (cAMP $\uparrow$)	嗅觉受体
Gq 家族			
Gq	α q	激活磷脂酶 C (IP ₃ $\uparrow$)	5-HT ₂ 5-羟色胺能受体 α -1 肾上腺素能受体 血管加压素 1 型受体 血管紧张素 II 受体 1 型 降钙素受体 组胺 H1 受体 代谢型谷氨酸受体, I 类 M1、M3 和 M5 毒蕈碱受体
G12/13 家族			
G12	α 12		
G13	α 13	Na ⁺ /H ⁺ 交换 $\uparrow$	
	$\beta\gamma$ 亚基	效应子/信号传递通	

[0129]

		路	
	$\beta\gamma$	开启 K^+ 通道 ($K^+ \uparrow$)	
	$\beta\gamma$	腺苷酸环化酶 (cAMP) $\uparrow$ 或 $\downarrow$	
	$\beta\gamma$	磷脂酶 C (IP_3)	

[0130] 一般而言,对本领域技术人员显而易见的是,本发明的结合结构域至少结合从生物学和/或治疗的观点来看最相关的那些GPCR:G蛋白复合物类型。因此可理解,根据目的和应用,包含在靶复合物中的GPCR和G蛋白可以是天然存在的或非天然存在的(即,人为改变的)。如本文中使用的,术语“天然存在的”意指天然生产的GPCR或G蛋白。特别是,GPCR和G蛋白的野生型多态变体和同种型,以及不同物种的直向同源物是天然存在的蛋白的例子,可见于例如而不限于哺乳动物中,更具体的是在人中,或病毒中,或植物中,或昆虫中等。因此,这类GPCR或G蛋白是自然界中可见的。如本文中使用的,术语“非天然存在的”意指GPCR或G蛋白不是天然存在的。在某些环境下,有利的是包含在复合物中的GPCR和/或G蛋白是非天然存在的蛋白。例如,仅出于示例的目的,为了增加获得通过本发明的结合结构域稳定的GPCR:G蛋白复合物的晶体的可能性,理想的是实施一些蛋白质改造,不影响或仅最低限度的影响配体结合亲和力。或可选的或额外的,为了增加GPCR和/或G蛋白的细胞表达水平,或增加稳定性,还可以考虑在目标GPCR和/或G蛋白中导入某些突变。非天然存在的GPCR的非限制例子包括但不限于相对于其氨基酸或核苷酸序列,通过突变而使其组成型活化的GPCR、具有环缺失的GPCR、具有N-和/或C-末端缺失的GPCR、相对于其氨基酸或核苷酸序列具有取代、插入或添加,或其任意组合的GPCR,或天然存在的GPCR的其他变体。类似的,非天然存在的G蛋白的非限制例子包括但不限于具有N-和/或C-末端缺失的G蛋白、相对于其氨基酸或核苷酸序列具有取代、插入或添加或其任意组合的G蛋白,或天然存在的G蛋白的其他变体。本发明范围内还包括了靶GPCR:G蛋白复合物,所述复合物包含嵌合的或杂交的GPCR,例如具有一种GPCR的N-和/或C-末端和第二种GPCR的环的嵌合GPCR,或包含与某部分(如T4溶菌酶)融合的GPCR(还参见实施例部分)。

[0131] 根据本发明范围内的特定实施方案,包含在GPCR:G蛋白复合物中的非天然存在的GPCR或G蛋白可具有与天然存在的GPCR或G蛋白至少80%相同,至少90%相同,至少95%相同,至少97%相同,或至少99%相同的氨基酸序列。为了进一步示例,以 β_2 -肾上腺素能受体作为本发明范围内的GPCR的特定的非限制性例子,根据上文应该明显的是除了人 β_2 肾上腺素能受体(例如,GenBank登录号NP_000015描述的序列)外,还涵盖了小鼠 β_2 肾上腺素能受体(例如,GenBank登录号NM007420描述的序列)或其他哺乳动物的 β_2 肾上腺素能受体。还考虑了特定物种的 β_2 肾上腺素能受体的野生型多态变体和某些其他的活性变体。例如,“人 β_2 肾上腺素能受体”具有与GenBank登录号NP_000015的天然存在的“人 β_2 肾上腺素能受体”至少80%相同,至少90%相同,至少95%相同,至少97%相同,或至少99%相同的氨基酸序列。类似的,并且以Gas、Gai和Gat作为本发明范围内的G蛋白亚基的特定的非限制性例子,根据上文应该明显的是除了人Gas或Gai或Gat外,还涵盖了小鼠Gas或Gai或Gat蛋白质,或其他哺乳动物的Gas或Gai或Gat蛋白质。还考虑了特定物种的Gas或Gai或Gat的野生型多态变体和某些其他的活性变体。例如,“人Gas”或“人Gai”或“人Gat”分别具有与Genbank登录号

P63092、P63096和P11488的天然存在的“人Gas”或“人Gai”或“人Gat”至少80%相同,至少90%相同,至少95%相同,至少97%相同,或至少99%相同的氨基酸序列。此外,存在许多G蛋白亚基的同种型,包括例如Gs和Gi蛋白的同种型(Gas:GNAS;Ga0:GNA01;Gai:GNAI1或GNAI2或GNAI3;Gβ:GNB1或GNB2或GNB3或GNB4或GNB5或GNB1L或GNB2L;Gγ:NGT1或NGT2或GNG2或GNG3或GNG4或GNG5或GNG7或GNG8或GNG10或GNG11或GNG12或GNG13;根据人基因的HGNC标准化命名;不同生物体的不同同种型的登录号可获得自 www.uniprot.org)。表5提供了G蛋白亚基的同种型的一些特定例子。本领域技术人员可以理解,在不同物种和生物体之间的不同G蛋白亚基的氨基酸序列几乎是100%相同,即使不是100%,也是保守的。值得注意的是,G蛋白的人、牛、大鼠和小鼠β亚基的氨基酸序列的序列比对揭示了这些生物体之间的氨基酸序列是100%保守的。类似的,G蛋白的人、小鼠和牛γ亚基的氨基酸序列是100%相同的。大鼠和小鼠的Gas氨基酸序列也是100%相同的,而人和牛的Gas仅分别有1或2个氨基酸不同。因此,预期针对和/或特异性结合GPCR:G蛋白复合物,特别是结合包含在复合物中的G蛋白,的结合结构域是交叉反应的。还明显的是,包含在靶复合物中的GPCR和G蛋白可以来自相同或不同的物种。优选的,GPCR和/或G蛋白是哺乳动物的蛋白质,或植物蛋白质,或微生物蛋白质,或病毒蛋白质,或昆虫蛋白质。更优选的,GPCR是人蛋白质。

[0132] 还预期本发明的结合结构域一般能够结合这样的GPCR:G蛋白复合物,所述复合物中包含在复合物中包含的特定GPCR和/或G蛋白的所有天然存在的或合成的类似物、变体、突变体、等位基因、部分、片段和同种型;或至少结合复合物中包含的含有一个或多个这样的抗原决定子或表位的特定GPCR和/或G蛋白的那些类似物、变体、突变体、等位基因、部分、片段和同种型,所述抗原决定子或表位与本发明的结合结构域结合包含在复合物中的特定GPCR和/或G蛋白结合的抗原决定子或表位基本相同。

[0133] 可使用各种方法确定在结合结构域和靶GPCR:G蛋白复合物之间的特异性结合(如上文定义),包括例如,酶联免疫吸附测定(ELISA)、表面等离子共振测定、噬菌体展示等,都是本领域的常规实践活动,例如在Sambrook等人,(2001),*Molecular Cloning, A Laboratory Manual*.第3版,Cold Spring Harbor Laboratory Press,Cold Spring Harbor,NY中予以讨论,并在实施例章节中进一步示例。出于该目的,可以理解通常将使用独特的标记或标签,如肽标记、核酸标记、化学物标记、荧光标记或放射频率标签,如本文进一步所述。

[0134] 根据特定的实施方案,相比分别仅与G蛋白和/或仅与GPCR结合,针对GPCR:G蛋白复合物的结合结构域以更高的亲和力结合复合物。在一个实施方案中,针对GPCR:G蛋白复合物的结合结构域特异性结合包含在复合物中的GPCR,而不结合G蛋白。优选的,在另一个实施方案中,针对GPCR:G蛋白复合物的结合结构域特异性结合包含在复合物中的G蛋白,而不结合GPCR。更具体而言,针对GPCR:G蛋白复合物的结合结构域特异性结合包含在Gs偶联受体和Gs蛋白的复合物中的为Gs蛋白的G蛋白。

[0135] 普遍已知的是,GPCR是结构复杂的膜蛋白,其响应天然的和合成的配体,而表现出一系列功能行为。在自然界中,结合了配体的GPCR可以与G蛋白关联成复合物,所述复合物呈现出特定的功能性构象状态,更具体而言是活性构象状态,导致特定的生物学活性。本发明提供了特殊的优势在于本文描述的结合结构域可以稳定与G蛋白复合并与多种天然或合成的配体结合的GPCR的多种这类活性构象。本领域技术人员将认识到,特异性结合配体:

GPCR:G蛋白复合物的结合结构域可稳定包含在复合物中的GPCR的特定构象。在优选的实施方案中,结合结构域能够稳定或者增加GPCR:G蛋白复合物的特定功能性构象状态的稳定性,优选其中GPCR处于活性构象状态。一般而言,所述GPCR的功能性构象状态可以是基础构象状态,或活性构象状态或失活构象状态。优选的,本发明的结合结构域能够使GPCR稳定在其活性构象状态和/或能够在结合GPCR:G蛋白复合物时,将GPCR锁定在活性构象状态,不论是否存在受体配体。

[0136] 在本发明的特别优选的实施方案中,考虑了结合结构域针对和/或特异性结合包含GPCR、G蛋白和受体配体(如本文定义)的复合物。更优选的,结合结构域针对和/或特异性结合由GPCR、G蛋白和受体配体组成的复合物。受体配体可以是小化合物、肽、抗体或抗体片段等,其在结合时触发应答。本文定义的受体配体,或仅仅配体,可以是orthosteric配体(天然的和合成的),其结合受体活性位点,并根据其功效或换言之根据其通过特定通路对受体信号传递的效应而分类。如本文中使用的,“激动剂”指通过结合受体,而增加受体的信号传递活性的配体。完全激动剂能够最大限度的刺激受体;部分激动剂即使在饱和浓度下也不能引发完全活性。部分激动剂还可以通过阻止结合更强力的激动剂,而作为“阻断剂(blocker)”发挥功能。“拮抗剂”指结合受体而不刺激任何活性的配体。“拮抗剂”还被称为“阻断剂”,因其阻止其他配体结合的能力并且因此阻断激动剂诱导的活性。此外,“反向激动剂”指除了阻断激动剂效应外,还使受体的基础活性或组成活性降低至低于未结合配体的受体的拮抗剂。优选的,本发明的结合结构域针对和/或特异性结合包含GPCR、G蛋白和受体配体的复合物,其中受体配体是激动剂。更具体而言,激动剂在orthosteric位点结合受体。

[0137] GPCR的信号传递活性(及其构象行为)也可以受结合另一种被称为变构调节剂的配体的影响。“变构调节剂(regulator)”或“变构调控剂(modulator)”或“效应子分子”结合在GPCR的变构位点(即,物理上不同于蛋白质活性位点的调节剂位点)。与orthosteric配体相反,变构调控剂是非竞争性的,因为它们在不同的位点结合受体,并且即使也结合内源性配体,也可修饰受体功能。因此,变构调控剂不限于像绝大部分药物那样简单的开启或关闭受体。反而更类似二聚体开关发挥作用,提供对活化或去活化的强度的控制,同时通过改变受体与其(内源性)配体的亲和力,允许机体保留其对触发受体活化的天然控制。增强蛋白质活性的变构调节剂在本文中被称为“变构激活剂”或“正变构调控剂”,而减少蛋白质活性的变构调节剂在本文中被称为“变构抑制剂”或“负变构调控剂”。因此,在一个特定的实施方案中,本发明的结合结构域针对和/或特异性结合包含GPCR、G蛋白和受体配体的复合物,其中受体配体是变构调控剂,优选正变构调控剂。更具体而言,正变构调控剂在变构位点结合受体。

[0138] 如所述,GPCR如何调控细胞生理学的经典观点是:配体(如激素、神经递质或感觉刺激素)的结合稳定了受体的活性构象状态,从而允许与异源三聚体G蛋白相互作用。除了与G蛋白相互作用外,结合了激动剂的GPCR关联GPCR激酶(GRK),导致受体磷酸化。GRK使GPCR磷酸化的普遍结果是减少GPCR与G蛋白的相互作用,和增加GPCR与抑制蛋白的相互作用,其在空间上限制了其他的G蛋白信号传递,导致受体脱敏。由于 β -抑制蛋白关闭了G蛋白信号,因此可以同时触发第二组平行的信号级联,如MAPK通路。GPCR还与一般的GPCR-相互作用蛋白家族外的多种蛋白质(G蛋白、GRK、抑制蛋白和其他受体)相互作用。这些GPCR-选

择性伙伴蛋白可以介导GPCR信号传递、通过G蛋白组织GPCR信号传递、指导GPCR运输、锚定GPCR至特别是亚细胞区域,和/或影响GPCR药理学(Ritter和Hall2009)。就此而言,本文使用的配体还可以是具有选择性刺激一小类受体的信号传递活性的能力的“偏向性(biased)配体”,例如选择性激活G蛋白或B-抑制蛋白功能。这类配体也被称为“偏向性配体”、“偏向性激动剂”或“功能上的选择性激动剂”。更特别的是,配体偏向性可以是不完善的偏向性,其特征是配体刺激多种受体活性,所述受体活性对不同的信号具有不同的相对功效(非绝对选择性);或者配体偏向性可以是完善的偏向性,其特征是配体刺激一种受体活性,而对另一种已知的受体活性没有任何刺激作用。因此,在一个特定的实施方案中,本发明的结合结构域针对和/或特异性结合包含GPCR、G蛋白和受体配体的复合物,其中受体配体是偏向性配体。

[0139] 此外,根据本发明的优选的实施方案,特别考虑了本发明的针对和/或特异性结合GPCR:G蛋白复合物的结合结构域,如上文所述,源自先天性或适应性免疫系统。优选的,所述结合结构域源自免疫球蛋白。优选的,根据本发明的结合结构域是抗体或抗体片段。术语“抗体”(Ab)一般指由免疫球蛋白基因或其功能片段编码的多肽,其特异性结合和识别抗原,并且是本领域技术人员已知的。常规的免疫球蛋白(抗体)结构单元包括四聚体。每个四聚体都包含两个相同的多肽链对,每对具有一条“轻链”(约25kDa)和一条“重链”(约50-70kDa)。每条链的N-末端定义了约100至110个或更多个氨基酸的可变区,其主要负责抗原识别。术语可变轻链(VL)和可变重链(VH)分别指这些轻链和重链。术语“抗体”意在包括完整的抗体,包括完整的单链抗体,和抗原结合片段。在一些实施方案中,抗原结合片段可以是结合抗原的抗体片段,包括但不限于Fab、Fab'和F(ab')₂、Fd、单链Fv(scFv)、单链抗体、二硫键连接的Fv(dsFv),和包含VL或VH结构域的片段,或由VL或VH结构域组成的片段,和任意上述的组合,或能够结合靶抗原的免疫球蛋白肽的任何其他功能部分。术语“抗体”还意在包括重链抗体,或其功能片段,如单结构域抗体,更具体而言,免疫球蛋白单可变结构域,如VHH或纳米抗体,如本文别处定义。

[0140] 优选的,本发明的结合结构域是免疫球蛋白单可变结构域。更优选的,所述结合结构域是包含这样的氨基酸序列的免疫球蛋白单可变结构域,所述氨基酸序列包含4个框架区(FR1至FR4)和3个互补决定区(CDR1至CDR3),优选根据下列通式(1):

[0141] FR1-CDR1-FR2-CDR2-FR3-CDR3-FR4 (1)

[0142] ,或其任何合适的片段(通常含有形成至少一个互补决定区的至少一些氨基酸残基)。

[0143] 包含4个FR和3个CDR的结合结构域是本领域技术人员已知的,并作为非限制性例子描述在Wesolowski等人(2009,Med.Microbiol.Immunol.198:157)中。免疫球蛋白单可变结构域的典型但非限制性的例子包括通常源自常规的四链抗体的轻链可变结构域序列(例如,V_L结构域序列)或重链可变结构域序列(例如,V_H结构域序列)。优选的,免疫球蛋白单可变结构域源自骆驼抗体,优选源自不含轻链的重链骆驼抗体,被称为V_HH结构域序列或纳米抗体(如本文另外所述)。

[0144] 如本文中使用的,术语“纳米抗体”(Nb)指源自天然存在的重链抗体的最小抗原结合片段或单可变结构域(V_HH),是本领域技术人员已知的。它们源自骆驼中可见的仅含重链的抗体(Hamers-Casterman等人,1993;Desmyter等人,1996)。在“骆驼”科中,发现了不含轻

链多肽的免疫球蛋白。“骆驼”包括旧世界骆驼(双峰驼(*Camelus bactrianus*)和单峰驼(*Camelus dromedarius*))和新世界骆驼(例如羊驼(*Lama paccos*)、家羊驼(*Lama glama*)、原驼(*Lama guanicoe*)和小羊驼(*Lama vicugna*))。所述单可变结构域重链抗体在本文中被称为纳米抗体或V_HH抗体。NanobodyTM和NanobodiesTM是Ablynx NV(比利时)的商标。Nb的小尺寸和独特的生物物理学特性使其在识别不常见的或隐藏表位时,和结合到蛋白质靶的空穴或活性位点中时,优于常规抗体片段。此外,可以将Nb设计为多特异性和/或多价的抗体,或与报告子分子连接(Conrath等人,2001)。Nb是稳定且僵硬的单结构域蛋白质,可以方便的生产和在胃肠系统中存活。因此,Nb可用于多种用途,包括药物发现和治疗(Saerens等人,2008),也可作为多用途和宝贵的工具,用于蛋白质的纯化、功能研究和结晶化(Conrath等人,2009)。作为结晶化分子伴侣结合天然靶的构象表位发挥作用的特定类型的纳米抗体被称为Xaperone,是本文特别考虑的。Xaperone是结构生物学中的独特工具。XaperoneTM是VIB和VUB(比利时)的商标。使用骆驼抗体片段作为结晶辅助手段的主要优点是Xaperone(1)结合隐蔽表位和使蛋白质锁定为独特的天然构象,(2)增加可溶性蛋白质和增溶性膜蛋白的稳定性,(3)降低可溶性蛋白质和增溶性膜蛋白的构象复杂程度,(4)增加极性表面,使其能够生长衍射晶体,(5)隔离凝聚或多聚化的表面,(6)允许亲和力捕获活性蛋白质。

[0145] 因此,本发明的免疫球蛋白单可变结构域,特别是本发明的纳米抗体,一般包括单条氨基酸链,所述氨基酸链通常包含根据通式(1)的4个“框架序列”或FR和3个“互补决定区”或CDR:

[0146] FR1-CDR1-FR2-CDR2-FR3-CDR3-FR4(1)。

[0147] 术语“互补决定区”或“CDR”指免疫球蛋白单可变结构域中的可变区,含有能够特异性结合抗原靶的氨基酸序列。这些CDR区负责纳米抗体对特定抗原决定子结构的基础特异性。这类区域也被称为“超变区”。免疫球蛋白单可变结构域具有3个CDR区,彼此互不连续(称为CDR1、CDR2、CDR3)。应该明确的是,免疫球蛋白单可变结构域的框架区也可能对与其抗原的结合作出贡献(Desmyter等人2002;Korotkov等人,2009)。本文中描述了根据本发明的这类免疫球蛋白单可变结构域的非限制性例子以及FR和CDR的特定组合(参见表2-3)。CDR序列(以及FR序列)的描述是基于IMGT关于V-结构域和V样结构域的特有编号体系的(Lefranc等人,2003)。可选的,可以通过使用Kabat编号体系描述FR和CDR序列,如用于Riechmann和Muyldermans(2000)文章中的骆驼V_HH结构域的体系。如本领域技术人员已知的,免疫球蛋白单可变结构域,特别是纳米抗体的特征尤其可以是在一条或多条框架序列中,存在一个或多个骆驼科印记(hallmark)残基(根据Kabat编号体系),例如W008/020079的第75页,表A-3中所述(其通过引用整合到本文中)。

[0148] 在优选的实施方案中,本发明提供了含这样的氨基酸序列的免疫球蛋白单可变结构域,所述氨基酸序列选自基本由4个框架区(分别是FR1至FR4)和3个互补决定区(分别是CDR1至CDR3)组成的氨基酸序列,其中所述氨基酸序列的CDR序列与SEQ ID NO:1-6,优选SEQ ID NO:1和/或4的免疫球蛋白单可变结构域的至少一条CDR序列(参见表3)具有至少70%氨基酸同一性,优选至少80%氨基酸同一性,更优选至少90%氨基酸同一性,如至少95%,至少96%,至少97%,至少98%,至少99%或者甚至100%氨基酸同一性。可理解,对于确定免疫球蛋白单可变结构域的一条或多条序列的CDR的氨基酸序列的氨基酸同一性的程度,忽略来自框架区的氨基酸残基。SEQ ID NO:1-6,优选SEQ ID NO:1和/或SEQ ID NO:4中

给出了本发明的免疫球蛋白单可变结构域的一些优选但非限制性的例子(参见表2)。

[0149] 应注意,本发明的免疫球蛋白单可变结构域,特别是纳米抗体就其最广的含义而言不限于特定的生物学来源或特定的制备方法。例如,一般可以如下获得本发明的免疫球蛋白单可变结构域,特别是纳米抗体:(1)分离天然存在的重链抗体的V_HH结构域;(2)表达编码天然存在的V_HH结构域的核苷酸序列;(3)使天然存在的V_HH结构域“人源化”或表达编码此类人源化V_HH结构域的核酸;(4)使来自任何动物物种(并且特别是来自哺乳动物物种,例如来自人)的天然存在的VH结构域“骆驼化”,或表达编码此类骆驼化VH结构域的核酸;;(5)如本领域所述使“结构域抗体”或“Dab”“骆驼化”或表达编码这类骆驼化VH结构域的核酸;(6)使用合成或半合成技术,制备蛋白质、多肽或其他本身已知的氨基酸序列;(7)使用本身已知的核酸合成技术,再表达所获得的核酸,制备编码纳米抗体的核酸;和/或(8)一种或多种上述技术的任意组合。

[0150] 一类优选的免疫球蛋白单可变结构域对应于天然存在的针对GPCR和G蛋白的靶复合物的重链抗体的V_HH结构域。虽然免疫球蛋白单可变结构域的原始文库(**naïve library**)或合成文库可含有针对靶复合物的构象结合物(binder),但本发明的优选实施方案包括用靶复合物免疫骆驼科,使动物的免疫系统暴露在复合物特有的构象表位下。可以用相互作用单体的混合物免疫动物。任选的,可以通过化学交联或添加稳定复合物的协作/变构配体/代谢物(orthosteric激动剂、变构活化剂、Ca⁺⁺、ATP.....),来稳定复合物。还可以通过共价修饰(磷酸化.....)复合物的成员(之一),使复合物稳定。在可选的实施方案中,还可以用GPCR和/或G蛋白单独(因此彼此不复合)免疫骆驼。任选的,还可以通过例如添加稳定GPCR和/或G蛋白的协作/变构配体/代谢物(orthosteric激动剂、变构活化剂、Ca⁺⁺、ATP等),来稳定GPCR和/或G蛋白。

[0151] 因此,一般可以通过用包含GPCR和G蛋白的靶复合物,或用一种或两种其组成成员蛋白,适当的免疫骆驼物种(即,使得提高免疫应答和/或针对靶复合物的重链抗体),从所述骆驼获得合适的生物学样品(如血样,或任何B细胞样品),和使用本身已知的任何合适的技术,从所述样品起始生成针对靶复合物的V_HH序列,来生成或获得这类V_HH序列。这类技术对本领域技术人员是显而易见的。可选的,可以从骆驼V_HH序列的原始文库中获得这类天然存在的V_HH结构域,例如通过使用一种或多种本身已知的筛选技术,用靶复合物或其至少一个部分、片段、抗原决定子或表位筛选这类文库。这类文库和技术描述在例如W09937681、W00190190、W003025020和W003035694中。可选的,可以使用源自原始V_HH文库的改良的合成或半合成文库,如通过技术例如随机诱变和/或CDR改组从原始V_HH文库中获得的V_HH文库,如W00043507中所述。用于获得针对靶的V_HH序列的另一种技术涉及使用本身已知的任何合适的技术,适当的免疫能够表达重链抗体的转基因哺乳动物(即,从而提高针对靶的免疫应答和/或重链抗体),从所述转基因哺乳动物获得合适的生物学样品(如血样,或任何B细胞样品),然后从所述样品开始生成针对靶的V_HH序列。例如,出于该目的,可以使用表达重链抗体的小鼠,和W002085945和W004049794中描述的其他方法和技术。

[0152] 本发明的特别优选的免疫球蛋白单可变结构域类型,特别是本发明的纳米抗体的特别优选的类型,包括具有与天然存在的V_HH结构域的氨基酸序列对应的氨基酸序列的免疫球蛋白单可变结构域,但所述氨基酸序列已经过“人源化”,即,用存在于人的常规4链抗体的VH结构域的相应位置上的一个或多个氨基酸残基取代了所述天然存在的V_HH序列(特

别是框架序列中)的氨基酸序列中的一个或多个氨基酸残基。可以以本身已知的方式实施,这对本领域技术人员是显而易见的,例如基于本文和现有技术关于本文所述人源化作用的进一步描述。此外,应该注意本发明的这类人源化免疫球蛋白单可变结构域是可以通过任何本身已知的合适方式获得的(即,如上文(1)-(8)点所示),因此,不严格限于使用包含天然存在的V_HH结构域的多肽作为起始材料获得的多肽。人源化免疫球蛋白单可变结构域,特别是纳米抗体可具有若干优点,如比相应的天然存在的V_HH结构域降低的免疫原性。这类人源化作用一般涉及用存在于人VH结构域(如人VH3结构域)的相同位置上的氨基酸残基取代天然存在的V_HH序列中的一个或多个氨基酸残基。应该选择这样的人源化取代,使获得的人源化免疫球蛋白单可变结构域仍然保留本文定义的免疫球蛋白单可变结构域的有利特性。本领域技术人员能够选择人源化取代或人源化取代的合适组合,所述取代或取代组合使一方面人源化取代提供的有利特性与另一方面天然存在的V_HH结构域的有利特性之间的平衡最优化,或者实现了理想或合适的平衡。

[0153] 本发明的免疫球蛋白单可变结构域,特别是本发明的纳米抗体的另一个特别优选的类型包括具有这样的氨基酸序列的免疫球蛋白单可变结构域,所述氨基酸序列与天然存在的VH结构域的氨基酸序列对应,但已经被“骆驼化”,即,用存在于重链抗体的V_HH结构域的相应位置上的一个或多个氨基酸残基取代了常规4链抗体的天然存在的VH结构域的氨基酸序列中的一个或多个氨基酸残基。这类“骆驼化”取代优选插入到形成VH-VL界面的氨基酸位置和/或存在于VH-VL界面中的氨基酸位置,和/或位于所谓的骆驼科的印记残基的氨基酸位置,如本文定义的(参见例如W09404678)。优选的,用作生成或设计骆驼化纳米抗体的起始材料或起点的VH序列优选是哺乳动物的VH序列,更优选是人的VH序列,如VH3序列。然而,应注意本发明的这类骆驼化免疫球蛋白单可变结构域是可以通过任何本身已知的合适方式获得的(即,如上文(1)-(8)点所示),因此,不严格限于使用包含天然存在的VH结构域的多肽作为起始材料获得的多肽。

[0154] 例如,可以如下实施“人源化”和“骆驼化”:分别提供编码天然存在的V_HH结构域或VH结构域的核苷酸序列,然后以本身已知的方式改变所述核苷酸序列中的一个或多个密码子,使新的核苷酸序列分别编码本发明的“人源化的”或“骆驼化的”免疫球蛋白单可变结构域。然后,可以以本身已知的方式表达该核酸,从而提供理想的本发明的免疫球蛋白单可变结构域。可选的,分别基于天然存在的V_HH结构域或VH结构域的氨基酸序列,可以设计本发明的理想的人源化或骆驼化免疫球蛋白单可变结构域的氨基酸序列,然后使用本身已知的肽合成技术从头合成它们。还可分别基于天然存在的V_HH结构域或VH结构域的氨基酸序列或核苷酸序列来设计编码本发明的理想的人源化或骆驼化免疫球蛋白单可变结构域的核苷酸序列,然后使用本身已知的核酸合成技术从头合成所述核苷酸序列,之后以本身已知的方式表达因此获得的核酸,从而提供理想的本发明的免疫球蛋白单可变结构域。用于从天然存在的VH序列,或优选从V_HH序列获得本发明的免疫球蛋白单可变结构域和/或其编码核酸的其他合适的方法和技术对本领域技术人员是显而易见的,可以例如包括以合适的方式组合一个或多个天然存在的VH序列(如一个或多个FR序列和/或CDR序列)的一个或多个部分,一个或多个天然存在的V_HH序列(如一个或多个FR序列和/或CDR序列)的一个或多个部分,和/或一个或多个合成或半合成的序列,从而提供理想的本发明的纳米抗体或其编码核苷酸序列或核酸。

[0155] 还位于本发明范围内的是如本文定义的本发明的免疫球蛋白单可变结构域,特别是纳米抗体,天然的或合成的类似物、突变体、变体、等位基因、部分或片段(在本文中统称“变体”),特别是SEQ ID NO:1-6的免疫球蛋白单可变结构域的变体(参见表2-3)。因此,根据本发明的一个实施方案,术语“本发明的免疫球蛋白单可变结构域”或“本发明的纳米抗体”就其最广的含义还覆盖了上述变体。一般而言,相比本文定义的本发明的免疫球蛋白单可变结构域,这类变体中已取代、缺失和/或添加了一个或多个氨基酸残基。这类取代、插入或缺失可在一个或多个FR和/或一个或多个CDR中进行,特别是SEQ ID NO:1-6的免疫球蛋白单可变结构域的FR和CDR的变体(参见表2-3)。如本文中使用的,变体是这样的序列,其中每个或任何一个框架区、每个或任何一个互补决定区都与参照序列的相应区域表现出至少80%同一性,优选至少85%同一性,更优选90%同一性,甚至更优选95%同一性,或者甚至更优选99%同一性(即,FR1_变体与FR1_参照物、CDR1_变体与CDR1_参照物、FR2_变体与FR2_参照物、CDR2_变体与CDR2_参照物、FR3_变体与FR3_参照物、CDR3_变体与CDR3_参照物、FR4_变体与FR4_参照物),如可以利用算法,如PILEUP和BLAST电子测量的(50,51)。用于实施BLAST分析的软件是可以通过美国国立生物技术信息中心公开获得的([http://www/ncbi.nlm.nih.gov/](http://www.ncbi.nlm.nih.gov/))。可以理解,为了确定免疫球蛋白单可变结构域的一条或多条序列的CDR的氨基酸序列的氨基酸同一性程度,忽略了形成框架区的氨基酸残基。类似的,为了确定本发明的免疫球蛋白单可变结构域的一条或多条序列的FR的氨基酸序列的氨基酸同一性程度,忽略了形成互补区的氨基酸残基。免疫球蛋白单可变结构域的这类变体可以是特别有利的,因为其具有改善的效价/亲和力。

[0156] 作为非限制性例子,取代可以是例如保守性取代(如本文所述)和/或氨基酸残基被天然存在于另一种V_HH结构域的相同位置上的另一种氨基酸残基取代。因此,本发明的范围内包括了任何一个或多个这样的取代、缺失或插入,或其任意的组合,所述取代、缺失或插入或其组合改善了本发明的免疫球蛋白单可变结构域的特性,或至少不过分减弱本发明的纳米抗体的所需特性或所需特性的平衡或组合(即所述免疫球蛋白单可变结构域不再适用于其预期的用途)。本领域技术人员一般能够基于本文公开的内容,任选的在有限程度的常规实验后,确定和选择合适的取代、缺失或插入,或其合适的组合,所述常规实验可涉及例如导入有限数量的可能取代和确定其对所获得的免疫球蛋白单可变结构域的特性的影响。

[0157] 根据特别优选的实施方案,本发明的免疫球蛋白单可变结构域,特别是纳米抗体的变体可以在1、2或3个CDR中具有1、2或3个氨基酸的取代、缺失或插入,更具体的是在(i) CDR1或CDR2或CDR3; (ii) CDR1和CDR2,或CDR1和CDR3,或CDR2和CDR3; (iii) CDR1和CDR2和CDR3中具有1、2或3个氨基酸的取代、缺失或插入,所述CDR的氨基酸序列列于表3中。更优选的,本发明的免疫球蛋白单可变结构域,特别是纳米抗体的变体可以在1、2或3个CDR中具有1、2或3个氨基酸的保守性取代(如本文定义),更具体的是在表3列举的CDR的氨基酸序列的(i) CDR1或CDR2或CDR3; (ii) CDR1和CDR2,或CDR1和CDR3,或CDR2和CDR3; (iii) CDR1和CDR2和CDR3中具有1、2或3个氨基酸的保守性取代。

[0158] 根据特定的实施方案,发明提供了包含这样的氨基酸序列的免疫球蛋白单可变结构域,所述氨基酸序列包含4个框架区(FR1至FR4)和3个互补决定区(CDR1至CDR3),根据下列通式(1):

[0159] FR1-CDR1-FR2-CDR2-FR3-CDR3-FR4 (1) ;

[0160] 并且其中CDR1选自:

[0161] a) SEQ ID NO:13-18,

[0162] b) 与SEQ ID NO:13-18具有至少80%氨基酸同一性的多肽,

[0163] c) 与SEQ ID NO:13-18具有3、2或1个氨基酸差异的多肽,

[0164] 并且其中CDR2选自:

[0165] a) SEQ ID NO:25-30,

[0166] b) 与SEQ ID NO:25-30具有至少80%氨基酸同一性的多肽,

[0167] c) 与SEQ ID NO:25-30具有3、2或1个氨基酸差异的多肽,

[0168] 并且其中CDR3选自:

[0169] a) SEQ ID NO:37-42,

[0170] b) 与SEQ ID NO:37-42具有至少80%氨基酸同一性的多肽,

[0171] c) 与SEQ ID NO:37-42具有3、2或1个氨基酸差异的多肽。

[0172] 在特别优选的实施方案中,本发明提供了包含这样的氨基酸序列的免疫球蛋白单可变结构域,所述氨基酸序列包含4个框架区 (FR1至FR4) 和3个互补决定区 (CDR1至CDR3), 根据下列通式 (1) :

[0173] FR1-CDR1-FR2-CDR2-FR3-CDR3-FR4 (1) ;

[0174] 其中CDR1是SEQ ID NO:13;其中CDR2是SEQ ID NO:25;且其中CDR3是SEQ ID NO:37。

[0175] 此外,根据用于表达本发明的结合结构域,特别是免疫球蛋白单可变结构域的宿主生物,可以这样设计缺失和/或取代,使得去除了一个或多个用于翻译后修饰的位点(如一个或多个糖基化位点),这是本领域技术人员能力范围内的。可选的,可以这样设计取代或插入,使得导入一个或多个用于连接官能团,例如允许位点特异性聚乙二醇化,的位点(如本文所述)。

[0176] 修饰的例子、以及本发明的结合结构域,特别是免疫球蛋白单可变结构域中可以被修饰的氨基酸残基的例子(即,在蛋白质骨架上,优选在侧链上)、可用于导入这类修饰的方法和技术,和这类修饰的潜在用途和优点,对本领域技术人员是显而易见的。例如,这类修饰可涉及向本发明的结合结构域,特别是免疫球蛋白单可变结构域,之中或其上导入(例如,通过共价连接或另一种合适的方式)一个或多个官能团、残基或部分,特别是赋予本发明的结合结构域一种或多种想要的特性或功能的一个或多个官能团、残基或部分。这类官能团和用于导入它们的技术的例子对本领域技术人员是显而易见的,一般可包括本领域中提及的所有官能团和技术,以及本身已知用于修饰药物蛋白质,特别是用于修饰抗体或抗体片段(包括ScFv和单结构域抗体)的官能团和技术,例如参考Remington的Pharmaceutical Sciences,第16版,Mack Publishing Co.,Easton,PA(1980)。这类官能团可以例如与本发明的结合结构域,特别是免疫球蛋白单可变结构域直接(例如共价的)连接,或任选的通过合适的接头或间隔子连接,这对于本领域技术人员仍然是显而易见的。一种最常用的用于增加药物蛋白质的半衰期和/或降低其免疫原性的技术包括连接合适的可药用的聚合物,如聚(乙二醇)(PEG)或其衍生物(如甲氧基聚(乙二醇)或mPEG)。一般而言,可以属于任何合适的聚乙二醇化形式,如本领域中用于抗体和抗体片段的聚乙二醇化(包

括但不限于(单)结构域抗体和ScFv);参考例如Chapman, Nat. Biotechnol., 54, 531-545 (2002); Veronese和Harris, Adv. Drug. Deliv. Rev. 54, 453-456 (2003); Harris和Chess, Nat. Rev. Drug. Discov., 2, (2003)和W004060965。用于蛋白质聚乙二醇化的各种试剂也是可商购,例如从Nektar Therapeutics, USA。优选使用定点聚乙二醇化,特别是通过半胱氨酸残基(参见例如Yang等人, Protein Engineering, 16, 10, 761-770 (2003))。例如,出于该目的,可以将PEG连接到天然存在于本发明的免疫球蛋白单可变结构域中的半胱氨酸残基上,可以修饰本发明的免疫球蛋白单可变结构域从而适当的导入一个或多个用于连接PEG的半胱氨酸残基,或可以融合包含一个或多个用于连接PEG的半胱氨酸残基的氨基酸序列至本发明的纳米抗体的N-和/或C-末端,上述所有都使用技术人员本身已知的蛋白质改造技术。优选的,对于本发明的免疫球蛋白单可变结构域,可以使用分子量大于5000,如大于10,000且小于200,000,如小于100,000的PEG;例如分子量范围是20,000-80,000。另外,通常较不优选的修饰包括N-连接或O-连接糖基化,通常作为共翻译和/或翻译后修饰的一部分,这取决于用于表达本发明的结合结构域,特别是免疫球蛋白单可变结构域的宿主细胞。用于增加结合结构域的半衰期的另一个技术包括改造为双功能结合结构域(例如,针对靶GPCR:G蛋白 复合物的一个免疫球蛋白单可变结构域和针对血清蛋白质如白蛋白的一个),或改造成结合结构域,特别是免疫球蛋白单可变结构域与肽(例如,针对血清蛋白质(如白蛋白)的肽)的融合物。

[0177] 另一种修饰可包括导入一个或多个可检测的标记或其他生成信号的基团或部分,取决于被标记的结合结构域的预期用途。合适的标记和用于连接、使用 and 检测标记的技术是本领域技术人员显而易见的,例如包括但不限于荧光标记(如荧光素、异硫氰酸盐、罗丹明、藻红蛋白、藻蓝蛋白、别藻蓝蛋白、邻苯二醛和荧光胺,和荧光金属,如Eu,或其他镧系元素系列的金属)、发磷光标记、化学发光标记或生物发光标记(如鲁米诺(luminal)、异氨基苯二酰肼、theromatic acridinium ester、咪唑、吡啶盐、草酸酯、二氧杂环丁烷或GFP,及其类似物)、放射性同位素、金属、金属螯合物或金属阳离子,或其他特别适合在体内、体外或原位诊断和成像使用的金属或金属阳离子,以及生色团和酶(如苹果酸脱氢酶、葡萄球菌核酸酶、 Δ -V-类固醇异构酶、酵母醇脱氢酶、 α -甘油磷酸脱氢酶、磷酸丙糖异构酶、生物素-亲和素过氧化物酶、辣根过氧化物酶、碱性磷酸酶、天冬酰胺酶、葡萄糖氧化酶、 β -半乳糖苷酶、核糖核酸酶、脲酶、过氧化氢酶、葡萄糖-VI-磷酸脱氢酶、葡糖淀粉酶和乙酰胆碱酯酶)。其他合适的标记对本领域技术人员是显而易见的,例如包括可以使用NMR或ESR光谱检测的部分。这类标记的本发明的结合结构域可用于例如体外、体内或原位测定(包括本身已知的免疫测定,如ELISA、RIA、EIA和其他“夹心测定”等)以及体内诊断和成像目的,取决于特定标记的选择。对本领域技术人员显而易见的是,另一种修饰可涉及导入螯合基团,例如螯合一种或多种上述金属或金属阳离子。例如,合适的螯合基包括但不限于二乙烯三胺五乙酸(DTPA)或乙二胺四乙酸(EDTA)。另一种修饰可包括导入作为特定结合对的一部分的官能团,如生物素-(链霉)抗生物素结合对。这类官能团可用于连接本发明的结合结构域与另一种与结合对的另一半结合的蛋白质、多肽或化学化合物,即,通过形成结合对。例如,本发明的免疫球蛋白单可变结构域可以与生物素缀合,并与另一种与抗生物素蛋白 或链霉抗生物素缀合的蛋白质、多肽、化合物或载体连接。例如,这类缀合的免疫球蛋白单可变结构域可用作报告子,例如在将可检测的信号生成剂与抗生物素蛋白或链霉抗生物素缀合的诊断

系统中。这类结合对还可以用于例如将本发明的免疫球蛋白单可变结构域与载体结合,包括适合制药目的的载体。一个非限制性例子是Cao和Suresh, *Journal of Drug Targetting*, 8, 4, 257 (2000) 描述的脂质体制剂。这类结合对还可以用于连接治疗活性剂与本发明的结合结构域。

[0178] 本发明范围内还涵盖了处于“多价”形态的本发明的结合结构域,特别是本发明的免疫球蛋白单可变结构域,所述结构域是通过化学的或重组DNA技术将两个或多个单价免疫球蛋白单可变结构域键合在一起而形成的。多价构建体的非限制性例子包括“二价”构建体、“三价”构建体、“四价”构建体等。包含在多价构建体中的免疫球蛋白单可变结构域可以是相同或不同的。在另一个特定的实施方案中,本发明的免疫球蛋白单可变结构域处于“多特异性”形态,是通过将两个或多个免疫球蛋白单可变结构域键合在一起而形成的,其中至少一个具有不同的特异性。多特异性构建体的非限制性例子包括“双特异性”构建体、“三特异性”构建体、“四特异性”构建体等。为了进一步示例,本发明的任何多价或多特异性的(如本文定义)免疫球蛋白单可变结构域都可合适的针对相同抗原上的两种或多种不同的表位,例如针对包含在GPCR:G蛋白复合物中的G蛋白的两个或多个不同的部分;或者可以针对两种或多种不同的抗原,例如针对GPCR的表位和G蛋白的表位。特别的是,本发明的单价免疫球蛋白单可变结构域是这样的,即其以小于500nM,优选小于200nM,更优选小于10nM,如小于500pM的亲和力结合靶GPCR:G蛋白复合物(如本文所述)的。本发明的多价或多特异性的免疫球蛋白单可变结构域还可对想要的GPCR:G蛋白复合物具有(或被改造和/或被选择为)增加的亲合力和/或改善的选择性,和/或通过使用这类多价或多特异性的免疫球蛋白单可变结构域可获得的任何其他理想的特性或理想特性的组合。

[0179] 此外,本发明的结合结构域,特别是本发明的免疫球蛋白单可变结构域一般可针对或特异性结合任何构象表位,所述表位由包含GPCR和G蛋白的复合物所呈现,或有所述复合物所接触,或是所述复合物的一部分。特异性结合“三维”表位或“构象”表位的结合结构域是特异性结合折叠的蛋白质或蛋白质复合物的三级或四级结构的结合结构域。这类结合结构域以明显降低的(即,降低至至少2、5、10、50或100分之一)亲和力结合线性(即,未折叠的、变性的)多肽链。此类结合结构域所结合的结构通常含有在蛋白质(复合物)的线性序列中不连续的氨基酸。换言之,这类结合结构域与多肽的结合依赖于折叠成特定三维构象的多肽。应该明确的是,本发明的结合结构域选择性识别的构象表位可以是GPCR特异性表位,或G蛋白特异性表位,或者GPCR:G蛋白复合物-特异性表位,其仅在关联组分蛋白时并且因此通过组合所述GPCR和所述G蛋白的氨基酸残基而形成。在一个实施方案中,本发明的结合结构域,特别是免疫球蛋白单可变结构域,特异性结合任何理想G蛋白或其部分的任何构象表位。在另一个实施方案中,所述构象表位可以是任何所需GPCR的胞内或胞外区的一部分,或膜内区,或结构域或环状结构。明显的是,在非关联形态的GPCR和/或G蛋白中,一些这类构象表位是可接触的,而其他构象表位则仅在形成复合物时可接触。根据一个特定的实施方案,本发明的结合结构域,特别是免疫球蛋白单可变结构域,特异性结合位于G蛋白的 α 和 β 亚基之间的界面上的构象表位,如本文别处的非限制性例子所述(参见实施例章节)。

[0180] 根据其他特定的实施方案,结合结构域结合这样的GPCR:G蛋白复合物,其中G蛋白处于其不合核苷酸的形态。根据其他特定的实施方案,本发明的结合结构域,特别是免疫球蛋白单可变结构域,在存在核苷酸(特别是鸟嘌呤核苷酸(如GDP或GTP),或其类似物,如不

可水解的GTP类似物(如GTP γ S)),或与氟化铝或氟化铍组合的GDP,或者在最小核苷酸类似物(如焦磷酸盐或forcarnet)的条件下,抑制或阻止GPCR:G蛋白复合物的解离。在缺少本发明的结合结构域的条件下,GPCR:G蛋白复合物通常在存在这些核苷酸的条件下解离。

[0181] 根据一个特定的方面,本发明还提供了针对或特异性结合G蛋白(即,仅G蛋白,不与GPCR复合)的结合结构域。在特定的实施方案中,本文所述的结合结构域针对和/或特异性结合Gs蛋白。根据该方面特定的实施方案,本发明的结合结构域阻止或抑制核苷酸(特别是鸟嘌呤核苷酸或类似物(如上文所述))与G蛋白结合。或者本发明的结合结构域还能够替代G蛋白中的鸟嘌呤核苷酸或类似物。实施例章节中提供了测定的非限制性例子,用于确定对鸟嘌呤核苷酸与蛋白质的结合的抑制程度或对蛋白质的鸟嘌呤核苷酸的替代程度,例如实施例3和8。此外,可以理解所有涉及本发明的结合结构域的特定实施方案(如上文所述)也适用于本发明的这一特定方面。

[0182] GPCR的功能多样性与导致构象谱的这些蛋白质的柔性内在关联。构象能量景观(energy landscape)与这样的因素内在相关,如存在结合的配体(效应子分子、激动剂、拮抗剂、逆向激动剂.....)、脂类环境或结合相互作用的蛋白质。因此,在一个实施方案中,本发明的结合结构域,特别是免疫球蛋白单可变结构域增加了包含GPCR和G蛋白的复合物在结合所述结合结构域时的稳定性。在其他实施方案中,本发明的结合结构域结合任何下述构象表位,诱导形成GPCR的功能性构象状态,特别是所述GPCR的活性构象状态,所述构象表位由包含GPCR:G蛋白复合物呈现,或可接近所述复合物,或者是所述复合物的一部分。更具体而言,本发明的结合结构域能够通过增加G蛋白对受体的亲和力,稳定包含在GPCR:G蛋白复合物中的GPCR的活性状态。同样的,本发明的结合结构域能够稳定结合了激动剂的GPCR:G蛋白复合物和/或增强激动剂对GPCR:G蛋白复合物的亲和力。优选的,通过 K_d 的减少所测量的,结合结构域能够增加G蛋白对GPCR的亲和力,和/或激动剂对GPCR:G蛋白复合物的亲和力至少2倍、至少5倍,更优选至少10倍。可选的,在任何比较存在复合物稳定的结合结构域的条件与缺少该结合结构域的条件下的受体和G蛋白的相互作用强度的测定设置中,或者任何其他本领域技术人员已知的亲和力或效价测量中,结合结构域能够诱导 EC_{50} 或 IC_{50} 改变至少2倍、至少5倍,更优选至少10倍。

[0183] 如本文中使用的,术语“功能性构象状态”指这样的事实,即,蛋白质,特别是膜蛋白(如GPCR)具有多种具动态活性范围的不同构象状态,特别是从无活性至最大活性(参见Kobilka和Deupi,2007的综述)。应该明确的是“功能性构象状态”并非意在覆盖蛋白质的变性状态。例如,基础构象状态可定义为在缺少配体时,受体的低能量状态。蛋白质经历转变成另一种构象状态的可能性是两种状态之间的能量差异和两种状态之间的能量壁垒高度的函数。在受体蛋白质(如GPCR)的情况下,配体结合的能量可用于改变两种状态之间的能量壁垒,或用于改变两种状态之间的相对能量水平,或同时。能量壁垒的改变影响两种状态之间的转换速率,而能量水平的改变影响受体在两种状态中的平衡分布。结合激动剂或部分激动剂可降低能量壁垒和/或降低更高活性构象状态相对于失活构象状态的能量。反向激动剂可增加能量壁垒和/或降低失活构象状态相对于活性构象状态的能量。偶联受体与其G蛋白可进一步改变能量景观。在配体结合测定中观察到的 $\beta 2AR$ 和Gs的协作相互作用形成了GPCR活化的三元复合物模型的基础(Delean等人,1980)。在由激动剂、受体和G蛋白组成的三元复合物中,受体对激动剂的亲和力增强,G蛋白对鸟嘌呤核苷酸的特异性改变,

相比GDP更偏爱GTP。

[0184] 应该注意到,整合膜蛋白(包括GPCR)的活性还受膜中的围绕它们的脂分子结构的影响。膜蛋白不是僵硬的实体,而是变形以确保与周围脂双层良好的疏水性匹配。一个重要的参数是脂双层的疏水厚度,由脂类的脂肪酰链长度定义。此外,在定义位于脂类头部基团区域中的膜蛋白的那些部分的结构时,脂类头部基团区域的结构也可能是重要的。在其他脂类中,棕榈酰基化和GPCR结合胆固醇在单体受体中也发挥了结构性作用,并对受体寡聚物的形成/稳定作出贡献(Lee2004;Chini和Parenti2009)。

[0185] 本发明的另一个方面涉及包含本发明的结合结构域的复合物。更具体而言,提供了包含本发明的结合结构域、GPCR、G蛋白和任选的受体配体的复合物。作为非限制性例子,可通过凝胶过滤纯化稳定的复合物,如 对含有纳米抗体、GPCR、G蛋白和受体配体的四级复合物的所进行的操作那样进行的(参见实施例章节)。在特定的实施方案中,复合物可以是结晶的。相应的,还提供了复合物的晶体,以及制备所述晶体的方法,更详细的描述在本文别处。

[0186] 在另一个方面,编码本发明的任何结合结构域,特别是免疫球蛋白单可变结构域,的氨基酸序列的核酸序列也是本发明的一部分,表4中提供了非限制性的例子。根据优选的实施方案,本发明涉及本发明的结合结构域,特别是免疫球蛋白单可变结构域的核酸序列,其中序列与SEQ ID NO:49-54的结合结构域的至少一条核酸序列的序列具有大于80%,优选大于90%,更优选大于95%,如99%或更高的序列同一性(如本文定义)(参见表4)。为了计算百分比序列同一性,应该忽略标签(例如,His标签或EPEA标签)的核酸序列。此外,核酸序列中可包含本文所述的核酸序列。

[0187] 此外,本发明还考虑了包含编码本发明的任何结合结构域,特别是免疫球蛋白单可变结构域的核酸序列的表达载体,以及表达这类表达载体的宿主细胞。合适的表达系统包括细菌或酵母的组成型和诱导型的表达系统、病毒表达系统(如杆状病毒、塞姆利基森林病毒和慢病毒),或在昆虫或哺乳动物细胞中瞬时转染。可以根据本领域技术人员已知的技术,克隆、表达和/或纯化本发明的结合结构域,特别是免疫球蛋白单可变结构域。

[0188] 因此,本发明涵盖了表达本发明的结合结构域,特别是免疫球蛋白单可变结构域的细胞或细胞培养物,所述结构域针对和/或能够特异性结合包含GPCR和G蛋白的复合物。根据本发明的细胞可以是任何原核或真核生物。优选的,细胞是内源性或重组表达目标GPCR和/或G蛋白的真核细胞,例如酵母细胞或昆虫细胞,或培养细胞系,例如哺乳动物细胞系,优选人细胞系。使用的细胞的本质通常取决于生产天然蛋白质的容易程度和成本、所需的糖基化特性、靶蛋白的来源、预期用途,或其任意组合。用于蛋白质生产的真核细胞或细胞系是本领域普遍已知的,包括具有修饰的糖基化通路的细胞系,下文提供了非限制性例子。

[0189] 适合容纳、表达和生产用于后续分离和/或纯化的蛋白质的动物或哺乳动物宿主细胞包括中华仓鼠卵巢细胞(CHO),如CHO-K1(ATCC CCL-61)、DG44(Chasin等人,1986, Som.Ceu Molec.Genet.,12:555-556;和Kolkekar等人,1997,Biochemistry,36:10901-10909)、CHO-K1Tet-On细胞系(Clontech)、标明为ECACC85050302的CHO(CAMR,Salisbury,Wiltshire,UK)、CHO克隆13(GEIMG,Genova,IT)、CHO克隆B(GEIMG,Genova,IT)、标明为ECACC93061607的CHO-K1/SF(CAMR,Salisbury,Wiltshire,UK)、标明为ECACC92052129的

RR-CHOK1 (CAMR, Salisbury, Wiltshire, UK)、二氢叶酸还原酶阴性的CHO细胞(CHO/-DHFR, Urlaub和Chasin, 1980, Proc. Natl. Acad. Sci. USA, 77:4216)和dp12.CHO细胞(美国专利号5,721,121):用SV40转化的猴肾CV1细胞(COS细胞,COS-7,ATCC CRL-1651):人胚胎肾细胞(例如,293细胞或293T细胞,或用于悬浮培养的亚克隆293细胞,Graham等人,1977, J. Gen. Virol., 36:59,或GnTI KO HEK293S细胞,Reeves等人,2002,PNAS,99:13419):幼仓鼠肾细胞(BHK,ATCC CCL-10):猴肾细胞(CV1,ATCC CCL-70):非洲绿猴肾细胞(VERO-76, ATCC CRL-1587;VERO,ATCC CCL-81):小鼠sertoli细胞(TM4,Mather,1980,Biol.Reprod., 23:243-251):人宫颈癌细胞(HELA,ATCC CCL-2):犬肾细胞(MDCK,ATCC CCL-34):人肺细胞(W138,ATCC CCL-75):人肝细胞瘤细胞(HEP-G2,HB8065):小鼠乳房肿瘤细胞(MMT060562, ATCC CCL-51):水牛大鼠肝细胞(BRL3A,ATCC CRL-1442):TRI细胞(Mather,1982,Annals NYAcad.Sci.,383:44-68):MCR5细胞;FS4细胞。根据特定的实施方案,细胞是选自Hek293细胞或COS细胞的哺乳动物细胞。

[0190] 示例性非哺乳动物细胞系包括但不限于Sf9细胞、杆状病毒-昆虫细胞系统(例如,综述Jarvis,Virology,第310卷,第1期,2003年5月25日,第1-7页)、植物细胞,如烟草细胞、番茄细胞、玉米细胞、藻类细胞、或酵母,如酵母属(*Saccharomyces*)物种、裂殖酵母属(*Schizosaccharomyces*)物种、汉逊酵母属(*Hansenula*)物种、耶氏酵母属(*Yarrowia*)物种或毕赤酵母属(*Pichia*)物种。根据特定的实施方案,真核细胞是来自酵母属物种(例如酿酒酵母(*Saccharomyces cerevisiae*))、裂殖酵母属物种(例如栗酒裂殖酵母(*Schizosaccharomyces pombe*))、汉逊酵母属物种(例如多形汉逊酵母(*Hansenula polymorpha*))、耶氏酵母属物种(例如解脂耶氏酵母(*Yarrowia lipolytica*))、克鲁维酵母属(*Kluyveromyces*)物种(例如乳酸克鲁维酵母(*Kluyveromyces lactis*))、毕赤酵母属物种(例如巴斯德毕赤酵母(*Pichia pastoris*))或Komagataella物种(例如Komagataella *pastoris*)的酵母细胞。根据特定的实施方案,真核细胞是毕赤酵母细胞,在最特定的实施方案中,是巴斯德毕赤酵母细胞。

[0191] 可以按Sambrook和Russel(Molecular Cloning,A Laboratory Manual,第3版,第3卷,第16章,第16.1-16.54节)概括的下列原理,转染靶细胞(例如哺乳动物细胞)。此外,还可以使用如腺病毒载体的试剂实施病毒转导。选择恰当的病毒载体系统、调节区和宿主细胞是本领域普通技术人员掌握的常识。根据标准实践将获得的转染细胞保持在培养物中,并冷冻用于后续用途。

[0192] 相应的,本发明的另一个方面涉及生产根据本发明的结合结构域的方法,方法至少包括步骤:

[0193] a) 在合适的细胞表达系统(如上文定义)中表达根据本发明的核酸,和任选的

[0194] b) 分离和/或纯化所述结合结构域。

[0195] 本文描述的结合结构域、复合物、细胞或细胞系可用于多种环境 and 应用,例如但不限于用于捕获和/或纯化GPCR:G蛋白复合物,和GPCR:G蛋白复合物的结晶化研究和高分辨率结构分析。因此,本发明的一个目的是利用根据本发明的结合结构域,特别是免疫球蛋白单可变结构域(如纳米抗体)作为稳定GPCR:G蛋白复合物的工具,并进一步涉及利用这些结合结构域作为与G蛋白复合的GPCR的共结晶化的辅助,或换言之促进GPCR:G蛋白复合物的晶体发生。额外的和/或可选的,结合结构域和优选的表达结合结构域的细胞系统(如本文

所述)可用于其他用途,如配体筛选、药物发现、免疫,下文别处将全面描述。

[0196] 稳定GPCR:G蛋白复合物和使GPCR锁定在G蛋白结合状态

[0197] 因此,根据一个方面,本发明涉及上述结合结构域的用途,用于稳定包含GPCR和G蛋白的复合物。根据优选的实施方案,被稳定的复合物进一步包括受体配体,更具体的激动剂。如本文中使用的,术语“稳定”或“增加稳定性”指在复合物的一个或两个构成蛋白质(特别是GPCR和/或G蛋白)的结构(构象状态)和/或特定生物学活性(胞内信号传递活性)方面,增加GPCR:G蛋白复合物的稳定性。在一个特别优选的实施方案中,本发明的结合结构域可用于稳定GPCR:G蛋白复合物,使GPCR锁定或固定在活性的或结合G蛋白的状态。采用这类活性的或结合G蛋白的状态的GPCR将执行其天然的生物学活性。上文已描述了确定(增加的)GPCR:G蛋白复合物的稳定性的方式,实施例章节中将进一步示例。

[0198] 可以理解,在GPCR的结构和/或特定生物学活性方面,具有增加的稳定性包括对其他变性剂或变性条件的稳定性,包括热、去垢剂、离液剂和极端pH。相应的,在其他实施方案中,根据本发明的结合结构域能够增加GPCR:G蛋白复合物在由稀释、浓缩、缓冲液组成、加热、冷却、冰冻、去垢剂、离液剂、pH等诱导的非生理条件下的稳定性。相应的,术语“热稳定”、“增加热稳定性”指GPCR:G蛋白复合物的功能特性而非热动力学特性,并且是指构成蛋白对由热和/或化学方式诱导的不可逆变性的抗性,所述方式包括但不限于加热、冷却、冰冻、化学变性剂、pH、去垢剂、盐、添加剂、蛋白酶或温度。不可逆变性导致蛋白质功能性构象的不可逆解折叠、生物学活性丢失和变性蛋白质凝聚。如本文中使用的,术语“(热)稳定”、“增加(热)稳定性”应用于包埋在脂质颗粒或脂层(例如,脂单层、脂双层等)中的GPCR:G蛋白复合物,和溶解在去垢剂中的GPCR:G蛋白复合物。

[0199] 对于增加的热稳定性,可以通过测量配体结合或使用光谱学方法(如对升高温度时解折叠敏感的荧光、CD或光散射),方便地确定。优选的是,结合结构域能够增加稳定性,如通过GPCR:G蛋白复合物的热稳定性增加至少2℃、至少5℃、至少8℃,和更优选至少10℃或15℃或20℃所测量的。根据另一个优选的实施方案,结合结构域能够增加含受体配体(更具体的是GPCR依赖性信号传递通路的激动剂或正变构调控剂)的GPCR:G蛋白复合物的热稳定性。根据另一个优选的实施方案,本发明的结合结构域能够增加GPCR:G蛋白复合物在存在去垢剂或离液剂的条件下的稳定性。优选的,结合结构域能够增加GPCR:G蛋白复合物对由热或化学方法诱导的变性的稳定性。对于增加的对热、去垢剂或离液剂的稳定性,通常在存在测试去垢剂或测试离液剂的条件,将GPCR:G蛋白孵育定义的时间,并使用例如配体结合或光谱学方法确定稳定性,任选在增加的温度下,如上文所讨论的。根据仍然另一个优选的实施方案中,根据本发明的结合结构域能够增加GPCR的功能性构象状态对极端pH的稳定性。优选的,结合结构域能够增加GPCR:G蛋白复合物对极端pH的稳定性。对于极端的pH,可选择典型的测试pH,例如在6至8、5.5至8.5、5至9、4.5至9.5的范围内,更具体的是在4.5至5.5的范围内(低pH)或8.5至9.5的范围内(高pH)。

[0200] 在特别优选的实施方案中,根据本发明的结合结构域可用于阻止复合物在存在核苷酸,特别是鸟嘌呤核苷酸或其类似物的条件下解离。更具体而言,鸟嘌呤核苷酸包括GDP和GTP,鸟嘌呤核苷酸的类似物包括但不限于GTP γ S或GDP与氟化铝或氟化铍的组合,或核苷酸片段(如焦磷酸盐或膦甲酸)。

[0201] 捕获和/或纯化GPCR:G蛋白复合物

[0202] 可理解,形成稳定的GPCR:G蛋白复合物的能力对于捕获和/或纯化GPCR:G蛋白复合物是特别有效的,其允许后续结晶化、配体表征和化合物筛选、免疫等。此外,特殊的优点是,本发明的结合结构域可以是可用于一系列GPCR:G蛋白复合物的普遍有效工具。

[0203] 相应的,本发明还考虑了捕获和/或纯化包含GPCR和G蛋白的复合物的方法,方法包括步骤:

[0204] a) 提供根据本发明的结合结构域,和

[0205] b) 允许结合结构域与包含GPCR和G蛋白的复合物结合,和

[0206] c) 任选的分离步骤b)中形成的复合物。

[0207] 在特定的实施方案中,本发明提供了捕获包含GPCR和G蛋白的复合物的方法,方法包括步骤:

[0208] a) 将含有多个GPCR和G蛋白的溶液用于具有根据本发明的固定化结合结构域的固体支持物,和

[0209] b) 形成结合结构域、GPCR和G蛋白的复合物,和

[0210] c) 去除微弱结合的或未结合的分子。

[0211] 本发明还考虑了纯化包含GPCR和G蛋白的复合物的方法,方法包括步骤:

[0212] a) 将包含GPCR和G蛋白的溶液与根据本发明的结合结构域接触,和

[0213] b) 形成包含结合结构域、GPCR和G蛋白的复合物,和

[0214] c) 分离步骤b)的复合物,

[0215] 其中,GPCR和G蛋白的复合物是基本纯化的。

[0216] 根据特定的实施方案,本文所述的结合结构域还可用于捕获进一步包含受体配体和/或一个或多个其它相互作用蛋白质的靶GPCR:G蛋白复合物。

[0217] 上述用于捕获/纯化靶GPCR:G蛋白复合物的方法包括但不限于基于亲和力的方法,如亲和层析、亲和纯化、免疫沉淀、蛋白质检测、免疫组化、表面展示等,是本领域普遍已知的。

[0218] 结晶化和解析GPCR:G蛋白复合物的结构

[0219] 包括GPCR在内的膜蛋白的结晶化仍然是艰巨的挑战。虽然出现了允许生成毫克量级的表达和纯化方法,但使这些分子获得稳定性可能是最难克服的障碍。首先,根据本发明的结合结构域可增加去垢剂溶解的GPCR:G蛋白复合物的稳定性,保护其免受蛋白水解的降解和/或聚集,以及促进纯化和浓缩正确折叠的蛋白质的均质样品。本领域的普通技术人员可认识到,这类样品是用于生成衍射晶体的优选起点。

[0220] 结晶化是通过X射线晶体学确定大分子结构的方法的另一个主要瓶颈。成功的结晶化需要形成核,及其之后生长为合适大小的晶体。晶体生长一般在过饱和溶液中自发发生,是均质成核的结果。在经典的稀疏矩阵筛选实验中,蛋白质可以结晶,其中对沉淀剂、添加剂和蛋白质的浓度广泛取样,可以鉴别对于特定蛋白质的适合成核和晶体生长的过饱和条件。稀疏矩阵筛选方法相关的是在蛋白质本身中产生结构差异,例如通过添加与蛋白质结合的配体,或通过产生不同的突变(优先在靶蛋白质的表面残基中),或通过尝试结晶靶蛋白质的不同种类的直向同源物(Chang1998)。本发明的一个预料之外的发现是特异性结合GPCR:G蛋白复合物的结合结构域的有效性,用于在结合时导入一定程度的结构改变,同时保留复合物的整体折叠。

[0221] 因为结晶化涉及不利的丢失了待装配在晶格中的分子的构象熵,所以降低仍然在溶液中的靶的构象熵的方法应该通过降低晶格形成的净熵罚分 (net entropic penalty), 增强了结晶的可能性。已证实“表面熵减”方法是非常有效的 (Derewenda2004)。类似的, 结合伙伴对象 (如离子、小分子配体和肽) 可以通过结合和稳定蛋白质的构象状态亚组降低构象异质性。虽然这类结合伙伴对象是有效的, 但并非所有的蛋白质都具有已知的结合伙伴对象, 甚至当已知结合伙伴对象时, 其亲和力、溶解度和化学稳定性可能与结晶试验不相容。因此, 令人惊讶的发现本发明的结合结构域可用作工具, 通过结合G蛋白的特定构象, 使靶GPCR:G蛋白复合物的构象异质性最小化, 增加获得良好有序晶体的可能性。

[0222] 用于高分辨率结构研究的GPCR结晶是特别困难的, 因为这类膜蛋白的两亲性表面。包埋在膜双分子层中, 蛋白质与磷脂的酰基链的接触位点是疏水性的, 而极性表面则暴露给脂类的极性头部基团和水相。为了获得良好有序的三维晶体——高分辨率X射线结构分析的必要前提——在去垢剂的帮助下溶解GPCR, 并作为蛋白质-去垢剂复合物对其纯化。去垢剂微团以带状 (belt-like) 方式覆盖了膜蛋白的疏水性表面 (Hunte和Michel2002; Ostermeier等人, 1995)。GPCR-去垢剂复合物形成三维晶体, 其中通过从去垢剂微团上突出的蛋白质的极性表面产生相邻蛋白质分子之间的接触 (Day等人, 2007)。显而易见的是, 去垢剂微团需要晶格中的空间。虽然微团之间的吸引性相互作用可以稳定晶体堆积 (Rasmussen等人, 2007; Dunn等人, 1997), 但这些相互作用不导致僵硬的晶体接触。由于包括GPCR在内的许多膜蛋白含有相对小或高度灵活的亲水性结构域, 因此, 一种增加得到良好有序晶体的可能性的对策是扩大蛋白质的极性表面和/或降低其柔性。最生理学的方法是使用天然的信号传递伙伴对象 (partner), 如G蛋白或抑制蛋白。不幸的是, GPCR与G蛋白或抑制蛋白的相互作用高度依赖于脂类, 难以形成具有足够稳定性的复合物进行用于晶体学。因此, 本发明的结合结构域通过结合G蛋白, 可用于扩大GPCR的极性表面, 补充可促进晶格的分子与G蛋白和纳米抗体的极性表面间初级接触的蛋白质表面的量。本发明的结合结构域还可以降低其胞外区的柔性, 以生长良好有序的晶体。免疫球蛋白单可变结构域, 包括纳米抗体尤其适合该目的, 因为它们结合构象表位, 并包含一个僵硬的球状结构域, 与常规抗体或来自其的片段 (如Fab) 不同, 缺少柔性接头区。

[0223] 因此, 根据优选的实施方案, 本发明提供了可用作使包含GPCR和G蛋白的复合物结晶, 最终解析结构的工具的结合结构域。更优选的, 利用本发明的结合结构域结晶的复合物还包括受体配体, 更具体的激动剂。在特别优选的实施方案中, 包含在复合物中的GPCR处于活性状态或构象。

[0224] 因此, 可以使用多种专门用于膜蛋白的结晶方法, 使与GPCR:G蛋白复合物以及任选地受体配体复合的结合结构域结晶, 许多方法见Caffrey (2003&2009) 的综述。一般而言, 方法是基于脂类的方法, 包括在结晶前向复合物中添加脂类。这类方法之前已用于结晶其他膜蛋白。许多这类方法, 包括脂质立方相结晶方法和bicelle结晶方法, 利用了脂类和去垢剂作为小泡的自发的自装配特性 (小泡融合方法)、盘状微团 (bicelle方法)、和脂质晶体或中间相 (中间相或立方相方法中)。脂质立方相结晶方法描述在例如: Landau等人, 1996; Gouaux1998; Rummel等人, 1998; Nollert等人, 2004, Rasmussen等人, 2011中, 通过引用这些方法的公开内容整合这些出版物。Bicelle结晶方法描述在例如: Faham等人, 2005; Faham等人, 2002中, 通过引用这些方法的公开内容整合这些出版物。

[0225] 根据另一个实施方案,本发明涉及本文所述结合结构域的用途,用于解析包含GPCR和G蛋白的靶复合物的结构,所述复合物任选还包括受体配体。本文中使用的“解析结构”指确定蛋白质的原子或原子坐标排列,通常是通过生物物理学方法进行的,如X射线晶体学。

[0226] 在X射线晶体学中,正确组合的衍射数据给出了单位晶胞中的分子的电子密度的3D傅里叶变换幅度。如果已知相位,则可以通过傅里叶合成简单的获得电子密度。对于蛋白质复合物,当具有已知结构(研究模型)的蛋白质比例低(小于50%的氨基酸含量)时,和/或当晶体表现出有限的衍射质量时,仅从分子置换(MR)成功推断相位信息是令人质疑的。虽然已证实了多个同晶型置换(MIR)和MR定相的组合可成功用于蛋白质复合物(例如,Ostermeier等人,1995;Li等人,1997;Hunte等人,2000),但生产良好的重原子衍生物的需要几乎总是困难重重。在过去的二十年里,反常色散数据的应用普遍超越了经典的MIR方法,所述反常色散数据主要使用硒代甲硫氨酸(SeMet)掺入(MAD或SAD)(Hendrickson1991)。实际上,使用Se-边缘能量的反常实验数据普遍提供了比MIR或基于模型的MR定相数据更好且偏见更少的相位信息。相应的,一个特定的实施方案涉及了根据本发明的结合结构域的用途,用于通过MR或MAD定相GPCR:G复合物。特别的是,免疫球蛋白单可变结构域,包括纳米抗体一般是大量表达的,适合SeMet掺入。为了进一步示例而非限制,通过仅在纳米抗体中导入所有的SeMet位点来定相包含GPCR、G蛋白和纳米抗体的复合物规避了在GPCR或G蛋白中掺入SeMet位点的要求。

[0227] 在许多情况下,获得衍射级的晶体是解决原子分辨率结构的主要障碍。因此,根据特定的实施方案,可以使用本文所述的结合结构域改善晶体的衍射质量,从而可以解析靶复合物的晶体结构。

[0228] 此外,获得GPCR靶的结构信息,例如用于帮助指导GPCR药物发现,是非常理想的。除了结晶更多的GPCR外,尤其需要用于获得与不同类型的配体结合的受体结构的方法,所述配体包括激动剂、拮抗剂、变构调节剂和/或G蛋白。本发明特别提供了用于获得GPCR:G蛋白复合物的晶体的一般性工具。特别的是,结合激动剂的GPCR:G蛋白复合物晶体可提供活性状态的GPCR的三维展示。这些结构将帮助澄清与配体结合和G蛋白相互作用位点相关联的构象改变,获得更精确的机制假设,最终获得新的治疗剂。考虑到配体活化的GPCR内在的构象柔性和结合激动剂的受体所表现出的更大的异质性,这类状态是难以稳定的。因此,这方面的工作可受益于结合了激动剂的受体构象与其异源三聚体的G蛋白结合后的复合物的稳定作用,所述稳定作用是通过添加这类复合物特异性的结合结构域。尤其适合的是结合G蛋白的形成了这类复合物的一部分的结合结构域,因为这类结合结构域可用作稳定所有通过相同的G蛋白(例如,Gs偶联受体、Gi偶联受体等)传递信号的GPCR的普遍工具。

[0229] 根据备选的实施方案,本发明涵盖了确定包含GPCR和G蛋白的复合物的晶体结构的方法,方法包括步骤:

[0230] a) 提供根据本发明的结合结构域,和

[0231] b) 允许结合结构域与包含GPCR和G蛋白的复合物结合,和

[0232] c) 使步骤b)中形成的复合物结晶。

[0233] 在上述确定晶体结构的方法的特定实施方案中,包含GPCR和G蛋白的靶复合物还包括结合GPCR的受体配体,更具体的激动剂。

[0234] 可以通过生物物理学方法进行所述晶体结构的确定,如X射线晶体学。方法还可包括获得晶体的原子坐标的步骤(如上文定义)。

[0235] 鉴别靶向GPCR:G蛋白复合物的化合物

[0236] 在化合物筛选、药物发现和先导物优化的过程中,需要更快、更有效、成本更低、尤其是信息量更丰富的筛选测定,所述测定提供了各种化合物特征及其对各种细胞通路的影响(即,功效、特异性、毒性和药物代谢)的同步信息。因此,需要快速和低价的筛选大量的化合物,从而鉴别目标GPCR的新的特异性配体,优选构象特异性配体,其可能是潜在的新药候选物。本发明通过提供使处于功能性构象状态的GPCR:G蛋白复合物稳定的结合结构域,解决了该难题,所述复合物之后可用作免疫原或选择试剂,用于在多种情况下筛选。根据本发明的结合结构域的主要优点是包含在GPCR:G蛋白复合物中的GPCR可以保持在稳定的功能性构象中,特别是活性状态构象中。例如,选择性结合该受体活性构象的文库化合物具有更大的倾向性表现为激动剂,因为GPCR的活性构象的orthosteric或变构稳定引发生物学应答。另一个优点是结合结构域增加了包含在复合物中的GPCR活性构象的热稳定性,因而保护GPCR免受由化合物筛选和药物发现中使用的非自然条件导致的不可逆变性或热变性,而不需要依赖具有增加的稳定性的突变GPCR。根据本发明的构象选择性结合结构域的另一个主要优点是其允许快速可靠的筛选和区分GPCR和GPCR依赖性通路的受体激动剂、反向激动剂、拮抗剂和/或调控剂,以及抑制剂,从而增加了鉴别具有理想药理学特性的配体的可能性。

[0237] 为了进一步示例,非常明确的概念是当与G蛋白复合时,绝大部分GPCR表现出更高的激动剂结合亲和力。这归功于被激动剂占据的受体和G蛋白之间的协作相互作用。抑制GPCR和G蛋白复合物的解离的本发明的结合结构域因而稳定R:G蛋白复合物的活性构象状态,因而增加了GPCR与激动剂的亲和力,减少了与反向激动剂的亲和力。因此,识别R:G复合物的活性功能性构象的结合结构域可例如用于高通量筛选测定中以筛选激动剂,因为相对于反向激动剂,其增加受体与激动剂的亲和力。识别G:R复合物的活性功能性构象的结合结构域还可用于高通量筛选测定中筛选偏向性激动剂,所述偏向性激动剂具有选择性刺激某类受体的信号传递活性的能力,例如相对于 β -抑制蛋白的功能,选择性激活G蛋白。根据特定的实施方案,可以使用特异性结合与GPCR复合的G蛋白的结合结构域(例如,表2-3)作为靶向多种GPCR的筛选程序的普遍工具,因为特定的G蛋白(例如Gs)可与多种GPCR(例如,Gs偶联受体,包括5-HT受体5-HT₄和5-HT₇型、ACTH受体、腺苷受体A_{2a}和A_{2b}型、精氨酸血管加压素受体2、 β -肾上腺素能受体 β_1 、 β_2 和 β_3 型、降钙素受体、降钙素基因相关肽受体、促肾上腺皮质激素释放激素受体、多巴胺受体D₁样家族(D₁和D₅)、FSH-受体、胃抑制型多肽受体、胰高血糖素受体、组胺H₂受体、黄体生成素/绒毛膜促性腺激素受体、黑皮质素受体、甲状旁腺激素受体1、前列腺素受体D₂和I₂型、分泌素受体、促甲状腺素受体等;还参见表1)形成复合物。

[0238] 因此,根据本发明的另一个方面涵盖了结合结构域的用途,或包含结合结构域的复合物、细胞、膜制品的用途,均如上文所述,用于GPCR:G蛋白复合物的构象特异性结合伙伴对象的筛选和/或鉴别程序,最终可导致潜在的新药候选物。

[0239] 根据一个实施方案,本发明考虑了鉴别能够选择性结合GPCR:G蛋白复合物的化合物的方法,方法包括步骤:

[0240] (i) 提供包含GPCR和G蛋白的复合物

[0241] (ii) 使复合物与针对和/或特异性结合复合物的结合结构域接触, 允许结合结构域结合复合物, 和

[0242] (iii) 提供测试化合物, 和

[0243] (iv) 评估测试化合物是否结合复合物, 和

[0244] (v) 选择选择性结合复合物的化合物。

[0245] 显而易见的是, 用于任何上述方法的结合结构域都能够稳定GPCR:G蛋白复合物的功能性构象状态, 和阻止复合物解离。优选的, GPCR:G蛋白复合物处于活性构象状态(如上文定义)。根据上述筛选方法的特别优选的实施方案, GPCR:G蛋白复合物还包括受体配体。

[0246] 应该注意到, 结合结构域的特别优选的实施方案如上文关于本发明之前方面所述。

[0247] 因此, 本发明的结合结构域可用于筛选测定。用于药物发现的筛选测定可以是固相或液相的测定, 例如结合测定, 如放射性配体结合测定。可以理解的是, 在一些情况下, 测试化合物的高通量筛选是优选的, 上述方法可用作“文库筛选”方法, 这是本领域技术人员普遍已知的术语。因此, 测试化合物可以是测试化合物的文库。特别的是, 用于治疗化合物的高通量筛选测定构成本发明的一部分, 所述治疗化合物如激动剂、拮抗剂或反向激动剂和/或调控剂。出于高通量的目的, 可以使用化合物文库, 如变构化合物文库、肽文库、抗体文库、基于片段的文库、合成化合物文库、天然化合物文库、噬菌体展示文库等。用于制备和筛选这类文库的方法是本领域已知的。在一个优选的实施方案中, 高通量筛选方法涉及提供含有大量潜在的治疗配体的组合的化学品或肽文库。然后在一个或多个测定中筛选这类“组合文库”或“化合物文库”, 如本文所述, 鉴别表现出理想的特征活性的文库成员(特定化学种类或亚类)。“化合物文库”是一类储藏的化学品集合, 通常最后用于高通量筛选中。“组合文库”是通过组合多种化学的“构件”(如试剂), 通过化学合成或生物学合成生成的多种化学化合物的集合。组合文库的制备和筛选是本领域技术人员普遍已知的。因而鉴别的化合物可作为常规的“先导化合物”发挥作用, 或者本身可用作潜在的或实际的治疗剂。因此, 在另一个实施方案中, 本文上述的筛选方法还包括修饰已表现出与构象活性的GPCR:G蛋白复合物结合的测试化合物, 以及确定修饰的测试化合物是否与保持在特定构象中的GPCR结合。

[0248] 当优选就构象特异性结合伙伴对象对靶GPCR:G蛋白复合物进行高通量筛选时, 这可以通过将根据本发明的结合结构域, 或被结合结构域稳定的GPCR:G蛋白复合物固定在可以被阵列或其他多路的合适固体表面或支持物上予以促进。合适的固体支持物的非限制性例子包括珠子、柱、载玻片、芯片或平板。更具体而言, 固体支持物可以是颗粒状(例如, 珠状或粒状, 一般用于抽提柱中)或片层形态(例如, 膜或滤纸、玻璃或塑料载片、微滴度测定板、测验片、毛细管填充装置等), 其可以是平板的、褶层的或中空的纤维或管。下列基质作为例子而非穷举给出, 这类例子包括硅石(多孔无定形硅石), 即, 含有由Biotage (Dyax Corp. 的分部) 供应的60A不规则硅石(32-63um或35-70um)的FLASH系列柱体; 琼脂糖或聚丙烯酰胺支持物, 如由Amersham Pharmacia Biotech供应的Sephacrose系列产品, 或由Bio-Rad供应的Affi-Gel支持物。此外, 还有大孔聚合物, 如Bio-Rad提供的压力稳定型Affi-Prep支持物。可利用的其他支持物包括: 葡聚糖、胶原、聚苯乙烯、甲基丙烯酸盐、海藻酸钙、可控孔玻璃、铝、钛和多孔陶瓷。可选的, 固体表面可包括一部分质量依赖型传感器, 例如, 表面等离

振子共振检测仪。可商购的支持物的其他例子讨论在例如Protein Immobilisation, R.F.Taylor编著, Marcel Dekker, Inc., New York, (1991) 中。

[0249] 固定可以是非共价或共价的。特别的是, 根据本发明的结合结构域, 或通过结合结构域稳定的GPCR:G蛋白复合物, 在固体表面的非共价固定或吸附可如下产生, 通过根据本领域技术人员已知的标准技术, 用任何识别与结合结构域或GPCR连接的分子标签的抗体、或链霉抗生物素或抗生物素蛋白、或金属离子包被表面(例如, 生物素标签、组氨酸标签等)。可选的, 可以使用常规的偶联化学, 通过共价交联, 将根据本发明的结合结构域, 或通过结合结构域稳定的GPCR:G蛋白复合物连接到固体表面。固体表面可天然的包括适用于共价连接的可交联残基, 或者可以将其包被或衍生而导入根据本领域普遍已知的方法的合适的可交联基团。在一个特定的实施方案中, 在通过不含化学间隔子臂的反应部分与理想基质直接共价偶联后, 固定化蛋白质保留了足够的功能。Jung等人, (2008) 中讨论了关于在固体支持物上固定抗体(片段)的方法的其他例子和更详细的信息; 类似的, Cooper (2004) 中综述了膜受体固定方法; 都通过引用整合到本文中。值得注意的是, 例如, 将特定氨基酸(在具有已知或提示结构的蛋白质中)突变为赖氨酸或半胱氨酸(或其他理想的氨基酸)可以为共价偶联提供特异性位点。还可能重新改造特定蛋白质, 改变涉及化学偶联的表面可利用氨基酸的分布(Kallwass等人, 1993), 这有效控制偶联蛋白质的方向。类似的方法可用于根据本发明的结合结构域, 以及构象稳定化的GPCR:G蛋白复合物, 从而提供在不添加其他肽尾部或含有天然或非天然氨基酸的结构域的条件下, 定向固定的手段。在抗体或抗体片段(如纳米抗体)的情况下, 在框架区中导入突变是优选的, 使对抗体(片段)的抗原结合活性的破坏最小化。

[0250] 常规上, 固定的蛋白质可用于免疫吸附方法, 如免疫测定(例如ELISA), 或免疫亲和纯化方法, 这是通过根据本领域常规的标准方法, 使固定的根据本发明的蛋白质与测试样品(即, 包含测试化合物等)接触来进行的。可选的, 特别是出于高通量的目的, 可以排列或多路传输固定的蛋白质。优选的, 固定的根据本发明的蛋白质可用于筛选和选择特异性结合构象稳定的GPCR:G蛋白复合物的化合物, 其中特别是处于活性构象状态的GPCR。

[0251] 可理解, 根据用途的类型或所需筛选的类型, 可以固定结合结构域, 或(构象稳定的)GPCR:G蛋白复合物, 或其构成蛋白。此外, 对GPCR:G蛋白稳定化结合结构域(靶向GPCR:G蛋白复合物的特定构象表位)的选择将决定蛋白质的方向, 并且因此, 化合物鉴别的理想结果, 例如, 特异性结合所述构象稳定化GPCR的胞外部分、膜内部分或胞内部分的化合物, 或特异性结合所述构象稳定化G蛋白的化合物。

[0252] 可选的, 测试化合物(或测试化合物的文库)可固定在固体表面上, 如芯片表面, 而在例如去垢剂溶液或膜样制品中提供了结合结构域和GPCR:G蛋白复合物(参见下文)。

[0253] 最优选的, 结合结构域、GPCR:G蛋白复合物或其构成蛋白, 或测试化合物都是没有固定的, 正如在例如溶液中的噬菌体展示选择过程或放射性配体结合测定中。在优选的实施方案中, 用于任何上述筛选方法的结合结构域、GPCR:G蛋白复合物(或分别的构成蛋白)是作为完整细胞或细胞(细胞器)提取物, 如膜提取物或其级分, 提供的, 或者可以被掺入到脂质层或小泡中(包括天然和/或合成的脂质)、高密度脂颗粒, 或任何纳米粒, 如纳米盘(nanodisk), 或作为VLP被提供, 从而保留了相应蛋白质充分的功能。Cooper (2004) 中详细综述了从膜片段或膜-去垢剂提取物形成的GPCR制品, 所述文献通过引用整合到本文中。可

选的,还可以将结合结构域、GPCR:G蛋白复合物或构成蛋白溶解在去垢剂中。实施例章节中还提供了溶解的受体制品的非限制性例子。

[0254] 可使用各种方法确定在稳定的GPCR:G蛋白复合物和测试化合物之间的结合,包括本领域的常规实践,例如酶联免疫吸附测定(ELISA)、表面等离子共振测定、基于芯片的测定、免疫细胞荧光、酵母双杂技术和噬菌体展示,例如在Sambrook等人,(2001), *Molecular Cloning, A Laboratory Manual*. 第3版, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY中。其他用于检测测试化合物和GPCR之间的结合的方法包括用离子喷雾质量光谱/HPLC方法超滤,或其他(生物)物理学和分析方法。还可以使用荧光能量共振转移(FRET)方法,例如本领域技术人员普遍已知的。可理解,可以使用与化合物关联的独特标记或标签,检测结合的测试化合物,所述标记或标签如肽标记、核酸标记、化学标记、荧光标记或放射性频率标签,如本文别处所述。

[0255] 可以使用任何本身已知的合适的体外测定、基于细胞的测定、体内测定和/或动物模型,或其任意组合,测试化合物和/或包含化合物的组合物的功效,这取决于所涉及的特定疾病或病症。

[0256] 在一个特定的实施方案中,确定了化合物是否改变GPCR与受体配体的结合(如本文定义)。可以使用本文所述的本领域已知的标准配体结合方法,测定GPCR与其配体的结合。例如,配体可以是放射性标记的或荧光标记的。可以在完整细胞上或从细胞获得的膜上或去垢剂水性溶解的受体上进行测定。通过化合物改变标记配体结合的能力来表征所述化合物(还参见实施例章节)。化合物可减少在GPCR及其配体之间的结合,或可增加在GPCR及其配体之间的结合,例如减少或增加至少2倍、3倍、4倍、5倍、10倍、20倍、30倍、50倍、100倍。因此,根据更特殊的实施方案,用于上述任何筛选方法的复合物还包括受体配体。优选的,受体配体选自小分子、多肽、抗体或来自其的任何片段、天然产物等。更优选的,受体配体是完全激动剂,或部分激动剂,或反向激动剂,或拮抗剂,如上文所述。

[0257] 除了确立与功能性构象状态的靶GPCR:G蛋白复合物结合外,理想的还确定化合物对GPCR:G蛋白复合物,特别是对GPCR及下游相互作用伙伴对象的生物学活性的功能效应。特别的是,根据本发明的结合结构域可用于筛选调控(增加或减少)GPCR:G蛋白复合物或其组分(GPCR或G蛋白)的生物学活性的化合物。所需的对生物学活性的调控取决于选择的GPCR。化合物可结合靶GPCR:G蛋白复合物,特别是结合它的一种或两种组分,导致调控(激活或抑制)下游的受体信号传递。该GPCR信号传递的调控可orthosterically或变构(allosterically)的发生。化合物可结合包含与G蛋白或其组分结合的GPCR的靶复合物,从而激活或增加受体信号传递;或可选的,从而减少或抑制受体信号传递。化合物还可以这样的方式结合靶复合物,使其阻断GPCR的构成活性。化合物还可以这样的方式结合靶复合物,使得它们介导变构调节(例如,在变构位点结合GPCR或G蛋白)。以该方式,化合物可通过结合GPCR:G蛋白复合物(例如,在变构位点)中的不同区域,调节受体功能。例如参考George等人,(2002), Kenakin (2002) 和Rios等人,(2001)。本发明的化合物还可以这样的方式结合靶复合物,使其延长GPCR介导的信号传递的持续期,或通过增加受体-配体亲和力,增强受体信号传递。此外,化合物还可以这样的方式结合靶复合物,使其抑制或增强GPCR功能性同聚物或异聚物的装配。

[0258] 此外,基于细胞的测定对于评估新生物学靶的作用机制和化学化合物的生物学活

性是关键。目前用于GPCR的基于细胞的测定包括测量通路激活(Ca^{2+} 释放、cAMP生成或转活性);通过用GFP标记GPCR和下游元件,测量蛋白质运输;和使用Förster共振能量转移(FRET)、生物发光共振能量转移(BRET)或酵母双杂方法,直接测量在蛋白质之间的相互作用。通过本领域普遍已知和常规使用的任何手段,将本发明的结合结构域在细胞内细胞内导入至细胞的相关区室(胞内或胞外),可导致新的或更好的基于细胞的测定。

[0259] 特别的是,需要将尚未鉴别出天然的激活配体的那些GPCR“去孤儿化”。使用根据本发明的结合结构域稳定处于功能性构象状态中的GPCR,能实现用于鉴别“孤儿”GPCR的配体的筛选方法,所述“孤儿”GPCR的天然配体是未知的。例如,采用各种方法“去孤儿化”,包括针对已知配体的家族进行阵列筛选。可以从生物学样品中鉴别孤儿GPCR的配体。因此,在特定的实施方案中,提供测试化合物作为生物学样品。特别的是,样品可以是个体采集的任何合适的样品。例如,样品可以是体液,如血液、血清、血浆、脊髓液。可选的,样品是组织或细胞提取物。

[0260] 用于任何上述筛选方法的测试化合物可选自多肽、肽、小分子、天然产物、肽模拟物、核酸、脂类、脂肽、碳水化合物、抗体或来自其的任何片段,如Fab、Fab'和F(ab')₂、Fd、单链Fv(scFv)、单链抗体、二硫键连接的Fv(dsFv)和包含VL或VH结构域的片段、重链抗体(hcAb)、单结构域抗体(sdAb)、微抗体、源自骆驼重链抗体的可变结构域(VHH或纳米抗体)、源自鲨鱼抗体(VNAR)的新型抗原受体的可变结构域、蛋白质支架(包括abody)、蛋白A、蛋白G、设计的锚蛋白重复结构域(DARPin)、III型纤连蛋白重复、anticalins、knottins、改造的CH2结构域(纳米抗体),如上文定义。

[0261] 测试化合物可任选的与可检测的标记共价或非共价的连接。合适的可检测标记和用于连接、使用 and 检测它们的技术是本领域技术人员显而易见的,包括但不限于任何通过光谱、光化学、生物化学、免疫化学、电学、光学或化学手段可检测的组合物。有效的标记包括磁珠(例如,dynabead)、荧光染料(例如,所有的Alexa Fluor染料、荧光素异硫氰酸盐、Texas红、罗丹明、绿色荧光蛋白等)、放射性标记(例如,³H、¹²⁵I、³⁵S、¹⁴C或³²P)、酶(例如,辣根过氧化物酶、碱性磷酸酶)和比色标记,如胶体金或有色玻璃或塑料(例如,聚苯乙烯、聚丙烯、胶乳等)珠子。检测这类标记的方法是本领域技术人员已知的。因此,例如,可以使用胶卷或闪烁计数器,检测放射性标记,可以使用光检测器检测发散的光照来检测荧光标记物。通常通过为酶提供底物,和检测由酶对底物的作用生产的反应产物,检测酶标记。通过将有色标记简单可视化,检测比色标记。之前在本发明涉及结合结构域的第一方面中也描述了其他合适的可检测标记。

[0262] 根据特别优选的实施方案,测试化合物是抗体或来自其的任何片段,如上所述,包括纳米抗体。例如但不限于,测试化合物可以是针对被根据本发明的结合结构域稳定的GPCR:G蛋白复合物产生的抗体(如本文中最广泛含义定义的)。在体内产生抗体的方法是本领域已知的。优选的,以本文之前描述的相似方式免疫动物。本发明还涉及用于选择特异性结合构象稳定化的GPCR:G蛋白复合物的抗体的方法,涉及在所述GPCR:G蛋白复合物上,筛选编码免疫球蛋白基因或其部分的表达文库,所述文库是在细菌、酵母、丝状噬菌体、核糖体或核糖体亚基或其他展示系统中表达的。

[0263] 本发明的特定方面涉及固定了根据本发明的结合结构域的固体支持物。因而这类固体支持物(如上文所述)可用于任何上述筛选方法。

[0264] 调控GPCR受体信号传递

[0265] 本发明的结合结构域还可用于调节GPCR信号传递,特别是G蛋白介导的GPCR信号传递,包括消除G蛋白介导的GPCR信号传递。术语“调节”、“调节的”意指蛋白质或蛋白质复合物,特别是GPCR:G蛋白复合物的活性的增加或减少。特别的是,本发明的结合结构域可以是变构调控剂或变构抑制剂。术语“变构调控剂”或“变构抑制剂”在本发明的上下文中指非竞争性调控剂或抑制剂,其通过结合受体的活性位点以外的位点产生效应,调节受体活性或使受体在信号转导中无效。“正变构调控剂(PAM)”增加了信号转导,而“负变构调控剂(NAM)”降低了信号转导。特别的是,变构抑制剂还可以消除信号转导。用于评估本发明的结合结构域对GPCR信号传递的调控的测定如上文所述。

[0266] 在这一方面,根据特定的实施方案,本发明的结合结构域,特别是免疫球蛋白单可变结构域还可用于前导物鉴别和设计肽模拟物(peptidomimetics)。利用生物学相关的肽或蛋白质结构作为前导物鉴别的起点,代表了现代药物发现中最强大的方法之一。肽模拟物是关键元素(药效团(pharmacophore))的3D空间模拟天然肽或蛋白质的化合物,同时保留了与生物学靶相互作用和产生相同生物学效应的能力。肽模拟物被设计为规避与天然肽相关的一些难题:例如对蛋白分解的稳定性(活性持续期)和低下的生物利用度。通常可以显著改善其他一些特性,如受体选择性或效价。

[0267] 治疗和诊断应用

[0268] 某些上述结合结构域可具有治疗用途,并可施用给具有病况的对象,从而治疗对象的病况。结合结构域的治疗用途是由所述结合结构域结合的靶GPCR:G蛋白复合物决定的,因为通过所述GPCR的信号传递与病况关联。在某些情况下,可以通过在结合配体的条件下激活GPCR。在其他实施方案中,可以例如突变GPCR使其组成型激活。可应用本发明结合结构域治疗GPCR介导的病况,如精神分裂症、偏头痛、反流、哮喘、支气管痉挛、前列腺肥大、溃疡、癫痫、咽痛、过敏、鼻炎、癌症例如前列腺癌、青光眼和中风。在NCBI的世界范围内的网站中可发现Man数据库中的On-line Mendelian Inheritance下的其他示例性的GPCR相关病况。因此,本发明的特定实施方案还考虑了本发明的结合结构域,或包含结合结构域的药物组合物,用于治疗GPCR相关性疾病或病症。可理解,治疗用途还取决于结合结构域所针对的GPCR:G蛋白复合物的特定构象表位。

[0269] 在固定的药物组合物中,本发明的结合结构域可与另一种药物混合,或可以分开施用,在其他药物物质之前、同时或之后。一般而言,这些规程涉及向患有GPCR相关性疾病或病症的个体施用有效量的结合结构域,其调节宿主中的GPCR的信号传递活性和治疗个体的病症。

[0270] 在一些实施方案中,当需要降低某些GPCR的活性时,可施用一种或多种减少GPCR活性的化合物,而当需要增加某些GPCR的活性时,可施用一种或多种增加GPCR活性的化合物。

[0271] 根据本发明的方法,多种个体是可治疗的。这类个体一般是哺乳动物或哺乳类,其中这类术语用于广义的描述落入哺乳纲中的生物,包括食肉 动物目(例如狗和猫)、啮齿目(例如,小鼠、豚鼠和大鼠)和灵长目(例如,人、黑猩猩和猴子)。在多个实施方案中,个体是人。本发明的治疗方法通常在患这类病症的个体上实施,或者在需要避免这类病症的个体上实施。

[0272] 在另一个方面,本发明还涉及包含治疗有效量的本发明的结合结构域和至少一种可药用的载体、佐剂或稀释剂的药物组合物。

[0273] “载体”或“佐剂”,特别是“可药用的载体”或“可药用的佐剂”是任何合适的赋形剂、稀释剂、载体和/或佐剂,其本身不诱导产生对接受组合物的个体有害的抗体,也不引起保护作用。因此,可药用的载体本质上是无毒和非治疗性的,是本领域技术人员已知的。合适的载体或佐剂通常包括一种或多种包括在下列非穷举列表中的化合物:代谢缓慢的大型大分子,如蛋白质、多糖、聚乳酸、聚乙醇酸、多聚氨基酸、氨基酸共聚物和失活的病毒颗粒。作为非限制性的例子,载体或佐剂可以是林格氏液、右旋糖溶液或Hank溶液。还可以使用不含水的溶液,如非挥发性油类和油酸乙酯。优选的赋形剂是在生理盐水中的5%右旋糖。赋形剂可含有少量的添加剂,如增强等渗性和化学稳定性的物质,包括缓冲剂和防腐剂。

[0274] 可以通过口服、吸入或肠胃外的施用,来施用本文所述的结合结构域或其可药用的盐。在特定的实施方案中,通过鞘内或脑室内施用,来递送纳米抗体。可以单独施用活性化合物,或优选将其配制成药物组合物。有效治疗某些表达结合结构域识别的抗原的疾病或病症的量取决于常见的因素,如要治疗的病症的性质和严重程度,和哺乳动物的体重。然而,单位剂量通常在0.01至50mg,例如0.01至10mg,或0.05至2mg结合结构域或其可药用的盐。单位剂量通常每天施用一次或一次以上,例如每天2、3或4次,更常见每天1至3次,使总每日剂量通常在0.0001至1mg/kg的范围内;因此,对于70kg的成人,合适的总每日剂量是0.01至50mg,例如0.01至10mg,或更常见0.05至10mg。更优选的是,以单位剂量组合物的形式施用化合物或其可药用的盐,如单位剂量的口服、肠胃外的或吸入的组合物。这类组合物是通过混合制备的,适用于口服、吸入或肠胃外的施用,其形态可以是片剂、胶囊剂、口服液体制品、散剂、颗粒剂、锭剂、可复水的散剂、可注射和可输注的溶液剂或混悬剂或栓剂或气雾剂。用于口服施用的片剂和胶囊剂通常以单位剂量存在,并含有常规赋形剂,如粘合剂、填充剂、稀释剂、压片剂、润滑剂、崩解剂、着色剂、矫味剂(flavouring)和湿润剂。可以根据本领域普遍已知的方法包被片剂。使用的合适的填充剂包括纤维素、甘露醇、乳糖和其他类似试剂。合适的崩解剂包括淀粉、聚乙烯吡咯烷酮和淀粉衍生物,如淀粉羟乙酸钠。合适的润滑剂包括例如硬脂酸镁。合适的可药用的湿润剂包括十二烷基硫酸钠。这些固体的口服组合物可以通过掺合、填充、压片等常规方法制备。重复的掺合操作可用于将活性剂分布到整个应用了大量填充剂的上述组合物中。这类操作当然是本领域常规的。口服的液体制品可以是例如含水或油性混悬剂、溶液剂、乳剂、糖浆剂或酏剂的形态,或者可以表现为用于在使用前用水或其他合适溶媒复水的干燥粉末。这类液体制品可含有常规的添加剂,如悬浮剂,例如山梨醇、糖浆、甲基纤维素、明胶、羟乙基纤维素、羧甲基纤维素、硬脂酸铝凝胶或氢化的食用脂肪;乳化剂,例如卵磷脂、脱水山梨醇单油酸酯或阿拉伯胶;不含水溶媒(可包括食用油),例如杏仁油、分馏的椰子油、油状酯类如丙三醇、丙二醇或乙醇的脂类;防腐剂,例如对羟基苯甲酸甲酯或丙酯或山梨酸,以及根据需要常规的矫味剂或着色剂。口服制剂还包括常规的缓释制剂,如具有肠溶衣的片剂或颗粒剂。优选的,用于吸入的组合物作为单独的或与惰性载体(如乳糖)组合的鼻烟、或气雾剂、或喷雾器的溶液、或作为用于吹入的超细粉末存在用于向呼吸道施用。在上述情况下,活性化合物的颗粒恰当的具有小于50微米,优选小于10微米的直径,例如1至5微米,如2至5微米的直径。有利的吸入剂量在0.05至2mg,例如0.05至0.5mg,0.1至1mg或0.5至2mg。对于肠胃外的施用,制备含有本发明的化

合物和无菌溶媒的流体单位剂型。根据溶媒和浓度,活性化合物可以是悬浮的或溶解的。通常通过将化合物溶解在溶媒中,在装入合适的小瓶或安瓿瓶之前过滤除菌,并密封,来制备肠胃外的溶液。有利的是,溶媒中也溶解了佐剂(如局部麻醉剂)、防腐剂 and 缓冲剂。为了增强稳定性,在装入小瓶后可以冷冻组合物,并在真空下去除水分。以基本相同的方式制备肠胃外的重悬剂,除了将化合物悬浮而非溶解在溶媒中,并在悬浮在无菌溶媒中之前通过暴露给环氧乙烷来除菌。有利的是,组合物中包括表面活性剂或湿润剂,以促进活性化合物的均匀分布。恰当时,可包括少量的支气管扩张药,例如拟交感胺,如异丙肾上腺素、乙基异丙肾上腺素、舒喘宁(salbutamol)、苯肾上腺素和麻黄碱;黄嘌呤衍生物,如茶碱和氨茶碱、和皮质类固醇如泼尼松龙和肾上腺刺激物,如ACTH。作为普通实践,组合物通常附有手写或印刷的用于相关医学处理的说明。

[0275] 可以如关于肽、多肽和蛋白质所述,将结合结构域,特别是免疫球蛋白单可变结构域递送到细胞内。如果抗原是胞外的或是胞外结构域,则结合结构域可通过与该结构域结合来行使其功能,而不需要胞内递送。本文所述的本发明的结合结构域可以靶向目标GPCR: G蛋白的胞内构象表位。为了使用这些结合结构域作为细胞内的有效且安全的治疗剂,可以通过本领域已知的蛋白质转导或递送系统增强胞内递送。由于其跨过生物膜转位的能力,蛋白质转导结构域(PTD)已吸引了药物递送领域的普遍关注。PTD是相对较短的(11-35个氨基酸)序列,赋予与其缀合、复合或融合的蛋白质和其他大分子负荷这一明显的转位活性(Sawant和Torchilin2010)。例如,源自HIV的TAT肽(YGRKKRRQRRR)已经广泛用于各种活性剂的胞内递送,所述活性剂的范围覆盖了从小分子至蛋白质、肽、多种药物纳米载体和成像剂。可选的,还可以使用受体介导的胞吞机制进行胞内药物递送。例如,转铁蛋白受体介导的内化通路是有效的细胞摄入通路,已研发其用于位点特异性的递送药物和蛋白质(Qian等人,2002)。这是通过缀合转铁蛋白与治疗药物或蛋白质化学地实现,或者通过融合治疗肽或蛋白质到转铁蛋白的结构中遗传实现。由于天然存在的蛋白质(如铁结合蛋白转铁蛋白)是可生物降解的、无毒且无免疫原性的,因此这些蛋白质在药物靶向领域非常有效。此外,由于在细胞表面存在大量的其受体,因此可以实现位点特异性靶向。其他递送系统包括但不限于基于聚合物和基于脂质体的递送系统。

[0276] 可以使用本身已知的任何合适的体外测定、基于细胞的测定、体内测定和/或动物模型,或其任意的组合,测试本发明的结合结构域以及包含所述结构域的复合物的功效,这取决于所涉及的特定疾病或病症。

[0277] 筛选、选择和生产结合结构域

[0278] 在另一个方面,本发明还涵盖了筛选针对和/或特异性结合包含GPCR和G蛋白的复合物的结合结构域的方法,方法包括步骤:

[0279] a) 提供多个结合结构域,和

[0280] b) 在所述多个结合结构域中,筛选与包含GPCR和G蛋白的复合物结合的结合结构域,和

[0281] c) 分离与复合物结合的结合结构域。

[0282] 在本发明这一方面的优选的实施方案中,生成并就其特异性结合包含GPCR和G蛋白,以及任选受体配体在内的复合物来筛选结合结构域。还可以生成结合结构域和筛选其对G蛋白的特异性结合。如本文所述,可以多种方式生成结合结构域。在免疫球蛋白单可变

结构域(如纳米抗体)的情况下,通常用包含与G蛋白和受体配体结合的GPCR的靶复合物免疫动物,如上文所述(例如V_HH序列,作为非限制性例子)及本文别处示例的。

[0283] 为了用靶复合物免疫动物,可以使用在宿主细胞中表达所述蛋白质的重组形态的常规方法生产和纯化靶复合物(即,GPCR和G蛋白)的蛋白质,并使用亲和力层析和/或基于抗体的方法纯化蛋白质。在特定的实施方案中,可应用杆状病毒/Sf-9系统表达,但也可以使用其他表达系统(例如,细菌、酵母或哺乳动物细胞系统)。表达和纯化GPCR的示例性方法描述在例如Kobilka (1995), Eroglu等人 (2002), Chelikani等人 (2006) 和书籍“Identification and Expression of G Protein-Coupled Receptors”, (Kevin R. Lynch (编著), 1998) 等中。还可以使用纯化的受体(例如, $\beta 2$ -AR或MOR)(用化学当量过量的G蛋白(Gs或Gi)重构为重组HDL颗粒), 来重构功能性GPCR:G蛋白复合物,如Whorton等人, (2009) 所述(关于 $\beta 2$ -AR:Gs)或Kuszk等人, (2009) 所述(关于MOR:Gi)。还可以在磷脂小泡中重构GPCR,并将其负载化学当量过量的G蛋白。类似的,用于在磷脂小泡中重构活性GPCR的方法是已知的,描述在:Luca等人 (2003), Mansoor等人 (2006), Niu等人, (2005), Shimada等人, (2002) 和Eroglu等人, (2003) 等中。在某些情况下,可以高密度的重构GPCR和磷脂(例如,每mg磷脂1mg受体)。在许多情况下,GPCR可以以两种方向位于磷脂小泡中(正常的方向,和“颠倒的”方向,后者的胞内环位于小泡外侧)。GPCR的其他免疫方法包括但不限于使用表达GPCR和/或G蛋白的完整细胞或来自其的膜。

[0284] 在特定的实施方案中,用于双功能交联剂交联的靶复合物免疫动物(还参见实施例章节)。可以使用本领域技术人员普遍已知的标准技术进行化学交联(参见例如, Hermanson, G.T. (2008) Bioconjugate Techniques, 第2版, Elsevier Inc., 第1202页)。

[0285] 可以使用本领域普遍已知的任何适合产生免疫应答的技术,免疫任何合适的动物,例如暖血动物,特别是哺乳动物,如兔、小鼠、大鼠、骆驼、绵羊、牛或猪,或鸟类,如鸡或火鸡。

[0286] 可以例如通过筛选在表面表达结合结构域的细胞(例如,从适当免疫的骆驼获得的B细胞)的组合、集合或文库,通过筛选结合结构域的(原始或免疫)文库,或者通过筛选编码结合结构域的氨基酸序列的核酸序列的(原始的或免疫)文库,实施对特异性结合靶复合物的构象表位的结合结构域的筛选,所述筛选都可以本身已知的方式实施,且所述方法还可以任选的包括一个或多个其他合适的步骤,例如但不限于亲和力成熟步骤、表达理想的氨基酸序列的步骤、筛选针对理想抗原结合和/或活性的步骤、确定理想的氨基酸序列或核苷酸序列的步骤、导入一个或多个个人源化取代的步骤、以合适的多价和/或多特异性模式格式化的步骤、筛选理想的生物学和/或生理学特性的步骤(即,使用本领域已知的合适的测定),和/或以任何合适的顺序的一个或多个这类步骤的组合。

[0287] 本发明的另一个方面涉及包含根据本发明的结合结构域的试剂盒。试剂盒还可包括试剂,如缓冲剂、分子标签、载体构建体、参照样品材料以及合适的固体支持物、细胞、核酸等,的组合。这类试剂盒可用于本文所述的本发明的任何应用。例如,试剂盒可包括用于化合物筛选用途的测试化合物(的文库)。

[0288] 最后,本发明的最后一个方面是根据本发明的任何结合结构域的用途,用于分离负责特异性结合GPCR:G蛋白复合物的构象表位的氨基酸序列和构建基于所述氨基酸序列的人工结合结构域。可理解,在根据本发明的结合结构域中,框架区和互补决定区是已知

的,对于结合GPCR:G蛋白复合物中相同构象表位的结合结构域衍生物的研究将允许获得参与结合所述构象表位的关键氨基酸。这一知识可用于构建最小的结合结构域和生成其衍生物,这是可以通过本领域技术人员已知的技术常规实行的。

[0289] 下列实施例意在促进对本发明的进一步理解。本文中参考了示例的实施方案描述本发明,但应理解,本发明不限于此。具有本领域常规技能和接受本文教导的人将认识到在本发明范围内的其他修饰和实施方案。因此,本发明仅受本文所附权利要求的限制。

实施例

[0290] 实施例1:形成和纯化稳定的激动剂- β 2AR-Gs三元复合物

[0291] 通过在2ml缓冲液(10mM HEPES,pH7.5,100mM NaCl,0.1%DDM,1mM EDTA,3mM $MgCl_2$,10 μ M BI-167107)中混合约100 μ M浓度的Gs异源三聚体与过量摩尔(约130 μ M)的结合了BI-167107的T4L- β 2AR(或 β 2AR-365),并在室温孵育3hr,形成稳定的复合物(参见图2)。从筛选和表征约50种不同的 β 2AR激动剂中鉴别出的BI-167107具有约30hr的解离半衰期,为结合了活性G蛋白的受体提供了比其他的完全激动剂(如异丙肾上腺素)更稳定作用(Rasmussen等人,2011)。为了维持复合物的不合核苷酸的高亲和力状态,在90min后加入腺苷三磷酸双磷酸酶(25mU/ml,NEB),水解结合受体时从Gs释放的 剩余GDP。腺苷三磷酸双磷酸酶水解GDP得到的GMP对复合物中的G蛋白具有非常低的亲和力。重新结合GDP可导致R:G复合物解离(图3a)。

[0292] 4℃下,DDM中的R:G复合物在48小时后表现出显著的解离(图4a)。筛选和表征了超过50种的两亲物(数据未显示),并且鉴别出MNG-3(NG-310,Affymetrix-Anatrace;Chae等人,2011)及其密切相关的类似物作为基本上稳定复合物的去垢剂(图4a,b)。通过在室温下将R:G混合物(2ml)添加到含有1%MNG-3的8ml缓冲液(20mM HEPES,pH7.5,100mM NaCl,10 μ M BI-167107)中1hr,将复合物交换入MNG-3中。

[0293] 在该阶段,混合物含有R:G复合物、非功能性Gs和过量的 β 2AR。为了从非功能性Gs中分离功能性R:G复合物和完成去垢剂交换,将R:G复合物固定在M1Flag树脂上,并在含有0.2%MNG-3的缓冲液(20mMHEPES,pH7.5,100mM NaCl,10 μ M BI-167107和3mM $CaCl_2$)中洗涤。为了阻止半胱氨酸键桥介导的R:G复合物聚集,在用50kDa MWCMillipore浓缩器浓缩前,向洗脱的蛋白质中加入100 μ M TCEP。在用含有0.02%MNG-3,10mM HEPES pH7.5,100mM NaCl,10 μ M BI-167107和100 μ M TCEP的缓冲液平衡的Superdex20010/300GL柱(GE Healthcare)上实施了最终的尺寸排阻层析步骤,以从R:G复合物中分离了过量的游离受体(图5b)。混合峰级分(图5b),将所述级分用100kDa MWC Viva-旋转浓缩器浓缩至约90mg ml^{-1} ,并通过SDS-PAGE/考马斯亮蓝染色(图5a)和凝胶过滤(图5c)分析。为了验证是纯的、均质的和去磷酸化的制品,通过离子交换层析常规分析R:G复合物(图5d)。

[0294] 实施例2:生成与激动剂: β 2AR:Gs三元复合物结合的纳米抗体

[0295] 从负染EM成像(数据未显示)中,观察到了Gs的 α 螺旋状结构域是柔性的。通过生成与激动剂- β 2AR-Gs三元复合物结合的纳米抗体解决该结构域的靶向稳定化。纳米抗体是单结构域抗体,源自家羊驼的仅含重链的抗体(Muyldermans,2001)。为了鉴别结合(负载激动剂的)受体偶 联型Gs-蛋白的纳米抗体,用二(磺基琥珀酸亚酰胺)戊二酸酯(BS2G,Pierce)交联的 β 2AR:Gs:BI167107三元复合物免疫了2只家羊驼(Llama glama)。两只动物

都用每隔一周50-100 μ g一针免疫四次。在结束免疫后,从免疫动物中分离外周血淋巴细胞,提取总RNA,制备cDNA。从约107淋巴细胞中分离总RNA,如Chomczynski和Sacchi (1987)所述。根据生产商 (Invitrogen) 说明,使用dN6引物和superscript RT,制备第一链的cDNA合成。如前所述,使用特异性引物,通过PCR从该cDNA中扩增编码VHH基因的片段 (Conrad等人, 2001)。使用巢式PCR,分别在VHH可读框的起始和末端改造PstI和BstEII。将VHH作为PstI-BstEII片段克隆到噬菌体展示载体pMESy4中。对于每只家羊驼,都构建独立的噬菌体展示文库,其容纳相应的纳米抗体库作为geneIII融合物 (Domanska等人, 2011)。通过对i) 包埋在ApoL生物素化高密度脂蛋白颗粒 (rHDL, Whorton等人, 2007) 中的 β 2AR:Gs:BI167107三元复合物或ii) BS2G交联型 β 2AR:Gs:BI167107三元复合物进行两轮生物淘选,富集R:G复合物特异性纳米抗体。对于第一轮生物淘选对策,在20mM Hepes (pH8.0), 100mM NaCl, 1mM EDTA, 100 μ M TCEP和100nM BI167107中,将含有 β 2AR:Gs:BI167107三元复合物的生物素化rHDL颗粒以1 μ g/孔固定在包被了亲和素 (neutravidin) 的Maxisorp平板 (Nunc) 上。对于第二轮生物淘选对策,将BS2G交联型 β 2AR:Gs:BI167107三元复合物以1 μ g/孔固相包被在Maxisorp平板上。对于每轮生物淘选,加入 10^{11} 个噬菌体固定抗原,并孵育1至2小时。之后,从含抗原的孔中去除未结合的噬菌体,用20mM HEPES, 100mM NaCl, pH8洗涤孔14次,最后用200 μ L的20mM Hepes (pH8.0), 100mM NaCl, 1mM EDTA, 100 μ M TCEP和100nM BI167107孵育10分钟,去除非特异性噬菌体。为了洗脱复合物特异性的噬菌体,用胰蛋白酶处理孔,将噬菌体重新覆盖并用于感染对数生长的TG1细胞 ($OD_{600} \pm 0.5$)。从每个富集文库中随机挑选48个集落,生长在1ml含氨苄青霉素和葡萄糖的2xTY中。用IPTG诱导培养物,诱导纳米抗体表达,制备含有部分纯化的纳米抗体的周质提取物。通过ELISA分析包含 在这些周质提取物中的纳米抗体与激动剂: β 2AR:Gs三元复合物的结合。

[0296] 通过对固定在亲和素包被的Maxisorb平板上的下述复合物和空白rHDL颗粒之间的比较ELISA,分析在含有 β 2AR:Gs:BI167107三元复合物的生物素化rHDL颗粒上富集的纳米抗体。通过对相同固相包被的复合物与未包被的孔之间的比较ELISA,分析在固相包被的BS2G交联型 β 2AR:Gs:BI167107三元复合物上富集的纳米抗体。从比较ELISA的打分阳性的集落中制备单克隆,提取DNA,并使用常规方法分析编码的纳米抗体基因的序列 (表2-3显示了氨基酸序列)。对于Nb35、Nb36和Nb37,通过分析性凝胶过滤进一步验证与 β 2AR:Gs:BI167107三元复合物的结合 (图3d、3e、3f、3g)。

[0297] 实施例3:Nb35、Nb36和Nb37结合Gs,并通过GTP γ S阻止复合物解离

[0298] 为了确定针对 β 2AR:Gs:BI167107三元复合物产生的纳米抗体 (表2) 是否结合受体或Gs,之后在ELISA中监控了这些纳米抗体与单独的纯化受体的结合。在用结合了激动剂的 β 2AR-356包被的固相 (Maxisorb, Nunc) 中,表2的所有纳米抗体被打分为负,所述结合了激动剂的 β 2AR-356高密度的重构成磷脂小泡 (Rasmussen等人, 2011)。在该ELISA中, β 2AR特异性纳米抗体——Nb80 (Rasmussen等人, 2011) 打分阳性。表2中描述的 β 2AR:Gs:BI167107结合物无一结合重构的单独受体,提示其结合Gs中包含的表位。尺寸排阻层析显示, Nb35和Nb37结合位于Gs异源三聚体上的分离的表位,形成R:G:Nb35:Nb37复合物 (图3d)。类似的, Nb36和Nb37结合位于Gs异源三聚体上的分离的表位,形成R:G:Nb36:Nb37复合物 (图3e)。

[0299] GDP、GTP和不能水解的GTP类似物破坏 β 2AR:Gs复合物 (图3a), 在体外和体内导致GPCR G蛋白复合物解离。在存在和不存在GTP γ S的条件下,通过分析级的尺寸排阻层析,分

析Nb35和不能水解的GTP类似物GTP γ S对激动剂: β_2 AR:Gs三元复合物的完整性的相互影响。发现纳米抗体35、36和37保护 β_2 AR:Gs:BI167107复合物免于被GTP γ S解离(图 3d、3e&3g)。

[0300] 实施例4. 纳米抗体辅助的 β_2 AR-Gs复合物结晶

[0301] G蛋白偶联受体(GPCR)负责了对激素和神经递质,以及视觉、嗅觉和味觉的大部分细胞应答。GPCR信号传递的范例是通过激动剂占据的受体活化异源三聚体GTP结合蛋白(G蛋白)。在理解GPCR信号传递的结构基础的尝试中,结晶了 β_2 AR-Gs复合物,通过X射线晶体学解析了它的结构。

[0302] 晶体形成的一个难题是在去垢剂溶液中制备稳定的 β_2 AR:Gs复合物。 β_2 AR和Gs在脂质双层中有效偶联,但不在用于溶解和纯化这些蛋白质的去垢剂中偶联(实施例1)。发现了可以通过在十二烷基麦芽糖苷溶液中混合纯化的GDP-Gs(约100 μ M终浓度)与摩尔过量的结合了纯化 β_2 AR的高亲和力激动剂(BI167107;Rasmussen等人,2011),制备相对稳定的 β_2 AR:Gs复合物。可以添加一种非选择性的嘌呤核苷三磷酸双磷酸酶,水解从与 β_2 AR形成复合物的Gs上释放的GDP。之后,通过连续的抗体亲和力层析和尺寸排阻层析纯化复合物。通过将复合物交换到新研发的麦芽糖新戊二醇去垢剂(NG-310,Anatrace)中,增强复合物的稳定性(Chae等人,2010)。该复合物可以在室温下孵育24小时,没有任何可注意到的降解;然而,使用稀疏矩阵筛选在去垢剂微团、bicelle和脂质立方相(LCP)中结晶复合物的初步尝试失败了。

[0303] 为了进一步评估复合物的品质,通过单粒电子显微镜(EM)分析蛋白质。结果验证了复合物是单分散的(数据未显示),但揭示了用于获得质量晶体衍射的其他可能的瓶颈。首先,用于稳定复合物的去垢剂形成了大的微团,导致在 β_2 AR:Gs复合物的胞外侧只留下极少的极性表面用于形成晶格接触。因此,用T4溶菌酶(T4L)替换了 β_2 AR的无结构的氨基末端。之前通过在跨膜区段(TM)5和6的细胞质末端之间插入T4L,使用T4L促进失活 β_2 AR的晶体发生(Rosenbaum等人,2007)。该融合蛋白(T4L- β_2 AR)表现出正常的配体结合和Gs偶联特性。在LCP中,使用修饰的单油酸酯(monolein)进行结晶化试验(7.7MAG,由Martin Caffrey提供),所述单油酸酯设计为与T4L- β_2 AR:Gs复合物的大的亲水性组分相容(Misquitta等人,2004)。虽然我们能够获得衍射7Å的小晶体,但却不能通过使用添加剂和其他修饰来改善其质量。

[0304] 单粒EM分析揭示的另一个晶体生成的潜在难题是放置Gas亚基的 α -螺旋状组分的位置时增加的可变性。Gas由2个结构域组成,ras样GTP酶结构域(GasRas),其与 β_2 AR和G β 亚基相互作用,和 α -螺旋状结构域(GasAH)(Sprang等人,1997)。两个Gas亚结构域的界面形成了结合核苷酸的口袋(图1),而EM2D平均值和3D重构提示了在缺少鸟嘌呤核苷酸条件下,GasAH相对于T4L- β_2 AR-GasRAS-G β γ 的复合物具有可变的位置(图1b)。

[0305] 在进一步促进复合物的晶体生成的工作中,我们尝试了共结晶复合物和纳米抗体35.Nb35保护复合物免于被GTP γ S解离,提示稳定了Gs:Nb相互作用(图3a)。按1:1.2的摩尔比,混合结合了BI-167107的T4L- β_2 AR:Gs复合物和Nb35(参见图6和7)。通过分析凝胶过滤验证了略微摩尔过量的Nb35(图7b)。利用之前报道过的双体注射器(twin-syringe)混合法(Caffrey2009),在按1:1蛋白质比脂类的比例(w/w)与含有10%胆固醇(C8667,Sigma)的7.7MAG(由Martin Caffrey提供)混合前,室温孵育混合物1hr。 β_2 AR:Gs:Nb复合物在7.7MAG中的浓度是约25mg/ml。通过LCP分散机器人(Gryphon,Art Robbins Instruments),按每滴

40n1,向24孔或96孔玻璃夹心平板中递送蛋白质:脂类混合物,并用0.8 μ l沉淀溶液覆盖全体。使用部分基于StockOptions Salt试剂盒试剂(Hampton Research)的内部筛选,初步鉴别了多个结晶前导物。在18至22%PEG400,100mM MES pH6.5(图1c),350至450mM硝酸钾,10mM膦甲酸(图3b),1mM TCEP和10 μ M BI167107中,生长用于收集数据的晶体。晶体在20℃下3-4天内达到实际大小(图8),并将其从海绵样中间相中挑取,并在不含额外的冷冻保护剂的条件下,在液氮中瞬间冰冻。

[0306] 实施例5:Nb35促进R:G复合物的晶体形成

[0307] BI-167107结合在空间群P2₁中结晶的T4L- β 2AR:Gs:Nb35复合物,每个不对称单元中具有单个复合物。图9a显示了晶体学堆积的相互作用。图9b显示了包括T4L和Nb35在内的完整复合物的结构,图9c显示了 β 2AR:Gs复合物自身。

[0308] 在备选的含水层和脂质层中排列结合了BI-167107的T4L- β 2AR:Gs:Nb35复合物,具有几乎只在复合物的可溶组分之间形成的晶格接触,留下受体分子悬浮在G蛋白层之间,在膜的平面上彼此广泛分开。在所有的可溶蛋白质之间形成了广泛的晶格接触,可能导致强烈的整体衍射和G蛋白的非常清楚的电子密度。

[0309] Nb35和T4L促进了结合BI-167107的T4L- β 2AR:Gs:Nb35复合物的晶体形成。Nb35结合Gs上的构象表位,压在G β 和G α 亚基与互补决定区(CDR,根据IMTG编号定义;Lefranc, 2003) 1的界面上,所述互补决定区1主要与G β 相互作用(图10a),而长的CDR3环同时与G β 和G α 亚基相互作用(图10b)。Nb35的一些框架区还与相同复合物的G α 相互作用(图10c)。来自一个复合物的其他框架区与来自两个相邻复合物的G α 亚基相互作用(图11),主要负责晶格内的晶体接触。T4L与受体的氨基末端形成相对稀疏的相互作用,但压在一个复合物的G β 亚基的氨基末端,另一个复合物的G γ 亚基与另一个复合物的G α 亚基的羧基末端。

[0310] 实施例6.活性状态 β 2AR的结构

[0311] β 2AR:Gs结构为GPCR的跨胞质膜信号转导机制和三元复合物的功能特性的结构基础提供了第一个高分辨率的透视。图12a比较了在 β 2AR:Gs复合物中和在结合了失活卡拉洛尔的 β 2AR中结合了激动剂的受体的结构。失活和活性结构之间的最大差异是在E268的Ca碳处测量时, TM6向外移动了**14 Å**。TM5螺旋的胞质末端向外移动和延伸更少,少7个残基。在第3个胞内环(ICL3)中,26个氨基酸的区段是乱序的。失活和活性结构之间的另一个显著差异是第2个胞内环(ICL2),其在失活的 β 2AR结构中形成延伸的环,而在 β 2AR:Gs复合物中形成 α -螺旋。在 β 2AR-Nb80结构中也观察到了这一螺旋(图12b);然而,这可能不是活性状态特有的特征,因为在高度均质的鸟类 β 1AR的失活结构中也观察到了这一结构(Warne等人, 2008)。

[0312] 在该 β 2AR-GsRas界面上, β 2AR的电子密度图的质量是最高的,而胞外部分则弱得多,可能是由于缺少与胞外表面的晶格接触(图9a)。结果是,不能确定地构建在结合配体的口袋中的高亲和力激动剂(BI-167107)的模型。然而,T4L- β 2AR:Gs复合物中的 β 2AR的整体结构与G蛋白模拟的纳米抗体(Nb80)稳定的 β 2AR的我们最近的活性状态结构非常相似。这些结构主要背离了在TM5和6的胞质末端(图12b),可能是由于存在取代了 β 2AR-Nb80结构中的ICL3的T4L。不论如何, β 2AR-Nb80复合物对激动剂异丙肾上腺素表现出与 β 2AR:Gs复合物(Rasmussen等人, 2011)同样高的亲和力,与配体结合口袋周围的高结构同源性一致。 β 2AR-Nb80晶体的电子密度图为配体结合口袋周围和配体结合口袋与Gs偶联界面之间的氨基酸

构象重排提供了更可靠的观察(Rasmussen等人,2011)。

[0313] 图12c显示了与 β_2 AR-Nb80复合物相比, β_2 AR:Gs复合物中的高保守序列基序的位置,所述基序包括D/ERY和NPxxY(还参见图13)。已建议这些保守序列对激活或维持受体处于活性失活是重要的(Hofmann等人,2009)。这些氨基酸的位置在这两个结构中基本相同,证实Nb80是非常好的G蛋白代替品。这两个结构间仅Arg131不同。在 β_2 AR-Nb80结构中,Arg131与Nb80相互作用,而在 β_2 AR:Gs结构中,Arg131压在Gs的Tyr391上(图13)。

[0314] 通过与(GasRas)广泛的相互作用,稳定了 β_2 AR的活性状态(图14)。与G β 或G γ 亚基没有任何直接的相互作用。 β_2 AR-GasRas界面的总包埋表面是**2576 Å²**(GasRas为**1300 Å²**, β_2 AR为**1276 Å²**)。该界面是由 β_2 AR的ICL2、TM5和TM6、以及由GasRas的 α 5-螺旋、 α N- β 1连接、 β 3链的顶部和 α 4-螺旋形成的(参见表6关于特异性相互作用)。涉及该相互作用的 β_2 AR序列表现出在G蛋白偶联中发挥作用;然而,当比对这些区段和其他GPCR时,没有用于Gs-偶联特异性的明显的共同序列。这可能不令人惊讶,考虑到 β_2 AR还偶联Gi,而许多GPCR偶联一个以上的G蛋白同种型。因此,G蛋白偶联特异性的结构基础必须涉及更细微的二级和三级结构特征。然而,值得注意的相互作用涉及Phe139,其位于ICL2螺旋的起始处,坐落于由 β 1链起始处的Gas His41、 β 3链起始处的Val1213和 α 5-螺旋中的Phe376、Arg380和Ile383形成的疏水性口袋中(图14c)。 β_2 AR突变体F139A展现了与Gs的偶联严重受损(Moro等人,1993)。在几乎所有的Gs偶联受体上,与Phe139对应的残基是Phe或Leu,但在已知与其他G蛋白偶联的GPCR中所述残基更多变。令人感兴趣的是,保守DRY序列的Asp130与ICL2螺旋中央的Tyr141之间的相互作用,稳定ICL2螺旋(图14c)。Tyr141表现出是胰岛素受体酪氨酸激酶的底物(Baltensperger等人,1996);然而,该磷酸化的功能性重要性目前未知。

[0315] 实施例7.活化的Gs的结构

[0316] β_2 AR:Gs复合物中最令人惊讶的观察结果是GasAH相对于GasRas的大位移(围绕结构域之间的连接约180°旋转)(图15a)。在Gs的晶体结构中,由GasRas和GasAH之间的界面形成结合核苷酸的口袋。结合鸟嘌呤核苷酸稳定了这两个结构域之间的相互作用。失去该结合鸟嘌呤核苷酸的稳定效应与在去垢剂溶解的复合物的单粒EM分析中观察到的GasAH的高柔性一致(数据未显示)。这也和复合物形成时,这两个结构域之间的界面上的氘交换增加一致(数据未显示)。最近,Hamm、Hubbell及其同事利用双电子电子共振(DEER)光谱,证实了在与光活化的视紫红质形成复合物后,位于Ras和Gi的 α -螺旋状结构域上的硝基氧探针之间的距离大改变(多达**20 Å**)(Van Eps2011)。因此,可能不需要惊讶的是,替换的是GasAH而非GasRas;然而,其在该晶体结构中的位置可能反映了对晶体堆积的相互作用而非对其生理学构象的影响。

[0317] 在 β_2 AR和核苷酸结合口袋之间的构象关联主要涉及Gs的氨基和羧基末端的螺旋(图14)。图15b关注了比较Gs- β_2 AR复合物与Gas-GTP γ S复合物的GasRAS结构时,经历最大构象改变的GasRAS区域(Sunahara 等人,1997)。观察到的最大差异是 α 5螺旋,其向受体移动了**6 Å**并转动至羧基末端突入 β_2 AR的跨膜核心中。与该移动相关,与Gas-GTP γ S结构中的鸟嘌呤环相互作用的 β 6- α 5环朝向外侧,偏离了核苷酸结合口袋(图15b-d)。 α 5螺旋的移动也与此螺旋和 β 6片层、 α N- β 1环和 α 1螺旋之间的相互作用的改变相关。 β 1链形成了 β_2 AR和核苷酸结合口袋之间的另一连接。该链的C-末端改变了Gly47周围的构象,在配合GTP结合形

态中的 γ -磷酸的 $\beta 1$ - $\alpha 1$ 环(P环)中还有其他改变(图15b-d)。晶体结构的观察结果与氘交换实验一致,其中,在形成不含核苷酸的 $\beta 2AR:Gs$ 复合物后,在 $\beta 1$ 片层中和 $\alpha 5$ 螺旋的氨基末端中存在增强的氘交换(数据未显示)。

[0318] 尚未确定Gs异源三聚体的结构,因此不能直接比较 $\beta 2AR:Gs$ 复合物形成前后的 $G\alpha s-G\beta \gamma$ 界面。基于结合GDP的Gi异源三聚体的结构(Wall等人,1995),在与 $\beta 2AR$ 形成复合物后,没有观察到 $G\alpha sRAS$ 和 $G\beta \gamma$ 之间相互作用的大的改变。这也与氘交换研究一致(数据未显示)。应注意的是,Nb35结合在 $G\alpha sRAS$ 和 $G\beta$ 之间的界面(图2b)。因此,不能排除Nb35可影响晶体结构中的 $G\alpha sRAS-G\beta \gamma$ 界面的相对方向的可能性。然而,单粒EM研究提供了Nb35不破坏 $G\alpha sAH$ 和 $G\alpha sRAS$ 之间的相互作用的证据(数据未显示)。

[0319] 实施例8:Nb35和Nb37结合Gs上的不同表位并抑制核苷酸结合

[0320] 为了研究纳米抗体(Nb35和Nb37)对Gs本身的效应,在20mM Tris-HCl, pH8.0, 3mM $MgCl_2$, 1mM DTT中添加纳米抗体和bodipy-GTP γ S-FL和各种Gs蛋白制品至终体积200 μ L。含有异源三聚体Gs蛋白的样品还包括0.1% DDM。Bodipy-GTP γ S-FL是稳定的荧光GTP类似物($\lambda_{ex} \approx 470nm$, $\lambda_{em} \approx 515nm$)。其荧光强度在结合G蛋白后增加,因此,Bodipy-GTP γ S-FL可用于实时测量核苷酸与G蛋白的结合(McEwen等人,2001)。在M5荧光平板读数器(Molecular Precision)上,以96孔微滴度板模式测量荧光。

[0321] 在第一个实验中(图16),孵育量增加的Nb37与1 μ M纯化的 $G\alpha s$ 和100nM Bodipy-GTP γ S-FL,短时间(300秒)的测量荧光增加,使水解产物bodipy-磷酸的积累最小化(Jameson等人,2005)。如前所述,纯化异源三聚体Gs蛋白的 $G\alpha s$ 亚基(Sunahara等人,1997)。根据该实验,Nb37似乎以剂量依赖性的方式阻断GTP γ S结合单独的 $G\alpha s$ 。这些结果还提示Nb37的结合表位限于异源三聚体Gs蛋白的 $G\alpha s$ 亚基。在类似的实验中(图17),孵育量增加的Nb35与1 μ M纯化的 $G\alpha s$ 和100nM Bodipy-GTP γ S-FL。与观察结果一致的是,Nb35结合包含 $G\alpha sRAS$ 和 $G\beta$ 元件的表位(参见实施例7),Nb35对GTP γ S与单独的 $G\alpha s$ 亚基结合没有影响。

[0322] 在另一个实验中(图18),孵育量增加的Nb35与1 μ M纯化的 $G\alpha s\beta \gamma$ 异源三聚体和100nM Bodipy-GTP γ S-FL。该实验提示,Nb35以剂量依赖性的方式阻断GTP γ S与游离的 $G\alpha s\beta \gamma$ 异源三聚体结合。

[0323] 实施例9.Nb35稳定其他的激动剂-GPCR-Gs复合物

[0324] $G\alpha s$ 亚基(或Gs蛋白)是通过激活腺苷酸环化酶,激活cAMP-依赖性通路的异源三聚体G蛋白亚基。与Gs偶联的G蛋白偶联受体包括:5-HT受体5-HT4和5-HT7型、ACTH受体、腺苷受体A2a和A2b型、精氨酸血管加压素受体2、 β -肾上腺素能受体 $\beta 1$ 、 $\beta 2$ 和 $\beta 3$ 型、降钙素受体、降钙素基因相关肽受体、促肾上腺皮质激素释放激素受体、多巴胺受体D₁样家族(D₁和D₅)、FSH-受体、胃抑制型多肽受体、胰高血糖素受体、胰高血糖素样肽1受体(GLP1-R)、组胺H₂受体、黄体生成素/绒毛膜促性腺激素受体、黑皮质素受体、甲状旁腺激素受体1、前列腺素受体D₂和I₂型、分泌素受体、促甲状腺素受体等。

[0325] 为了确定结合 $\beta 2AR:Gs:BI167107$ 中的Gs的纳米抗体是否也稳定其他的GPCR:Gs:激动剂复合物,制备了与Gs、Nb35、精氨酸血管加压素(AVP:NT4LV2R:Gs)复合的精氨酸血管加压素受体2(登录号P30518;V2R_HUMAN)的复合物,并证实了该复合物在SEC中的稳定性。精氨酸血管加压素(AVP)也被称为血管加压素、精氨加压素或抗利尿素激素,是激活精氨酸血管加压素受体2的天然配体。

[0326] 通过在0.1ml缓冲液(10mM HEPES,pH7.5,100mM NaCl,0.1%DDM,1mM EDTA,3mM MgCl₂,10μM AVP)中混合约90μM浓度的His标记的Gs异源三聚体与结合AVP的NT4LV2R(90μM)和Nb35(100μM),并在室温孵育2小时,形成稳定的复合物(图19a)。之后,在两步连续的亲和力和纯化步骤中纯化AVP:NT4LV2R:Gs复合物。通过SDS-PAGE监控纯化(图19)。首先,在向反应混合物加入含300μl1%MNG的10mM HEPES,pH7.5,100mM NaCl缓冲液后,将复合物用于Ni-NTA柱。在用缓冲液充分洗涤后,在0.2%MNG,10mM HEPES,pH7.5,100mM NaCl AVP10μM和200mM咪唑中洗脱复合物。之后,将复合物用于FLAG-标签的亲和力柱,在含有0.01%MNG的相同缓冲液中充分洗涤,并用FLAG-肽洗脱。通过SDS-PAGE监控该过程(图19b),并且该过程显示可以相应的纯化含有NT4LVTR、GaS、Gβ、Gγ和Nb35的复合物。在10mM HEPES,pH7.5,100mM NaCl,0.01%MNG,1mMEDTA,3mM MgCl₂,1μM AVP中,通过superdex200柱上的SEC进一步纯化了该复合物。通过SDS-PAGE监控该过程(图19c),显示可以相应的纯化含有NT4LVTR、GaS、Gβ、Gγ和Nb35的单分散性复合物。

[0327] 为了验证AVP:NT4LV2R:Gs复合物的稳定性,在冰上孵育纯化的样品24hr,并将其重新用于SEC,验证它的单分散性特征及其MW(图20)。与预期相同,过量的拮抗剂SR121463(10μM)破坏了AVP:NT4LV2R:Gs复合物。

[0328] 实施例10.使用纳米抗体稳定的GPCR:G蛋白复合物,对激动剂或正变构调控剂的改善的筛选

[0329] 选择性稳定GPCR的非主要构象异构体的纳米抗体将允许更有效的筛选与该特定的低丰度构象异构体选择性相互作用的配体。此外,构象异构体选择性的纳米抗体还可用于暴露变构的或隐藏的可用于药物研发的位点,或反向遮蔽对药物筛选不理想的结合口袋。当特定的构象异构体是活性状态时,鉴别的配体具有高的可能性作为激动剂行动,该结果受Costanzi&Vilar(2011)所述的电脑模拟入坞(in silico docking)实验的支持。实际上,其结果提示,活化的结构对激动剂的鉴别比对拮抗剂的鉴别更有利,而失活的结构对受体阻断剂的鉴别比对激动剂的鉴别更有利。

[0330] 提供的证据是,通过结合Gas和Gβγ亚基的界面,Nb35稳定了在活化的β2AR和G蛋白之间的复合物。鉴别与β2AR的低丰度的活性状态的构象异构体选择性相互作用的配体的筛选测定的例子可以是放射性配体测定,其利用Nb35使β2AR群体更加移向其活性状态。可以按Seifert及其同事(1998)所述,以极少的修正来类似的执行这类放射性配体测定。本文描述了涉及β2AR作为选择的通过通过任何与Gas亚基相互作用的其他GPCR替换β2AR的靶的测定,允许执行类似的筛选方法,鉴别针对该特定GPCR的激动型配体。可以筛选小分子(化合物MW通常在250至1000Da之间)或者甚至基于片段(化合物Mw通常<250Da)的文库来鉴别候选激动剂。稳定非主要的构象异构体将显著增加片段文库筛选的性能,尤其是因为在基于片段的药物筛选中,初始命中通常具有低的效价/亲和力。Nb35将选择性的增加对这样的化合物的亲和力,所述化合物是对选择性的、可用于药物研发的构象异构体特异性的,因此对鉴别从头片段具有深远影响。

[0331] 在孵育缓冲液(50mM Hepes pH7.4,1mM CaCl₂,5mM MgCl₂,100mM NaCl和0.5%w/v BSA)中,将恰当量(通常10μg)的来自HEK293T细胞的人β2AR均质化膜提取物与Nb35或不相关的纳米抗体(不稳定活性GPCR构象)在30℃平行孵育1h,所述细胞含有锚定了G蛋白亚基的膜。纳米抗体是外源提供的,相比肾上腺素能受体是大量摩尔过量的(例如,≥1μM)。之

后,将结合了纳米抗体的膜加入到含有文库化合物和2nM的3H-二氢阿普洛尔(DHA)拮抗型放射性配体的96孔板中。用孵育缓冲液调节每个孔的总体积至100 μ l,在30℃进一步孵育反应混合物另外1小时。之后,使用预先浸泡在0.3%聚乙烯亚胺中的GF/C玻璃纤维96孔过滤板(Perkin Elmer),收获与膜结合的放射性配体。用冰冷的洗涤缓冲液(50mM Tris-HCl pH7.4)洗涤过滤板,在50℃干燥30分钟。在加入25 μ l闪烁液(MicroScintTM-0,Perkin Elmer)后,在Wallac MicroBeta TriLux闪烁计数器中测量放射活性(cpm)。在存在Nb35而不施用不相关纳米抗体时,cpm显著减少的那些文库化合物被认为是激动型配体。通过初步文库筛选鉴别的候选激动剂命中将以剂量应答的方式重新筛选。使用Graphpad Prism软件,计算每种候选激动剂在存在Nb35和不相关的纳米抗体的条件下的%放射性配体替换曲线的IC₅₀值。为了证实从头鉴别的化合物的有效的激动作用,在细胞的 β 2AR信号传递测定中评估这些化合物的剂量依赖性效应。这类测定的一个实例依赖于检测第二信使分子,如G α s介导的信号传递后的cAMP(例如,HitHunter cAMP测定技术,DiscoverX)。除了使用膜提取物和外源应用的Nb35外,可以在用Nb35作为胞内抗体共转染的、表达 β 2AR的细胞系(或来源膜)上实施放射性配体测定,从而使 β 2AR群体移向它的活性状态。可选的,可以使用重组G蛋白和 β 2AR通过Nb35稳定 β 2AR的活性状态。

[0332] 实施例的材料和方法

[0333] 表达和纯化 β 2AR,Gs异源三聚体和纳米抗体-35

[0334] 在用重组杆状病毒(BestBac,Expression Systems)感染的Sf-9昆虫细胞培养物中,表达在第365位截短的N-末端融合的T4溶菌酶- β 2AR构建体(T4L- β 2AR,下文详述),根据前述方法(Kobilka等人,1995),将构建体溶解在正十二烷基- β -D-麦芽糖苷(DDM)中(参见图6关于纯化概述)。大部分分析实验和氩交换实验使用在第365位残基后截短的 β 2AR构建体(β 2AR-365;SEQ ID NO:55)。M1Flag亲和层析(Sigma)作为最初的纯化步骤,再进行阿普洛尔-Sepharose层析,用于选择功能性受体。使用后续的M1Flag亲和层析步骤交换结合了受体的阿普洛尔与高亲和力激动剂BI-167107。洗脱结合了激动剂的受体,针对缓冲液(20mM HEPES pH7.5,100mM NaCl,0.1%DDM和10 μ M BI-167107)透析,用入磷酸酶(New England Biolabs)处理,用50kDa截留分子量(MWCO)的Millipore浓缩仪浓缩至约50mg ml⁻¹。在旋转浓缩之间,将 β 2AR-365构建体而非T4L- β 2AR用PNGaseF(New England Biolabs)处理,去除氨基末端N-连接糖基化。通过SDS-PAGE/考马斯亮蓝染色常规分析纯化的受体(参见图5a)。

[0335] 在生长于Insect Xpress无血清培养基(Lonza)中的High5昆虫细胞(Invitrogen)内表达牛G α s短(Bovine G α sshort)、与His₆标签融合的大鼠G β ₁和大鼠G γ ₂(参见表5)。培养物生长至150万格细胞/ml的密度,然后,用3种分开的苜蓿银纹夜蛾(Autographa californica)核型多角体病毒感染,每种病毒含有1种1:1复合感染的G蛋白亚基的基因(病毒由Dr.Alfred Gilman馈赠)。在孵育40-48小时后,通过离心收获感染的细胞,并重悬在每升培养体积75ml裂解缓冲液(50mM HEPES pH8.0,65mM NaCl,1.1mM MgCl₂,1mM EDTA,1x PTT(35 μ g/ml苯甲基磺酰氟(Phenylmethanesulfonyl Fluoride),32 μ g/ml甲苯磺酰基苯丙氨酰甲基氯酮(tosylphenylalanyl chloromethyl ketone),32 μ g/ml甲苯磺酰基赖氨酰甲基氯酮),1x LS(3.2 μ g/ml亮抑酶肽和3.2 μ g/ml大豆胰蛋白酶抑制剂)、5mM β -ME和10 μ M GDP)中。在氮空化贮罐(nitrogen cavitation bomb,Parr Instrument Company)中,用600psig

N₂对悬浮液加压40分钟。在降压后,将裂解物离心,去除核与未裂解的细胞,然后,在180,000x g超速离心40分钟。使用Dounce均质器,将沉淀的膜重悬在每升培养体积30ml洗涤缓冲液(50mM HEPES pH8.0,50mM NaCl,100μM MgCl₂,1x PTT,1x LS,5mMβ-ME,10μMGDP)中,并在180,000x g再次离心40分钟。将洗净的沉淀团块重悬在小体积的洗涤缓冲液中,并用液氮瞬间冰冻。

[0336] 融解冷冻的膜,用新鲜的洗涤缓冲液稀释至5mg/ml的总蛋白质浓度。向悬浮液中加入终浓度1.0%的胆酸钠去垢剂,加入终浓度5mM的MgCl₂,每升培养体积加入0.05mg纯化的蛋白质磷酸酶5(内部制备)。样品在冰上搅拌40分钟,然后在180,000x g离心40分钟,去除不可溶的残渣。用Ni-NTA上样缓冲液(20mM HEPES pH8.0,363mM NaCl,1.25mM MgCl₂,6.25mM咪唑,0.2%Anzergent3-12,1x PTT,1x LS,5mMβ-ME,10μM GDP)稀释上清液5倍,缓慢的仔细添加缓冲液,避免使胆酸浓度 过快下降至低于其临界微团浓度。每升培养体积加入3ml在Ni-NTA洗涤缓冲液1(20mM HEPES pH8.0,300mM NaCl,2mM MgCl₂,5mM咪唑,0.2%胆酸,0.15%Anzergent3-12,1x PTT,1x LS,5mMβ-ME,10μM GDP)中预平衡的Ni-NTA树脂(Qiagen),在冰上搅拌样品20分钟。将树脂收集到重力柱,并用4x柱体积的Ni-NTA洗涤缓冲液1、Ni-NTA洗涤缓冲液2(20mM HEPES pH8.0,50mM NaCl,1mM MgCl₂,10mM咪唑,0.15%Anzergent3-12,0.1%DDM,1x PTT,1x LS,5mMβ-ME,10μM GDP)和Ni-NTA洗涤缓冲液3(20mM HEPES pH8.0,50mM NaCl,1mM MgCl₂,5mM咪唑,0.1%DDM,1x PTT,1x LS,5mMβ-ME,10μM GDP)洗涤。用Ni-NTA洗脱缓冲液(20mM HEPES pH8.0,40mM NaCl,1mM MgCl₂,200mM咪唑,0.1%DDM,1x PTT,1x LS,5mMβ-ME,10μMGDP)洗脱蛋白质。混合含蛋白质的级分,并加入MnCl₂至终浓度100μM。每升培养体积加入50μg纯化的入蛋白磷酸酶(内部制备),在冰上孵育洗脱液并伴随搅拌30分钟。将洗脱液通过0.22μm滤膜,并直接上样到在MonoQ缓冲液A(20mM HEPES pH8.0,50mM NaCl,100μMMgCl₂,0.1%DDM,5mMβ-ME,1x PTT)中平衡的MonoQ HR16/10柱(GE Healthcare)上。用150ml缓冲液A,以5ml/min洗涤柱,用超过350ml、达28%的MonoQ缓冲液B(除了1M NaCl外,与缓冲液A相同)的线性梯度洗脱结合的蛋白质。将级分收集到沾有足够GDP的试管中,使终浓度为10μM。用具有10kDa NMWL再生的纤维素膜(Millipore)的搅拌型超滤池,将含Gs的级分浓缩至2ml。在用S200缓冲液(20mM HEPES pH8.0,100mM NaCl,1.1mM MgCl₂,1mM EDTA,0.012%DDM,100μM TCEP,2μM GDP)平衡的Superdex200prep grade XK16/70柱(GE Healthcare)上运行浓缩的样品。混合含有纯Gs的级分,加入甘油至10%终浓度,然后使用30kDa MWC0再生纤维素Amicon离心超滤装置,浓缩蛋白质至至少10mg/ml。然后,将浓缩的样品等分、瞬间冷冻,并储藏在-80°。来自8升细胞培养体积中的最终的、纯化的Gs异源三聚体的典型产量是6mg。

[0337] 根据前述方法,在大肠杆菌(*Escherichia coli*)菌株WK6的周质中表达纳米抗体-35(Nb35)(SEQ ID NO:1),提取,并通过镍亲和层析纯化(Rasmussen等人,2011),之后使用Mono S10/100GL柱(GE Healthcare)进行离子交换层析(图7a)。用缓冲液(10mM HEPES pH7.5,100mM NaCl)透析选定的Nb35级分,用10kDa MWC0 Millipore浓缩仪浓缩至约65mg ml⁻¹。

[0338] 蛋白质改造

[0339] 为了增加获得R:G复合物晶体的可能性,使用两种对策着手增加受体胞外表面的极性表面面积。方法是用之前使用的球状蛋白质T4溶菌酶(T4L)替换柔性和假设无结构的

N-末端,结晶并解析结合了卡拉洛尔的受体(Rosenbaum等人,2007)。本文使用的构建体(T4L- β 2AR)含有可切割的信号序列,后接M1Flag表位(DYKDDDDA;SEQ ID NO:70)、TEV蛋白酶识别序列(ENLYFQG;SEQ ID NO:71)、从N2至Y161的噬菌体T4溶菌酶(包含C54T和C97A突变),以及与人 β 2AR序列的D29至G365融合的2个剩余的丙氨酸接头(SEQ ID NO:69定义的T4L- β 2AR融合构建体)。将 β 2AR在N187的PNGaseF不可接触的糖基化位点突变为Glu。用Thr分别替换第一个胞外环中的M96和M98以增加另外的T4L- β 2AR的低表达水平。苏氨酸突变不影响配体结合 ^3H -二氢阿普洛尔的亲和力,但导致与异丙肾上腺素的亲和力少量减少,约2倍(数据未显示)。应注意,本文使用的野生型参照物 β 2AR由SEQ ID NO:72定义。

[0340] 显微晶体学数据收集和加工

[0341] 在Advanced Photon Source beamline23ID-B上实施数据收集。筛选了数百个晶体,使用自20个最强衍射晶体通常10度的衍射楔汇编最终的数据集。使用HKL2000(Otwinowski等人,1997)实施所有的数据整理。虽然在许多情况下,在初步框架中可见超过**3 Å**的衍射,但放射性破坏和各向异性衍射产生在较高分辨率的壳(shell)中的较低的完整性。由UCLA 各向异性衍射服务商(Strong等人,2006)对最终的数据集的分析提示:沿a*倒易轴的衍射比沿其他方向的衍射更好。基于沿着各倒易空间(reciprocal space)轴的F/sigF截止值为3,在用于精修之前,将反射进行各向异性截断,沿着a*、b*和c*的分辨率限为2.9、3.2和3.2埃。由于高分辨率壳的较低完整性,报道了该结构的整体分辨率仅为**3.3 Å**,但应该注意到,在精修和图计算的过程中,包括了一些达到**2.9 Å**的衍射数据。

[0342] 结构解析和精修

[0343] 使用Phaser(McCoy等人,2007a,b),通过分子替换解析结构。在解析结构中,发现分子替换研究的顺序是关键性的。在使用的顺序中,研究模型是:来自Gi异源三聚体G蛋白(PDB ID:1GP2)、Gs α ras结构域(PDB ID:1AZT)、活性状态 β 2肾上腺素能受体(PDB ID:3P0G)、 β 2结合纳米抗体(PDB ID:3P0G)、T4溶菌酶(PDB ID:2RH1)、Gs α 螺旋状结构域(PDB ID:1AZT)结构的 β 和 γ 亚基。在通过分子替换确定最初结构后,在Phenix(Afonine等人,2005)和BUSTER(Blanc等人,2004)中实施刚体精修和模拟退火,然后在Coot(Emsley等人,2004)中约束精修和手工重建。在迭代精修和手工调整后,使用DEN方法在CNS中精修结构。虽然该结构的分辨率超过了DEN通常最有效的分辨率,但若干解析较差区域的存在提示了整合用于指导精修的额外信息可以提供更好的结果。使用的DEN参照物模型是如上所述作为分子替换研究的模型,例外的是NB35,它是非常有序的且对其不能获得更高分辨率的结构。使用PyMOL(The PyMOL Molecular Graphics System,1.3版, **Schrödinger**,LLC.)制图。表7给出了精修的统计学。

[0344] 结合

[0345] 如前所述,从感染杆状病毒的Sf9细胞制备表达 β 2AR或 β 2AR-Gs肽融合物的膜,和实施 ^3H -双氢阿普洛尔(^3H -DHA)结合(Swaminath等人,2002)。对于竞争结合,在收获至GF/B滤膜之前,用 ^3H -DHA(最终1.1nM)和递增浓度的(-)-异丙肾上腺素(ISO)孵育膜1hr。以双位点结合模型拟合竞争数据,使用GraphPad棱镜计算ISO的高和低 K_i 和分数。

[0346] 纯化NT4LV2R

[0347] 使用杆状病毒系统(PfastBac),在Sf9细胞中表达N-末端融合的T4LV2R构建体

(NT4L-V2R; SEQ ID NO:73)。按 4×10^6 细胞/ml的密度感染细胞,并在27℃振荡培养瓶48h。收获后,在包含10mM Tris-HCl pH7.5,1mM EDTA,1μM Tolvaptan (Sigma)和2mg ml⁻¹用于阻断反应性半胱氨酸的碘乙酰胺的缓冲液中通过渗透压休克,裂解细胞。在包含0.5%十二烷基麦芽糖苷(DDM)、0.3%Cholate、0.03%胆固醇半琥珀酸酯(CHS)、20mM HEPES pH7.5,0.5M NaCl,30%v/v甘油、2mgml⁻¹碘乙酰胺和1μM Tolvaptan的溶解缓冲液中,用Dounce均质器从Sf9膜中提取NT4L-V2R。在离心后,向上清液中添加镍-NTA琼脂糖,搅拌2h,然后用100g旋转(spin)分批洗涤,每次5min,洗涤缓冲液为0.1%DDM,0.03%Cholate,0.01%CHS,20mM HEPES pH7.5和0.5M NaCl。将树脂倒入玻璃柱内,在补充了300mM咪唑的洗涤缓冲液中洗脱结合的受体。使用抗-Flag M1亲和力树脂进一步纯化NT4L-V2R,与激动剂AVP交换配体。将Ni-NTA树脂洗脱液上样到抗-Flag M1树脂上,在存在10μMAVP的条件下充分洗涤。然后,在存在1μM AVP的条件下,从抗-Flag M1亲和力树脂上用0.2mgml⁻¹Flag肽和2mM EDTA洗脱受体,并使用100kDa MWC0浓缩仪浓缩。

[0348] 表2. 纳米抗体列表

[0349]

纳米抗体 参考编号	纳米抗体 简写标注	SEQ ID NO	序列(包括C末端组氨酸标签和EPEA标签)
CA4435	Nb35	1	QVQLQESGGGLVQPGGSLRLSCAASGFTFSNY KMNWVRQAPGKGLEWVSDISQSGASISYTGSVK GRFTISRDNANTLYLQMNSLKPEDTAVYYCAR CPAPFTRDCFDVTSTTYAYRGQGTQVTVSSHHH HHHEPEA
CA4433	Nb33	2	QVQLQESGGGLVQPGGSLRLSCAASGFTFSNY VMNWVRQAPGKGLEWVSDISNGGGTTSYASSV KGRFTISRDNANTLYLQMNSLKPADTAVYYCA RCPAPFTNDCMDITSTTYAYRGQGTQVTVSSHH HHHEPEA
CA4436	Nb36	3	QVQLQESGGGSVQAGGSLRLSCTVSGTIFSVTV MGWYRQAPGKQRELVAGFTNTRNTNYVDSVKG RFTISKDSAKNTMYLQMNSLKPEDTAVYYCNVR RWGGTNWNDYWGGGTQVTVSSHHHHHHHEPEA
CA4437	Nb37	4	QVQLQESGGGFVQAGGSLRLSCAASGSIFSNT MAWFRQAPGKERELVAASPTGGSTAYKDSVKG RFTISRDSAKNTVLLQMNVLKPEDTAVYYCHLRQ NNRGSWFHYWGQGTQVTVSSHHHHHHHEPEA
CA4440	Nb40	5	QVQLQESGGGLVQAGGSLRLSCAVSGTIFDITP MGWYRQTPGKQREVVADLTSRGTTNYADSVKG RFTISRDNAAKMLYLQMNSLKSDDTGVYYCNVK RWGGIGWNDYWGGGTQVTVSSHHHHHHHEPEA
CA4441	Nb41	6	QVQLQESGGGLVQSGGSLRLSCVASGFRFSNF PMMWVRQAPGKGLEWVSLISIGGSTTNYADSVK GRFTISRDNANTLFLQMNSLKPEDTAVYYCAKY LGRLVPPTTEGQGTQVTVSSHHHHHHHEPEA

[0350]

表 3. 纳米抗体的 FR 和 CDR 的组合

纳米抗体参 考编号	纳米抗体简 写标注	FR1	CDR1	FR2	CDR2	FR3	CDR3	FR4
CA4435	Nb35	QVQLQESG GGLVQPGG SLRLSCAAS (SEQ ID NO 7)	GFTFSNYK (SEQ ID NO 13)	MINWVRQAPG KGLEWVSD (SEQ ID NO 19)	ISQSGASI (SEQ ID NO 25)	SYTGSVKGRFTI SRDNAKNTLYLQ MNSLKPEDTAVY YC (SEQ ID NO 31)	ARCPAPFTRDCFD VTSTTYAY (SEQ ID NO 37)	RQGQTQVTVSS (SEQ ID NO 43)
CA4433	Nb33	QVQLQESG GGLVQPGG SLRLSCAAS (SEQ ID NO 8)	GFTFSNYV (SEQ ID NO 14)	MINWVRQAPG KGLEWVSD (SEQ ID NO 20)	ISNGGGTT (SEQ ID NO 26)	SYASSVKGRFTI SRDNAKNTLYLQ MNSLKPEDTAV YYC (SEQ ID NO 32)	ARCPAPFTNDCM DITSTTYAY (SEQ ID NO 38)	RQGQTQVTVSS (SEQ ID NO 44)
CA4436	Nb36	QVQLQESG GGSVQAGG SLRLSCTVS (SEQ ID NO 9)	GTIFSVTV (SEQ ID NO 15)	MGWYRQAPG KQRELVAG (SEQ ID NO 21)	FTNTRNT (SEQ ID NO 27)	NYVDSVKGRFTI SKDSAKNTMYL QMNSLKPEDTA VYYC (SEQ ID NO 33)	NVRRWGGTNWN DY (SEQ ID NO 39)	WGQGTQVTVSS (SEQ ID NO 45)
CA4437	Nb37	QVQLQESG GGFVQAGG SLRLSCAAS (SEQ ID NO 10)	GSIFSNT (SEQ ID NO 16)	MAWFRQAPGK ERELVAA (SEQ ID NO 22)	SPTGGST (SEQ ID NO 28)	AYKDSVKGRFTI SRDSAKNTVLLQ MNSLKPEDTAVY YC (SEQ ID NO 34)	HLRQNNRGSWFH Y (SEQ ID NO 40)	WGQGTQVTVSS (SEQ ID NO 46)
CA4440	Nb40	QVQLQESG GGLVQAGG SLRLSCAVS (SEQ ID NO 11)	GTIFDITP (SEQ ID NO 17)	MGWYRQTPG KQREVVAD (SEQ ID NO 23)	LTSRGTT (SEQ ID NO 29)	NYADSVKGRFTI SRDNAKKMLYL QMNSLKSDDTG VYYC (SEQ ID NO 35)	NVKRWGGIGWND Y (SEQ ID NO 41)	WGQGTQVTVSS(S EQ ID NO 47)
CA4441	Nb41	QVQLQESG GGLVQSGG SLRLSCVAS (SEQ ID NO 12)	GFRFSNFP (SEQ ID NO 18)	MMWVRQAPG KGLEWVSL (SEQ ID NO 24)	ISIGGST (SEQ ID NO 30)	NYADSVKGRFTI SRDNAKNTLFLQ MNSLKPEDTAVY YC (SEQ ID NO 36)	AKYLGRLVPPTT (SEQ ID NO 42)	EGQGTQVTVSS(SE Q ID NO 48)

[0351] 表4. 纳米抗体的核酸序列

[0352]

纳米抗体 参考编号	纳米抗体 简写标注	SEQ ID NO	纳米抗体的核苷酸序列 (包括 His 标签和 EPEA 标签的核苷酸序列, 其为下划线的)
CA4435	Nb35	49	CAGGTGCAGCTGCAGGAGTCTGGAGGAGGCTTG GTGCAGCCTGGGGGGTCTCTGAGACTCTCCTGTG CGGCCTCTGGATTACCTTCAGCAATTATAAAATG AACTGGGTCCGCCAGGCTCCAGGAAAGGGGCTC GAGTGGGTCTCAGATATTTCTCAGAGTGGTGCTA GCATAAGTTACACAGGCTCCGTGAAGGGCCGATT CACCATCTCCAGAGACAACGCCAAGAACACGCTG TATCTACAAATGAACAGCCTGAAGCCTGAGGACA CGGCCGTCTATTACTGTGCCAGATGTCCGGCCCC ATTACAGAGAGATTGTTTTGACGTGACTAGTACCA CGTATGCCTACAGGGGCCAGGGGACCCAGGTCA CCGTCTCCTCACACCACCATCACCATCACGAACC <u>TGAAGCCTAG</u>
CA4433	Nb33	50	CAGGTGCAGCTGCAGGAGTCTGGAGGGGGCTTG GTGCAGCCTGGGGGGTCTCTGAGACTCTCCTGTG CAGCCTCTGGATTCACTTTCAGTAATATGTCATG AACTGGGTCCGCCAGGCTCCAGGAAAGGGGCTC GAGTGGGTCTCAGATATTTCTAATGGCGGTGGTA CCACAAGTTATGCAAGCTCCGTGAAGGGCCGATT CACCATCTCCAGAGACAACGCCAAGAACACGCTG TATCTGCAAATGAACGGCCTGAAGCCTGCGGACA CGGCCGTCTATTACTGTGCAAGATGTCCGGCCCC ATTACGAACGATTGTATGGACATAACTAGTACCA CGTATGCCTACAGGGGCCAGGGGACCCAGGTCA CCGTCTCCTCACACCACCATCACCATCACGAACC <u>TGAAGCCTAG</u>
CA4436	Nb36	51	CAGGTGCAGCTGCAGGAGTCTGGAGGAGGCTCG GTGCAGGCTGGGGGGTCTCTGAGACTCTCCTGTA CAGTCTCTGGAACCATCTTCAGTGTCAGTGTGTCATG GGCTGGTACCGCCAGGCTCCAGGGAAGCAGCGC GAGTTGGTCCGAGGTTTTACTAATACTAGAAACAC AACTATGTAGACTCCGTGAAGGGCCGCTTCACC ATCTCCAAAGACAGCGCCAAGAACACGATGTATC TACAAATGAACAGCCTGAAACCTGAGGACACAGC CGTCTATTACTGTAATGTACGTCCGTGGGGCGGT ACGAATTGGAATGACTACTGGGGCCAGGGGAACC AGGTCAACCGTCTCCTCACACCACCATCACCATCA <u>CGAACCTGAAGCCTAG</u>
CA4437	Nb37	52	CAGGTGCAGCTGCAGGAGTCTGGAGGGGGCTTC GTGCAGGCTGGGGGGTCTCTGAGACTCTCCTGTG CAGCCTCTGGAAGCATCTTCAGTAAGAATACCATG GCCTGGTCCGCCAGGCTCCAGGGAAGGAGCGA GAGTTGGTCCGAGCTAGTCTACGGGTGGTAGCA CAGCGTATAAAGACTCCGTGAAGGGCCGATTAC CATCTCCAGAGACAGCGCCAAGAACACGCTGTTG CTGCAAATGAACGTCTGAAACCTGAGGATACTG CCGTCTATTACTGTCATCTACGTCAAAATAACCGT GGTTCTTGGTTCCTACTACTGGGGCCAGGGGACCC AGGTCAACCGTCTCCTCACACCACCATCACCATCA <u>CGAACCTGAAGCCTAG</u>
CA4440	Nb40	53	CAGGTGCAGCTGCAGGAGTCTGGGGGAGGCTTG GTGCAGGCTGGGGGGTCTCTGAGACTCTCTTGT GCAGTCTCTGGTACGATCTTCGATATCACTCCCAT GGGCTGGTACCGCCAGACTCCAGGGAAGCAGCG CGAAGTGGTCCGAGATCTTACTAGTCGCGGTACC

[0353]

			ACAAATTACGCAGACTCCGTGAAGGGCCGGTTCA CCATCTCCAGAGACAACGCCAAGAAAATGTTGTAT CTGCAAATGAACAGCCTGAAATCTGACGACACAG GCGTGTATTACTGTAACGTGAAACGGTGGGGAGG TATTGGCTGGAACGACTACTGGGGCCAGGGGAC CCAGGTCACCGTCTCCTCACACCACCATCACCAT CACGAACCTGAAGCCTAG
CA4441	Nb41	54	CAGGTGCAGCTGCAGGAGTCTGGAGGAGGCTTG GTGCAGTCTGGGGGTCTCTGAGACTCTCCTGTG TAGCCTCTGGATTGAGATTGAGTAACCTTCCTATG ATGTGGGTCCGCCAGGCCCCAGGAAAGGGGCTC GAGTGGGTCTCGCTGATTAGCATTGGTGGTAGTA CCACGAATTATGCGGACTCCGTGAAGGGCCGATT CACCATCTCCAGAGACAACGCCAAGAACAACGCTG TTTCTGCAAATGAACAGCCTGAAACCTGAGGACA CGGCCGTGTATTACTGTGCAAATATCTTGGTCG GCTGGTCCCACCGACTACTGAGGGCCAGGGGAC CCAGGTCACCGTCTCCTCACACCACCATCACCAT CACGAACCTGAAGCCTAG

[0354] 表5:G蛋白亚基的同种型的实例

[0355]

蛋白质/亚基	登录号 (SEQ ID NO)	同种型	AA序列
人 Gαs短	P63092 (56)	GNAS2_HUMAN	MGCLGNSKTEDQRNEEKAQREANKKIEKQLQKDKQ VYRATHRLLLLGAGESGKSTIVKQMRILHVNGFNGD SEKATKVQDIKNNLKEAETIVAAMSNLVPPVELANP ENQFRVDYILSVMNVPDFDFPPEFYEHAALWEDE GVRACYERSNEYQLIDCAQYFLDKIDVIKQADYVPSD QDLLRCRVLTSGIFETKFQVDKVNFMFDVGGQQRD ERRKWIQCFNDVTAHFVVASSYNMVIREDNQTNR LQEALNLFKSIWNNRWLRTISVILFNKQDLLAEKVL AGKSKIEDYFPEFARYTTPEDATPEPGEDPRVTRAKY FIRDEFILISTASGDGRHYCYPHFTCAVDTENIRRVFN DCRDIIQRMHLRQYELL
人 Gαi	P63096 (57)	GNAI1_HUMAN	MGCTLSAEDKAAVERSKMIDRNLRDGEKAAREVKL LLGAGESGKSTIVKQMKIIHEAGYSEEECKQYKAVV YSNTIQSIIAIRAMGRLLKIDFGDSARADDARQLFVLA GAAEEGFMTAELAGVIKRLWKDSGVQACFNRSREY QLNDSAAYYLNLDRIAQPNYIPTQQDVLRTVRKTT GIVETHFTFKDLHFKMFDVGGQSRERKKWIHCFEGV TAIIFCVALSDYDLVLAEDEEMNRMHESMKLFDSICN NKWFTDTSIILFNKKDLFEKIKKSPLTICYPEYAGSN TYEEAAAYIQCFEDLNKRKDTKEIYTHFTCATDTKN VQVFVDAVTDVIIKNNLKDCGLF
人 Gαt	P11488	GNAT1_HUMAN	MGAGASAEKHSRELEKKLKEDAEDKARTVKLLLLG AGESGKSTIVKQMKIIHQDGYSLEECLEFIIYIYNTLQ SILAIVRAMTTLNIQYGDSARQDDARKLMHMADTIE

[0356]

蛋白质/亚基	登录号 (SEQ ID NO)	同种型	AA 序列
	(58)		EGTMPKEMSDIIQRLWKDSGIQACFERASEYQLNDS AGYYLSDLERLVTPGYVPTEQDVLRSRVKTTGIIETQF SFKDLNFRMFDVGGQRSEKWKWIHCFEVTCIIIFIAA LSAYDMVLVEDDEVNRMHESLHLFNSICNHRYPATT SIVLFLNKKDVFFKIKKAHLSICFPDYDGPNTYEDAG NYIKVQFLELNMRRDVKEIYSHMTCATDTQNVKFVF DAVTDIIKENLKDCLGF
牛 Gas 短	P04896 (59)	GNAS2_BOVIN	MGCLGNSKTEDQRNEEKAQREANKKIEKQLQKDKQ VYRATHRLLLLGAGESGKSTIVKQMRILHVNGFNGE GGEEDPQAARSNSDGEKATKVQDIKNNLKEAETIV AAMSNLVPPVELANPENQFRVDYILSVMNVPNDFD PPEFYEHAKALWEDEGVRACYERSNEYQLIDCAQYF LDKIDVIKQDDYVPSDQDLLRCRVLTSGIFETKFQVD KVNFMFDVGGQRDERRKWIQCFNDVTAIIFVVAS SSYNMVIREDNQTNRLQEALNLFKSIWNNRWLRTIS VILFLNKQDLLAEKVLGKSKIEDYFPEFARYTTPEDA TPEPGEDPRVTRAKYFIRDEFRLISTASGDGRHYCYP HFTCAVDTENIRRVFNDCRDIIQRMHLRQYELL
大鼠 Gas 短	P63095 (60)	GNAS2_RAT	MGCLGNSKTEDQRNEEKAQREANKKIEKQLQKDKQ VYRATHRLLLLGAGESGKSTIVKQMRILHVNGFNGE GGEEDPQAARSNSDGEKATKVQDIKNNLKEAETIV AAMSNLVPPVELANPENQFRVDYILSVMNVPNDFD PPEFYEHAKALWEDEGVRACYERSNEYQLIDCAQYF LDKIDVIKQADYVPSDQDLLRCRVLTSGIFETKFQVDK VNFHMFVGGQRDERRKWIQCFNDVTAIIFVVASS SYNMMVIREDNQTNRLQEALNLFKSIWNNRWLRTISV ILFLNKQDLLAEKVLGKSKIEDYFPEFARYTTPEDAT PEPGEDPRVTRAKYFIRDEFRLISTASGDGRHYCYPH FTCAVDTENIRRVFNDCRDIIQRMHLRQYELL
小鼠 Gas 短	P63094 (61)	GNAS2_MOUSE	MGCLGNSKTEDQRNEEKAQREANKKIEKQLQKDKQ VYRATHRLLLLGAGESGKSTIVKQMRILHVNGFNGE GGEEDPQAARSNSDGEKATKVQDIKNNLKEAETIV AAMSNLVPPVELANPENQFRVDYILSVMNVPNDFD PPEFYEHAKALWEDEGVRACYERSNEYQLIDCAQYF LDKIDVIKQADYVPSDQDLLRCRVLTSGIFETKFQVDK VNFHMFVGGQRDERRKWIQCFNDVTAIIFVVASS SYNMMVIREDNQTNRLQEALNLFKSIWNNRWLRTISV ILFLNKQDLLAEKVLGKSKIEDYFPEFARYTTPEDAT PEPGEDPRVTRAKYFIRDEFRLISTASGDGRHYCYPH FTCAVDTENIRRVFNDCRDIIQRMHLRQYELL
牛 Gβ	P62871 (62)	GBB1_BOVIN	MSELDQLRQAEQLKNQIRDARKACADATLSQITN NIDPVGRIQMRTTRTLRGHLAKIYAMHWGTD SRL VSASQDGKLIWDSYTTNKVHAIPLRSSWVMTCAYA PSGNYVACGGLDNICSIIYNLKTREGNVVRSRELAGHT

[0357]

蛋白质/亚基	登录号 (SEQ ID NO)	同种型	AA序列
			GYLSCCRFLDDNQIVTSSGDTTCALWDIETGQQTTT FTGHTGDVMSLSLAPDTRLFVSGACDASAKLWDVR EGMCRQTFTGHESDINAICFFPNGNAFATGSDDATC RLFDLRADQELMTYSHDNIICGITSVSFSKSGRLLLAG YDDFNCNVWDALKADRAGVLAGHDNRVSLGVT DGMVATGSWDSFLKIWN
人 Gβ	P62873 (63)	GBB1_HUMAN	MSELDQLRQAEQLKNQIRDARKACADATLSQITN NIDPVGRIQMRTTRTLRGHLAKIYAMHWGTDSSRL VSASQDGKLIWDSYTTNKVHAIPRSSWVMTCA YAPSGNYVACGGLDNICSIYNLKTREGNVRVS RELAGHTGYLSCCRFLDDNQIVTSSGDTTCALWDIETGQQTTT FTGHTGDVMSLSLAPDTRLFVSGACDASAKLWDVR EGMCRQTFTGHESDINAICFFPNGNAFATGSDDATC RLFDLRADQELMTYSHDNIICGITSVSFSKSGRLLLAG YDDFNCNVWDALKADRAGVLAGHDNRVSLGVT DGMVATGSWDSFLKIWN
大鼠 Gβ	P54311 (64)	GBB1_RAT	MSELDQLRQAEQLKNQIRDARKACADATLSQITN NIDPVGRIQMRTTRTLRGHLAKIYAMHWGTDSSRL VSASQDGKLIWDSYTTNKVHAIPRSSWVMTCA YAPSGNYVACGGLDNICSIYNLKTREGNVRVS RELAGHTGYLSCCRFLDDNQIVTSSGDTTCALWDIETGQQTTT FTGHTGDVMSLSLAPDTRLFVSGACDASAKLWDVR EGMCRQTFTGHESDINAICFFPNGNAFATGSDDATC RLFDLRADQELMTYSHDNIICGITSVSFSKSGRLLLAG YDDFNCNVWDALKADRAGVLAGHDNRVSLGVT DGMVATGSWDSFLKIWN
小鼠 Gβ	P62874 (65)	GBB1_MOUSE	MSELDQLRQAEQLKNQIRDARKACADATLSQITN NIDPVGRIQMRTTRTLRGHLAKIYAMHWGTDSSRL VSASQDGKLIWDSYTTNKVHAIPRSSWVMTCA YAPSGNYVACGGLDNICSIYNLKTREGNVRVS RELAGHTGYLSCCRFLDDNQIVTSSGDTTCALWDIETGQQTTT FTGHTGDVMSLSLAPDTRLFVSGACDASAKLWDVR EGMCRQTFTGHESDINAICFFPNGNAFATGSDDATC RLFDLRADQELMTYSHDNIICGITSVSFSKSGRLLLAG YDDFNCNVWDALKADRAGVLAGHDNRVSLGVT DGMVATGSWDSFLKIWN
牛 Gy	P63212 (66)	GBG2_BOVIN	MASNNTASIAQARKLVEQLKMEANIDRIKVS KAAADLMAYCEAHAKEDPLTPVPA SENPFREKKFFCAIL
小鼠 Gy	P63213 (67)	GBG2_MOUSE	MASNNTASIAQARKLVEQLKMEANIDRIKVS KAAADLMAYCEAHAKEDPLTPVPA SENPFREKKFFCAIL

[0358]

蛋白质/亚基	登录号 (SEQ ID NO)	同种型	AA序列
人 Gy	P59768 (68)	GBG2_HUMAN	MASNNTASIAQARKLVEQLKMEANIDRIKVSAAAAD LMAYCEAHAKEDPLLTPVPAENP FREKKFFCAIL

[0359] 表6. 在R:G界面内的潜在分子间相互作用

	β_2 AR 中的原子	Gs α 中的原子	距离 (Å)
TM3	[ARG 131 CG]	[TYR 391 CE2]	3.6
	[ALA 134 O]	[HIS 387 ND1]	3.1
	[ALA 134 CB]	[HIS 387 ND1]	3.8
	[ALA 134 CB]	[HIS 387 CG]	3.8
	[ILE 135 O]	[GLN 384 NE2]	2.9
	[ILE 135 CD1]	[LEU 388 CD1]	3.5
	[ILE 135 CD1]	[LEU 393 CD1]	3.5
	[THR 136 O]	[ARG 380 NH2]	3.0
IL2	[PRO 138 O]	[ILE 383 CD1]	3.4
	[PRO 138 CG]	[GLN 384 CG]	3.3
	[PHE 139 CD2]	[HIS 41 NE2]	3.6
	[PHE 139 CB]	[VAL 217 CG2]	3.7
	[PHE 139 CE1]	[PHE 376 CZ]	3.3
	[PHE 139 CZ]	[CYS 379 C]	3.9
	[PHE 139 CZ]	[ARG 380 N]	3.2
	[TYR 141 CD2]	[HIS 387 NE2]	3.9
	[GLN 142 OE1]	[HIS 387 NE2]	3.5
	[SER 143 OG]	[ALA 39 CB]	3.6
TM5	[VAL 222 CG1]	[LEU 393 CD1]	3.7
	[GLU 225 OE2]	[GLN 384 NE2]	3.0
	[ALA 226 CA]	[LEU 388 CD2]	3.6
	[ALA 226 CB]	[LEU 393 O]	3.8
	[GLN 229 NE2]	[ASP 381 OD1]	3.6
	[GLN 229 NE2]	[GLN 384 OE1]	2.8
	[GLN 229 OE1]	[ARG 385 NE]	3.4
	[GLN 229 CG]	[LEU 388 CD2]	3.9
	[LEU 230 CG]	[LEU 394 CD1]	3.5
	[LYS 232 NZ]	[ASP 381 OD1]	3.1
	[ILE 233 CD1]	[TYR 358 OH]	3.4
	[ILE 233 CG1]	[ARG 385 NH1]	3.5
	[ARG 239 NE]	[THR 350 OG1]	3.5
TM6	[ALA 271 CB]	[LEU 393 O]	3.0
	[THR 274 CG2]	[GLU 392 O]	3.4
	[THR 274 OG1]	[LEU 393 CD2]	3.3
	[LEU 275 CD2]	[LEU 393 CD2]	3.6

[0361] 表7.数据收集和精修统计

[0362]

数据采集*	
晶体数	20
空间群	P 2 ₁
晶胞维度	
a, b, c (Å)	119.3, 64.6, 131.2
α, β, γ (°)	90.0, 91.7, 90.0
分辨率 (Å)	41 - 3.2 [3.26 - 3.20]
R _{merge} (%)	15.6 (55.3)
<I> / <σI>	10.8 (1.8)
完整度 (%)	91.2 (53.9)
冗余度	6.5 [5.0]
精修	
分辨率 (Å)	41 - 3.2
无反射	31075 (1557 在测试集中)
R _{work} /R _{free} (%)	22.5 / 27.7
无原子	10277
无蛋白质残基	1318
各向异性B张量	B ₁₁ = -7.0 / B ₂₂ = 4.7 / B ₃₃ = 2.3 / B ₁₂ = 2.1
未建模序列 ^a	
β ₂ 肾上腺素能受体	29 ^b , 176-178, 240-264, 342-365
G _s α _s ras 结构域	1-8, 60-88, 203-204, 256-262
G _s γ	1-4, 63-68
T4 溶菌酶	161 ^c
平均B-因子(Å ²)	
β ₂ 肾上腺素能受体	133.5
G _s α _s ras 结构域	82.8
G _s α _s 螺旋结构域	123.0
G _s β	64.2
G _s γ	85.2
纳米抗体 35	60.7
T4 溶菌酶	113.7
自理想值的R.m.s. 偏差	
键长 (Å)	0.007
键角 (°)	0.72
Ramachandran 统计 ^d	
最大允许区(Favored regions) (%)	95.8
允许区(allowed regions) (%)	4.2
非允许区(Outliers) (%)	0

[0363] *括号中是最高的壳统计。^a由于电子密度解析较差,从模型中省略了这些区域。这些残基范围内不包括未建模的纯化标签。^b构建体省略了β2AR 的第1-28位残基,T4L与跨膜螺旋1的氨基末端融合,有利于结晶。^c构建体省略了T4L残基。^d如MolProbity定义的。

[0364] 参考文献

[0365] -Afonine,P.V.,Grosse-Kunstleve,R.W.,&Adams,P.D. (2005) .Arobust bulk-

solvent correction and anisotropic scaling procedure. *Acta crystallographica*. Section D, Biological crystallography, 61, 850–5.

[0366] -Baltensperger, K. 等人 The beta-adrenergic receptor is a substrate for the insulin receptor tyrosine kinase. *J Biol Chem* 271, 1061–1064 (1996)。

[0367] -Binz 等人, *Nature Biotech.*, 22:575–582 (2004)。

[0368] -Blanc, E., Roversi, P., Vöhrlein, C., Flensburg, C., Lea, S. M., Bricogne, G., 等人 (2004). Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT. *Acta crystallographica. Section D, Biological crystallography*, 60, 2210–21。

[0369] -Caffrey (2003). Membrane protein crystallization. *J Struct. Biol.* 2003 142:108–32。

[0370] -Caffrey, M. & Cherezov, V. Crystallizing membrane proteins using lipidic mesophases. *Nat Protoc* 4, 706–731, (2009)。

[0371] -Chae, P. S. 等人 Maltose-neopenytl glycol (MNG) amphiphiles for solubilization, stabilization and crystallization of membrane proteins. *Nat Methods* 7, 1003–1008 (2010)。

[0372] -Chelikani 等人 *Protein Sci.* 2006 15:1433–40。

[0373] -Chini, B., & Parenti, M. (2009). G-protein-coupled receptors, cholesterol and palmitoylation: facts about fats. *Journal of molecular endocrinology*, 42 (5), 371–9。

[0374] -Chomczynski, P. 和 Sacchi, N., 1987. Single step method of RNA isolation by acid guanidium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* 162, 页 156。

[0375] -Conrath K, Pereira AS, Martius CE, Timóteo CG, Tavares P, Spinelli S, Kinne J, Flaudrops C, Cambillau C, Muyldermans S, Moura I, Moura JJ, Tegoni M, Desmyter A. Camelid nanobodies raised against an integral membrane enzyme, nitric oxide reductase. *Protein Sci.* 2009 Mar; 18 (3): 619–28。

[0376] -Conrath K. E., M. Lauwereys, M. Galleni 等人, *Antimicrob Agents Chemother* 45 (10), 2807 (2001)。

[0377] -Costanzi S, Vilar S (2011). In Silico screening for agonists and blockers of the beta (2) adrenergic receptor: Implications of inactive and activated state structures. *Journal of computational chemistry* 33: 561–572。

[0378] -Day P. W., Rasmussen S. G., Parrot C., Fung J. J., Masood A., Kobilka T. S., Yao X. J., Choi H. J., Weis W. I. 和 Rohrer D. K. 等人, A monoclonal antibody for G protein-coupled receptor crystallography, *Nat Methods* 4 (2007), 页 927–929。

[0379] -Delean, A., J. M. Stadel, 等人 (1980). "A TERNARY COMPLEX MODEL EXPLAINS THE AGONIST-SPECIFIC BINDING-PROPERTIES OF THE ADENYLATE CYCLASE-COUPLED BETA-ADRENERGIC-RECEPTOR." *JOURNAL OF BIOLOGICAL CHEMISTRY* 255 (15): 7108–7117。

[0380] -Derewenda Z. S. Rational protein crystallization by mutational surface

engineering, Structure (Camb) 12 (2004), 页529-535。

[0381] -Domanska, K. 等人 Atomic structure of a nanobody-trapped domain-swapped dimer of an amyloidogenic beta2-microglobulin variant. Proc Natl Acad Sci U S A 108, 1314-1319, (2011)。

[0382] -Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 60, 2126-2132, (2004)。

[0383] -Eroglu 等人 EMB020023:491~96

[0384] -Eroglu 等人 Proc. Natl. Acad. Sci. 2003100:10219-10224

[0385] -Faham 等人 Crystallization of bacteriorhodopsin from bicelle formulations at room temperature. Protein Sci. 200514:836-40. 2005

[0386] -Faham 等人, Bicelle crystallization: a new method for crystallizing membrane proteins yields a monomeric bacteriorhodopsin structure. J Mol Biol. 2002Feb8; 316(1):1-6。

[0387] -Foord, S. M., T. I. Bonner, 等人 (2005). "International Union of Pharmacology. XLVI. G protein-coupled receptor list." Pharmacological reviews 57 (2):279-288。

[0388] -Fredriksson, R., M. C. Lagerstrom, 等人 (2003). "The G-protein-coupled receptors in the human genome form five main families. Phylogenetic analysis, paralogon groups, and fingerprints." Molecular Pharmacology 63 (6):1256-1272。

[0389] -Gebauer & Skerra (2009) Current opinion in chemical biology 13, 245-255

[0390] -George 等人, Nat Rev Drug Discov 1:808-820 (2002)

[0391] -Gouaux, It's not just a phase: crystallization and X-ray structure determination of bacteriorhodopsin in lipidic cubic phases. Structure. 19986:5-10;

[0392] -Hamers-Casterman, C., T. Atarhouch, S. Muyldermans 等人 Naturally occurring antibodies devoid of light chains. Nature 363, 446-448, doi:10.1038/363446a0 (1993)。

[0393] -Hendrickson W A. Determination of macromolecular structures from anomalous diffraction of synchrotron radiation. Science. 1991 Oct 4; 254 (5028):51-8。

[0394] -Hofmann K. P., P. Scheerer, P. W. Hildebrand 等人, Trends Biochem Sci 34 (11), 540 (2009)。

[0395] -Hunte C. and Michel H., Crystallisation of membrane proteins mediated by antibody fragments, Curr Opin Struct Biol 12 (2002), 页503-508。

[0396] -Jameson, E. E. 等人 (2005). Real-time detection of basal and stimulated G protein GTPase activity using fluorescent GTP analogues. J Biol Chem 280, 7712-7719。

[0397] -Kenakin, Trends Pharmacol Sci 25:186-192 (2002)

[0398] -Kobilka 等人 (2007) Trends in pharmacological sciences 28, 397-406。

- [0399] -Kobilka,B.K.Amino and carboxylterminal modifications to facilitate the production and purification of a G protein-coupled receptor.Anal Biochem231,269-271 (1995)。
- [0400] -Koide等人,J.Mol Biol,284:1141-1151 (1998))
- [0401] -Kolakowski,L.F. (1994) ."GCRDB-A G-PROTEIN-COUPLED RECEPTOR DATABASE." Receptors&Channels2 (1) :1-7。
- [0402] -Kolb,P.,D.M.Rosenbaum,等人 (2009) ."Structure-based discovery of beta (2)-adrenergic receptor ligands."Proceedings of the National Academy of Sciences of the United States of America106 (16) :6843-6848。
- [0403] -Kusmak,A.J.,S.Pitchaya,等人 (2009) ."Purification and functional reconstitution of monomeric mu-opioid receptors:allosteric modulation of agonist binding by Gi2."The Journal of biological chemistry284:26732-26741。
- [0404] -**Lagerström**,M.C.和H.B.**Schiöth** (2008) ."Structural diversity of G protein-coupled receptors and signifiHcance for drug discovery."Nature reviews.Drug discovery7:339-357。
- [0405] -Landau等人,Lipidic cubic phases:a novel concept for the crystallization of membrane proteins.Proc.Natl.Acad.Sci.199693:14532-5
- [0406] -Lee,A.G. (2004) .How lipids affect the activities ofintegral membrane proteins.Biochimica et biophysica acta,1666 (1-2) ,62-87。
- [0407] -Lee GM,Craik CS (2009) .Trapping moving targets with small molecules.Science.Apr10;324 (5924) :213-5。
- [0408] -Lefranc,M.P.,C.Pommie,等人 (2003) ."IMGT unique numbering for immunoglobulin and T cell receptor variable domains and Ig superfamily V-like domains."Developmental and Comparative Immunology27 (1) :55-77。
- [0409] -Li H.,Dunn J.J.,Luft B.J.and Lawson C.L.,Crystal structure of Lyme disease antigen outer surface protein A complexed with an Fab,Proc Natl Acad Sci U S A94 (1997) ,页3584-3589
- [0410] -Luca等人Proc.Natl.Acad.Sci.2003100:10706-11
- [0411] -Lynch Kevin R. (编) Identification and Expression of G Protein-Coupled Receptors published by John Wiley&Sons (March1998)
- [0412] -Mansoor等人Proc.Natl.Acad.Sci.2006103:3060-3065
- [0413] -Marchese等人Genomics23:609-618,1994
- [0414] -McCoy,A.J.Solving structures of protein complexes by molecular replacement with Phaser.Acta Crystallogr D Biol Crystallogr63,32-41 (2007) 。
- [0415] -McCoy,A.J.,Grosse-Kunstleve,R.W.,Adams,P.D.,Winn,M.D.,Storoni,L.C.,Read,R.J.,等人 (2007) .Phaser crystauographic software.Journal of applied crystallography,40 (Pt4) ,658-674。
- [0416] -McEwen,D.P.,Gee,K.R.,Kang,H.C.&Neubig,R.R. (2001) Fluorescent BODIPY-GTP analogs:real-time measurement ofnucleotide binding to G proteins.Ana/

Biochem291,109-117。

[0417] -Misquitta,L.V.等人Membrane protein crystallization in lipidic mesophases with tailored bilayers.Structure12,2113-2124,(2004)。

[0418] -Moro,O.,Lameh,J.,Hogger,P.&Sadegh,W.Hydrophobic amino acid in the i2loop plays a key role in receptor-G protein coupling.J Biol Chem268,22273-22276(1993)。

[0419] -Niu等人,Biophys J.200589:1833-1840

[0420] -Nollert等人Lipidic cubic phases as matrices for membrane protein crystallization Methods.200434:348-53

[0421] -Ostermeier C.,Iwata S.,Ludwig B.and Michel H.,Fv fragment-mediated crystallization of the membrane protein bacterial cytochrome c oxidase,Nat Struct Biol2(1995),页842-846。

[0422] -Otwinowski,Z.,&Minor,W.(1997).Processing of X-ray diffraction data collected in oscillation mode.Methods in enzymology,276,307-325.

[0423] -Palczewski,K.等人Crystal structure of rhodopsin:A G protein-coupled receptor[see comments].Science289,739-745(2000)。

[0424] -Probst等人1992,DNA Cell Biol.199211:1-20:

[0425] -Qian ZM,Li H,Sun H和Ho K(2002).Targeted drug delivery via the transferring receptor-mediated endocytosis pathway.Pharmacol Rev54,561-587。

[0426] -Rasmussen等人(2011)Nature469,175-180。

[0427] -Rasmussen S.G.,Choi H.J.,Rosenbaum D.M.,Kobilka T.S.,Thian F.S.,Edwards P.C.,Burghammer M.,Ratnala V.R.,Sanishvili R.和Fischetti R.F.等人,Crystal structure of the human beta2adrenergic G-protein-coupled receptor,Nature450(2007),页383-387。

[0428] -Riechmann和Muyldermans J.Immunol. Methods2000;240:185-195。

[0429] -Rios等人,Pharmacol Ther92:71-87(2001)。

[0430] -Ritter,S.L.,&Hall,R.a.(2009).Fine-tuning of GPCR activity by receptor-interacting proteins.Nature reviews.Molecular cell biology,10(12),819-30.Nature Publishing Group.doi:10.1038/nrm2803.)

[0431] -Rosenbaum D.M.,S.G.Rasmussen,and B.K.Kobilka,Nature459(7245),356(2009)。

[0432] -Rosenbaum,D.M.等人Structure and function of an irreversible agonist-beta(2)adrenoceptor complex.Nature469,236-240(2011)。

[0433] -Rosenbaum,D.M.,V.Cherezov,等人(2007).“GPCR engineering yields high-resolution structural insights into beta2-adrenergic receptor function.” Science318:1266-1273。

[0434] -Rummel等人,Lipidic Cubic Phases:New Matrices for the Three-Dimensional Crystallization of Membrane Proteins.J.Struct.Biol.1998121:82-91;

[0435] -Sawant R,Torchilin V.Intracellular transduction using cell-

penetrating peptides. *Mol Biosyst.* 2010 Apr; 6 (4) :628-40. Epub 2009 Dec 21.

[0436] -Seifert等人(1998). Reconstitution of beta2-adrenoceptor-GTP-binding-protein interaction in Sf9 cells--high coupling efficiency in a beta2-adrenoceptor-G(s alpha) fusion protein. *Eur. J. Biochem.* 255:369-382.

[0437] -Shimada等人 *J. Biol. Chem.* 2002 277:31774-80

[0438] -Skerra, J. *Molecular Recognition*, 13:167-187 (2000)

[0439] -Sprang, S. R. *G protein mechanisms: insights from structural analysis. Annu Rev Biochem* 66, 639-678 (1997)。

[0440] -Starovasnik等人, *Proc. Natl. Acad. Sci. USA*, 94:10080-10085 (1997)

[0441] -Steyaert J, Kobilka BK (2011). Nanobody stabilization of G protein-coupled receptor conformational states. *Curr Opin Struct Biol.* Aug; 21 (4) :567-72.

[0442] -Strong, M. 等人 Toward the structural genomics of complexes: crystal structure of a PE/PPE protein complex from *Mycobacterium tuberculosis*. *Proc Natl Acad Sci U S A* 103, 8060-8065, (2006)。

[0443] -Sunahara, R. K., Tesmer, J. J., Gilman, A. G. & Sprang, S. R. Crystal structure of the adenylyl cyclase activator Gs alpha [see comments]. *Science* 278, 1943-1947 (1997)。

[0444] -Swaminath, G., Steenhuis, J., Kobilka, B. & Lee, T. W. Allosteric modulation of beta2-adrenergic receptor by Zn(2+). *Mol Pharmacol* 61, 65-72. (2002)。

[0445] -Tatsuya Haga (编), *G Protein-Coupled Receptors*, published by CRC Press (September 24, 1999)

[0446] -Van Eps, N. 等人 Interaction of a G protein with an activated receptor opens the interdomain interface in the alpha subunit. *Proc Natl Acad Sci U S A* (2011)。

[0447] -Wall, M. A. 等人 The structure of the G protein heterotrimer Gialb1g2. *Cell* 83, 1047-1058 (1995)。

[0448] -Warne, T. 等人 Structure of a beta1-adrenergic G-protein-coupled receptor. *Nature* 454, 486-491, (2008)。

[0449] -Watson, S. (编) *G-Protein Linked Receptor Factsbook*, published by Academic Press (第一版; 1994)。

[0450] -Wess Jurgen (编) *Structure-Function Analysis of G Protein-Coupled Receptors* published by Wiley-Liss (1st edition; October 15, 1999)

[0451] -Whorton, M. R. 等人 A monomeric G protein-coupled receptor isolated in a high-density lipoprotein particle efficiently activates its G protein. *Proc Natl Acad Sci U S A* 104, 7682-7687 (2007)。

[0452] -Whorton, M. R., S. G. F. Rasmussen, 等人 (2009). "The effect of ligand efficacy on the formation and stability of a GPCR-G protein complex." *PNAS* 106:1-6。

序列表

<110> 非营利性组织佛兰芒综合大学生物技术研究所(VIB VZW)
 布鲁塞尔自由大学(VRIJE UNIVERSITEIT BRUSSEL)
 利兰斯坦福青年大学托管委员会(TH E BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY)
 密执安大学董事会(TH E REGENTS OF TH E UNIVERSITY OF MICHIGAN)

<120> 针对GPCR:G蛋白复合物的结合结构域及来自其的用途

<130> JS/GPROT/380

<150> US 61/571, 159

<151> 2011-06-21

<150> EP 11181357.2

<151> 2011-09-15

<160> 73

<170> PatentIn 版本 3.5

<210> 1

<211> 138

<212> PRT

<213> 家羊驼(Lama glama)

<400> 1

Gln Val Gln Leu Gln Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
 1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Asn Tyr
 20 25 30

Lys Met Asn Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
 35 40 45

Ser Asp Ile Ser Gln Ser Gly Ala Ser Ile Ser Tyr Thr Gly Ser Val
 50 55 60

Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys Asn Thr Leu Tyr
 65 70 75 80

Leu Gln Met Asn Ser Leu Lys Pro Glu Asp Thr Ala Val Tyr Tyr Cys
 85 90 95

Ala Arg Cys Pro Ala Pro Phe Thr Arg Asp Cys Phe Asp Val Thr Ser
 100 105 110

Thr Thr Tyr Ala Tyr Arg Gly Gln Gly Thr Gln Val Thr Val Ser Ser
 115 120 125

His His His His His His Glu Pro Glu Ala
 130 135

<210> 2

<211> 138

<212> PRT

<213> 家羊驼

<400> 2

Gln Val Gln Leu Gln Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
 1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Asn Tyr
 20 25 30

Val Met Asn Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
 35 40 45

Ser Asp Ile Ser Asn Gly Gly Gly Thr Thr Ser Tyr Ala Ser Ser Val
 50 55 60

Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys Asn Thr Leu Tyr
 65 70 75 80

Leu Gln Met Asn Gly Leu Lys Pro Ala Asp Thr Ala Val Tyr Tyr Cys

[0001]

		85		90		95
Ala Arg Cys	Pro Ala Pro Phe Thr	Asn Asp Cys Met Asp Ile Thr Ser				
	100	105		110		
Thr Thr Tyr	Ala Tyr Arg Gly Gln Gly Thr Gln Val Thr Val Ser Ser					
	115	120		125		
His His His His His His	Glu Pro Glu Ala					
	130	135				
<210>	3					
<211>	129					
<212>	PRT					
<213>	家羊驼					
<400>	3					
Gln Val Gln Leu Gln Glu Ser Gly Gly Gly Ser Val Gln Ala Gly Gly						
1	5	10		15		
Ser Leu Arg Leu Ser Cys Thr Val Ser Gly Thr Ile Phe Ser Val Thr						
	20	25		30		
Val Met Gly Trp Tyr Arg Gln Ala Pro Gly Lys Gln Arg Glu Leu Val						
	35	40		45		
Ala Gly Phe Thr Asn Thr Arg Asn Thr Asn Tyr Val Asp Ser Val Lys						
	50	55		60		
Gly Arg Phe Thr Ile Ser Lys Asp Ser Ala Lys Asn Thr Met Tyr Leu						
65	70	75		80		
Gln Met Asn Ser Leu Lys Pro Glu Asp Thr Ala Val Tyr Tyr Cys Asn						
	85	90		95		
Val Arg Arg Trp Gly Gly Thr Asn Trp Asn Asp Tyr Trp Gly Gln Gly						
	100	105		110		
Thr Gln Val Thr Val Ser Ser His His His His His His Glu Pro Glu						
	115	120		125		
Ala						
<210>	4					
<211>	129					
<212>	PRT					
<213>	家羊驼					
<400>	4					
Gln Val Gln Leu Gln Glu Ser Gly Gly Gly Phe Val Gln Ala Gly Gly						
1	5	10		15		
Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Ser Ile Phe Ser Lys Asn						
	20	25		30		
Thr Met Ala Trp Phe Arg Gln Ala Pro Gly Lys Glu Arg Glu Leu Val						
	35	40		45		
Ala Ala Ser Pro Thr Gly Gly Ser Thr Ala Tyr Lys Asp Ser Val Lys						
	50	55		60		
Gly Arg Phe Thr Ile Ser Arg Asp Ser Ala Lys Asn Thr Val Leu Leu						
65	70	75		80		
Gln Met Asn Val Leu Lys Pro Glu Asp Thr Ala Val Tyr Tyr Cys His						
	85	90		95		
Leu Arg Gln Asn Asn Arg Gly Ser Trp Phe His Tyr Trp Gly Gln Gly						
	100	105		110		
Thr Gln Val Thr Val Ser Ser His His His His His His Glu Pro Glu						
	115	120		125		
Ala						

[0002]

<210> 5
 <211> 129
 <212> PRT
 <213> 家羊驼

<400> 5

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

Ala

<210> 6
 <211> 129
 <212> PRT
 <213> 家羊驼

<400> 6

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

Ala

<210> 7
 <211> 25
 <212> PRT
 <213> 家羊驼

<400> 7

Gln Val Gln Leu Gln Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly

[0003]

	1	5	10	15
	Ser Leu Arg Leu Ser Cys Ala Ala Ser			
		20	25	
	<210> 8			
	<211> 25			
	<212> PRT			
	<213> 家羊驼			
	<400> 8			
	Gln Val Gln Leu Gln Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly			
	1	5	10	15
	Ser Leu Arg Leu Ser Cys Ala Ala Ser			
		20	25	
	<210> 9			
	<211> 25			
	<212> PRT			
	<213> 家羊驼			
	<400> 9			
	Gln Val Gln Leu Gln Glu Ser Gly Gly Gly Ser Val Gln Ala Gly Gly			
	1	5	10	15
	Ser Leu Arg Leu Ser Cys Thr Val Ser			
		20	25	
	<210> 10			
	<211> 25			
	<212> PRT			
	<213> 家羊驼			
	<400> 10			
[0004]	Gln Val Gln Leu Gln Glu Ser Gly Gly Gly Phe Val Gln Ala Gly Gly			
	1	5	10	15
	Ser Leu Arg Leu Ser Cys Ala Ala Ser			
		20	25	
	<210> 11			
	<211> 25			
	<212> PRT			
	<213> 家羊驼			
	<400> 11			
	Gln Val Gln Leu Gln Glu Ser Gly Gly Gly Leu Val Gln Ala Gly Gly			
	1	5	10	15
	Ser Leu Arg Leu Ser Cys Ala Val Ser			
		20	25	
	<210> 12			
	<211> 25			
	<212> PRT			
	<213> 家羊驼			
	<400> 12			
	Gln Val Gln Leu Gln Glu Ser Gly Gly Gly Leu Val Gln Ser Gly Gly			
	1	5	10	15
	Ser Leu Arg Leu Ser Cys Val Ala Ser			
		20	25	
	<210> 13			
	<211> 8			
	<212> PRT			
	<213> 家羊驼			
	<400> 13			
	Gly Phe Thr Phe Ser Asn Tyr Lys			

	1		5
	<210>	14	
	<211>	8	
	<212>	PRT	
	<213>	家羊驼	
	<400>	14	
	Gly Phe Thr Phe Ser Asn Tyr Val		
	1	5	
	<210>	15	
	<211>	8	
	<212>	PRT	
	<213>	家羊驼	
	<400>	15	
	Gly Thr Ile Phe Ser Val Thr Val		
	1	5	
	<210>	16	
	<211>	8	
	<212>	PRT	
	<213>	家羊驼	
	<400>	16	
	Gly Ser Ile Phe Ser Lys Asn Thr		
	1	5	
	<210>	17	
	<211>	8	
	<212>	PRT	
	<213>	家羊驼	
[0005]	<400>	17	
	Gly Thr Ile Phe Asp Ile Thr Pro		
	1	5	
	<210>	18	
	<211>	8	
	<212>	PRT	
	<213>	家羊驼	
	<400>	18	
	Gly Phe Arg Phe Ser Asn Phe Pro		
	1	5	
	<210>	19	
	<211>	17	
	<212>	PRT	
	<213>	家羊驼	
	<400>	19	
	Met Asn Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val Ser		
	1	5	10 15
	Asp		
	<210>	20	
	<211>	17	
	<212>	PRT	
	<213>	家羊驼	
	<400>	20	
	Met Asn Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val Ser		
	1	5	10 15
	Asp		

<210> 21
 <211> 17
 <212> PRT
 <213> 家羊驼

<400> 21

Met Gly Trp Tyr Arg Gln Ala Pro Gly Lys Gln Arg Glu Leu Val Ala
 1 5 10 15

Gly

<210> 22
 <211> 17
 <212> PRT
 <213> 家羊驼

<400> 22

Met Ala Trp Phe Arg Gln Ala Pro Gly Lys Glu Arg Glu Leu Val Ala
 1 5 10 15

Ala

<210> 23
 <211> 17
 <212> PRT
 <213> 家羊驼

<400> 23

Met Gly Trp Tyr Arg Gln Thr Pro Gly Lys Gln Arg Glu Val Val Ala
 1 5 10 15

Asp

[0006]

<210> 24
 <211> 17
 <212> PRT
 <213> 家羊驼

<400> 24

Met Met Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val Ser
 1 5 10 15

Leu

<210> 25
 <211> 8
 <212> PRT
 <213> 家羊驼

<400> 25

Ile Ser Gln Ser Gly Ala Ser Ile
 1 5

<210> 26
 <211> 8
 <212> PRT
 <213> 家羊驼

<400> 26

Ile Ser Asn Gly Gly Gly Thr Thr
 1 5

<210> 27
 <211> 7
 <212> PRT
 <213> 家羊驼

<400> 27

Phe Thr Asn Thr Arg Asn Thr
1 5

<210> 28
<211> 7
<212> PRT
<213> 家羊驼

<400> 28

Ser Pro Thr Gly Gly Ser Thr
1 5

<210> 29
<211> 7
<212> PRT
<213> 家羊驼

<400> 29

Leu Thr Ser Arg Gly Thr Thr
1 5

<210> 30
<211> 8
<212> PRT
<213> 家羊驼

<400> 30

Ile Ser Ile Gly Gly Ser Thr Thr
1 5

<210> 31
<211> 38
<212> PRT
<213> 家羊驼

[0007]

<400> 31

Ser Tyr Thr Gly Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn
1 5 10 15

Ala Lys Asn Thr Leu Tyr Leu Gln Met Asn Ser Leu Lys Pro Glu Asp
20 25 30

Thr Ala Val Tyr Tyr Cys
35

<210> 32
<211> 38
<212> PRT
<213> 家羊驼

<400> 32

Ser Tyr Ala Ser Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn
1 5 10 15

Ala Lys Asn Thr Leu Tyr Leu Gln Met Asn Gly Leu Lys Pro Ala Asp
20 25 30

Thr Ala Val Tyr Tyr Cys
35

<210> 33
<211> 38
<212> PRT
<213> 家羊驼

<400> 33

Asn Tyr Val Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Lys Asp Ser
1 5 10 15

Ala Lys Asn Thr Met Tyr Leu Gln Met Asn Ser Leu Lys Pro Glu Asp
20 25 30

[0008]

Thr Ala Val Tyr Tyr Cys
35

<210> 34
<211> 38
<212> PRT
<213> 家羊驼

<400> 34

Ala Tyr Lys Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Ser
1 5 10 15

Ala Lys Asn Thr Val Leu Leu Gln Met Asn Val Leu Lys Pro Glu Asp
20 25 30

Thr Ala Val Tyr Tyr Cys
35

<210> 35
<211> 38
<212> PRT
<213> 家羊驼

<400> 35

Asn Tyr Ala Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn
1 5 10 15

Ala Lys Lys Met Leu Tyr Leu Gln Met Asn Ser Leu Lys Ser Asp Asp
20 25 30

Thr Gly Val Tyr Tyr Cys
35

<210> 36
<211> 38
<212> PRT
<213> 家羊驼

<400> 36

Asn Tyr Ala Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn
1 5 10 15

Ala Lys Asn Thr Leu Phe Leu Gln Met Asn Ser Leu Lys Pro Glu Asp
20 25 30

Thr Ala Val Tyr Tyr Cys
35

<210> 37
<211> 21
<212> PRT
<213> 家羊驼

<400> 37

Ala Arg Cys Pro Ala Pro Phe Thr Arg Asp Cys Phe Asp Val Thr Ser
1 5 10 15

Thr Thr Tyr Ala Tyr
20

<210> 38
<211> 21
<212> PRT
<213> 家羊驼

<400> 38

Ala Arg Cys Pro Ala Pro Phe Thr Asn Asp Cys Met Asp Ile Thr Ser
1 5 10 15

Thr Thr Tyr Ala Tyr
20

<210> 39
<211> 13
<212> PRT
<213> 家羊驼

<400> 39

Asn Val Arg Arg Trp Gly Gly Thr Asn Trp Asn Asp Tyr
1 5 10

<210> 40
<211> 13
<212> PRT
<213> 家羊驼

<400> 40

His Leu Arg Gln Asn Asn Arg Gly Ser Trp Phe His Tyr
1 5 10

<210> 41
<211> 13
<212> PRT
<213> 家羊驼

<400> 41

Asn Val Lys Arg Trp Gly Gly Ile Gly Trp Asn Asp Tyr
1 5 10

<210> 42
<211> 12
<212> PRT
<213> 家羊驼

<400> 42

Ala Lys Tyr Leu Gly Arg Leu Val Pro Pro Thr Thr
1 5 10

[0009]

<210> 43
<211> 11
<212> PRT
<213> 家羊驼

<400> 43

Arg Gly Gln Gly Thr Gln Val Thr Val Ser Ser
1 5 10

<210> 44
<211> 11
<212> PRT
<213> 家羊驼

<400> 44

Arg Gly Gln Gly Thr Gln Val Thr Val Ser Ser
1 5 10

<210> 45
<211> 11
<212> PRT
<213> 家羊驼

<400> 45

Trp Gly Gln Gly Thr Gln Val Thr Val Ser Ser
1 5 10

<210> 46
<211> 11
<212> PRT
<213> 家羊驼

<400> 46

Trp Gly Gln Gly Thr Gln Val Thr Val Ser Ser

	1	5	10
	<210> 47		
	<211> 11		
	<212> PRT		
	<213> 家羊驼		
	<400> 47		
	Trp Gly Gln Gly Thr Gln Val Thr Val Ser Ser		
	1 5 10		
	<210> 48		
	<211> 11		
	<212> PRT		
	<213> 家羊驼		
	<400> 48		
	Glu Gly Gln Gly Thr Gln Val Thr Val Ser Ser		
	1 5 10		
	<210> 49		
	<211> 417		
	<212> PRT		
	<213> 家羊驼		
	<400> 49		
	Cys Ala Gly Gly Thr Gly Cys Ala Gly Cys Thr Gly Cys Ala Gly Gly		
	1 5 10 15		
	Ala Gly Thr Cys Thr Gly Gly Ala Gly Gly Ala Gly Gly Cys Thr Thr		
	20 25 30		
	Gly Gly Thr Gly Cys Ala Gly Cys Cys Thr Gly Gly Gly Gly Gly Gly		
	35 40 45		
[0010]	Thr Cys Thr Cys Thr Gly Ala Gly Ala Cys Thr Cys Thr Cys Cys Thr		
	50 55 60		
	Gly Thr Gly Cys Gly Gly Cys Cys Thr Cys Thr Gly Gly Ala Thr Thr		
	65 70 75 80		
	Cys Ala Cys Cys Thr Thr Cys Ala Gly Cys Ala Ala Thr Thr Ala Thr		
	85 90 95		
	Ala Ala Ala Ala Thr Gly Ala Ala Cys Thr Gly Gly Gly Thr Cys Cys		
	100 105 110		
	Gly Cys Cys Ala Gly Gly Cys Thr Cys Cys Ala Gly Gly Ala Ala Ala		
	115 120 125		
	Gly Gly Gly Gly Cys Thr Cys Gly Ala Gly Thr Gly Gly Gly Thr Cys		
	130 135 140		
	Thr Cys Ala Gly Ala Thr Ala Thr Thr Thr Cys Thr Cys Ala Gly Ala		
	145 150 155 160		
	Gly Thr Gly Gly Thr Gly Cys Thr Ala Gly Cys Ala Thr Ala Ala Gly		
	165 170 175		
	Thr Thr Ala Cys Ala Cys Ala Gly Gly Cys Thr Cys Cys Gly Thr Gly		
	180 185 190		
	Ala Ala Gly Gly Gly Cys Cys Gly Ala Thr Thr Cys Ala Cys Cys Ala		
	195 200 205		
	Thr Cys Thr Cys Cys Ala Gly Ala Gly Ala Cys Ala Ala Cys Gly Cys		
	210 215 220		
	Cys Ala Ala Gly Ala Ala Cys Ala Cys Gly Cys Thr Gly Thr Ala Thr		
	225 230 235 240		
	Cys Thr Ala Cys Ala Ala Ala Thr Gly Ala Ala Cys Ala Gly Cys Cys		
	245 250 255		
	Thr Gly Ala Ala Gly Cys Cys Thr Gly Ala Gly Gly Ala Cys Ala Cys		

```

                260                265                270
Gly Gly Cys Cys Gly Thr Cys Thr Ala Thr Thr Ala Cys Thr Gly Thr
   275                280                285
Gly Cys Cys Ala Gly Ala Thr Gly Thr Cys Cys Gly Gly Cys Cys Cys
   290                295                300
Cys Ala Thr Thr Cys Ala Cys Gly Ala Gly Ala Gly Ala Thr Thr Gly
   305                310                315                320
Thr Thr Thr Thr Gly Ala Cys Gly Thr Gly Ala Cys Thr Ala Gly Thr
   325                330                335
Ala Cys Cys Ala Cys Gly Thr Ala Thr Gly Cys Cys Thr Ala Cys Ala
   340                345                350
Gly Gly Gly Gly Cys Cys Ala Gly Gly Gly Gly Ala Cys Cys Cys Ala
   355                360                365
Gly Gly Thr Cys Ala Cys Cys Gly Thr Cys Thr Cys Cys Thr Cys Ala
   370                375                380
Cys Ala Cys Cys Ala Cys Cys Ala Thr Cys Ala Cys Cys Ala Thr Cys
   385                390                395                400
Ala Cys Gly Ala Ala Cys Cys Thr Gly Ala Ala Gly Cys Cys Thr Ala
   405                410                415

```

Gly

<210> 50
 <211> 417
 <212> PRT
 <213> 家羊驼

<400> 50

[0011]

```

Cys Ala Gly Gly Thr Gly Cys Ala Gly Cys Thr Gly Cys Ala Gly Gly
   1                5                10                15
Ala Gly Thr Cys Thr Gly Gly Ala Gly Gly Gly Gly Cys Thr Thr
   20                25                30
Gly Gly Thr Gly Cys Ala Gly Cys Cys Thr Gly Gly Gly Gly Gly Gly
   35                40                45
Thr Cys Thr Cys Thr Gly Ala Gly Ala Cys Thr Cys Thr Cys Cys Thr
   50                55                60
Gly Thr Gly Cys Ala Gly Cys Cys Thr Cys Thr Gly Gly Ala Thr Thr
   65                70                75                80
Cys Ala Cys Thr Thr Thr Cys Ala Gly Thr Ala Ala Cys Thr Ala Thr
   85                90                95
Gly Thr Cys Ala Thr Gly Ala Ala Cys Thr Gly Gly Gly Thr Cys Cys
  100                105                110
Gly Cys Cys Ala Gly Gly Cys Thr Cys Cys Ala Gly Gly Ala Ala Ala
  115                120                125
Gly Gly Gly Gly Cys Thr Cys Gly Ala Gly Thr Gly Gly Gly Thr Cys
  130                135                140
Thr Cys Ala Gly Ala Thr Ala Thr Thr Thr Cys Thr Ala Ala Thr Gly
  145                150                155                160
Gly Cys Gly Gly Thr Gly Gly Thr Ala Cys Cys Ala Cys Ala Ala Gly
  165                170                175
Thr Thr Ala Thr Gly Cys Ala Ala Gly Cys Thr Cys Cys Gly Thr Gly
  180                185                190
Ala Ala Gly Gly Gly Cys Cys Gly Ala Thr Thr Cys Ala Cys Cys Ala
  195                200                205

```

Thr Cys Thr Cys Cys Ala Gly Ala Gly Ala Cys Ala Ala Cys Gly Cys
 210 215 220
 Cys Ala Ala Gly Ala Ala Cys Ala Cys Gly Cys Thr Gly Thr Ala Thr
 225 230 235 240
 Cys Thr Gly Cys Ala Ala Ala Thr Gly Ala Ala Cys Gly Gly Cys Cys
 245 250 255
 Thr Gly Ala Ala Gly Cys Cys Thr Gly Cys Gly Gly Ala Cys Ala Cys
 260 265 270
 Gly Gly Cys Cys Gly Thr Cys Thr Ala Thr Thr Ala Cys Thr Gly Thr
 275 280 285
 Gly Cys Ala Ala Gly Ala Thr Gly Thr Cys Cys Gly Gly Cys Cys Cys
 290 295 300
 Cys Ala Thr Thr Cys Ala Cys Gly Ala Ala Cys Gly Ala Thr Thr Gly
 305 310 315 320
 Thr Ala Thr Gly Gly Ala Cys Ala Thr Ala Ala Cys Thr Ala Gly Thr
 325 330 335
 Ala Cys Cys Ala Cys Gly Thr Ala Thr Gly Cys Cys Thr Ala Cys Ala
 340 345 350
 Gly Gly Gly Gly Cys Cys Ala Gly Gly Gly Gly Ala Cys Cys Cys Ala
 355 360 365
 Gly Gly Thr Cys Ala Cys Cys Gly Thr Cys Thr Cys Cys Thr Cys Ala
 370 375 380
 Cys Ala Cys Cys Ala Cys Cys Ala Thr Cys Ala Cys Cys Ala Thr Cys
 385 390 395 400
 Ala Cys Gly Ala Ala Cys Cys Thr Gly Ala Ala Gly Cys Cys Thr Ala
 405 410 415

[0012]

Gly

<210> 51
 <211> 390
 <212> PRT
 <213> 家羊驼

<400> 51

Cys Ala Gly Gly Thr Gly Cys Ala Gly Cys Thr Gly Cys Ala Gly Gly
 1 5 10 15
 Ala Gly Thr Cys Thr Gly Gly Ala Gly Gly Ala Gly Gly Cys Thr Cys
 20 25 30
 Gly Gly Thr Gly Cys Ala Gly Gly Cys Thr Gly Gly Gly Gly Gly Gly
 35 40 45
 Thr Cys Thr Cys Thr Gly Ala Gly Ala Cys Thr Cys Thr Cys Cys Thr
 50 55 60
 Gly Thr Ala Cys Ala Gly Thr Cys Thr Cys Thr Gly Gly Ala Ala Cys
 65 70 75 80
 Cys Ala Thr Cys Thr Thr Cys Ala Gly Thr Gly Thr Cys Ala Cys Thr
 85 90 95
 Gly Thr Cys Ala Thr Gly Gly Gly Cys Thr Gly Gly Thr Ala Cys Cys
 100 105 110
 Gly Cys Cys Ala Gly Gly Cys Thr Cys Cys Ala Gly Gly Gly Ala Ala
 115 120 125
 Gly Cys Ala Gly Cys Gly Cys Gly Ala Gly Thr Thr Gly Gly Thr Cys
 130 135 140
 Gly Cys Ala Gly Gly Thr Thr Thr Thr Ala Cys Thr Ala Ala Thr Ala
 145 150 155 160

[0013]

Cys Thr Ala Gly Ala Ala Ala Cys Ala Cys Ala Ala Ala Cys Thr Ala
 165 170 175
 Thr Gly Thr Ala Gly Ala Cys Thr Cys Cys Gly Thr Gly Ala Ala Gly
 180 185 190
 Gly Gly Cys Cys Gly Cys Thr Thr Cys Ala Cys Cys Ala Thr Cys Thr
 195 200 205
 Cys Cys Ala Ala Ala Gly Ala Cys Ala Gly Cys Gly Cys Cys Ala Ala
 210 215 220
 Gly Ala Ala Cys Ala Cys Gly Ala Thr Gly Thr Ala Thr Cys Thr Ala
 225 230 235 240
 Cys Ala Ala Ala Thr Gly Ala Ala Cys Ala Gly Cys Cys Thr Gly Ala
 245 250 255
 Ala Ala Cys Cys Thr Gly Ala Gly Gly Ala Cys Ala Cys Ala Gly Cys
 260 265 270
 Cys Gly Thr Cys Thr Ala Thr Thr Ala Cys Thr Gly Thr Ala Ala Thr
 275 280 285
 Gly Thr Ala Cys Gly Thr Cys Gly Gly Thr Gly Gly Gly Gly Cys Gly
 290 295 300
 Gly Thr Ala Cys Gly Ala Ala Thr Thr Gly Gly Ala Ala Thr Gly Ala
 305 310 315 320
 Cys Thr Ala Cys Thr Gly Gly Gly Gly Cys Cys Ala Gly Gly Gly Gly
 325 330 335
 Ala Cys Cys Cys Ala Gly Gly Thr Cys Ala Cys Cys Gly Thr Cys Thr
 340 345 350
 Cys Cys Thr Cys Ala Cys Ala Cys Cys Ala Cys Cys Ala Thr Cys Ala
 355 360 365
 Cys Cys Ala Thr Cys Ala Cys Gly Ala Ala Cys Cys Thr Gly Ala Ala
 370 375 380
 Gly Cys Cys Thr Ala Gly
 385 390
 <210> 52
 <211> 390
 <212> PRT
 <213> 家羊驼
 <400> 52
 Cys Ala Gly Gly Thr Gly Cys Ala Gly Cys Thr Gly Cys Ala Gly Gly
 1 5 10 15
 Ala Gly Thr Cys Thr Gly Gly Ala Gly Gly Gly Gly Cys Thr Thr
 20 25 30
 Cys Gly Thr Gly Cys Ala Gly Gly Cys Thr Gly Gly Gly Gly Gly Gly
 35 40 45
 Thr Cys Thr Cys Thr Gly Ala Gly Ala Cys Thr Cys Thr Cys Cys Thr
 50 55 60
 Gly Thr Gly Cys Ala Gly Cys Cys Thr Cys Thr Gly Gly Ala Ala Gly
 65 70 75 80
 Cys Ala Thr Cys Thr Thr Cys Ala Gly Thr Ala Ala Gly Ala Ala Thr
 85 90 95
 Ala Cys Cys Ala Thr Gly Gly Cys Cys Thr Gly Gly Thr Thr Cys Cys
 100 105 110
 Gly Cys Cys Ala Gly Gly Cys Thr Cys Cys Ala Gly Gly Gly Ala Ala
 115 120 125
 Gly Gly Ala Gly Cys Gly Ala Gly Ala Gly Thr Thr Gly Gly Thr Cys

130	135	140
Gly Cys Ala Gly Cys Thr Ala Gly Thr Cys Cys Thr Ala Cys Gly Gly 145	150	155 160
Gly Thr Gly Gly Thr Ala Gly Cys Ala Cys Ala Gly Cys Gly Thr Ala 165	170	175
Thr Ala Ala Ala Gly Ala Cys Thr Cys Cys Gly Thr Gly Ala Ala Gly 180	185	190
Gly Gly Cys Cys Gly Ala Thr Thr Cys Ala Cys Cys Ala Thr Cys Thr 195	200	205
Cys Cys Ala Gly Ala Gly Ala Cys Ala Gly Cys Gly Cys Cys Ala Ala 210	215	220
Gly Ala Ala Cys Ala Cys Gly Gly Thr Gly Thr Thr Gly Cys Thr Gly 225	230	235 240
Cys Ala Ala Ala Thr Gly Ala Ala Cys Gly Thr Cys Cys Thr Gly Ala 245	250	255
Ala Ala Cys Cys Thr Gly Ala Gly Gly Ala Thr Ala Cys Thr Gly Cys 260	265	270
Cys Gly Thr Cys Thr Ala Thr Thr Ala Cys Thr Gly Thr Cys Ala Thr 275	280	285
Cys Thr Ala Cys Gly Thr Cys Ala Ala Ala Ala Thr Ala Ala Cys Cys 290	295	300
Gly Thr Gly Gly Thr Thr Cys Thr Thr Gly Gly Thr Thr Cys Cys Ala 305	310	315 320
Cys Thr Ala Cys Thr Gly Gly Gly Gly Cys Cys Ala Gly Gly Gly Gly 325	330	335
[0014] Ala Cys Cys Cys Ala Gly Gly Thr Cys Ala Cys Cys Gly Thr Cys Thr 340	345	350
Cys Cys Thr Cys Ala Cys Ala Cys Cys Ala Cys Cys Ala Thr Cys Ala 355	360	365
Cys Cys Ala Thr Cys Ala Cys Gly Ala Ala Cys Cys Thr Gly Ala Ala 370	375	380
Gly Cys Cys Thr Ala Gly 385	390	
<210> 53		
<211> 390		
<212> PRT		
<213> 家羊驼		
<400> 53		
Cys Ala Gly Gly Thr Gly Cys Ala Gly Cys Thr Gly Cys Ala Gly Gly 1	5	10 15
Ala Gly Thr Cys Thr Gly Gly Gly Gly Ala Gly Gly Cys Thr Thr 20	25	30
Gly Gly Thr Gly Cys Ala Gly Gly Cys Thr Gly Gly Gly Gly Gly Gly 35	40	45
Thr Cys Gly Cys Thr Gly Ala Gly Ala Cys Thr Cys Thr Cys Thr Thr 50	55	60
Gly Thr Gly Cys Ala Gly Thr Cys Thr Cys Thr Gly Gly Thr Ala Cys 65	70	75 80
Gly Ala Thr Cys Thr Thr Cys Gly Ala Thr Ala Thr Cys Ala Cys Thr 85	90	95
Cys Cys Cys Ala Thr Gly Gly Gly Cys Thr Gly Gly Thr Ala Cys Cys 100	105	110

Gly Cys Cys Ala Gly Ala Cys Thr Cys Cys Ala Gly Gly Gly Ala Ala
 115 120 125
 Gly Cys Ala Gly Cys Gly Cys Gly Ala Ala Gly Thr Gly Gly Thr Cys
 130 135 140
 Gly Cys Ala Gly Ala Thr Cys Thr Thr Ala Cys Thr Ala Gly Thr Cys
 145 150 155 160
 Gly Cys Gly Gly Thr Ala Cys Cys Ala Cys Ala Ala Thr Thr Ala
 165 170 175
 Cys Gly Cys Ala Gly Ala Cys Thr Cys Cys Gly Thr Gly Ala Ala Gly
 180 185 190
 Gly Gly Cys Cys Gly Gly Thr Thr Cys Ala Cys Cys Ala Thr Cys Thr
 195 200 205
 Cys Cys Ala Gly Ala Gly Ala Cys Ala Ala Cys Gly Cys Cys Ala Ala
 210 215 220
 Gly Ala Ala Ala Ala Thr Gly Thr Thr Gly Thr Ala Thr Cys Thr Gly
 225 230 235 240
 Cys Ala Ala Ala Thr Gly Ala Ala Cys Ala Gly Cys Cys Thr Gly Ala
 245 250 255
 Ala Ala Thr Cys Thr Gly Ala Cys Gly Ala Cys Ala Cys Ala Gly Gly
 260 265 270
 Cys Gly Thr Gly Thr Ala Thr Thr Ala Cys Thr Gly Thr Ala Ala Cys
 275 280 285
 Gly Thr Gly Ala Ala Ala Cys Gly Gly Thr Gly Gly Gly Gly Ala Gly
 290 295 300
 Gly Thr Ala Thr Thr Gly Gly Cys Thr Gly Gly Ala Ala Cys Gly Ala
 305 310 315 320
 Cys Thr Ala Cys Thr Gly Gly Gly Gly Cys Cys Ala Gly Gly Gly Gly
 325 330 335
 Ala Cys Cys Cys Ala Gly Gly Thr Cys Ala Cys Cys Gly Thr Cys Thr
 340 345 350
 Cys Cys Thr Cys Ala Cys Ala Cys Cys Ala Cys Cys Ala Thr Cys Ala
 355 360 365
 Cys Cys Ala Thr Cys Ala Cys Gly Ala Ala Cys Cys Thr Gly Ala Ala
 370 375 380
 Gly Cys Cys Thr Ala Gly
 385 390

[0015]

<210> 54
 <211> 390
 <212> PRT
 <213> 家羊驼

<400> 54

Cys Ala Gly Gly Thr Gly Cys Ala Gly Cys Thr Gly Cys Ala Gly Gly
 1 5 10 15
 Ala Gly Thr Cys Thr Gly Gly Ala Gly Gly Ala Gly Cys Thr Thr
 20 25 30
 Gly Gly Thr Gly Cys Ala Gly Thr Cys Thr Gly Gly Gly Gly Gly Gly
 35 40 45
 Thr Cys Thr Cys Thr Gly Ala Gly Ala Cys Thr Cys Thr Cys Cys Thr
 50 55 60
 Gly Thr Gly Thr Ala Gly Cys Cys Thr Cys Thr Gly Gly Ala Thr Thr
 65 70 75 80

[0016]

Cys Ala Gly Ala Thr Thr Cys Ala Gly Thr Ala Ala Cys Thr Thr Thr
 85 90 95
 Cys Cys Thr Ala Thr Gly Ala Thr Gly Thr Gly Gly Gly Thr Cys Cys
 100 105 110
 Gly Cys Cys Ala Gly Gly Cys Cys Cys Cys Ala Gly Gly Ala Ala Ala
 115 120 125
 Gly Gly Gly Gly Cys Thr Cys Gly Ala Gly Thr Gly Gly Gly Thr Cys
 130 135 140
 Thr Cys Gly Cys Thr Gly Ala Thr Thr Ala Gly Cys Ala Thr Thr Gly
 145 150 155 160
 Gly Thr Gly Gly Thr Ala Gly Thr Ala Cys Cys Ala Cys Gly Ala Ala
 165 170 175
 Thr Thr Ala Thr Gly Cys Gly Gly Ala Cys Thr Cys Cys Gly Thr Gly
 180 185 190
 Ala Ala Gly Gly Gly Cys Cys Gly Ala Thr Thr Cys Ala Cys Cys Ala
 195 200 205
 Thr Cys Thr Cys Cys Ala Gly Ala Gly Ala Cys Ala Ala Cys Gly Cys
 210 215 220
 Cys Ala Ala Gly Ala Ala Cys Ala Cys Gly Cys Thr Gly Thr Thr Thr
 225 230 235 240
 Cys Thr Gly Cys Ala Ala Ala Thr Gly Ala Ala Cys Ala Gly Cys Cys
 245 250 255
 Thr Gly Ala Ala Ala Cys Cys Thr Gly Ala Gly Gly Ala Cys Ala Cys
 260 265 270
 Gly Gly Cys Cys Gly Thr Gly Thr Ala Thr Thr Ala Cys Thr Gly Thr
 275 280 285
 Gly Cys Ala Ala Ala Ala Thr Ala Thr Cys Thr Thr Gly Gly Thr Cys
 290 295 300
 Gly Gly Cys Thr Gly Gly Thr Cys Cys Cys Ala Cys Cys Gly Ala Cys
 305 310 315 320
 Thr Ala Cys Thr Gly Ala Gly Gly Gly Cys Cys Ala Gly Gly Gly Gly
 325 330 335
 Ala Cys Cys Cys Ala Gly Gly Thr Cys Ala Cys Cys Gly Thr Cys Thr
 340 345 350
 Cys Cys Thr Cys Ala Cys Ala Cys Cys Ala Cys Cys Ala Thr Cys Ala
 355 360 365
 Cys Cys Ala Thr Cys Ala Cys Gly Ala Ala Cys Cys Thr Gly Ala Ala
 370 375 380
 Gly Cys Cys Thr Ala Gly
 385 390
 <210> 55
 <211> 1095
 <212> PRT
 <213> 人工
 <220>
 <223> Beta2AR构建体
 <400> 55
 Gly Gly Gly Cys Ala Ala Cys Cys Cys Gly Gly Gly Ala Ala Cys Gly
 1 5 10 15
 Gly Cys Ala Gly Cys Gly Cys Cys Thr Thr Cys Thr Thr Gly Cys Thr
 20 25 30
 Gly Gly Cys Ala Cys Cys Cys Ala Ala Thr Ala Gly Ala Ala Gly Cys
 35 40 45

[0017]

Cys Ala Thr Gly Cys Gly Cys Cys Gly Gly Ala Cys Cys Ala Cys Gly
 50 55 60
 Ala Cys Gly Thr Cys Ala Cys Gly Cys Ala Gly Cys Ala Ala Ala Gly
 65 70 75 80
 Gly Gly Ala Cys Gly Ala Gly Gly Thr Gly Thr Gly Gly Gly Thr Gly
 85 90 95
 Gly Thr Gly Gly Gly Cys Ala Thr Gly Gly Gly Cys Ala Thr Cys Gly
 100 105 110
 Thr Cys Ala Thr Gly Thr Cys Thr Cys Thr Cys Ala Thr Cys Gly Thr
 115 120 125
 Cys Cys Thr Gly Gly Cys Cys Ala Thr Cys Gly Thr Gly Thr Thr Thr
 130 135 140
 Gly Gly Cys Ala Ala Thr Gly Thr Gly Cys Thr Gly Gly Thr Cys Ala
 145 150 155 160
 Thr Cys Ala Cys Ala Gly Cys Cys Ala Thr Thr Gly Cys Cys Ala Ala
 165 170 175
 Gly Thr Thr Cys Gly Ala Gly Cys Gly Thr Cys Thr Gly Cys Ala Gly
 180 185 190
 Ala Cys Gly Gly Thr Cys Ala Cys Cys Ala Ala Cys Thr Ala Cys Thr
 195 200 205
 Thr Cys Ala Thr Cys Ala Cys Thr Thr Cys Ala Cys Thr Gly Gly Cys
 210 215 220
 Cys Thr Gly Thr Gly Cys Thr Gly Ala Thr Cys Thr Gly Gly Thr Cys
 225 230 235 240
 Ala Thr Gly Gly Gly Cys Cys Thr Gly Gly Cys Ala Gly Thr Gly Gly
 245 250 255
 Thr Gly Cys Cys Cys Thr Thr Thr Gly Gly Gly Gly Cys Cys Gly Cys
 260 265 270
 Cys Cys Ala Thr Ala Thr Thr Cys Thr Thr Ala Thr Gly Ala Ala Ala
 275 280 285
 Ala Thr Gly Thr Gly Gly Ala Cys Thr Thr Thr Thr Gly Gly Cys Ala
 290 295 300
 Ala Cys Thr Thr Cys Thr Gly Gly Thr Gly Cys Gly Ala Gly Thr Thr
 305 310 315 320
 Thr Thr Gly Gly Ala Cys Thr Thr Cys Cys Ala Thr Thr Gly Ala Thr
 325 330 335
 Gly Thr Gly Cys Thr Gly Thr Gly Cys Gly Thr Cys Ala Cys Gly Gly
 340 345 350
 Cys Cys Ala Gly Cys Ala Thr Thr Gly Ala Gly Ala Cys Cys Cys Thr
 355 360 365
 Gly Thr Gly Cys Gly Thr Gly Ala Thr Cys Gly Cys Ala Gly Thr Gly
 370 375 380
 Gly Ala Thr Cys Gly Cys Thr Ala Cys Thr Thr Thr Gly Cys Cys Ala
 385 390 395 400
 Thr Thr Ala Cys Thr Thr Cys Ala Cys Cys Thr Thr Thr Cys Ala Ala
 405 410 415
 Gly Thr Ala Cys Cys Ala Gly Ala Gly Cys Cys Thr Gly Cys Thr Gly
 420 425 430
 Ala Cys Cys Ala Ala Gly Ala Ala Thr Ala Ala Gly Gly Cys Cys Cys
 435 440 445
 Gly Gly Gly Thr Gly Ala Thr Cys Ala Thr Thr Cys Thr Gly Ala Thr
 450 455 460

[0018]

Gly Gly Thr Gly Thr Gly Gly Ala Thr Thr Gly Thr Gly Thr Cys Ala
 465 470 475 480
 Gly Gly Cys Cys Thr Thr Ala Cys Cys Thr Cys Cys Thr Thr Cys Thr
 485 490 495
 Thr Gly Cys Cys Cys Ala Thr Thr Cys Ala Gly Ala Thr Gly Cys Ala
 500 505 510
 Cys Thr Gly Gly Thr Ala Cys Cys Gly Gly Gly Cys Cys Ala Cys Cys
 515 520 525
 Cys Ala Cys Cys Ala Gly Gly Ala Ala Gly Cys Cys Ala Thr Cys Ala
 530 535 540
 Ala Cys Thr Gly Cys Thr Ala Thr Gly Cys Cys Gly Ala Gly Gly Ala
 545 550 555 560
 Gly Ala Cys Cys Thr Gly Cys Thr Gly Thr Gly Ala Cys Thr Thr Cys
 565 570 575
 Thr Thr Cys Ala Cys Gly Ala Ala Cys Cys Ala Ala Gly Cys Cys Thr
 580 585 590
 Ala Thr Gly Cys Cys Ala Thr Thr Gly Cys Cys Thr Cys Thr Thr Cys
 595 600 605
 Cys Ala Thr Cys Gly Thr Gly Thr Cys Cys Thr Thr Cys Thr Ala Cys
 610 615 620
 Gly Thr Thr Cys Cys Cys Cys Thr Gly Gly Thr Gly Ala Thr Cys Ala
 625 630 635 640
 Thr Gly Gly Thr Cys Thr Thr Cys Gly Thr Cys Thr Ala Cys Thr Cys
 645 650 655
 Cys Ala Gly Gly Gly Thr Cys Thr Thr Thr Cys Ala Gly Gly Ala Gly
 660 665 670
 Gly Cys Cys Ala Ala Ala Ala Gly Gly Cys Ala Gly Cys Thr Cys Cys
 675 680 685
 Ala Gly Ala Ala Gly Ala Thr Thr Gly Ala Cys Ala Ala Ala Thr Cys
 690 695 700
 Thr Gly Ala Gly Gly Gly Cys Cys Gly Cys Thr Thr Cys Cys Ala Thr
 705 710 715 720
 Gly Thr Cys Cys Ala Gly Ala Ala Cys Cys Thr Thr Ala Gly Cys Cys
 725 730 735
 Ala Gly Gly Thr Gly Gly Ala Gly Cys Ala Gly Gly Ala Thr Gly Gly
 740 745 750
 Gly Cys Gly Gly Ala Cys Gly Gly Gly Gly Cys Ala Thr Gly Gly Ala
 755 760 765
 Cys Thr Cys Cys Gly Cys Ala Gly Ala Thr Cys Thr Thr Cys Cys Ala
 770 775 780
 Ala Gly Thr Thr Cys Thr Gly Cys Thr Thr Gly Ala Ala Gly Gly Ala
 785 790 795 800
 Gly Cys Ala Cys Ala Ala Ala Gly Cys Cys Cys Thr Cys Ala Ala Gly
 805 810 815
 Ala Cys Gly Thr Thr Ala Gly Gly Cys Ala Thr Cys Ala Thr Cys Ala
 820 825 830
 Thr Gly Gly Gly Cys Ala Cys Thr Thr Thr Cys Ala Cys Cys Cys Thr
 835 840 845
 Cys Thr Gly Cys Thr Gly Gly Cys Thr Gly Cys Cys Cys Thr Thr Cys
 850 855 860
 Thr Thr Cys Ala Thr Cys Gly Thr Thr Ala Ala Cys Ala Thr Thr Gly
 865 870 875 880

[0019]

Thr Gly Cys Ala Thr Gly Thr Gly Ala Thr Cys Cys Ala Gly Gly Ala
 885 890 895
 Thr Ala Ala Cys Cys Thr Cys Ala Thr Cys Cys Gly Thr Ala Ala Gly
 900 905 910
 Gly Ala Ala Gly Thr Thr Thr Ala Cys Ala Thr Cys Cys Thr Cys Cys
 915 920 925
 Thr Ala Ala Ala Thr Thr Gly Gly Ala Thr Ala Gly Gly Cys Thr Ala
 930 935 940
 Thr Gly Thr Cys Ala Ala Thr Thr Cys Thr Gly Gly Thr Thr Thr Cys
 945 950 955 960
 Ala Ala Thr Cys Cys Cys Cys Thr Thr Ala Thr Cys Thr Ala Cys Thr
 965 970 975
 Gly Cys Cys Gly Gly Ala Gly Cys Cys Cys Ala Gly Ala Thr Thr Thr
 980 985 990
 Cys Ala Gly Gly Ala Thr Thr Gly Cys Cys Thr Thr Cys Cys Ala Gly
 995 1000 1005
 Gly Ala Gly Cys Thr Thr Cys Thr Gly Thr Gly Cys Cys Thr Gly
 1010 1015 1020
 Cys Gly Cys Ala Gly Gly Thr Cys Thr Thr Cys Thr Thr Thr Gly
 1025 1030 1035
 Ala Ala Gly Gly Cys Cys Thr Ala Thr Gly Gly Gly Ala Ala Thr
 1040 1045 1050
 Gly Gly Cys Thr Ala Cys Thr Cys Cys Ala Gly Cys Ala Ala Cys
 1055 1060 1065
 Gly Gly Cys Ala Ala Cys Ala Cys Ala Gly Gly Gly Gly Ala Gly
 1070 1075 1080
 Cys Ala Gly Ala Gly Thr Gly Gly Ala Thr Ala Ala
 1085 1090 1095
 <210> 56
 <211> 380
 <212> PRT
 <213> 人(Homo sapiens)
 <400> 56
 Met Gly Cys Leu Gly Asn Ser Lys Thr Glu Asp Gln Arg Asn Glu Glu
 1 5 10 15
 Lys Ala Gln Arg Glu Ala Asn Lys Lys Ile Glu Lys Gln Leu Gln Lys
 20 25 30
 Asp Lys Gln Val Tyr Arg Ala Thr His Arg Leu Leu Leu Leu Gly Ala
 35 40 45
 Gly Glu Ser Gly Lys Ser Thr Ile Val Lys Gln Met Arg Ile Leu His
 50 55 60
 Val Asn Gly Phe Asn Gly Asp Ser Glu Lys Ala Thr Lys Val Gln Asp
 65 70 75 80
 Ile Lys Asn Asn Leu Lys Glu Ala Ile Glu Thr Ile Val Ala Ala Met
 85 90 95
 Ser Asn Leu Val Pro Pro Val Glu Leu Ala Asn Pro Glu Asn Gln Phe
 100 105 110
 Arg Val Asp Tyr Ile Leu Ser Val Met Asn Val Pro Asp Phe Asp Phe
 115 120 125
 Pro Pro Glu Phe Tyr Glu His Ala Lys Ala Leu Trp Glu Asp Glu Gly
 130 135 140
 Val Arg Ala Cys Tyr Glu Arg Ser Asn Glu Tyr Gln Leu Ile Asp Cys

	145	150	155	160
	Ala Gln Tyr Phe	Leu Asp Lys Ile Asp Val Ile Lys Gln Ala Asp Tyr		
		165	170	175
	Val Pro Ser Asp Gln Asp Leu Leu Arg Cys Arg Val Leu Thr Ser Gly			
		180	185	190
	Ile Phe Glu Thr Lys Phe Gln Val Asp Lys Val Asn Phe His Met Phe			
		195	200	205
	Asp Val Gly Gly Gln Arg Asp Glu Arg Arg Lys Trp Ile Gln Cys Phe			
		210	215	220
	Asn Asp Val Thr Ala Ile Ile Phe Val Val Ala Ser Ser Ser Tyr Asn			
		225	230	235
	Met Val Ile Arg Glu Asp Asn Gln Thr Asn Arg Leu Gln Glu Ala Leu			
		245	250	255
	Asn Leu Phe Lys Ser Ile Trp Asn Asn Arg Trp Leu Arg Thr Ile Ser			
		260	265	270
	Val Ile Leu Phe Leu Asn Lys Gln Asp Leu Leu Ala Glu Lys Val Leu			
		275	280	285
	Ala Gly Lys Ser Lys Ile Glu Asp Tyr Phe Pro Glu Phe Ala Arg Tyr			
		290	295	300
	Thr Thr Pro Glu Asp Ala Thr Pro Glu Pro Gly Glu Asp Pro Arg Val			
		305	310	315
	Thr Arg Ala Lys Tyr Phe Ile Arg Asp Glu Phe Leu Arg Ile Ser Thr			
		325	330	335
	Ala Ser Gly Asp Gly Arg His Tyr Cys Tyr Pro His Phe Thr Cys Ala			
		340	345	350
[0020]	Val Asp Thr Glu Asn Ile Arg Arg Val Phe Asn Asp Cys Arg Asp Ile			
		355	360	365
	Ile Gln Arg Met His Leu Arg Gln Tyr Glu Leu Leu			
		370	375	380
	<210>	57		
	<211>	354		
	<212>	PRT		
	<213>	人		
	<400>	57		
	Met Gly Cys Thr Leu Ser Ala Glu Asp Lys Ala Ala Val Glu Arg Ser			
		1	5	10
	Lys Met Ile Asp Arg Asn Leu Arg Glu Asp Gly Glu Lys Ala Ala Arg			
		20	25	30
	Glu Val Lys Leu Leu Leu Leu Gly Ala Gly Glu Ser Gly Lys Ser Thr			
		35	40	45
	Ile Val Lys Gln Met Lys Ile Ile His Glu Ala Gly Tyr Ser Glu Glu			
		50	55	60
	Glu Cys Lys Gln Tyr Lys Ala Val Val Tyr Ser Asn Thr Ile Gln Ser			
		65	70	75
	Ile Ile Ala Ile Ile Arg Ala Met Gly Arg Leu Lys Ile Asp Phe Gly			
		85	90	95
	Asp Ser Ala Arg Ala Asp Asp Ala Arg Gln Leu Phe Val Leu Ala Gly			
		100	105	110
	Ala Ala Glu Glu Gly Phe Met Thr Ala Glu Leu Ala Gly Val Ile Lys			
		115	120	125
	Arg Leu Trp Lys Asp Ser Gly Val Gln Ala Cys Phe Asn Arg Ser Arg			
		130	135	140

Glu Tyr Gln Leu Asn Asp Ser Ala Ala Tyr Tyr Leu Asn Asp Leu Asp
 145 150 155 160
 Arg Ile Ala Gln Pro Asn Tyr Ile Pro Thr Gln Gln Asp Val Leu Arg
 165 170 175
 Thr Arg Val Lys Thr Thr Gly Ile Val Glu Thr His Phe Thr Phe Lys
 180 185 190
 Asp Leu His Phe Lys Met Phe Asp Val Gly Gly Gln Arg Ser Glu Arg
 195 200 205
 Lys Lys Trp Ile His Cys Phe Glu Gly Val Thr Ala Ile Ile Phe Cys
 210 215 220
 Val Ala Leu Ser Asp Tyr Asp Leu Val Leu Ala Glu Asp Glu Glu Met
 225 230 235 240
 Asn Arg Met His Glu Ser Met Lys Leu Phe Asp Ser Ile Cys Asn Asn
 245 250 255
 Lys Trp Phe Thr Asp Thr Ser Ile Ile Leu Phe Leu Asn Lys Lys Asp
 260 265 270
 Leu Phe Glu Glu Lys Ile Lys Lys Ser Pro Leu Thr Ile Cys Tyr Pro
 275 280 285
 Glu Tyr Ala Gly Ser Asn Thr Tyr Glu Glu Ala Ala Ala Tyr Ile Gln
 290 295 300
 Cys Gln Phe Glu Asp Leu Asn Lys Arg Lys Asp Thr Lys Glu Ile Tyr
 305 310 315 320
 Thr His Phe Thr Cys Ala Thr Asp Thr Lys Asn Val Gln Phe Val Phe
 325 330 335
 Asp Ala Val Thr Asp Val Ile Ile Lys Asn Asn Leu Lys Asp Cys Gly
 340 345 350

[0021]

Leu Phe

<210> 58
 <211> 350
 <212> PRT
 <213> 人

<400> 58

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

Pro Gly Tyr Val Pro Thr Glu Gln Asp Val Leu Arg Ser Arg Val Lys
 165 170 175
 Thr Thr Gly Ile Ile Glu Thr Gln Phe Ser Phe Lys Asp Leu Asn Phe
 180 185 190
 Arg Met Phe Asp Val Gly Gly Gln Arg Ser Glu Arg Lys Lys Trp Ile
 195 200 205
 His Cys Phe Glu Gly Val Thr Cys Ile Ile Phe Ile Ala Ala Leu Ser
 210 215 220
 Ala Tyr Asp Met Val Leu Val Glu Asp Asp Glu Val Asn Arg Met His
 225 230 235 240
 Glu Ser Leu His Leu Phe Asn Ser Ile Cys Asn His Arg Tyr Phe Ala
 245 250 255
 Thr Thr Ser Ile Val Leu Phe Leu Asn Lys Lys Asp Val Phe Phe Glu
 260 265 270
 Lys Ile Lys Lys Ala His Leu Ser Ile Cys Phe Pro Asp Tyr Asp Gly
 275 280 285
 Pro Asn Thr Tyr Glu Asp Ala Gly Asn Tyr Ile Lys Val Gln Phe Leu
 290 295 300
 Glu Leu Asn Met Arg Arg Asp Val Lys Glu Ile Tyr Ser His Met Thr
 305 310 315 320
 Cys Ala Thr Asp Thr Gln Asn Val Lys Phe Val Phe Asp Ala Val Thr
 325 330 335
 Asp Ile Ile Ile Lys Glu Asn Leu Lys Asp Cys Gly Leu Phe
 340 345 350

<210> 59
 <211> 394
 <212> PRT
 <213> 欧洲牛(Bos taurus)

<400> 59

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

	180	185	190
Ser Asp Gln Asp Leu Leu Arg Cys Arg Val Leu Thr Ser Gly Ile Phe			
195	200	205	
Glu Thr Lys Phe Gln Val Asp Lys Val Asn Phe His Met Phe Asp Val			
210	215	220	
Gly Gly Gln Arg Asp Glu Arg Arg Lys Trp Ile Gln Cys Phe Asn Asp			
225	230	235	240
Val Thr Ala Ile Ile Phe Val Val Ala Ser Ser Ser Tyr Asn Met Val			
245	250	255	
Ile Arg Glu Asp Asn Gln Thr Asn Arg Leu Gln Glu Ala Leu Asn Leu			
260	265	270	
Phe Lys Ser Ile Trp Asn Asn Arg Trp Leu Arg Thr Ile Ser Val Ile			
275	280	285	
Leu Phe Leu Asn Lys Gln Asp Leu Leu Ala Glu Lys Val Leu Ala Gly			
290	295	300	
Lys Ser Lys Ile Glu Asp Tyr Phe Pro Glu Phe Ala Arg Tyr Thr Thr			
305	310	315	320
Pro Glu Asp Ala Thr Pro Glu Pro Gly Glu Asp Pro Arg Val Thr Arg			
325	330	335	
Ala Lys Tyr Phe Ile Arg Asp Glu Phe Leu Arg Ile Ser Thr Ala Ser			
340	345	350	
Gly Asp Gly Arg His Tyr Cys Tyr Pro His Phe Thr Cys Ala Val Asp			
355	360	365	
Thr Glu Asn Ile Arg Arg Val Phe Asn Asp Cys Arg Asp Ile Ile Gln			
370	375	380	

[0023]

Arg Met His Leu Arg Gln Tyr Glu Leu Leu
385 390

<210> 60
 <211> 394
 <212> PRT
 <213> 褐家鼠(Rattus norvegicus)

<400> 60

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

Ala Cys Tyr Glu Arg Ser Asn Glu Tyr Gln Leu Ile Asp Cys Ala Gln
 165 170 175
 Tyr Phe Leu Asp Lys Ile Asp Val Ile Lys Gln Ala Asp Tyr Val Pro
 180 185 190
 Ser Asp Gln Asp Leu Leu Arg Cys Arg Val Leu Thr Ser Gly Ile Phe
 195 200 205
 Glu Thr Lys Phe Gln Val Asp Lys Val Asn Phe His Met Phe Asp Val
 210 215 220
 Gly Gly Gln Arg Asp Glu Arg Arg Lys Trp Ile Gln Cys Phe Asn Asp
 225 230 235 240
 Val Thr Ala Ile Ile Phe Val Val Ala Ser Ser Ser Tyr Asn Met Val
 245 250 255
 Ile Arg Glu Asp Asn Gln Thr Asn Arg Leu Gln Glu Ala Leu Asn Leu
 260 265 270
 Phe Lys Ser Ile Trp Asn Asn Arg Trp Leu Arg Thr Ile Ser Val Ile
 275 280 285
 Leu Phe Leu Asn Lys Gln Asp Leu Leu Ala Glu Lys Val Leu Ala Gly
 290 295 300
 Lys Ser Lys Ile Glu Asp Tyr Phe Pro Glu Phe Ala Arg Tyr Thr Thr
 305 310 315 320
 Pro Glu Asp Ala Thr Pro Glu Pro Gly Glu Asp Pro Arg Val Thr Arg
 325 330 335
 Ala Lys Tyr Phe Ile Arg Asp Glu Phe Leu Arg Ile Ser Thr Ala Ser
 340 345 350
 Gly Asp Gly Arg His Tyr Cys Tyr Pro His Phe Thr Cys Ala Val Asp
 355 360 365
 Thr Glu Asn Ile Arg Arg Val Phe Asn Asp Cys Arg Asp Ile Ile Gln
 370 375 380
 Arg Met His Leu Arg Gln Tyr Glu Leu Leu
 385 390

[0024]

<210> 61
 <211> 394
 <212> PRT
 <213> 小鼠(Mus musculus)

<400> 61

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

130	135	140
Glu Phe Tyr Glu His Ala Lys Ala Leu Trp Glu Asp Glu Gly Val Arg 145	150	155
Ala Cys Tyr Glu Arg Ser Asn Glu Tyr Gln Leu Ile Asp Cys Ala Gln 165	170	175
Tyr Phe Leu Asp Lys Ile Asp Val Ile Lys Gln Ala Asp Tyr Val Pro 180	185	190
Ser Asp Gln Asp Leu Leu Arg Cys Arg Val Leu Thr Ser Gly Ile Phe 195	200	205
Glu Thr Lys Phe Gln Val Asp Lys Val Asn Phe His Met Phe Asp Val 210	215	220
Gly Gly Gln Arg Asp Glu Arg Arg Lys Trp Ile Gln Cys Phe Asn Asp 225	230	235
Val Thr Ala Ile Ile Phe Val Val Ala Ser Ser Ser Tyr Asn Met Val 245	250	255
Ile Arg Glu Asp Asn Gln Thr Asn Arg Leu Gln Glu Ala Leu Asn Leu 260	265	270
Phe Lys Ser Ile Trp Asn Asn Arg Trp Leu Arg Thr Ile Ser Val Ile 275	280	285
Leu Phe Leu Asn Lys Gln Asp Leu Leu Ala Glu Lys Val Leu Ala Gly 290	295	300
Lys Ser Lys Ile Glu Asp Tyr Phe Pro Glu Phe Ala Arg Tyr Thr Thr 305	310	315
Pro Glu Asp Ala Thr Pro Glu Pro Gly Glu Asp Pro Arg Val Thr Arg 325	330	335
[0025] Ala Lys Tyr Phe Ile Arg Asp Glu Phe Leu Arg Ile Ser Thr Ala Ser 340	345	350
Gly Asp Gly Arg His Tyr Cys Tyr Pro His Phe Thr Cys Ala Val Asp 355	360	365
Thr Glu Asn Ile Arg Arg Val Phe Asn Asp Cys Arg Asp Ile Ile Gln 370	375	380
Arg Met His Leu Arg Gln Tyr Glu Leu Leu 385	390	
<210> 62		
<211> 340		
<212> PRT		
<213> 欧洲牛		
<400> 62		
Met Ser Glu Leu Asp Gln Leu Arg Gln Glu Ala Glu Gln Leu Lys Asn 1	5	10
Gln Ile Arg Asp Ala Arg Lys Ala Cys Ala Asp Ala Thr Leu Ser Gln 20	25	30
Ile Thr Asn Asn Ile Asp Pro Val Gly Arg Ile Gln Met Arg Thr Arg 35	40	45
Arg Thr Leu Arg Gly His Leu Ala Lys Ile Tyr Ala Met His Trp Gly 50	55	60
Thr Asp Ser Arg Leu Leu Val Ser Ala Ser Gln Asp Gly Lys Leu Ile 65	70	75
Ile Trp Asp Ser Tyr Thr Thr Asn Lys Val His Ala Ile Pro Leu Arg 85	90	95
Ser Ser Trp Val Met Thr Cys Ala Tyr Ala Pro Ser Gly Asn Tyr Val 100	105	110

[0026]

Ala Cys Gly Gly Leu Asp Asn Ile Cys Ser Ile Tyr Asn Leu Lys Thr
 115 120 125
 Arg Glu Gly Asn Val Arg Val Ser Arg Glu Leu Ala Gly His Thr Gly
 130 135 140
 Tyr Leu Ser Cys Cys Arg Phe Leu Asp Asp Asn Gln Ile Val Thr Ser
 145 150 155 160
 Ser Gly Asp Thr Thr Cys Ala Leu Trp Asp Ile Glu Thr Gly Gln Gln
 165 170 175
 Thr Thr Thr Phe Thr Gly His Thr Gly Asp Val Met Ser Leu Ser Leu
 180 185 190
 Ala Pro Asp Thr Arg Leu Phe Val Ser Gly Ala Cys Asp Ala Ser Ala
 195 200 205
 Lys Leu Trp Asp Val Arg Glu Gly Met Cys Arg Gln Thr Phe Thr Gly
 210 215 220
 His Glu Ser Asp Ile Asn Ala Ile Cys Phe Phe Pro Asn Gly Asn Ala
 225 230 235 240
 Phe Ala Thr Gly Ser Asp Asp Ala Thr Cys Arg Leu Phe Asp Leu Arg
 245 250 255
 Ala Asp Gln Glu Leu Met Thr Tyr Ser His Asp Asn Ile Ile Cys Gly
 260 265 270
 Ile Thr Ser Val Ser Phe Ser Lys Ser Gly Arg Leu Leu Leu Ala Gly
 275 280 285
 Tyr Asp Asp Phe Asn Cys Asn Val Trp Asp Ala Leu Lys Ala Asp Arg
 290 295 300
 Ala Gly Val Leu Ala Gly His Asp Asn Arg Val Ser Cys Leu Gly Val
 305 310 315 320
 Thr Asp Asp Gly Met Ala Val Ala Thr Gly Ser Trp Asp Ser Phe Leu
 325 330 335
 Lys Ile Trp Asn
 340
 <210> 63
 <211> 340
 <212> PRT
 <213> 人
 <400> 63
 Met Ser Glu Leu Asp Gln Leu Arg Gln Glu Ala Glu Gln Leu Lys Asn
 1 5 10 15
 Gln Ile Arg Asp Ala Arg Lys Ala Cys Ala Asp Ala Thr Leu Ser Gln
 20 25 30
 Ile Thr Asn Asn Ile Asp Pro Val Gly Arg Ile Gln Met Arg Thr Arg
 35 40 45
 Arg Thr Leu Arg Gly His Leu Ala Lys Ile Tyr Ala Met His Trp Gly
 50 55 60
 Thr Asp Ser Arg Leu Leu Val Ser Ala Ser Gln Asp Gly Lys Leu Ile
 65 70 75 80
 Ile Trp Asp Ser Tyr Thr Thr Asn Lys Val His Ala Ile Pro Leu Arg
 85 90 95
 Ser Ser Trp Val Met Thr Cys Ala Tyr Ala Pro Ser Gly Asn Tyr Val
 100 105 110
 Ala Cys Gly Gly Leu Asp Asn Ile Cys Ser Ile Tyr Asn Leu Lys Thr
 115 120 125
 Arg Glu Gly Asn Val Arg Val Ser Arg Glu Leu Ala Gly His Thr Gly

130	135	140
Tyr Leu Ser Cys Cys Arg Phe Leu Asp Asp Asn Gln Ile Val Thr Ser		
145	150	155
Ser Gly Asp Thr Thr Cys Ala Leu Trp Asp Ile Glu Thr Gly Gln Gln		
165	170	175
Thr Thr Thr Phe Thr Gly His Thr Gly Asp Val Met Ser Leu Ser Leu		
180	185	190
Ala Pro Asp Thr Arg Leu Phe Val Ser Gly Ala Cys Asp Ala Ser Ala		
195	200	205
Lys Leu Trp Asp Val Arg Glu Gly Met Cys Arg Gln Thr Phe Thr Gly		
210	215	220
His Glu Ser Asp Ile Asn Ala Ile Cys Phe Phe Pro Asn Gly Asn Ala		
225	230	235
Phe Ala Thr Gly Ser Asp Asp Ala Thr Cys Arg Leu Phe Asp Leu Arg		
245	250	255
Ala Asp Gln Glu Leu Met Thr Tyr Ser His Asp Asn Ile Ile Cys Gly		
260	265	270
Ile Thr Ser Val Ser Phe Ser Lys Ser Gly Arg Leu Leu Leu Ala Gly		
275	280	285
Tyr Asp Asp Phe Asn Cys Asn Val Trp Asp Ala Leu Lys Ala Asp Arg		
290	295	300
Ala Gly Val Leu Ala Gly His Asp Asn Arg Val Ser Cys Leu Gly Val		
305	310	315
Thr Asp Asp Gly Met Ala Val Ala Thr Gly Ser Trp Asp Ser Phe Leu		
325	330	335

[0027]

Lys Ile Trp Asn
340

<210> 64
 <211> 340
 <212> PRT
 <213> 褐家鼠

<400> 64

Met Ser Glu Leu Asp Gln Leu Arg Gln Glu Ala Glu Gln Leu Lys Asn		
1	5	10
Gln Ile Arg Asp Ala Arg Lys Ala Cys Ala Asp Ala Thr Leu Ser Gln		
20	25	30
Ile Thr Asn Asn Ile Asp Pro Val Gly Arg Ile Gln Met Arg Thr Arg		
35	40	45
Arg Thr Leu Arg Gly His Leu Ala Lys Ile Tyr Ala Met His Trp Gly		
50	55	60
Thr Asp Ser Arg Leu Leu Val Ser Ala Ser Gln Asp Gly Lys Leu Ile		
65	70	75
Ile Trp Asp Ser Tyr Thr Thr Asn Lys Val His Ala Ile Pro Leu Arg		
85	90	95
Ser Ser Trp Val Met Thr Cys Ala Tyr Ala Pro Ser Gly Asn Tyr Val		
100	105	110
Ala Cys Gly Gly Leu Asp Asn Ile Cys Ser Ile Tyr Asn Leu Lys Thr		
115	120	125
Arg Glu Gly Asn Val Arg Val Ser Arg Glu Leu Ala Gly His Thr Gly		
130	135	140
Tyr Leu Ser Cys Cys Arg Phe Leu Asp Asp Asn Gln Ile Val Thr Ser		
145	150	155
		160

Ser Gly Asp Thr Thr Cys Ala Leu Trp Asp Ile Glu Thr Gly Gln Gln
 165 170 175
 Thr Thr Thr Phe Thr Gly His Thr Gly Asp Val Met Ser Leu Ser Leu
 180 185 190
 Ala Pro Asp Thr Arg Leu Phe Val Ser Gly Ala Cys Asp Ala Ser Ala
 195 200 205
 Lys Leu Trp Asp Val Arg Glu Gly Met Cys Arg Gln Thr Phe Thr Gly
 210 215 220
 His Glu Ser Asp Ile Asn Ala Ile Cys Phe Phe Pro Asn Gly Asn Ala
 225 230 235 240
 Phe Ala Thr Gly Ser Asp Asp Ala Thr Cys Arg Leu Phe Asp Leu Arg
 245 250 255
 Ala Asp Gln Glu Leu Met Thr Tyr Ser His Asp Asn Ile Ile Cys Gly
 260 265 270
 Ile Thr Ser Val Ser Phe Ser Lys Ser Gly Arg Leu Leu Leu Ala Gly
 275 280 285
 Tyr Asp Asp Phe Asn Cys Asn Val Trp Asp Ala Leu Lys Ala Asp Arg
 290 295 300
 Ala Gly Val Leu Ala Gly His Asp Asn Arg Val Ser Cys Leu Gly Val
 305 310 315 320
 Thr Asp Asp Gly Met Ala Val Ala Thr Gly Ser Trp Asp Ser Phe Leu
 325 330 335
 Lys Ile Trp Asn
 340

[0028]

<210> 65
 <211> 340
 <212> PRT
 <213> 小鼠

<400> 65

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

```

180      185      190
Ala Pro Asp Thr Arg Leu Phe Val Ser Gly Ala Cys Asp Ala Ser Ala
195      200      205
Lys Leu Trp Asp Val Arg Glu Gly Met Cys Arg Gln Thr Phe Thr Gly
210      215      220
His Glu Ser Asp Ile Asn Ala Ile Cys Phe Phe Pro Asn Gly Asn Ala
225      230      235      240
Phe Ala Thr Gly Ser Asp Asp Ala Thr Cys Arg Leu Phe Asp Leu Arg
245      250      255
Ala Asp Gln Glu Leu Met Thr Tyr Ser His Asp Asn Ile Ile Cys Gly
260      265      270
Ile Thr Ser Val Ser Phe Ser Lys Ser Gly Arg Leu Leu Leu Ala Gly
275      280      285
Tyr Asp Asp Phe Asn Cys Asn Val Trp Asp Ala Leu Lys Ala Asp Arg
290      295      300
Ala Gly Val Leu Ala Gly His Asp Asn Arg Val Ser Cys Leu Gly Val
305      310      315      320
Thr Asp Asp Gly Met Ala Val Ala Thr Gly Ser Trp Asp Ser Phe Leu
325      330      335
Lys Ile Trp Asn
340

```

<210> 66
 <211> 71
 <212> PRT
 <213> 欧洲牛

<400> 66

[0029]

```

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

```

<210> 67
 <211> 71
 <212> PRT
 <213> 小鼠

<400> 67

```

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

```

<210> 68
 <211> 71

<212> PRT

<213> 人

<400> 68

Met Ala Ser Asn Asn Thr Ala Ser Ile Ala Gln Ala Arg Lys Leu Val
1 5 10 15

Glu Gln Leu Lys Met Glu Ala Asn Ile Asp Arg Ile Lys Val Ser Lys
20 25 30

Ala Ala Ala Asp Leu Met Ala Tyr Cys Glu Ala His Ala Lys Glu Asp
35 40 45

Pro Leu Leu Thr Pro Val Pro Ala Ser Glu Asn Pro Phe Arg Glu Lys
50 55 60

Lys Phe Phe Cys Ala Ile Leu
65 70

<210> 69

<211> 514

<212> PRT

<213> 人工

<220>

<223> T4L-Beta2AR 融合构建体

<400> 69

Asp Tyr Lys Asp Asp Asp Asp Ala Glu Asn Leu Tyr Phe Gln Gly Asn
1 5 10 15

Ile Phe Glu Met Leu Arg Ile Asp Glu Gly Leu Arg Leu Lys Ile Tyr
20 25 30

Lys Asp Thr Glu Gly Tyr Tyr Thr Ile Gly Ile Gly His Leu Leu Thr
35 40 45

Lys Ser Pro Ser Leu Asn Ala Ala Lys Ser Glu Leu Asp Lys Ala Ile
50 55 60

Gly Arg Asn Thr Asn Gly Val Ile Thr Lys Asp Glu Ala Glu Lys Leu
65 70 75 80

Phe Asn Gln Asp Val Asp Ala Ala Val Arg Gly Ile Leu Arg Asn Ala
85 90 95

Lys Leu Lys Pro Val Tyr Asp Ser Leu Asp Ala Val Arg Arg Ala Ala
100 105 110

Leu Ile Asn Met Val Phe Gln Met Gly Glu Thr Gly Val Ala Gly Phe
115 120 125

Thr Asn Ser Leu Arg Met Leu Gln Gln Lys Arg Trp Asp Glu Ala Ala
130 135 140

Val Asn Leu Ala Lys Ser Arg Trp Tyr Asn Gln Thr Pro Asn Arg Ala
145 150 155 160

Lys Arg Val Ile Thr Thr Phe Arg Thr Gly Thr Trp Asp Ala Tyr Ala
165 170 175

Ala Asp Glu Val Trp Val Val Gly Met Gly Ile Val Met Ser Leu Ile
180 185 190

Val Leu Ala Ile Val Phe Gly Asn Val Leu Val Ile Thr Ala Ile Ala
195 200 205

Lys Phe Glu Arg Leu Gln Thr Val Thr Asn Tyr Phe Ile Thr Ser Leu
210 215 220

Ala Cys Ala Asp Leu Val Met Gly Leu Ala Val Val Pro Phe Gly Ala
225 230 235 240

Ala His Ile Leu Thr Lys Thr Trp Thr Phe Gly Asn Phe Trp Cys Glu
245 250 255

[0030]

Phe Trp Thr Ser Ile Asp Val Leu Cys Val Thr Ala Ser Ile Glu Thr
 260 265 270
 Leu Cys Val Ile Ala Val Asp Arg Tyr Phe Ala Ile Thr Ser Pro Phe
 275 280 285
 Lys Tyr Gln Ser Leu Leu Thr Lys Asn Lys Ala Arg Val Ile Ile Leu
 290 295 300
 Met Val Trp Ile Val Ser Gly Leu Thr Ser Phe Leu Pro Ile Gln Met
 305 310 315 320
 His Trp Tyr Arg Ala Thr His Gln Glu Ala Ile Asn Cys Tyr Ala Glu
 325 330 335
 Glu Thr Cys Cys Asp Phe Phe Thr Asn Gln Ala Tyr Ala Ile Ala Ser
 340 345 350
 Ser Ile Val Ser Phe Tyr Val Pro Leu Val Ile Met Val Phe Val Tyr
 355 360 365
 Ser Arg Val Phe Gln Glu Ala Lys Arg Gln Leu Gln Lys Ile Asp Lys
 370 375 380
 Ser Glu Gly Arg Phe His Val Gln Asn Leu Ser Gln Val Glu Gln Asp
 385 390 395 400
 Gly Arg Thr Gly His Gly Leu Arg Arg Ser Ser Lys Phe Cys Leu Lys
 405 410 415
 Glu His Lys Ala Leu Lys Thr Leu Gly Ile Ile Met Gly Thr Phe Thr
 420 425 430
 Leu Cys Trp Leu Pro Phe Phe Ile Val Asn Ile Val His Val Ile Gln
 435 440 445
 Asp Asn Leu Ile Arg Lys Glu Val Tyr Ile Leu Leu Asn Trp Ile Gly
 450 455 460
 Tyr Val Asn Ser Gly Phe Asn Pro Leu Ile Tyr Cys Arg Ser Pro Asp
 465 470 475 480
 Phe Arg Ile Ala Phe Gln Glu Leu Leu Cys Leu Arg Arg Ser Ser Leu
 485 490 495
 Lys Ala Tyr Gly Asn Gly Tyr Ser Ser Asn Gly Asn Thr Gly Glu Gln
 500 505 510
 Ser Gly

[0031]

<210> 70
 <211> 8
 <212> PRT
 <213> 人工

<220>
 <223> MI Flag 表位

<400> 70

Asp Tyr Lys Asp Asp Asp Asp Ala
1 5

<210> 71
 <211> 7
 <212> PRT
 <213> 人工

<220>
 <223> TEV 蛋白酶识别序列

<400> 71

Glu Asn Leu Tyr Phe Gln Gly
1 5

<210> 72

<211> 413
 <212> PRT
 <213> 人

<400> 72

```

Met Gly Gln Pro Gly Asn Gly Ser Ala Phe Leu Leu Ala Pro Asn Gly
1      5      10      15
Ser His Ala Pro Asp His Asp Val Thr Gln Glu Arg Asp Glu Val Trp
20     25     30
Val Val Gly Met Gly Ile Val Met Ser Leu Ile Val Leu Ala Ile Val
35     40     45
Phe Gly Asn Val Leu Val Ile Thr Ala Ile Ala Lys Phe Glu Arg Leu
50     55     60
Gln Thr Val Thr Asn Tyr Phe Ile Thr Ser Leu Ala Cys Ala Asp Leu
65     70     75     80
Val Met Gly Leu Ala Val Val Pro Phe Gly Ala Ala His Ile Leu Met
85     90     95
Lys Met Trp Thr Phe Gly Asn Phe Trp Cys Glu Phe Trp Thr Ser Ile
100    105    110
Asp Val Leu Cys Val Thr Ala Ser Ile Glu Thr Leu Cys Val Ile Ala
115    120    125
Val Asp Arg Tyr Phe Ala Ile Thr Ser Pro Phe Lys Tyr Gln Ser Leu
130    135    140
Leu Thr Lys Asn Lys Ala Arg Val Ile Ile Leu Met Val Trp Ile Val
145    150    155    160
Ser Gly Leu Thr Ser Phe Leu Pro Ile Gln Met His Trp Tyr Arg Ala
165    170    175
Thr His Gln Glu Ala Ile Asn Cys Tyr Ala Asn Glu Thr Cys Cys Asp
180    185    190
Phe Phe Thr Asn Gln Ala Tyr Ala Ile Ala Ser Ser Ile Val Ser Phe
195    200    205
Tyr Val Pro Leu Val Ile Met Val Phe Val Tyr Ser Arg Val Phe Gln
210    215    220
Glu Ala Lys Arg Gln Leu Gln Lys Ile Asp Lys Ser Glu Gly Arg Phe
225    230    235    240
His Val Gln Asn Leu Ser Gln Val Glu Gln Asp Gly Arg Thr Gly His
245    250    255
Gly Leu Arg Arg Ser Ser Lys Phe Cys Leu Lys Glu His Lys Ala Leu
260    265    270
Lys Thr Leu Gly Ile Ile Met Gly Thr Phe Thr Leu Cys Trp Leu Pro
275    280    285
Phe Phe Ile Val Asn Ile Val His Val Ile Gln Asp Asn Leu Ile Arg
290    295    300
Lys Glu Val Tyr Ile Leu Leu Asn Trp Ile Gly Tyr Val Asn Ser Gly
305    310    315    320
Phe Asn Pro Leu Ile Tyr Cys Arg Ser Pro Asp Phe Arg Ile Ala Phe
325    330    335
Gln Glu Leu Leu Cys Leu Arg Arg Ser Ser Leu Lys Ala Tyr Gly Asn
340    345    350
Gly Tyr Ser Ser Asn Gly Asn Thr Gly Glu Gln Ser Gly Tyr His Val
355    360    365
Glu Gln Glu Lys Glu Asn Lys Leu Leu Cys Glu Asp Leu Pro Gly Thr
370    375    380

```

[0032]

Glu Asp Phe Val Gly His Gln Gly Thr Val Pro Ser Asp Asn Ile Asp
 385 390 395 400
 Ser Gln Gly Arg Asn Cys Ser Thr Asn Asp Ser Leu Leu
 405 410
 <210> 73
 <211> 539
 <212> PRT
 <213> 人工序列
 <220>
 <223> N-末端 T4LV2R构建体
 <400> 73
 Asp Tyr Lys Asp Asp Asp Ala Gly Leu Met Ala Ser Thr Thr Ser
 1 5 10 15
 Ala Val Pro Gly His Pro Ser Leu Pro Ser Leu Pro Ser Asn Ser Ser
 20 25 30
 Gln Glu Arg Pro Leu Asp Thr Asn Ile Phe Glu Met Leu Arg Ile Asp
 35 40 45
 Glu Gly Leu Arg Leu Lys Ile Tyr Lys Asp Thr Glu Gly Tyr Tyr Thr
 50 55 60
 Ile Gly Ile Gly His Leu Leu Thr Lys Ser Pro Ser Leu Asn Ala Ala
 65 70 75 80
 Lys Ser Glu Leu Asp Lys Ala Ile Gly Arg Asn Thr Asn Gly Val Ile
 85 90 95
 Thr Lys Asp Glu Ala Glu Lys Leu Phe Asn Gln Asp Val Asp Ala Ala
 100 105 110
 Val Arg Gly Ile Leu Arg Asn Ala Lys Leu Lys Pro Val Tyr Asp Ser
 115 120 125
 Leu Asp Ala Val Arg Arg Ala Ala Leu Ile Asn Met Val Phe Gln Met
 130 135 140
 Gly Glu Thr Gly Val Ala Gly Phe Thr Asn Ser Leu Arg Met Leu Gln
 145 150 155 160
 Gln Lys Arg Trp Asp Glu Ala Ala Val Asn Leu Ala Lys Ser Arg Trp
 165 170 175
 Tyr Asn Gln Thr Pro Asn Arg Ala Lys Arg Val Ile Thr Thr Phe Arg
 180 185 190
 Thr Gly Thr Trp Asp Ala Tyr Arg Asp Pro Leu Leu Ala Arg Ala Glu
 195 200 205
 Leu Ala Leu Leu Ser Ile Val Phe Val Ala Val Ala Leu Ser Asn Gly
 210 215 220
 Leu Val Leu Ala Ala Leu Ala Arg Arg Gly Arg Arg Gly His Trp Ala
 225 230 235 240
 Pro Ile His Val Phe Ile Gly His Leu Cys Leu Ala Asp Leu Ala Val
 245 250 255
 Ala Leu Phe Gln Val Leu Pro Gln Leu Ala Trp Lys Ala Thr Asp Arg
 260 265 270
 Phe Arg Gly Pro Asp Ala Leu Cys Arg Ala Val Lys Tyr Leu Gln Met
 275 280 285
 Val Gly Met Tyr Ala Ser Ser Tyr Met Ile Leu Ala Met Thr Leu Asp
 290 295 300
 Arg His Arg Ala Ile Cys Arg Pro Met Leu Ala Tyr Arg His Gly Ser
 305 310 315 320
 Gly Ala His Trp Asn Arg Pro Val Leu Val Ala Trp Ala Phe Ser Leu
 325 330 335

[0033]

[0034]

Leu Leu Ser Leu Pro Gln Leu Phe Ile Phe Ala Gln Arg Asn Val Glu
 340 345 350
 Gly Gly Ser Gly Val Thr Asp Cys Trp Ala Cys Phe Ala Glu Pro Trp
 355 360 365
 Gly Arg Arg Thr Tyr Val Thr Trp Ile Ala Leu Met Val Phe Val Ala
 370 375 380
 Pro Thr Leu Gly Ile Ala Ala Cys Gln Val Leu Ile Phe Arg Glu Ile
 385 390 395 400
 His Ala Ser Leu Val Pro Gly Pro Ser Glu Arg Pro Gly Gly Arg Arg
 405 410 415
 Arg Gly Arg Arg Thr Gly Ser Pro Gly Glu Gly Ala His Val Ser Ala
 420 425 430
 Ala Val Ala Lys Thr Val Arg Met Thr Leu Val Ile Val Val Val Tyr
 435 440 445
 Val Leu Cys Trp Ala Pro Phe Phe Leu Val Gln Leu Trp Ala Ala Trp
 450 455 460
 Asp Pro Glu Ala Pro Leu Glu Gly Ala Pro Phe Val Leu Leu Met Leu
 465 470 475 480
 Leu Ala Ser Leu Asn Ser Cys Thr Asn Pro Trp Ile Tyr Ala Ser Phe
 485 490 495
 Ser Ser Ser Val Ser Ser Glu Leu Arg Ser Leu Leu Cys Cys Ala Arg
 500 505 510
 Gly Arg Thr Pro Pro Ser Leu Gly Pro Gln Asp Glu Ser Cys Thr Thr
 515 520 525
 Ala Ser Ser Ser Leu Ala Lys Asp Thr Ser Ser
 530 535

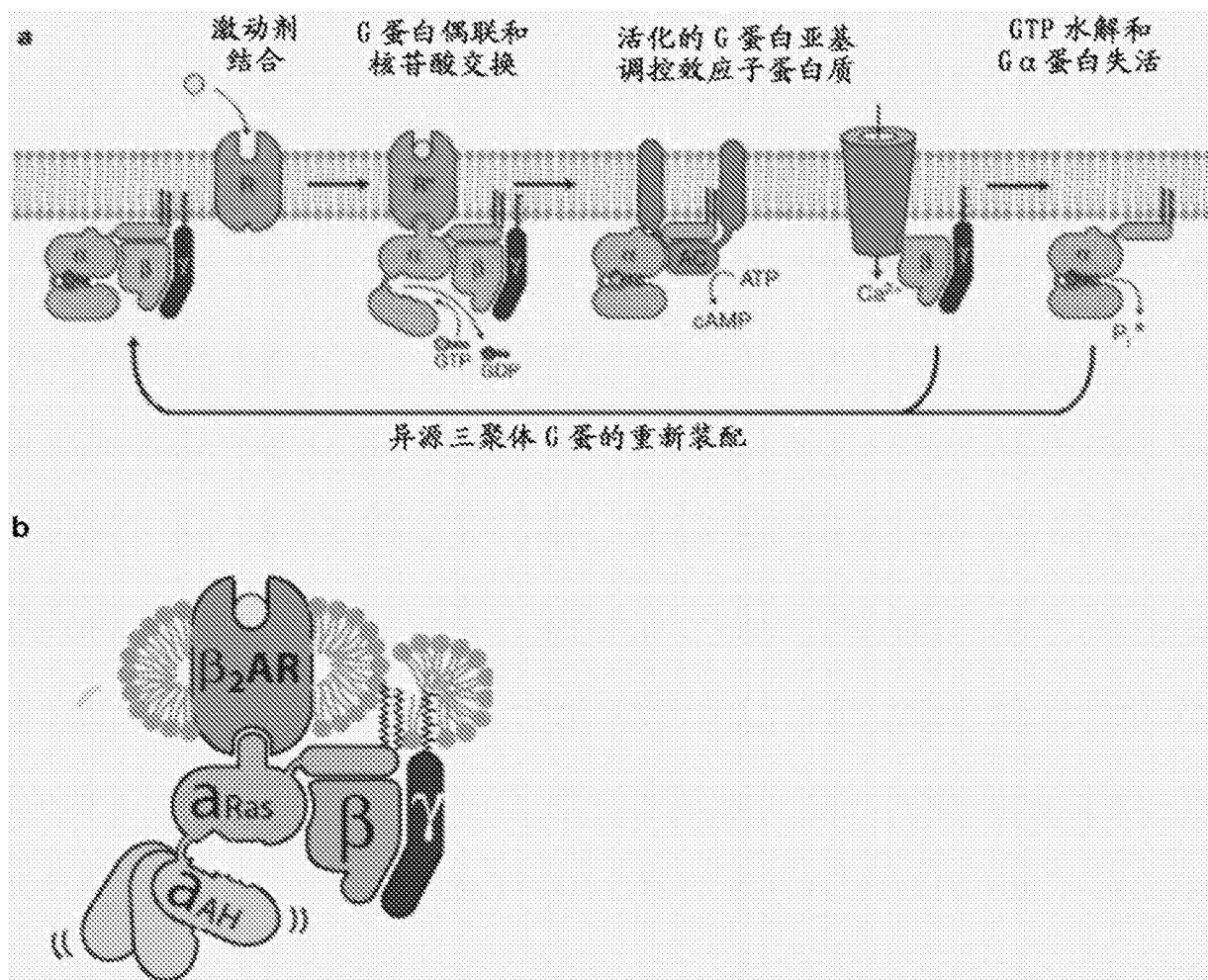


图1

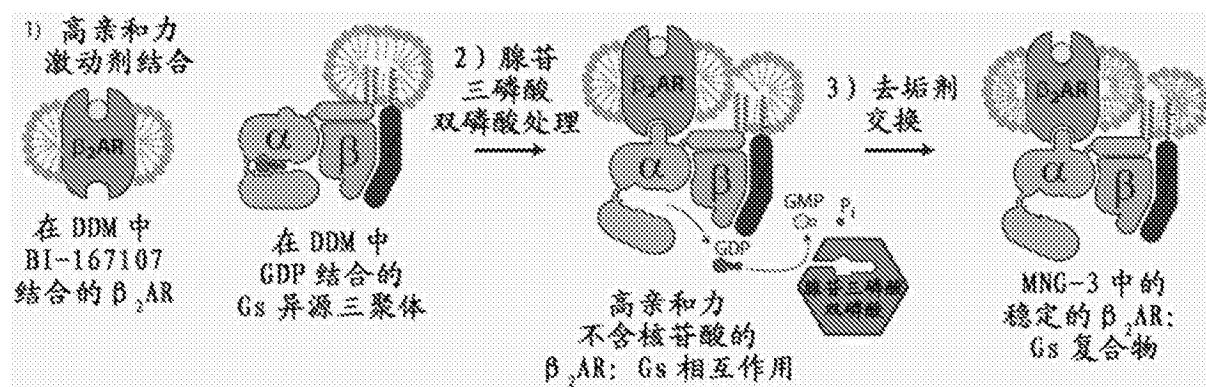
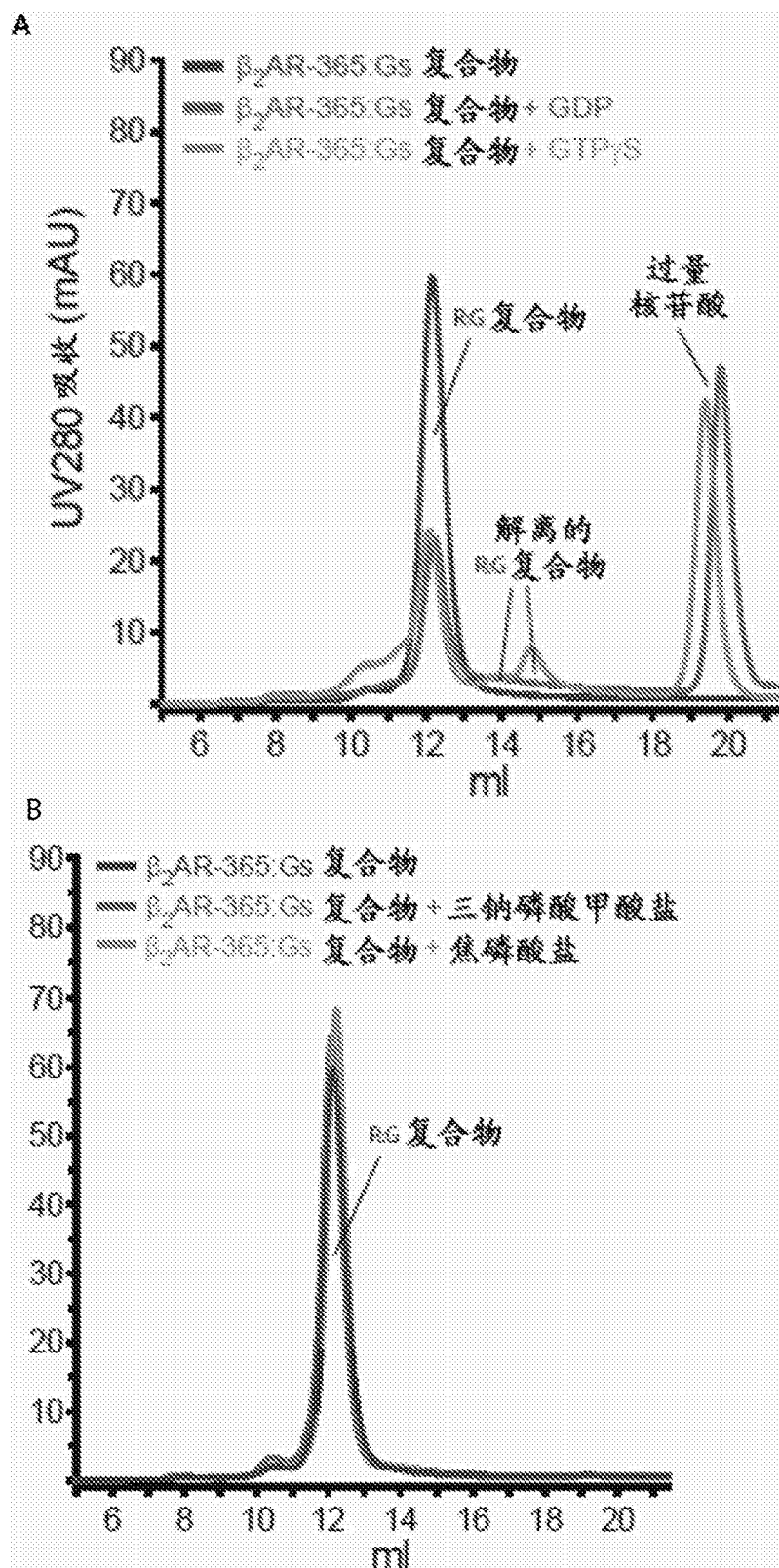
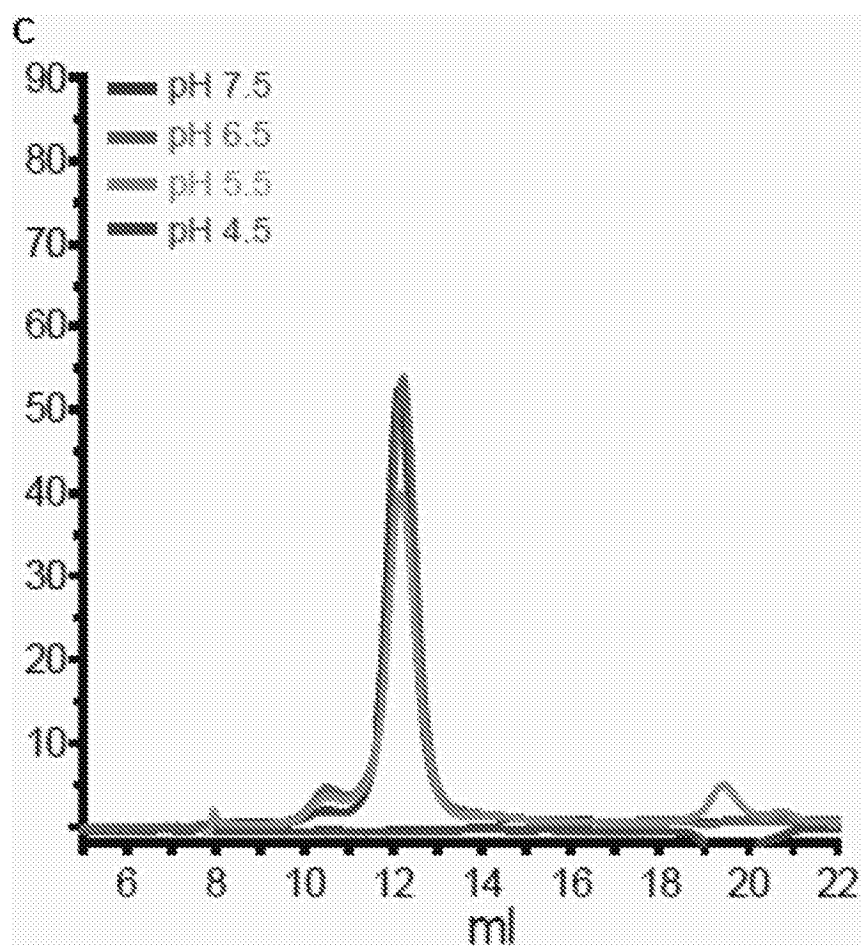
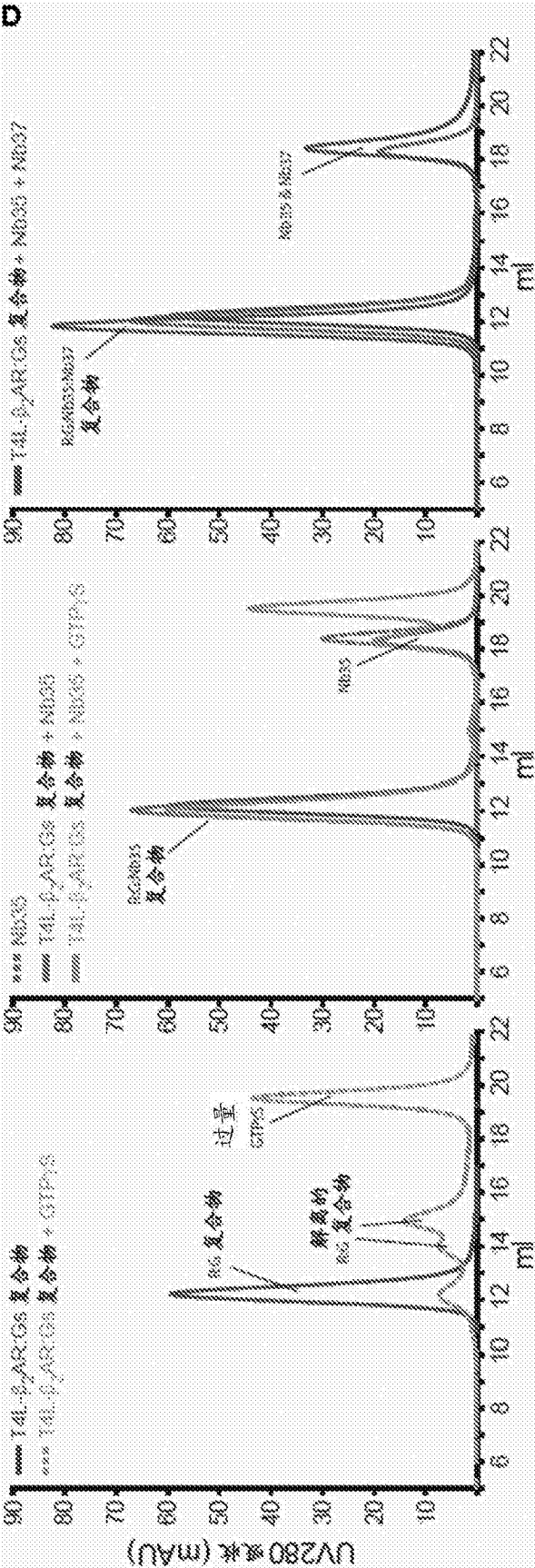


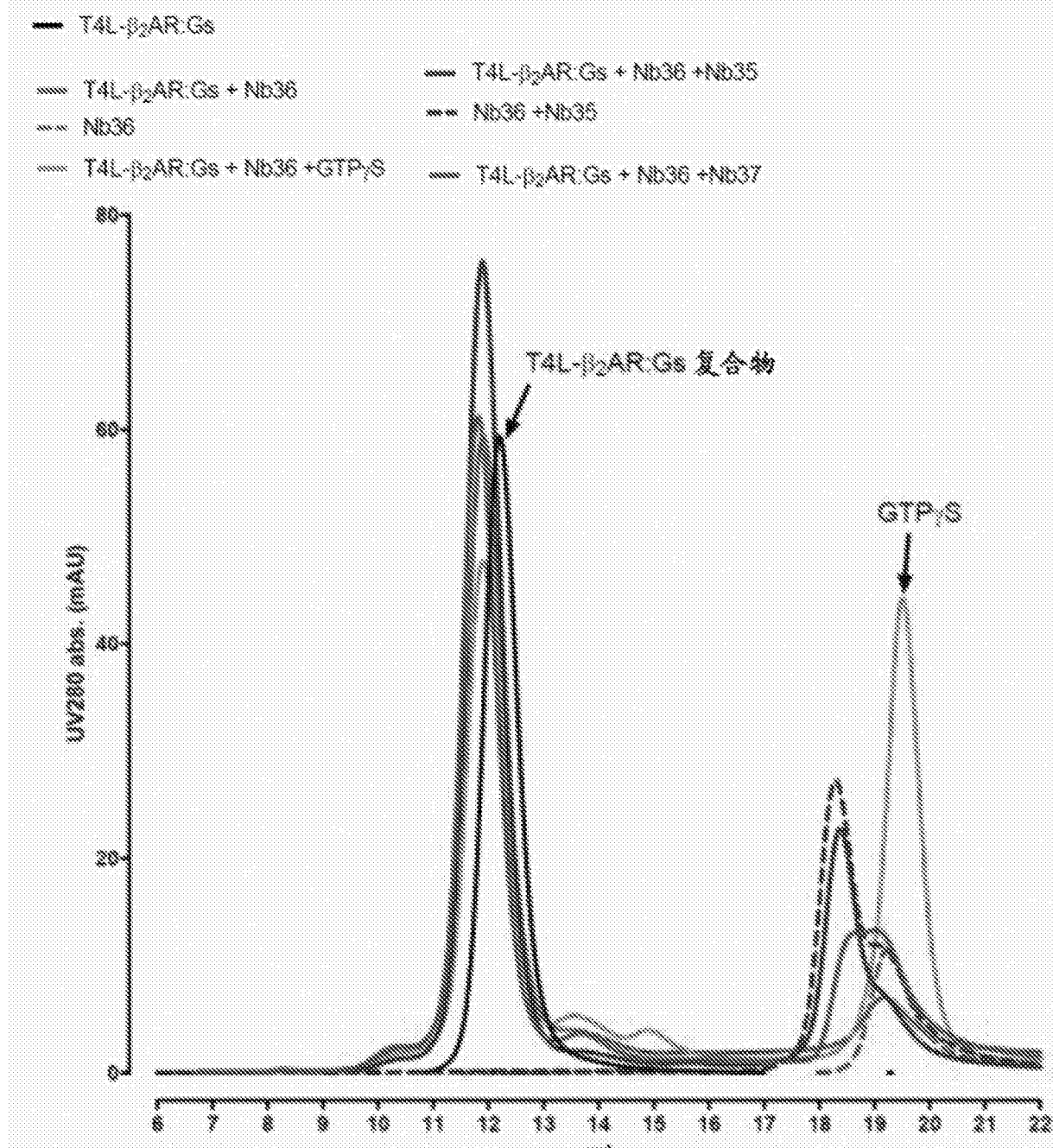
图2







E



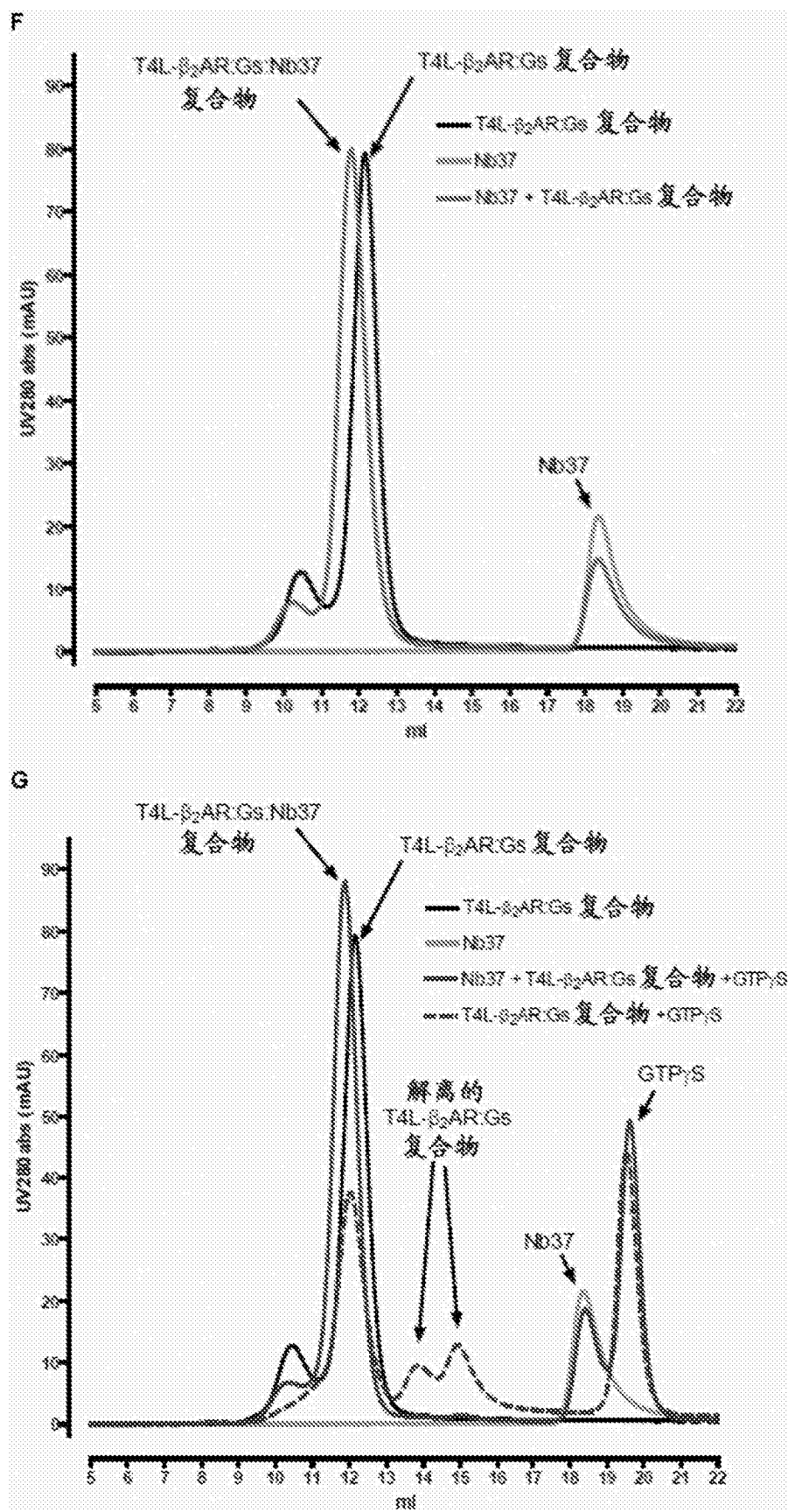


图3

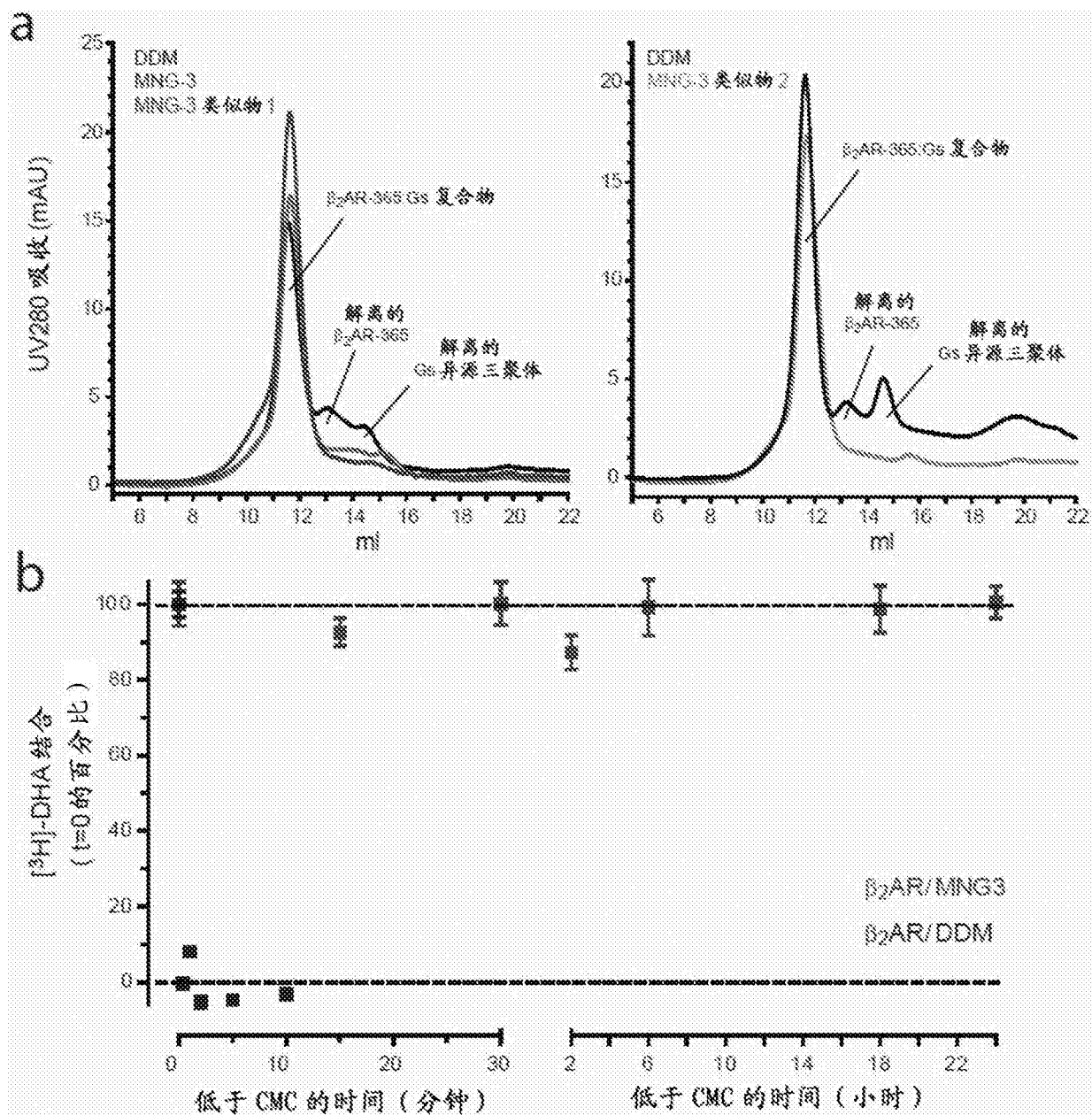


图4

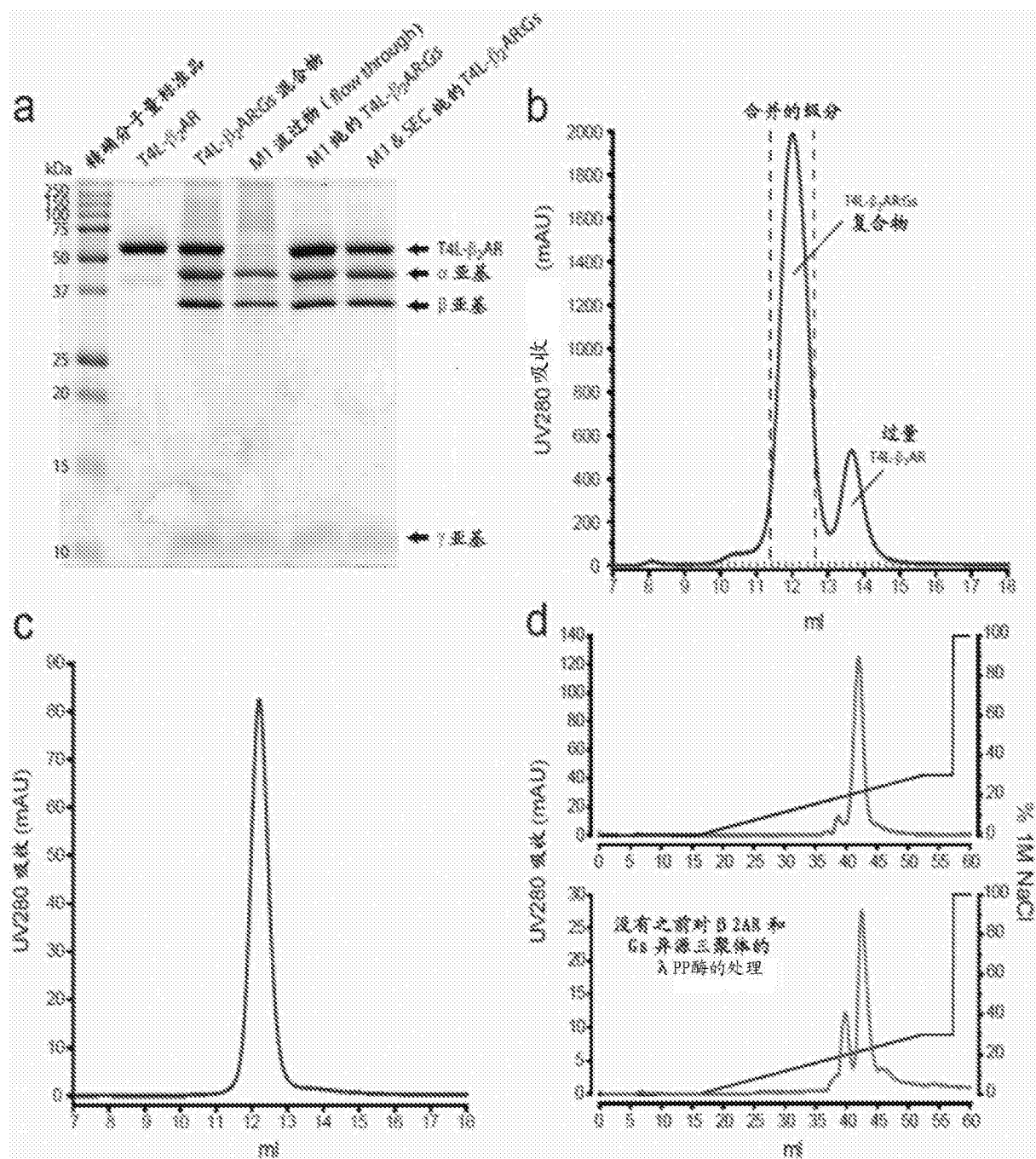


图5

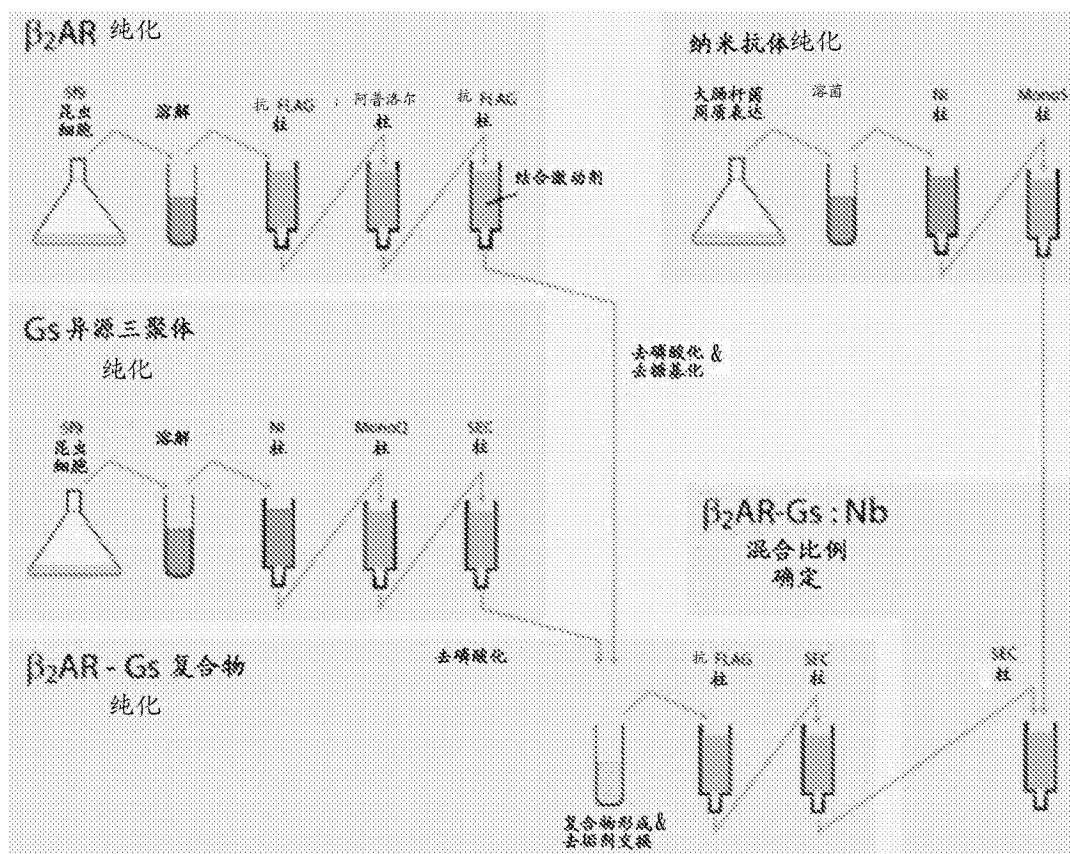


图6

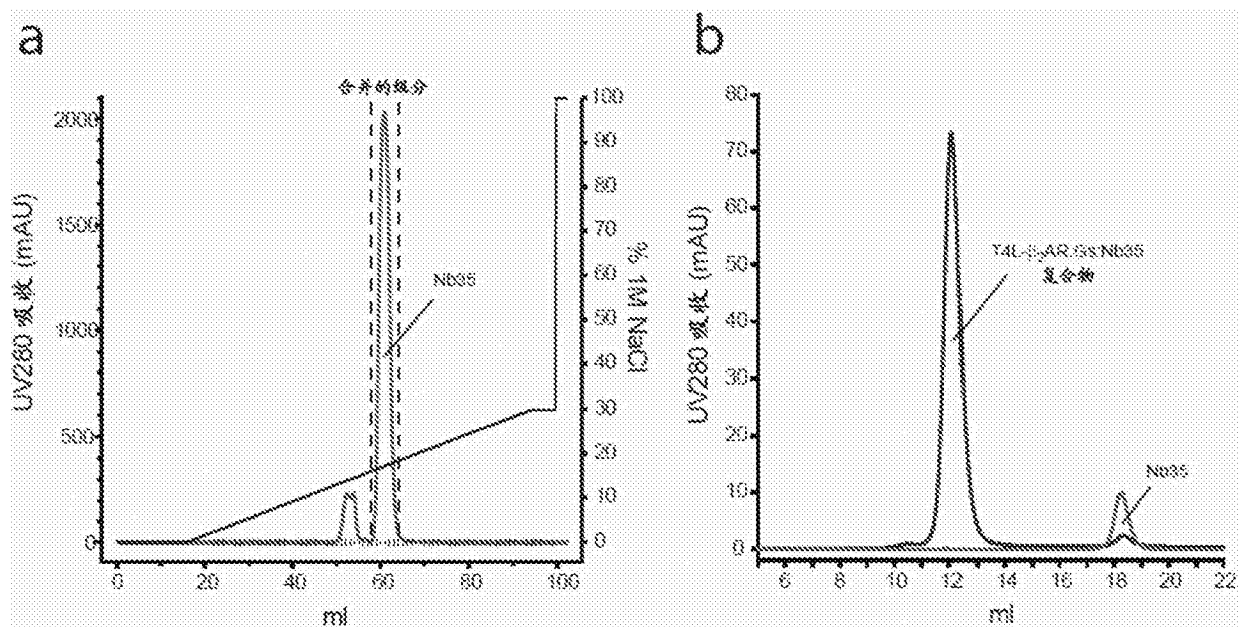


图7

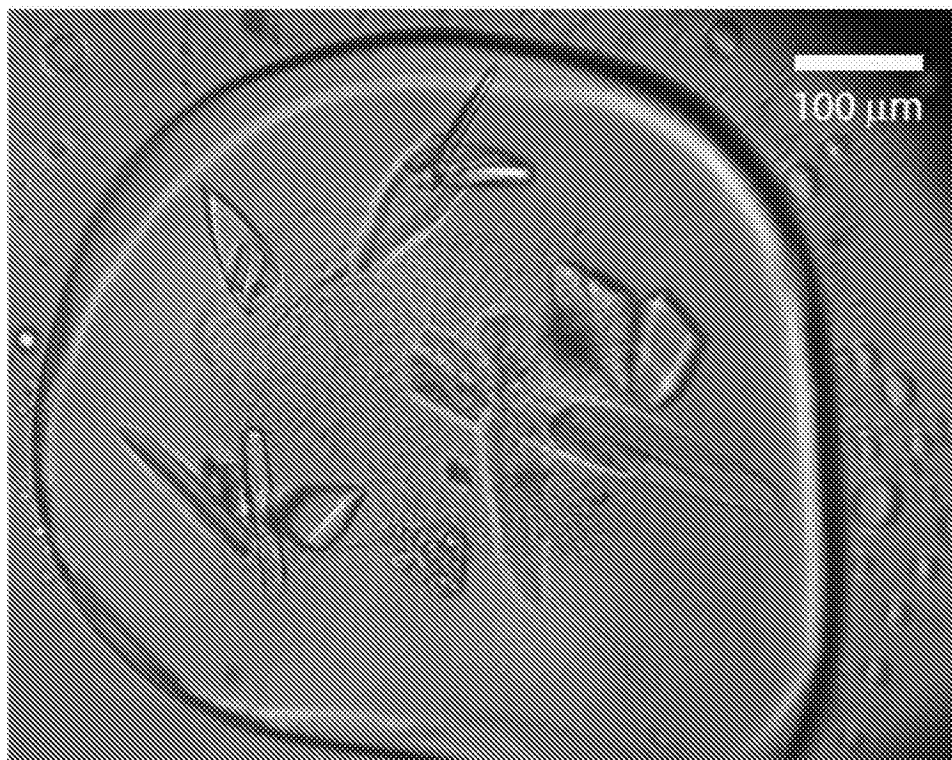
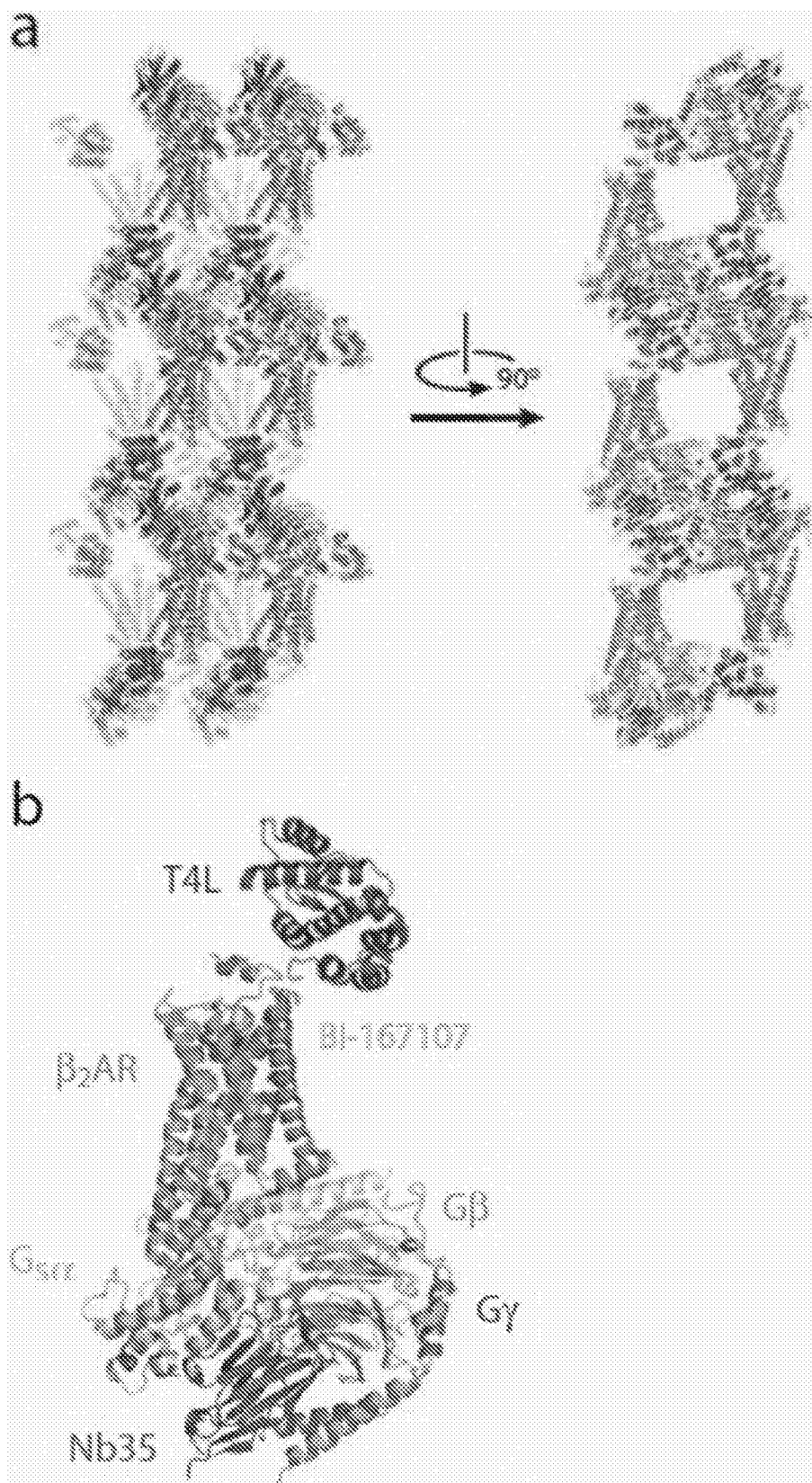


图8



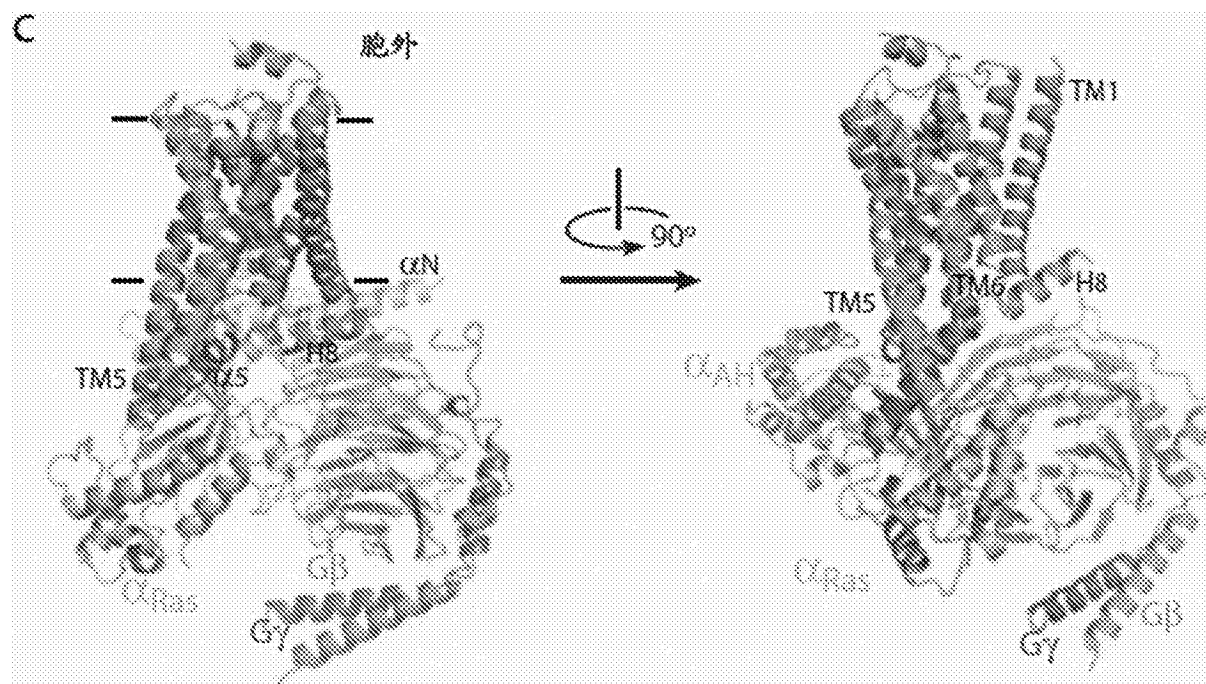


图9

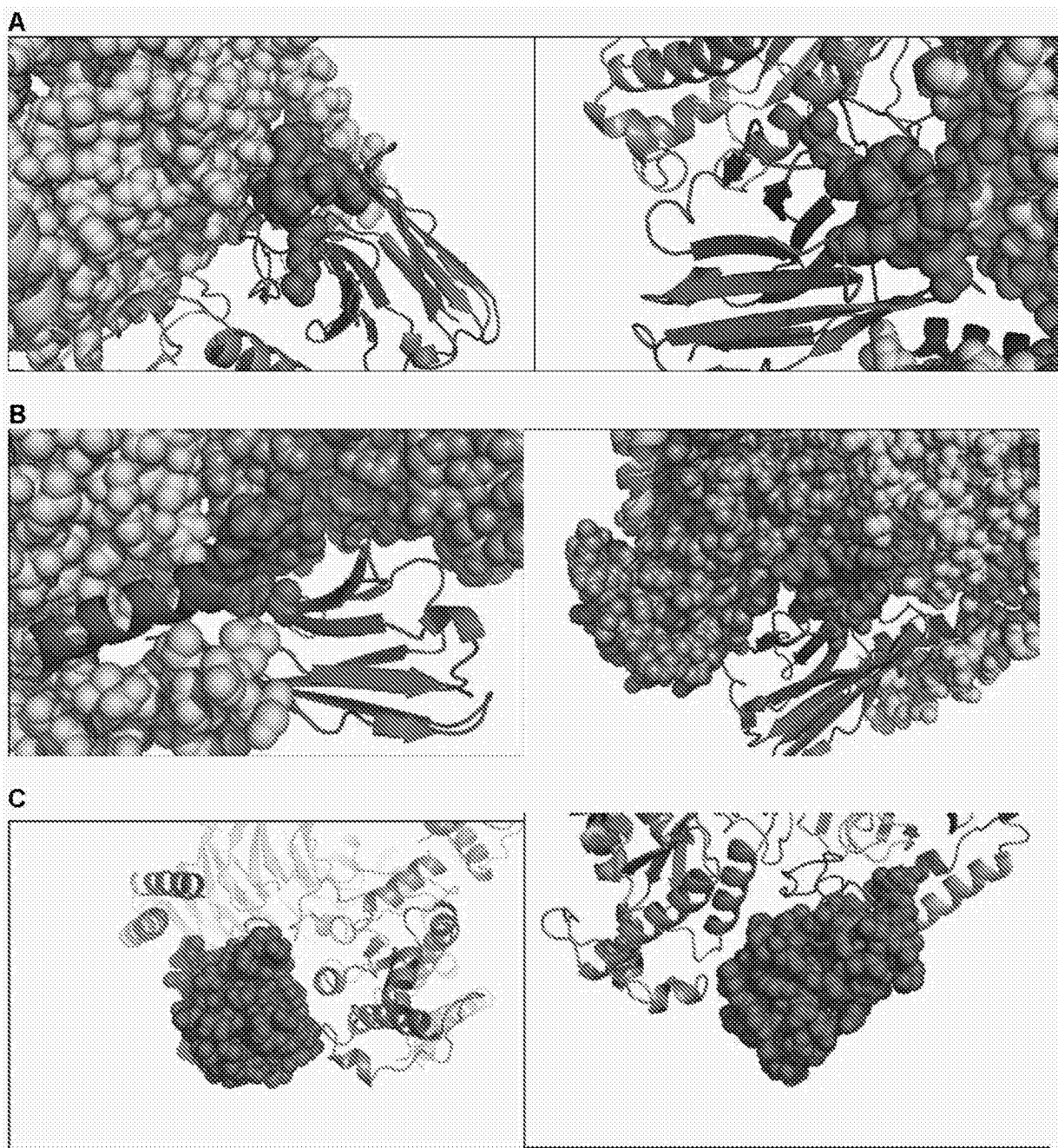


图10

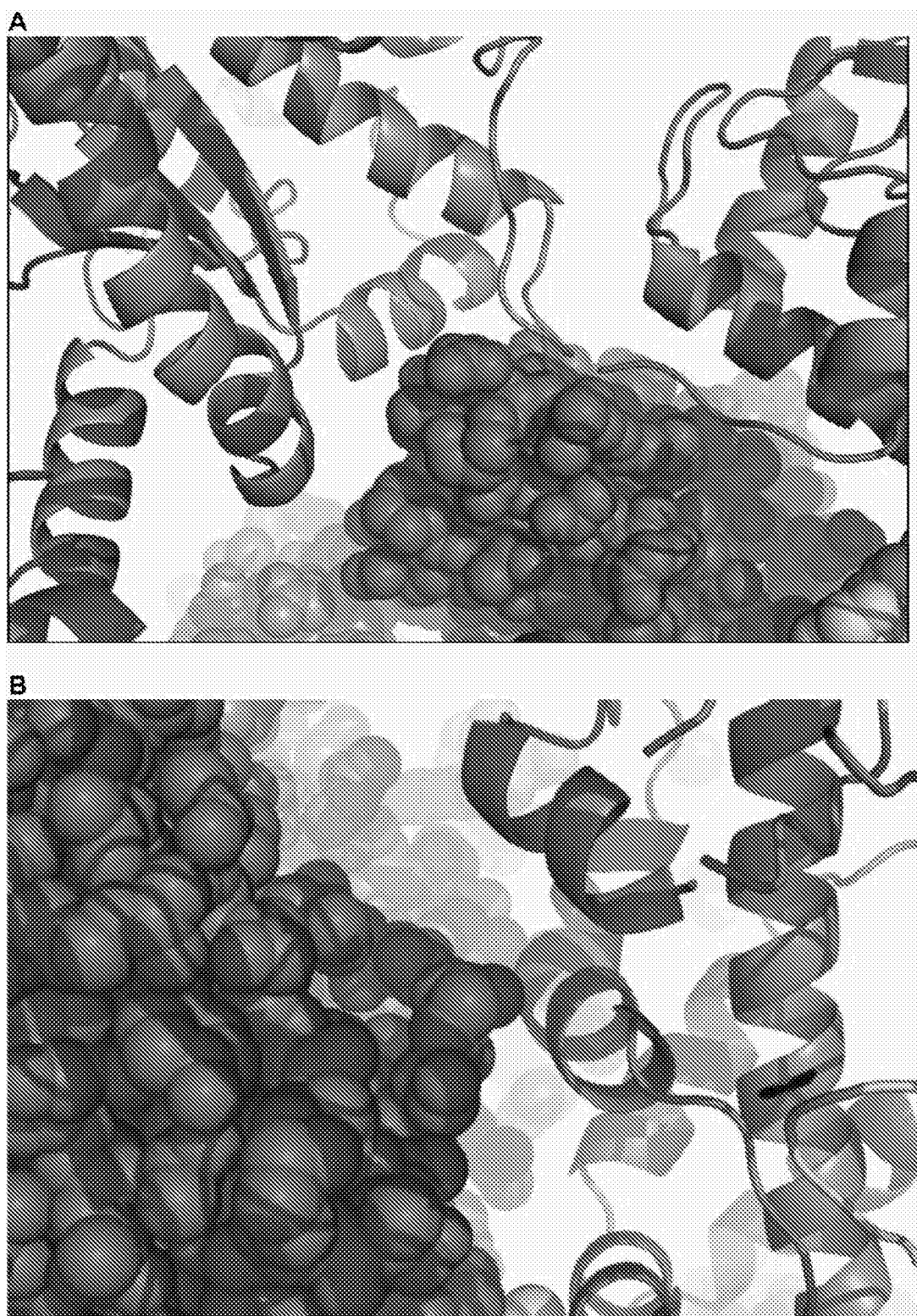
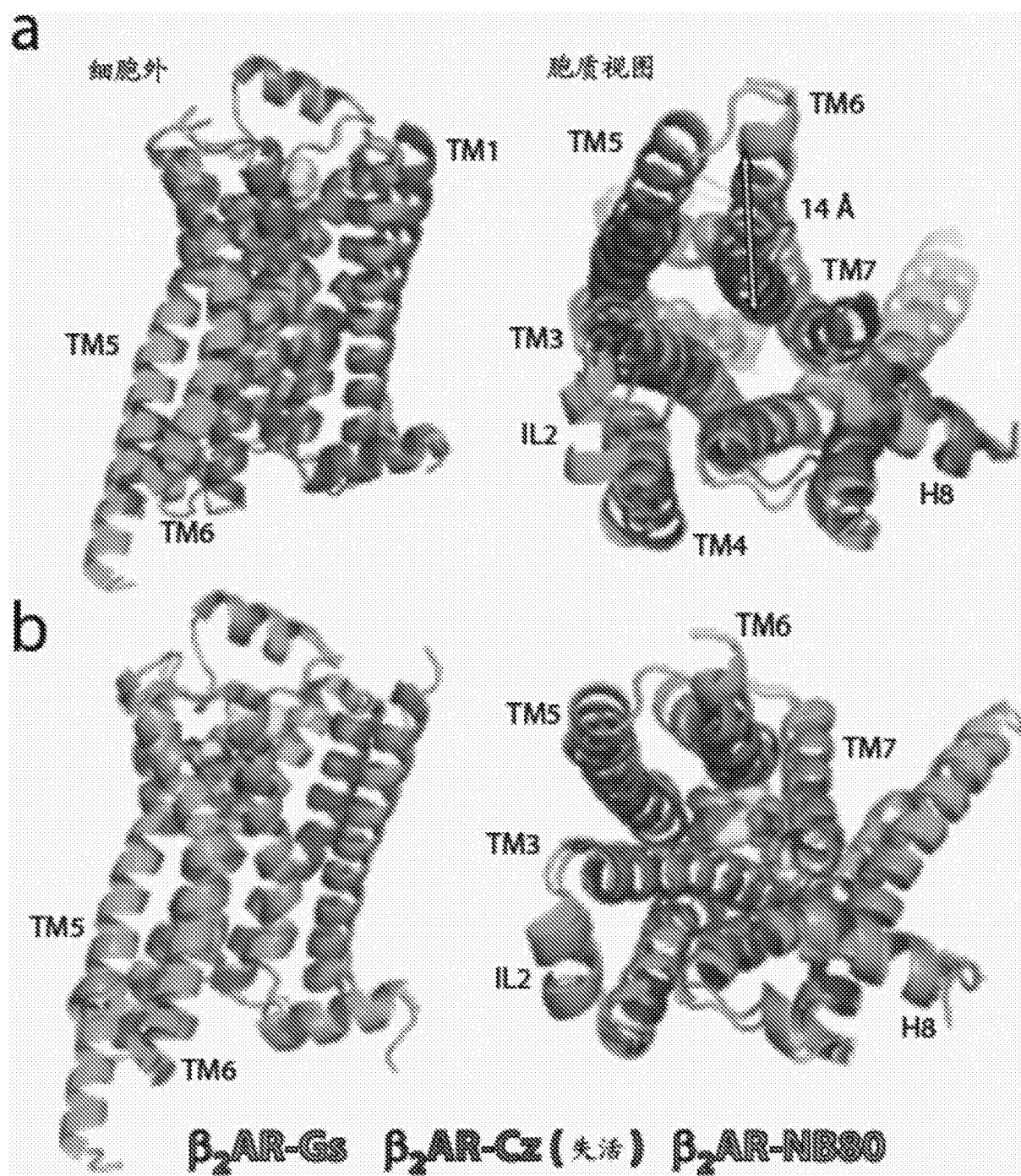


图11



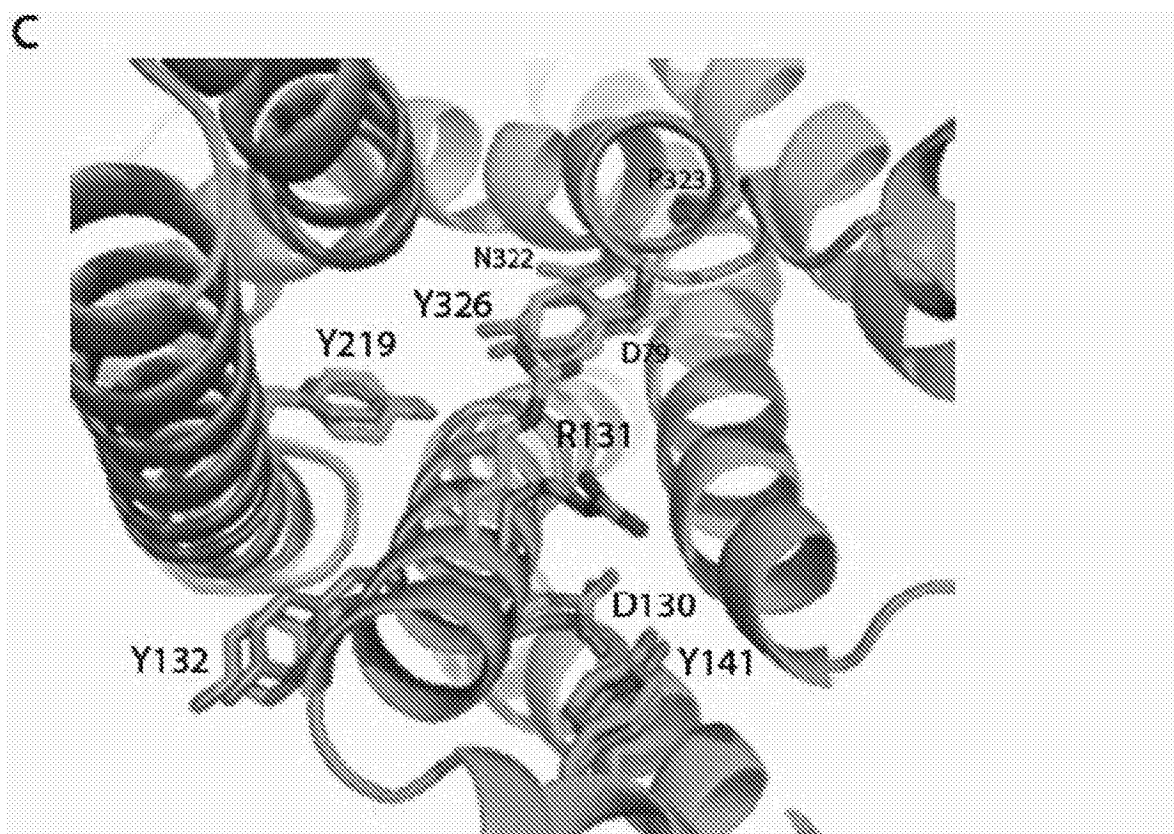
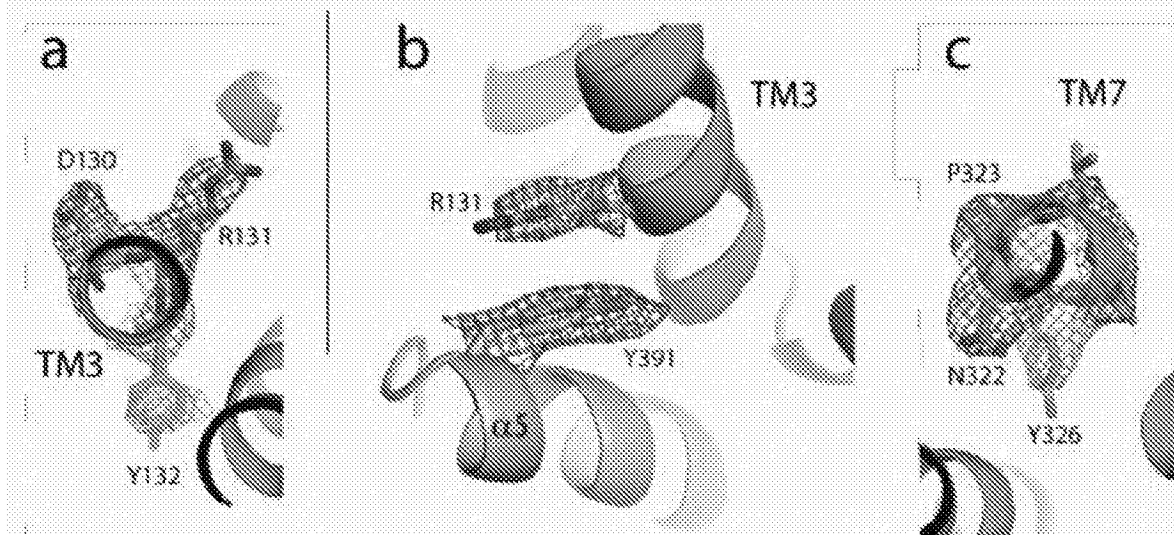


图 12



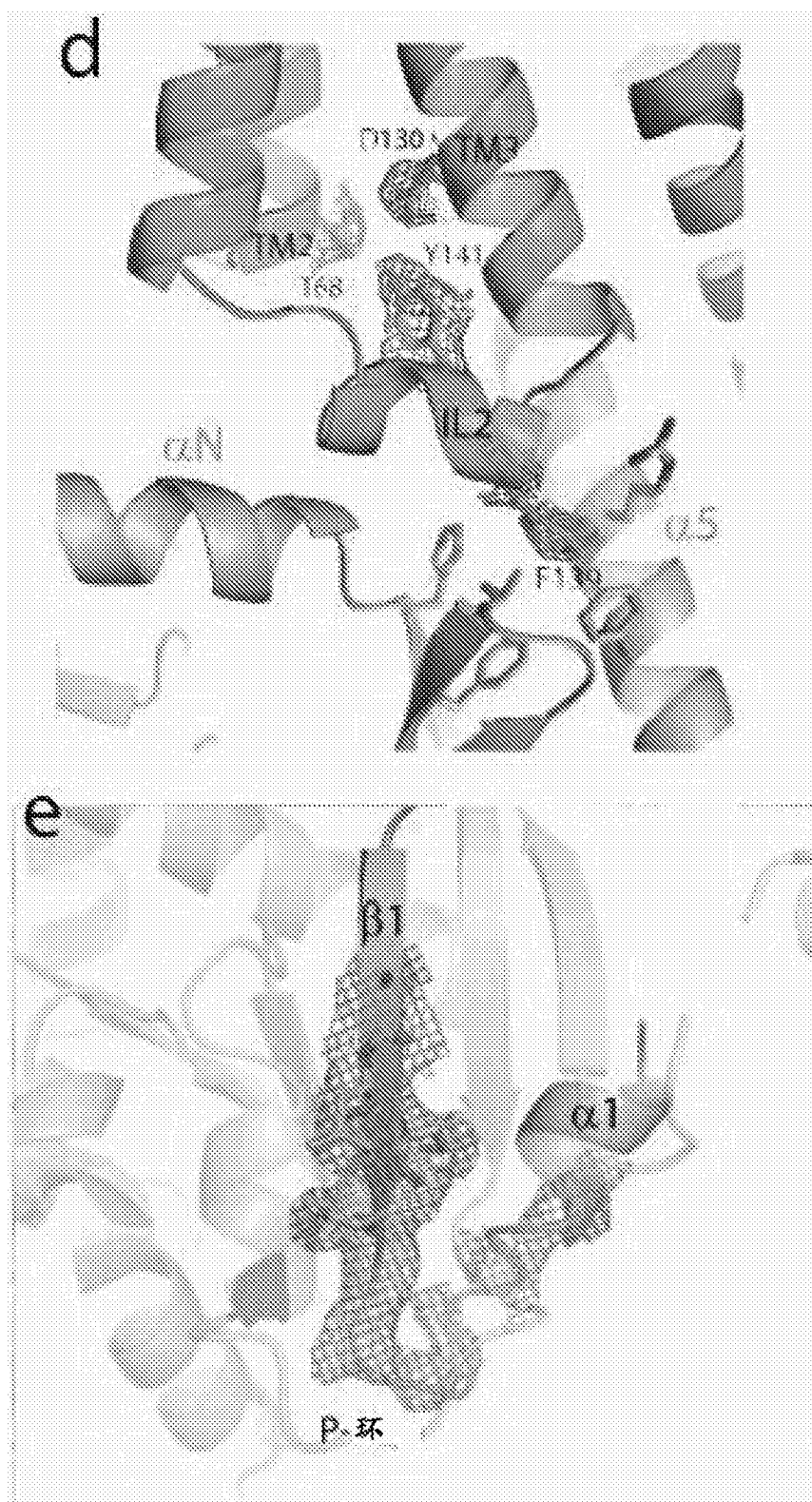


图13

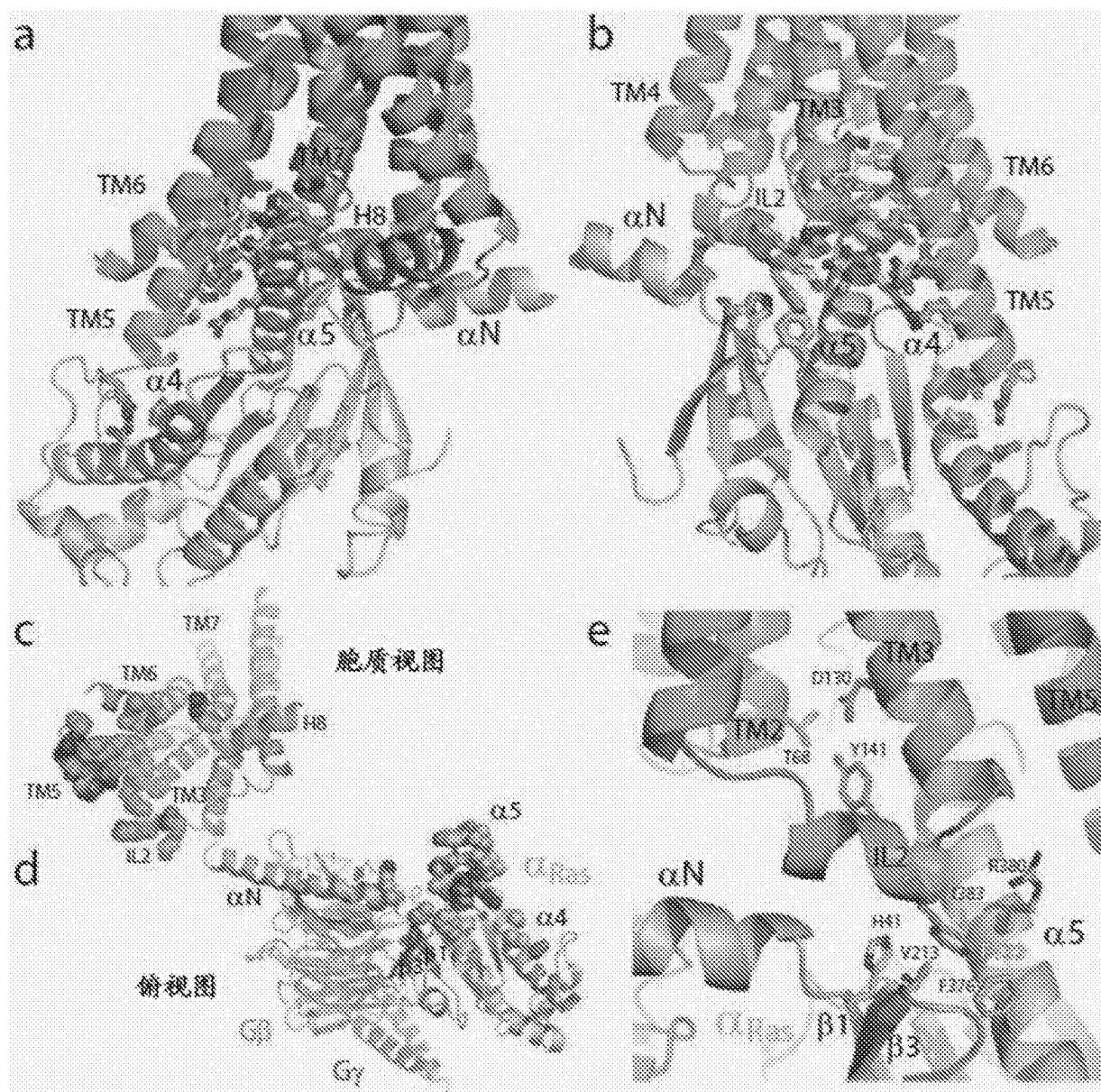


图14

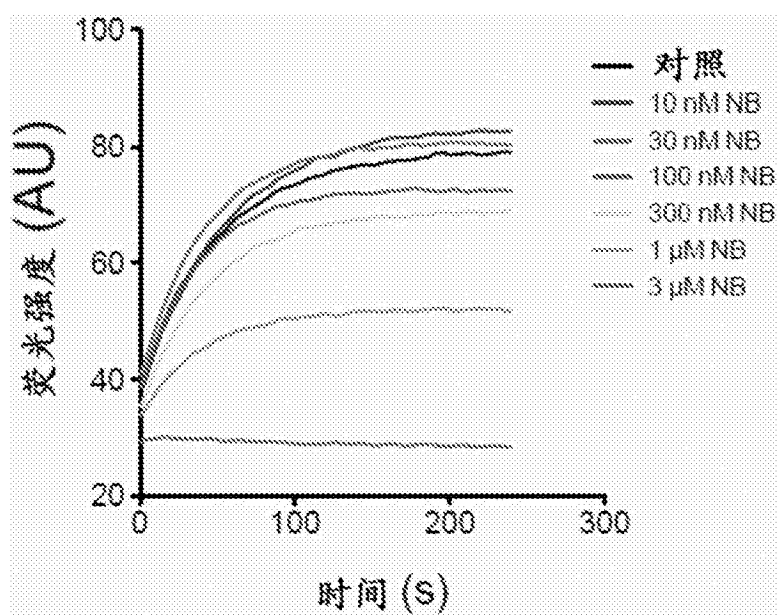


图16

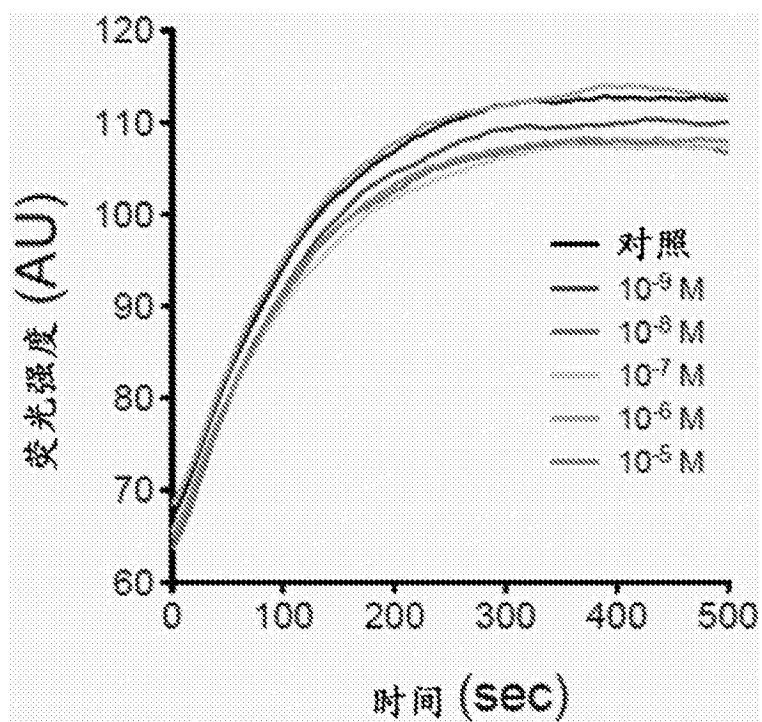


图17

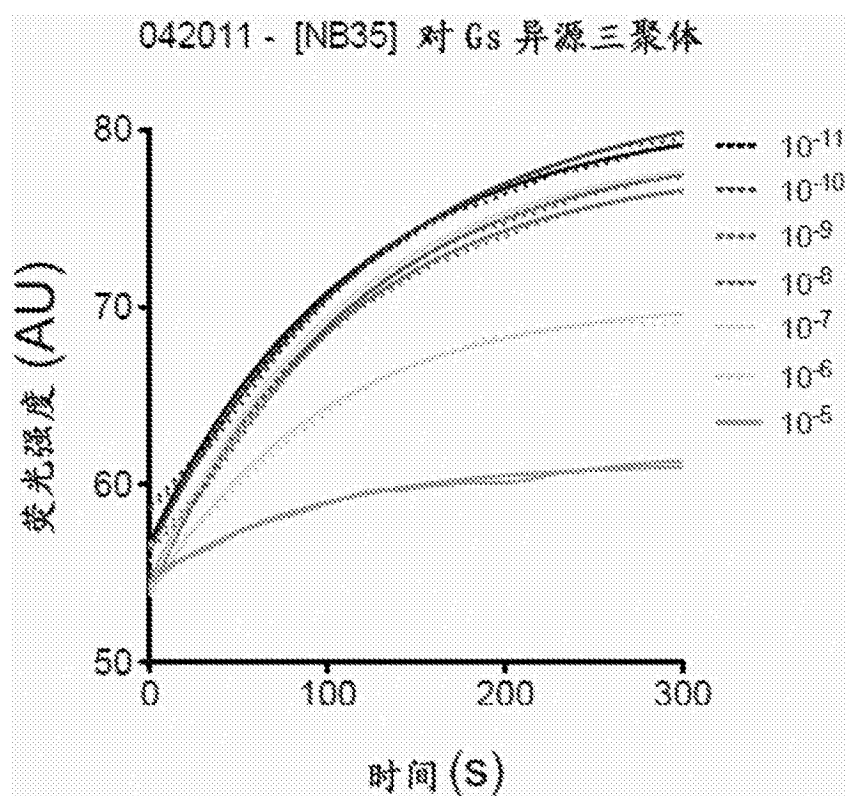


图18

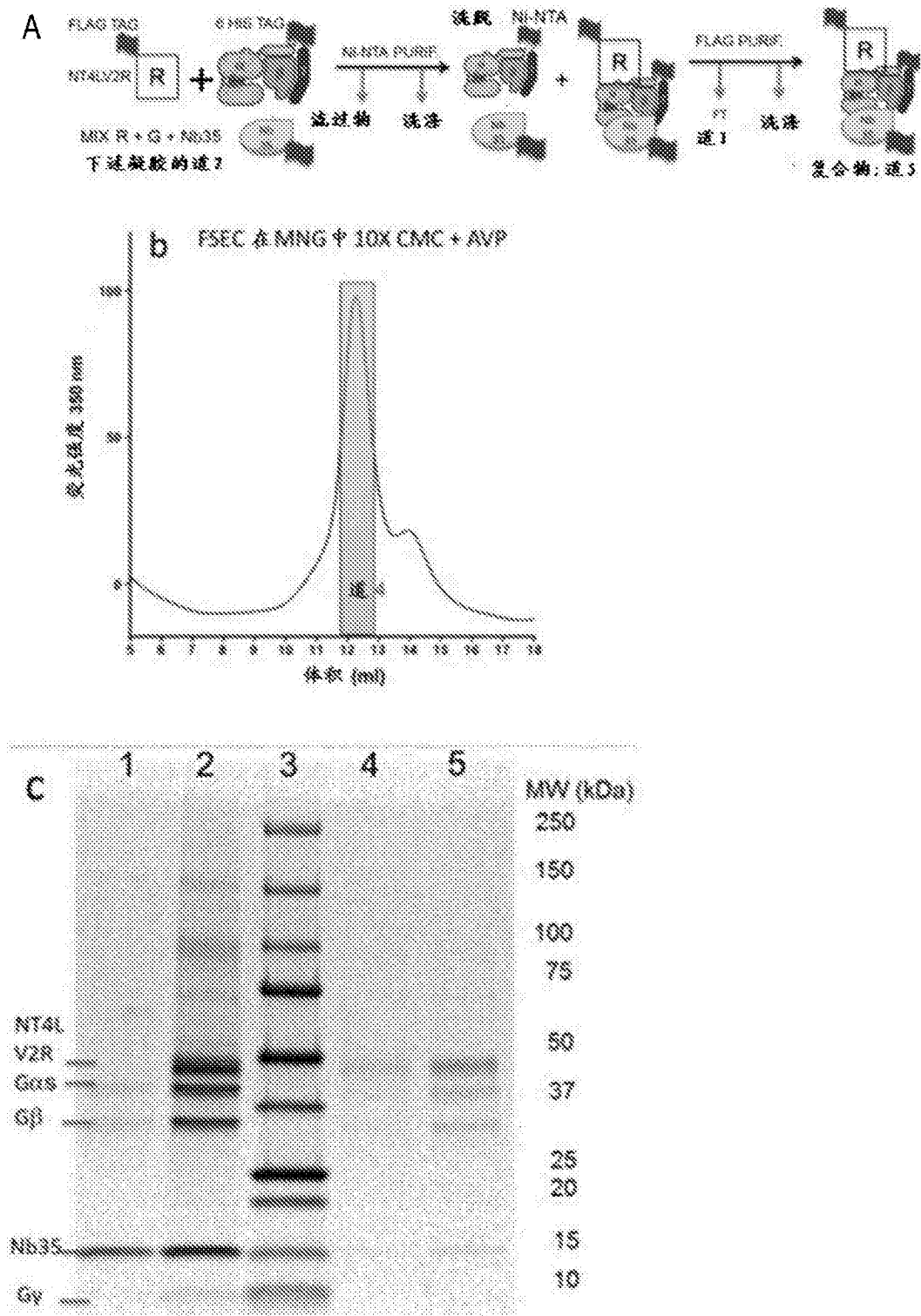


图19

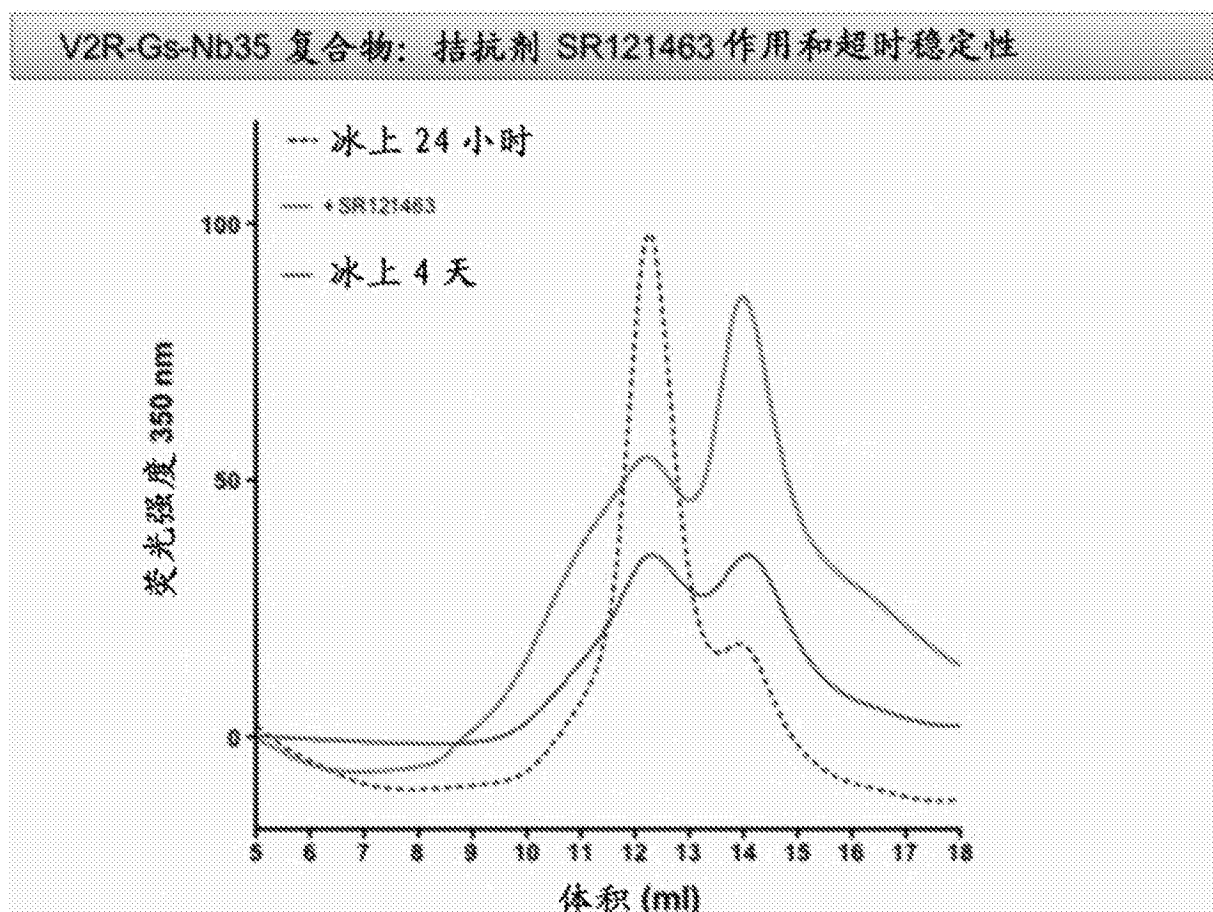


图20