



(12) 发明专利申请

(10) 申请公布号 CN 101845090 A

(43) 申请公布日 2010.09.29

(21) 申请号 200910080734.4

(22) 申请日 2009.03.26

(71) 申请人 中国人民解放军总医院  
地址 100853 北京市复兴路 28 号  
申请人 北京表源生物技术有限公司

(72) 发明人 唐佩福 刘长振 黄鹏 张世谦  
高斌

(74) 专利代理机构 北京康信知识产权代理有限  
责任公司 11240  
代理人 余刚 张英

(51) Int. Cl.

*C07K 14/47* (2006.01)

*C30B 29/54* (2006.01)

*G01N 33/68* (2006.01)

*G01N 33/53* (2006.01)

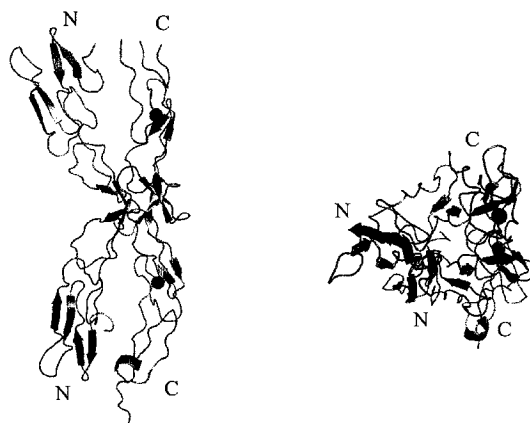
权利要求书 2 页 说明书 12 页 附图 6 页

(54) 发明名称

鼠 RANK 胞外蛋白晶体

(57) 摘要

本发明提供了一种鼠 RANK 胞外蛋白晶体,以及该 RANK 胞外蛋白晶体在医药领域中的应用。



1. 一种胞外区域蛋白的晶体,其中所述晶体包括鼠 RANK 胞外区域的氨基酸序列。
2. 根据权利要求 1 所述的晶体,其中所述胞外区域是指鼠 RANK 单体 N-末端从 26 位氨基酸至 210 位氨基酸。
3. 根据权利要求 1 所述的晶体,其中所述晶体由鼠 RANK 单体链 A 及单体链 B 形成的二聚体构成。
4. 根据权利要求 1-3 任一项所述的晶体,其中所述晶体具有空间群  $P2_12_12_1$ , 晶胞参数为:  $a=39.82 \text{ \AA}$ ,  $b=94.25 \text{ \AA}$  和  $c=102.42 \text{ \AA}$ 。
5. 根据权利要求 1-3 中任一项所述的晶体,其中所述二聚体以反向平行形式存在;所述单体链 A 及单体链 B 含有四个 CRD 结构域,存在于所述四个 CRD 结构域中的多个二硫键使得所述两个单体链形成细长的带状;所述两个单体链的重叠所形成的均方根差为  $2.13 \text{ \AA}$ , 除去溶剂后的二聚体所形成的总面积为  $3231 \text{ \AA}^2$ ;所述两个单体之间所包含的面积被等分,所述二聚体的长度与所述 RANK 的单体链长度近似。
6. 根据权利要求 5 所述的晶体,其中在所述二聚体中,所述四个 CRD 结构域具有不同的均方根差值,所述均方根值的范围为对于 C  $\alpha$  原子从 CRD3 的  $0.35 \text{ \AA}$  至 CRD2 的  $1.05 \text{ \AA}$ 。
7. 根据权利要求 5 所述的晶体,其中所述单体链 A 以及单体链 B 主要在所述 CRD2 和 CRD3 结构域发生相互作用,涉及到所述相互作用的重要残基为所述单体链 A 的 Gln-145、Asn-147 和 Lys-171 以及所述单体链 B 的 Asn-147、Gln-145 和 Glu-84。
8. 根据权利要求 1-3 中任一项所述的晶体,其中所述单体链 A 以及所述单体链 B 分别存在一个钠离子,所述钠离子可与所述两个单体链上的 Cys-134、Ala-135、Phe-138、Val-163 以及骨架上的 Ser-161 侧链的羰基基团螯合。
9. 一种制备 RANK 胞外蛋白晶体的方法,包括:(1) 将纯化的 RANK 在 pH 7.00 的 1M TRIS 缓冲溶液中浓缩成 10mg/mL;  
(2) 在温度为 294K 的条件下,利用高通量纳升坐滴蒸汽扩散方法形成 RANK 胞外蛋白的晶体;3) 所述 RANK 晶体在 20% PEG 3350, 0.2M 酒石酸钠条件下生长。
10. 根据权利要求 9 所述的方法,其中所述 RANK 晶体在 30% 的 PEG 5000, 0.2M 的硫酸钠,以及 pH6.5 的 0.1M MES 的条件下生长。
11. 根据权利要求 9 所述的方法,其中所述 RANK 晶体在含有 100nL 蛋白质溶液和 100nL pH 为 6.5 的 10% PEG 3350, 15% PEG 5000, 0.1M 硫酸铵, 0.1M 酒石酸钠和 0.05M MES 的池液的液滴中生长。
12. 权利要求 1-3 中任一项所述的鼠 RANK 胞外蛋白晶体在筛选治疗骨吸收类疾病或抗骨吸收药物中的应用,包括:  
根据所述蛋白晶体的结构,通过计算机模拟来筛选可能结合到特定部位的任一种多肽、蛋白质、抗体、免疫结合物、无机物或有机物;  
根据该蛋白晶体的结构,通过计算机模拟来筛选可能结合到特定部位的已知多肽、蛋白质、抗体、免疫结合物、无机物或有机物;  
其中与所述 RANK 胞外蛋白序列具有至少 50% 同源性的任何 RANK 胞外蛋白结合的多肽、蛋白质、抗体、免疫结合物、无机物或有机物或者与 RANK 胞外蛋白的亲合力  $PD_2$  值  $> 4.0$  的多肽、蛋白质、抗体、免疫结合物、无机物或有机物为候选物;

通过生物学实验确定所述候选化合物在治疗骨吸收类疾病或抗骨吸收药物中的应用。

## 鼠 RANK 胞外蛋白晶体

### 技术领域

[0001] 本发明涉及一种蛋白质晶体,尤其涉及一种鼠 RANK 胞外蛋白晶体,本发明还提供了一种制备 RANK 胞外蛋白晶体的方法,以及该 RANK 胞外蛋白晶体在医药领域中的应用。

### 背景技术

[0002] 肿瘤坏死因子 (TNF) 家族中的细胞因子在诸如炎症反应、器官发生、宿主防御、自身免疫和细胞凋亡等多种生物功能中扮演着重要的角色。这些生物调节因子是通过受体-配体相互作用,从而引发细胞内的信号传递以及细胞表型的改变而发挥相应的生理作用的。作为 TNF 家族成员的 RANKL 以及其受体 RANK 在生物体内的多种活动中扮演着重要的角色,其对于骨的重塑过程起到了重要的调节作用,并且 RANK 以及 RANKL 分子之间的相互作用还可调节 T 细胞/树突状细胞的信号传导、树突状细胞的存活以及淋巴组织的发育,近年来受到了广泛的关注。

[0003] 在人体骨的重塑过程中,破骨细胞是骨重塑的“启动子”,成骨细胞是骨重塑的“调节者”,骨吸收类疾病共同的病理特征为破骨细胞调节异常导致骨形成和骨吸收之间丧失平衡,因此深入了解破骨细胞的激活机制将对于成功防治骨吸收类疾病的发生有着至关重要的意义。1997 年,几个不同的研究小组发现了肿瘤坏死因子受体-配体超家族新成员:骨保护蛋白 (osteoprotegerin, OPG)、RANKL 以及 RANK;随后,多项研究相继显示:RANKL、OPG 具有调节破骨细胞分化、发育、影响其功能的作用,RANK 是 RANKL、OPG 发挥作用的关键,它们形成一个骨调节轴,来调节骨的重塑过程。

[0004] RANKL 为一种 II 型跨膜蛋白,其多为三个单体聚合成的复合体,主要分布在活化 T 细胞、骨髓基质细胞及成骨细胞等细胞中,目前认为哺乳动物中存在三种 RANKL 蛋白:RANKL1、RANKL2、RANKL3,其中前两种为膜结合蛋白,后一种为被分泌到胞外的可溶性蛋白。人和大鼠的 RANKL 分子分别包含 317、316 个氨基酸残基,具有 87% 的同源性;人类 RANKL 基因位于 13q14,其启动子区包含成骨细胞分化的关键转录因子 cbaf-1 的活性。RANKL mRNA 可在多种组织表达,以骨骼和淋巴组织中最高,在骨骼中,RANKL 主要表达于骨小梁和骨髓;在骨髓基质细胞和成骨细胞均可检测到 RANKL mRNA。2004 年 Ito S 等人以及国际专利申请 NO. WO 03/014077 都披露了 RANKL 的三维结构,其全部内容结合于此作为参考。各项研究表明:每一个 RANKL 单体包括一个夹层结构,其由两个平面反平行  $\beta$ -折叠片层构成。第一个片层由  $\beta$ -链 A'、A、H、C 及 F 组成,第二个片层由 B'、B、G、D 及 E 组成,其中第一片层为内层  $\beta$ -片层,该片层富含疏水性氨基酸,并通过疏水作用与另外两个 RANKL 相互结合,而 B'、B、G、D 及 E 链形成外层  $\beta$ -片层,该片层多为带电及极性氨基酸,负责与受体 RANK 结合;每一个 RANKL 单体中的  $\beta$ -夹层结构的一个边缘、链 E 及 F 包裹相邻单体的 AHCF  $\beta$ -片层的内部疏水层从而形成 RANKL 三聚体结构。RANKL 三聚体的晶体结构特征是:具有  $P2_12_12_1$  的空间群,晶胞参数为  $a=65.3\text{\AA}\pm 0.2\text{\AA}$ ,  $b=82.0\text{\AA}\pm 0.2\text{\AA}$ ,  $c=99.5\text{\AA}\pm 0.2\text{\AA}$ ,以及空间群 R3,胞晶为  $a=b=150.6\text{\AA}\pm 0.2\text{\AA}$ ,  $c=139.5\text{\AA}\pm 0.2\text{\AA}$ 。其形状就像被截平了的锥形,靠近膜的一端宽于远端的顶点处。RANKL 三聚体高  $55\text{\AA}$ ,底部直径约为  $55\text{\AA}$ ,顶点处的直径

约为35 Å。

[0005] RANK 是 RANKL 的受体,属于 TNF 受体家族,是一种 I 型跨膜蛋白,鼠 RANK 分子在破骨细胞前体细胞和成熟破骨细胞内表达量较高。RANK 基因定位于 18q22.1,其编码的蛋白质在体内以跨膜蛋白型存在,其表达于单核-巨噬细胞系、破骨细胞前体细胞、T 淋巴细胞、B 淋巴细胞、树突状细胞等,表达于破骨细胞前体细胞膜表面的 RANK 介导 RANKL 的信号,发挥调节破骨细胞分化的功能。

[0006] 当配体 RANKL 与受体 RANK 的胞外区域结合后,导致了 RANK 的激活,活化后的 RANK 通过其胞内区域的 TRAF 结合域与 TRAF1、2、3、5、6 等蛋白结合,并且通过这些 TRAF 蛋白激活细胞内的两条重要的信号通路:JNK/SAPK 通路和 NF $\kappa$ B 信号通路。这两条信号通路的激活诱导了与破骨细胞分化相关的基因的表达,最终促进了破骨细胞的分化与成熟。而且,各种骨代谢调节因子、激素参与 RANKL-RANK-OPG 轴信号的调节。与其它信号途径一样,RANKL-RANK-OPG 轴信号也存在正反馈和负反馈调节。 $1\alpha,25-(OH)_2D_3$ 、甲状旁腺素、前列腺素 E2、白介素 1、IL-6 等骨吸收因子分别通过维生素 D 受体介导通路、蛋白激酶 A 介导通路、gp130/STAT 介导通路上调成骨细胞/骨髓基质细胞 RANKL mRNA 表达;IL-1、TNF- $\alpha$  分别通过激活破骨细胞表面的 IL-1R 和 TNFR1 加强 RANKL 信号;TGF- $\beta$  则可增加破骨细胞表面的 RANK 上调 RANKL 信号。

[0007] 作为 RANKL-RANK-OPG 信号系统调控破骨细胞分化的中心环节,RANKL 与 RANK 之间的相互作用及其促使 RANK 活化的微观机制一直以来是骨吸收类疾病研究领域的热点问题。其焦点问题之一是 RANKL、RANK 蛋白以及 RANK/RANKL 复合体的晶体结构。虽然目前已经获得了 RANKL 蛋白的晶体结构,但是仍然没有获得 RANK 蛋白以及 RANK/RANKL 复合体的晶体结构,正是由于对这些结构的认识不足,所以学者们对于 RANKL-RANK-OPG 信号体系如何被调控,以及受调控后 RANKL-RANK-OPG 执行信号传递的微观机制尚无明确的认识。这严重影响了有关骨吸收类疾病致病机制的探索以及抗骨吸收药物的研究。

[0008] 分析目前还未获得 RANK 蛋白以及 RANK/RANKL 复合体的晶体结构的原因主要有以下两点:1. RANK 蛋白较难于制备和纯化。在 RANK 被发现至今的时间里,一直无 RANK 蛋白 E. coli 表达以及纯化的相关报道。2. 相对于包含 RANKL 在内的 TNF 蛋白家族成员所具有的较刚性的结构,包括 RANK 在内的 TNFR 蛋白家族成员其结构具有较大的可变性,因而更难于获得结晶。

## 发明内容

[0009] 本发明涉及一种蛋白质晶体,其中所述晶体包括鼠 RANK 胞外区域的氨基酸序列。鼠 RANK 的胞外区域是指 RANK 单体 N-末端从 26 位氨基酸至 210 位氨基酸。

[0010] 本发明所提供的晶体由鼠 RANK 单体链 A 及单体链 B 形成的二聚体构成。

[0011] 根据本发明所提供的鼠 RANK 胞外蛋白晶体,该晶体具有空间群  $P2_12_12_1$ ,晶胞参数为: $a=39.82\text{ \AA}$ , $b=94.25\text{ \AA}$ 和 $c=102.42\text{ \AA}$ 。

[0012] 在本发明所提供的晶体中,单体链 A 及单体链 B 以反向平行形式,即单体链 A 的 N 端对应于单体链 B 的 C 端而形成二聚体;该二聚体含有四个 CRD,存在于这四个 CRD 结构域中的多个二硫键使得单体链 A 和 B 形成细长的带状;两个单体的重叠所形成的均方根差为  $2.13\text{ \AA}$ ,除去溶剂后的二聚体所形成的总面积为  $3231\text{ \AA}^2$ ;由两个单体之间所包含的面积

被这两个单体等分；形成的二聚体的长度与 RANK 单体的长度近似。

[0013] 在本发明所提供的晶体中，在所述二聚体中，每个不同的 CRD 结构域具有不同的均方根差，相对于 C $\alpha$  原子而言，该均方根差的范围为从 CRD3 的 0.35 Å 至 CRD2 的 1.05 Å。

[0014] 在本发明所提供的晶体中，单体链 A 以及单体链 B 主要在 CRD2 和 CRD3 结构域发生相互作用，涉及到这种相互作用的重要残基为单体链 A 的 Gln-145, Asn-147 和 Lys-171 以及单体链 B 的 Asn-147, Gln-145 和 Glu-84。

[0015] 在本发明所提供的晶体中，单体链 A 及单体链 B 中分别存在一个钠离子，该钠离子可与单体链 A 和 B 上的 Cys-134, Ala-135, Phe-138, Val-163 以及骨架上的 Ser-161 侧链的羰基基团螯合。

[0016] 本发明还提供了一种制备 RANK 胞外蛋白晶体的方法，包括：(1) 将纯化的 RANK 在 pH 7.00 的 1M TRIS 缓冲溶液中浓缩成 10mg/mL；(2) 在温度为 294K 的条件下，利用高通量纳升坐滴蒸汽扩散方法形成 RANK 胞外蛋白的晶体；(3) 所述 RANK 晶体在 20% PEG 3350, 0.2M 酒石酸钠条件下生长。本发明提供的该制备方法尤其适用于鼠 RANK 胞外蛋白晶体的制备。

[0017] 根据本发明提供的制备 RANK 晶体的方法，其中所述 RANK 晶体还可以在 30% 的 PEG 5000, 0.2M 的硫酸钠，以及 pH6.5 的 0.1M MES 的条件下生长。该生长条件尤其适用于鼠 RANK 胞外蛋白晶体的生长。

[0018] 根据本发明提供的制备方法，其中所述 RANK 晶体可以在含有 100nL 蛋白质溶液和 100nL pH 为 6.5 的 10% PEG 3350, 15% PEG5000, 0.1M 硫酸铵, 0.1M 酒石酸钠和 0.05M MES 的池液的液滴中生长。该生长条件尤其适用于鼠 RANK 胞外蛋白晶体的生长。

[0019] 本发明所提供的鼠 RANK 胞外蛋白晶体在筛选治疗骨吸收类疾病或抗骨吸收药物中的应用，包括：根据该蛋白晶体的结构，通过计算机模拟来筛选可能结合到特定部位的任一种多肽、蛋白质、抗体、免疫结合物、无机物或有机物；根据该蛋白晶体的结构，通过计算机模拟来筛选可能结合到特定部位的已知多肽、蛋白质、抗体、免疫结合物、无机物或有机物；其中与该鼠 RANK 胞外蛋白序列具有至少 50% 同源性的任何 RANK 胞外蛋白结合的多肽、蛋白质、抗体、免疫结合物、无机物或有机物或者与 RANK 胞外蛋白的亲和力  $PD_2$  值  $> 4.0$  的多肽、蛋白质、抗体、免疫结合物、无机物或有机物为候选物；通过生物学实验确定所述候选化合物在治疗骨吸收类疾病或抗骨吸收药物中的应用。

## 附图说明

[0020] 图 1A 示出了未结合状态的鼠 RANK 胞外区域二聚体的三维结构；

[0021] 图 1B 示出了鼠 RANK 单体链 A 和单体链 B 的重叠；

[0022] 图 1C 示出了参与到胞外区域二聚体相互作用的重要残基；

[0023] 图 1D 是在图 1C 的基础上加有深色背景，以更清楚的识别二聚体相互作用中的重要残基；

[0024] 图 2 示出了未结合与结合 RANKL 状态下 RANK 在三维构象上的差异；

[0025] 图 3 示出了存在于鼠 RANK 单体链中的钠离子以及其与其它氨基酸之间的相互作用。

## 具体实施方式

[0026] 一、表达及纯化

[0027] 1. 鼠 RANK 胞外蛋白的表达与纯化

[0028] 鼠 RANK 的氨基酸序列如下：

[0029] MAPRARRRRQLPAPLLALCVLLVPLQVTLQVTPPCTQERHYEHLGRCCSRCEPGKYLSSKCTPTSDSVC  
LPCGPDEYLDTWNEEDKCLLHKVCDAGKALVAVDPGNHTAPRRCACTAGYHWNSDCECCRRNTECAPGFGAQHPLQL  
NKDTVCTPCLLGGFSDVFSSTDKCKPWTNCTLLGKLEAHQGTTESDVVCSSSMTLRRPPKEAQAYLPSLIVLLLFIS  
 VVVVAATIFGVYYRKGGKALTANLWNWVNDACSSLSGNKESSGDRCAGSHSATSSQQEVCEGILLMTREEKMPEDG  
 AGVCGPVCAAGGPWAEVRDSRTFTLVSEVETQGDLRKIPTETEYDTRPSQPSTGSLLLIQGSKSIPPFQEPLVQ  
 ENDSLQCFGTGTESTVDSEGCDFTEPPSRTDSMPVSPEKHLTKEIEGDSCLPWVSSNSTDGYTGSNTPGEDHEPF  
 PGLSKCGPLPQCAYSMGFPSEAAAASMAEAGVRPQDRADERGASGSGSSPSDQPPASGNTGNSNSTFI SSGQVMNF  
 KGDIIVVVYSQTSQEGPGSAEPESEPVGRPVQEETLAHRDSFAGTAPRFPDVCATGAGLQEQGAPRQKDGTSRPVQE  
 QGGAQTSLHTQSGQCAE

[0030] 划线部分为鼠 RANK 胞外区域的氨基酸序列

[0031] 材料

[0032] 寡聚核苷酸由 Sangon Biotech(China) 制备而成；从 Fermentas 购得限制性内切酶、T4DNA 连接酶以及第一链 cDNA 合成试剂盒；从 Tiangen Biotech(China) 购得 pfu DNA 聚合酶；从 Sigma 购得谷胱苷肽（还原型和氧化型）；谷胱苷肽琼脂糖凝胶 4B 获自通用电气医疗集团（GE Healthcare）；TRIZOL 试剂购自 Invitrogen, Inc.；所有其它得化学物质均为分析纯。

[0033] 质粒及菌株

[0034] 编码成熟鼠 RANK 胞外区域（残基 26-210）的 cDNA 为鼠 RAW264.7 细胞的 mRNA 经逆转录 PCR 技术所获得，并且克隆到 pET28 载体（Novagen）中。用于 GST-RANKL 表达的质粒（残基 159-361）获赠于 Daved H. Fremont 教授（Department of Pathology and Immunology, Washington University School of Medicine, USA.），用 E. coli 菌株 BL21-Gold（购买于 Stratagene 公司）表达重组蛋白。

[0035] 表达及纯化

[0036] 在 E. coli 菌株 BL21-Gold(DE3) 中，重组 RANK 主要以包涵体形式表达。该包涵体通过超声处理而被纯化并且重新溶解在 6M 的盐酸胍中。重组 RANK 的重新折叠通过如下步骤而实现：该包涵体稀释在包含 20mM 的  $\text{Na}_2\text{HPO}_4$  (pH 7.3)、1M L-精氨酸、20%甘油、10mM 还原型谷胱苷肽以及 1mM 氧化型谷胱苷肽的复性 (IB) 溶液中，然后在含有 20mM 的  $\text{Na}_2\text{HPO}_4$  (pH 7.3)、0.5M L-精氨酸、10%甘油的重新折叠缓冲液 1 中 4℃ 下透析 12 小时，然后在含有 20mM 的  $\text{Na}_2\text{HPO}_4$  (pH 7.3)、0.2M L-精氨酸、5%甘油的重新折叠缓冲液 2 中 4℃ 下透析 12 小时，最后在 20mM 的  $\text{Na}_2\text{HPO}_4$  (pH 7.3)、0.2M L-精氨酸中 4℃ 下透析 12 小时，在 20000g 离心 10 分钟后，将上清液用琼脂 75(Superdex 75) 色谱柱（购自 Amersham pharmacia 公司）进行纯化，并且收集正确折叠的鼠 RANK，并用 SDS-PAGE 分析。

[0037] 2. 鼠 RANKL 的表达和纯化

[0038] 鼠 RANKL 的氨基酸序列如下：

[0039] MRRASRDYGYLRSSEEMGSGPVPHEGPHLPAPSAPAPAPPPAASRSMFLALLGLGLGQVVCIALF

LYFRAQMDPNRI SEDSTHCFYRILRLHENADLQDSTLESEDTLPDSCRRMKQAFQGA VQKELQHI VGPQRFSGAPA  
MMEGSWLDVAQRGKPEAQPF AHLTINAASIPSGSHKVTLS SWYHDRGWAKISNMTLSNGKLRVNQDGFYYLYANI  
CFRHETSGSVPTDYLQLMVYVVKTSIKIPSSHNL MKGGSTKNWSGNSEFHFYSINVGGFFKLRAGEEISIQVSN  
PSLLDPDQDATYFGAFKVQDID

[0040] 划线部分为鼠 RANKL 胞外区域的氨基酸序列

[0041] 鼠 RANKL 的可溶性胞外区域是以谷胱苷肽巯基转移酶融合蛋白形式表达的;融合蛋白利用谷胱苷肽琼脂糖凝胶 4B(购自 Amersham pharmacia 公司)进行亲和纯化,用 PreScission 蛋白酶进行切割;酶切所得的 RANKL 用琼脂糖 75(Superdex 75)色谱柱(购自 Amersham pharmacia 公司)纯化。

[0042] 二、RANK 胞外区域蛋白的结晶化

[0043] 将最终纯化的 RANK 在 pH 7.00 的 1M TRIS 缓冲溶液中浓缩成 10mg/mL。在温度为 294K 的条件下,利用高通量纳升坐滴蒸汽扩散方法(sitting drop vapour diffusion method)形成 RANK 胞外区域的晶体网格(crystallization screen),并且利用牛津蛋白产物工具中的自动结晶图像系统进行监控。

[0044] 得到的晶体初始在两个条件下成长:1)20%的 PEG 3350,0.2M 酒石酸钠;2)30%的 PEG 5000、0.2M 的硫酸钠,以及 pH6.5 的 0.1M 生物缓冲液 MES。在条件 1)下的晶体生长很难重复,在条件 2)下很容易再生长,但它们多以多倍体形式存在。通过实验发现,在含有 100nL 蛋白质溶液和 100nL pH 为 6.5 的 10% PEG 3350,15% PEG 5000,0.1M 硫酸铵,0.1M 酒石酸钠和 0.05M MES 的池液的液滴中,可以得到晶体生长最好的鼠 RANK 胞外区域蛋白晶体。

[0045] 三、鼠 RANK/RANKL 胞外区域复合体晶体的形成与结构测定

[0046] 以 1:1 的比例混合两种蛋白而制的 RANK/RANKL 复合体蛋白溶液。在温度为 294K 的条件下,利用高通量纳升坐滴蒸汽扩散方法形成 RANK/RANKL 胞外区域的晶体网格(crystallization screen),并且利用牛津蛋白产物工具中的自动结晶图像系统进行监控(Mayo 等 2005, Walter 等 2005)。在以下两个优选条件下得到了生长情况最好的晶体:1)0.1M 磷酸二氢钠,2M 氯化钠,0.1M 二氢磷酸钾和 0.1M MES(pH 6.5);2)20%的 PEG 3350 以及 0.2M 的甲酸铵,结晶液滴中含有 100nL 蛋白溶液以及 100nL 池液。

[0047] 在条件 1 下得到 RANK/RANKL 复合体晶体的 X-射线数据是在 ESRF,利用光束线 ID23-EH2 得到的。在 0.873Å 下,从单晶的 2 个位置收集到衍射震荡为 1.0° 的 180 个图像。在波长 1.060Å 下,利用钻石的光束线 103 得到了条件 2 下生长的复合体晶体的 X-射线数据。25%甘油作为冷冻保护剂加入到结晶液滴中,晶体被冷冻起来并且在整个数据采集阶段,采用氮气的冷冻流来保持晶体处于 100K。利用 HKL2000(Ref.)对数据图像进行索引,积分和合并。采用计算机程序分析得到 RANK/RANKL 复合体晶体中 RANK 的构型模型。

[0048] 四、RANK 晶体结构的测定

[0049] 在法国 Grenoble 的欧洲分子生物学实验室(EMBL)法国分部(ESRF),利用光束线 BM14 测得 RANK 的 X-ray 数据。在波长 0.954Å 下,从单晶中收集到衍射震荡为 1.0° 的 180 个图像。25%甘油作为冷冻保护剂加入到结晶液滴中,晶体被冷冻起来并且在整个数据采集阶段,采用氮气的冷冻流将晶体保持在 100K 的温度。利用 HKL2000(Ref.)对数据图像进行索引,积分和合并。表 1 中给出了 X-ray 衍射数据。

[0050] 五、结果解析：

[0051] 1. 晶体结构测试结果

[0052] 表 1X-ray 衍射结果

数据	鼠 RANK 胞外区域蛋白晶体
X-射线源	BM14, ESRF
波长 (Å)	0.9537
空间群	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
晶胞 (a,b,c[Å])	39.82, 94.25, 102.42
分辨距离 (Å)	30.0-2.00 (2.07-2.00)
独立反射	26472 (2582)
完整性 (%)	99.6 (9.2)
冗余值	6.9(6.7)
[0053] 平均值 $I/\sigma I$	17.7(2.3)
Rmerge	0.096 (0.644)
精确统计	
分辨距离 (Å)	30.0-2.00
反射数目 (运作、测试)	24896、1332
R-因子 ( $R_{work}/R_{free}$ )	0.216/0.265
原子数目 (蛋白/水/其他)	2433/221/47
键长均方根差 (Å)	0.007
键角均方根差 (°)	1.5
平均 B 因子[Å <sup>2</sup> ]	31/35/49

[0054] 括号中的数据是高分辨率的结果。 $R_{work}$  和  $R_{free}$  定义为  $R = \frac{\sum_{hkl} ||F_{obs}| - |F_{calc}||}{\sum_{hkl} |F_{obs}|}$ , 其中 h, k, l 是指反射指数 (用于  $R_{work}$  精细化; 5%, 不用于  $R_{free}$  的精细),  $F_{obs}$  和  $F_{calc}$  是结构因子, 从测量的强度中推论出来并通过模型计算出来。

[0055] 平均 B 因子用于蛋白, 水和其他原子

[0056] 鼠 RANK 胞外区域蛋白晶体的三维晶体结构的参数为: 空间群组为 P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub>, 晶胞的尺寸为  $a=39.82\text{Å}$ ,  $b=94.25\text{Å}$  和  $c=102.42\text{Å}$ 。

[0057] 2. 鼠 RANK 胞外区域蛋白的结构解析

[0058] 2.1 解析方法

[0059] 在利用同步辐射收集相关数据后, 利用鼠 RANK/RANKL 复合体晶体中的 RANK 模型作为起始模型, 利用 CCP4 软件包中 MOLREP 程序通过分子置换法首次解析了鼠 RANK 胞外区域蛋白的三级结构。RANK 单体具有作为其家族的代表性结构的 TNF 样 jelly 折叠。带有 2,6-His 标记的 RANK 的分子质量为 23.5kDa。从建立的部分模型中计算出来的初步衍射电子密度图谱清楚的显示出了 RANK 多肽, 并且该图谱显示出能够构建 RANK 胞外区域蛋白分子的 4CRD 区域的特性。

[0060] 利用复合体中的 RANK 模型解析未结合配体的 RANK 胞外区域的晶体结构。但是, 由于 RANK 薄而狭长的形状以及分子的柔性大, 利用分子替代方法对结构的解析并不容易。在利用程序 PHASER(ccp4 ref) 定位中心 2CRD 区域后, 才得到未结合配体的 RANK 的正确解析结果。对于从 30 至  $4.0\text{Å}$  的数据, 初始的 R 因子是 0.54。

[0061] 2.2 未结合配体的鼠 RANK 胞外区域的三级结构

[0062] 对于未结合配体的鼠 RANK 胞外区域蛋白结构,在一个不对称晶胞中的两个分子是通过局部的 2 次对称旋转轴相关联的,该未结合配体的 RANK 模型与 RANKL-RANK 复合体中的 RANK 模型叠加后发现在第一至第四 CRD 区域之间存在很大的区别,这说明存在较大的构型差异。在修正后,利用所有的数据至  $2.0\text{\AA}$  的分辨率,构建出第一和第四 CRD 区域。

[0063] RANK 最终精细化结构具有良好的晶体学 R 因子以及立体化学,以上结果均已在表 1 中示出。

[0064] 未结合配体的 RANK 胞外区域蛋白以反向平行(头-对-尾)的二聚体形式存在,即由单体链 A 及单体链 B 构成(如图 1A)。每个二聚体具有四个 CRD,存在于 CRD 结构域中的多个二硫键使得 RANK 分子形成细长的带状。利用 Vega ZZ 2.1 软件分析可知两个单体链的均方根差(RMSD)为  $2.13\text{\AA}$ 。不同的结构域具有不同的 RMSD,相对于  $C\alpha$  原子,RMSD 的范围从 CRD3 中的  $0.35\text{\AA}$  至 CRD2 中的  $1.05\text{\AA}$ (如图 1B)。两个反向平行的分子互相重叠,形成的二聚体的长度接近于单体的长度。除去溶剂后的二聚体形成的总面积约为  $3231\text{\AA}^2$ ,并且两个单体包含的面积被平分。利用 Pymol 软件分析可知在 RANK 二聚体中,主要在两个单体链的 CRD2 和 CRD3 之间通过范德化力、氢键和盐键,形成亚单位之间的相互作用。参与到这种相互作用的重要残基是单体链 A 的 Gln-145、Asn-147 和 Lys-171 以及单体链 B 的 Asn-147、Gln-145 和 Glu-84,它们之间形成极性相互作用(见图 1C 及图 1D)。

[0065] 利用 Pymol 软件分析,结合状态以及未结合配体状态的 RANK 分子之间的差别如下:(1) 复合体中的 RANK 受体单体是被拉直的,与它们的未结合状态相比,弯曲度减少。未结合状态与复合体状态相比,位置 Lys-91 处的扭转角度发生 60 度的变化。(2) 螺旋构象中的环 100S(残基 102-112)在 RANK 与 RANKL 结合的过程中发生了移动,该移动起始于位置 Val-102,扭转角度发生约 50 度的改变,并在位置 Cys-113 处终止。这种改变有助于稳定受体与配体结合的最适宜的构型。(3) 在螺旋结构中,环 120S 仅发生很小的移动,并且在受体的两种状态中,扭转角度仅发生 15 度的改变。(4) 通过结合,RANK 的 C 末端发生了一些构型的改变。两个区域,即残基 182-185,以及残基 194-196 的自由卷曲构型转化成复合体中的 2 个较短的  $\beta$ -片层构型,从而形成了较为有序的结构域。(5) 参与到形成 RANK 二聚体中单体链 A 及单体链 B 之间的相互作用的大多数重要残基,如 Gln-145, Asn-147 以及 Lys-171 都位于 RANK/RANKL 复合体的外螺旋区域,并不直接参加 RANK 与 RANKL 的相互作用。

[0066] 在本研究中,未结合配体的 RANK 胞外区域蛋白在 pH6.5 下结晶形成反向平行二聚体,这一点向人们示出,当受体没有与配体结合时, TNF 受体家族成员的胞外区域之间可形成二聚体的特性。但是, RANK 的构型并不同于 TNFR-1 二聚体的两个反向平行的构型。在 pH7.5 下,在 TNFR-1 二聚体的反向平行构型中,两个单体仅在每个分子的 N 末端部分重叠,而本发明提供的 RANK 二聚体的反向平行构型产生了非常强烈的相互作用。在 pH 3.7 下,尽管反向平行的 TNFR-1 二聚体中的两个单体是完全重叠的,但两个单体以左手双螺旋的形式缠绕在一起,而在反向平行的 RANK 二聚体中,这两个单体形成右手双螺旋。未结合配体的 RANK 构型的差异也表明了 TNF 受体家族成员的特异性。

[0067] 有趣的是,研究发现,在形成的二聚体中,在 CRD3 结构域以及 CRD4 结构域之间的螺旋区域中分别观察到了一个与之结合的金属离子,根据结合位点的平均键长,计算得

到的该金属离子的平均钙键 - 化合价总和 (CBVS, calcium bond-valence sum) 为 1.2, 对于钠离子, CBVS 的预期值为 1.6, 钾离子为 0.6, 钙离子为 2.0, 锰离子为 3.2, 镁离子为 4.2 (Mulle 等, 2003), 所以推断该金属离子为钠离子; 通过在波长 1.698 Å 下收集到的数据计算出的异常绕射差异图 (anomalous difference map), 该推断被进一步证实, 目前为止在其它 TNF 受体家族成员中尚未有相似报道。该钠离子可与 Cys-134、Ala-135、Phe-138 和 Val-163 以及骨架 Ser-161 侧链的羰基基团整合。(见图 3)

[0068] 结合以上说明, 本发明揭示出了鼠 RANK 胞外区域蛋白的晶体结构。从上述晶体结构中, 不难找到单体链 A 和单体链 B 相互作用中的重要残基。根据 NCBI 网站 (www.ncbi.nlm.nih.gov) 上 Blast 的分析结果, 鼠与人的 RANK 胞外区域的同源性为 83%, 所以利用这些信息, 我们可以继续探寻 RANK/RANKL 所引起的生物活性的激活途径, 设计出能够调节相关生物激活途径的药物。本发明提供了鼠 RANK 胞外蛋白晶体在筛选治疗骨吸收类疾病或抗骨吸收药物中的应用, 包括: 根据该蛋白晶体的结构, 通过计算机模拟来筛选可能结合到特定部位的任一种多肽、蛋白质、抗体、免疫结合物、无机物或有机物; 根据该蛋白晶体的结构, 通过计算机模拟来筛选可能结合到特定部位的已知多肽、蛋白质、抗体、免疫结合物、无机物或有机物; 其中与该鼠 RANK 胞外蛋白序列具有至少 50% 同源性的任何 RANK 胞外蛋白结合的多肽、蛋白质、抗体、免疫结合物、无机物或有机物或者与 RANK 胞外蛋白的亲合力  $PD_2$  值 > 4.0 的多肽、蛋白质、抗体、免疫结合物、无机物或有机物为候选物; 通过生物学实验确定所述候选化合物在治疗骨吸收类疾病或抗骨吸收药物中的应用。

[0069] 上文中, 基于特定优选的实施方式描述了本发明, 然而在不违背所附权利要求中所限定的本发明的精神和范围的情况下, 可增加多种变化和修改, 这对于本领域中的普通技术人员应当是显而易见的。

[0070] 序列表

[0071] <110> 中国人民解放军总医院

[0072] 北京表源生物技术有限公司

[0073] <120> 鼠 RANK 胞外蛋白晶体

[0074] <130> P21358BYSW

[0075] <160> 2

[0076] <170> PatentIn version 3.2

[0077] <210> 1

[0078] <211> 625

[0079] <212> PRT

[0080] <213> 鼠 RANK 蛋白

[0081] <400> 1

[0082] Met Ala Pro Arg Ala Arg Arg Arg Arg Gln Leu Pro Ala Pro Leu Leu

[0083] 1 5 10 15

[0084] Ala Leu Cys Val Leu Leu Val Pro Leu Gln Val Thr Leu Gln Val Thr

[0085] 20 25 30

[0086] Pro Pro Cys Thr Gln Glu Arg His Tyr Glu His Leu Gly Arg Cys Cys

[0087] 35 40 45

[0088]	Ser Arg Cys Glu Pro Gly Lys Tyr Leu Ser Ser Lys Cys Thr Pro Thr
[0089]	50 55 60
[0090]	Ser Asp Ser Val Cys Leu Pro Cys Gly Pro Asp Glu Tyr Leu Asp Thr
[0091]	65 70 75 80
[0092]	Trp Asn Glu Glu Asp Lys Cys Leu Leu His Lys Val Cys Asp Ala Gly
[0093]	85 90 95
[0094]	Lys Ala Leu Val Ala Val Asp Pro Gly Asn His Thr Ala Pro Arg Arg
[0095]	100 105 110
[0096]	Cys Ala Cys Thr Ala Gly Tyr His Trp Asn Ser Asp Cys Glu Cys Cys
[0097]	115 120 125
[0098]	Arg Arg Asn Thr Glu Cys Ala Pro Gly Phe Gly Ala Gln His Pro Leu
[0099]	130 135 140
[0100]	Gln Leu Asn Lys Asp Thr Val Cys Thr Pro Cys Leu Leu Gly Phe Phe
[0101]	145 150 155 160
[0102]	Ser Asp Val Phe Ser Ser Thr Asp Lys Cys Lys Pro Trp Thr Asn Cys
[0103]	165 170 175
[0104]	Thr Leu Leu Gly Lys Leu Glu Ala His Gln Gly Thr Thr Glu Ser Asp
[0105]	180 185 190
[0106]	Val Val Cys Ser Ser Ser Met Thr Leu Arg Arg Pro Pro Lys Glu Ala
[0107]	195 200 205
[0108]	Gln Ala Tyr Leu Pro Ser Leu Ile Val Leu Leu Leu Phe Ile Ser Val
[0109]	210 215 220
[0110]	Val Val Val Ala Ala Ile Ile Phe Gly Val Tyr Tyr Arg Lys Gly Gly
[0111]	225 230 235 240
[0112]	Lys Ala Leu Thr Ala Asn Leu Trp Asn Trp Val Asn Asp Ala Cys Ser
[0113]	245 250 255
[0114]	Ser Leu Ser Gly Asn Lys Glu Ser Ser Gly Asp Arg Cys Ala Gly Ser
[0115]	260 265 270
[0116]	His Ser Ala Thr Ser Ser Gln Gln Glu Val Cys Glu Gly Ile Leu Leu
[0117]	275 280 285
[0118]	Met Thr Arg Glu Glu Lys Met Val Pro Glu Asp Gly Ala Gly Val Cys
[0119]	290 295 300
[0120]	Gly Pro Val Cys Ala Ala Gly Gly Pro Trp Ala Glu Val Arg Asp Ser
[0121]	305 310 315 320
[0122]	Arg Thr Phe Thr Leu Val Ser Glu Val Glu Thr Gln Gly Asp Leu Ser
[0123]	325 330 335
[0124]	Arg Lys Ile Pro Thr Glu Asp Glu Tyr Thr Asp Arg Pro Ser Gln Pro
[0125]	340 345 350
[0126]	Ser Thr Gly Ser Leu Leu Leu Ile Gln Gln Gly Ser Lys Ser Ile Pro

[0127]	355	360	365	
[0128]	Pro Phe Gln Glu Pro Leu Glu Val Gly Glu Asn Asp Ser Leu Ser Gln			
[0129]	370	375	380	
[0130]	Cys Phe Thr Gly Thr Glu Ser Thr Val Asp Ser Glu Gly Cys Asp Phe			
[0131]	385	390	395	400
[0132]	Thr Glu Pro Pro Ser Arg Thr Asp Ser Met Pro Val Ser Pro Glu Lys			
[0133]	405	410	415	
[0134]	His Leu Thr Lys Glu Ile Glu Gly Asp Ser Cys Leu Pro Trp Val Val			
[0135]	420	425	430	
[0136]	Ser Ser Asn Ser Thr Asp Gly Tyr Thr Gly Ser Gly Asn Thr Pro Gly			
[0137]	435	440	445	
[0138]	Glu Asp His Glu Pro Phe Pro Gly Ser Leu Lys Cys Gly Pro Leu Pro			
[0139]	450	455	460	
[0140]	Gln Cys Ala Tyr Ser Met Gly Phe Pro Ser Glu Ala Ala Ala Ser Met			
[0141]	465	470	475	480
[0142]	Ala Glu Ala Gly Val Arg Pro Gln Asp Arg Ala Asp Glu Arg Gly Ala			
[0143]	485	490	495	
[0144]	Ser Gly Ser Gly Ser Ser Pro Ser Asp Gln Pro Pro Ala Ser Gly Asn			
[0145]	500	505	510	
[0146]	Val Thr Gly Asn Ser Asn Ser Thr Phe Ile Ser Ser Gly Gln Val Met			
[0147]	515	520	525	
[0148]	Asn Phe Lys Gly Asp Ile Ile Val Val Tyr Val Ser Gln Thr Ser Gln			
[0149]	530	535	540	
[0150]	Glu Gly Pro Gly Ser Ala Glu Pro Glu Ser Glu Pro Val Gly Arg Pro			
[0151]	545	550	555	560
[0152]	Val Gln Glu Glu Thr Leu Ala His Arg Asp Ser Phe Ala Gly Thr Ala			
[0153]	565	570	575	
[0154]	Pro Arg Phe Pro Asp Val Cys Ala Thr Gly Ala Gly Leu Gln Glu Gln			
[0155]	580	585	590	
[0156]	Gly Ala Pro Arg Gln Lys Asp Gly Thr Ser Arg Pro Val Gln Glu Gln			
[0157]	595	600	605	
[0158]	Gly Gly Ala Gln Thr Ser Leu His Thr Gln Gly Ser Gly Gln Cys Ala			
[0159]	610	615	620	
[0160]	Glu			
[0161]	625			
[0162]	<210>2			
[0163]	<211>316			
[0164]	<212>PRT			
[0165]	<213> 鼠 RANKL 蛋白氨基酸序列			

[0166] <400>2  
 [0167] Met Arg Arg Ala Ser Arg Asp Tyr Gly Lys Tyr Leu Arg Ser Ser Glu  
 [0168] 1 5 10 15  
 [0169] Glu Met Gly Ser Gly Pro Gly Val Pro His Glu Gly Pro Leu His Pro  
 [0170] 20 25 30  
 [0171] Ala Pro Ser Ala Pro Ala Pro Ala Pro Pro Pro Ala Ala Ser Arg Ser  
 [0172] 35 40 45  
 [0173] Met Phe Leu Ala Leu Leu Gly Leu Gly Leu Gly Gln Val Val Cys Ser  
 [0174] 50 55 60  
 [0175] Ile Ala Leu Phe Leu Tyr Phe Arg Ala Gln Met Asp Pro Asn Arg Ile  
 [0176] 65 70 75 80  
 [0177] Ser Glu Asp Ser Thr His Cys Phe Tyr Arg Ile Leu Arg Leu His Glu  
 [0178] 85 90 95  
 [0179] Asn Ala Asp Leu Gln Asp Ser Thr Leu Glu Ser Glu Asp Thr Leu Pro  
 [0180] 100 105 110  
 [0181] Asp Ser Cys Arg Arg Met Lys Gln Ala Phe Gln Gly Ala Val Gln Lys  
 [0182] 115 120 125  
 [0183] Glu Leu Gln His Ile Val Gly Pro Gln Arg Phe Ser Gly Ala Pro Ala  
 [0184] 130 135 140  
 [0185] Met Met Glu Gly Ser Trp Leu Asp Val Ala Gln Arg Gly Lys Pro Glu  
 [0186] 145 150 155 160  
 [0187] Ala Gln Pro Phe Ala His Leu Thr Ile Asn Ala Ala Ser Ile Pro Ser  
 [0188] 165 170 175  
 [0189] Gly Ser His Lys Val Thr Leu Ser Ser Trp Tyr His Asp Arg Gly Trp  
 [0190] 180 185 190  
 [0191] Ala Lys Ile Ser Asn Met Thr Leu Ser Asn Gly Lys Leu Arg Val Asn  
 [0192] 195 200 205  
 [0193] Gln Asp Gly Phe Tyr Tyr Leu Tyr Ala Asn Ile Cys Phe Arg His His  
 [0194] 210 215 220  
 [0195] Glu Thr Ser Gly Ser Val Pro Thr Asp Tyr Leu Gln Leu Met Val Tyr  
 [0196] 225 230 235 240  
 [0197] Val Val Lys Thr Ser Ile Lys Ile Pro Ser Ser His Asn Leu Met Lys  
 [0198] 245 250 255  
 [0199] Gly Gly Ser Thr Lys Asn Trp Ser Gly Asn Ser Glu Phe His Phe Tyr  
 [0200] 260 265 270  
 [0201] Ser Ile Asn Val Gly Gly Phe Phe Lys Leu Arg Ala Gly Glu Glu Ile  
 [0202] 275 280 285  
 [0203] Ser Ile Gln Val Ser Asn Pro Ser Leu Leu Asp Pro Asp Gln Asp Ala  
 [0204] 290 295 300

---

[0205] Thr Tyr Phe Gly Ala Phe Lys Val Gln Asp Ile Asp  
[0206] 305 310 315

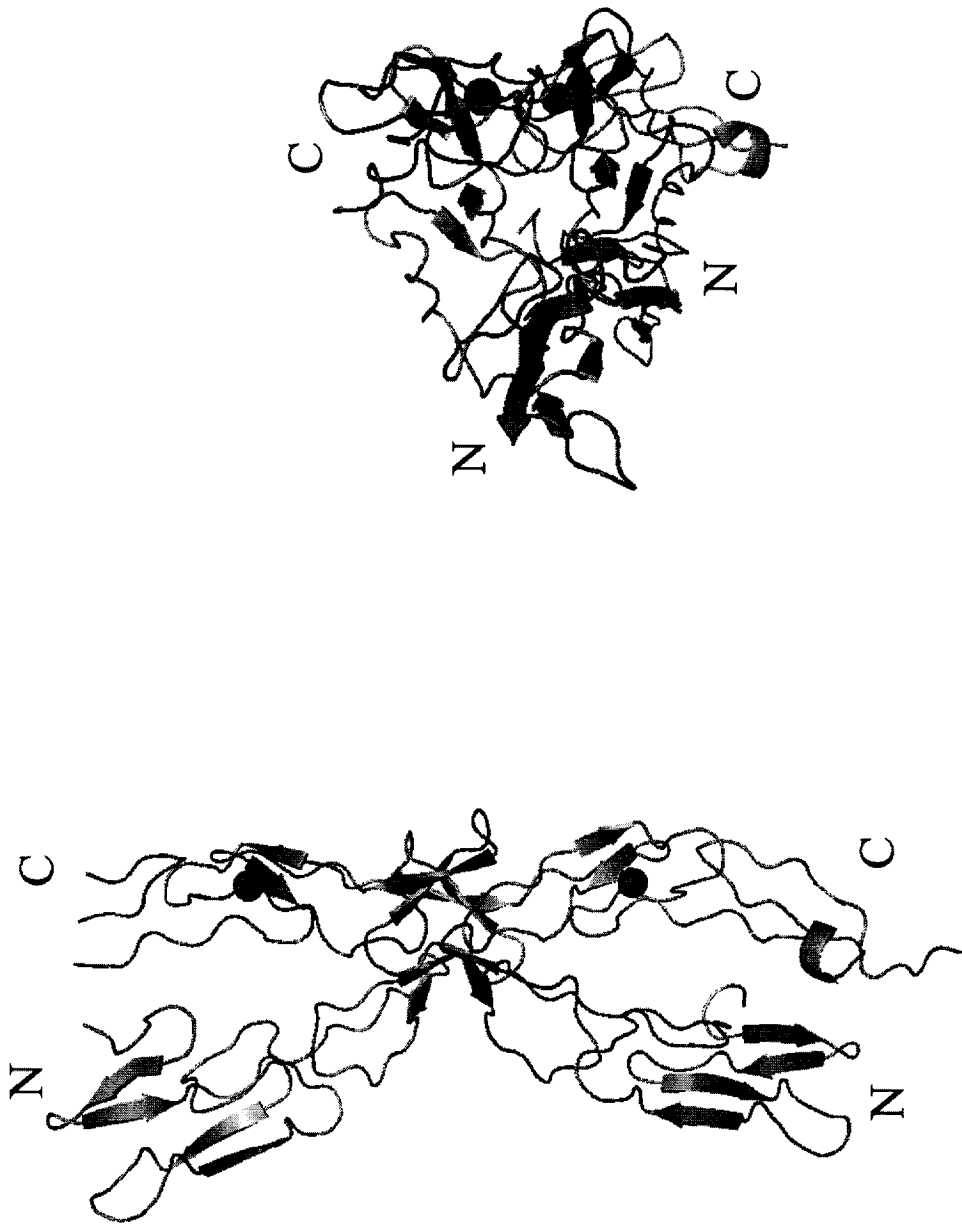


图 1A

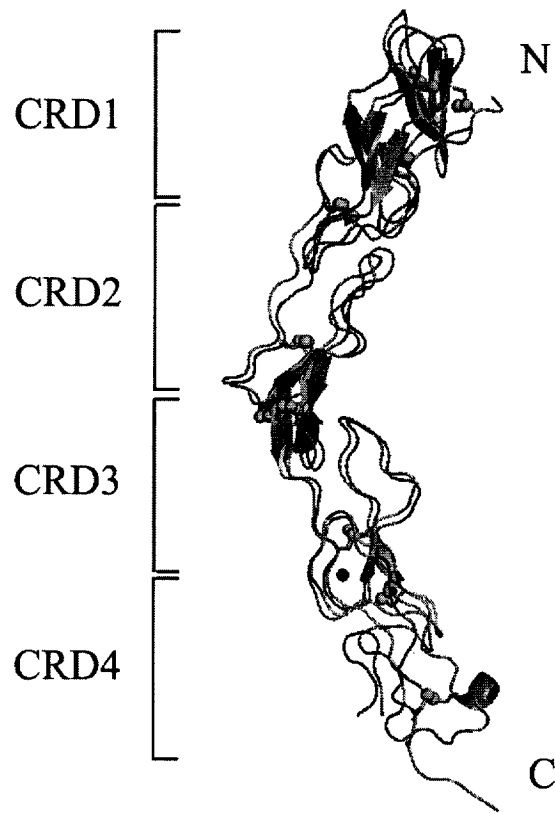
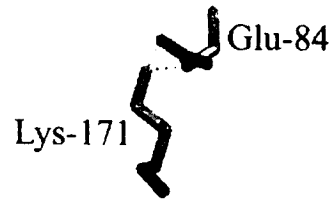


图 1B



单体链 A

单体链 B

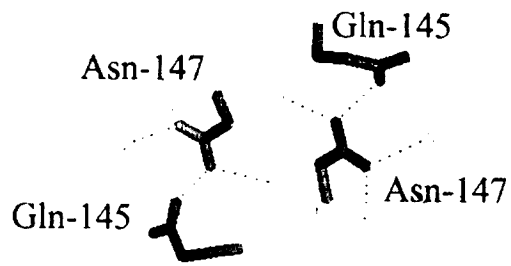


图 1C

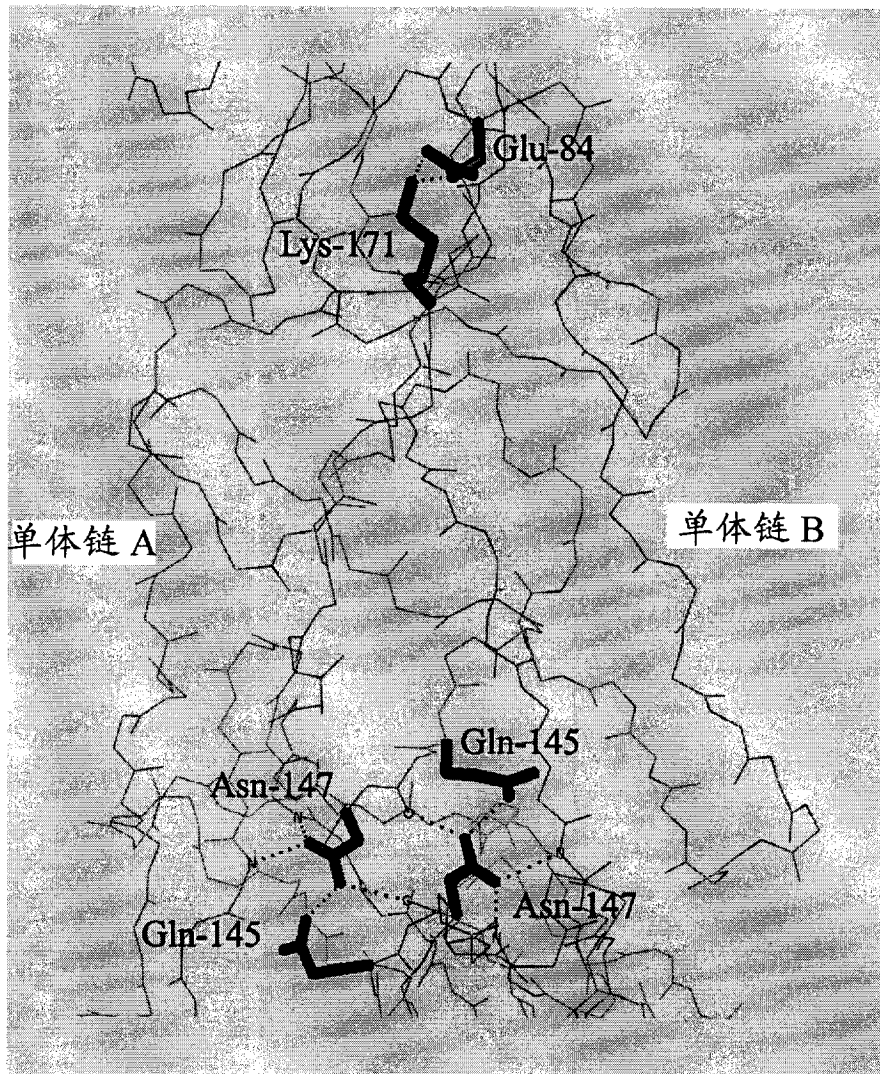


图 1D

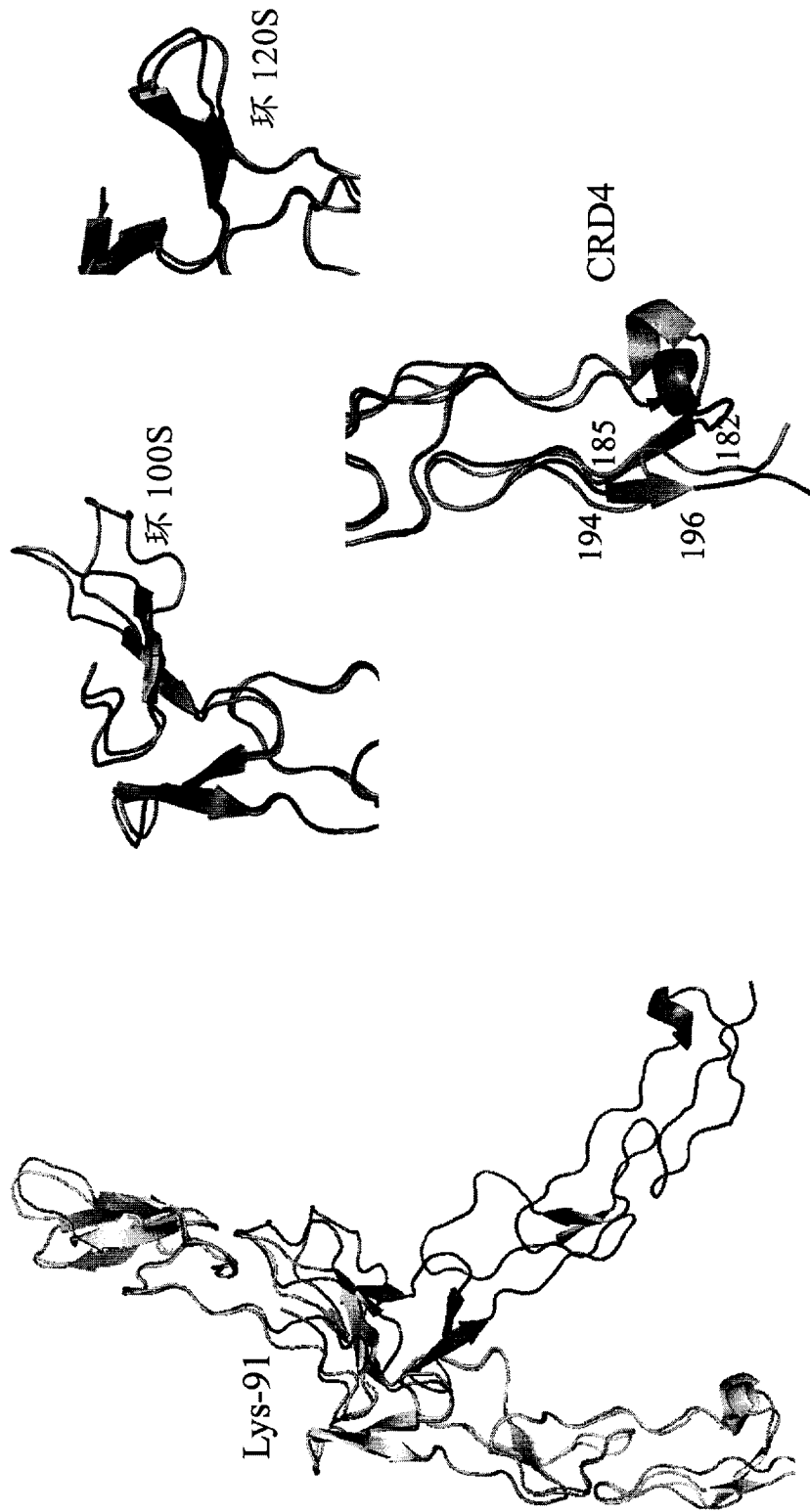


图 2

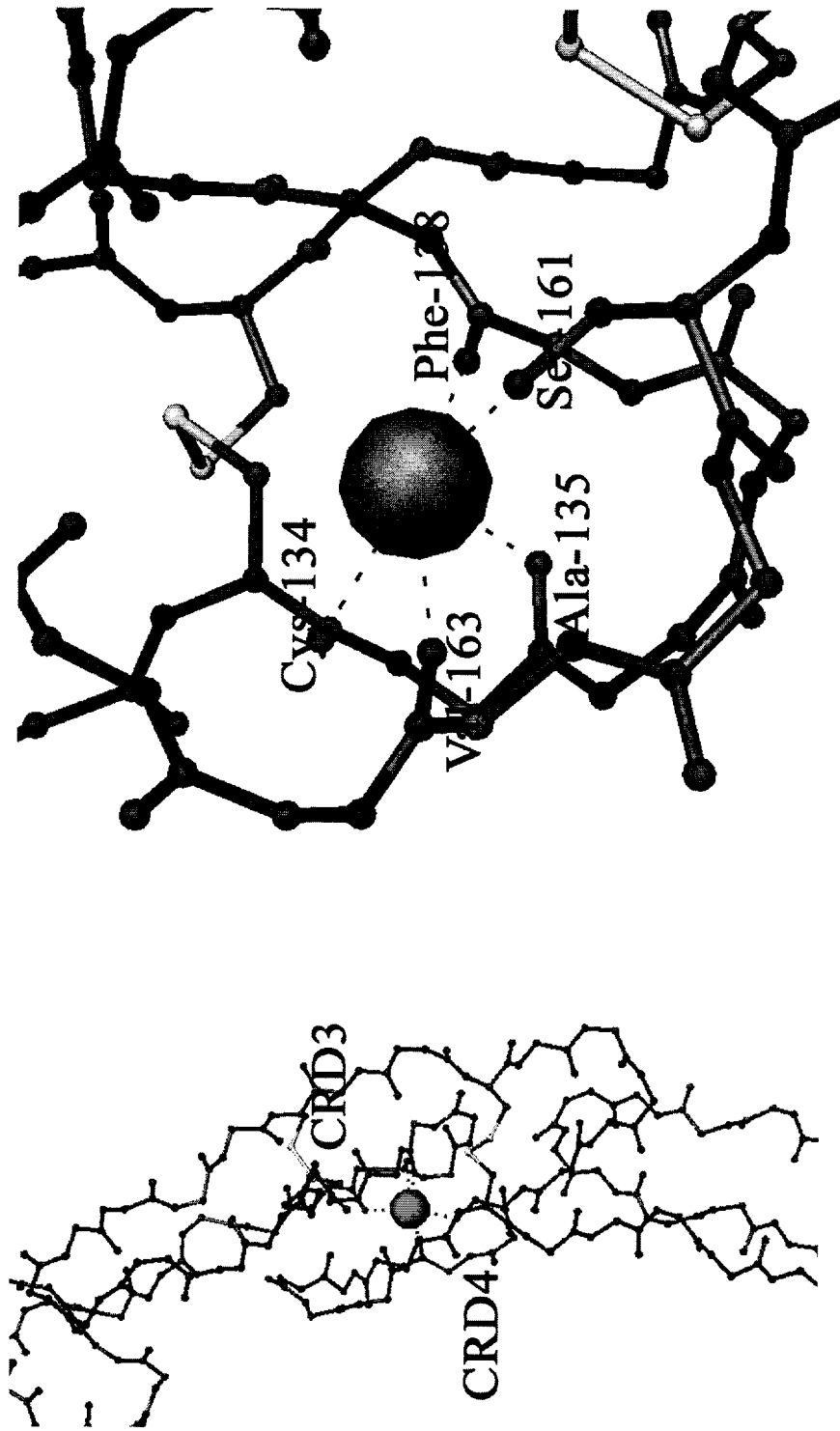


图 3

专利名称(译)	鼠RANK胞外蛋白晶体		
公开(公告)号	<a href="#">CN101845090A</a>	公开(公告)日	2010-09-29
申请号	CN200910080734.4	申请日	2009-03-26
[标]申请(专利权)人(译)	中国人民解放军总医院		
申请(专利权)人(译)	中国人民解放军总医院		
当前申请(专利权)人(译)	中国人民解放军总医院		
[标]发明人	唐佩福 刘长振 黄鹏 张世谦 高斌		
发明人	唐佩福 刘长振 黄鹏 张世谦 高斌		
IPC分类号	C07K14/47 C30B29/54 G01N33/68 G01N33/53		
代理人(译)	余刚 张英		
其他公开文献	CN101845090B		
外部链接	<a href="#">Espacenet</a> <a href="#">SIPO</a>		

摘要(译)

本发明提供了一种鼠RANK胞外蛋白晶体，以及该RANK胞外蛋白晶体在医药领域中的应用。

数据	鼠RANK胞外区域蛋白晶体
X-射线源	BM14, ESRF
波长 (Å)	0.9537
空间群	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>
晶胞 (a,b,c[Å])	39.82, 94.25, 102.42
分辨率 (Å)	30.0-2.00 (2.07-2.00)
独立反射	26472 (2582)
完整性 (%)	99.6 (9.2)
冗余值	6.9(6.7)
平均值 $I/\sigma I$	17.7(2.3)
Rmerge	0.096 (0.644)
精确统计	
分辨率 (Å)	30.0-2.00
反射数目 (运作、测试)	24896、1332
R-因子 (R <sub>work</sub> /R <sub>free</sub> )	0.216/0.265
原子数目 (蛋白/水/其他)	2433/221/47
键长均方根差 (Å)	0.007
键角均方根差 (°)	1.5
平均 B 因子[Å <sup>2</sup> ]	31/35/49