

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2014 年 6 月 26 日 (26.06.2014)



(10) 国际公布号
WO 2014/094444 A1

- (51) 国际专利分类号:
C40B 40/10 (2006.01) C40B 50/02 (2006.01)
C40B 30/02 (2006.01) C07K 7/06 (2006.01)
- (21) 国际申请号: PCT/CN2013/081761
- (22) 国际申请日: 2013 年 8 月 19 日 (19.08.2013)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201210561815.8 2012 年 12 月 20 日 (20.12.2012) CN
- (71) 申请人: 天津大学 (TIANJIN UNIVERSITY)
[CN/CN]; 中国天津市南开区卫津路 92 号天津大学,
Tianjin 300072 (CN)。
- (72) 发明人: 孙彦 (SUN, Yan); 中国天津市南开区卫津
路 92 号天津大学, Tianjin 300072 (CN)。 赵韦韦
(ZHAO, Weiwei); 中国天津市南开区卫津路 92 号天
津大学, Tianjin 300072 (CN)。 刘夫锋 (LIU,
Fufeng); 中国天津市南开区卫津路 92 号天津大学,
Tianjin 300072 (CN)。 史清洪 (SHI, Qinghong); 中
国天津市南开区卫津路 92 号天津大学, Tianjin
300072 (CN)。

- (74) 代理人: 北京三聚阳光知识产权代理有限公司
(SUNSHINE INTELLECTUAL PROPERTY INTER-
NATIONAL CO., LTD.); 中国北京市海淀区海淀南
路甲 21 号中关村知识产权大厦 A 座 5 层 503 室,
Beijing 100080 (CN)。
- (81) 指定国 (除另有指明, 要求每一种可提供的国家保
护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG,
BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR,
CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB,
GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP,
KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU,
LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ,
NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA,
RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST,
SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ,
VC, VN, ZA, ZM, ZW。
- (84) 指定国 (除另有指明, 要求每一种可提供的地区保
护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA,
RW, SD, SL, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ,
BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH,
CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE,
IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO,
RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM,
GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

[见续页]

(54) Title: AFFINITY LIGAND POLYPEPTIDE LIBRARY OF IMMUNOGLOBULIN G CONSTRUCTED BASED ON PRO-
TEIN A AFFINITY MODEL AND USE THEREOF

(54) 发明名称: 基于蛋白 A 亲和模型构建的免疫球蛋白 G 的亲和配基多肽库及其应用

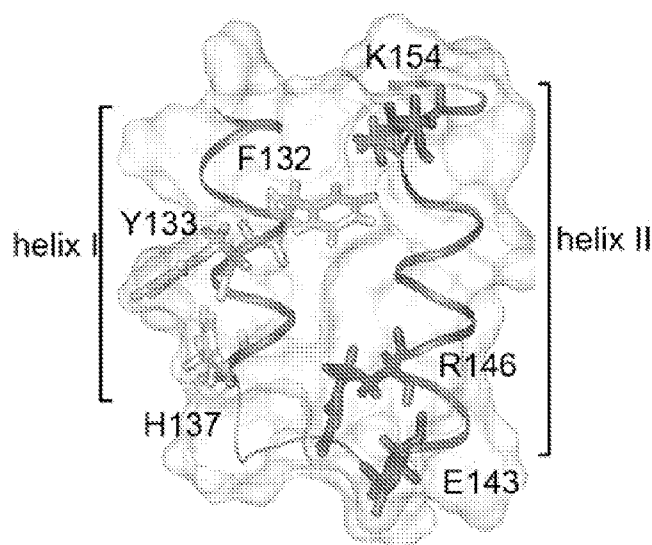


图 2 / FIG.2

(57) Abstract: Disclosed in the present invention are an affinity ligand polypeptide library of immunoglobulin G constructed based on a protein A affinity model and the use thereof. On the basis of a human IgG-protein A composite structure, a key residue of protein A having a relatively high affinity effect on a human IgG is obtained by parsing, and based on which an affinity polypeptide library of IgG is constructed. Further, a candidate polypeptide is gradually screened using an amino acid positioning method, molecular docking and molecular dynamics simulation means. A polypeptide affinity ligand capable of effectively separating and purifying IgG is determined through an affinity chromatography test method.

(57) 摘要: 本发明公开了一种基于蛋白 A 亲和模型构建的免疫球蛋白 G 的亲和配基多肽库及其应用。在人 IgG-蛋白 A 复合物结构基础上解析获得与人 IgG 具有较高亲和作用的蛋白 A 的关键残基, 在其基础上构建 IgG 的亲和多肽库。进一步利用氨基酸定位法、分子对接和分子动力学模拟手段逐步筛选候选多肽。通过亲和色谱试验方法确定能有效分离纯化 IgG 的多肽亲和配基。



根据细则 4.17 的声明:

— 发明人资格(细则 4.17(iv))

本国际公布:

— 包括国际检索报告(条约第 21 条(3))。

基于蛋白A亲和模型构建的免疫球蛋白G的亲和配基多肽库及其应用

技术领域

本发明涉及利用分子模拟仿生设计目标蛋白质的亲和配基技术，以及利用亲和色谱技术纯化目标蛋白质，属于生物技术中的计算机模拟和蛋白质分离纯化技术领域。

背景技术

抗体（免疫球蛋白，Ig）是位于动物血液和组织液中，由B淋巴细胞在对抗原的免疫应答中产生的一类糖蛋白。抗体与抗原具有的高亲和性使其广泛应用于生物学研究和临床治疗领域。特别是随着基因技术和克隆技术的逐渐成熟，单克隆抗体已成为治疗炎症、肿瘤和传染类疾病的有效药物。目前，已有大约20种单克隆抗体药物被FDA批准上市，至少300个还在研发之中，2006年单克隆抗体类药物的产值达到了206亿美元。抗体在医疗领域中日益显现的重要性亟需高效、稳定和低廉的生产工艺。由于抗体表达的复杂性以及对医用抗体的高质量要求，抗体纯化已成为整个生产过程的关键步骤。其中IgG是血清中主要的抗体成分，约占血清Ig的75%，也是需求量最大的一类抗体。

抗体纯化通常采用盐析、凝胶过滤色谱、疏水作用色谱、离子交换色谱和亲和色谱等分离手段进行多步纯化。其中亲和色谱因能对目标分子进行高效特异性纯化已成为抗体纯化后期最常用的色谱方法之一。亲和色谱是利用亲和配基能与目标分子特异且可逆结合的特性，从复杂的生物样品中分离纯化目标分子，具有选择性强，纯化效率高的优点。亲和色谱法的纯化效果取决于目标分子与配基之间的亲和性。因此，针对特定的目标分子，开发合适的亲和配基是构建一个亲和色谱体系所首要解决的问题。

金黄色葡萄球菌蛋白A(SpA)、蛋白G和蛋白L作为亲和配基已广泛用于制备高纯度抗体。这类配基的优点在于特异性高，而缺陷是过高的亲和力需要较苛刻的洗脱条件，容易导致目标蛋白变性和配基脱落，吸附容量较低；此外这类配基的制备也比较困难，价格昂贵，蛋白经固定化后一般会失去部分活性。这些弱点使得上述蛋白类配基的应用受到限制。

亲和肽配基的研究始于1986年Geysen对合成肽库的研究，他提出：含有关键残基的短肽能够模拟蛋白质上的决定簇。而且在多数情况下，几个关键残基与对应目标分子间的非共价作用构成了复合物结合的主要作用力。这两个观点奠定了亲和肽配基的理论基础。肽配基通常仅由很少的氨基酸组成，不会在产品使用时引起免疫中毒反应。而且其分子量小，即使从固定相上脱落、掺入产品中也很容易从终产物中除去。此外肽配基与蛋白质的作用条件温和，有利于控制分离条件，避免目标蛋白的变性。和蛋白类等具有高亲和力的配基相比，亲和肽配基也具有足够高的亲和力结合目标蛋白。肽配基的构象和理化性质更稳定，能承受分离操作中较强的酸碱洗脱和再生条件，可实现在GMP条件下大规模无菌生产。近些年来涌现出的一批肽配基都对抗体具有较好的分离纯化效果，如TG19318、Peptide H、A1P、A2P、8/7、线性肽配基（HWRGWV等）。

尽管多肽作为亲和配基具有如此多的优越性，但自然界中与目标蛋白有亲和力的多肽数量十分有限。上述所讨论的小分子配基虽然在抗体纯化研究中显现出了很大优势，但与SpA亲和介质相比，也存在着不足，如特异性和亲和力较差等。因此在亲和色谱的实际应用中，配基的筛选和设计至关重要。如何选择合适的多肽序列作为亲和配基，以及如何提高多肽的亲和力和选择性的问题，已成为影响多肽亲和色谱应用的关键。现有的筛选和设计方法主要分为实验筛选和理性设计两类方法。实验筛选是基于组合库技术进行高通量实验筛选，根据构建多肽文库方法的不同，亲和配基的筛选技术主要分为：组合化学合成肽库筛选，如上述提到的TG19318/D-TG19318、Peptide H、MAbSorbent A1P、A2P和8/7；噬菌体展示肽库筛选，如HWRGWV、HYFKFD和HFRRHL；核糖体展示肽库筛选这三类。理性设计主要是基于目标蛋白或已有配基的结构和性质设计新配基。随着计算机技术、计算化学和药物化学的发展，

亲和配基的设计已进入以计算机辅助设计为主导的理性设计阶段。计算机辅助配基设计的各种虚拟设计方法包括分子对接, 3D-QSAR, 药效团模型, 分子动力学(molecular dynamics, MD)模拟和从头设计等。

分子对接是两个分子之间通过几何匹配和能量匹配而相互识别的过程。分子对接计算是把配基分子放在目标蛋白结合位点的位置, 然后按照几何互补、能量互补和化学环境互补的原则来实时评价配基与目标蛋白结合的好坏, 并找到两个分子之间最佳的结合构象。由于分子对接考虑了目标蛋白与配基相互作用的信息, 因此从原理上讲, 分子对接是一种基于受体的直接设计方法。近年来随着蛋白晶体结构信息的快速增长以及小分子数据库的不断更新, 分子对接已经成为基于结构设计中的最为重要的方法。常用软件有 DOCK, Autodock 和 FlexX 等。

分子动力学是建立在牛顿力学基础上的一种分子模拟方法, 用于研究多粒子体系中各粒子的运动过程。MD 模拟的基本步骤可以分为如下四步: (1) 初始化; (2) 计算原子受力; (3) 更新原子坐标和速度。根据上一步的原子坐标、速度和受力, 即可得到原子在下一时刻的坐标和速度。不断循环进行 (2) 和 (3) 步得到体系状态随模拟时间的变化情况; (4) 分析轨迹。常用的分子动力学模拟软件主要有 GROMACS、NAMD、AMBER 和 CHARMM 等。通过分析 MD 模拟轨迹可以获得模拟体系的各种性质, 包括构象、能量、动力学性质以及配基-目标蛋白质之间的相互作用力等。

采用多种理性设计方法, 通过设计合理的组合策略可以实现降低成本和更高精度的配基设计。在前期采用一些计算速度快但精度有限的方法以富集可能的候选分子, 如分子对接法。随后再采用 MD 模拟等计算量大但更精确的方法进一步挑选最佳的配基分子。在后期阶段则采用比较耗时和高成本的实验方法做最后验证。

发明内容

本发明的目的在于提出一种基于蛋白 A 亲和模型构建免疫球蛋白 G 的新型亲和配基多肽库及设计方法的应用。本发明所述的抗体亲和短肽配基的仿生设计方法是首次建立的, 并经验证是有效的。

本发明基于蛋白 A 亲和模型构建免疫球蛋白 G 的新型亲和配基多肽库, 其依据为 6 个 SpA 关键残基: F132、Y133、H137、E143、R146 和 K154。为了配基的固定, 在多肽序列中间位置添加一个半胱氨酸。该肽库共有以下 8 种序列:

FYCHXXXE、FYXHCXXE、FYCHXXR、FYXHCXR、FYXCRXE、YFXCRXE、HXYFCXR 和 HXYFCXK;

其中 X 代表除去半胱氨酸以外的 19 种氨基酸。

本发明的新型的亲和多肽设计方法的应用, 是在上述肽库的基础上, 进一步利用氨基酸定位方法, 确定 X 所代表的氨基酸种类。

X 所代表的氨基酸种类如下表所示:

多肽构建模式& 条数	X 代表的氨基酸种类
FYCHXXXE 990	X_1: A, R, N, D, Q, E, H, I, L, K, M, T, W, Y, V (15) X_2: A, Q, E, I, L, K (6) X_3: A, R, D, Q, E, L, K, M, P, S, T (11)
FYXHCXXE 594	X_1: A, R, N, D, I, P, T, W, Y (9) X_2: A, Q, E, I, L, K (6) X_3: A, R, D, Q, E, L, K, M, P, S, T (11)
FYCHXXR 112	X_1: N, D, Q, E, H, I, K, M, F, P, T, W, Y, V (14) X_2: Q, E, H, I, K, M, F, W (8)
FYXHCXR	X_1: A, R, N, D, I, P, T, W, Y (9)

72	X_2: Q, E, H, I, K, M, F, W (8)
FYXCRXE	X_1: A, R, N, D, Q, E, G, H, I, L, K, M, F, P, W, Y, V (17)
153	X_2: N, D, Q, E, L, M, S, T, W (9)
YFXCRXE	X_1: N, D, E, H, I, L, K, M, P, S, T, Y, V (13)
117	X_2: N, D, Q, E, L, M, S, T, W (9)
HXYFCXR	X_1: A, R, N, D, I, P, T, W, Y (9)
54	X_2: Q, H, K, S, W, Y (6)
HXYFCXK	X_1: A, R, N, D, I, P, T, W, Y (9)
81	X_2: A, R, Q, H, K, F, W, Y, V (9)

具体的短肽序列如表 5 所示。

将多肽库进行分子对接筛选、均方根偏差比较以及分子动力学模拟复筛，获得与 hIgG 具有较高亲和性的多肽配基：FYWHCLDE、FYFCRWE、FYIHCLPE、FYHCKKE、FYCHWALE、FYCHWQDE、FYCHTIDE、FYRHCQRE、FYCHHKTE、FYCHLQKE、FYCHRKAE、FYCHNQDE、FYCHRQEE 和 FYNHCASE。

需要强调的是，上述仿生肽库中包含的所有多肽分子（2173 条），理论上都有可能是 hIgG 的亲肽配基；而进行分子对接筛选以及分子动力学模拟复筛，是为了富集与 hIgG 有较高亲和性的多肽分子，有效缩小候选多肽数量便于进行后续实验验证；经过依次筛选最后得到的 14 个多肽分子，是与 hIgG 具有较高亲和性的概率最大的多肽分子。虽然计算机模拟技术发展至今不断趋于成熟和完善，但对生物分子间的相互作用的预测还不能达到与实际情况完全一致的水平，因此分子模拟软件就不可避免地具有各种局限性（如利用不同的分子模拟软件或采用不同的参数就可能会产生不同的结果），只能做到与实际情况尽可能地趋于接近。所以应用分子模拟软件进行筛选时，会出现两种情况：一是筛选得到的多肽分子是 hIgG 的亲肽配基的命中率随筛选次数不断增大；二是漏筛的、排除掉的可作为 hIgG 的亲肽配基的多肽分子也会随筛选次数增多。所以，不排除多肽库中其余的 2161 条多肽分子也可能是 hIgG 的多肽亲和配基。

本发明利用分子模拟和实验手段获得 IgG 亲和短肽配基的方法如下 1-5 步：

1. 应用分子力学/泊松-波尔兹曼溶剂可及表面积(molecular mechanics-Poisson-Boltzmann surface area, MM/PBSA)自由能计算和自由能分解方法计算得到蛋白 A 与 hIgG1 结合（图 1）的关键残基，得到蛋白 A 简化的亲和结合模型（图 2）：首先使用 MM/PBSA 方法计算了 SpA 和 hIgG1 复合物的绝对结合自由能。然后，采用一种基于 MM/PBSA 的自由能分解方法解析 SpA 和 hIgG1 的高亲和性分子机理并分析 SpA-hIgG1 复合物的结合表面的残基对结合自由能的贡献。根据每个残基的自由能贡献和残基作用对的分析确定 SpA 和 hIgG1 相互作用的热点残基。

2. 根据亲和结合模型中 6 个离散的热点残基的构象和相对位置，应用 Autodock vina（简称 vina）分子对接软件计算出多肽序列的长度和需要插入的氨基酸残基种类；在多肽序列中间位置加入一个半胱氨酸（Cys），便于方便地、有选择性地多肽固定到色谱介质—Thiopropyl（硫丙基）Sephacrose 6B 上，从而使多肽两端的热点残基有充分的自由度，使其能与 hIgG 发挥关键的亲和作用，不至于在多肽固定到介质上后而减弱其与 hIgG 的亲和作用。最后得到包含一系列七肽和八肽序列的多肽库。

具体操作是从 hIgG1-SpA 复合物的晶体结构(PDB ID: 1FC2)中分别选取 Fc 片段和 SpA 的 B 结构域的三维坐标文件；从 SpA B 结构域的坐标文件中再单独取出 6 个热点残基(F132, Y133, H137, E143, R146 和 K154)的坐标结构，计算两两热点残基之间相应 C 和 N 端之间的距离，根据插入氨基酸数目的原则，计算热点残基之间应插入的氨基酸残基数，确定多肽序列模式，见表 1；然后将多肽序列中间位置的氨基酸确定为 Cys，见表 2；之后选取两两热点残基之间对应的 Fc 片段，使其包含在 vina 对接软件的格栅盒子之中。选择好盒子之后，将 19 种氨基酸（除去 Cys）依次与盒子包含的 Fc 片段进行对接，选择打分最高的前 20 个构象。按照氨基酸选取的基本原则进行筛选：1) 构象合适，即氨基酸残基

对接构象的 C 和 N 末端必须与相应热点氨基酸残基的 N 和 C 端首尾相接形成肽键；2) 亲和结合自由能 < -2.0 kcal/mol。结果：得到多肽中未知氨基酸残基 (X) 的种类，见表 3。利用 perl 脚本调用 charmm 构建多肽库，其共包含 2173 条序列，见表 5，每条序列都包含 4 个热点残基。

3. 利用 vina 和 Rosetta Flexpepdock (简称 Flexpepdock) 分子对接软件将多肽库中的多肽分子依次与 hIgG1 的 Fc 片段进行对接，然后根据打分排名获得与 Fc 片段具有较高结合能的多肽分子。这里需要说明的是，1) 打分的分数标准是根据打分分布人为设定的，是根据经验值选择的最有可能富集有效分子的数值；2) 利用分子对接软件进行筛选是为了富集与 hIgG 具有亲和性的多肽分子的数目，剔除可能性较低的分子，从而节省后续验证实验的操作成本；而不是将未通过筛选标准的多肽分子作为非 hIgG 亲和肽配基摒弃掉，这些多肽分子只是不符合本次研究设定的标准。相关技术人员可根据实际需要，适当地提高或降低这一分数标准，虽然会改变筛选得到的分子数目（提高或降低筛选标准会使进入下一轮筛选的候选分子数目相应地减少或增加，后者会减少漏筛的可能性，但另一方面也会增加整个研究过程的复杂性。最佳情况是能找到一个平衡点，使得既不会漏筛掉过多的潜在分子，也不会使最后得到的多肽库中无效或效果弱的分子较多），但这与本发明目的一致，也符合本发明精神。vina 对接盒子恰好包含 Fc 和 SpA 的结合区域—CH2 和 CH3 之间的“一致性结合位点 (consensus-binding site, CBC)”，固定好盒子后，依次将多肽库里的多肽分子与 vina 盒子包含的 Fc 片段进行对接。选取结合自由能（打分分数）低于 -6.5 kcal/mol 的多肽分子，共计 754 个多肽。在这里值得注意的是，配基筛选采用的是打分分数低于 -6.5 kcal/mol 的标准，选取这一标准的理由会在下文中有详细说明。利用 GROMACS 分子模拟软件自带的 g_rms 程序计算 vina 对接得到的 754 条多肽序列与 SpA 中相应的热点残基之间的均方根偏差 (rmsd)，据此比较 vina 对接后多肽中的热点残基与 SpA 中的热点残基构象之间的差异。选择 Encad 全原子力场，对除去氢原子之外的热点残基计算 rmsd 值。rmsd 值越小，表示多肽包含的热点残基的对接构象与 SpA 中热点残基构象越接近。结果表明，rmsd 值主要分布在 $0.2\sim 0.6$ nm 之间，选择 $\text{rmsd} < 0.4$ nm 的 150 个多肽进行下一步的研究。在这里值得注意的是，筛选肽配基的 rmsd 值标准也是可根据实际情况进行适当更改的，相关技术人员在相似的研究背景和情况下，可参照本发明设置合适的 rmsd 值标准使得筛选得到的多肽分子数目适中又不至于漏筛掉过多的潜在分子。将 vina 对接得到的多肽-Fc 复合物构象作为初始构象，利用 Flexpepdock 对上一轮筛选得到 150 个多肽与 Fc 进行逐个对接。Flexpepdock 对接参数为 flexpepdock.flags:

```
-pep_refine
-use_input_sc
-ex1
-ex2aro
-ignore_unrecognized_res
-nstruct 1
-out:suffix pepdock
-lowres_preoptimize
-scorefile flexpepdock.sc
```

进行两次平行分子对接，以减少只进行一次对接的结果的随机性。结果发现只有少数几个多肽不能与 Fc 结合，大部分多肽序列都能结合到 Fc 上。对接打分分数 (结合界面能量分数, I_{sc}) 分布在 $-4\sim -22$ 之间，其中 I_{sc} 在 $-14\sim -16$ 之间的分布最多，选取 $I_{sc} \leq -16$ 为筛选标准，这样既可使候选多肽数目不太多从而避免计算资源浪费（考虑到下一轮分子动力学模拟复筛比较消耗计算机资源，速度慢），又不至于漏筛掉过多的潜在分子。当然相关技术人员可根据自己研究的实际情况更改为更适合自己的 I_{sc} 筛选标准。共有 15 个多肽分子与 hIgG1 的 Fc 片段之间的 I_{sc} 在两次平行分子对接中均小于 -16 ，挑选这 15 条多肽分子进行下一轮筛选。

4. 将步骤 3 中筛选得到的 15 条多肽分子与 hIgG1 的 Fc 片段进行 MD 模拟，再次筛选多肽配基，得到在模拟时间尺度内能与 Fc 片段稳定结合的多肽分子，即这些多肽分子有可能与 hIgG 具有较高的亲和作用。将步骤 3 中利用 Flexpepdock 对接获得的多肽-hIgG1 复合物的构象作为初始构象，利用

GROMACS 4.5.3 软件包, 选择 G53a6 力场, 将 15 个候选多肽的 pdb 坐标结构利用 pdb2gmx 命令转化为 GROMACS 专用的 gro 结构; 利用 editconf 命令将多肽-蛋白复合物置于矩形水盒子中心, 使它们距离盒子边缘最小为 0.9 nm; 然后用 genbox 命令来向模拟盒子中添加水分子, 水分子采用 SPC216 水模型; 然后用 grompp 命令将 mdp 文件中标明的参数整合为结构和拓扑文件从而生成 tpr 文件, 以及应用 genion 命令加入平衡系统净电荷所需的离子种类和数量; 之后进行能量最小化, 去除体系中的原子间的碰撞和不正确的几何构型; 接下来用 mdrun 命令依次进行 100 ps 的正则 (NVT) 系综和等温等压 (NPT) 系综下的限制动力学平衡, 最后仍然利用 mdrun 命令进行 20 ns 的无限制 MD 模拟。相关技术人员可根据自己研究的实际情况更改相应参数, 使其更适合特定的、不同的研究情况。

5. 合成筛选得到的多肽分子, 将其固定到 Thiopropyl Sepharose 6B 色谱介质上制成亲和介质并装柱, 进行亲和色谱实验验证。研究脉冲进样 hIgG 溶液以及对人血清中 hIgG 的分离和电泳情况。

本方法首先利用 MM/PBSA 自由能计算和自由能分解方法获得与人 IgG 具有高亲和性的蛋白 A 的热点残基, 并以此为基础构建蛋白 A 与 hIgG 结合的亲和结合模型。然后利用氨基酸定位、分子对接、均方根偏差比较以及在多肽序列中间位置添加一个半胱氨酸的方法, 构建多肽库, 之后将多肽库进行分子对接筛选以及分子动力学模拟复筛, 获得与 hIgG 具有较高亲和性的多肽配基, 已经确定的能有效分离纯化 hIgG 的多肽分子有两个, 分别是 FYWHCLDE 和 FYCHWALE。并应用亲和色谱、聚丙烯酰胺凝胶电泳 (SDS-PAGE) 和分光光度法等一系列实验手段进行了 IgG 的分离纯化及其表征的研究。

值得提出的是, 多肽库中的全部 2173 条多肽分子理论上都有可能成为 hIgG 的亲和配基, 但由于分子模拟自身的局限性, 不可能完全模拟实际情况, 而只能是与实际情况尽可能地趋于接近, 所以最终需要实验进行实际情况的验证。由于时间人力物力的限制, 本发明不可能将全部多肽依次进行实验验证。本发明只在最终筛选得到的 14 条多肽中随机挑选两个进行实验表征, 主要目的在于阐述如何用实验方法进行多肽作为 hIgG 亲和配基有效性的验证, 而不是指只有本发明提到的两个多肽有效。相关技术人员可根据本发明所述的实验方法对多肽库中其它多肽进行验证。

附图说明

图 1 为蛋白 A 与 hIgG1 Fc 片段复合物的三维结构; helix I 和 helix II 是蛋白 A 的 B 结构域的两个螺旋结构, 它们与 Fc 片段的 CH2 和 CH3 铰链区结合, 结合位点称为“一致性结合位点(CBC)”

图 2 为蛋白 A 的简化亲和结合模型; helix I 包含三个热点残基 (F132、Y133 和 H137); helix II 包含三个热点残基 (E143、R146 和 K154)

图 3 是氨基酸插入图示; 虚线框内表示插入的氨基酸残基

图 4 为多肽分子与 Fc 的 vina 对接打分数值分布; 结合自由能数值越负表示配基与 Fc 之间的亲和性越高

图 5 为 SpA 的热点残基与多肽对接构象的相对应的热点残基之间的均方根偏差数值分布, 数值越小, 表示两者的构象越接近

图 6 为肽分子与 Fc 片段的第一次 Flexpepdock 对接打分数值分布; 结合界面能量分数数值越负, 表示结合越牢固

图 7 是肽配基亲和柱分离纯化人血清样品中的 IgG。平衡缓冲液: 20 mM 磷酸钠盐缓冲液, pH 6.0; 再生缓冲液: 0.1 M Gly-HCl 缓冲液, pH 2.4。A 图表示偶联有肽配基 FYWHCLDE 的亲和柱对人血清样品中 IgG 的分离, 洗脱缓冲液: 50 mM 柠檬酸钠盐缓冲液, pH 3.0; B 图表示偶联有肽配基 FYCHWALE 的亲和柱对人血清样品中 IgG 的分离, 洗脱缓冲液: 平衡缓冲液, 包含 0.2 M NaCl, pH 6.0

图 8 亲和色谱洗脱结果的 SDS-PAGE 分析, 从左至右电泳条带依次为标准蛋白 Marker (30-120 kDa), 人血清样品 (feedstock), 流出峰 (F-T) 和洗脱峰 (Elution) 组分。A 图为 FYWHCLDE 的电泳图, B 图为 FYCHWALE 的电泳图

具体实施方式

下面结合附图对本发明作进一步详细地描述, 该实施例是用于解释、而不以任何方式限制本发明。

实施例 1 获得 SpA 的简化亲和结合模型

首先使用 MM/PBSA 方法计算金黄色葡萄球菌蛋白 SpA 和 hIgG1 复合物的绝对结合自由能。然后, 采用一种基于 MM/PBSA 的自由能分解方法解析 SpA 和 hIgG1 高亲和性的分子机理并分析 SpA-hIgG1 复合物的结合表面的残基对结合自由能的贡献。根据每个残基的自由能贡献和对残基作用对的分析确定 SpA 和 hIgG1 相互作用的热点残基。最后, 基于以上解析获得的分子机理和热点残基构建了 SpA 的亲和结合模型。

SpA-hIgG1 复合物分子体系被用于分子动力学模拟。每个复合物模型包含一条 Fc 片段的单链和 SpA 的 B 结构域。SpA-hIgG1 的模型结构取自蛋白质数据库(PDB ID: 1FC2) (图 1)。分子动力学模拟使用 CHARMM 软件以及 CHARMM27 全原子力场完成。首先采用 TIP3P 水分子模型将 SpA-hIgG1 复合物溶解于长方体盒子中(100×80×60 Å)。添加 Na⁺或 Cl⁻作为反离子以中和体系电荷。体系经能量最小化后, 采用 NPT 系综平衡 200 ps, 最后采用 NVT 系综继续模拟 15 ns。体系的温度设为 298 K, 并采用 Nose-Hoover 法维持恒温。所有的模拟都应用周期性边界条件。计算非键作用时采用 12 Å 的距离截断, 应用 Particle Mesh Ewald (PME)计算长程静电作用。SHAKE 用于限制所有氢原子, 时间步长设为 2 fs。由于该体系只包含 Fc 片段的一条单链, 故在 Fc 片段的两个终端残基上施加一个距离约束以限制 CH2 和 CH3 域的弯曲。最后从平衡阶段的最后 3 ns 轨迹中以 40 ps 为间隔均匀取样, 共得到 75 帧构象供后续分析。

采用 MM/PBSA 方法计算 SpA-hIgG1 复合物的结合自由能(ΔG_{bind})。 ΔG_{bind} 为气相作用能(ΔG_{gas})、溶剂化作用能(ΔG_{sol})和熵作用项(-TS)的总和。

$$\Delta G_{bind} = \langle \Delta G_{gas} \rangle + \langle \Delta G_{sol} \rangle - \langle T\Delta S \rangle$$

括号<...>表示所有模拟得到的能量项的平均值。T 代表绝对温度, S 是溶质熵。

G_{gas} 包含分子间静电作用项(Gelec), 范德华作用 (vdw) 项(G_{vdw})和内能项(G_{inter})。

由于本研究采用“相同轨道方法”进行取样分析, 因此内能项 (G_{inter})为 0。故 ΔG_{gas} 对结合自由能的贡献是 ΔG_{elec} 和 ΔG_{vdw} 的总和。

$$\Delta G_{gas} = \Delta G_{elec} + \Delta G_{vdw}$$

溶剂化作用能包括以下两部分: 静电溶剂化能(G_{PB})和非极性溶剂化能(G_{np})。

$$G_{sol} = G_{PB} + G_{np}$$

G_{PB} 是使用 CHARMM 程序的 PBEQ 模块求解线性泊松-玻尔兹曼方程(PB)得到的。在所有的 PB 计算中, 溶质和溶剂介电常数分别设为 1 和 80。离子强度为 0, 溶剂分子半径设为 1.4 Å。 G_{np} 是溶剂空穴作用项和溶质-溶剂范德华作用项之和, $G_{np} = \gamma \times SASA + b$, 常数 γ 和 b 分别设为 0.00542 kcal/(mol.Å²)和 0.92 kcal/mol。SASA 代表溶剂可及表面积。

溶质熵(S)可分解为平动熵(S_{trans})、转动熵(S_{rot})和振动熵(S_{vib})。这三项由统计力学计算得到。 S_{trans} 和 S_{rot} 分别是分子的质量和转动惯量的函数。振动熵(S_{vib})是使用 CHARMM 程序的 VIBRAN 模块做标准模式分析(normal mode analysis, NMA)得到。

每个残基的自由能贡献分解为极性(G_{polar})和非极性作用($G_{nonpolar}$)。

其中每一项继续分解为两个能量项之和。在接下来的分析中, G_{polar} 被认为是残基的静电作用贡献, $G_{nonpolar}$ 可认为是疏水作用贡献。需要注意的是, $G_{residue}$ 只是分解为 G_{polar} 和 $G_{nonpolar}$ (即不包含熵)。 $G_{residue} = G_{polar} + G_{nonpolar}$, $G_{polar} = G_{elec} + G_{PB}$, $G_{nonpolar} = G_{vdw} + G_{np}$ 。每个残基的静电作用能(G_{polar})等于分子间静电作用能(G_{elec})和溶剂化静电作用能(G_{PB})之和。线性 PB 方程允许将静电溶剂化能分解为每个原子的贡献。SpA 上每个残基的范德华作用能贡献等于该残基与 Fc 片段之间范德华作用能的一半, 对于 Fc 片段上每个残基的范德华作用能反之亦然。每个残基的非极性溶剂化作用能与该残基的溶剂可及表面积成比例。通过以上分析即可计算得到 SpA-hIgG1 复合物的结合自由能(ΔG_{bind})以及每个残基的自由能贡献, 进而得到 SpA 与 hIgG1 相互作用的热点残基。

在 SpA-hIgG1 复合物中, 与 Fc 片段接触的 SpA 残基离散地分布于螺旋 I(K126 至 H137)、螺旋 II(E144 至 D155)和不规则卷曲(L138 至 E143)上。在本研究中, 热点残基被定义为对结合自由能贡献较大的残基以及参与形成重要的分子间作用以补偿不利的溶剂化作用的残基。采用 ± 2.5 kcal/mol 的标准来识

别对自由能贡献较大的残基。为了识别复合物中的重要残基来指导亲和配基的理性设计，本研究所采用的标准高于其他文献中常用的 ± 2.0 kcal/mol 的标准。SpA 中只有 F132, Y133, H137, R146 和 K154 这 5 个残基对结合自由能贡献较大。SpA 的 E143 残基虽然只有较小的能量贡献(-1.2 kcal/mol)，但其自由能偏差却高达 ± 6.7 kcal/mol。此外，E143 的侧链带有一个带负电荷的羧基，主要提供静电作用。因此，E143 的自由能贡献对残基的构象十分敏感。SpA 的 E143 与 hIgG1 的 K317 之间有很强的分子间静电吸引作用(-24.0 kcal/mol)。相反，它与 hIgG1 的 D280 和 D315 残基有较弱的静电排斥作用。因此，虽然 SpA 的 E143 没有直接贡献较多的结合自由能，但却为 hIgG1 的 K317 创造了一个有利的局部结合环境。故 E143 也被认为是 SpA 的热点残基之一。

螺旋 I 与 Fc 片段通过疏水作用结合。特别是螺旋 I 的热点残基 F132 和 Y133 提供了大部分的疏水作用。MD 模拟结果表明 F132 和 Y133 的芳香侧链紧密结合在 Fc 片段疏水性口袋的浅槽。相反，螺旋 II 通过静电作用与 Fc 片段结合。螺旋 II 包含 3 个极性热点残基(E143, R146 和 K154)，与分布在结合位点的疏水区域周围的极性残基(H310, Q311, D315, K317 和 K338)有很强的静电吸引作用。因此，为开发新的拟蛋白 A 亲和配基，在理性设计过程中需要综合考虑疏水作用和特殊的静电作用。根据 SpA 与 hIgG1 的亲和机理和 SpA 的热点残基分布，构造了一个 SpA 的简化的结合模型，6 个 SpA 关键残基为 F132, Y133, H137, E143, R146 和 K154，见图 2。此模型可以作为 IgG 的拟蛋白 A 亲和配基理性设计的起始点。

实施例 2 多肽库的构建

1. 确定多肽序列的长度

已知：

肽键的长度 $\approx 1.33\text{\AA}$ ；

氨基酸主链长度 $\approx 2.78\text{\AA}$ ；

插入一个氨基酸残基需 2 个肽键长度和一个氨基酸的主链长度 $\approx 2 \times 1.33 + 2.78 = 5.44\text{\AA}$ ；

插入两个氨基酸残基需 3 个肽键长度和两个氨基酸主链长度 $\approx 3 \times 1.33 + 2 \times 2.78 = 9.55\text{\AA}$ ；

图 3 为插入一个氨基酸的示例（虚线框内线条代表增添的键）。若两个热点残基之间的 C 和 N 端之间的距离（insertion distance） $< 5.44\text{\AA}$ 时，就可以插入一个氨基酸残基；若 $5.44\text{\AA} \leq \text{insertion distance} < 9.55\text{\AA}$ ，则考虑插入两个氨基酸残基；若 $\text{insertion distance} > 9.55\text{\AA}$ ，则需要插入三个或以上的氨基酸残基。

根据 SpA 两个热点残基相应的 N 和 C 端之间的距离（利用 Visual Molecular Dynamics, VMD 软件计算得到），确定需插入的氨基酸个数。最终确定的多肽构建模式如表 1 所示。

表 2

表 1		表 2	
序号	多肽构建模式	序号	多肽构建模式
1	FYXHXXXE	1	FYCHXXXE
2	FYXHXXR	2	FYXHCXXE
3	FYXXRXE	3	FYCHXXR
4	YFXXRXE	4	FYXHCXR
5	HXYFXXR	5	FYXCRXE
6	HXYFXXK	6	YFXCRXE
		7	HXYFCXR
		8	HXYFCXK

其中，一种模式是八肽，其余都是七肽。考虑到多肽配基的 N 和 C 两末端都由热点残基组成，它们和 IgG 均具有较高的亲和作用。为使配基在固定化后仍然保持其与抗体之间的高亲和性，考虑在配基的中间位置插入 Cys，这样就可将多肽通过二硫键偶联在 Thiopropyl Sepharose 6B 介质上。因此，在多肽序列的中间位置插入一个 Cys 以便于多肽的固定化。这样就可保证多肽固定于介质上后，其两端

的关键残基具有充分的自由度，能够充分发挥两端关键残基的亲合作用，从而保持其与 IgG Fc 片段之间的亲合作用力。据此，根据 SpA 热点残基以及残基之间需插入的残基个数，我们最终得到八种多肽的构建模式（见表 2）。

2. 确定热点残基空隙处插入的氨基酸残基种类

利用片段定位法（fragment location method），片段定位法是利用分子模拟方法来确定特定原子或片段在结合腔中的最佳位置的一种模拟方法。首先确定结合腔中不同的作用区域，如：静电区域、疏水区域、氢键供体区和氢键受体区。再根据化学环境匹配的原则，在结合腔内放置与之化学特征匹配的配体分子。例如，受体疏水区附近放置的配体分子片段也是疏水性基团，如苯环、脂肪疏水链等；受体上的正/负电区域应该与配体的负/正电区域匹配。首先选择与蛋白 A 的六个热点残基之间的空缺位置所对应的 Fc 片段区域，利用分子对接软件 vina 将 19 种氨基酸（除去 Cys，便于之后的多肽配基固定化）与 Fc 片段区域依次进行对接，得到与该空缺位置具有较高亲和性的氨基酸种类，再将这些氨基酸残基与热点残基按顺序连接起来得到候选多肽配基。

利用氨基酸定位法确定热点残基之间应插入氨基酸的种类，以此为基础构建多肽库。

选取氨基酸的基本原则如下：1) 构象合适，即氨基酸残基对接构象的 C 和 N 末端必须与其它氨基酸残基的相应 N、C 端首尾相接；2) 亲和结合自由能 < -2.0 kcal/mol。表 3 所示为确定的多肽中未知氨基酸残基（X）的种类。

表 3

多肽构建模式& 条数	X 代表的氨基酸种类
FYCHXXXE 990	X_1: A, R, N, D, Q, E, H, I, L, K, M, T, W, Y, V (15) X_2: A, Q, E, I, L, K (6) X_3: A, R, D, Q, E, L, K, M, P, S, T (11)
FYXHCXXE 594	X_1: A, R, N, D, I, P, T, W, Y (9) X_2: A, Q, E, I, L, K (6) X_3: A, R, D, Q, E, L, K, M, P, S, T (11)
FYCHXXR 112	X_1: N, D, Q, E, H, I, K, M, F, P, T, W, Y, V (14) X_2: Q, E, H, I, K, M, F, W (8)
FYXHCXR 72	X_1: A, R, N, D, I, P, T, W, Y (9) X_2: Q, E, H, I, K, M, F, W (8)
FYXCRXE 153	X_1: A, R, N, D, Q, E, G, H, I, L, K, M, F, P, W, Y, V (17) X_2: N, D, Q, E, L, M, S, T, W (9)
YFXCRXE 117	X_1: N, D, E, H, I, L, K, M, P, S, T, Y, V (13) X_2: N, D, Q, E, L, M, S, T, W (9)
HXYFCXR 54	X_1: A, R, N, D, I, P, T, W, Y (9) X_2: Q, H, K, S, W, Y (6)
HXYFCXK 81	X_1: A, R, N, D, I, P, T, W, Y (9) X_2: A, R, Q, H, K, F, W, Y, V (9)

利用 perl 脚本调用 Charmm 软件得到多肽库，总共包含 2173 条序列，每条序列都包含 4 个热点残基。

实施例 3 多肽分子与 Fc 片段的对接

1. 多肽与 Fc 片段的 vina 对接

由于单个残基与 Fc 有较强的亲合作用不一定代表由残基组成的多肽与 Fc 有较强的相互作用，因此继续用 vina 将多肽库中的全部多肽与 Fc 片段的 CBC 依次对接。对接结果显示所有多肽都能与 Fc 片段结合，且所有多肽的预测结合自由能在 -4.5~-8.2 kcal/mol 之间，此范围符合亲和配基的适中亲和力要求。

(结合常数在 $10^4 \sim 10^8 \text{ M}^{-1}$ 之间)。其中结合自由能在 -6.5 kcal/mol 左右的多肽分布最多。筛选时为避免漏选, 选取结合自由能低于 -6.5 kcal/mol 的多肽分子, 共计 754 个多肽。对接打分数值的分布见图 4。

为了能够充分发挥热点残基与 IgG 之间的亲和作用, 要尽量寻找多肽中的热点残基的对接构象与 SpA 中相应的热点残基构象一致的多肽序列, 这样的多肽序列能够仿效 SpA 与 Fc 片段形成高效的亲和结合作用。为了分析短肽配基中热点残基与 SpA 中相应残基之间构象的差别, 本研究利用 GROMACS 分子模拟软件自带的 `g_rms` 程序计算 vina 对接得到的 754 条多肽序列与 SpA 中相应的热点残基之间的 `rmsd`。选择 Encad 全原子力场, 对除去氢原子之外的关键氨基酸残基计算 `rmsd` 值。`rmsd` 值越小, 表示多肽包含的热点残基的对接构象与 SpA 中相应热点残基的构象越接近。结果表明, `rmsd` 值分布在 $0.2 \sim 0.6 \text{ nm}$ 之间 (图 5)。选择 `rmsd` $< 0.4 \text{ nm}$ 的 150 个多肽序列进行下一步的研究。

2. FlexPepDock 复筛多肽分子

一般的分子对接软件如 Autodock, DOCK, PatchDock, ParDock, MEDdock 等只适用于含有一定数目的可旋转键的小分子的对接, 由于多肽侧链较多, 比小分子含有更多的自由度, 因此以上对接软件在用于多肽与蛋白质之间的分子对接筛选时有一定的限制。Flexpepdock 是一种针对研究多肽与目标蛋白对接的新型软件, 该软件采用蒙特卡洛最小化方法, 充分考虑多肽的主、侧链柔性以及受体蛋白的侧链柔性。只需知道结合位点以及多肽-蛋白的大致结合模型, 就能够利用该软件高精度地预测多肽-受体蛋白质之间的结合构象。FlexPepDock 的打分函数是一种全原子能量函数, 包括 Lennard-Jones 全原子吸引和排斥项, Lazaridis-Jarplus 溶剂项以及氢键等。大量研究证实评价多肽-蛋白结合的强弱主要是看结合界面能量分数 (I_{sc})。

利用 Flexpepdock 对上一轮筛选得到 150 个多肽与 Fc 进行逐个对接。结果发现只有少数几个多肽不能与 Fc 结合, 大部分多肽序列都能结合到 Fc 上。这些待选多肽的 I_{sc} 在 $-5 \sim -22$ 之间, 分数越负, 表示结合作用越强。挑选 $I_{sc} < -16$ 的待选多肽序列, 共进行了两次平行对接, 第一次对接结果中有 34 个多肽序列的 I_{sc} 的绝对值大于 16 (图 6)。第二次对接结果中, 有 38 个多肽序列的 I_{sc} 的绝对值大于 16。在这两次 Flexpepdock 对接结果中, I_{sc} 绝对值均大于 16 的多肽序列有 15 条。表 4 中 `vina_score` 表示 vina 对接的打分, I_{sc-1} 表示第一次 Flexpepdock 对接的打分, I_{sc-2} 表示第二次平行 Flexpepdock 对接的打分数。

表 4

Peptide	vina_score	I_{sc-1}	I_{sc-2}
FYWHCLDE	-7.3	-21.3	-22.0
FYFCRWE	-7.0	-16.9	-19.3
FYIHCLPE	-6.9	-18.2	-18.8
FYYHCKKE	-6.5	-16.7	-18.8
FYCHWALE	-7.6	-17.1	-18.1
FYCHWQDE	-7.7	-20.1	-17.8
FYCHTIDE	-6.8	-17.8	-17.4
FYRHCQRE	-6.5	-19.7	-16.9
FYCHHKTE	-6.6	-17.0	-16.9
FYCHLQKE	-6.7	-17.2	-16.8
FYCHRKAE	-6.9	-17.6	-16.8
FYCHNQDE	-6.6	-19.9	-16.7
FYCHRQEE	-6.7	-17.6	-16.6
FYNHCASE	-7.1	-16.8	-16.1

FYTHCAKE

-7.0

-17.1

-16.1

因此，我们挑选这 15 条多肽分子做进一步的分子动力学模拟分析其与 IgG 之间的亲和性。

实施例 4 分子动力学(MD)模拟

MD 模拟是研究蛋白动力学行为的非常有效的工具，蛋白快速的内部运动、较慢的构象变化以及折叠过程都可以应用 MD 模拟进行研究。为了进一步验证 IgG—多肽配基复合物之间的亲和作用力，不仅要考虑其静态结构（利用分子对接），也要研究其动态行为。接下来就利用 MD 模拟来研究上述 15 条候选多肽分子与 IgG 的 Fc 片段结合的动态信息。

所有的 MD 模拟都是用 GROMACS 4.5.3 软件包进行的，选择 GROMOS96 53a6 力场，将多肽-蛋白复合物置于矩形水盒子中心，距离盒子边缘至少 0.9 nm，库仑力和 Lennard-Jones 作用的截断值均为 0.9 nm，PME 方法用于计算长程静电相互作用，其网格间距和内插级数分别为 0.12 nm 和 4，应用 v-rescale 恒温器和 Parrinello-Rahman 恒压器分别控制体系温度（300 K）和压力（1 bar）。积分步长为 2 fs，应用 LINCS 算法限制所有的共价键。先进行 50,000 步最陡下降法将体系进行能量最小化，然后依次进行 100 ps 的 NVT 和 NPT 系综下的限制动力学平衡，最后进行 20 ns 的无限制动力学模拟。每 500 ps 保存一帧构象，共进行 20 ns 的 MD 模拟，最后得到 40 个不同模拟时间的构象。所有的 MD 模拟都是在 Dawning A620r-F 服务器上进行。

首先利用 VMD 软件计算多肽与 Fc 片段之间的相互作用和相对位置的变化，发现 15 个多肽分子当中只有 FYTHCAKE 从 Fc 片段上解离下来，并在 MD 模拟过程中逐渐远离对方。分析 FYTHCAKE 和 Fc 作用时发现 C_α 原子的 rmsd 值随模拟时间的变化，发现在整个 20 ns 的模拟过程中，多肽和 Fc 的 C_α 的 rmsd 值剧烈波动，一直没有达到稳定状态。将其最终构象与初始构象（Flexpepdock 对接构象）相比，其 C_α 原子的 rmsd 值达到 0.45 nm。并分析 FYTHCAKE 与 Fc 原子之间的接触数和最小距离随模拟时间的变化，发现在 15 ns 左右，多肽与 Fc 开始发生脱离，接触数从 600 开始急剧下降，相应地最小距离迅速增大，在 18 ns 左右两者已经完全脱离。说明 FYTHCAKE 和 IgG 之间的亲和作用不高。相反，在整个分子动力学模拟过程中其他多肽分子均能与 Fc 保持结合状态，说明这 14 个多肽分子可能是 IgG 的有效亲和配基，分别是 FYWHCLDE、FYFCRWE、FYIHCLPE、FYYHCKKE、FYCHWALE、FYCHWQDE、FYCHTIDE、FYRHCQRE、FYCHHKTE、FYCHLQKE、FYCHRKAE、FYCHNQDE、FYCHRQEE 和 FYNHCASE。

实施例 5 亲和色谱实验验证

1. 肽配基的固定和亲和柱的制备

多肽分子委托吉尔生化有限公司(上海)进行合成，产品经高效液相色谱(HPLC)纯化，纯度为 96.36%。称取 1 g 干介质（Thiopropyl Sepharose 6B 购自 GE Healthcare），用大约 200 mL 过膜水在布氏漏斗中清洗 15 min，滤除水分后从中分别称取 1g 湿介质加入到两个 25 mL 锥形瓶中，然后分别加入 6 mL 和 10.7 mL 交联缓冲液（0.1 M Tris-HCl, pH 7.5, 0.5 M NaCl, 1 mM EDTA）使介质充分预平衡。从分子动力学模拟得到的 14 条多肽中随机挑选多肽分子（FYWHCLE 和 FYCHWALE），分别称取上述肽粉，溶于 500 μL 50%乙二醇溶液中。将肽溶液和对应的锥形瓶中的介质充分混合，使得肽溶液初始浓度为 1.0 mg/mL，25℃ 170 rpm 水浴摇床反应 2 小时。每隔 1h 离心取 500 μL 反应上清液，用于反相高压液相色谱（RP-HPLC）检测反应液中剩余的肽含量。待反应上清液中肽含量不再变化后，加入 8 mg Cys 继续反应 0.5 h，封闭介质上的未反应基团以免干扰蛋白与配基的亲和作用。反应完毕后，离心，移去上清液，然后将介质用清洗缓冲液（10 mM 磷酸盐缓冲液，pH 7.2，150 mM NaCl）在布氏漏斗中反复清洗多次，洗去未固定的多肽。最后将介质用平衡缓冲液（20 mM 柠檬酸钠盐缓冲液，pH 5.0~5.5；或 20 mM 磷酸钠盐缓冲液，pH 5.5~6.0）悬浮为 1 mL，脱气后缓慢灌入玻璃柱内（Tricorn 型色谱柱，Tricorn 5×5，GE 公司），先用平衡缓冲液以流速 0.1 mL/min 冲洗柱子，待柱压稳定后将流速升高至 0.2 mL/min，按此操作直至流速升高至 1.0 mL/min，观察凝胶柱高度不再下降时，将高度调节器旋至凝胶柱上表面，即可。

2. 溶液中肽含量的检测

反应液中肽含量的检测采用反相高压液相色谱法，具体检测参数如下：流动相 A，含 0.1% 三氟乙酸的水溶液；流动相 B，含 0.1% 三氟乙酸的乙腈；上样量：10 μL ；流速：0.5 mL/min；检测波长：220 nm。其中多肽 FYWHCLDE 几乎能全部固定到介质上，FYCHWALE 的固定化率为 93.7%。

3. 亲和色谱实验

(1)平衡：用平衡缓冲液冲洗直至基线走平，流速 0.5 mL/min，继续冲洗 5 倍柱体积开始上样。

(2)上样：除非特别声明，蛋白样品均以平衡缓冲液配制，每次上样后用 5-10 倍柱体积的平衡缓冲液（20 mM 柠檬酸钠盐 buffer，pH $\leq$ 5.5；20 mM 磷酸钠盐 buffer，pH $\geq$ 5.5）进行平衡。上样体积为 100 μL ，蛋白样品浓度为 1.0 mg/mL，流速：0.5 mL/min。

(3)洗脱：用洗脱缓冲液冲洗 5-10 个柱体积，流速：0.8 mL/min。

(4)保存：先用 10 倍柱体积的纯水冲洗柱子，然后再用 10 倍柱体积的 20%乙醇溶液冲洗柱子，并将其保存于该溶液中。所有缓冲液均要过膜脱气。

在 pH 5.0~6.0 时，FYWHCLDE 肽配基对 hIgG 有较强吸吸附作用。相反，对 BSA 只有在 pH 5.5 时才有较大吸附。FYWHCLDE 肽配基与 hIgG 之间的亲和作用主要涉及特异性静电相互作用。在 pH 5.5~7.0 时，配基 FYCHWALE 对 hIgG 都有较大吸附，但在 pH 6.0 时几乎能吸附全部的 hIgG。因此磷酸钠盐缓冲液（20 mM，pH 6.0）为 FYCHWALE 吸附 hIgG 的最佳吸附条件。肽配基与 hIgG 之间的亲和作用主要涉及静电相互作用，因此吸附体系的离子强度要适宜，不能太高，以免引起肽配基对抗体蛋白的吸附力减弱。FYCHWALE 在 pH 5.5 时能吸附大约一半的 BSA，pH 6.0 时吸附 1/3 的 BSA，pH 6.5 时只吸附很少量的 BSA。以上结果说明两条多肽在 pH 6.0 时对 hIgG 都有特异性吸附。

4. 蛋白含量测定

混合蛋白溶液中蛋白总量的测定均采用 Bradford 法，操作如下：配制一系列浓度为 15、30、45、60、75 和 90 $\mu\text{g/mL}$ 的 BSA 溶液各 1 mL，空白对照为 1mL 蒸馏水。测定每个数据点时，向蛋白溶液和对照溶液中各加入考马斯亮蓝 G-250 溶液 3 mL，充分混和，1-2 分钟内于 595 nm 下测定吸光值，蛋白浓度对吸光值作图即可得到蛋白含量标准曲线。为了制作 IgG 浓度的标准曲线，分别配制 0~2.0 mg/mL 浓度的 hIgG 溶液，用紫外分光光度法测 280 nm 下的吸收值，然后绘制标准曲线。IgG 标准曲线用于测定只含有 IgG 的纯蛋白溶液。测定色谱各组分中蛋白在 280 nm 和 595 nm 下的吸光值，然后根据标准曲线计算其中的蛋白含量。

5. 对血清样品中 IgG 的色谱分离纯化

亲和介质经平衡缓冲液（20 mM 磷酸钠盐缓冲液，pH 6.0 或 6.5）充分平衡后，即各项指标基线都达到平衡，将人血清与吸附缓冲液按 1:9 混合得到血清样品，进样 500 μL 。继续用平衡缓冲液冲洗 5 个柱体积，然后用含有 0.5 或 0.2 M NaCl 的平衡缓冲液进行洗脱；或是 50 mM 柠檬酸钠盐缓冲液（pH 3.0）进行洗脱。待洗脱峰完全分离后，用 0.1 M Gly-HCl 缓冲液（pH 2.4）进行亲和介质再生。分离结果如图 7 所示。

6. 十二烷基硫酸钠—聚丙烯酰胺凝胶电泳(SDS-PAGE)

主要溶液：

1). 10%的分离胶，4 mL：1330 μL 的 30%的凝胶（丙烯酰胺：甲叉双丙=29:1）；1670 μL 的去离子水；1000 μL 的 1.5M 的 Low Tris 缓冲液；40 μL 的 10%过硫酸胺（APS）；10 μL 的 TEMED。

2). Low Tris 缓冲液：称取 Tris 18.17 g，2.5 mL 浓 HCl，SDS 0.4 g，加水溶解，定容至 100 mL，调 pH 至 8.8。

3). 10%过硫酸胺：称取 APS 0.1 g，加水溶解，定容至 1 mL。

4). Upp Tris 缓冲液：称取 Tris 3.03 g，2 mL 浓 HCl，SDS 0.2 g，加水溶解，定容至 50 mL，调 pH 至 6.8。

5). 5%的浓缩胶，4 mL：665 μL 的 30%的凝胶；2335 μL 的去离子水；1000 μL 的 1.0 M 的 Upp Tris；40 μL 的 10%APS；16 μL 的 TEMED。

6). 非还原型电泳样品缓冲液（Non-reducing Sample buffer）：取 Upp Tris 5.5 mL，甘油 8.8 mL，再称取 SDS 2 g，溴酚蓝 5-10 mg，加水溶解，定容至 50 mL。

7). PAGE 电泳染色液: 称取考马斯亮蓝 (R-250) 0.625 g, 甲醇 250 mL 和乙酸 50 mL 混合, 定容至 500 mL。

8). PAGE 电泳脱色液: 取甲醇 150 mL 和乙酸 50 mL 混合, 定容至 500 mL。

9). 电泳缓冲液 (PAGE buffer): 称取 Tris 7.5 g, 十二烷基硫酸钠 (SDS) 1.0 g 和甘氨酸 36 g, 溶于水, 定容 500 mL, 调节 pH 8.3。使用方法: 每次取 100 mL, 用水稀释 5 倍; 电泳缓冲液要没过平台, 并在电泳结束后回收。

样品准备: 取 25 μ L 的原样品溶液和 25 μ L 的样品 buffer 混合均匀。Maker 进样 5 μ L, 样品进样 10 μ L。取 SDS-PAGE 的分离胶加到 SE/250 型垂直板电泳槽中, 分离胶加入量要控制在距玻璃板 1.5 cm 处, 然后轻轻在分离胶上覆盖一层水, 待分离胶凝固后, 倒出水, 加入浓缩胶并插入梳子。待浓缩胶凝固后拔出梳子, 用微量进样器向梳子孔内加入待测样品。样品中蛋白质含量以 10 μ g 为宜。选择在 10 mA 的电流下运行, 直到指示带迁移到浓缩胶底部, 然后再将电流调至 25 mA, 直到指示带迁移到凝胶底部, 结束电泳。电泳结束后, 采用考马斯亮蓝 (R-250) 染色法, 该方法灵敏度高, 最低检测极限为 0.3~1.0 μ g。将凝胶浸入染色液中, 室温染色 4h, 这样可以同时达到固定和染色的目的。再利用脱色液脱色, 直到凝胶背景接近无色为止。所得凝胶经凝胶成像仪照相, 用 Gel-Pro 分析软件对凝胶结果进行分析。

对分离人血清得到的流出峰和洗脱峰组分进行电泳, 电泳图如图 8 所示。用 Gel-Pro 分析软件计算得到 IgG 纯度, 用 Bradford 法计算回收率。人血清中 IgG 含量为 18.07%, FYWHCLDE 纯化的洗脱峰组分中 IgG 含量 88.86%, 回收率 65.5%。FYCHWALE 纯化的洗脱峰组分中 IgG 含量分别为 89.6% (pH 6.5 吸附体系) 和 88.5% (pH 6.0 吸附体系), IgG 得率达 71% (pH 6.0 吸附体系)。和报导中提到的其他 hIgG 亲和配基的分离纯化效果相比, 本发明利用的肽配基纯化得到的 hIgG 的纯度以及得率都较高, 因此是 hIgG 的有效亲和配基。

本发明所介绍的分子模拟设计人 IgG 亲和肽配基的方法与其它的构建新型肽配基的方法如组合化学合成肽库筛选、噬菌体展示肽库筛选以及核糖体展示肽库筛选相比, 有如下优点: 第一, 不需要复杂昂贵的原材料, 比实验手段操作简单, 成本低廉; 第二, 可以构建数量庞大的候选分子库, 实现真正意义上的大规模高通量筛选。筛选得到的多肽分子不仅是人 IgG 的亲和配基, 也适用于其他物质, 如也能对猪和羊 IgG 等进行有效地分离纯化。本发明所使用的分子对接筛选、分子动力学模拟复筛和最后的实验验证, 对多肽库中的所有多肽分子都是适用的, 由于各种限制因素, 不能将肽库中的全部多肽分子都进行实验验证, 只能通过各种筛选方法进行有效候选多肽分子的富集。但有一点值得肯定的是, 除了最终得到的 14 条多肽序列, 散布在多肽库中的其他多肽分子也可能是 IgG 的亲和肽配基, 相关技术人员如有需要可挑选并进行实验验证。但根据已取得的实验数据, 表明肽库中的多肽分子是人 IgG 的有效候选亲和肽配基, 希望有效果较佳的肽配基能够得到商业化, 进而造福人类。

表 5

FYCHAAAE	FYCHAARE	FYCHAADE	FYCHAAQE	FYCHAAEE	FYCHAALE	FYCHAAKE
FYCHAAME	FYCHAAPE	FYCHAASE	FYCHAATE	FYCHAQAE	FYCHAQRE	FYCHAQDE
FYCHAQQE	FYCHAQEE	FYCHAQLE	FYCHAQKE	FYCHAQME	FYCHAQPE	FYCHAQSE
FYCHAQTE	FYCHAEAE	FYCHAEERE	FYCHAEDE	FYCHAEQE	FYCHAEERE	FYCHAELE
FYCHAEKE	FYCHAE ME	FYCHAEPE	FYCHAESE	FYCHAE TE	FYCHAI AE	FYCHAI RE
FYCHAI DE	FYCHAI QE	FYCHAI EE	FYCHAI LE	FYCHAI KE	FYCHAI ME	FYCHAI PE
FYCHAI SE	FYCHAI TE	FYCHAL AE	FYCHAL RE	FYCHAL DE	FYCHAL QE	FYCHAL EE
FYCHAL LE	FYCHAL KE	FYCHAL ME	FYCHAL PE	FYCHAL SE	FYCHAL TE	FYCHAK AE
FYCHAK RE	FYCHAK DE	FYCHAK QE	FYCHAK EE	FYCHAK LE	FYCHAK KE	FYCHAK ME
FYCHAK PE	FYCHAK SE	FYCHAK TE	FYCHRA AE	FYCHRA RE	FYCHRA DE	FYCHRA QE
FYCHRA EE	FYCHRA LE	FYCHRA KE	FYCHRA ME	FYCHRA PE	FYCHRA SE	FYCHRA TE

WO 2014/094444

PCT/CN2013/081761

FYCHRQAE	FYCHRQRE	FYCHRQDE	FYCHRQQE	FYCHRQEE	FYCHRQLE	FYCHRQKE
FYCHRQME	FYCHRQPE	FYCHRQSE	FYCHRQTE	FYCHREAE	FYCHRERE	FYCHREDE
FYCHREQE	FYCHREEE	FYCHRELE	FYCHREKE	FYCHREME	FYCHREPE	FYCHRESE
FYCHRETE	FYCHRIAE	FYCHRIRE	FYCHRIDE	FYCHRIQE	FYCHRIEE	FYCHRILE
FYCHRIKE	FYCHRIME	FYCHRIPE	FYCHRISE	FYCHRITE	FYCHRLAE	FYCHRLRE
FYCHRLDE	FYCHRLQE	FYCHRLLE	FYCHRLKE	FYCHRLME	FYCHRLPE	FYCHRLSE
FYCHRLSE	FYCHRLTE	FYCHRKAE	FYCHRKRE	FYCHRKDE	FYCHRKQE	FYCHRKKE
FYCHRKLE	FYCHRKKE	FYCHRKME	FYCHRKPE	FYCHRKSE	FYCHRKTE	FYCHNKTE
FYCHNAAE	FYCHNARE	FYCHNADE	FYCHNAQE	FYCHNAEE	FYCHNALE	FYCHNAKE
FYCHNAME	FYCHNAPE	FYCHNASE	FYCHNATE	FYCHNQAE	FYCHNQRE	FYCHNQDE
FYCHNQQE	FYCHNQEE	FYCHNQLE	FYCHNQKE	FYCHNQME	FYCHNQPE	FYCHNQSE
FYCHNQTE	FYCHNEAE	FYCHNERE	FYCHNEDE	FYCHNEQE	FYCHNEEE	FYCHNELE
FYCHNEKE	FYCHNEME	FYCHNEPE	FYCHNESE	FYCHNETE	FYCHNIAE	FYCHNIRE
FYCHNIDE	FYCHNIQE	FYCHNIEE	FYCHNILE	FYCHNIKE	FYCHNIME	FYCHNIPE
FYCHNISE	FYCHNITE	FYCHNLAE	FYCHNLRE	FYCHNLDE	FYCHNLQE	FYCHNLEE
FYCHNLLE	FYCHNLKE	FYCHNLME	FYCHNLPE	FYCHNLSE	FYCHNLTE	FYCHNKAE
FYCHNKRE	FYCHNKDE	FYCHNKQE	FYCHNKKE	FYCHNKLE	FYCHNKKE	FYCHNKME
FYCHNKPE	FYCHNKSE	FYCHDKTE	FYCHDKPE	FYCHDKSE	FYCHDKRE	FYCHDKDE
FYCHDAAE	FYCHDARE	FYCHDADE	FYCHDAQE	FYCHDAEE	FYCHDALE	FYCHDAKE
FYCHDAME	FYCHDAPE	FYCHDASE	FYCHDATE	FYCHDQAE	FYCHDQRE	FYCHDQDE
FYCHDQQE	FYCHDQEE	FYCHDQLE	FYCHDQKE	FYCHDQME	FYCHDQPE	FYCHDQSE
FYCHDQTE	FYCHDEAE	FYCHDERE	FYCHDEDE	FYCHDEQE	FYCHDEEE	FYCHDELE
FYCHDEKE	FYCHDEME	FYCHDEPE	FYCHDESE	FYCHDETE	FYCHDIAE	FYCHDIRE
FYCHDIDE	FYCHDIQE	FYCHDIEE	FYCHDILE	FYCHDIKE	FYCHDIME	FYCHDIPE
FYCHDISE	FYCHDITE	FYCHDLAE	FYCHDLRE	FYCHDLDE	FYCHDLQE	FYCHDLEE
FYCHDLLE	FYCHDLKE	FYCHDLME	FYCHDLPE	FYCHDLSE	FYCHDLTE	FYCHDKAE
FYCHDKQE	FYCHDKKE	FYCHDKLE	FYCHDKKE	FYCHDKME	FYCHQKTE	FYCHQKPE
FYCHQAAE	FYCHQARE	FYCHQADE	FYCHQAQE	FYCHQAE	FYCHQALE	FYCHQAKE
FYCHQAME	FYCHQAPE	FYCHQASE	FYCHQATE	FYCHQQAE	FYCHQQRE	FYCHQQDE
FYCHQQQE	FYCHQQEE	FYCHQQLE	FYCHQQKE	FYCHQQME	FYCHQQPE	FYCHQQSE
FYCHQQTE	FYCHQEAE	FYCHQERE	FYCHQEDE	FYCHQEQE	FYCHQEEE	FYCHQELE
FYCHQEKE	FYCHQEME	FYCHQEPE	FYCHQESE	FYCHQETE	FYCHQIAE	FYCHQIRE
FYCHQIDE	FYCHQIQE	FYCHQIEE	FYCHQILE	FYCHQIKE	FYCHQIME	FYCHQIPE
FYCHQISE	FYCHQITE	FYCHQLAE	FYCHQLRE	FYCHQLDE	FYCHQLQE	FYCHQLEE
FYCHQLLE	FYCHQLKE	FYCHQLME	FYCHQLPE	FYCHQLSE	FYCHQLTE	FYCHQKAE
FYCHQKQE	FYCHQKEE	FYCHQKLE	FYCHQKKE	FYCHQKME	FYCHQKSE	FYCHQKRE

WO 2014/094444

PCT/CN2013/081761

FYCHQKDE	FYCHEKTE	FYCHEKPE	FYCHEKDE	FYCHEKME	FYCHEKSE	FYCHEKRE
FYCHEAAE	FYCHEARE	FYCHEADE	FYCHEAQE	FYCHEAEE	FYCHEALE	FYCHEAKE
FYCHEAME	FYCHEAPE	FYCHEASE	FYCHEATE	FYCHEQAE	FYCHEQRE	FYCHEQDE
FYCHEQQE	FYCHEQEE	FYCHEQLE	FYCHEQKE	FYCHEQME	FYCHEQPE	FYCHEQSE
FYCHEQTE	FYCHEEAE	FYCHEERE	FYCHEEDE	FYCHEEQE	FYCHEEEE	FYCHEELE
FYCHEEKE	FYCHEEME	FYCHEEPE	FYCHEESE	FYCHEETE	FYCHEIAE	FYCHEIRE
FYCHEIDE	FYCHEIQE	FYCHEIEE	FYCHEILE	FYCHEIKE	FYCHEIME	FYCHEIPE
FYCHEISE	FYCHEITE	FYCHELAE	FYCHELRE	FYCHELDE	FYCHELQE	FYCHELEE
FYCHELLE	FYCHELKE	FYCHELME	FYCHELPE	FYCHELSE	FYCHELTE	FYCHEKAE
FYCHEKQE	FYCHEKEE	FYCHEKLE	FYCHEKKE	FYCHHKTE	FYCHHKPE	FYCHHKDE
FYCHHAAE	FYCHHARE	FYCHHADE	FYCHHAQE	FYCHHAEE	FYCHHALE	FYCHHAKE
FYCHHAME	FYCHHAPE	FYCHHASE	FYCHHATE	FYCHHQAE	FYCHHQRE	FYCHHQDE
FYCHHQQE	FYCHHQEE	FYCHHQLE	FYCHHQKE	FYCHHQME	FYCHHQPE	FYCHHQSE
FYCHHQTE	FYCHHEAE	FYCHHERE	FYCHHEDE	FYCHHEQE	FYCHHEEE	FYCHHELE
FYCHHEKE	FYCHHEME	FYCHHEPE	FYCHHESE	FYCHHETE	FYCHHIAE	FYCHHIRE
FYCHHIDE	FYCHHIQE	FYCHHIEE	FYCHHILE	FYCHHIKE	FYCHHIME	FYCHHIPE
FYCHHISE	FYCHHITE	FYCHHLAE	FYCHHLRE	FYCHHLDE	FYCHHLQE	FYCHHLEE
FYCHHLLLE	FYCHHLKE	FYCHHLME	FYCHHLPE	FYCHHLSE	FYCHHLTE	FYCHHKAE
FYCHHKQE	FYCHHKKE	FYCHHKLE	FYCHHKKE	FYCHHKME	FYCHHKSE	FYCHHKRE
FYCHIAAE	FYCHIARE	FYCHIADE	FYCHIAQE	FYCHIAEE	FYCHIALE	FYCHIAKE
FYCHIA ME	FYCHIAPE	FYCHIASE	FYCHIA TE	FYCHIQAE	FYCHIQRE	FYCHIQDE
FYCHIQQE	FYCHIQEE	FYCHIQLE	FYCHIQKE	FYCHIQME	FYCHIQPE	FYCHIQSE
FYCHIQTE	FYCHIEAE	FYCHIERE	FYCHIEDE	FYCHIEQE	FYCHIEEE	FYCHIELE
FYCHIEKE	FYCHIE ME	FYCHIEPE	FYCHIESE	FYCHIE TE	FYCHIIAE	FYCHIIRE
FYCHIIDE	FYCHIIQE	FYCHIIEE	FYCHIILE	FYCHIIKE	FYCHII ME	FYCHIIPE
FYCHIISE	FYCHIITE	FYCHILAE	FYCHILRE	FYCHILDE	FYCHILQE	FYCHILEE
FYCHILLE	FYCHILKE	FYCHILME	FYCHILPE	FYCHILSE	FYCHILTE	FYCHIKAE
FYCHIKQE	FYCHIKKE	FYCHIKLE	FYCHIKKE	FYCHIKME	FYCHIKSE	FYCHIKRE
FYCHIKTE	FYCHIKPE	FYCHIKDE	FYCHLAAE	FYCHLARE	FYCHLADE	FYCHLAQE
FYCHLAME	FYCHLAPE	FYCHLASE	FYCHLATE	FYCHLQAE	FYCHLQRE	FYCHLQDE
FYCHLQQE	FYCHLQEE	FYCHLQLE	FYCHLQKE	FYCHLQME	FYCHLQPE	FYCHLQSE
FYCHLQTE	FYCHLEAE	FYCHLERE	FYCHLEDE	FYCHLEQE	FYCHLEEE	FYCHLELE
FYCHLEKE	FYCHLE ME	FYCHLEPE	FYCHLESE	FYCHLE TE	FYCHLIAE	FYCHLIRE
FYCHLIDE	FYCHLIQE	FYCHLIEE	FYCHLILE	FYCHLIKE	FYCHLIME	FYCHLIPE
FYCHLISE	FYCHLITE	FYCHLLAE	FYCHLLRE	FYCHLLDE	FYCHLLQE	FYCHLLEE
FYCHLLLE	FYCHLLKE	FYCHLLME	FYCHLLPE	FYCHLLSE	FYCHLLTE	FYCHLKAE

WO 2014/094444

PCT/CN2013/081761

FYCHLKQE	FYCHLKEE	FYCHLKLE	FYCHLKKE	FYCHLKME	FYCHLKSE	FYCHLKRE
FYCHLAEE	FYCHLALE	FYCHLAKE	FYCHLKTE	FYCHLKPE	FYCHLKDE	FYCHKAAE
FYCHKAME	FYCHKAPE	FYCHKASE	FYCHKATE	FYCHKQAE	FYCHKQRE	FYCHKQDE
FYCHKQQE	FYCHKQEE	FYCHKQLE	FYCHKQKE	FYCHKQME	FYCHKQPE	FYCHKQSE
FYCHKQTE	FYCHKEAE	FYCHKERE	FYCHKEDE	FYCHKEQE	FYCHKEEE	FYCHKELE
FYCHKEKE	FYCHKEME	FYCHKEPE	FYCHKESE	FYCHKETE	FYCHKIAE	FYCHKIRE
FYCHKIDE	FYCHKIQE	FYCHKIEE	FYCHKILE	FYCHKIKE	FYCHKIME	FYCHKIPE
FYCHKISE	FYCHKITE	FYCHKLAE	FYCHKLRE	FYCHKLDE	FYCHKLQE	FYCHKLEE
FYCHKLLE	FYCHKLKE	FYCHKLME	FYCHKLPE	FYCHKLSE	FYCHKLTE	FYCHKKAE
FYCHKKQE	FYCHKKEE	FYCHKKLE	FYCHKKKE	FYCHKKME	FYCHKKSE	FYCHKKRE
FYCHKARE	FYCHKADE	FYCHKAQE	FYCHKAEE	FYCHKALE	FYCHKAKE	FYCHKKTE
FYCHKKPE	FYCHKKDE	FYCHMAAE	FYCHMKPE	FYCHMKDE	FYCHMARE	FYCHMADE
FYCHMAME	FYCHMAPE	FYCHMASE	FYCHMATE	FYCHMQAE	FYCHMQRE	FYCHMQDE
FYCHMQQE	FYCHMQEE	FYCHMQLE	FYCHMQKE	FYCHMQME	FYCHMQPE	FYCHMQSE
FYCHMQTE	FYCHMEAE	FYCHMERE	FYCHMEDE	FYCHMEQE	FYCHMEEE	FYCHMELE
FYCHMEKE	FYCHMEME	FYCHMEPE	FYCHMESE	FYCHMETE	FYCHMIAE	FYCHMIRE
FYCHMIDE	FYCHMIQE	FYCHMIEE	FYCHMILE	FYCHMIKE	FYCHMIME	FYCHMIPE
FYCHMISE	FYCHMITI	FYCHMLAE	FYCHMLRE	FYCHMLDE	FYCHMLQE	FYCHMLLE
FYCHMLLE	FYCHMLKE	FYCHMLME	FYCHMLPE	FYCHMLSE	FYCHMLTE	FYCHMKAE
FYCHMKQE	FYCHMKEE	FYCHMKLE	FYCHMKKE	FYCHMKME	FYCHMKSE	FYCHMKRE
FYCHMAQE	FYCHMAEE	FYCHMALE	FYCHMAKE	FYCHMKTE	FYCHTADE	FYCHTAQE
FYCHTAME	FYCHTAAE	FYCHTKPE	FYCHTKDE	FYCHTARE	FYCHTQRE	FYCHTQDE
FYCHTQQE	FYCHTAPE	FYCHTASE	FYCHTATE	FYCHTQAE	FYCHTQPE	FYCHTQSE
FYCHTQTE	FYCHTQEE	FYCHTQLE	FYCHTQKE	FYCHTQME	FYCHTEEE	FYCHTELE
FYCHTEKE	FYCHTEAE	FYCHTERE	FYCHTEDE	FYCHTEQE	FYCHTIAE	FYCHTIRE
FYCHTIDE	FYCHTEME	FYCHTEPE	FYCHTESE	FYCHTETE	FYCHTIME	FYCHTIPE
FYCHTISE	FYCHTIQE	FYCHTIEE	FYCHTILE	FYCHTIKE	FYCHTLQE	FYCHTLEE
FYCHTLLE	FYCHTITE	FYCHTLAE	FYCHTLRE	FYCHTLDE	FYCHTLTE	FYCHTKAE
FYCHTKQE	FYCHTLKE	FYCHTLME	FYCHTLPE	FYCHTLSE	FYCHTKSE	FYCHTKRE
FYCHTAE	FYCHTKEE	FYCHTKLE	FYCHTKKE	FYCHTKME	FYCHWKPE	FYCHWKDE
FYCHWAME	FYCHTALE	FYCHTAKE	FYCHTKTE	FYCHWAAE	FYCHWQRE	FYCHWQDE
FYCHWQQE	FYCHWAPE	FYCHWASE	FYCHWATE	FYCHWQAE	FYCHWQPE	FYCHWQSE
FYCHWQTE	FYCHWQEE	FYCHWQLE	FYCHWQKE	FYCHWQME	FYCHWEEE	FYCHWELE
FYCHWEKE	FYCHWEAE	FYCHWERE	FYCHWEDE	FYCHWEQE	FYCHWIAE	FYCHWIRE
FYCHWIDE	FYCHWEME	FYCHWEPE	FYCHWESE	FYCHWETE	FYCHWIME	FYCHWIPE
FYCHWISE	FYCHWIQE	FYCHWIEE	FYCHWILE	FYCHWIKE	FYCHWLQE	FYCHWLEE

WO 2014/094444

PCT/CN2013/081761

FYCHWLLE	FYCHWITE	FYCHWLAE	FYCHWLRE	FYCHWLDE	FYCHWLTE	FYCHWKAE
FYCHWKQE	FYCHWLKE	FYCHWLME	FYCHWLPE	FYCHWLSE	FYCHWKSE	FYCHWKRE
FYCHWARE	FYCHWKEE	FYCHWKLE	FYCHWKKE	FYCHWKME	FYCHWAKE	FYCHWKTE
FYCHYAME	FYCHWADE	FYCHWAQE	FYCHWAEE	FYCHWALE	FYCHYQRE	FYCHYQDE
FYCHYQQE	FYCHYAPE	FYCHYASE	FYCHYATE	FYCHYQAE	FYCHYQPE	FYCHYQSE
FYCHWQTE	FYCHYQEE	FYCHYQLE	FYCHYQKE	FYCHYQME	FYCHWEEE	FYCHWELE
FYCHWEKE	FYCHWEAE	FYCHWERE	FYCHWEDE	FYCHWEQE	FYCHWIAE	FYCHWIRE
FYCHYIDE	FYCHWEME	FYCHWEPE	FYCHWESE	FYCHWETE	FYCHYIME	FYCHYIPE
FYCHYISE	FYCHYIQE	FYCHYIEE	FYCHYILE	FYCHYIKE	FYCHYLQE	FYCHYLEE
FYCHYLLE	FYCHYITE	FYCHYLAE	FYCHYLRE	FYCHYLDE	FYCHYLTE	FYCHYKAE
FYCHYKQE	FYCHYLKE	FYCHYLME	FYCHYLPE	FYCHYLSE	FYCHYKSE	FYCHYKRE
FYCHYARE	FYCHYKEE	FYCHYKLE	FYCHYKKE	FYCHYKME	FYCHYAKE	FYCHYKTE
FYCHYAAE	FYCHYADE	FYCHYAQE	FYCHYAEI	FYCHYALE	FYCHVKDE	FYCHVQSE
FYCHWQTE	FYCHYKPE	FYCHYKDE	FYCHVAAE	FYCHVKPE	FYCHWEEE	FYCHWELE
FYCHWEKE	FYCHWEAE	FYCHWERE	FYCHWEDE	FYCHWEQE	FYCHWIAE	FYCHWIRE
FYCHVIDE	FYCHWEME	FYCHWEPE	FYCHWESE	FYCHWETE	FYCHVIME	FYCHVIPE
FYCHVISE	FYCHVIQE	FYCHVIEE	FYCHVILE	FYCHVIKE	FYCHVLQE	FYCHVLEE
FYCHVLLE	FYCHVITE	FYCHVLAE	FYCHVLRE	FYCHVLDE	FYCHVLTE	FYCHVKAE
FYCHVKQE	FYCHVLKE	FYCHVLME	FYCHVLPE	FYCHVLSE	FYCHVKSE	FYCHVKRE
FYCHVARE	FYCHVKEE	FYCHVKLE	FYCHVKKE	FYCHVKME	FYCHVAKE	FYCHVKTE
FYCHVAME	FYCHVADE	FYCHVAQE	FYCHVAEE	FYCHVALE	FYCHVQRE	FYCHVQDE
FYCHVQQE	FYCHVAPE	FYCHVASE	FYCHVATE	FYCHVQAE	FYCHVQPE	FYCHVQEE
FYCHVQLE	FYCHVQKE	FYCHVQME	FYAHCAAE	FYAHCARE	FYAHCADE	FYAHCAQE
FYAHCAEE	FYAHCALE	FYAHCAKE	FYAHCAME	FYAHCAPE	FYAHCASE	FYAHCATE
FYAHCQEE	FYAHCQLE	FYAHCQKE	FYAHCQME	FYAHCQPE	FYAHCQSE	FYAHCQTE
FYAHCQAE	FYAHCQRE	FYAHCQDE	FYAHCQQE	FYAHCEAE	FYAHCERE	FYAHCEDE
FYAHCEEE	FYAHCELE	FYAHCEKE	FYAHCEME	FYAHCEPE	FYAHCESE	FYAHCETE
FYAHCEQE	FYAHCIAE	FYAHCIRE	FYAHCIDE	FYAHCIQE	FYAHCIEE	FYAHCILE
FYAHCIKE	FYAHCIME	FYAHCIPE	FYAHCISE	FYAHCITE	FYAHCLAE	FYAHCLRE
FYAHCLDE	FYAHCLQE	FYAHCLLE	FYAHCLLE	FYAHCLKE	FYAHCLME	FYAHCLPE
FYAHCLSE	FYAHCLTE	FYAHCKAE	FYAHCKRE	FYAHCKSE	FYAHCKTE	FYAHCKPE
FYAHCKDE	FYAHCKQE	FYAHCKEE	FYAHCKLE	FYAHCKKE	FYAHCKME	FYRHCAAE
FYRHCAEE	FYRHCALE	FYRHCAKE	FYRHCAME	FYRHCAPE	FYRHCASE	FYRHCATE
FYRHCQEE	FYRHCQLE	FYRHCQKE	FYRHCQME	FYRHCQPE	FYRHCQSE	FYRHCQTE
FYRHCQAE	FYRHCQRE	FYRHCQDE	FYRHCQQE	FYRHCEAE	FYRHCERE	FYRHCEDE
FYRHCEEE	FYRHCELE	FYRHCEKE	FYRHCEME	FYRHCEPE	FYRHCESE	FYRHCETE

WO 2014/094444

PCT/CN2013/081761

FYRHCEQE	FYRHCIAE	FYRHCIRE	FYRHCIDE	FYRHCIQE	FYRHCIEE	FYRHCILE
FYRHCIKE	FYRHCIME	FYRHCIPE	FYRHCISE	FYRHCITE	FYRHCLAE	FYRHCLRE
FYRHCLDE	FYRHCLQE	FYRHCLLE	FYRHCLLE	FYRHCLKE	FYRHCLME	FYRHCLPE
FYRHCLSE	FYRHCLTE	FYRHCKAE	FYRHCKRE	FYRHCKSE	FYRHCKTE	FYRHCKPE
FYRHCARE	FYRHCADE	FYRHCAQE	FYRHCKDE	FYRHCKQE	FYRHCKEE	FYRHCKLE
FYRHCKKE	FYRHCKME	FYNHCAAE	FYNHCKKE	FYNHCKME	FYNHCARE	FYNHCADE
FYNHCAEE	FYNHCALE	FYNHCAKE	FYNHCAME	FYNHCAPE	FYNHCASE	FYNHCATE
FYNHCQEE	FYNHCQLE	FYNHCQKE	FYNHCQME	FYNHCQPE	FYNHCQSE	FYNHCQTE
FYNHCQAE	FYNHCQRE	FYNHCQDE	FYNHCQQE	FYNHCEAE	FYNHCERE	FYNHCEDE
FYNHCEEE	FYNHCELE	FYNHCEKE	FYNHCEME	FYNHCEPE	FYNHCESE	FYNHCETE
FYNHCEQE	FYNHCIAE	FYNHCIRE	FYNHCIDE	FYNHCIQE	FYNHCIEE	FYNHCILE
FYNHCIKE	FYNHCIME	FYNHCIPE	FYNHCISE	FYNHCITE	FYNHCLAE	FYNHCLRE
FYNHCLDE	FYNHCLQE	FYNHCLLE	FYNHCLLE	FYNHCLKE	FYNHCLME	FYNHCLPE
FYNHCLSE	FYNHCLTE	FYNHCKAE	FYNHCKRE	FYNHCKSE	FYNHCKTE	FYNHCKPE
FYNHCAQE	FYNHCKDE	FYNHCKQE	FYNHCKEE	FYNHCKLE	FYDHCAAE	FYDHCKKE
FYDHCAEE	FYDHCALE	FYDHCAKE	FYDHCAME	FYDHCAPE	FYDHCASE	FYDHCATE
FYDHCQEE	FYDHCQLE	FYDHCQKE	FYDHCQME	FYDHCQPE	FYDHCQSE	FYDHCQTE
FYDHCQAE	FYDHCQRE	FYDHCQDE	FYDHCQQE	FYDHCEAE	FYDHCERE	FYDHCED
FYDHCEEE	FYDHCELE	FYDHCEKE	FYDHCEME	FYDHCEPE	FYDHCESE	FYDHCETE
FYDHCEQE	FYDHCIAE	FYDHCIRE	FYDHCIDE	FYDHCIQE	FYDHCIEE	FYDHCILE
FYDHCIKE	FYDHCIME	FYDHCIPE	FYDHCISE	FYDHCITE	FYDHCCLAE	FYDHCCLRE
FYDHCCLDE	FYDHCCLQE	FYDHCCLLE	FYDHCCLLE	FYDHCCLKE	FYDHCCLME	FYDHCCLPE
FYDHCCLSE	FYDHCCLTE	FYDHCCKAE	FYDHCCKRE	FYDHCCKSE	FYDHCCKTE	FYDHCCKPE
FYDHCCKME	FYDHCARE	FYDHCAD	FYDHCAD	FYDHCCKDE	FYDHCCKQE	FYDHCCKEE
FYDHCCKLE	FYIHCAAE	FYIHCKKE	FYIHCKME	FYIHCARE	FYIHCAD	FYIHCAD
FYIHCAEE	FYIHCALE	FYIHCAKE	FYIHCAME	FYIHCAD	FYIHCASE	FYIHCATE
FYIHCQEE	FYIHCQLE	FYIHCQKE	FYIHCQME	FYIHCQPE	FYIHCQSE	FYIHCQTE
FYIHCQAE	FYIHCQRE	FYIHCQDE	FYIHCQQE	FYIHCAD	FYIHCERE	FYIHCED
FYIHCCEEE	FYIHCCELE	FYIHCCEKE	FYIHCCEME	FYIHCCEPE	FYIHCCESE	FYIHCCETE
FYIHCCEQE	FYIHCIAE	FYIHCIRE	FYIHCIDE	FYIHCIQE	FYIHCIEE	FYIHCILE
FYIHCICE	FYIHCIME	FYIHCICE	FYIHCISE	FYIHCITE	FYIHCCLAE	FYIHCCLRE
FYIHCCLDE	FYIHCCLQE	FYIHCCLLE	FYIHCCLLE	FYIHCCLKE	FYIHCCLME	FYIHCCLPE
FYIHCCLSE	FYIHCCLTE	FYIHCCKAE	FYIHCCKRE	FYIHCCKSE	FYIHCCKTE	FYIHCCKPE
FYIHCCKDE	FYIHCCKQE	FYIHCCKEE	FYIHCCKLE	FYPHCAAE	FYPHCKKE	FYPHCKME
FYPHCAEE	FYPHCALE	FYPHCAKE	FYPHCAME	FYPHCAPE	FYPHCASE	FYPHCATE
FYPHCQEE	FYPHCQLE	FYPHCQKE	FYPHCQME	FYPHCQPE	FYPHCQSE	FYPHCQTE

WO 2014/094444

PCT/CN2013/081761

FYPHCQAE	FYPHCQRE	FYPHCQDE	FYPHCQQE	FYPHCEAE	FYPHCERE	FYPHCEDE
FYPHCEEE	FYPHCELE	FYPHCEKE	FYPHCEME	FYPHCEPE	FYPHCESE	FYPHCETE
FYPHCEQE	FYPHCIAE	FYPHCIRE	FYPHCIDE	FYPHCIQE	FYPHCIEE	FYPHCILE
FYPHCIKE	FYPHCIME	FYPHCIPE	FYPHCISE	FYPHCITE	FYPHCLAE	FYPHCLRE
FYPHCLDE	FYPHCLQE	FYPHCLLE	FYPHCLKE	FYPHCLME	FYPHCLPE	
FYPHCLSE	FYPHCLTE	FYPHCKAE	FYPHCKRE	FYPHCKSE	FYPHCKTE	FYPHCKPE
FYPHCKDE	FYPHCKQE	FYPHCKEE	FYPHCKLE	FYPHCARE	FYPHCADE	FYPHCAQE
FYTHCAAE	FYTHCKKE	FYTHCKME	FYTHCKDE	FYTHCKQE	FYTHCKEE	FYTHCKLE
FYTHCAEE	FYTHCALE	FYTHCAKE	FYTHCAME	FYTHCAPE	FYTHCASE	FYTHCATE
FYTHCQEE	FYTHCQLE	FYTHCQKE	FYTHCQME	FYTHCQPE	FYTHCQSE	FYTHCQTE
FYTHCQAE	FYTHCQRE	FYTHCQDE	FYTHCQQE	FYTHCEAE	FYTHCERE	FYTHCEDE
FYTHCEEE	FYTHCELE	FYTHCEKE	FYTHCEME	FYTHCEPE	FYTHCESE	FYTHCETE
FYTHCEQE	FYTHCIAE	FYTHCIRE	FYTHCIDE	FYTHCIQE	FYTHCIEE	FYTHCILE
FYTHCIKE	FYTHCIME	FYTHCIPE	FYTHCISE	FYTHCITE	FYTHCLAE	FYTHCLRE
FYTHCLDE	FYTHCLQE	FYTHCLLE	FYTHCLKE	FYTHCLME	FYTHCLPE	
FYTHCLSE	FYTHCLTE	FYTHCKAE	FYTHCKRE	FYTHCKSE	FYTHCKTE	FYTHCKPE
FYTHCARE	FYTHCADE	FYTHCAQE	FYWHCAAE	FYWHCKKE	FYWHCKME	FYWHCKDE
FYWHCAEE	FYWHCALE	FYWHCAKE	FYWHCAME	FYWHCAPE	FYWHCASE	FYWHCATE
FYWHCQEE	FYWHCQLE	FYWHCQKE	FYWHCQME	FYWHCQPE	FYWHCQSE	FYWHCQTE
FYWHCQAE	FYWHCQRE	FYWHCQDE	FYWHCQQE	FYWHCEAE	FYWHCERE	FYWHCEDE
FYWHCEEE	FYWHCELE	FYWHCEKE	FYWHCEME	FYWHCEPE	FYWHCESE	FYWHCETE
FYWHCEQE	FYWHCIAE	FYWHCIRE	FYWHCIDE	FYWHCIQE	FYWHCIEE	FYWHCILE
FYWHCIKE	FYWHCIME	FYWHCIPE	FYWHCISE	FYWHCITE	FYWHCLAE	FYWHCLRE
FYWHCLDE	FYWHCLQE	FYWHCLLE	FYWHCLKE	FYWHCLME	FYWHCLPE	
FYWHCLSE	FYWHCLTE	FYWHCKAE	FYWHCKRE	FYWHCKSE	FYWHCKTE	FYWHCKPE
FYWHCKQE	FYWHCKEE	FYWHCKLE	FYWHCARE	FYWHCADE	FYWHCAQE	FYYHCAAE
FYYHCAEE	FYYHCALE	FYYHCAKE	FYYHCAME	FYYHCAPE	FYYHCASE	FYYHCATE
FYYHCQEE	FYYHCQLE	FYYHCQKE	FYYHCQME	FYYHCQPE	FYYHCQSE	FYYHCQTE
FYYHCQAE	FYYHCQRE	FYYHCQDE	FYYHCQQE	FYYHCEAE	FYYHCERE	FYYHCEDE
FYYHCEEE	FYYHCELE	FYYHCEKE	FYYHCEME	FYYHCEPE	FYYHCESE	FYYHCETE
FYYHCEQE	FYYHCIAE	FYYHCIRE	FYYHCIDE	FYYHCIQE	FYYHCIEE	FYYHCILE
FYYHCIKE	FYYHCIME	FYYHCIPE	FYYHCISE	FYYHCITE	FYYHCLAE	FYYHCLRE
FYYHCLDE	FYYHCLQE	FYYHCLLE	FYYHCLKE	FYYHCLME	FYYHCLPE	
FYYHCLSE	FYYHCLTE	FYYHCKAE	FYYHCKRE	FYYHCKSE	FYYHCKTE	FYYHCKPE
FYYHCKKE	FYYHCKME	FYYHCKDE	FYYHCKQE	FYYHCKEE	FYYHCKLE	FYYHCARE
FYYHCADE	FYYHCAQE	FYCHNQNR	FYCHNER	FYCHNHR	FYCHNIR	FYCHNKR

WO 2014/094444

PCT/CN2013/081761

FYCHNMR	FYCHNFR	FYCHNFWR	FYCHDQR	FYCHDER	FYCHDHR	FYCHDIR
FYCHDKR	FYCHDMR	FYCHDFR	FYCHDFWR	FYCHQQR	FYCHQER	FYCHQHR
FYCHQKR	FYCHQMR	FYCHQFR	FYCHQFWR	FYCHQIR	FYCHEQR	FYCHEER
FYCHEKR	FYCHEMR	FYCHEFR	FYCHEFWR	FYCHEIR	FYCHEHR	FYCHHQR
FYCHHER	FYCHHKR	FYCHHMR	FYCHHFR	FYCHHFWR	FYCHHIR	FYCHHHR
FYCHIQR	FYCHIER	FYCHIKR	FYCHIMR	FYCHIFR	FYCHIFWR	FYCHIIR
FYCHIHHR	FYCHKQR	FYCHKER	FYCHKKR	FYCHKMR	FYCHKFR	FYCHKFWR
FYCHKIR	FYCHKHR	FYCHMQR	FYCHMER	FYCHMKR	FYCHMMR	FYCHMFR
FYCHMFWR	FYCHMIR	FYCHMHR	FYCHFQR	FYCHFER	FYCHFKR	FYCHFMR
FYCHFFR	FYCHFFWR	FYCHFIR	FYCHFHR	FYCHPQR	FYCHPER	FYCHPKR
FYCHPMR	FYCHPFR	FYCHPFWR	FYCHPIR	FYCHPHR	FYCHTQR	FYCHTER
FYCHTKR	FYCHTMR	FYCHTFR	FYCHTFWR	FYCHTIR	FYCHTHR	FYCHWQR
FYCHWER	FYCHWKR	FYCHWMR	FYCHWFR	FYCHWFWR	FYCHWIR	FYCHWHR
FYCHYQR	FYCHYER	FYCHYKR	FYCHYMR	FYCHYFR	FYCHYFWR	FYCHYIR
FYCHYHR	FYCHVQR	FYCHVER	FYCHVKR	FYCHVMR	FYCHVFR	FYCHVFWR
FYCHVIR	FYCHVHR	FYAHCQR	FYAHCER	FYAHCHR	FYAHCIR	FYAHCKR
FYAHCMR	FYAHCFR	FYAHCWR	FYRHCQR	FYRHCECER	FYRHCHR	FYRHCIIR
FYRHCKR	FYRHCMR	FYRHCFR	FYRHCWR	FYNHCQR	FYNHCER	FYNHCHR
FYNHCIR	FYNHCKR	FYNHCMR	FYNHCFR	FYNHCWR	FYDHCQR	FYDHCER
FYDHCCHR	FYDHCIR	FYDHCCKR	FYDHCMR	FYDHCFR	FYDHCWR	FYIHCQR
FYIHCER	FYIHCHR	FYIHCIR	FYIHCKR	FYIHCMR	FYIHCFR	FYIHCWR
FYPHCER	FYPHCHR	FYPHCIR	FYPHCKR	FYPHCMR	FYPHCFR	FYPHCWR
FYPHCQR	FYTHCQR	FYTHCER	FYTHCHR	FYTHCIR	FYTHCKR	FYTHCMR
FYTHCFR	FYTHCWR	FYWHCQR	FYWHCER	FYWHCHR	FYWHCIR	FYWHCKR
FYWHCMR	FYWHCFR	FYWHCWR	FYYHCQR	FYYHCER	FYYHCHR	FYYHCIR
FYYHCKR	FYYHCMR	FYYHCFR	FYYHCWR	FYACRNE	FYACRDE	FYACRQE
FYACREE	FYACRLE	FYACRME	FYACRSE	FYACRTE	FYACRWE	FYRCRNE
FYRCRDE	FYRCRQE	FYRCREE	FYRCRLE	FYRCRME	FYRCRSE	FYRCRTE
FYRCRWE	FYNCRNE	FYNCRDE	FYNCRQE	FYNCRREE	FYNCRLE	FYNCRME
FYNCRSE	FYNCRTE	FYNCRWE	FYDCRNE	FYDCRDE	FYDCRQE	FYDCREE
FYDCRLE	FYDCRME	FYDCRSE	FYDCRTE	FYDCRWE	FYQCRNE	FYQCRDE
FYQCRQE	FYQCRREE	FYQCRLE	FYQCRME	FYQCRSE	FYQCRTE	FYQCRWE
FYECRQE	FYECREE	FYECRLE	FYECRME	FYECRSE	FYECRTE	FYECRWE
FYECRNE	FYECRDE	FYGCRNE	FYGCRDE	FYGCRQE	FYGCRREE	FYGCRLE
FYGCRME	FYGCRSE	FYGCRTE	FYGCRWE	FYHCRNE	FYHCRDE	FYHCRQE
FYHCREE	FYHCRLE	FYHCRME	FYHCRSE	FYHCRTE	FYHCRWE	FYICRNE

WO 2014/094444

PCT/CN2013/081761

FYICRDE	FYICRQE	FYICREE	FYICRLE	FYICRME	FYICRSE	FYICRTE
FYICRWE	FYLCRNE	FYLCRDE	FYLCRQE	FYLCREE	FYLCRLE	FYLCRME
FYLCRSE	FYLCRTE	FYLCRWE	FYKCRNE	FYKCRDE	FYKCRQE	FYKCRREE
FYKCRLE	FYKCRME	FYKCRSE	FYKCRTE	FYKCRWE	FYMCRNE	FYMCRDE
FYMCRQE	FYMCREE	FYMCRLE	FYMCRME	FYMCRSE	FYMCRTE	FYMCRWE
FYFCRQE	FYFCREE	FYFCRLE	FYFCRME	FYFCRSE	FYFCRTE	FYFCRWE
FYFCRNE	FYFCRDE	FYPCRNE	FYPCRDE	FYPCRQE	FYPCREE	FYPCRLE
FYPCRME	FYPCRSE	FYPCRTE	FYPCRWE	FYWCRNE	FYWCRDE	FYWCRQE
FYWCRREE	FYWCRLE	FYWCRME	FYWCRSE	FYWCRTE	FYWCRWE	FYYCRNE
FYYCRDE	FYYCRQE	FYYCREE	FYYCRLE	FYYCRME	FYYCRSE	FYYCRTE
FYYCRWE	FYVCRNE	FYVCRDE	FYVCRQE	FYVCRREE	FYVCRLE	FYVCRME
FYVCRSE	FYVCRTE	FYVCRWE	YFNCRNE	YFNCRDE	YFNCRQE	YFNCRREE
YFNCRLE	YFNCRME	YFNCRSE	YFNCRTE	YFNCRWE	YFDCRNE	YFDCRDE
YFDCRQE	YFDCREE	YFDCRLE	YFDCRME	YFDCRSE	YFDCRTE	YFDCRWE
YFECRQE	YFECREE	YFECRLE	YFECRME	YFECRSE	YFECRTE	YFECRWE
YFECRNE	YFECRDE	YFHCRNE	YFHCRDE	YFHCRQE	YFHCRREE	YFHCRLE
YFHCRME	YFHCRSE	YFHCRTE	YFHCRWE	YFICRNE	YFICRDE	YFICRQE
YFICREE	YFICRLE	YFICRME	YFICRSE	YFICRTE	YFICRWE	YFLCRNE
YFLCRDE	YFLCRQE	YFLCREE	YFLCRLE	YFLCRME	YFLCRSE	YFLCRTE
YFLCRWE	YFKCRNE	YFKCRDE	YFKCRQE	YFKCREE	YFKCRLE	YFKCRME
YFKCRSE	YFKCRTE	YFKCRWE	YFMCRNE	YFMCRDE	YFMCRQE	YFMCRREE
YFMCRLE	YFMCRME	YFMCRSE	YFMCRTE	YFMCRWE	YFPCRNE	YFPCRDE
YFPCRQE	YFPCREE	YFPCRLE	YFPCRME	YFPCRSE	YFPCRTE	YFPCRWE
YFSCRQE	YFSCREE	YFSCRLE	YFSCRME	YFSCRSE	YFSCRTE	YFSCRWE
YFSCRNE	YFSCRDE	YFTCRNE	YFTCRDE	YFTCRQE	YFTCREE	YFTCRLE
YFTCRME	YFTCRSE	YFTCRTE	YFTCRWE	YFYCRNE	YFYCRDE	YFYCRQE
YFYCREE	YFYCRLE	YFYCRME	YFYCRSE	YFYCRTE	YFYCRWE	YFVCRNE
YFVCRDE	YFVCRQE	YFVCRREE	YFVCRLE	YFVCRME	YFVCRSE	YFVCRTE
YFVCRWE	HAYFCQR	HAYFCHR	HAYFCKR	HAYFCSR	HAYFCWR	HAYFCYR
HRYFCQR	HRYFCHR	HRYFCKR	HRYFCSR	HRYFCWR	HRYFCYR	HNYFCQR
HNYFCHR	HNYFCKR	HNYFCSR	HNYFCWR	HNYFCYR	HDYFCQR	HDYFCHR
HDYFCKR	HDYFCSR	HDYFCWR	HDYFCYR	HIYFCQR	HIYFCHR	HIYFCKR
HIYFCSR	HIYFCWR	HIYFCYR	HPYFCQR	HPYFCHR	HPYFCKR	HPYFCSR
HPYFCWR	HPYFCYR	HTYFCQR	HTYFCHR	HTYFCKR	HTYFCSR	HTYFCWR
HTYFCYR	HWYFCQR	HWYFCHR	HWYFCKR	HWYFCSR	HWYFCWR	HWYFCYR
HYYFCQR	HYYFCHR	HYYFCKR	HYYFCSR	HYYFCWR	HYYFCYR	HAYFCAK

WO 2014/094444

PCT/CN2013/081761

HAYFCRK	HAYFCQK	HAYFCHK	HAYFCKK	HAYFCFK	HAYFCWK	HAYFCYK
HAYFCVK	HRYFCAK	HRYFCRK	HRYFCQK	HRYFCHK	HRYFCKK	HRYFCFK
HRYFCWK	HRYFCYK	HRYFCVK	HNYFCAK	HNYFCRK	HNYFCQK	HNYFCHK
HNYFCKK	HNYFCFK	HNYFCWK	HNYFCYK	HNYFCVK	HDYFCAK	HDYFCRK
HDYFCQK	HDYFCHK	HDYFCKK	HDYFCFK	HDYFCWK	HDYFCYK	HDYFCVK
HIYFCQK	HIYFCHK	HIYFCKK	HIYFCFK	HIYFCWK	HIYFCYK	HIYFCVK
HIYFCAK	HIYFCRK	HPYFCAK	HPYFCRK	HPYFCQK	HPYFCHK	HPYFCKK
HPYFCFK	HPYFCWK	HPYFCYK	HPYFCVK	HTYFCAK	HTYFCRK	HTYFCQK
HTYFCHK	HTYFCKK	HTYFCFK	HTYFCWK	HTYFCYK	HTYFCVK	HWYFCAK
HWYFCRK	HWYFCQK	HWYFCHK	HWYFCKK	HWYFCFK	HWYFCWK	HWYFCYK
HWYFCVK	HYYFCAK	HYYFCRK	HYYFCQK	HYYFCHK	HYYFCKK	HYYFCFK
HYYFCWK	HYYFCYK	HYYFCVK				

1. 基于蛋白 A 亲和模型构建免疫球蛋白 G 的新型亲和配基多肽库，在多肽序列中间位置添加一个半胱氨酸 Cys 的方法，其特征是一共有以下 8 种序列：

FYCHXXXE、FYXHCXXE、FYCHXXR、FYXHCXR、FYXCRXE、YFXCRXE、HXYFCXR 和 HXYFCXK；

其中 X 代表除去半胱氨酸以外的 19 种氨基酸。

2. 如权利要求 1 所述的多肽库，其特征是构建依据为 6 个 SpA 关键残基：F132、Y133、H137、E143、R146 和 K154。

3. 以权利要求 1 和 2 所述的构建多肽库的设计方法的应用，其特征是：在肽库的范围内，利用氨基酸定位法构建候选多肽分子库。

4. 如权利要求 3 所述的设计方法的应用，其特征是根据方法应用得到如下的候选多肽分子库：

多肽构建模式 &条数	X 代表的氨基酸种类
FYCHXXXE 990	X_1: A, R, N, D, Q, E, H, I, L, K, M, T, W, Y, V (15) X_2: A, Q, E, I, L, K (6) X_3: A, R, D, Q, E, L, K, M, P, S, T (11)
FYXHCXXE 594	X_1: A, R, N, D, I, P, T, W, Y (9) X_2: A, Q, E, I, L, K (6) X_3: A, R, D, Q, E, L, K, M, P, S, T (11)
FYCHXXR 112	X_1: N, D, Q, E, H, I, K, M, F, P, T, W, Y, V (14) X_2: Q, E, H, I, K, M, F, W (8)
FYXHCXR 72	X_1: A, R, N, D, I, P, T, W, Y (9) X_2: Q, E, H, I, K, M, F, W (8)
FYXCRXE 153	X_1: A, R, N, D, Q, E, G, H, I, L, K, M, F, P, W, Y, V (17) X_2: N, D, Q, E, L, M, S, T, W (9)
YFXCRXE 117	X_1: N, D, E, H, I, L, K, M, P, S, T, Y, V (13) X_2: N, D, Q, E, L, M, S, T, W (9)
HXYFCXR 54	X_1: A, R, N, D, I, P, T, W, Y (9) X_2: Q, H, K, S, W, Y (6)
HXYFCXK 81	X_1: A, R, N, D, I, P, T, W, Y (9) X_2: A, R, Q, H, K, F, W, Y, V (9)

5. 如权利要求 4 所述的设计方法的应用，其特征是将候选多肽库包含的多肽分子进行分子对接筛选、均方根偏差比较以及分子动力学模拟复筛，获得有可能与 hIgG 具有较高亲和性的多肽配基：FYWHCLDE、FYFCRWE、FYIHCLPE、FYYHCKKE、FYCHWALE、FYCHWQDE、FYCHTIDE、FYRHCQRE、FYCHHKTE、FYCHLQKE、FYCHRKAE、FYCHNQDE、FYCHRQEE 和 FYNHCASE。

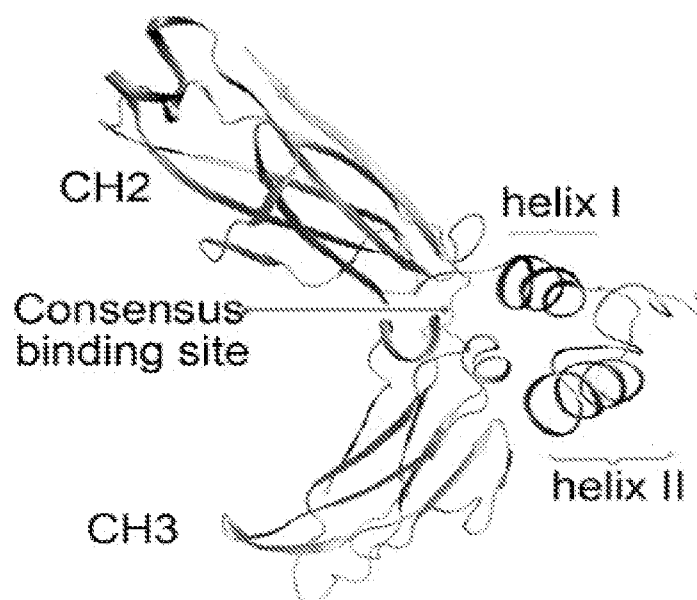


图 1

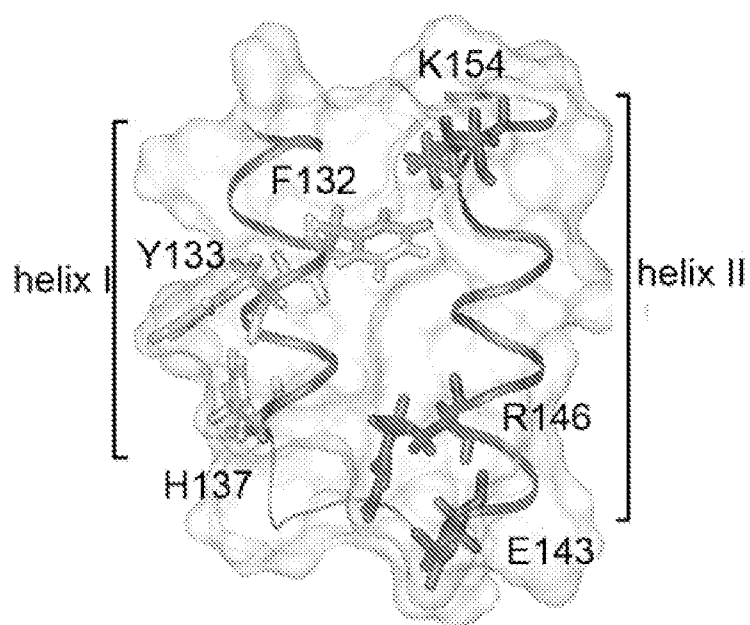


图 2



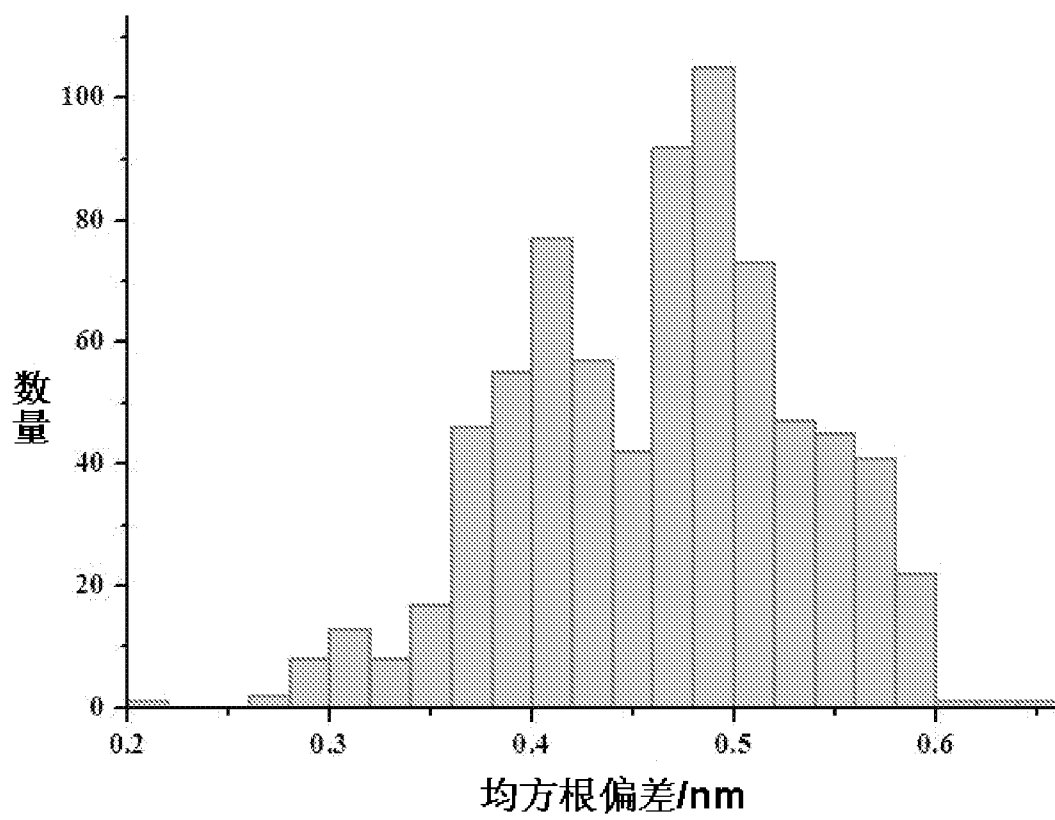


图 5

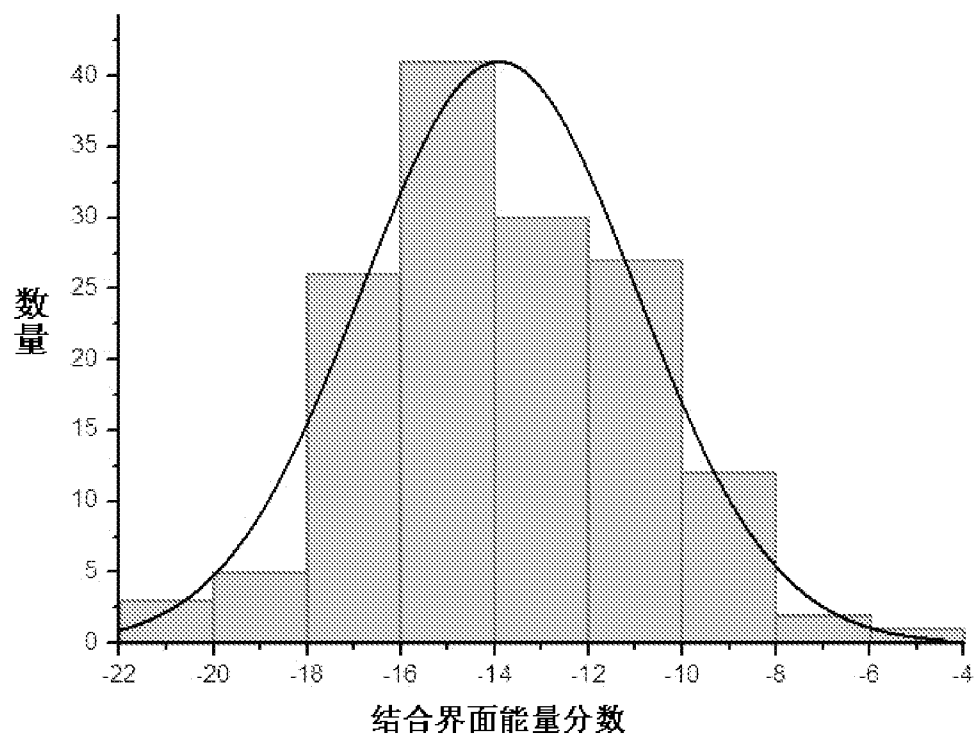


图 6

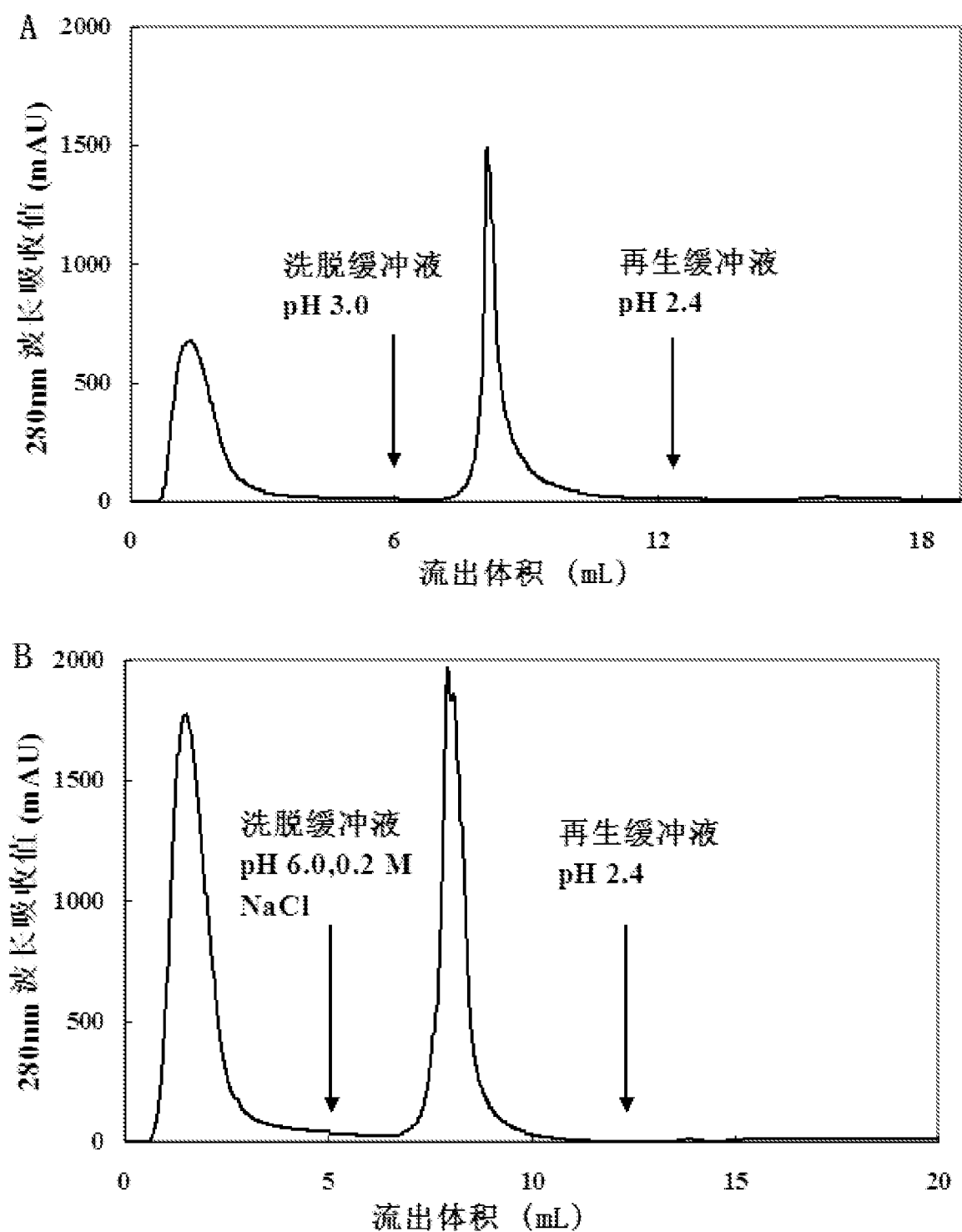


图 7

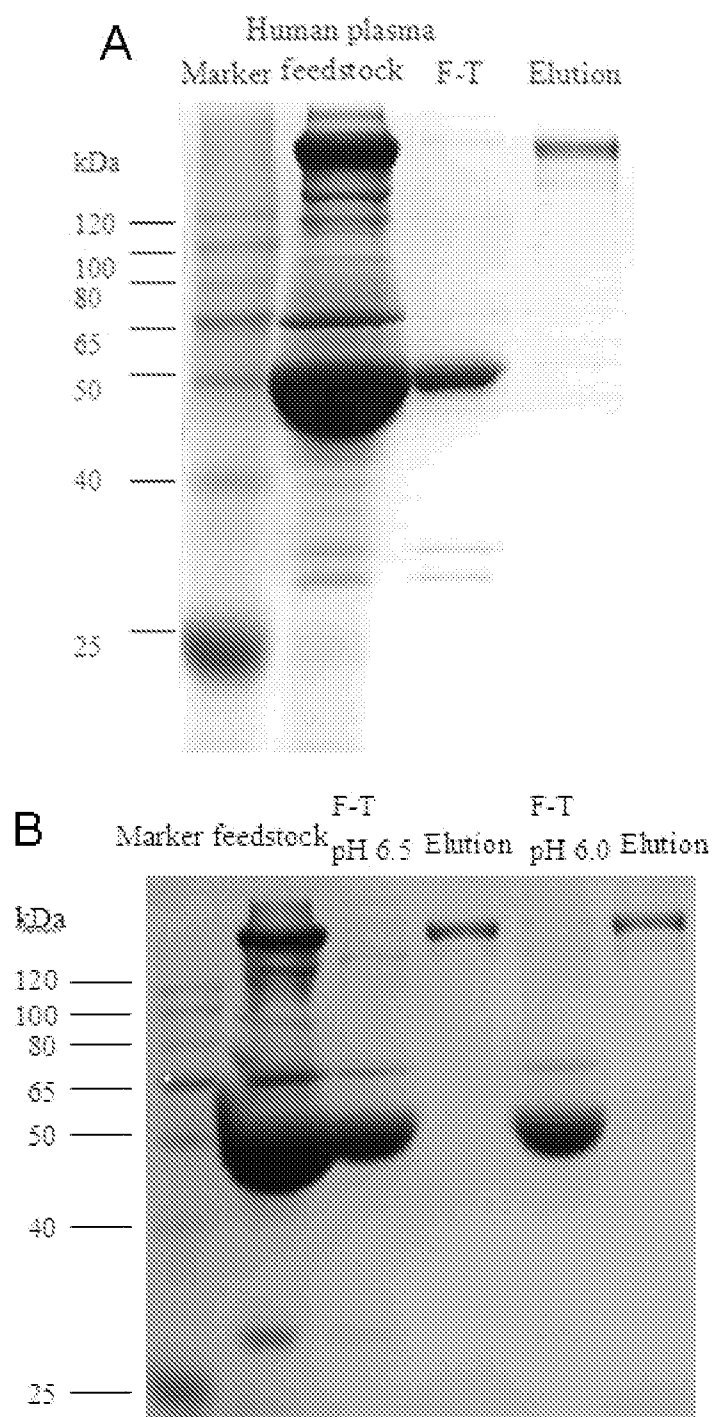


图 8

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2013/081761

A. CLASSIFICATION OF SUBJECT MATTER

See the extra sheet

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC: C40B, C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNABS, CNKI, CNTXT and search terms: antibody, immunoglobulin, IgG, protein A, SpA, ligand, peptide etc.;

DWPI, CPEA, SIPOABS, EPTXT, WOTXT, USTXT, JPTXT, ELSEVIER, EMBASE and search terms: antibod+, immunoglobulin, IgG, protein A, SpA, ligand?, peptide? etc.;

STN, GENBANK, EMBL, National Bio-sequence Database of Chinese Patent and search terms: polypeptide sequences in claim 1

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	EHRLICH, G.K. et al., Identification of model peptides as affinity ligands for the purification of humanized monoclonal antibodies by means of phage display, JOURNAL OF BIOCHEMICAL AND BIOPHYSICAL METHODS, 2001, vol. 49, pages 443-454	1-5
A	US 7408030 B2 (UNIV NORTH CAROLINA STAT et al.), 05 August 2008 (05.08.2008), the whole document	1-5

☒ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:	“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
“A” document defining the general state of the art which is not considered to be of particular relevance	“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
“E” earlier application or patent but published on or after the international filing date	“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	“&” document member of the same patent family
“O” document referring to an oral disclosure, use, exhibition or other means	
“P” document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 14 November 2013 (14.11.2013)	Date of mailing of the international search report 05 December 2013 (05.12.2013)
Name and mailing address of the ISA/CN: State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No.: (86-10) 62019451	Authorized officer LI, Chen Telephone No.: (86-10) 62411100

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2013/081761**C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	LI, Rongxiu et al., Design, synthesis, and application of a Protein A mimetic, NATURE BIOTECHNOLOGY, February 1998, vol. 16, pages 190-195, the whole document	1-5
A	FASSINA, G. et al., Protein A Mimetic Peptide Ligand for Affinity Purification of Antibodies, JOURNAL OF MOLECULAR RECOGNITION, 1996, vol. 9, pages 564-569, the whole document	1-5
P, X	CN 103014880 A (TIANJIN UNIVERSITY), 03 April 2013 (03.04.2013), claims 1-5	1-5

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2013/081761

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
US 7408030 B2	05.08.2008	US 2006153834 A1	13.07.2006
		US 2009005261 A1	01.01.2009
CN 103014880 A	03.04.2013	None	

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2013/081761

CONTINUATION: CLASSIFICATION OF SUBJECT MATTER

C40B 40/10 (2006.01) i

C40B 30/02 (2006.01) i

C40B 50/02 (2006.01) i

C07K 7/06 (2006.01) i

A. 主题的分类		
见附加页		
按照国际专利分类(IPC)或者同时按照国家分类和 IPC 两种分类		
B. 检索领域		
检索的最低限度文献(标明分类系统和分类号)		
IPC: C40B, C07K		
包含在检索领域中的除最低限度文献以外的检索文献		
在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))		
CNABS, CNKI, CNTXT 和检索词: 抗体, 免疫球蛋白, IgG, 蛋白 A, SpA, 配基, 配体, 肽等;		
DWPI, CPEA, SIPOABS, EPTXT, WOTXT, USTXT, JPTXT, ELSEVIER, EMBASE 和检索词: antibod+, immunoglobulin, IgG, protein A, SpA, ligand?, peptide?等;		
STN, GENBANK, EMBL, 中国专利生物序列检索系统和检索词: 权利要求 1 中的多肽序列		
C. 相关文件		
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
A	EHRLICH, G. K.等, Identification of model peptides as affinity ligands for the purification of humanized monoclonal antibodies by means of phage display, Journal of Biochemical and Biophysical Methods, 2001, 第 49 卷, 第 443-454 页 全文	1-5
A	US 7408030 B2 (UNIV NORTH CAROLINA STAT 等) 05.8 月 2008 (05.08.2008) 全文	1-5
☒ 其余文件在 C 栏的续页中列出。 ☒ 见同族专利附件。		
* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件 “T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件		
国际检索实际完成的日期 14.11 月 2013 (14.11.2013)		国际检索报告邮寄日期 05.12 月 2013 (05.12.2013)
ISA/CN 的名称和邮寄地址: 中华人民共和国国家知识产权局 中国北京市海淀区蓟门桥西土城路 6 号 100088 传真号: (86-10)62019451		受权官员 李晨 电话号码: (86-10) 62411100

C(续). 相关文件		
类 型	引用文件, 必要时, 指明相关段落	相关的权利要求
A	LI, Rongxiu 等, Design, synthesis, and application of a Protein A mimetic, Nature Biotechnology, 2 月 1998, 第 16 卷, 第 190-195 页 全文	1-5
A	FASSINA, Giorgio 等, Protein A Mimetic Peptide Ligand for Affinity Purification of Antibodies, Journal of Molecular Recognition, 1996, 第 9 卷, 第 564-569 页 全文	1-5
P, X	CN 103014880 A (天津大学) 03.4 月 2013 (03.04.2013) 权利要求 1-5	1-5

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2013/081761

检索报告中引用的 专利文件	公布日期	同族专利	公布日期
US 7408030 B2	05.08.2008	US 2006153834 A1 US 2009005261 A1	13.07.2006 01.01.2009
CN 103014880 A	03.04.2013	无	

续：主题的分类

C40B 40/10 (2006.01) i

C40B 30/02 (2006.01) i

C40B 50/02 (2006.01) i

C07K 7/06 (2006.01) i