

人 IgG 四亚类对噬菌体展示 Ig 结合蛋白单结构域随机组合文库的体外进化筛选

祁培培¹, 丁莹莹¹, 吴莉莉², 陈秋莉¹, 王锦红¹, 刘超¹, 廖文婷¹, 张婧²,
曹洁¹, 潘卫¹

1 中国人民解放军第二军医大学微生物学教研室, 上海 200433

2 安徽医科大学病理生理学教研室, 安徽 合肥 230032

祁培培, 丁莹莹, 吴莉莉, 等. 人 IgG 四亚类对噬菌体展示 Ig 结合蛋白单结构域随机组合文库的体外进化筛选. 生物工程学报, 2012, 28(9): 1093–1105.

Qi PP, Ding YY, Wu LL, et al. *In vitro* evolutionary selection of a combinatorial phage library displaying randomly-rearranged various binding domains of SpA and SpG with four human IgG subclasses. Chin J Biotech, 2012, 28(9): 1093–1105.

摘 要: 金黄色葡萄球菌蛋白 A (Staphylococcal protein A, SpA) 和链球菌蛋白 G (Streptococcal protein G, SpG) 是细菌产生的特异结合宿主抗体的细菌免疫球蛋白结合蛋白 (Immunoglobulin(Ig)-binding proteins, IBPs) 的代表分子。SpA 和 SpG 均包含由多个序列高度同源的结合结构域重复组成的抗体结合区, 各单结构域都具有完全的结合 IgG 的功能。为研究这些单结构域随机组合能否产生具有新结合特性的组合分子, 将 SpA 的 A、B、C、D、E 以及 SpG 的 B2、B3 共 7 个单结合结构域随机组合构建成噬菌体展示文库后, 应用人 IgG1、2、3、4 为诱饵分子对该文库进行 4 轮筛选, 获得了 SpA 天然分子中不存在的单结构域排列组合分子 D-C。在筛选过程中, 阴性对照噬菌体的逐渐减少、展示两个结构域以上的噬菌体比例不断增多, 尤其是 D-C 组合的选择性富集和其随机连接肽的严格筛选都显示了筛选的有效性和 D-C 组合的重要性。噬菌体 ELISA 进一步证实 D-C 与人 IgG 四亚类的结合能力远强于天然 SpA 分子。该研究应用分子进化技术首次获得了一种与人 IgG 四亚类具有结合优势的新型组合分子 D-C, 不仅可为 IgG 纯化、制备、检测等方面的应用提供新的候选分子, 还为细菌 IBP 结构功能的进一步研究提供新的手段。

关键词: 噬菌体, 免疫球蛋白结合蛋白, 进化筛选, IgG, 亚类

Received: March 7, 2012; **Accepted:** April 25, 2012

Supported by: National Natural Science Foundation of China (Nos. 30872405, 30872246, 30972632, 30972799), Shanghai Committee of Science and Technology (No. 08JC1405200), Chinese National Key Special Project for the Prevention and Control of Major Infectious Diseases (No. 2009ZX10004-105), Chinese National Key Special Project for Major New Drug Discovery (No. 2011ZX09506-001).

Corresponding author: Wei Pan. Tel: +86-21-81870989; E-mail: pwpanwei@yahoo.com.cn

Jie Cao. Tel: +86-21-81870993; E-mail: jessie_cao@yeah.net

国家自然科学基金 (Nos. 30872405, 30872246, 30972632, 30972799), 上海市基础研究重点项目 (No. 08JC1405200), 国家科技重大专项课题传染病专项 (No. 2009ZX10004-105), 国家科技重大专项课题重大新药专项 (No. 2011ZX09506-001) 资助。

In vitro evolutionary selection of a combinatorial phage library displaying randomly-rearranged various binding domains of SpA and SpG with four human IgG subclasses

Peipei Qi¹, Yingying Ding¹, Lili Wu², Qiuli Chen¹, Jinhong Wang¹, Chao Liu¹,
Wenting Liao¹, Jing Zhang², Jie Cao¹, and Wei Pan¹

¹ Department of Microbiology, Second Military Medical University, Shanghai 200433, China

² Department of Pathophysiology, Anhui Medical University, Hefei 230032, Anhui, China

Abstract: Protein A and protein G are two well-defined immunoglobulin (Ig)-binding proteins (IBPs), which show affinity for specific sites on Ig of mammalian hosts. Protein A and protein G contained several highly homologous IgG-binding domains which had been demonstrated to have function to bind to IgG. Whether combinations of Ig-binding domains of various IBPs could produce useful novel binding properties remains interesting. We constructed a combinatorial phage library which displayed randomly-rearranged A, B, C, D and E domains of protein A, B2 and B3 domains of protein G. Four rounds molecular evolution of this library directed by all four human IgG subclasses respectively generated a common arrangement of D-C respectively which didn't exist in SpA. The dynamic loss of control phages and increase of the phages displaying two or more binding domains, especially the selective enrichment of D-C and strict selection of its linking peptides demonstrated the efficient molecular evolutions and the significance of the selected D-C arrangement. The phage binding assays confirmed that D-C possessed a binding advantage with four human IgG subclasses compared to SpA. In this work, a novel combination of Ig-binding domains, D-C, was obtained and presented the novel Ig binding properties which provided a novel candidate molecule for the purification, production and detection of IgG antibodies and a new approach for the further study of structures and functions of IBPs.

Keywords: phage, immunoglobulin (Ig)-binding proteins, evolutionary selection, IgG subclass

细菌免疫球蛋白结合蛋白 (Immunoglobulin (Ig)-binding proteins, IBPs) 是一类由细菌产生的特异结合宿主抗体的细菌免疫球蛋白结合蛋白, 是细菌重要致病因子之一^[1-3]。研究最多的 IBPs 有金黄色葡萄球菌蛋白 A (Staphylococcal protein A, SpA)、链球菌 (C 和 G 群) 蛋白 G (Streptococcal protein G, SpG) 和部分大消化链球菌表面的蛋白 L (*Peptostreptococcus magnus* protein L, PpL)^[4-8]。研究表明, 细菌 IBPs 均包含由多个序列高度同源的结合结构域头尾相连重复组成的抗体结合区, 每个单结构域具有与全

分子完全相同的结合特性, 形成 Ig 结合功能的基本单位。SpA 含有 5 个高度同源的重复结构域, 自 N 端起分别为 E、D、A、B、C, 单个结构域约 58 个氨基酸残基大小, 形成 3 股反向平行排列 α -螺旋的空间结构, 与哺乳动物 IgG 抗体重链中的第 2 及第 3 恒定区 (CH2 γ , CH3 γ) 间的分界面结合, 该结合以螺旋 1 和螺旋 2 与 Fc 的疏水作用为主并通过两个分子之间的 4 对氢键得到进一步稳定^[9-11]。Protein A 还可以结合由人 VHIII 基因家族编码的 Fab 片段, 因此还能结合其他类型抗体^[12-14]。SpG 分子量为 65 kDa, 包含 3 个高

度同源的 Ig 结合结构域 B1、B2 和 B3^[15] (为区别于 SpA 中的 B, 以下统称 G1、G2、G3), 与 SpA 不同, 其单结构域由 4 个串联的 β 折叠和 1 个 α 螺旋组成^[16-17], 也结合于 IgGFc 的 CH2 γ 和 CH3 γ 间的分界面上, 与 SpA 的结合部位部分重叠。因其折叠结构含量很高, Protein G 结合 Fc 位点位于其螺旋和连结螺旋至 β -折叠 3 的环区中^[18]。此外, Protein G 可以结合 IgG 的 Fab 段, 其识别部位定位于 γ 链的第 1 个恒定区 (CH1 γ), 所以它仅结合 IgG, 并具有更高的亲和力^[19]。以 SpA 为代表的 IBP 已被广泛用于抗体的纯化、制备与鉴定、酶联免疫吸附法检测与诊断、免疫沉淀试验、临床体外免疫吸附治疗, 成为非常有价值的科研、诊断和治疗工具^[20-22]。

SpA 和 SpG 的这些单结构域的核苷酸序列高度同源但又存在细微差异, 这些差异是否对 IgG 的结合特性产生影响? SpA 和 SpG 为什么需要多个单结构域的重复? 这些问题尚没有明确答案。以金黄色葡萄球菌为例, 该细菌广泛存在于自然界中, 为多种动物体表正常菌群的组成菌种, 同时也可在特定条件下感染致病。然而, 人和其他动物 IgG 分子作为 SpA 和 SpG 作用的靶分子序列虽然同源, 但进化差异已明显存在, 并且在某些物种中存在多种亚类的 IgG, 人的 IgG 就存在四亚类。因此, 理论上不同的 IBP 单结构域的重复可产生具有不同结合特性组合以应对不同物种、不同亚类 IgG 分子的这种多样性。根据这一思路, 我们尝试应用基于噬菌体展示分子进化方法对该问题进行探索性研究。先前, 本实验室应用自主建立的基于噬菌体展示技术的分子进化平台, 对噬菌体展示 Protein A、Protein G、Protein L 等 IBPs 的单结构域随机组合文库并

进行体外进化筛选获得了 A-G2^[23]、A-L^[23] (Protein L 的 B3 结构域)、LD3^[24]及 LD5^[24-25]等多种新型进化免疫球蛋白结合分子 (Novel evolved Ig-binding molecules, NEIBMs), 其中 LD3 和 LD5 为天然 IBPs 中不存在的单结合结构域组合, 这种组合被证实产生了天然 IBPs 所没有的抗体结合特性, 即对 Ig Fab κ 轻链和 VHIII 重链的双结合特性, 该结合特性被证实大大提高了与 IgM 的结合力, 并被成功地应用于 HCV 的 IgM 检测, 充分显示了分子进化方法用于该研究的可行性。在本研究中, 我们构建了以 SpA 的 A、B、C、D、E 5 个结构域和 SpG 的 G2、G3 结构域 7 个单结构域随机组合的噬菌体展示文库, 并应用人 IgG1、IgG2、IgG3 和 IgG4 四亚类分子进行体外进化筛选, 首次获得了一种与人 IgG 四亚类均具有结合优势的新型组合分子 D-C。

1 材料与方法

1.1 材料

1.1.1 质粒载体和菌株

噬菌粒载体 pCANTAB5S 由第二军医大学微生物学教研室保存^[26]; pCANTAB5S-SpA 噬菌粒、pMD-18T-G2 和 pMD-18T-G3 质粒分别克隆有 SpA 的 EDABC 五结构域基因 (GenBank Accession No. P02976)、SpG 的 B2 和 B3 结构域的基因 (GenBank Accession No. P06654) 为本室构建保存; *E. coli* TG1 和辅助噬菌体 M13K07 为本室保存。

1.1.2 试剂

限制性内切酶 *Xba* I、碱性磷酸酶 (CIP) 购自宝生物工程 (上海) 公司; 高效连接液购自 TOYOBO 公司; 质粒抽提试剂盒及 DNA *Taq* 酶

购自上海申能博彩生物公司；质粒小提试剂盒和琼脂糖凝胶 DNA 回收试剂盒购自天根生化科技(北京)有限公司；氨苄青霉素 (Amp) 与卡那霉素 (Kana) 均购自华美生物公司；人 IgG1、2、3、4 四亚类分子购自 Sigma 公司；鼠抗噬菌体酶联单抗 mouse anti-M13mAb-HRP 购自 Amersham Pharmacia 公司。

1.2 方法

1.2.1 引物设计

根据本室保存的碱基序列设计 12 条引物，序列见表 1。应用 12 条引物进行 PCR 扩增获得下游带有 3 个氨基酸大小的随机连接肽编码序列 (NNSNNSNNS) 的 A、B、C、D、E、G2、G3 七个结构域的 DNA 片段，且片段两端均引入 *Xba* I

酶切位点 (TCTAGA) 和 53 bp 保护序列。此外，用于 PCR 扩增检测噬菌粒 pCANTAB5S 中克隆片段及测序的上游引物 pCANTAB5S-1 序列为：5'-CAACGTGAAAAAATTATTATTCGC-3'，而下游引物 pCANTAB5S-6 序列为：5'-GTAAATGAA TTTTCTGTATGAGG-3'。以上引物均委托上海生物工程科技服务有限公司合成。

1.2.2 SpA 和 SpG 七个单结构域随机组合噬菌体展示文库的构建

以 pCANTAB5S-SpA 噬菌粒为模板，分别以 uA 和 dAD、uB 和 dB、uC 和 dC、uD 和 dAD、uE 和 dE 为上下游引物，PCR 扩增 SpA 的 A、B、C、D、E 五结构域片段；以 uG2 和 dG2 为上下游引物，分别以 pMD-18T-G2、pMD-18T-G3 质

表 1 七个免疫球蛋白结合结构域片段的扩增引物

Table 1 Primers for amplification of DNA fragments encoding seven Ig-binding domains of SpA and SpG

Primer name	Primer sequence (5'-3')
uA	TCGTCAGACGCCGTACCTGCTCTAGA*GCTGACAACAATTCAAC
uD	TCGTCAGACGCCGTACCTGCTCTAGA*GCTGATGCGCAACAAAAT
dAD	TCGTCAGACGCCGTACCTGCTCTAGA* <u>SNNSNNSNN</u> **TTTCGGTGCTTGAGATTC
uB	TCGTCAGACGCCGTACCTGCTCTAGA*GCGGATAACAAATTCAAC
dB	TCGTCAGACGCCGTACCTGCTCTAGA* <u>SNNSNNSNN</u> **TTTTGGTGCTTGATGCATC
uC	TCGTCAGACGCCGTACCTGCTCTAGA*GCTGACAACAATTCAAC
dC	TCGTCAGACGCCGTACCTGCTCTAGA* <u>SNNSNNSNN</u> **TTTTGGTGCTTGAGCATC
uE	TCGTCAGACGCCGTACCTGCTCTAGA*GCTCAACAAAATGCTT
dE	TCGTCAGACGCCGTACCTGCTCTAGA* <u>SNNSNNSNN</u> **TTTTGGAGCTTGAGAGTC
uG2	TCGTCAGACGCCGTACCTGCTCTAGA*ACCTACAAACTGGTTATC
dG2	TCGTCAGACGCCGTACCTGCTCTAGA* <u>SNNSNNSNN</u> **TTCGGTAACGGTGAAGGT
Sec	ATACCTAGCCATTACCCTAGCCACCTAGTCCGTCGTCAGACGCCGTACCTGCTCTAGA*

Note: * *Xba* I restriction sites are in italic; ** nucleotide sequences of random linking peptide are underlined (N = A/T/C/G, S = G/C).

粒为模板, PCR 扩增 SpG 的 G2、G3 结构域片段; 为延长保护碱基以提高酶切效率, 再以 Sec 为上下游引物, 分别以上述 7 个单结构域片段的 PCR 产物为模板, PCR 扩增得到在两端分别引入约 50 bp 个保护碱基的 7 个单结构域片段。将 7 个片段的 PCR 产物纯化后用 *Xba* I 酶消化, 与同样酶切并用 CIP 酶去磷酸化处理的 pCANTAB5S 连接后转化 *E. coli* TG1 超级感受态细菌 (其制备见文献[27])。分别取 0.1、1、10 μ L 转化后的细菌涂布 LB (含 100 mg/L Amp) 平皿, 并计算库容。随机挑取 23 个转化子对文库中片段插入情况进行 PCR 鉴定, 将插入多个结构域片段的转化子委托上海杰李生物技术有限公司 (以下简称上海杰李) 进行序列测定。剩余菌液继续加入 10 mL 2 \times YT (Amp) 培养基, 37 $^{\circ}$ C、250 r/min 振荡培养过夜。加 500 μ L 1.3×10^{12} TU/mL 的 M13KO7 辅助噬菌体, 37 $^{\circ}$ C、250 r/min 振荡培养 1 h, 加入 Kana 后, 37 $^{\circ}$ C, 250 r/min 振荡培养 8 h。1 000 \times g 离心 10 min, 上清经 0.22 μ m 滤膜过滤, 即为展示 SpA 和 SpG 单结构域随机组合噬菌体文库, 滴度测定和无菌实验见文献[28]和[29]。

1.2.3 人 IgG1、IgG2、IgG3 和 IgG4 对噬菌体文库的体外分子进化筛选

应用人 IgG 四亚类分别对上述构建的噬菌体文库进行 4 轮体外分子进化筛选, 详细步骤参见文献[23]。根据本实验室工作基础, 筛选过程中文库空噬菌体所占比例的减少和展示多结构域 (≥ 2 个结构域) 噬菌体的增加情况是监测筛选是否有效的重要指标, 但因原代噬菌体文库中零插入噬菌体比例很少, 我们将原代库噬菌体与

零插入噬菌体在相同滴度下等体积混匀作为筛选的初始文库 (经 PCR 鉴定, 空噬菌体占 74%, 展示一个结构域的噬菌体占 9%, 展示多结构的噬菌体占 17%)。此外, 为深入了解整个进化过程中噬菌体克隆展示序列的变化情况, 本研究从人 IgG1、2、3、4 四组的各代平板上的阳性单克隆中都分别随机挑选 10 个插入多结构域片段的单克隆委托上海杰李进行序列测定; 若有插入多结构域片段的单克隆总数不足 10 个, 则将实际数目的单克隆进行测序。

1.2.4 ELISA 鉴定优势组合分子噬菌体的结合活性

根据序列分析结果, 将筛选获得的优势组合 D-C 的单克隆, 划 LB 平板 (Amp), 37 $^{\circ}$ C 过夜培养。为保证结果的真实可靠, 我们随机挑取 5 个菌落分别接种于 5 管 2 \times YT 培养基 (Amp) 中, 经 M13KO7 拯救, 制备成 5 管具有相同展示物却分别来自不同单克隆的噬菌体, 测定滴度后用作平行组实验。相同方法制备 pCANTAB5S-SpA 噬菌体和 pCANTAB5S 噬菌体, 分别作为本实验阳性、阴性对照。取 100 μ L 噬菌体加入用人 IgG1 包被的 ELISA 板条中, 37 $^{\circ}$ C、2 h 后用 PBST 洗 5 次, 加入鼠抗噬菌体酶联单抗结合, 37 $^{\circ}$ C、2 h 后 PBST 洗 10 次, TMB 显色后读取 OD_{450} 数值, 检测单克隆噬菌体与人 IgG1 的结合活性。同样的方法检测与人 IgG2、3 和 4 结合活性。

1.2.5 统计学处理

采用 SPSS 17.0 统计软件对 ELISA 检测数据进行统计学处理, 计量资料多组间比较采用单因素方差分析, 两两比较采用 *q* 检验, 以 $P<0.05$ 为有统计学意义。

2 结果

2.1 噬菌体展示 SpA 和 SpG 单结构域随机组合文库的构建

经 PCR 扩增得到 SpA 的 A、B、C、D、E 和 SpG 的 G2、G3 共 7 个 IBP 单结构域的 DNA 片段，纯化后经 *Xba* I 酶消化，得到两端带有 *Xba* I 酶切位点的 7 个单结构域片段 (图 1)。

将上述 7 个 IBPs 单结构域片段与经同样酶切并 CIP 处理后的噬菌粒 pCANTAB5S 连接，连接产物转化超级感受态细胞 *E. coli* TG1，构建的 SpA 和 SpG 单结构域随机组合文库的库容量即转化数为 8.2×10^6 CFU，原代噬菌体文库的滴度为 1.3×10^{12} TU/mL。

从转化平板上随机挑取 23 个转化子，对原代文库结构域片段插入情况的 PCR 检测结果为：零

插入的转化子即 5S 载体自连的为 3 个，百分比为 13%；插入单个结构域片段的为 14 个，占 61%；插入 2 个的为 4 个，占 17%；插入 3 个的为 2 个，占 9% (图 2)。原代文库的结构域片段插入率达 87%。此外 PCR 鉴定的 6 个插入多个结构域片段的阳性克隆序列测定结果显示：在插入到噬菌粒上共 14 单结构域片段中，A:B:C:D:E:G2:G3 个数为 2:2:1:2:0:4:3，虽然 E 未测出，但根据本实验室研究基础，原代库中插入的各片段具有随机性；在 14 单结构域片段中，插入片段的正反向之比为 8:6，与 1:1 相近，总插入片段的正反向也具有随机性；在 14 条随机连接肽序列 (NNS)₃ 中，第 1 位 N 上 A:T:C:G 为 15:8:13:6，第 2 位 N 上 A:T:C:G 为 7:9:14:12，所有 S 位上的 G:C 为 26:16，均具有随机性 (表 2)。因此我们构建的文库符合体外进化筛选要求。

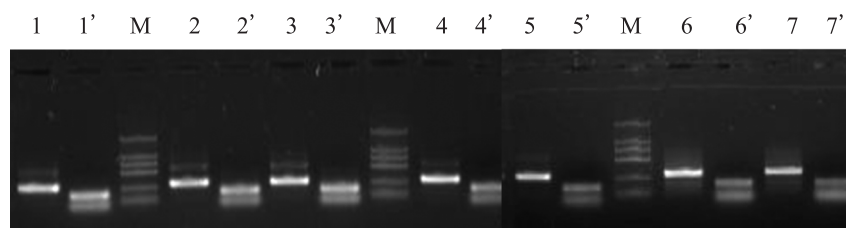


图 1 七个 SpA、SpG Ig 结合单结构域 DNA 片段的 *Xba* I 酶消化电泳图

Fig. 1 Electrophoretic profile of *Xba* I digestion of seven DNA fragments of Ig-binding mono-domains of SpA and SpG. M: DL2000 DNA marker (2 000 bp, 1 500 bp, 1 000 bp, 750 bp, 500 bp, 250 bp, 100 bp); 1-7: DNA fragments of A, B, C, D, E, G2, G3 respectively; 1', 2', 3', 4', 5', 6', 7': A, B, C, D, E, G2, G3 fragments after *Xba* I digestion respectively.



图 2 SpA、SpG 单结构域随机组合原代文库插入片段大小的 PCR 鉴定

Fig. 2 Identification of the size of inserted fragments in the original library displaying randomly-rearranged various binding domains of SpA and SpG by PCR. M: DL2000 DNA marker (2 000 bp, 1 500 bp, 1 000 bp, 750 bp, 500 bp, 250 bp, 100 bp); 1-23: the number of 23 single clones; C: the phagemid pCANTAB5S for positive control; B: the negative control.

2.2 人 IgG1、IgG2、IgG3 和 IgG4 对噬菌体展示 SpA 和 SpG 单结构域随机组合文库的体外进化筛选

根据本实验室对同类文库的进化筛选经验^[23-24]，因 IgG 具有 2 条 Fc 链，可以同时结合 2 个单结构域，文库中展示两结构域噬菌体所占比例的增加和空噬菌体的减少情况是检测筛选

是否有效的重要指标。从图 3 可知，在所有人 IgG 四亚类进化筛选中，展示两结构域片段克隆所占的比例持续稳定地增加，最后文库中基本全部进化为两结构域组合分子；相反地，零插入克隆和插入单个结构域片段克隆所占比例不断降低，最后基本被淘汰消失，这种变化充分显示了筛选的有效性。

表 2 SpA、SpG 单结构域随机组合原代文库中噬菌体克隆插入片段的序列分析
Table 2 Sequence analyses of inserted fragments on phage clones in the original library displaying randomly-rearranged various binding domains of SpA and SpG

No.	Combinatorial form of single domains	Nucleotide sequence of random linking peptide
1	B _{9n} -G2 ^R _{9n} -G2 _{9n}	GCG CTG TTC-ACG CGC GCG-AGC ACG TGG
2	G3 _{9n} -D ^R _{9n} -A ^R _{9n}	CTG ATG CTC-CTC ACC AGG-AGG CCG TGG
3	A ^R _{9n} -G2 _{9n}	CCC TCG CTG-CGC AAG TCG
4	D _{9n} -G3 _{9n}	CGC GAG TCC-AAC AGG GAG
5	B _{9n} -G2 ^R _{9n}	TGG ACG AGC-AAG ACG GAC
6	G3 ^R _{9n} -C _{9n}	TAC ACG CTC-CCC GTC CGG

R: reverse sequence of original sequence; 9n: the sequence of random linking peptides composed of nine nucleotides.

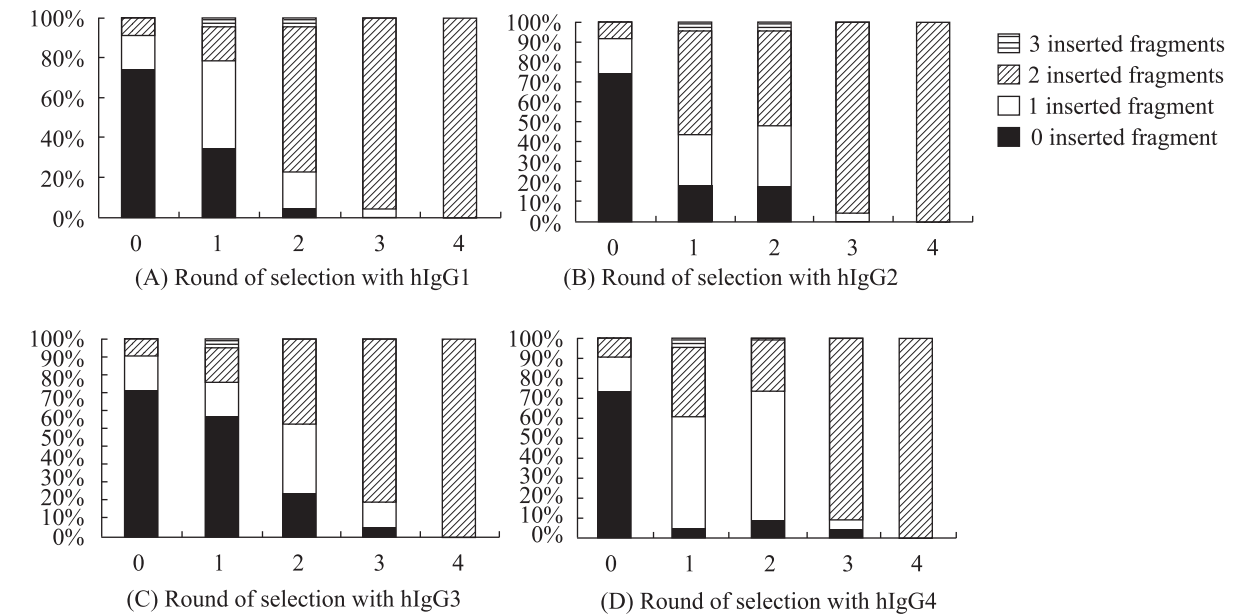


图 3 经人 IgG 四亚类筛选后各代噬菌体文库中 23 个单克隆噬菌体插入片段的组成比例变化
Fig. 3 Proportion of phage clones with different sizes of inserted fragments in 23 phage clones after each round of selection with four human IgG subclasses respectively (A-D).

为深入了解在整个进化过程中两结构域的序列变化情况,本研究从人 IgG 四亚类筛选的各代文库中都分别随机挑选 10 个插入多结构域片段的单克隆进行测序,而在 IgG1 组的 F1 代、IgG3 组的 F1 和 F2 代、IgG4 组的 F1 和 F2 代中因插入多结构域片段的单克隆数目较少,则按实际数目测序。测序结果显示:1) 第 1 轮筛选中四亚类组共测得 25 种多结构域组合分子,其中能正确展示的均为正向插入的多结构域组合为 17 种,

包括 D_{9n2}-C_{9n2} (人 IgG2 组的第 1 轮筛选)。2) 第 2 轮筛选中四亚类组均出现一种新的组合分子 D_{9n1}-C_{9n1} (以下简称 D-C),然而所占的比例在四亚类中却差别很大, IgG1 最高,占 50%, IgG2 和 IgG4 次之,均为 30%,而 IgG3 最低,仅为 10%。3) 人 IgG 四亚类组在第 3、4 轮筛选最终都得到相同的优势组合分子, D-C。第 4 轮筛选后的文库中随机挑取 10 个克隆除 IgG3 组出现了 30%的 A_{9n}-G_{39n},其余均为 D-C,见表 3。

表 3 人 IgG1、IgG2、IgG3 和 IgG4 诱导的进化筛选各轮获得展示片段的测序结果

Table 3 Sequence analyses of inserted fragments on phage clones in the four rounds of post-selection libraries with four human IgG subclasses

Phage libraries	1st round	2nd round	3th round	4th round
hIgG1	D9n-D9n-A9n; C9n-C9n; G3 ^R 9n-A9n	D _{9n1} -C _{9n1} (5*); C9n-D9n; D9n-G39n; B ^R 9n-C ^R 9n-E ^R 9n; A9n-G39n; A9n-C9n	D _{9n1} -C _{9n1} (9); A ^R 9n-E9n	D _{9n1} -C _{9n1} (10)
hIgG2	A9n-G39n; E9n-B9n-C9n; A9n-G39n; D9n2-C9n2; D9n-B9n; A9n-G39n; G3 ^R 9n-B9n; A9n1-B9n; A9n2-B9n2; B ^R 9n-A ^R 9n	D _{9n1} -C _{9n1} (3); D9n-C ^R 9n; D9n-G29n; E9n-C ^R 9n; G2 ^R 9n-C ^R 9n; B9n-D9n; B9n-C9n-G39n-A9n; E ^R 9n-C ^R 9n	D _{9n1} -C _{9n1} (10)	D _{9n1} -C _{9n1} (10)
hIgG3	C9n-A9n; B9n-G39n; D9n-B9n; G29n-B ^R 9n; E9n-G39n	D _{9n1} -C _{9n1} ; B ^R 9n-G29n; B ^R 9n-C ^R 9n; G39n-C9n; G29n-B9n; G3 ^R 9n-E ^R 9n; C ^R 9n-E ^R 9n; D9n-B ^R 9n; D ^R 9n-B9n	D _{9n1} -C _{9n1} (4); A9n-G39n(5); A ^R 9n-E9n	D _{9n1} -C _{9n1} (7); A9n-G39n(3)
hIgG4	D9n-G39n; A9n-G39n; A9n2-B9n2; A9n-C9n; D9n-E9n; A9n3-B9n3; G3 ^R 9n-A9n-A9n	D _{9n1} -C _{9n1} (3); D9n-E9n A9n-C9n; D9n-B9n	D _{9n1} -C _{9n1} (10)	D _{9n1} -C _{9n1} (10)

*: the number of phage clones with the same inserted fragment; R: reverse sequence of original sequence; 9n: the sequence of random linking peptides composed of nine nucleotides; D_{9n1}-C_{9n1}: DGCGCACACG-CATGTTGTAG; D_{9n2}-C_{9n2}: D_{TCCCCCGTC}-C_{GACAAGAAC}; A_{9n1}-B_{9n1}: A_{ACCCCCATC}-B_{AACTAGGAG}; A_{9n2}-B_{9n2}: A_{CGCCACAAC}-B_{GACACGAAG}; A_{9n3}-B_{9n3}: A_{TACTACTCG}-B_{TAGATGACG}.

2.3 ELISA 检测 D-C 噬菌体与人 IgG1、IgG2、IgG3 和 IgG4 的结合优势

将 D-C、pCANTAB5S-SpA 和 pCANTAB5S 三组分别制备 5 个单克隆噬菌体,并将其滴度均调整为约 1.0×10^{12} TU/mL 后在相同条件下进行 ELSIA 比较 3 组噬菌体分别对人 IgG 四亚类的结合活性。ELISA 结果显示: D-C 噬菌体和 SpA 噬菌体对人 IgG 四亚类的结合活性均远强于阴性对照组; D-C 噬菌体对人 IgG 四亚类的结合活性又均强于 SpA 噬菌体,两组的差异有统计学意义(图 4),且在 人 IgG 四亚类中 D-C 噬菌体对 IgG1 的结合活性最强, IgG3 组最弱。

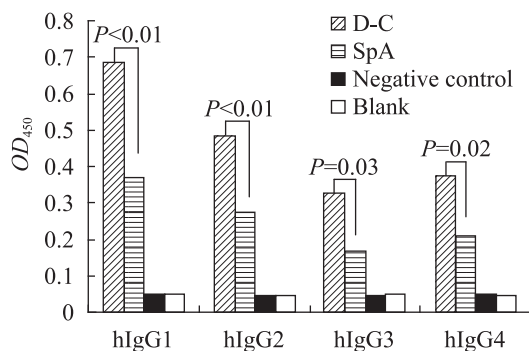


图 4 ELISA 比较 D-C 噬菌体和 SpA 噬菌体对人 IgG1、IgG2、IgG3 和 IgG4 的结合活性

Fig. 4 Comparison of the binding activities of D-C and SpA phage clones with four human IgG subclasses respectively by ELISA.

3 讨论

SpA 和 SpG 是细菌表面与宿主抗体特异结合的蛋白,是极具价值的诊断、治疗及抗体制备工具。它们的各单结构域分别以天然的形式与哺乳动物 IgG 的 Fc 结合。为探索 SpA 和 SpG 的各单结构域重组后能否获得对人 IgG 各亚类优势结合的 NEIBMs,本研究构建展示 SpA、SpG

单结构域随机组合的噬菌体展示文库,并应用人 IgG 四亚类分子进行体外进化筛选。我们预期能获得与各人 IgG 亚类特异结合的不同的 NEIBMs,然而,本研究的结果却与预期相反,4 种人 IgG 亚类所诱导的分子进化均得到了同一种优势分子, D-C 组合。

我们所构建的文库库容为 8.2×10^6 CFU, 原始噬菌体文库的滴度为 1.3×10^{12} TU/mL, 对原始文库中随机挑取阳性单克隆的 PCR 检测显示已插入结构域片段的阳性克隆总计达 87%。插入多结构域片段的阳性克隆序列测定结果显示原始文库中的各片段插入情况、片段正反向比例、随机连接肽序列 (NNS)₃ 均具有随机性。理论上,文库中两结构域组合的方式应为 $7 \times 2 \times 7 \times 2 = 196$, 各单结构域随机连接肽理论上的组合方式应为 $(2 \times 4 \times 4)^3 \times (2 \times 4 \times 4)^3 = 1048576$ 。可见,该文库虽然不能将所有的组合方式完全包含,但已远远超出了对 2 个及 2 个以上结构域组合充分展示的要求。这些结果表明我们所构建的文库在库容、随机性和多样性方面能够满足筛选要求。

根据本实验室对同类文库的 IgG 及 IgG Fc 筛选经验^[23-24],因 IgG 具有两条 Fc 链,可同时结合两个单结构域,文库中展示两结构域噬菌体比例的增加情况是监测筛选是否有效的最重要的指标。在人四 IgG 亚类的进化筛选中,展示两结构域噬菌体的比例随着筛选的进行不断增多,筛选到第 3、4 代时,文库中基本已全部进化为两结构域组合分子,这种变化充分显示筛选的有效性。

该次筛选中我们还特别检测了两结构域组合分子组成和序列在进化过程中的变化情况。出

乎我们意料的是,人 IgG1、IgG2、IgG3、IgG4 筛选最终都得到具有相同的优势组合分子 D-C, 并且其随机连接肽序列也完全相同。该优势组合分子 D-C 均在第 2 轮筛选中出现,然而所占的比例在四亚类中却很大差别, IgG1 最高, 占 50%, IgG2 和 IgG4 次之, 均为 30%, 而 IgG3 最低, 仅为 10%。有意思的是, 这样的差别与 phage ELISA 所反映的各亚类 IgG 与 D-C 的结合能力完全相符。人 IgG1、IgG2、IgG4 诱导的进化 3、4 轮即已完全, D-C 成为唯一的克隆, 而 IgG3 所诱导的筛选到第 4 轮筛选时出现了 D-C 和 A_{9n}-G3_{9n} 两种优势分子, 其中 D-C 占 70%, 为相对优势分子。另外, 原代文库和四亚类的第 1 轮筛选后文库两个结构域组合中均未发现 D-C, 表明其在原代文库中未形成优势。同时四亚类的第 1、2 轮筛选我们发现至少存在 E-G3、D-G2、C-C、A-B 等 28 种不同的正确展示的两个结构域组合分子, 显示了 D-C 的超强结合能力。的确, 我们采用严格对照的噬菌体 ELISA 实验证明了 D-C 噬菌体与人 IgG 四亚类的结合均远远强于 SpA 噬菌体。因此两结构域组合分子序列的动态变化情况很好地反映了筛选的有效性和 D-C 的意义。另外, 在人 IgG2 组第 1 轮筛选中还出现了具有不同于 D-C 的随机连接肽序列的 D_{9n2}-C_{9n2}, 然而在后续测序中再未出现过, 最终 D-C 成为四亚类筛选的优势组合分子, 这说明随机连接肽的序列对 D-C 的结合功能有影响, 的确, 四亚类 IgG 最终筛选到的 D-C 随机连接肽序列完全一致, 被严格选择。

上述结果均说明和证明了 D-C 与 IgG 的结合优势, 然而, 为什么 D-C 组合具有结合优势目前尚无明确解释。SpA 中五结构域的天然组合顺序

是 E-D-A-B-C, D-C 组合不存在其中。我们应用 5 种 SpA 的单结构域和 2 种 SpG 的单结构域共 7 种 IgG Fc 单结合结构域构建噬菌体随机组合文库, 因此理论上能产生同时正确结合两条 IgG Fc 的双结构域正确组合为 49 种 (7×7), 远远超过天然 SpA 和 SpG 所能产生出来的 6 种两个结构域组合 (SpG 分子中 3 个结合结构域的组合顺序是 G1-G2-G3, 可产生 2 种组合)。此外, 文库中两个结构域组合是通过 3 个氨基酸的随机连接肽连接, 这种方式比天然分子的直接连接对结合 IgG Fc 双链更具有优势 (随机连接肽的序列对两个结构域组合的结合功能有影响, 见上段讨论)。因此, 产生出结合能力超过天然两个结构域组合的分子完全可能。金黄色葡萄球菌感染多种宿主时, SpA 能与多物种、多亚类 IgG 结合, 其 5 个单结构域的串联可大大扩展其结合谱, 然而, 由于基因容量的限制, 使其无法容纳更多的串联体而限制其产生具有最强结合能力的组合。因此, 我们所得到的在 SpA 天然分子中不存在的具有强结合能力的 D-C 也具合理性。目前, 我们正在进行原核表达 D-C 融合蛋白, 以期在分子水平上进一步确证 D-C 的超强结合能力。

与其他 3 组相比, 人 IgG3 组筛选中 D-C 富集偏慢, 在第 4 轮筛选时除此还筛得一定比例的另一组合分子 A_{9n}-G3_{9n} (30%)。噬菌体 ELISA 结果也显示 D-C 与 IgG3 的结合活性在四亚类中最弱。根据人 IgG1 (Gene ID:3500)、IgG2 (Gene ID:3501)、IgG3 (Gene ID:3502) 和 IgG4 (Gene ID:3503) 重链区 CH2-CH3 的氨基酸序列比对结果 (图 5), 我们将 SpA 与人 IgG 的 CH2-CH3 区域的作用位点^[30-31]进行标注 (阴影区域) 后发现与其他三亚类不同, 人 IgG3 在第 334、335 位结

合位点分别为 R、F，我们推测其影响了 IgG3 的 Fc 构象以及与 SpA 结合，同时也引起了上述 IgG3 与其他 IgG 组进化筛选的不同。

本研究用 IgG 四亚类对展示了 SpA、SpG 共 7 个单结构域随机组合噬菌体文库成功地进行了

体外分子进化筛选,得到了天然 SpA 分子中不存在的具有强结合活性的 D-C 组合,不仅可为 IgG 纯化、制备、检测等方面的应用提供了新的候选分子,还为细菌 IBP 结构功能的进一步研究提供了新的手段。

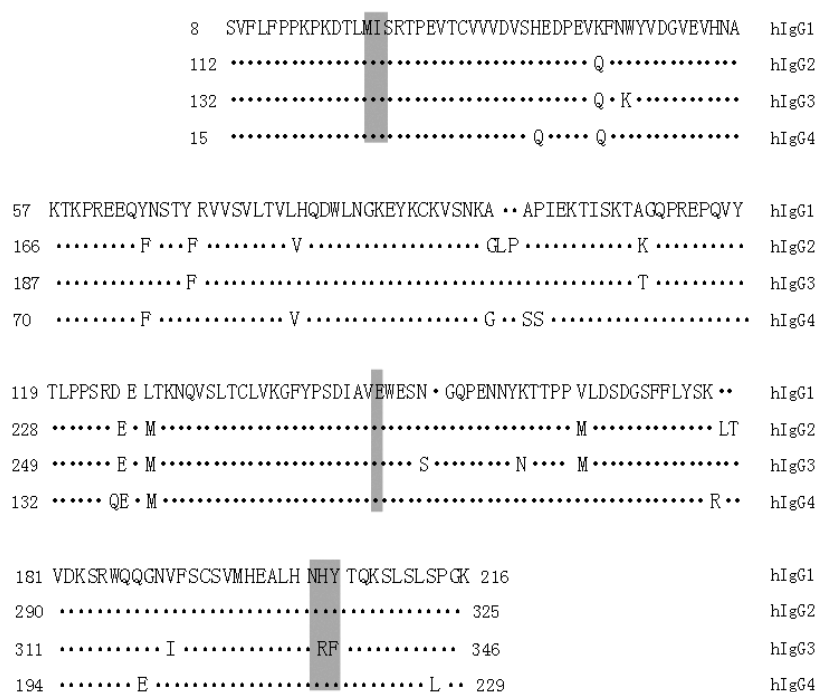


图 5 人 IgG1、IgG2、IgG3 和 IgG4 四亚类 Fc 的 CH2-CH3 段序列比对分析

Fig. 5 Alignment of amino acid sequences in CH2-CH3 of four human IgG subclasses. The binding sites of human IgG Fc with SpA are shaded.

REFERENCES

- [1] Housden NG, Harrison S, Roberts SE, et al. Immunoglobulin-binding domains: protein L from *Peptostreptococcus magnus*. *Biochem Soc T*, 2003, 31(3): 716–718.
- [2] Graille M, Stura EA, Housden NG, et al. Complex between *Peptostreptococcus magnus* protein L and a human antibody reveals structural convergence in the interaction modes of Fab binding proteins. *Structure*, 2001, 9(8): 679–687.
- [3] Kumar A, Tassopoulos AM, Li Q, et al. *Staphylococcus aureus* protein A induced inflammatory response in human corneal epithelial cells. *Biochem Biophys Res Commun*, 2007, 354(4): 955–961.
- [4] Atkins KL, Burman JD, Chamberlain ES, et al. *S. aureus* IgG-binding proteins SpA and Sbi: host specificity and mechanisms of immune complex formation. *Mol Immunol*, 2008, 45(6): 1600–1611.
- [5] Nomellini JF, Duncan G, Dorocicz IR, et al. S-Layer-mediated display of the immunoglobulin

- G-binding domain of streptococcal protein G on the surface of *Caulobacter crescentus*: development of an immunoactive reagent. *Appl Environ Microbiol*, 2007, 73(10): 3245–3253.
- [6] Goward CR, Scawen MD, Murphy JP, et al. Molecular evolution of bacterial cell-surface proteins. *Trends Biochem Sci*, 1993, 18(4): 136–140.
- [7] Brockwell DJ, Beddard GS, Paci E, et al. Mechanically unfolding the small, topologically simple protein L. *Biophys J*, 2005, 89(1): 506–519.
- [8] Stura EA, Graille M, Housden NG, et al. Protein L mutants for the crystallization of antibody fragments. *Acta Crystallogr D Biol Crystallogr*, 2002, 58(Pt 10 Pt 1): 1744–1748.
- [9] Uhlén M, Guss B, Nilsson B, et al. Complete sequence of the staphylococcal gene encoding protein A. A gene evolved through multiple duplications. *J Biol Chem*, 1984, 259(3): 1695–1702.
- [10] Lei HX, Wu C, Wang ZX, Zhou YQ, et al. Folding processes of the B domain of protein A to the native state observed in all-atom ab initio folding simulations. *J Chem Phys*, 2008, 128(23): 235105.
- [11] Moks T, Abrahmsén L, Nilsson B, et al. Staphylococcal protein A consists of five IgG-binding domains. *Eur J Biochem*, 1986, 156(3): 637–643.
- [12] Graille M, Stura EA, Corper AL, et al. Crystal structure of a *Staphylococcus aureus* protein A domain complexed with the Fab fragment of a human IgM antibody: structural basis for recognition of B-cell receptors and superantigen activity. *Proc Natl Acad Sci USA*, 2000, 97(10): 5399–5404.
- [13] Sasso EH, Silverman GJ, Mannik M. Human IgM molecules that bind staphylococcal protein A contain VHIII H chains. *J Immunol*, 1989, 142(8): 2778–2783.
- [14] Sasso EH, Silverman GJ, Mannik M. Human IgA and IgG F(ab')₂ that bind to staphylococcal protein A belong to the VHIII subgroup. *J Immunol*, 1991, 147(6): 1877–1883.
- [15] Stone GC, Sjöbring U, Björck L, et al. The Fc binding site for streptococcal protein G is in the C γ 2-C γ 3 interface region of IgG and is related to the sites that bind staphylococcal protein A and human rheumatoid factors. *J Immunol*, 1989, 143(2): 565–570.
- [16] Gronenborn AM, Clore GM. Identification of the contact surface of a streptococcal protein G domain complexed with a human Fc fragment. *J Mol Biol*, 1993, 233(3): 331–335.
- [17] Sloan DJ, Hellinga HW. Dissection of the protein G B1 domain binding site for human IgG Fc fragment. *Protein Sci*, 1999, 8(8): 1643–1648.
- [18] Kato K, Lian LY, Barsukov IL, et al. Model for the complex between protein G and an antibody Fc fragment in solution. *Structure*, 1995, 3(1): 79–85.
- [19] Watanabe H, Matsumaru H, Ooishi A, et al. Optimizing pH response of affinity between protein G and IgG Fc: how electrostatic modulations affect protein-protein interactions. *J Biol Chem*, 2009, 284(18): 12373–12383.
- [20] Zhou J, Liao YX, Chen Z, et al. Development of a SPA-ELISA method for detecting anti-coronavirus IgG antibodies in serum samples from fulvous fruit bats. *J South Med Univ*, 2008, 28(5): 736–738.
周杰, 廖玉学, 陈忠, 等. SpA-ELISA 检测棕果蝠血清冠状病毒抗体的研究. *南方医科大学学报*, 2008, 28(5): 736–738.
- [21] Poullin P, Announ N, Mugnier B, et al. Protein A-immunoabsorption (Prosorba[®] column) in the treatment of rheumatoid arthritis. *Joint Bone Spine*, 2005, 72(2): 101–103.
- [22] Hober S, Nord K, Linholt M. Protein A chromatography for antibody purification. *J Chromatogr B Analyt Technol Biomed Life Sci*, 2007, 848(1): 40–47.
- [23] Yang H, Cao J, Li LQ, et al. Evolutional selection of a combinatorial phage library displaying randomly-rearranged various single domains of immunoglobulin (Ig)-binding proteins (IBPs) with four kinds of Ig molecules. *BMC Microbiol*, 2008,

- 8(1): 137.
- [24] Jiang SH, Wang JF, Xu R, et al. Alternate arrangement of PpL B3 domain and SpA D domain creates synergistic double-site binding to VH3 and Vkappa regions of fab. *DNA Cell Biol*, 2008, 27(8): 423–431.
- [25] Cao J, Chen QL, Zhang HQ, et al. Novel evolved immunoglobulin (Ig)-binding molecules enhance the detection of IgM against hepatitis C virus. *PLoS One*, 2011, 6(4): e18477.
- [26] Xu R, Pan W, Shen YJ, et al. Construction of a novel phagemid pCANTAB5S. *Acta Univ Med Anhui*, 2004, 39(2): 83–86.
徐容, 潘卫, 沈毅珺, 等. 新型噬菌体展示载体 pCANTAB5S 的构建. *安徽医科大学学报*, 2004, 39(2): 83–86.
- [27] Ge YB, Yang XF, Du ZM, et al. Constructing a phage-displayed random mutation library of HIV-1 Tat38-61 at the sites of 51 and 55 amino acids in basic region. *Chin J Biotech*, 2011, 27(5): 1755–763.
葛宜兵, 杨旭芳, 杜哲明, 等. 噬菌体展示 HIV-1 Tat38-61 碱性区 51 和 55 位随机突变体文库的构建. *生物工程学报*, 2011, 27(5): 755–763.
- [28] Xu R, Shen YJ, Deng SH, et al. Phage display of random combinatorial libraries of Ig-binding mono-domains of protein a and protein l and Ig affinity screening. *Prog Biochem Biophys*, 2005, 32(6): 535–543.
徐容, 沈毅珺, 邓松华, 等. 噬菌体展示 protein A 及 protein L Ig 结合单结构域随机组合文库及 Ig 亲和筛选. *生物化学与生物物理进展*, 2005, 32(6): 535–543.
- [29] Shen YJ, Pan W, Xu Y, et al. Phage display of recombinant human lymphotoxin mutation libraries and receptor affinity screening. *Prog Biochem Biophys*, 2005, 32(1): 75–80.
沈毅珺, 潘卫, 许燕, 等. 噬菌体展示重组人淋巴毒素突变体库及受体亲和筛选. *生物化学与生物物理进展*, 2005, 32(1): 75–80.
- [30] DeLano WL, Ultsch MH, de Vos AM, et al. Convergent solutions to binding at a protein-protein interface. *Science*, 2000, 287(5456): 1279–1283.
- [31] Verkhivker GM, Bouzida D, Gehlhaar DK, et al. Monte Carlo simulations of the peptide recognition at the consensus binding site of the constant fragment of human immunoglobulin G: the energy landscape analysis of a hot spot at the intermolecular interface. *Proteins*, 2002, 48(3): 539–557.