

for the Control of Pharmaceutical and Biological Products, Beijing 100050, China).

Abstract On the basis of the transition-state theory, catalytic antibodies could be induced in body by using a transition-state analogue as hapten which was selected according to the specific reaction mechanism. A fuller description is given on the principle of hapten design for generation

of catalytic antibodies, the relationship between catalytic antibodies and non-catalytic antibodies and the comparison of catalytic antibodies with enzymes. The application prospects of catalytic antibody in medicine science and the limitations in the application are also discussed.

Key words catalytic antibody, antibody, enzyme, transition-state analogue

多肽噬菌体展示

唐为钢 甘人宝 王克夷

(中国科学院上海生物化学研究所, 上海 200031)

摘要 噬菌体展示技术已被广泛地应用于生物学研究的各个方面. 利用它可融合表达多肽、蛋白质结构域和蛋白质. 尤其是多肽噬菌体展示, 已被作为一种便利的研究工具去发现和研究那些与受体、酶、凝集素、抗体、核酸以及其他生物分子亲和的多肽配基和酶的底物专一性, 该技术在药物的发现, 疫苗的设计等医学领域也有着潜在的应用价值.

关键词 噬菌体展示, 固定化的靶分子, 多肽噬菌体展示

最近几年, 噬菌体展示技术发展迅猛, 它在抗体的抗原决定簇、蛋白质的抑制剂和激活剂、酶的底物专一性、蛋白质折叠、药物和疫苗的设计、基因的表达调控、免疫系统内的分子识别等研究领域有着广泛的应用. 所谓噬菌体展示 (phage display), 即用分子克隆的方法将外源核酸片段以融合蛋白的形式表达于噬菌体颗粒的外表面上. 一般地, 噬菌体展示实验操作可概括为两步: a. 建库; b. 筛库. 首先通过人工合成、cDNA 法、DNase I 随机水解法等来制备各种序列不一的核酸片段, 然后将其克隆于载体 (噬菌体或噬菌粒) 中, 随后通过感染或超感染大肠杆菌, 使其分泌出融合表达有外源片段的噬菌体, 合称为建库. 之后, 用固定化的靶分子 (immobilized target) 去筛选与它亲和的噬菌体. 称之为筛库. 噬菌体展示技术具有两个明显的优点, 其一是噬菌体易于扩增, 其二是融合表达的外源多肽或蛋白质

的氨基酸序列可通过测噬菌体 DNA 的序列而推知. 噬菌体展示可融合表达多肽、蛋白质结构域和蛋白质, 展示于噬菌体上的融合多肽称为噬菌体多肽. 本文将讨论占有重要地位的融合表达多肽的噬菌体展示——多肽噬菌体展示. 多肽噬菌体展示库依据外源融合多肽的构象是否被限制, 可分为两类, 一类为构象限制的肽库, 即用一对二硫键将外源融合多肽环化; 另一类为构象不限制的肽库, 即没有束缚外源融合多肽构象的二硫键. 依据外源融合多肽是以单拷贝形式还是多拷贝形式表达于噬菌体外膜上, 可相应地分为单价多肽噬菌体展示和多价多肽噬菌体展示. 以下将从两个方面并举出代表性的实例来介绍和评析渗透到各个生物学研究领域的多肽噬菌体展示. 首先, 将概述用来筛选与各种生物分子结合的多肽配基的多肽噬菌体展示库, 称之为配基噬菌体展示

库. 随后, 将介绍用来研究酶的底物专一性的底物噬菌体展示库.

1 配基噬菌体展示库

用来筛选配基噬菌体展示库的固定化的靶分子越来越广, 使得噬菌体展示技术的应用范围不断地拓展. 下文将根据用于筛库的固定化的靶分子的不同, 来概述用于筛选多肽配基的配基噬菌体展示库.

1.1 抗体

Scott 等构建了 FUSE5 噬菌体作为表达随机六肽的载体, 在随机六肽的两侧分别插入 DGA 和 GAAGA 两段氨基酸序列, 以尽可能减少噬菌体外膜蛋白对融合于其 N 端的六肽的构象的影响. Scott 等用两个单抗 M33 和 A2 (两者专一性结合蚯蚓肌红蛋白的抗原决定簇——DFLEKL) 来筛选随机六肽库, 筛到的六肽多数与抗原决定簇 DFLEKL 相似, 它们前三位氨基酸序列为 DFL, 而后三位氨基酸序列各有差异. 但他们也筛到了一种与抗原决定簇 DFLEKL 没有什么相似性的六肽 CRFVWC, ELISA 实验表明它与单抗 M33 结合很强, 这种与原始抗原决定簇 (此例中为 DFLEKL) 没有什么相似性的氨基酸序列称为 Mimotope^[1]. 此实验表明, 在事先不知道单抗的抗原决定簇的情况下, 通过筛选随机肽库可能得知单抗的抗原决定簇, 而且有时可筛到一些与原始抗原决定簇的氨基酸序列相差甚远的多肽, 丰富了抗体-抗原相互作用方面的结构信息.

1.2 凝集素

凝集素伴刀豆球蛋白 (ConA) 专一性结合甲基- α -甘露糖 (MaM), Scott 等用 ConA 去筛选随机六肽库, 得到的噬菌体多肽具有特征序列 YPY, ConA 与这些噬菌体多肽的结合可被 ConA 的天然配基 MaM 所抑制. 用另外一些类似的结合 D-甘露糖的凝集素进行结合实验, 这些噬菌体多肽同它们的结合很微弱, 表明这些噬菌体多肽与 ConA 结合的专一性很高^[2]. Oldenberg 等^[3]用 ConA 从随机八肽库

中也筛到了具有特征序列 YPY 的噬菌体多肽. 可见, ConA 的配基并不局限于糖类, 有可能从随机肽库中筛到模拟糖类的多肽.

1.3 主要组织相容性复合物

主要组织相容性复合物 II 类分子是具有高度多形性的膜糖蛋白, 与蛋白质的裂解肽段结合并将其呈递在外表面上, 以便被 CD4⁺ T 细胞识别. Hammer 等用三种相关的主要组织相容性复合物 II 类 alleles (HLA-DRB1 * 0101, * 1401, * 1101) 去筛选随机九肽库, 相应地得到了三种基元 YX₂MXAX₂L, WX₂MXTLX₂, WX₂MXRX₃, X 表示任一常见氨基酸 (下同). 三者在 1, 4 位上有两个锚定残基 (anchor residues) 类似或相同, 而在 6 位上显示出各自特异的锚定残基^[4,5]. 可见, 多肽噬菌体库对于研究多肽与主要组织相容性复合物的相互作用也有所裨益.

1.4 酶

用枯草杆菌蛋白酶部分水解牛胰核糖核酸酶, 得到两个片段: S-肽 (20 个氨基酸) 和 S-蛋白质 (104 个氨基酸). 两者单个都没有酶活力, 只有结合后才表现出酶活性, S-肽和 S-蛋白质体系被认为是研究受体与配基相互作用的很好的模型. Smith 等^[6]用 S-蛋白质去筛选随机六肽库, 筛到的噬菌体多肽具有特征序列 (F/Y)NF(E/V)(I/V)(L/V), 这一特征序列与 S-肽的序列相差甚远, 但用人工合成的具有该特征序列的多肽 YNFEVL 进行抑制实验, 表明它竞争性地抑制 S-肽与 S-蛋白质的结合, 是 S-肽的拮抗剂. 该实验对于拮抗剂、激动剂以及药物的发现和研究不无启示.

1.5 核酸

Krook 等^[7]分别在两种缓冲液体系 (pH 5.5 和 pH 7.5) 中用单链七聚胞苷酸来筛选随机六肽库. 在 pH 5.5 缓冲体系中筛到了三个与单链七聚胞苷酸强结合的序列不相似的噬菌体多肽 (PPPLYF, RFCDTs, RSR-LIW). 近年来, 人们对与 DNA 特异结合的多肽兴趣大增, 因为这些多肽可能用于选择性地调控细胞某些基因的表达, 为抗病毒和肿瘤药

物的设计提供信息, 而用随机肽库来研究核酸与多肽之间的相互作用比较方便.

1.6 受体

血小板上的凝血酶受体是含七次跨膜 α 螺旋的膜蛋白, 属于 G-蛋白偶联受体超级家族. 该受体被凝血酶切去胞外 N 端的一部分后, 其 N 端的六个氨基酸作为缚着的配基 (tethered ligand) 与受体结合并激活之. Doorbar 等^[8]利用含凝血酶受体的血小板来筛选随机肽库, 从中筛到了能免疫沉淀凝血酶受体的噬菌体多肽 (MSRPACPNDKYE). 人工合成该噬菌体多肽, 它可拮抗该受体的激动剂所引发的血小板凝集、酪氨酸磷酸化和释放 5-羟色胺. 而且, 他们也筛到了能激活该受体的噬菌体多肽.

1.7 钙调蛋白

钙调蛋白在胞内信号转导途径中的作用非常重要, 它可与许多靶蛋白或酶结合并调节它们的功能. Dedman 等^[9]用钙依赖的钙调蛋白去筛选随机十五肽库, 从中筛到了 28 种噬菌体多肽, 这些多肽都含有一色氨酸, 其中 W-P 出现的频率很高. 人工合成的这些噬菌体多肽可与结合钙的钙调蛋白结合, 并且与已知的钙调蛋白的抑制剂竞争结合钙调蛋白.

1.8 SH 同源域

在参与胞内信号转导途径的蛋白中, 有许多都含有一段由 50~60 个氨基酸组成的结构域, 即 SH3 同源域. SH3 同源域介导了信号转导途径中蛋白质与蛋白质的相互作用. 以往的研究表明 SH3 同源域的识别序列为 PXXP, 而 Rickles 等^[10]利用多肽噬菌体展示库分别进一步地确定了 5 种蛋白 (Src, Fyn, Lyn, P13K 和 Abl) 的 SH3 同源域的特征识别序列, 依次为 XXXRPLPPLPXP, XXXRPLPP (I/L) PXX, RXXRPLPPLPXP, RXXRPLPPLPPP, PPPYPPPP (I/V) PXX, 这对 SH3 同源域介导的信号转导途径的研究以及针对该途径的药物设计提供了有价值的信息.

1.9 整联蛋白

整联蛋白 (integrin) 是一类与细胞粘附

等有关的跨膜糖蛋白, 已知的整联蛋白都是异源二聚体. $\alpha_5\beta_1$ 整联蛋白能识别纤连蛋白 (fibronectin) 的 RGD 序列而与之结合. Koivunen 等^[11]用 $\alpha_5\beta_1$ 整联蛋白从随机六肽库中筛到了 32 种不同的 $\alpha_5\beta_1$ 整联蛋白的多肽配基, 其中 28 种含 RGD 基元, 3 种含与 RGD 相关的序列, 还有一种噬菌体多肽 (SLIDIP) 与 RGD 差别很大. 特别是筛到的噬菌体多肽 CRGDCL, 因其序列中含两个半胱氨酸, 所以它的构象可能被二硫键限制, 它能以高于其他含 RGD 序列的线性六肽 10 倍的效率, 抑制表达有 $\alpha_5\beta_1$ 整联蛋白的细胞与纤连蛋白的结合.

1.10 分子伴侣

分子伴侣 Bip 位于真核细胞的内质网中, 可协助新生肽穿过内质网膜以及随后在内质网腔内的折叠和装配. Bip 的作用依赖于它能识别一大批氨基酸序列不大相似的新生肽, 并且能够识别正确折叠和未折叠的蛋白质结构. Blond & Elguindi 等^[12]从随机八肽库和十二肽库中筛到了与 Bip 结合的噬菌体多肽, 这些多肽具有特征序列 Hy (W/X) HyXHyXHy, 该序列在奇数位上交替出现大的疏水氨基酸 (Hy), Hy 常为色氨酸、亮氨酸、或苯丙氨酸. 由此 Blond & Elguindi 提出了 Bip 的多肽结合位点是由相互交替的疏水氨基酸构成的模型. 人工合成具有该特征序列的多肽可与 Bip 结合并且激活它的 ATPase 活力. 由此可见, 多肽噬菌体库可作为研究蛋白质折叠的一种有力工具.

2 底物噬菌体展示库

底物噬菌体展示库已成功地用于研究蛋白水解酶和蛋白激酶的底物专一性, 以下将分别介绍研究这两类酶的底物专一性的底物噬菌体展示库.

2.1 蛋白水解酶

Matthews 等^[13]利用单价噬菌体展示库来研究蛋白水解酶的底物专一性, 他们构建了两个底物噬菌体展示库, 中间的插入片段分别为 GPGG (X)₅GPGG 和 GPAA (X)₅AAPG, 利

用它们来研究枯草杆菌蛋白酶和人 Xa 因子的底物专一性。通过对筛到的噬菌体的测序而推知, 枯草杆菌蛋白酶的水解底物含有组氨酸, 而且组氨酸多分布在随机多肽序列的第 2 位或第 4 位上; 抗枯草杆菌蛋白酶水解的多肽序列的特征为含一个或多个脯氨酸, 而且不含组氨酸。人 Xa 因子的底物序列的特征是几乎都含至少一个精氨酸, 其中有 4 个克隆含 2 个精氨酸。

Matthews 等^[14]还利用底物噬菌体展示库研究了弗林蛋白酶 (furin) 的底物专一性。弗林蛋白酶是哺乳动物的一种对许多组成型表达 (constitutively expressed) 的蛋白质前体进行剪切的酶。通过六轮筛选, 由测序而推知弗林蛋白酶的识别序列为 RXXR, 而且筛到的许多克隆在第 2 个精氨酸的前一位往往是赖氨酸, 或精氨酸, 或脯氨酸。

2.2 蛋白激酶

真核细胞 M 期 (分裂期) 蛋白激酶可磷酸化多种蛋白, 这些蛋白的磷酸化对细胞分裂可能非常重要。Westendorf 等^[15]利用了随机十五肽库和能识别 M 期真核细胞的 40 多种蛋白中的含磷酸化氨基酸残基的序列并与之结合的单克隆抗体 MPM2 来研究 M 期蛋白激酶的底物专一性。他们首先用 M 期蛋白激酶去磷酸化随机十五肽库, 那些被 M 期蛋白激酶磷酸化的噬菌体多肽因与单克隆抗体 MPM2 结合而被筛选出来, 通过测序而推知, M 期蛋白激酶的特征磷酸化识别位点是由连续 5 个氨基酸残基组成的多肽序列, 由在该序列中的每个位置上出现频率最高的氨基酸组成的序列是 LTPLK。

综合上述, 多肽噬菌体展示已广泛地应用于生物学的诸多研究领域, 而且, 很有可能从多肽噬菌体展示文库中筛到有临床应用价值的噬菌体多肽, 再通过人工合成、半合成等方法制备大量的多肽, 将其应用于临床。可以预见, 多肽噬菌体展示应用于生物学基础研究和应用开发的广度和深度将进一步得到拓展。

参 考 文 献

- 1 Scott J K, Smith G P. Searching for peptide ligands with an epitope library. *Science*, 1990, **249**: 386~ 390
- 2 Scott J K, Loganathan D, Easley R B *et al.* A family of concanavalin A-binding peptides from a hexapeptide epitope library. *Proc Natl Acad Sci USA*, 1992, **89**: 5398~ 5402
- 3 Oldenberg K R, Loganathan D, Goldstein I J *et al.* Peptide ligands for a sugar-binding protein isolated from a random peptide library. *Proc Natl Acad Sci USA*, 1992, **89**: 5393~ 5397
- 4 Hammer J, Takacs B, Sinigaglia F. Identification of a motif for HLA-DR 1 binding peptides using M13 display libraries. *J Exp Med*, 1992, **176**: 1007~ 1013
- 5 Hammer J, Valsasini P, Tolba K *et al.* Promiscuous and allele specific anchors in HLA-DR-binding peptides. *Cell*, 1993, **74**: 197~ 203
- 6 Smith G P, Schultz D A, Ladbury J E. A ribonuclease S-peptide antagonist discovered with a bacteriophage display library. *Gene*, 1993, **128**: 37~ 42
- 7 Krook M, Moshach K, Lindbladh C. Selection of peptides with affinity for single stranded DNA using a phage display library. *Biochem Biophys Res Commun*, 1994, **204** (2): 849~ 854
- 8 Doorbar J, Winter G. Isolation of a peptide antagonist to the thrombin receptor using phage display. *J Mol Biol*, 1994, **244**: 361~ 369
- 9 Dedman J R, Kaetzel M A, Chan H C *et al.* Selection of targeted biological modifiers from a bacteriophage library of random peptides. *J Biol Chem*, 1993, **268** (31): 23025~ 23030
- 10 Rickles R T, Botfield M C, Weng Z *et al.* Identification of Src, Fyn, Lyn, P13K and Abl SH3 domain ligands using phage display libraries. *EMBO J*, 1994, **13** (23): 5598~ 5604
- 11 Koivunen E, Gray D A, Ruoslahti E. Selection of peptides binding to the $\alpha_5\beta_1$ integrin from phage display library. *J Biol Chem*, 1993, **268** (27): 20205~ 20210
- 12 Blond-Elguindi S, Cwirla S E, Dower W J *et al.* Affinity panning of a library of peptides displayed on bacteriophages reveals the binding specificity of Bip. *Cell*, 1993, **75**: 717~ 728
- 13 Matthews D J, Wells J A. Substrate phage selection of protease substrates by monovalent phage display. *Science*, 1993, **260**: 1113~ 1117
- 14 Matthews D J, Goodman L J, Gorman C M *et al.* A survey of furin substrate specificity using substrate phage display. *Protein Science*, 1994, **3**: 1197~ 1205
- 15 Westendorf J M, Rao P N, Greace L. Cloning of cDNAs for M-phase phosphoproteins recognized by the MPM2 monoclonal antibody and determination of the phosphorylated epitope. *Proc Natl Acad Sci USA*, 1994, **91**: 714~ 718

Peptide Phage Display. TANG Weigang, GAN Renbao, WANG Keyi (*Shanghai Institute of Biochemistry, The Chinese Academy of Sciences, Shanghai 200031, China*).

Abstract Recently, phage display techniques have been applied to various biological aspects. Phage can be exploited to display peptides, protein domains and proteins on its surface. Especially peptide phage display is now well estab-

lished as a tool for the search for peptides binding to antibodies, enzymes, receptors, lectins, nucleic acids or other target molecules and in studying the substrate specificity of enzymes. It has great potential for application in the areas of drug discovery, vaccine development and other pharmaceutical development.

Key words phage display, immobilized target, peptide phage display

外显子捕捉法快速分离新基因的原理和方法

朱冠山¹⁾ 缪为民 焦炳华

(第二军医大学微生物学教研室, 上海 200433)

摘要 从大片段基因组 DNA 中快速分离转录序列是疾病基因定位克隆 (positional cloning) 中的关键步骤. 外显子捕捉法是一种较为成功的方法. 文章对该法的基本原理、方法步骤及应用成果等作了较为详细的介绍.

关键词 外显子捕捉法, 定位克隆, 基因组研究

近年来, 由于人基因组微卫星 DNA 多态标志的广泛应用, 使人们可以通过家系分析的方法对众多单基因遗传, 甚至多基因遗传的疾病基因进行精细的染色体定位. 而人类基因组计划的飞速发展, 使人们几乎完成了各条染色体的酵母人工染色体 (yeast artificial chromosome, YAC) 克隆及高分辨物理图谱. 这些成果大大方便了疾病基因的定位克隆 (positional cloning)^[1]. 当前, 在克隆某个疾病基因位点时, 人们根据其染色体定位, 能够迅速筛选到含有该位点的 YAC、粘粒 (cosmid) 或 p1 噬菌体等克隆. 但由于这些基因组大片段克隆长达几十至数百 kb, 因此, 如何快速从这些克隆中分离出基因转录序列, 则是决定定位克隆成功与否的关键步骤. 目前已经发展了数种方法, 其中由 Buckler 等^[2]建立的外显子捕捉法是较为成功的一种. 本文结合我们目前的工作拟对此法的原理及方法作一介绍.

1 基本原理

大多数真核生物的基因被较长的内含子序列分隔成数个较小的外显子. 这些基因最初被转录为 pre-mRNA, 之后经过在细胞核内拼接 (splicing) 过程, 将其中的内含子序列去除, 产生成熟的 mRNA 转运到细胞浆. 间隔外显子的拼接是由拼接体来完成的, 这种拼接体由 snRNP 和一些蛋白因子组成, 通过与 pre-mRNA 特定位点 (即 5' 拼接供位和 3' 拼接受位) 结合来实现拼接. 目前对拼接供位和受位的识别特征已有所了解, 已知拼接位点两侧有短的保守核苷酸序列, 其中 5' 端保守的拼接供位序列是 AG/GTRAGT, 3' 端保守的拼接受位序列是 NCAG/G ("R" 代表嘌呤核苷酸,

¹⁾第二军医大学第一附属医院传染病学教研室, 上海 200433.
收稿日期: 1996-05-07, 修回日期: 1996-09-20