

## 噬菌体展示随机十五肽库的构建

王筱蒙<sup>1</sup>, 邱 郑<sup>1</sup>, 张 芳<sup>2</sup>, 王 昊<sup>1\*</sup>

(<sup>1</sup>中国药科大学天然药物活性物质与功能国家重点实验室,南京 210009; <sup>2</sup>南京中医药大学中药学,南京 210023)

**摘要** 为了满足药物筛选的需要,本实验以 *pCANTAB 5E* 噬菌粒为载体,构建表达于丝状噬菌体尾丝蛋白 P3N 末端的随机十五肽库。通过自行设计引物与模板,人工合成编码随机十五肽的寡核苷酸片段及含有酶切位点的引物。通过 PCR 技术进行模板扩增获得编码随机十五肽的基因。将扩增后的目的基因经过 *Sfi I*, *Not I* 两个限制性内切酶双酶切后,与经同样双酶切的 5'去磷酸化的噬菌粒载体连接,将重组产物电转入大肠埃希菌 TG1 感受态细胞。集合菌落后进行库容和多样性检验。所建肽库库容可达  $5 \times 10^8$ 。随机挑取 20 个克隆测序,其核苷酸序列及推断出的氨基酸序列均随机。成功构建了库容量与多样性都满足筛选要求的噬菌体展示随机十五肽库。

**关键词** 噬菌体展示;肽库;随机十五肽

中图分类号 Q782 文献标志码 A 文章编号 1000-5048(2014)02-0242-05

doi:10.11665/j.issn.1000-5048.20140220

## Construction of a phage display random fifteen-peptide library

WANG Xiaomeng<sup>1</sup>, QIU Zheng<sup>1</sup>, ZHANG Fang<sup>2</sup>, WANG Min<sup>1\*</sup>

<sup>1</sup>State Key Laboratory of Natural Medicines, China Pharmaceutical University, Nanjing 210009;

<sup>2</sup>Traditional Chinese Medicine, Nanjing University of Chinese Medicine, Nanjing 210023, China

**Abstract** A phage display random fifteen-peptide library expressed in the filamentous phage tail fiber protein P3N terminal was constructed using *pCANTAB 5E* vector. First, a random oligonucleotides fragment containing fifteen codons was designed and synthesized. The fragment was amplified by PCR. Then, the amplified target gene was digested by two restriction endonuclease of *Sfi I* and *Not I*, and connected with *pCANTAB 5E* phagemid vector. Finally, the recombinant vector was transformed into electro-competent cells of *E. coli* TG1. The capacity of peptide library was up to  $5 \times 10^8$ . Twenty clones were randomly selected for sequencing; the nucleotide sequence and deduced amino acid sequences were randomized. The random fifteen-peptide library was successfully established for screening requirements.

**Key words** phage display; peptide library; random fifteen-peptide

This Study was supported by the National Natural Science Foundation of China (No. 81102899)

噬菌体展示技术(phage display technology)是由 Smith 于 1985 年创建,于 20 世纪 90 年代发展起来并得到广泛应用的新技术<sup>[1]</sup>。基本原理是噬菌体作为一种表达载体,可以将外源蛋白或多肽呈现在噬菌体表面,利用外源蛋白或多肽与待筛选分子的特异性亲和作用,通过吸附-洗脱-扩增的筛选富集过程,得到与待筛选分子特异性结合的外源蛋白或多肽。噬菌体展示技术不但可以将外源基因编码的

多肽或蛋白质呈现于噬菌体表面,而且能够在大肠埃希菌系统中进行扩增<sup>[2-4]</sup>,从而将基因型和表型、分子结合活性和噬菌体的可扩增性巧妙的结合起来,具有强大的筛选能力,是一种高效的筛选体系。

进行噬菌体展示技术操作的前提是获得相应的展示文库,主要包括展示随机肽库和抗体库两大类,其中噬菌体展示随机肽库是由随机排列的氨基酸组成的多肽展示在噬菌体表面所构成的分子文

\* 收稿日期 2013-11-15 \*通信作者 Tel:025-83271395 E-mail:minwang@cpu.edu.cn

基金项目 国家自然科学基金资助项目(No. 81102899)

库<sup>[8]</sup>,它提供了海量的结构与功能信息,只要库的容量足够大,几乎任何一种分子都能从肽库中获得与之结合的相应的配体。随机肽库可以用来研究蛋白质之间的相互识别、蛋白质的折叠及空间构象预测、肽与有机物之间的特异相互作用,以及酶与底物的结合、抗原抗体的相互作用、激素与受体的结合以及建立在这些信息基础上的实际应用,如疫苗设计、基因定位、小分子药物设计等<sup>[3]</sup>。

目前商品化的随机展示肽库的市场份额主要由美国 New England Biolabs 公司、安玛西亚及默克公司共同瓜分,国内尚无商品化的产品。这些肽库有一个共同的特点就是都价格不菲,而且使用载体为 M13 噬菌体载体。克隆在 M13 噬菌体载体间隔区的外源 DNA 有不稳定的趋向,因此商业肽库无法多次扩增使用。另外在已有一定研究基础的情况下,通常希望获得靶点蛋白的更多信息,此时通用性的商品化肽库往往不能满足要求,需要自行构建肽库进行进一步研究。因此,构建基于噬菌粒载体、稳定的、可反复扩增用于靶点筛选的、具有自主知识产权的肽库,是进行靶点研究的最佳途径。目前已自行构建十肽库、七肽库等报道,但构建十五肽库则未见报道。

本研究设计随机 15 肽的模板核苷酸序列,利用 PCR 扩增形成带有限制性内切酶切点的双链随机 15 肽模板 DNA,并与噬菌粒载体 pCANTAB 5E 连接,电转化大肠埃希菌,构建随机肽库。并从克隆插入正确率、肽库容量及多样性等方面对构建的肽库进行质量评价。

## 1 材 料

### 1.1 试剂、载体及菌株

*Sfi* I、*Not* I 限制性内切酶(美国 NEB 公司);牛小肠碱性磷酸酶(CIAP)(大连 TaKaRa 公司);质粒小提试剂盒、琼脂糖凝胶回收试剂盒、PCR 产物纯化试剂盒(Biomiga 公司);*Taq* DNA 聚合酶、DNA Marker DL2000(南京百斯凯公司);PCR 引物合成(上海生工生物技术服务有限公司);DNA 测序(上海美吉生物医药科技有限公司)。

噬菌粒 pCANTAB 5E、大肠埃希菌 TG1 均为本实验室保存。

### 1.2 仪 器

PCR 仪(德国 Eppendorf 公司);EPS-300 电泳

仪、HE-90 小号水平电泳槽、Tanon-2500 全自动凝胶图像处理系统(上海天能科学仪器有限公司);电击仪(美国 Bio-Rad 公司)。

## 2 方 法

### 2.1 随机十五肽核苷酸的合成

编码随机十五肽核苷酸片段的设计策略是采用(NNK)<sub>15</sub>编码方式,其中编码随机十五肽核苷酸序列的前两个核苷酸为任意核苷酸 N(A/T/C/G),最后一个核苷酸根据密码子的简并性设计为 K(G/T)。根据载体上 *Sfi* I、*Not* I 酶切位点情况及构建随机肽库的要求,设计两条寡核苷酸片段。链 1: 5'-GCTGGCCCAGCCGGCC (*Sfi* I) (NNK)<sub>15</sub> TA-AGCGGCCGC (*Not* I) AGGTGCATT-3', 在 5' 和 3' 端引入 *Sfi* I、*Not* I 酶切位点,用下划线表示;链 2: 5'-AATGCACCTGCGGCCGCTTA-3', 与链 1 的 3' 端互补。使最后合成的双链 DNA 的 5' 端和 3' 端具有完整的 *Sfi* I 和 *Not* I 核酸内切酶位点。用灭菌去离子水将合成的 DNA 干粉溶解,得到浓度为 20 μmol/L 的溶液,分装后于 -20 °C 冻存。

### 2.2 十五肽 DNA 片段的聚合和酶切

利用合成的模板(“2.1”项中的链 1)和引物(“2.1”项中的链 2)进行 PCR 扩增,从而得到双链 DNA。PCR 反应条件为:94 °C 预变性 5 min,94 °C 变性 1 min,56 °C 退火 30 s,72 °C 延伸 1 min,循环 30 次后,72 °C 延伸 10 min。经 1% 琼脂糖电泳检测扩增目的片段的大小,然后对 PCR 产物进行纯化,等待下一步酶切。将纯化好的 PCR 产物在 50 °C 水浴中用 *Sfi* I 酶切 3 h 后,加入 *Not* I,在 37 °C 水浴中酶切 3 h。用 PCR 产物纯化试剂盒回收酶切后的片段,并定量,分装后于 -20 °C 冻存。

### 2.3 pCANTAB 5E 噬菌粒载体的制备和酶切

用质粒小提试剂盒从宿主大肠埃希菌中提取 pCANTAB 5E 噬菌粒,在 50 °C 水浴中用 *Sfi* I 酶切 3 h 后,加入 *Not* I 在 37 °C 水浴中酶切 3 h。双酶切产物在 1% 的琼脂糖凝胶中于电泳 20 ~ 30 min 后,切下双酶切后的载体条带,用胶回收试剂盒回收后,用 CIAP 去磷酸化处理,再在 1% 的琼脂糖凝胶中电泳 20 ~ 30 min。切下目的条带,用胶回收试剂盒回收,定量,分装后于 -20 °C 冻存。

### 2.4 电转感受态细胞的制备

高质量的感受态细胞对于构建高容量文库至关

重要。本实验选用 *E. coli* TG1 作为感受态细胞菌株。挑取 TG1 单克隆于 LB 培养基 50 mL 中,37 °C, 振摇过夜。将过夜菌液每瓶 22 mL 分装于两个 500 mL LB 培养液中,25 °C, 220 r/min, 振摇 1 h 后, 每隔 20 min 测定一次对照组的  $A_{600}$ 。待  $A_{600}$  达到 0.35 左右时, 将实验组的菌 500 mL 平均分装在 3 个 250 mL 的离心杯中, 冰浴 20~25 min(不超过 0.5 h, 否则性菌毛会丢失)。4 °C, 4 000 r/min 离心 15 min, 弃去上清液, 得到菌体。用冰冷的双蒸水 500 mL 重悬(切勿剧烈晃动)。4 °C, 4 000 r/min 离心 10 min, 弃去上清液, 得到菌体。用冰冷的 10% 甘油 250 mL 重悬。4 °C, 4 000 r/min 离心 20 min, 弃去上清液, 得到菌体。用冰冷的 10% 甘油 10 mL 重悬。4 °C, 3 000 r/min 离心 20 min, 弃去上清液, 得到菌体。用冰冷的 10% 甘油 1 mL 重悬。将重悬的菌液每管 100  $\mu$ L 分装于 500  $\mu$ L 的 EP 管中, 可以立即转化, 也可以冻存于 -70 °C 长期保存。

## 2.5 十五肽 DNA 片段与噬菌粒载体的连接和转化

将双酶切并去磷酸化的噬菌粒载体和双酶切十五肽 DNA 片段按物质的量比 1:3, 1:5, 1:7, 1:10 的比例做一次小量连接和电转化, 确定最佳连接比例为 1:7。用不同浓度的连接产物进行转化以确定最佳转化产物量为 0.1  $\mu$ g(5  $\mu$ L)。利用 PCR 产物回收试剂盒纯化连接产物以除去连接酶和离子。取纯化连接产物 5  $\mu$ L 与感受态细胞 50  $\mu$ L, 混匀后加入预冷的 0.1 cm 电转化杯并在冰上放置 1~2 min, 然后在电转化仪(1.3 kV, 200  $\Omega$ , 25  $\mu$ F) 上进行电转化。电击后, 加入 SOC 培养基 1 mL, 如此转化 10 次, 最后将各次转化产物合并, 37 °C, 200 r/min 振摇 45~50 min 使细胞复苏。取少量菌体适量稀释后涂布 SOB(含 50  $\mu$ g/mL Amp) 小平板(直径 10 cm), 其余 8 000 r/min 离心 3 min 后涂布 SOB(含 100  $\mu$ g/mL Amp) 大平板(直径 15 cm), 37 °C 倒置培养 12~16 h, 以小平板进行计数, 大平板上的菌落用 LB 培养基溶解, 并用涂布棒均匀刮下来, 加甘油至终浓度 20%, 分装成小管, -70 °C 冻存。

## 2.6 随机十五肽库的容量测定和多样性鉴定

为测定随机十五肽库的容量, 在建库的过程中, 转化并复苏细胞后, 取少量菌液从  $10^{-1}$  到  $10^{-4}$  做 10 倍系列稀释, 分别涂布于 SOB 平板上, 次日根据单菌落数以确定库容量。为鉴定随机十五肽库的多样性, 从上述平板上随机挑取 20 个单菌落,

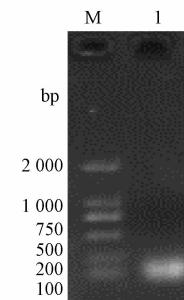
各提取重组质粒, 并对重组质粒进行 PCR 验证。

设计 PCR 验证引物: 上游引物 5'-CAGGAAA-CAGCTATGACC-3'(M13R), 下游引物 5'-GTAAAT-GAATTTCCTGTATGAGG-3'(S6)。对 PCR 阳性的重组质粒, 送去进行 DNA 序列测定, 并对构建的随机肽库的寡核苷酸的插入率、肽库的库容及多样性进行评价。

## 3 结 果

### 3.1 随机十五肽核苷酸的合成

利用合成的模板和引物进行 PCR 扩增, PCR 产物经 1% 琼脂糖凝胶电泳分析, 可见在 100 bp 左右有目的条带, 与预期片段大小接近, 见图 1。将 PCR 产物纯化后酶切, 并与载体 pCANTAB 5E 连接。连接产物电转化大肠埃希菌 TG1, 构建随机肽库。

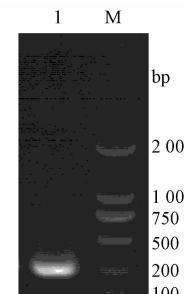


M:DNA marker; 1:PCR product

Figure 1 PCR amplification of DNA templates coding for random fifteen peptides

### 3.2 重组载体的 PCR 检验

从 SOB 平板上随机挑取若干(>25 个)长势较好的单菌落, 提取重组质粒, 并以重组质粒为模板, 用设计的两条引物 M13R 和 S6(序列见“2.6 项”)进行 PCR 验证, 对 PCR 产物进行 1% 琼脂糖凝胶电泳分析, 可见约 200 bp 的条带, 与预期片段大小接近(图 2), 其中 PCR 阳性率达 95%。



M:DNA marker; 1:PCR product of recombinant vector

Figure 2 PCR analysis of recombinant vector extracting from the selected clones

### 3.3 肽库的多样性

从PCR验证的阳性结果中选取24个克隆,进行序列测定,除去4个碱基数目不正确的乱码序列,测序结果的阳性率达83%,将其中20个碱基数目正确的核苷酸序列及其对应的氨基酸列于表1,20个DNA序列以及对应的氨基酸都是互不相同的。从表2的碱基分布情况分析,4种核苷酸的

出现情况与预先设计的NNK相符,第一、第二核苷酸分别为A/T/G/C中的任一种,其中每一种核苷酸和出现的频率的理论值为25%,第三位核苷酸为G/T中的任一种,G或T的出现频率为50%。实际值与理论值基本接近,但存在一定偏差,原因可能在于样本数量较小,这种偏差将随着测序样本的增加而逐渐减小。

**Table 1** Amino acid sequence corresponding to the selected clones nucleotide of random fifteen peptide library

Strain	Sequence
1	TCT TTG AGT CAT TTG AGT TTT TTG TGT CGG GGT CGT TTG CTT GGT S L S H L S F L C R G R L L G
2	TGT ACG GTG AAG TAT CTG CTT GCT GTT CTT CGG TGG TGT CAT TGG C T V K Y L L A V L R W C H W
3	TCG CTG TCG GGG ATT GTT CTG TCG TGT CTG TTG GTT GGT GTT CTG S L S G I V L S C L L V G V L
4	TAT ATG CTT CGT GTT ACT GTT ACG TGG TAT GGG TAT GTT CGT TCT Y M L R V T V T W Y G Y V R S
5	CGT AGT TCG GAG GCT GGT ACG CTG TCT CGC GAT TTG TCG ACT GGT R S S E A G T L S A D L S T G
6	TTG GCT GTT GGT ATT ATT TCT ACT CAT CTG GTG TGG CAT TGT GGT L A V G I I S T H L V W H C G
7	AGG TGG TTG ATT GCG TGT CCT ATT GTG TGG CCT CTG CAT CGT GAT R W L I A C P I V W P V H R D
8	CAG GTG TCG GCT GGT GTG CCT AGT CTG CAT CTG TCG AGT GAT CGT Q V S A G V A S L H L S S D R
9	ATG CTT CGG GTG CAG ATT CGT AGT GAT CGT GTG GCT GCT GAT CTG M L R V Q I R S D R V A A D L
10	AGT GTG CGG TGG ACG CCG CGT AGT ATG GCT GAT ATG CGG GTT ACG S V R W T P R S M A D M R V T
11	CCT AGT GAT CAG CGG GTG TAT TCG CTG GCT TGG ATG GCT TTG CGT P S D Q R V Y S L A W M A L R
12	ATT CGT CAG TCG TCG TCG GAG CTG GCT ACT CAG TCG CGT AGG TCG I R Q S S S E L A T Q S R R S
13	GTG TCT GCT ATG TCG GAT GAG TCG TCG CGC TGG TCT CGC TCT ACT V S A M S D E S S A W S A S T
14	ATG CGT CTG GAT AGG CGT TGG GCT ATT CTG TCG ATG GCT GAT ATT M R L D R R W A I L S M A D I
15	CGG GAT GTG TCG TTG TGT AGG GAT TCG GAT CGT TAT GCT TGT CTG R D V S L C R D S D R Y A C L
16	GAT GAT GAG GTT CAT CGC ACT GGT ACT GCT GAT GCT TGT ATT TCT D D E V H A S G T A D A C I S
17	GGT ATT TCG ACT GGT GCT ACT AGT CGT GTG CTG GTT GCT TCG TTG G I S T G A T S R V L V A S L
18	TCT GAG TGT GAT ACT GAG CTG CGT ACT CGT CGT TCG GAT TAT GTG S E C D T E L L S R R S D Y V
19	GCG CAT CGT TCG GCT CAT ATG TCG GAT GGT ATG GCT CTG CTG GAT A H R S A H M S D G M A L L D
20	TGT TGG CTG AAG CGT CAG CCT GCT GTG CTG CGG TGG TCT CAT TGG C W L K R Q P A V L R W S H W

**Table 2** Nucleotide distribution of selected clones in fifteen random peptide library(%)

Codon position	A	T	G	C
1	17.7	25.7	31.0	25.6
2	17.0	28.7	26.3	28.0
3	0	51.3	48.7	0

### 3.4 库容量

所构建的随机十五肽库的库容量,根据转化后在Amp抗性的平板上的菌落数(约 $6.5 \times 10^8$ ),结合PCR验证阳性率(95%)和测序结果的阳性率(达到83%以上)。根据公式:实际库容量 = 平板

总菌落数 $\times$ PCR验证阳性率 $\times$ 测序结果的阳性率,确定其实际库容量为 $5 \times 10^8$ 克隆。结果表明,所构建的噬菌体随机肽库不论从基本框架,还是从库容及多样性方面均满足设计要求。

#### 4 讨论

评价噬菌体展示随机肽库构建成功与否,主要是从库容及多样性两个方面进行。本研究所构建的随机十五肽库达到 $5 \times 10^8$ 以上实际库容,足以满足一般的筛选要求。而构建高库容量的肽库有几点需要注意,首先就是要采用高转化率的电转化方式进行转化,这里更为重要的是电转感受态细胞的制备,本实验通过不断摸索制备过程中细菌培养的温度、时间,连接产物的量,连接产物与感受态细胞混合物的孵育时间,电击的电压等条件,最终确定了最优转化条件(见“2.4项”和“2.5项”)。

另外,目的片段与载体的酶切酶连等也是影响最终库容的关键因素,本实验为了确保酶连效率,对双酶切后的载体又进行CIAP去磷酸化处理以防止载体自连。借助于丝状噬菌体这样一个相对稳定而且便于进行生化和遗传操作的载体,使肽库技术在各个方面得到了充分发展<sup>[5,9]</sup>。随机肽库的库容和多样性是衡量肽库质量与利用性的关键指标,文献报道噬菌体短肽库的库容一般在 $1 \times 10^6 \sim 1 \times 10^8$  cfu之间<sup>[6-7]</sup>,本研究应用pCANTAB 5E载体构建的十五肽库,实际库容达到 $5 \times 10^8$ ,基本满足筛选要求。目前大多数构建的肽库为七肽库,也有较长的如八肽,十肽等。在同样库容的条件下,随机肽越长,则多样性、特异性越

高,包含的信息越多,即包含更多的随机序列<sup>[8]</sup>。本实验构建较长的随机十五肽库,为今后开展肽库的研究与应用,特别是多肽类药物的筛选打下基础。

#### 参考文献

- [1] Smith GP. Filamentous fusion phage; novel expression vector that display cloned antigens on the viron surface[J]. *Science*, 1985, **228**(4 705):1 315 - 1 317.
- [2] Scott JK, Smith GP. Searching for peptide ligands with an epitope library[J]. *Science*, 1990, **249**(4 967):386 - 390.
- [3] Scheuermann J, Dumelin CE, Melkko S, et al. DNA-encoded chemical libraries[J]. *J Biotechnol*, 2006, **126**(4):568 - 581.
- [4] Jiang LD, Li LJ, Yu JY, et al. The application of phage display peptide library[J]. *J Clin Med Prac*(实用临床医药杂志), 2011, **15**(5):126 - 128.
- [5] Yuan L, Wang YZ, Xiao N, et al. Construction of a phage display random decapeptide library and panning for mimic antigen epitope of ScFv A1 against WSSV[J]. *Virol Sin*(中国病毒学), 2006, **21**(6):614 - 618.
- [6] Yang H, Zhang YQ, Han W, et al. Construction of random decapeptide library and screening identification of angiogenin binding peptides[J]. *Prog Biochem Biophys*(生物化学与生物物理进展), 2000, **27**(6):641 - 645.
- [7] Burritt JB, Bond CW, Doss KW, et al. Filamentous phage display of oligopeptide libraries[J]. *Anal Biochem*, 1996, **238**(1):1 - 13.
- [8] Hu Y, Yang JY, Hong SW. Construction of peptide library by phage display[J]. *Infect Dis Inf*(传染病信息), 2001, **14**(1):24 - 25.
- [9] Zheng JM, Zhang Y. The research and application of phage display peptide display[J]. *Mod Animal Husbandry*(当代畜禽养殖业), 2010(1):7 - 9.