



基于遗传算法的计算机辅助策略 优化孕酮适配体*

唐春花 江寒冰 杨洁 卢晓玲 陈美仑 魏铮 刘一杰 余鹏**

(中南大学湘雅药学院, 长沙 410013)

摘要 目的 本研究致力于优化孕酮 (progesterone, P4) 适配体的亲和力和选择性。方法 基于遗传算法 (genetic algorithm, GA) 的计算机辅助优化策略 (*in silico* maturation, ISM), 进行了4轮GA操作 (含交叉变异、单点突变和双点突变操作), 构建了初始文库和G1、G2、G3代ssDNA作为新的候选适配体库, 采用分子对接对候选适配体进行筛选和分析, 并使用迭代策略不断优化适配体。此外, 还提出了一种较为准确预测ssDNA三级结构的方法, 首先使用Mfold预测二级结构, 继而使用RNAComposer建立与ssDNA相对应的RNA三级结构, 输出的PDB文件使用Discovery Studio将RNA修改为DNA, 最后使用Molecular Operating Environment对结构进行能量最小化处理。结果 到G2代, 在局部搜索空间对P4S-0进行优化, 筛选出P4G1-14、P4G2-20、P4G1-6、P4G1-7和P4G2-14这5条适配体作为P4的最佳候选适配体。采用AuNPs比色法初步验证优化后适配体的亲和力, 继而构建了基于适配体结构开关的荧光法测定适配体的解离常数 (equilibrium dissociation constant, K_D), 并以此方法对适配体的选择性 (对双酚A、雌二醇、睾酮和皮质醇) 进行了评估。结论 通过ISM优化后的适配体, 对P4的亲和力较原适配体有了较大提升, 仍保留着识别结构类似分子的选择性。

关键词 遗传算法, 适配体, 亲和力, 分子对接

中图分类号 R917

DOI: 10.16476/j.pibb.2023.0021

孕酮 (progesterone, P4) 是一种天然的雌激素, 与女性的月经周期、怀孕及胎儿发育有着密切的关系^[1], 然而, 激素滥用导致的水源、食品中的P4含量过高却会给人们带来极大的身体危害^[2]。因此, 不论是在临床治疗用药, 还是在环境保护方面, 对P4浓度的监测都十分重要。常见的P4检测方法有免疫法和色谱法, 这两种方法灵敏度高, 特异性强, 但存在着价格昂贵、操作繁琐、专业性强、不适于现场检测等缺点^[3]。核酸适配体是一段单链寡聚核苷酸, 它可以与靶标分子特异性结合, 因此可作为新型识别分子应用于分析检测领域。适配体可以通过体外选择过程从随机核苷酸文库中获得, 这种方式称为指数富集的配体系统进化技术 (systematic evolution of ligands by exponential enrichment, SELEX)^[4]。考虑到初始文库的多样性有限, 且在PCR过程中存在偏好性, 可能导致潜在的高亲和力适配体丢失, 有研究者根据其筛选经验预估SELEX的成功率约为50%^[5]。

计算机辅助优化策略 (*in silico* maturation, ISM) 可作为一种针对适配体筛选后优化、改善适配体功能的方法。有研究者指出, 应用ISM对SELEX筛选后的序列进行重筛选可在一定程度上弥补SELEX技术的局限性^[6-7]。如模拟生物进化算法的遗传算法 (genetic algorithm, GA), 将亲代适配体通过选择、交叉和变异产生子代, 从有希望的亲本序列中进化出改良的适配体, 通过反复选择、进化, 在经ISM优化后进行体外评价。2005年, Ikebukuro等^[8]首次将基于GA的ISM应用于核酸适配体的优化, 用于抑制凝血酶的DNA适配体的重筛选。虽然ISM已经被应用于筛选多个靶标的

* 中南大学中央高校基本科研业务费专项基金 (2021zzts1002) 和湘基金委 [2022] 2号自然科学基金 (2022JJ80105) 资助项目。

** 通讯联系人。

Tel: 0731-82650329, E-mail: peng.yu@csu.edu.cn

收稿日期: 2023-01-16, 接受日期: 2023-05-04

高亲和力适配体且不需要重复 SELEX 过程^[7-13], 但迄今为止, 还没有研究者尝试使用 ISM 来提高适配体对 P4 的亲和力。

P4 属于小分子化合物, 已有研究指出针对小分子靶标选择的适配体比其他靶标选择的适配体亲和力低^[14]。筛选小分子靶标适配体的主要挑战之一是靶标-核酸复合物和游离核酸的分离, 由于靶标的分子质量较小, 并不会导致游离核酸和靶标-核酸复合物之间存在较大的质量差异, 分离过程更加复杂, 这也是传统 SELEX 技术在筛选小分子靶标适配体时面临的主要挑战^[15]。针对小分子适配体实验筛选的困难, ISM 则可作为一种有效的补充选择, 来改善小分子靶标适配体的功能。

虽然 ISM 的核酸文库容量比传统 SELEX 的容量小, 倘若每一个子代中的所有序列 (10~50 个序列) 都进行合成和评估, 对于高通量筛选来说, ISM 过程仍然是昂贵、繁琐和耗时的。同时, 研究人员也不可能通过确定每次传代过程中适配体-靶标复合物的结构特征, 来选择最好的亲本适配体进行传代。在 ISM 中, 毫无疑问以有效的、节约时间和成本的手段来识别适配体-靶标复合物之间的相互作用对筛选高质量核酸适配体起着非常重要的作用。而分子对接则可以作为研究适配体与靶标结合时产生何种亲和力以及亲和力大小的一个理想的工具^[16-17]。目前学界面临的主要问题之一是在分子对接时 ssDNA 适配体及其靶点的研究缺乏准确预测 ssDNA 三维结构的工具。最近也有研究者提出了通过先预测与 DNA 相对应的 RNA 三维结构, 再将其转换为 DNA 的方法来预测 ssDNA 的三维结构^[18]。

此外, 由于结合试验的灵敏度较低, 目前还没有标准化的分析方法来表征小分子适配体对靶标的

亲和力大小^[19-20]。近来有研究人员提出, 对于适配体-靶标复合物的解离常数 (equilibrium dissociation constant, K_D) 的测定, 大致方法可分为基于分离策略 (如透析法、超滤法、凝胶电泳法、毛细管电泳法等) 或均相策略 (荧光强度法、荧光各向异性法、圆二色谱法、紫外可见分光光度法及表面等离子体共振法等)^[21-22]。纳米金 (AuNPs) 比色法是一种现象明显、易制备、操作简单、耗时短的表征方法, 目前已有广泛的应用, 只是 AuNPs 比色法一般并不用于 K_D 值的准确测定, 而是多用于亲和性、特异性检测与传感器的制备^[22-23]。荧光检测法在测定适配体-靶标复合物时具有诸多优势, 它可以在适配体与靶标处于平衡时进行测定, 避开了结合速率和解离速率的问题。因此, 本研究在结合已有研究的基础上, 首先使用 AuNPs 比色法对经 ISM 优化后的适配体亲和力进行初步验证, 随后根据核酸适配体结构转换信号检测靶标的原理, 将标记有羧基荧光素 (carboxy fluorescein, FAM) 的核酸适配体与标记猝灭基团 (black hole quencher 1, BHQ1) 的适配体互补链配对杂交, 通过 P4 竞争适配体的结合位点使荧光恢复, 建立了一种基于核酸适配体结构转换开关的荧光共振能量转移 (fluorescence resonance energy transfer, FRET) 方法对适配体的 K_D 值进行测定^[24]。

2017 年, P4 适配体经过截短后, 这里称为 P4S-0, 在适配体传感器的构建中得到了应用^[25]。本研究描述了一种基于 GA 的 ISM 与分子对接相结合的方法, 来筛选和优化出高亲和力和选择性的 P4 适配体 (图 1); 此外, 还提出使用 Mfold 和 RNAComposer 分别预测 ssDNA 的二级结构和三级结构, 使用 Discovery Studio 将 RNA 修改为 DNA,

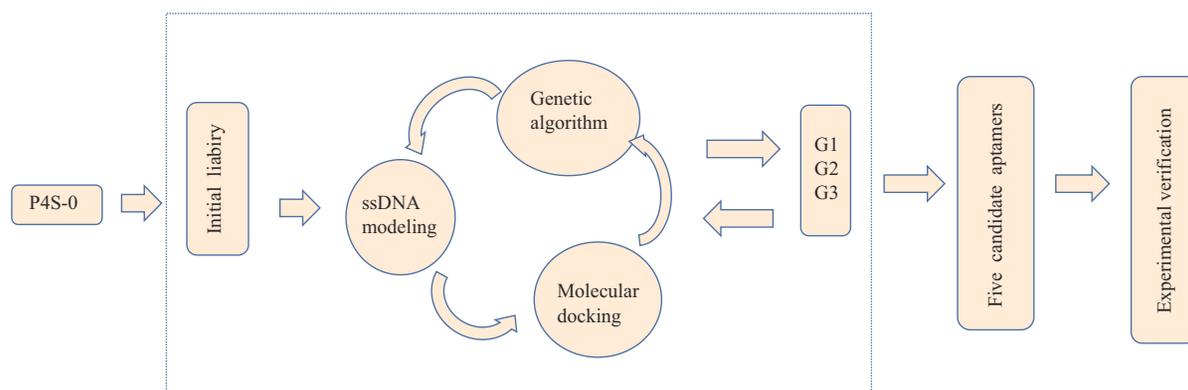


Fig. 1 The genetic algorithm-based *in silico* maturation strategy combined with molecular docking to obtain progesterone-binding aptamer with improved affinity

最后对结构进行能量最小化处理, 作为分子对接中的 ssDNA 适配体的三维模型的输入结构; 在亲和力验证阶段, 首先使用 AuNPs 比色法对适配体进行初步表征, 随后采用基于 FRET 的荧光法对适配体的 K_D 值进行测定, 优化后的候选适配体与其互补序列如表 1 所示。

Table 1 Candidate aptamers and their complementary sequences

Name	Sequence (5'-3')
P4S-0	GATTAACATTAGCCACCGCCACC
FDNA ₀	FAM-CTAATGTTAATC
BDNA ₀	GTGGGCGGTGGG-BHQ1
P4G1-14	GGTTAACATAAGCTAACCGACCTCC
FDNA ₁	FAM-CTTATGTTAACC
BDNA ₁	GAGGTCGGTTAG-BHQ1
P4G2-20	GATTAACAGACGGCCCTCGACCTCC
FDNA ₂	FAM-CGTCTGTTAATC
BDNA ₂	GAGGTCGAGGGC-BHQ1
P4G1-6	GATTAACATAAGCTAACCGCCGTC
FDNA ₃	FAM-CTTATGTTAATC
BDNA ₃	GACGGCGGTTAG-BHQ1
P4G1-7	GATTAACATAAGCACACCGCCGTC
FDNA ₄	FAM-CTTATGTTAAGC
BDNA ₄	GACGGCGGTGTG-BHQ1
P4G2-14	GGTTAACATGAGCTAACCGCCGTC
FDNA ₅	FAM-CTCATGTTAACC
BDNA ₅	GACGGCGGTTAG-BHQ1

1 材料与方法

1.1 P4适配体的计算机辅助优化方法

1.1.1 初始文库的生成

轮盘赌算法原理与操作过程: GA 中的轮盘赌是一种简单的选择方法, 即每个个体的选择概率和其适应度值成比例, 适应度越大, 选中概率也越大, 类似于商场的抽奖轮盘^[26]。此处以每条适配体与 P4 进行分子对接打分的分值作为算法中的适应度值, 打分越高, 那么适应度越高, 该序列成为亲本的可能性越大。后续每次参与变异的亲本序列、选择变异的类型与选择发生变异的位置等都离不开轮盘赌算法。

首先利用 GA 构造 DNA 序列库, 使用给定的 P4 适配体序列 P4S-0 产生初始亲代 DNA 序列库。初始序列库由 3 种类型的适配体组成, 包括分别由 1 个点突变、2 个点突变和不同交叉 (单点、双点

和多点交叉) 产生的 3 组适配体, 共 60 条序列。所有 GA 过程均在 Microsoft Visual Studio (2022)^[27] 中使用 C++ 语言编写的程序实现, 设定恒定的交叉百分比 (0.8) 和突变率 (0.2)^[28]。执行此程序后, 得到了初始文库的 60 条序列。

1.1.2 基于 GA 的 P4 适配体迭代优化步骤

产生第一代序列 (G1): 初始文库包括 60 条序列, 根据初始文库中序列的分子对接结果, 选择打分排名前 5 的序列, 根据对接结果的打分值对初始序列以不同的适应度值 (即出现率) 进行复制, 并随机配对选择亲本, 通过轮盘赌选择方法对序列进行单点、双点和多点交叉 (程序随机产生 1 个 0 到 2 的随机整数, 0 对应发生单点交叉, 1 对应发生两点交叉, 2 对应发生多点交叉), 产生 20 条不重复的子代序列作为 G1。将得到的适配体再次进行分子对接。

产生第二代序列 (G2): 将 G1 和初始文库中对接打分排名前 5 的序列作为亲代, 将单碱基突变随机引入序列以产生 1 组 20 条新序列作为 G2。具体操作为: 同样以不同的出现率按照轮盘赌方法选择亲代, 然后再产生 1~25 内 (序列长度为 25) 的随机数决定亲代的哪个位置发生单碱基突变, 突变后的碱基种类也是由随机数的产生决定, 随机产生 0~3 之间的整数, 每个整数对应一种碱基 (0 对应 A、1 对应 T、2 对应 C、3 对应 G), 如果产生的碱基与原碱基相同, 则重新产生随机数决定。此时产生了 20 条新序列作为 G2, 继续进行分子对接。

产生第三代序列 (G3): 选择初始文库、G1、G2 代中打分排名前 5 的序列, 引入双碱基突变以产生 20 条新序列作为 G3。具体操作为: 根据不同的出现率和轮盘赌方法, 来选择每次突变的亲代, 选择亲代后产生两个不同的随机整数来选择发生碱基突变的序列位置 (随机数范围就是序列长度范围 1~25), 再产生 0~3 的随机整数来表示突变后的碱基 (0 对应碱基 A、1 对应碱基 T、2 对应碱基 C、3 对应碱基 G), 若突变后的碱基与亲代序列的原位置碱基相同, 就重新产生随机数来重新选择, 直到选择到不同的碱基为止, 突变后将序列与亲代、已经产生过的突变序列进行比较, 如果重复, 那么将重新在该亲代上选择突变位置和突变后的碱基。此时产生了 20 条新序列作为 G3 代, 同样进行分子对接。

所有 GA 优化步骤和适配体对接打分结果见图 S1 和表 S1, 在初始文库、G1、G2 和 G3 代中选择

5条打分最高的序列作为候选适配体。

1.1.3 适配体二级结构的预测

使用 Mfold Web 服务器 (<http://www.unafold.org/hybrid2-same.php>) 对初始文库以及 G1、G2、G3 代序列的二级结构进行预测^[29]。预测在温度为 25°C, 离子浓度与后续荧光实验缓冲液 (pH 7.6) 的离子组分相同时, 即在 120 mmol/L Na⁺、20 mmol/L Mg²⁺ 下热稳定最好的结构。

1.1.4 适配体三级结构的建模和优化

对 ssDNA 适配体三级结构的预测主要包括 3 个步骤。首先, ssDNA 的二级结构已经通过 Mfold Web 服务器预测出, 将预测的二级结构通过 RNA Composer 服务器建立相对应的 3D RNA 模型^[30]。随后, 使用 Discovery Studio 4.5^[31] 将 ssRNA 结构中的尿嘧啶转换为胸腺嘧啶, 将核糖核酸转换为脱氧核糖核酸, 即转换为 ssDNA 的三级结构。随后, 使用 Molecular Operating Environment (MOE, 2014)^[32] 将 ssDNA 的三级结构进行能量最小化处理, 在 Amber 力场下使用最陡下降法使斜率小于 0.000 01 kcal/mol²。

通过对蛋白质数据库 (<https://www.rcsb.org/>) 中的 24 个 ssDNA 序列进行核磁共振解析, 评价该方法的准确性。用 Visual Molecular Dynamics (VMD)^[33] 软件比较了核磁共振求解的糖-磷酸盐骨架结构和用上述方法预测的糖-磷酸盐骨架结构之间的均方根偏差 (RMSD) 来确定结构的相似度。

1.1.5 分子对接

使用分子对接软件 AutoDockTools (ADT,

1.5.6)^[34] 对 GA 产生的 ssDNA 适配体与 P4 的亲合力大小进行评估。将能量最小化后的 ssDNA 适配体结构视为刚性结构, 随后对 P4 小分子进行检测适配体中心与扭转键的操作。为了覆盖适配体结构中的所有活性位点, 创建大小在 126 Å×126 Å×126 Å (x, y, z), 间距为 0.375 Å 以内的对接盒子。使用拉马克遗传算法构象搜索, 建立了 100 次遗传算法运行的所有对接模拟, 种群规模为 1 502 500 000 次能量评估的最大数量, 每次运行 27 000 代。根据结合能、相互作用类型 (主要是氢键、疏水作用) 和结合位点等分子对接结果对最佳配合物进行评分和筛选。

1.1.6 对接结果的图像表征

使用 Pymol^[35] 将适配体-P4 复合物的结构进行图像表征。

1.2 实验方法

为了验证经 ISM 优化后的适配体亲和力与特异性, 将适配体进行了如下的实验评估。首先, 使用 AuNPs 比色法初步验证, 比较候选适配体的亲和力大小^[22], 原理如图 2 所示。随后构建如图 3 的适配体结构开关准确测定候选适配体的 K_D 值。根据适配体结构开关检测靶标的原理, 在候选适配体的两条互补链上从 3' 和 5' 端分别标记荧光基团 FAM 和猝灭基团 BHQ1, 当候选适配体与两条互补链碱基配对杂交时, 标记了荧光基团的 FDNA 与标记了猝灭基团的 BDNA 相互靠近, 荧光猝灭; 通过 P4 竞争结合候选适配体 FDNA 的结合位点, 使两条互补链从适配体上解离下来, 荧光恢复。以此核酸适配体结构开关法评估候选适配体的亲和力和特异性。

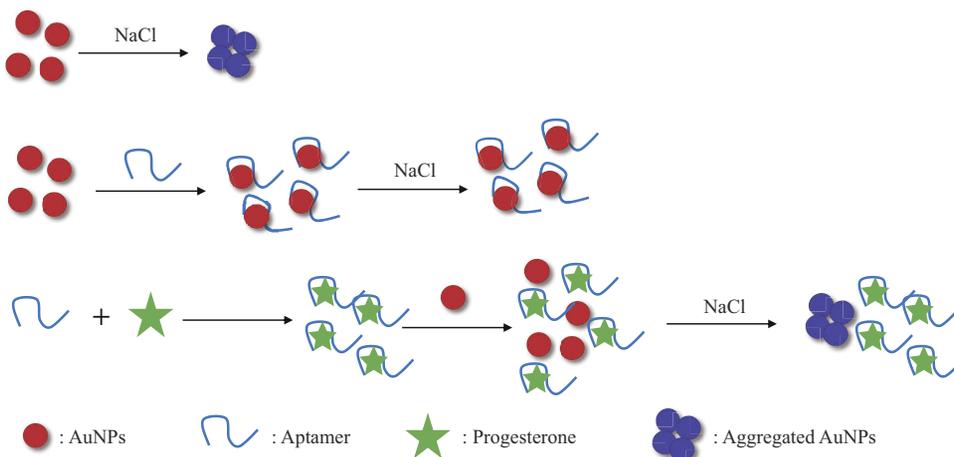


Fig. 2 The schematic for the evaluation of candidate aptamer based on colorimetric method

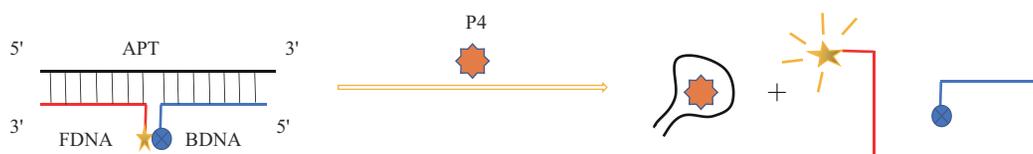


Fig. 3 The schematic for the evaluation of candidate aptamer based on FRET and the aptamer structure switch

1.2.1 实验试剂与仪器

孕酮（源叶生物科技有限公司，上海），双酚A、雌二醇、睾酮和皮质醇（阿拉丁生化科技股份有限公司，上海）、适配体及其互补序列（生工生物工程公司，上海）、Tris-HCl（福州奥研实验器材有限责任公司，福州），氯化钠、氯化镁和吐温-20（Tween-20）（阿拉丁生化科技股份有限公司，上海）、纳米金胶体（先丰纳米，南京）。

F-7100 荧光分光光度计（HITACHI，日本），设置激发波长为 480 nm，激发和发射狭缝宽度设定为 5 nm 和 5 nm；Synerg Lx 多功能酶标仪（博腾仪器有限公司，美国）；ZQPW-70A 全温振荡培养箱（莱玻特瑞仪器设备有限公司）；AUY220-0.1 mg 分析天平（岛津，日本）；TGL-20M 台式高速冷冻离心机（迈克实验仪器有限公司，湖南）；Q-LAB20-DV 超纯水仪（启沁环保科技有限公司，湖南）。

1.2.2 AuNPs 比色法初步验证候选适配体的亲和力

AuNPs 比色法的实验原理如图 2 所示，先对实验中 NaCl 和适配体的浓度进行优化，具体见文档 S1 和图 S2, S3。条件优化后进行适配体亲和力的定性分析，分别量取 25 μ l 3 μ mol/L 适配体溶液与 25 μ l 系列浓度梯度 P4 溶液（0、0.5、1.0、2.0、4.0、6.0、8.0、10.0 mg/L），震荡混匀，在室温下孵育 30 min 后，再加入 150 μ l AuNPs 溶液和 10 μ l 2.5 μ mol/L NaCl 溶液，室温孵育 5 min 后观察混合溶液的颜色变化，记录溶液吸光度变化情况，计算吸光度比值 A_{620}/A_{520} ，以 P4 浓度为横坐标，吸光度比值 A_{620}/A_{520} 为纵坐标，绘制工作曲线。

1.2.3 基于适配体结构开关的 FRET 法测量适配体的 K_D 值

a. 适配体 K_D 值的测量

构建适配体结构开关（图 3），并优化适配体与互补链的浓度、结构开关形成时间、P4 孵育时间以及探索溶液中 Mg^{2+} 浓度和 Tween 20 含量对结构开关的影响，上述条件优化见文档 S1 和

图 S4-S8。依据上述条件优化后的结果，测定在不同浓度 P4 的情况下结构开关体系的荧光信号变化强度。分别配制 0、10、20、30、50、100、500、1 000 nmol/L 的 P4 溶液。将缓冲液稀释好的 0.1 μ mol/L FDNA、0.2 μ mol/L 适配体、0.3 μ mol/L BDNA 各 50 μ l 置于 95°C 水浴锅中加热 10 min，取出后迅速转移到碎冰中冷却降温。然后取适配体、FDNA、BDNA 一起孵育 15 min，后加入 50 μ l P4 溶液在 37°C 恒温下孵育 30 min，使二者得到充分结合，在测定前加入 300 μ l 纯净水将反应体系补足到 500 μ l，利用荧光分光光度计测定每组荧光信号值。以 P4 浓度为横坐标 x ，以荧光信号恢复百分比 $(F_1 - F_2)/F_2$ 为纵坐标 y ，其中 F_1 是加入 P4 之后的荧光值， F_2 是阴性对照组（P4 为 0）的荧光值。通过 GraphPad Prism 8.0 拟合不同浓度 P4 与适配体结合的非线性饱和曲线，根据曲线 $y = B_{max}/(K_D + x)$ 计算 K_D 值， B_{max} 指的是适配体与 P4 的最大结合率，是常量。根据此方法测定以上对接打分最高的 5 条候选适配体与靶标复合物的 K_D 值。

b. 通过实验评估适配体的选择性

经 ISM 优化后，适配体对 P4 的特异性同样需要鉴定。因此，本实验选取了双酚 A、雌二醇、睾酮、皮质醇作为干扰物，将这些物质配制成浓度为 0.1 μ mol/L 的溶液。0.1 μ mol/L FDNA、0.2 μ mol/L P4 适配体和 0.3 μ mol/L BDNA 各 50 μ l 先孵育 15 min，然后再加入 0.1 μ mol/L 各物质的溶液 50 μ l，在测定前加入 300 μ l ddH₂O 将反应体系补足到 500 μ l，利用荧光分光光度计测定出不同物质的荧光信号变化强度，以不同物质种类为横坐标 x ，以荧光信号恢复值 $(F_1 - F_2)$ 为纵坐标 y 作图。

2 结果与讨论

2.1 候选适配体的热力学性质

初始文库包含 60 条 ssDNA，包含 3 种不同类型的二级结构，包括简单发夹环（hairpin loop）、内环（internal loop）和多分支环（multibranch

loop)。子代 G1、G2、G3 中的适配体二级结构更加多样化 (图4)。初始文库中二级结构形成的最小自

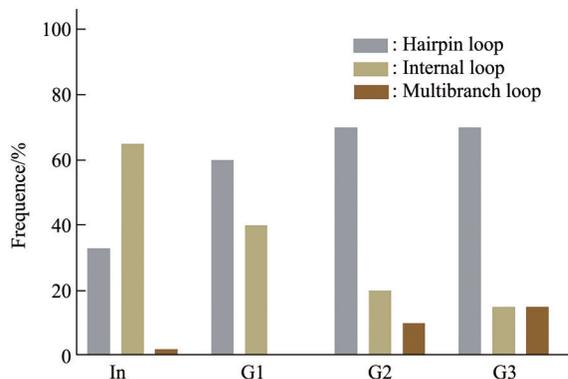


Fig. 4 Frequency of secondary structures in groups from the initial library and following generations (G1, G2, G3)

由能在 $-0.01 \sim -4.42$ kcal/mol, G1 代中为 $-0.03 \sim -10.21$ kcal/mol, G2 代为 $-0.19 \sim -4.98$ kcal/mol, G3 代为 $-0.27 \sim -10.8$ kcal/mol。初始适配体 P4S-0 具有内环结构 (图 5), 其自由能约为 -0.35 kcal/mol。初始库和子代文库中选择的几条候选适配体二级结构如图 5 所示, 结果表明, 不同的 GA 操作可以成功地创建一个虚拟文库, 不同子代的核酸序列在二级结构和热力学性质上具有多样性。

2.2 适配体的三维建模

通常, 虚拟筛选需要得到受体和配体分子的三

维结构。由于目前缺乏对 ssDNA 进行准确三维建模的计算工具, 且难以通过实验确定文库中所有核酸序列的三维结构, 许多研究人员不得不转而寻找较为可靠的预测 ssDNA 三维结构的方法。Jeddi 和 Saiz^[18] 在 2017 年首次报道了一种预测 ssDNA 适配体三维结构的方法, 这里本文学习这种建模思路并进行了一定修改。主要预测步骤包括: a. 构建二级结构; b. 构建等效的三维 ssRNA 结构; c. 根据报道的步骤, 将三维 ssRNA 结构转换为 ssDNA 结构。本文在步骤 b 中使用了 RNAComposer 而不是文献中的 Assemble2/Chimera; 此外, 并非使用 VMD, 而是使用 Discovery Studio 4.5 将 ssRNA 模型转换为等效的 ssDNA 3D 结构, 并使用 MOE 软件包进行能量优化。通过对 24 个具有已知结构的 ssDNA 进行 3D 建模, 每个预测的 ssDNA 结构都与从 PDB 数据库下载的相应 ssDNA 结构对齐, 并计算了每对结构之间的相应均方根偏差 (RMSD) 以测量相似程度。图 6 显示了 24 个预测的 3D 结构 (ssDNA-黄色) 和从 PDB 数据库下载的 24 个 NMR 结构 (ssDNA-绿色)^[36-56] 的叠加, 以及 RMSD 的计算值。结果表明, 本方法能够准确预测各种 ssDNA 的结构, 序列范围是 7~27 个核苷酸, 得到的 RMSD 值范围在 $2.51 \sim 9.02$ Å 的, 通常在 4 Å 附近, 平均值为 5.6 Å。然后, 利用本方法预测所有新的适配体的三维结构, 并用于分子对接研究。

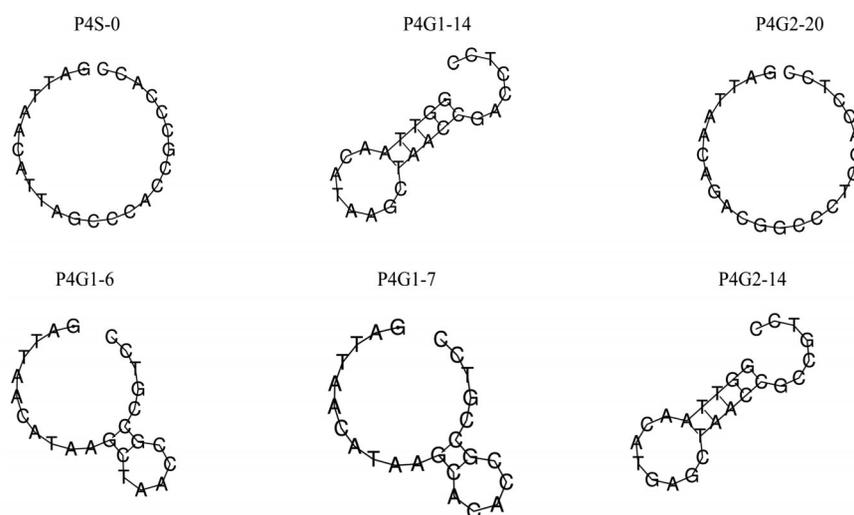


Fig. 5 *In silico* prediction of secondary structure with the minimum free energy of the selected aptamers

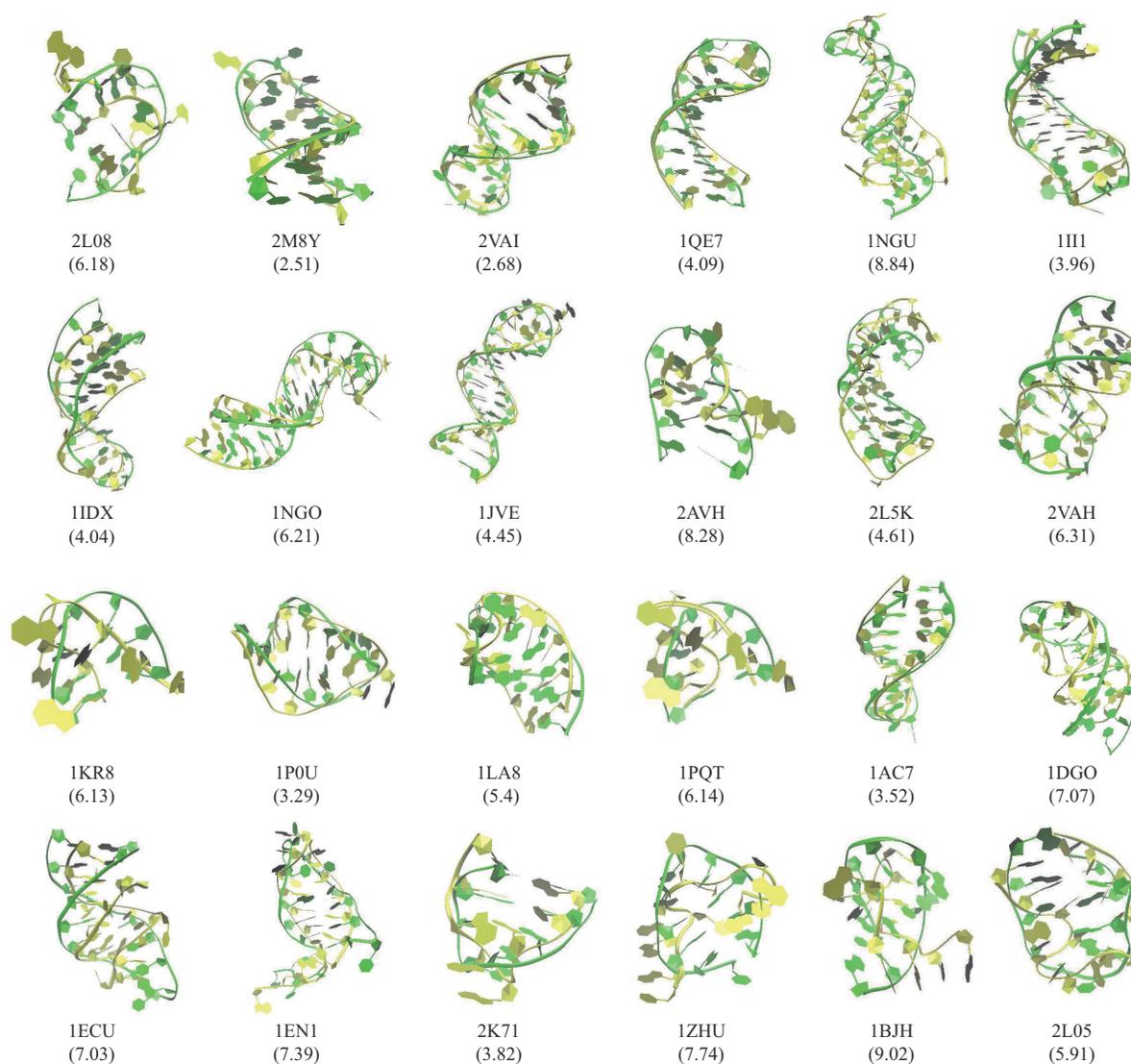


Fig. 6 Alignment of the 3D structures predicted by our pipeline (yellow) and the corresponding experimental ones obtained from PDB (green) for the 24 ssDNA

The PDB ID and RMSD values (Å) are presented as label and in parenthesis, respectively.

2.3 使用计算机辅助策略优化P4适配体

利用分子对接技术研究 P4 与 120 条候选适配体的结合能、结合位点和相互作用模式，其中包括来自初始文库和 G1、G2、G3 代的序列。在初始文库和不同子代中，ssDNA 适配体的预测结合能分别在 $-4.26 \sim -6.7$ kcal/mol 和 $-4.46 \sim -7.48$ kcal/mol (表 S1)。

初始适配体 P4S-0 与 P4 的相互作用主要是疏水作用和氢键作用，对接能量为 -4.63 kcal/mol，主要是 P4S-0 的这几个残基——DA2、DT3 和 DT4 与 P4 进行相互作用。P4S-0 的 DA2 残基与 P4 的 20 位酮羰基形成氢键 (图 S9)。

在初始文库和后几代的打分排名靠前的序列中，大多数具有简单的简单发夹环结构 (图 5)，其自由能等于或低于 P4S-0。对接结果指出，发夹环适配体的结合位点主要位于同一环区，而多分支环适配体的结合位点主要位于同一环区的循环区域外。通常，未配对残基可以作为结合位点参与靶标识别，因为这些残基更灵活，有更多可用的供体或受体^[57]。此外，对接结果也表明，候选适配体的平均结合能从初始库 (-5.08 kcal/mol) 到 G1 (-5.98 kcal/mol)、G2 (-5.93 kcal/mol)、G3 (-5.85 kcal/mol) 大幅降低，G3 代的结合能又有一些升高。在经过这几轮 GA 后的局部搜索后，似乎

对适配体序列进行了适当优化, 而进一步的修改可能导致了序列的分化, 导致功能的丧失。

在 G1 打分最高的序列中, 适配体 P4G1-14、P4G1-6 和 P4G1-7 对 P4 的亲合力最高。适配体 P4G1-14、P4G1-6 和 P4G1-7 的结合能分别为 -7.48 、 -6.9 和 -6.86 kcal/mol。P4G1-14 与 P4 的结合模式为沟槽结合, 主要作用力有氢键和疏水作用。P4G1-6 与 P4 主要通过疏水作用和氢键产生相互作用, 其中疏水作用占主导, 结合模式为沟槽结合。P4G1-7 与 P4 的结合位点在非环区, 作用模式也是沟槽结合。

在 G2 代中, P4G2-14 和 P4G2-20 的对接打分最高, 对接能量分别为 -6.76 和 -6.95 kcal/mol。P4G2-14 与 P4 的相互作用模式为插入结合, P4G2-20 与 P4 的相互作用模式为沟面结合, 相互作用力中疏水作用占主导。以上适配体与 P4 具体相互作用的具体碱基和 2D 拓扑图 (图 S9, S10), 可以更直观地体现适配体与目标物的具体作用位点、作用力的方向等。

基于热力学性质和对接分析, 筛选出 P4G1-14、P4G2-14、P4G1-6、P4G1-7 和 P4G2-20 这 5 条适配体最佳候选适配体, 随后通过实验进行评价, 实验部分以 P4S-0 为对照。

2.4 初步验证候选适配体的亲和力

在优化后的 NaCl 浓度、适配体浓度的条件下, 基于 AuNPs 比色法对原适配体 P4S-0 与候选适配体 (P4G1-14、P4G2-20、P4G1-6、P4G1-7 和 P4G2-14) 进行了亲和力比较。在与系列浓度梯度的 P4 溶液 (0、0.5、1.0、2.0、4.0、6.0、8.0、10.0 mg/L) 充分混合均匀后, 观察溶液颜色变化, A_{620}/A_{520} 比值越高, 说明适配体与靶标的亲和力越强。随着 P4 浓度的升高, 在浓度为 0~4 mg/L 时, 各核酸适配体吸光度比值 A_{620}/A_{520} 逐渐上升, 说明此时 P4 浓度不够, AuNPs 上仍吸附有过量的核酸适配体, 还未完全与 P4 结合, 受核酸适配体的保护, AuNPs 产生聚集不明显。随着 P4 浓度的升高, 越来越多 AuNPs 上吸附的核酸适配体脱落与 P4 结合, AuNPs 产生的聚集效应越来越明显, 直到反应完全, A_{620}/A_{520} 基本不变, 此时 AuNPs、核酸适配体与 P4 三者的结合达到了平衡。另外, 从图 7 可以看出, 在同等条件下, 候选适配体 P4G1-14 的吸光度比值 A_{620}/A_{520} 较高, 明显高于其他适配体, 说明

P4G1-14 与 P4 的亲合力最强, 原始适配体 P4S-0 的吸光度比值最低, 其余 4 条候选适配体的亲和力均高于 P4S-0。从上述分析可初步得出, 优化后的候选适配体亲和力均较 P4S-0 有了较大提升。只是 AuNPs 比色法并不能对适配体的 K_D 值进行准确测定。

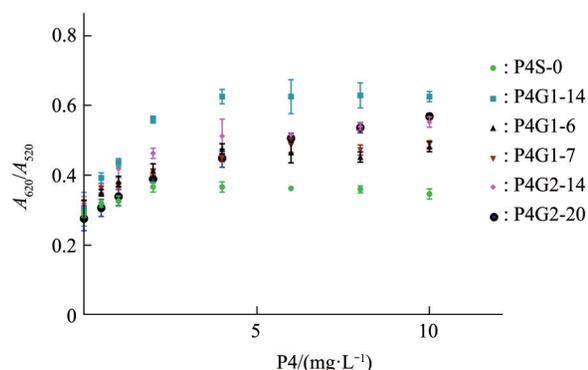


Fig. 7 Variation of A_{620}/A_{520} of AuNPs solution of candidate aptamers with increasing P4 concentration

2.5 适配体的 K_D 值

2.5.1 适配体 K_D 值的测定

使用基于适配体结构开关的 FRET 法测定适配体靶标复合物的 K_D 值, K_D 值越小, 说明适配体与靶标的亲和力越高。以上候选适配体与原始适配体的 K_D 测定结果如图 8 所示, 并与 AutoDOCK 对接打分结果进行比较 (表 2)。结果表明, 经以上 GA 优化后的候选适配体相比于原始适配体 P4S-0, 对 P4 的亲合力有了较大提高, 其中, P4G1-14 的亲合力提升最明显, 其 K_D 值为 4.457, 相比于 P4S-0 亲和力提升了约 3 倍, 其次是 P4G2-20、P4G1-6、P4G1-7 和 P4G1-14, K_D 值分别为 6.432、8.582、10.45、11.34, 此方法得出的 K_D 值亲和和大小顺序与对接打分的结果基本一致。这种实验的优势在于, 检测荧光信号可以在适配体和 P4 处于平衡状态时进行 K_D 测定, 这不仅避开了结合速率或解离速率的问题, 同时允许在各种缓冲溶液中进行测定, 这可能是其他分离模式的分析技术难于实现的^[58]。但是, 由于小分子适配体的亲和力鉴定目前缺乏标准化的分析方法, 此方法仅可作为判断的部分依据, 如果有其他亲和力鉴定方式作为辅助将会使结果更加可靠。

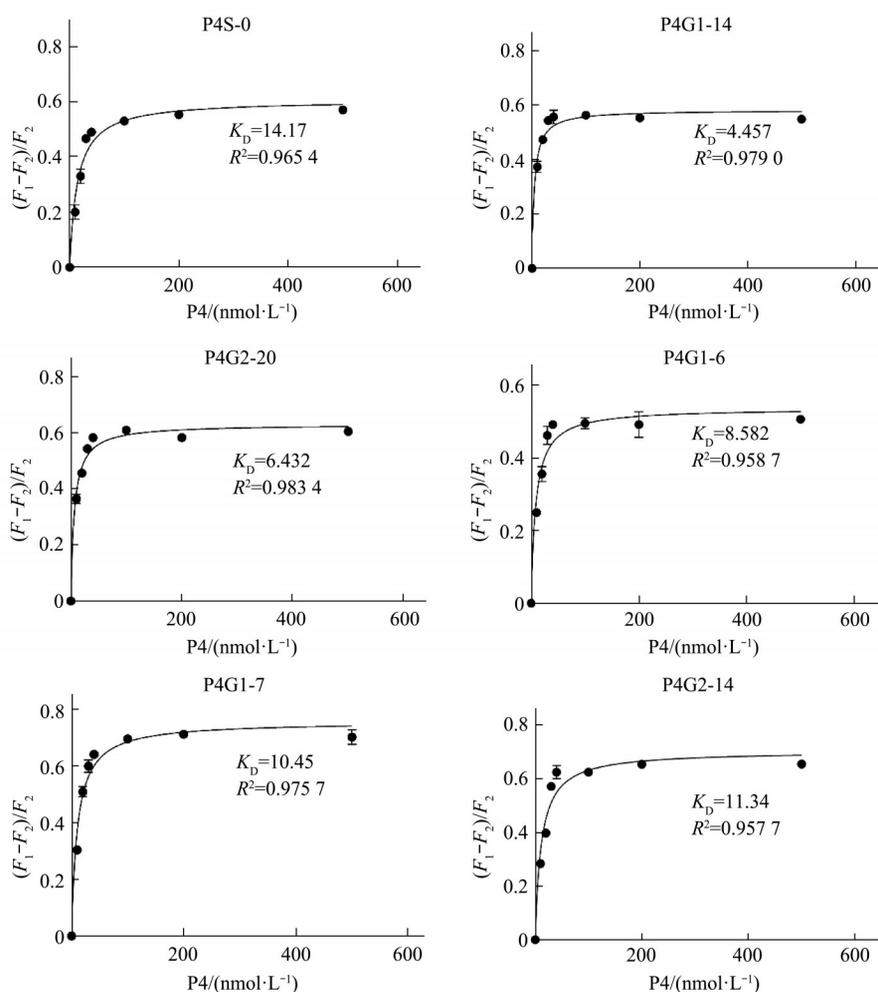


Fig. 8 Saturation curve of candidate aptamers binding to P4

Table 2 The K_D value of aptamer was measured by fluorescence detection method based on the structure switch of aptamer and compared with the docking score

Aptamer	K_D /(nmol·L ⁻¹)	Score/(kcal·mol ⁻¹)
P4S-0	14.17	-4.63
P4G1-14	4.457	-7.48
P4G2-20	6.432	-6.95
P4G1-6	8.582	-6.90
P4G1-7	10.45	-6.86
P4G2-14	11.34	-6.76

此外，方法中的测试结果可能会受到FDNA与BDNA的解链温度的影响，当这两条链GC含量高时，解链温度会随之升高^[59]，使得链解离更加困难，导致测得的解离常数偏小，这一点会对实验结果造成影响。例如P4S-0的BDNA链中GC含量约

为80%，而P4G1-14的BDNA中GC含量为58%，解链温度差距的影响可能会对实验结果造成一定偏差。本实验没有设置第二种有效的方法测定 K_D 值，无法与FRET法的结果相佐证，是本实验的局限性所在。

2.5.2 通过实验评估适配体的特异性

如图9所示，几个候选适配体对P4的响应值都是最高的，几种干扰物都只能引起部分响应。其中，双酚A的干扰性最小，而雌二醇、睾酮和皮质醇由于拥有与P4一样的甾体结构，所以候选适配体保存着对这几种甾体物质的识别能力，对雌二醇、睾酮和皮质醇的响应值较双酚A高。从图9中可以看出，在P4与其他干扰物共存的情况下，候选适配体仍对P4具有较高的特异性，候选适配体可用于进一步的传感器设计开发。

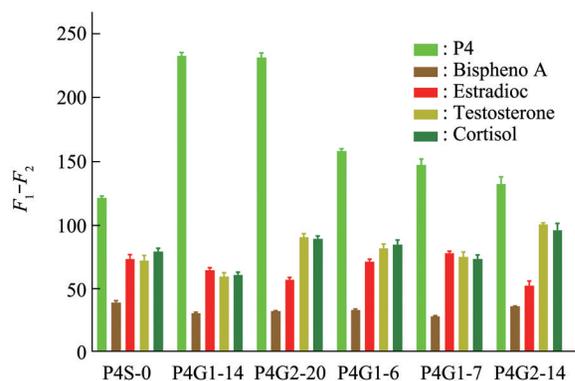


Fig. 9 Fluorescence response of aptamer structural switch to different targets

F_1 represented the fluorescence of solution containing aptamer, FNDA, BDNA and P4, F_2 represented the fluorescence of solution containing aptamer, FDNA and BDNA.

3 结 论

本研究成功地使用基于GA的ISM策略, 与分子对接结合, 提高了之前报道的P4S-0适配体对P4的亲合力。计算机打分结果与实验的一致性证实了此策略是设计或优化适配体功能的有效工具。分子对接技术可在分子水平上阐明受体与配体之间的相互作用, 有助于设计或优化所需的适配体。然而, 缺乏对生物大分子3D结构预测软件是计算机辅助技术发展的重要限制。本文通过先分别预测序列的二级结构以及相对应的RNA三级结构, 后将RNA转换为DNA并经过能量优化得到ssDNA的三级结构, 能够较准确预测DNA适配体的3D结构, 首次预测了原始适配体P4S-0与P4的结合位点和相互作用类型。此外, 对不同适配体-P4复合物的分析表明, 较低的结合能、较简单的二级结构、较低的DG和相互作用类型是得到高亲合力、高选择性适配体的重要因素。

得到经ISM优化的适配体后, 对适配体的亲合力和特异性验证是本研究的关键步骤。本研究首先采取了AuNPs比色法对候选适配体的亲合力进行初步验证, 由于AuNPs比色法并不能准确测定适配体的 K_D 值, 于是构建了基于结构开关的FRET法准确测量适配体的 K_D 值。基于适配体结构开关检测适配体对P4亲合力的方法, 可以避免结合速率或解离速率的问题, 同时允许在各种缓冲溶液中进行测定, 在表征小分子结合适配体方面所表现出的优势可以用于开发高效的荧光适配体传感器。但该

方法的缺陷在于, 对于小分子配体的亲合力鉴定没有一种标准化的分析方法, 应当多引入几种适配体亲合力测定方法对计算机优化适配体的结果加以佐证, 且此方法中互补链的不同解链温度会对结果造成偏差, 是FRET法的局限性所在, 同时, 也应该补充灵敏度实验以更直观地体现优化后的适配体拥有更强的功能。

其他改善适配体功能的策略, 如结合序列的共轭或者构建多价适配体、稳定适配体的结构或将疏水基团引入适配体, 可通过此方法流程进行研究。

附件 见本文网络版 (<http://www.pibb.ac.cn> 或 <http://www.cnki.net>):

PIBB_20230021_Doc S1.pdf
 PIBB_20230021_Figure S1.pdf
 PIBB_20230021_Figure S2.pdf
 PIBB_20230021_Figure S3.pdf
 PIBB_20230021_Figure S4.pdf
 PIBB_20230021_Figure S5.pdf
 PIBB_20230021_Figure S6.pdf
 PIBB_20230021_Figure S7.pdf
 PIBB_20230021_Figure S8.pdf
 PIBB_20230021_Figure S9.pdf
 PIBB_20230021_Figure S10.pdf
 PIBB_20230021_Table S1.pdf

参 考 文 献

- [1] 苏竞梅. 关于孕酮的那些事儿. 人人健康, 2015, 404(15): 65
 Su JM. Health for Everyone, 2015, 404(15): 65
- [2] Jimenez G C, Eissa S, Ng A, *et al.* Aptamer-based label-free impedimetric biosensor for detection of progesterone. Anal Chem, 2015, 87(2): 1075-1082
- [3] 杜高尚. 基于核酸适配体的孕酮快速检测新方法研究[D]. 上海: 上海交通大学, 2017
 Dy G S. Study on the Aptamer-Based Novel Methods for Rapid Detection of Progesterone[D]. Shanghai: Shanghai Jiao Tong University, 2017
- [4] Tuerk C, Gold L. Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. Science, 1990, 249(4968): 505-510
- [5] Mayer G, Ahmed M S L, Dolf A, *et al.* Fluorescence-activated cell sorting for aptamer SELEX with cell mixtures. Nat Protoc, 2010, 5(12): 1993-2004
- [6] Kim N, Gan H H, Schlick T. A computational proposal for designing structured RNA pools for *in vitro* selection of RNAs. RNA, 2007, 13(4): 478-492
- [7] Bini A, Mascini M, Mascini M, *et al.* Selection of thrombin-

- binding aptamers by using computational approach for aptasensor application. *Biosens Bioelectron*, 2011, **26**(11): 4411-4416
- [8] Ikebukuro K, Okumura Y, Sumikura K, *et al.* A novel method of screening thrombin-inhibiting DNA aptamers using an evolution-mimicking algorithm. *Nucleic Acids Res*, 2005, **33**(12): e108
- [9] Hasegawa H, Savory N, Abe K, *et al.* Methods for improving aptamer binding affinity. *Molecules*, 2016, **21**(4): 421
- [10] Savory N, Abe K, Sode K, *et al.* Selection of DNA aptamer against prostate specific antigen using a genetic algorithm and application to sensing. *Biosens Bioelectron*, 2010, **26**(4): 1386-1391
- [11] Torkamanian-Afshar M, Nematzadeh S, Tabar zad M, *et al.* *In silico* design of novel aptamers utilizing a hybrid method of machine learning and genetic algorithm. *Mol Divers*, 2021, **25**(3): 1395-1407
- [12] Hasegawa H, Savory N, Abe K, *et al.* Methods for improving aptamer binding affinity. *Molecules*, 2016, **21**(4): 421
- [13] Ikebukuro K, Yoshida W, Noma T, *et al.* Analysis of the evolution of the thrombin-inhibiting DNA aptamers using a genetic algorithm. *Biotechnol Lett*, 2006, **28**(23): 1933-1937
- [14] Ruscito A, Derosa M C. Small-molecule binding aptamers: selection strategies, characterization, and applications. *Front Chem*, 2016, **4**:14
- [15] Pfeiffer F, Mayer G. Selection and biosensor application of aptamers for small molecules. *Front Chem*, 2016, **4**:21
- [16] Chushak Y, Stone M O. *In silico* selection of RNA aptamers. *Nucleic Acids Res*, 2009, **37**(12): e87
- [17] Buglak A A, Samokhvalov A V, Zherdev A V, *et al.* Methods and applications of *in silico* aptamer design and modeling. *Int J Mol Sci*, 2020, **21**(22): 8420
- [18] Jeddi I, Saiz L. Three-dimensional modeling of single stranded DNA hairpins for aptamer-based biosensors. *Sci Rep*, 2017, **7**(1): 1178
- [19] Mckeague M, De Girolamo A, Valenzano S, *et al.* Comprehensive analytical comparison of strategies used for small molecule aptamer evaluation. *Anal Chem*, 2015, **87**(17): 8608-8612
- [20] 苏艺, 蒋灵丽, 林俊生. 小分子靶标与其核酸适配体亲和力的表征方法. *中国生物工程杂志*, 2019, **39**(11): 96-104
Su Y, Jiang L L, Lin J S. *China Biotechnology*, 2019, **39**(11): 96-104
- [21] Svobodova M, Skouridou V, Botero M L, *et al.* The characterization and validation of 17 beta-estradiol binding aptamers. *J Steroid Biochem Mol Biol*, 2017, **167**: 14-22
- [22] Alsager O A, Kumar S, Zhu B C, *et al.* Ultrasensitive colorimetric detection of 17 beta-estradiol: the effect of shortening DNA aptamer sequences. *Anal Chem*, 2015, **87**(8): 4201-4209
- [23] Mei Z L, Chu H Q, Chen W, *et al.* Ultrasensitive one-step rapid visual detection of bisphenol A in water samples by label-free aptasensor. *Biosens Bioelectron*, 2013, **39**(1): 26-30
- [24] Nutiu R, Li Y F. Structure-switching signaling aptamers. *J Am Chem Soc*, 2003, **125**(16): 4771-4778
- [25] Alhadrami H A, Chinnappan R, Eissa S, *et al.* High affinity truncated DNA aptamers for the development of fluorescence based progesterone biosensors. *Anal Biochem*, 2017, **525**(1): 78-84
- [26] 陈雯祎, 李琪. 一种基于轮盘赌选择的改进遗传算法. *福建电脑*, 2016, **32**(5): 50-51
Chen W W, Li Q. *Journal of Fujian Computer*, 2016, **32**(5): 50-51
- [27] Beirne J P, Gilmore A, Roddy A, *et al.* A bespoke target selection tool to guide biomarker discovery in tubo-ovarian cancer. *Comp Struct Biotechnol J*, 2022, **20**: 3359-3371
- [28] Mousivand M, Anfossi L, Bagherzadeh K, *et al.* *In silico* maturation of affinity and selectivity of DNA aptamers against aflatoxin B-1 for biosensor development. *Anal Chim Acta*, 2020, **1105**(1):178-186
- [29] Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res*, 2003, **31**(13): 3406-3415
- [30] Popena M, Szachniuk M, Antczak M, *et al.* Automated 3D structure composition for large RNAs. *Nucleic Acids Res*, 2012, **40**(14): e112
- [31] Oghara N, Haley-Vicente D. Protein target discovery and characterization - DS modeling and discovery studio streamline target discovery. *Genet Eng News*, 2002, **22**(21): 77
- [32] Johnson G T, Hertig S. A guide to the visual analysis and communication of biomolecular structural data. *Nat Rev Mol Cell Biol*, 2014, **15**(10): 690-698
- [33] Giorgino T. PLUMED-GUI: an environment for the interactive development of molecular dynamics analysis and biasing scripts. *Comput Phys Commun*, 2014, **185**(3): 1109-1114
- [34] Morris G M, Huey R, Lindstrom W, *et al.* AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J Comput Chem*, 2009, **30**(16): 2785-2791
- [35] Yuan S G, Chan H C S, Hu Z Q. Using PyMOL as a platform for computational drug design. *Wiley Interdiscip Rev Comput Mol Sci*, 2017, **7**(2): 10
- [36] Del Mundo I M A, Sifers K E, Fountain M A, *et al.* Structural basis for bifunctional zinc(II) macrocyclic complex recognition of thymine bulges in DNA. *Inorg Chem*, 2012, **51**(9): 5444-5457
- [37] Lim K W, Phan A T. Structural basis of DNA quadruplex-duplex junction formation. *Angew Chem Int Ed Engl*, 2013, **52**(33): 8566-8569
- [38] Wilton D J, Ghosh M, Chary K V A, *et al.* Structural change in a B-DNA helix with hydrostatic pressure. *Nucleic Acids Res*, 2008, **36**(12): 4032-4037
- [39] Ghosh M. Structural characterisation of a uracil containing hairpin DNA by NMR and molecular dynamics. *Nucleic Acids Res*, 1999, **27**(19): 3938-3944
- [40] Shiflett P R, Taylor-McCabe K J, Michaczyk R, *et al.* Structural studies on the hairpins at the 3' untranslated region of an anthrax toxin gene. *Biochemistry*, 2003, **42**(20): 6078-6089
- [41] Ghosh M, Rumpal N, Varshney U, *et al.* Structural basis for poor uracil excision from hairpin DNA. *Eur J Biochem*, 2002, **269**(7): 1886-1894
- [42] Ulyanov N B, Bauer W R, James T L. High-resolution NMR

- structure of an AT-rich DNA sequence. *J Biomol NMR*, 2002, **22**(3): 265-280
- [43] Hazel P, Parkinson G N, Neidle S. Topology variation and loop structural homology in crystal and simulated structures of a bimolecular DNA quadruplex. *J Am Chem Soc*, 2006, **128**(16): 5480-5487
- [44] Baouendi M, Cognet J A H, Ferreira C S M, *et al.* Solution structure of a truncated anti-MUC1 DNA aptamer determined by mesoscale modeling and NMR. *FEBS J*, 2012, **279**(3): 479-490
- [45] Padrta P, Stefl R, Kralik L, *et al.* Refinement of d(GCGAAGC) hairpin structure using one- and two-bond residual dipolar couplings. *J Biomol NMR*, 2002, **24**(1): 1-14
- [46] Chin K H, Chou S H. Sheared-type ganti·csyn base-pair: a unique d(GXC) loop closure motif. *J Mol Biol*, 2003, **329**(2): 351-361
- [47] Weisenseel J P, Reddy G R, Marnett L J, *et al.* Structure of the 1,N-2-propanodeoxyguanosine adduct in a three-base DNA hairpin loop derived from a palindrome in the *Salmonella typhimurium* hisD3052 gene. *Chem Res Toxicol*, 2002, **15**(2): 140-152
- [48] Zhu L, Chou S H, Reid B R. A single G-to-C change causes human centromere TGGAA repeats to fold back into hairpins. *Proc Natl Acad Sci USA*, 1996, **93**(22): 12159-12164
- [49] Van Dongen M J P, Mooren M M W, Willems E F A, *et al.* Structural features of the DNA hairpin d(ATCCTA-GTTA-TAGGAT): formation of a G-A base pair in the loop. *Nucleic Acids Res*, 1997, **25**(8): 1537-1547
- [50] Ghosh M, Vinay Kumar N, Varshney U, *et al.* Structural basis for uracil DNA glycosylase interaction with uracil: NMR study. *Nucleic Acids Res*, 2000, **28**(9): 1906-1912
- [51] Wu J H, Chang C, Pei J M, *et al.* Solution structure of E2F binding DNA fragment GCGCGAAAC-T-GTTTCGCGC studied by molecular dynamics simulation and two dimensional NMR experiment. Protein Data Bank, 2000. doi: 10.2210/pdb1ECU/pdb
- [52] Johnson P E, Turner R B, Wu Z R, *et al.* A mechanism for plus-strand transfer enhancement by the HIV-1 nucleocapsid protein during reverse transcription. *Biochemistry*, 2000, **39**(31): 9084-9091
- [53] Santini G P H, Cognet J A H, Xu D, *et al.* Nucleic acid folding determined by mesoscale modeling and NMR spectroscopy: solution structure of d(GC GAA GC). *J Phys Chem B*, 2009, **113**(19): 6881-6893
- [54] Zhu L, Chou S H, Xu J, *et al.* Structure of a single-cytidine hairpin loop formed by the DNA triplet GCA. *Nat Struct Mol Biol*, 1995, **2**(11): 1012-1017
- [55] Chou S H, Zhu L, Gao Z, *et al.* Hairpin loops consisting of single adenine residues closed by sheared A·A and G·G pairs formed by the DNA triplets AAA and GAG: solution structure of the d(GTACAAAGTAC) hairpin. *J Mol Biol*, 1996, **264**(5): 981-1001
- [56] Burley S K, Berman H M, Bhikadiya C, *et al.* Protein Data Bank: the single global archive for 3D macromolecular structure data. *Nucleic Acids Res*, 2019, **47**(D1): D520-D528
- [57] Carothers J M, Oestreich S C, Davis J H, *et al.* Informational complexity and functional activity of RNA structures. *J Am Chem Soc*, 2004, **126**(16): 5130-5137
- [58] 刘继红, 李淑芳, 尹海燕, 等. 核酸适配体-靶标复合物解离常数测定方法研究进展. *食品安全质量检测学报*. 2016, **7**(3): 1210-1219
- Liu J H, Li S F, Yin H Y, *et al.* *Journal of Food Safety & Quality*. 2016, **7**(3): 1210-1219
- [59] De Rache A, Mergny J L. Assessment of selectivity of G-quadruplex ligands *via* an optimised FRET melting assay. *Biochimie*, 2015, **115**(1): 194-202

Optimization of Progesterone Aptamers *in silico* Based on Genetic Algorithm*

TANG Chun-Hua, JIANG Han-Bing, YANG Jie, LU Xiao-Ling, CHEN Mei-Lun,
WEI Zheng, Liu Yi-Jie, YU Peng**

(College of Pharmaceutical Sciences, Central South University, Changsha 410013, China)

Abstract Objective The purpose of this study is to improve the affinity and selectivity of aptamers of progesterone (P4). **Methods** The computer-aided optimization strategy (*in silico* maturation, ISM) based on genetic algorithm (GA) was used. The 4 rounds of GA (including crossover mutation, single point mutation and double point mutation operations) were performed to construct the initial library and G1, G2 and G3 generation ssDNA oligonucleotides as new candidate aptamer libraries. The candidate aptamers were isolated and analyzed by molecular docking, and continuously optimized using an iterative strategy. In addition, a new strategy is proposed to predict the tertiary structure of ssDNA more accurately. Their secondary and tertiary structures were modeled by Mfold and RNA composer respectively. Output PDB files were modified from RNA to DNA in the Discovery Studio software. Finally, the structure is energy minimized by Molecular Operating Environment. **Results** In the second generation, P4S-0 sequence was optimized in the local search space, and P4G1-14, P4G2-20, P4G1-6, P4G1-7 and P4G2-14 were selected as the best aptamers for P4. AuNPs colorimetric method was applied to verify the affinity of the optimized aptamer. And then a fluorescence method based on the structure switch of the aptamer was constructed to determine the dissociation constant (K_D) of the aptamer, and the selectivity of the aptamer (for bisphenol A, estradiol, testosterone and cortisol) was evaluated. **Conclusion** The aptamer optimized by ISM has a greater affinity for P4 than the original aptamer, and still retains the selectivity for recognizing molecules with similar structure.

Key words genetic algorithm, aptamer, affinity, molecular docking

DOI: 10.16476/j.pibb.2023.0021

* This work was supported by grants from the Fundamental Research Funds for the Central Universities of Central South University (2021zzts1002) and the No.2 (2022) Natural Science Foundation of Hunan Foundation of China (2022JJ80105).

** Corresponding author.

Tel: 86-731-82650329, E-mail: peng.yu@csu.edu.cn

Received: January 16, 2023 Accepted: May 4, 2023