

蛋白质 - 核酸对接方法研究进展 *

刘董敏¹ 常 珊¹ 胡建平² 田绪红^{1△}

(1 华南农业大学信息学院 广东 广州 510642 2 乐山师范学院化学与生命科学学院 四川 乐山 614004)

摘要 :分子对接技术作为预测蛋白质 - 核酸复合物结构的有效方法 ,为研究在生物学过程中蛋白质 - 核酸的相互作用提供了重要的工具。本文首先分析了当前蛋白质 - 核酸对接研究中的主要困难 ,例如构象变化和核糖磷酸骨架的带电性问题。然后从构象搜索、打分函数、柔性策略三个方面比较和总结了蛋白质 - 核酸对接中主要的计算方法。最后回顾了蛋白质 - 核酸对接计算模型的应用 ,并对未来的工作进行了展望。

关键词 :蛋白质 - 核酸对接 ;构象搜索 ;打分函数 ;柔性对接

中图分类号 :Q615 ,Q75 **文献标识码** :A **文章编号** :1673-6273(2012)05-979-05

Progress in the Methods of Protein-nucleic Acid Docking*

LIU Dong-min¹, CHANG Shan¹, HU Jian-ping², TIAN Xu-hong^{1△}

(1 College of Informatics, South China Agricultural University, Guangzhou 510642, China;

2 College of Chemistry and Life Science, Leshan Teacher's College, Sichuan Leshan 614000, China)

ABSTRACT: As an effective approach on the prediction of protein-nucleic acid complexes structures, molecular docking technology plays a central role in the study of protein-nucleic acid interactions during biological processes. In this article, we firstly analyzed the main difficulties in the protein-nucleic acid docking, such as the conformational changes and the highly charged sugar-phosphate backbone. Meanwhile, we compared and concluded the major computational methods of the conformational searching, scoring function and flexible strategy in the protein-nucleic acid docking. Finally, the existing applications of the protein-nucleic acid docking were given and the prospects of the future study were made.

Key words: Protein-nuclei acid docking; Conformational search; Scoring function; Flexible docking

Chinese Library Classification(CLC): Q615 ,Q75 **Document code**: A

Article ID:1673-6273(2012)05-979-05

前言

蛋白质和核酸是构成生命体最为重要的两类生物大分子 ,蛋白质与核酸的相互作用是许多细胞功能的核心 ,它在基因表达调控、蛋白质翻译和细胞分裂等细胞过程中发挥着极其重要的作用^[1,2]。研究蛋白质 - 核酸复合物的相互作用不但可以帮助我们洞悉核酸在生物学过程中所发挥的基础作用 ,而且还可以辅助我们设计针对核酸的药物^[3]。但是 ,通过实验方法直接测定蛋白质 - 核酸复合物结构仍然相当困难^[4,5] ,如图 1 ,截止到 2011 年 6 月 ,蛋白质数据库(Protein Data Bank, PDB)中数据总数已超过 70000 个 ,而其中蛋白质 - 核酸复合物结构却不足 3500 个 ,如果考虑结构之间的同源性 ,则保留下来的结构就更少了。因此从理论模型角度研究蛋白质 - 核酸的相互作用特点 ,发展精确预测蛋白质 - 核酸复合物结构的计算方法具有重要的学术价值。在复合物结构预测的模拟方法中 ,分子对接已成为其中非常重要和应用最广的方法之一 ,并可为实验工作提供有益的参考依据^[6,7]。

相对于蛋白质 - 蛋白质对接 ,蛋白质 - 核酸对接的研究进

展要落后许多 ,它的主要困难在于 :第一 ,缺乏蛋白质的氨基酸和核酸的碱基对之间有关识别模式的信息^[8]。在蛋白质 - 蛋白质对接中 ,对接双方作用面上的大量信息可以引导对接并且限制构象空间的搜索^[9]。然而到现在为止 ,核酸在对接过程中作用面上的识别信息仍然非常模糊。第二 ,DNA 和 RNA 的带电性更加复杂 ,尤其是核糖磷酸盐骨架的带电性问题大大增加了评估系统静电能量稳定性的难度。第三 ,核酸在结合过程中发生的构象变化比较大。在蛋白质跟核酸的结合过程中 ,除了蛋白质的边链会发生构象变化外 ,DNA 的螺旋结构也有可能发生全局的构象变化 ,例如弯曲和解螺旋^[10]。

1 蛋白质 - 核酸对接方法概述

针对特定的蛋白质 - 核酸体系 ,国内外一些研究小组在相关研究领域展开了积极的研究工作 ,并取得了一定成果。Knegtel 和合作者开发的蒙特卡洛模拟程序模拟了抑制剂和 DNA 复合物的对接^[11] ,在限制 DNA 的构象变化条件下考虑了蛋白质侧链的柔性。Rohs 等人也用一种全原子蒙特卡洛模拟方法进行 DNA 和药物对接的研究 ,在没有任何先导结合位点

* 基金项目 :国家自然科学基金项目(11047183) ,广东高校优秀青年创新人才培育项目(LYM10042) ,广东省科技计划项目(2010B080701070)

作者简介 :刘董敏(1987-) ,女 ,硕士 ,主要研究方向 :计算机辅助药物设计 ,分子模拟与生物信息学。

△通讯作者 :田绪红 ,电话 :020-85288302 ,E-mail: tianxuhong@scau.edu.cn。

(收稿日期 :2011-07-07 接受日期 :2011-07-31)

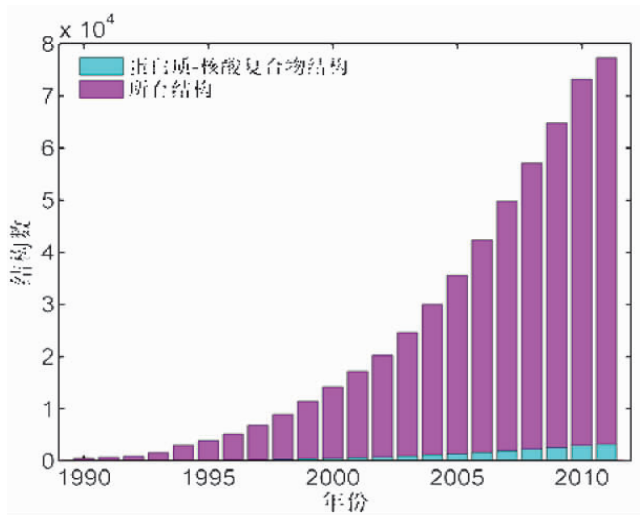


图 1 PDB 库数据统计
Fig.1 The statistics of Protein Data Bank

选择的情况下考察 DNA 和药物的结合位点和对接模式，并加入分子柔性的处理^[12]。Sternberg 等人将实验室开发的 FTDOCK 应用在蛋白质 -DNA 的对接上^[13]，通过为 DNA 分配特殊的电场扩大 DNA 螺旋槽中化学基团局部电荷的作用，并且引入氨基酸和核苷酸之间的经验势能来为他们的相互作用打分^[10]。Liu 等人研究的半柔性对接算法，使用基于结构的预测方法来克服基于序列的预测算法中的假阳性问题^[14]。Van Dijk 和合作者通过改进 HADDOCK 算法实现了蛋白质 -DNA 的柔性对接^[15]，该算法用非结构化的实验数据启动对接，用综合的打分函数筛选出得分最好的结构，随后用分子动力学模拟的方法考察蛋白

质和 DNA 在对接面的构象变化，寻找具有最低能量的构象。De Luca 等采用改进的 ESCHER^[16]实现蛋白质 -DNA 对接，通过设置一张特殊的电荷表来解决 DNA 电荷的计算问题^[17]。

相比起蛋白质 -DNA 的对接研究，蛋白质 -RNA 对接的研究更加困难。在蛋白质 -RNA 的对接中，无论是结构模式还是结合位点都比蛋白质 -DNA 的复杂。最近，随着蛋白质 -RNA 复合物的结构逐步增多，一些研究人员针对蛋白质和 RNA 相互作用的特性，设计了新的对接计算模型，并取得一定的成果。在对蛋白质 -RNA 相互作用的数据进行统计分析后，Ellis 等人发现，在蛋白质 -RNA 复合物中，范德华能量要比氢键能量发挥更大的作用，在蛋白质 -RNA 的结合位点上，RNA 的鸟嘌呤更倾向于和苯丙氨酸、苏氨酸、酪氨酸、天冬氨酸和精氨酸发生接触^[18]。Guilbert 等人设计了一种新的算法 MORDOR，通过在受体表面放置配体柔性探针，解决 RNA 在对接过程中的柔性问题^[19]。Lang 等人在 DOCK 的基础上，综合 GB/SA (Generalized Born/Surface Area) 和 PB/SA (Poisson-Boltzmann / Surface Area) 技术对对接后的构象重新打分，使其适用于蛋白质 -RNA 的对接^[20]。Foloppe 等人发展了一种快速经验函数对 RNA 和配体之间的相互作用进行打分，并把它应用在 RiboDock 中^[21]。Zhao^[22]和 Pfeffer^[23]的实验室发展了两种基于知识的打分函数，随后用实验和计算值之间的相关函数预测蛋白质 -RNA 的相互作用。

下面本文将针对蛋白质 - 核酸对接中的构象搜索、打分函数和分子柔性三个方面进行讨论与总结。表 1 列出了目前使用较为广泛的一些蛋白质 - 核酸的分子对接程序，并对基本情况进行了简单的介绍。

表 1 蛋白质 - 核酸对接程序
Table1. The protein-nucleic acids docking programs

参考程序 Reference procedure	研发单位(人) Explorer	适用范围 Applicability	算法及特点 Algorithm and features
HADDOCK ^[9]	Utrecht Univ. (Bonvin)	Protein-DNA	基于实验数据(比如 NMR 的化学位移、点突变等) 对接程序，考虑了对接柔性 Based on experimental data (such as NMR chemical shift, point mutation and so on) , docking procedure, takes dock flexibility into account
FTDock ^[10]	BioMol. Mod. Can. Res. UK (Sternberg)	Protein-DNA	FFT 算法，几何匹配，为 DNA 分配特殊电场 FFT algorithm, geometry matching, distribute specially electric field for DNA
ESCHER ^[16]	R.T.V. Univ. Bio. Dep. Italy(Ausiello)	Protein-DNA	几何匹配，氢键作用，静电作用，为 DNA 设置一张特殊的电荷表 Geometry matching, hydrogen bonding, static action, set a piece of specific charge table for DNA
DOT ^[24]	S. Diego Super-comput. Cen. (Mandell)	Protein-DNA	FFT 搜索，打分函数只有范德华和静电项 FFT searching, scoring function only vanderwaals and static items
HADDOCK ^[9]	Utrecht Univ. (Bonvin)	Protein-DNA	MC 搜索，在打分函数中加入基于 DNA 序列特性的构象能量函数 MC searching, acceded constellation energy function based on DNA sequence characteristics in scoring function
FTDock ^[10]	BioMol. Mod. Can. Res. UK (Sternberg)	Protein-RNA	RMSD 驱动能量最小化，在受体表面放置配体柔性探针 RMSD drive energy minimize, place ligand flexibility probe on receptor surface
ESCHER ^[16]	R.T.V. Univ. Bio. Dep. Italy(Ausiello)	Protein-RNA	综合 GB/SA 和 PB/SA 技术设计打分函数 Design scoring function synthesize GB / SA and PB / SA technology
GOLD\GLIDE ^[27]	S. Diego Super-comput. Cen. (Mandell)	Protein-RNA	GA 搜索，考虑边链柔性，高通量数据库筛选 GA searching, consider side chain flexibility, high flux reactor database screen

1.1 构象搜索

在蛋白质-核酸对接中,目前常用的构象搜索方法主要有快速傅里叶变换算法、蒙特卡洛算法和遗传算法。

在 FTDOCK 中,DNA 作为固定分子,蛋白质围绕着 DNA 进行平移和旋转,从而搜索形状和静电作用互补的对接复合物。为了加速搜索的进度,引入 Katchalski-Katzir 等人的快速傅里叶相关算法^[28]。对于蛋白质-DNA 对接,静电的计算至关重要。这里针对 DNA 使用了特殊的电场从而扩大 DNA 螺旋槽中化学基团局部电荷的作用,减弱磷酸的电荷作用。并且计算中仅考虑 DNA 上带电原子附近 2Å 距离以内格点的静电势。对于蛋白质,电荷被分配到主链原子以及 Asp、Glu、Arg、Lys、Asn、Gln、His 残基上^[10]。

Liu 等研究的半柔性对接算法使用蒙特卡洛模拟方法进行构象搜索^[14],该方法把 DNA 看作是刚性体,最多经过 200 次的蒙特卡洛模拟就可以把所有构象搜索出来。Rohs 等用蒙特卡洛变量来定义药物和 DNA 之间的相互运动和内部柔性^[12]。每个核苷酸都用刚性体的平移和旋转表示,并以变量定义 DNA 骨架各部分的构象变化以考虑对接的柔性。在模拟的每次循环中,核苷酸和离子的蒙特卡洛变量都发生改变,药物和离子在 DNA 螺旋结构附近自由移动,直至系统能量达到平稳的最低值。

遗传算法模仿自然界生物的遗传和进化,对多极值优化问题的解集合进行选择、杂交和突变操作,使得解集合不断优化,直至收敛于全局最优解或近似全局最优解。把遗传算法用到分子对接中,受体一般不动,而配体相对于受体的一系列平移和旋转对应着基因,并把受体和配体相互作用能的高低来表示配体的适应度。Li 等用 GOLD 和 Glide 程序去识别不同 RNA 结构的同源配体,利用遗传算法进行构象搜索,并考察对接过程中的边链柔性 and 局部骨架的移动^[27]。

1.2 打分函数

蛋白质-核酸对接的打分函数主要分为三类:基于物理的打分函数、基于经验的打分函数和基于知识的打分函数。针对核酸带电的特性,蛋白质-核酸对接的打分函数会对打分项进行改良,并且综合以上打分函数的方法,以发展出新的打分方法以正确地反映蛋白质-核酸复合物的特性。

基于物理的打分函数是利用“热力学方程”进行自由能预测并打分,具体来说是基于力场计算结合自由能。这种方法已被 DOCK^[29]等对接方法所采纳。这种打分函数的计算十分耗时,因此这类方法多采用预计算的网格近似法来进行加速。如 Lang 等人在 DOCK 的基础上,综合 GB/SA (Generalized Born / Surface Area) 和 PB/SA (Poisson-Boltzmann / Surface Area) 等方法计算对接构象的自由能^[29]。

基于经验函数所制定的打分函数考虑了多种因素的贡献。HADDOCK 中使用的打分函数通过分别计算静电能,范德华能量,溶剂化自由能, AIR (Ambiguous Interaction Restraints) 能量和溶剂可达表面,并且赋予他们不同的权值然后叠加得出。Morozov 等在用结构化模型预测蛋白质-DNA 结合特性时提出了一种由蛋白质-DNA 相互作用能量和 DNA 重构能量组成的打分函数^[25]。它在 Rosetta 蛋白质-氨基酸模型^[25,29,30]的基础上,加入了 DNA 序列特性的构象能量,用伦纳德-琼斯势的线

性组合模拟蛋白质-DNA 的交互作用,用依赖于方向的氢键势描述氨基酸边链之间和氨基酸边链与 DNA 底物氢键之间的相互作用,用动力学模型^[30]模拟静电和溶剂化作用。而 DNA 依赖于序列的构象能量则采用 Olson 等人发展的高效谐振势模型^[32]计算,它通过把所有碱基成对势和碱基梯度能量相加和计算得出。

基于知识的打分函数是利用波尔兹曼分布对已有的蛋白质结构数据库分析而获得,具体来说是分析实验测得的复合物结构,从中提取相互作用规律。Xu 等人通过对已知的蛋白质-配体复合物进行统计分析,设计了一种基于 DFIRE 的能量函数^[33],并把其应用在蛋白质-DNA 复合物的对接中。Liu 等研究的半柔性对接算法的打分函数由对接亲和力、范德华能量和限制能量三部分组成^[14]。对接亲和力用一种基于知识的氨基酸-核苷酸势能^[34]来计算,从而把之前一直被人们忽略的由于 DN-A 结构变形和局部结构交互作用所引起的能量变化考虑在内。Pfeffer 等人研究的 DrugScoreRNA 是在统计 670 种核酸-配体和核酸-蛋白质复合物的晶体结构后,设计的一种依赖于距离的成对势打分函数。最近,吴建盛等人也通过统计分析 PDB 库里面的 1,289 个蛋白质-DNA 和 413 个蛋白质-RNA 复合物结构,研究了蛋白质-核酸复合物界面的氨基酸与核苷酸偏好性^[35]。

1.3 柔性策略

核酸在与蛋白质结合的过程中其形状和结合界面上的结构会出现较大的构象变化,这成为核酸-蛋白质对接工作中的难点。核酸的柔性主要体现在由于碱基对和磷酸糖骨架的构象变化所引起的弯曲和扭曲上。在 Liu 等的蛋白质-核酸对接工作中,采用多构象模型解决柔性对接的问题^[14]。他们搜索所有已知的 DNA 结构并且把它们分类。随后在每一类结构中选取一个代表对接到蛋白质上,并为对接后的复合物打分,最后筛选出最好的对接结果。而 Lindahl 等则提出用低频率标准模型振幅优化方法改良对接之后的蛋白质-DNA 复合物^[36]。它以弹性网络模型作为工具,计算全原子标准模型来模拟大分子复合物在形成过程中的结构柔性问题,并在 5-10 的最低频率标准模型方向上改进受体和配体的非键能量。经测试证明该方法即使在不增加额外的实验数据基础上,都可以达到优化对接后结构的目标。

研究蛋白质-核酸复合物结合界面的特性,往往可以帮助我们发现蛋白质-核酸对接柔性的规律。Ellis 等通过对 26 种蛋白质-DNA 复合物结合界面的分析,把他们的结合模式分成三类:单头式的、双头式的和包围式的。他们实验室还对蛋白质-RNA 复合物在对接过程中蛋白质所发生的构象变化进行了考察,发现在不同情况下蛋白质残基在结合位点、非结合位点上都有可能发生不同程度的构象变化^[18]。Terribilini 的实验室在分析总结了蛋白质-RNA 复合物的结构后,发现对于有结合槽的蛋白质, α -helix、10-helix、 β -ribbon、irregular loop 结构会结合到 RNA 螺旋结构的凹槽位里面,为了适应底物的序列特性和槽位的形状或者尺寸, RNA 的螺旋结构有时还会发生弯曲。而对于用 β -sheet 表面生成检测 RNA 底物单螺旋结构口袋的蛋白质,则通过限制底物在一个合适的方向上来间接地结合那些完全非结构化的 RNA。另外,热力学研究表明,蛋白质-RNA 结合特性通常表现为某些因子的函数,包括氢键作用、非极性

接触和蛋白质-RNA 结合表面的互相适应^[37]。

2 蛋白质-核酸对接应用

蛋白质-核酸相互作用与识别在许多细胞活动中发挥关键作用并且与多种疾病密切相关^[38]。生命信息主要存储在脱氧核糖核酸(DNA)中,而蛋白质是生命活动的具体执行者。DNA 转录成核糖核酸(RNA)然后翻译成蛋白质。蛋白质又作为酶参与细胞内的各种生化反应,控制着遗传信息的复制、转录与翻译。很多威胁人类健康的重大疾病,如癌症、免疫系统疾病、心血管疾病等往往都与蛋白质-核酸之间的特异性识别有着密切的联系。因此,研究蛋白质-核酸的相互作用与识别是人们了解细胞活动的机理进而解开生命奥秘的关键所在,具有重要的科学意义与应用前景^[2]。目前,一些研究小组已经把蛋白质-核酸对接方法应用在蛋白质-核酸的相互作用与识别的研究上。

2.1 蛋白质-DNA 对接的应用

Roberts 等人把快速大分子对接程序 DOT^[24]应用在 3 个翼状螺旋转录因子和 DNA 的对接上^[39]。这 3 个翼状螺旋转录因子具有相似的 DNA 结合基序,但是结合的方法却是不同的,他们与线性 B 型 DNA 模型进行对接时,都实现了:第一,分辨出了转录因子和 DNA 结合的不同方法;第二,识别出了蛋白质-DNA 的结合位点和 DNA 的结合方向;第三,在前三位排名的方案中找到最少一个正确的结合模式。Hu 等人采用 DOT 程序获得整合酶二聚体与病毒 DNA 的结合模式,用 MD 模拟和统计方法研究了整合酶二聚体和 DNA 结合后的运动模式、相关运动和病毒 DNA 构象的变化,并简要分析了水分子在整合酶和病毒 DNA 识别中的重要作用^[40,41]。Adesokan 等也把 DOT 应用在 HIV-1 整合酶和 DNA 病毒的对接中^[42]。在对接结果的构象中,发现了脱氧腺苷 3 端的羟基有向二价的金属离子闭合的趋势,这与实验上催化的过程相吻合。另外,他们认为在对接结果中找到的第二个连接的 DNA 结合位点是非特异性宿主 DNA 结合位点。同时,De Luca 等也用改进的 ESCHER 程序对 HIV-1 整合酶和 DNA 病毒进行对接,模拟结果重现了突变和交叉耦合的实验结果,证明了该对接方法正确地预测了蛋白质-DNA 相互作用中的底物特性^[17]。Fan 等用 DOT 模拟连接组蛋白 H5 和核小体的不同的对接位点,以考察连接组蛋白家族的不同结构导致的功能差别^[43]。Liu 等把他们的半柔性对接程序应用在 141 个转录因子-DNA 复合物的对接中,同时对大肠杆菌和铁的摄取调节器复合物的对接模式进行预测,均取得了非常好的对接结果^[14]。Aloy 把 FTDOCK 应用在识别 DNA 的转录阻抑物的结构模型中,解释了为什么需要某些确定的碱基序列来识别阻抑物,以及哪些阻抑物能与 DNA 发生作用^[44]。

2.2 蛋白质-RNA 对接的应用

TAT 蛋白和 TAR RNA 的结合对于 HIV-1 病毒的复制来说是必不可少的。Filikov 等人用 DOCK 和 ICM 对接模型成功地对 HIV-1 TAR RNA 复合物进行了虚拟筛选^[45]。Daldrop 等人用一种新的 RNA 配体对接方法寻找嘌呤核开关的合适配体^[46]。通过虚拟筛选,他们在大分子领域里找到了四个具有结合亲和力的新配体。人类的核蛋白端粒酶已经被证实为抗癌药物的靶标,而 hTR-P2b 是端粒酶 RNA 的活性所在。Pinto 等人用

MORDOR 对接程序筛选出端粒酶 RNA 的合适配体,并且发现更多的配体倾向于与 hTR-P2b RNA 结合,而不是与 A-site rRNA 结合^[47]。Holt 等成功地把 Autodock 和 Surflex 应用到道诺霉素和远霉素与核酸的对接中,产生了与晶体结构相近的复合物^[48]。Varani 等人也尝试把 DOCK 和 AutoDOCK 应用在 RNA 药物对接上^[49]。Barbault 等人用对接的方法研究 16 种新的氨基糖苷类成分^[50],为了解决 RNA 的柔性问题,对接结果进行小分子的动态模拟以提高和实验数据之间的相关性,从而建立结构活性的关系。

3 结语与展望

蛋白质-核酸的相互作用可使我们深入洞悉许多基本的生物学过程,而用计算的方法对蛋白质-核酸的复合物结构进行预测可以帮助我们了解这些系统的识别机制,因此,研究蛋白质-核酸的对接工作具有非常重要的意义。目前,相关研究工作的难点还是主要集中在核苷酸本身的强带电性以及复合物中构象的多变性问题上。但是,已经有不少的研究人员在针对某类蛋白质-核酸复合物体系的结合特性进行分析后,把蛋白质-配体与蛋白质-蛋白质对接的方法移植到蛋白质-核酸对接上。随着 CAPRI (Critical Assessment of Prediction of Interactions)复合物结构预测竞赛^[51,52]的快速发展,分子对接的柔性问题也成为了研究者高度关注和重点攻克的问题,这对解决蛋白质-核酸对接中的构象多变性问题具有极大的推动作用。相信将来蛋白质-核酸相互作用的研究会取得更大的发展,成为生物大分子结构-功能关系研究的重要方向。

参考文献(References)

- [1] Berger JM, Muller CW. Protein-nucleic acid interactions: some assembly required - Editorial overview[J]. Current Opinion in Structural Biology, 2007, 17(1): 77-79
- [2] Bujalowski W. Thermodynamic and kinetic methods of analyses of protein-nucleic acid interactions. From simpler to more complex systems [J]. Chemical Reviews, 2006, 106(2): 556-606
- [3] Gan JH, Sheng J, Huang Z. Chemical and structural biology of nucleic acids and protein-nucleic acid complexes for novel drug discovery[J]. Science China-Chemistry, 2011, 54(1): 3-23
- [4] Berman HM. The Protein Data Bank: a historical perspective [J]. Acta Crystallographica Section A, 2008, 64: 88-95
- [5] Berman HM, Battistuz T, Bhat TN, et al. The Protein Data Bank[J]. Acta Crystallographica Section D-Biological Crystallography, 2002, 58: 899-907
- [6] Vajda S, Kozakov D. Convergence and combination of methods in protein-protein docking [J]. Current Opinion in Structural Biology, 2009, 19(2): 164-170
- [7] Ritchie DW. Recent progress and future directions in protein-protein docking[J]. Current Protein & Peptide Science, 2008, 9(1): 1-15
- [8] Gao M, Skolnick J. From Nonspecific DNA-Protein Encounter Complexes to the Prediction of DNA-Protein Interactions[J]. Plos Computational Biology, 2009, 5(4): 55-68
- [9] van Dijk M, van Dijk ADJ, Hsu V, et al. Information-driven protein-DNA docking using HADDOCK: it is a matter of flexibility [J]. Nucleic Acids Research, 2006, 34(11): 3317-3325

- [10] Sternberg MJE, Gabb HA, Jackson RM. Predictive docking of protein-protein and protein-DNA complexes[J]. *Current Opinion in Structural Biology*, 1998, 8(2): 250-256
- [11] Knegtel RMA, Antoon J, Rullmann C, et al. MONTY: a Monte Carlo approach to protein-DNA recognition[J]. *Journal of Molecular Biology*, 1994, 235(1): 318-324
- [12] Rohs R, Bloch I, Sklenar H, et al. Molecular flexibility in ab initio drug docking to DNA: binding-site and binding-mode transitions in all-atom Monte Carlo simulations[J]. *Nucleic Acids Research*, 2005, 33(22): 7048-7057
- [13] Smith GR, Sternberg MJE. Prediction of protein-protein interactions by docking methods[J]. *Current Opinion in Structural Biology*, 2002, 12(1): 28-35
- [14] Liu ZJ, Guo JT, Li T, et al. Structure-based prediction of transcription factor binding sites using a protein-DNA docking approach[J]. *Proteins-Structure Function and Bioinformatics*, 2008, 72(4): 1114-1124
- [15] van Dijk M, Bonvin A. Pushing the limits of what is achievable in protein-DNA docking: benchmarking HADDOCK's performance[J]. *Nucleic Acids Research*, 2010, 38(17): 5634-5647
- [16] Ausiello G, Cesareni G, Helmer Citterich M. ESCHER: A new docking procedure applied to the reconstruction of protein tertiary structure [J]. *Proteins-Structure Function and Genetics*, 1997, 28(4): 556-567
- [17] De Luca L, Pedretti A, Vistoli G, et al. Analysis of the full-length integrase - DNA complex by a modified approach for DNA docking [J]. *Biochemical and Biophysical Research Communications*, 2003, 310(4): 1083-1088
- [18] Ellis JJ, Jones S. Evaluating conformational changes in protein structures binding RNA[J]. *Proteins-Structure Function and Bioinformatics*, 2008, 70(4): 1518-1526
- [19] Guilbert C, James TL. Docking to RNA via root-mean-square-deviation-driven energy minimization with flexible ligands and flexible targets [J]. *Journal of Chemical Information and Modeling*, 2008, 48(6): 1257-1268
- [20] Lang PT, Brozell SR, Mukherjee S, et al. DOCK 6: Combining techniques to model RNA-small molecule complexes[J]. *Rna-a Publication of the Rna Society*, 2009, 15(6): 1219-1230
- [21] Foloppe N, Chen IJ, Davis B, et al. A structure-based strategy to identify new molecular scaffolds targeting the bacterial ribosomal A-site [J]. *Bioorganic & Medicinal Chemistry*, 2004, 12(5): 935-947
- [22] Zhao XY, Liu XF, Wang YY, et al. An improved PMF scoring function for universally predicting the interactions of a ligand with protein, DNA, and RNA [J]. *Journal of Chemical Information and Modeling*, 2008, 48(7): 1438-1447
- [23] Pfeffer P, Gohlke H. DrugScore (RNA) - Knowledge-based scoring function to predict RNA-ligand interactions [J]. *Journal of Chemical Information and Modeling*, 2007, 47(5): 1868-1876
- [24] Mandell JG, Roberts VA, Pique ME, et al. Protein docking using continuum electrostatics and geometric fit[J]. *Protein Engineering*, 2001, 14(2): 105-113
- [25] Morozov AV, Havranek JJ, Baker D, et al. Protein-DNA binding specificity predictions with structural models [J]. *Nucleic Acids Research*, 2005, 33(18): 5781-5798
- [26] Gray JJ, Moughon S, Wang C, et al. Protein-protein docking with simultaneous optimization of rigid-body displacement and side-chain conformations[J]. *Journal of Molecular Biology*, 2003, 331(1): 281-299
- [27] Li YZ, Shen J, Sun XG, et al. Accuracy Assessment of Protein-Based Docking Programs against RNA Targets[J]. *Journal of Chemical Information and Modeling*, 2010, 50(6): 1134-1146
- [28] Katchalski-katzir E, Shariv I, Eisenstein M, et al. Molecular surface recognition: Determination of geometric fit between proteins and their ligands by correlation techniques [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 1992, 89(6): 2195-2199
- [29] Havranek JJ, Baker D. Motif-directed flexible backbone design of functional interactions[J]. *Protein Science*, 2009, 18(6): 1293-1305
- [30] Ashworth J, Baker D. Assessment of the optimization of affinity and specificity at protein-DNA interfaces [J]. *Nucleic Acids Research*, 2009, 37(10): e73
- [31] Mongan J, Simmerling C, McCammon JA, et al. Generalized Born model with a simple, robust molecular volume correction[J]. *Journal of Chemical Theory and Computation*, 2007, 3(1): 156-169
- [32] Zheng GH, Czaplá L, Srinivasan AR, et al. How stiff is DNA?[J]. *Physical Chemistry Chemical Physics*, 2010, 12(6): 1399-1406
- [33] Xu BS, Yang YD, Liang HJ, et al. An all-atom knowledge-based energy function for protein-DNA threading, docking decoy discrimination, and prediction of transcription-factor binding profiles[J]. *Proteins-Structure Function and Bioinformatics*, 2009, 76(3): 718-730
- [34] Liu ZJ, Mao FL, Guo JT, et al. Quantitative evaluation of protein-DNA interactions using an optimized knowledge-based potential [J]. *Nucleic Acids Research*, 2005, 33(2): 546-558
- [35] 吴建盛, 胡栋, 伍洪涛, 等. 蛋白质-核酸复合物界面氨基酸与核苷酸偏好性分析[J]. *生物物理学报*, 2010, 26(3): 234-244
Wu Jian-sheng, Hu Dong, Wu Hong-tao, et al. Comprehensive Analysis of Propensities of Amino Acids and Nucleotides Usage at Protein-Nucleic Acid Interfaces[J]. *Acta Biophysica Sinica*, 2010, 26(3): 234-244(In Chinese)
- [36] Lindahl E, Delarue M. Refinement of docked protein-ligand and protein-DNA structures using low frequency normal mode amplitude optimization[J]. *Nucleic Acids Research*, 2005, 33(14): 4496-4506
- [37] Terribilini M, Lee JH, Yan CH, et al. Prediction of RNA binding sites in proteins from amino acid sequence [J]. *Rna-a Publication of the Rna Society*, 2006, 12(8): 1450-1462
- [38] von Hippel PH. From "simple" DNA-protein interactions to the macromolecular machines of gene expression[J]. *Annual Review of Biophysics and Biomolecular Structure*, 2007, 36: 79-105
- [39] Roberts VA, Case DA, Tsui V. Predicting interactions of winged-helix transcription factors with DNA[J]. *Proteins-Structure Function and Bioinformatics*, 2004, 57(1): 172-187
- [40] Hu JP, Ke GT, Chang S, et al. Studies on the binding modes of HIV-1 integrase with viral DNA via molecular docking method[J]. *Chemical Journal of Chinese Universities-Chinese*, 2008, 29(7): 1432-1437
- [41] Hu JP, Ke GT, Chang S, et al. Conformational Change of HIV-1 Viral DNA after Binding with Integrase[J]. *Acta Physico-Chimica Sinica*, 2008, 24(10): 1803-1810

(下转第 987 页)

- [12] Jia Rong-chen, Jyhs-HL, Chung Nan-weng, et al. Identification of a novel adhesin-like glycoprotein from *Mycoplasma hyopneumoniae* [J]. *Veterinary Microbiology*, 1998, 62:97-110
- [13] 曹培丽, 李媛, 陈超等. 猪肺炎支原体 P52 蛋白的原核表达及抗血清制备[J]. *中国预防兽医学报*, 2009, 31(7):572-576
- Cao Pei-li, Li Yuan, Chen Chao, et al. Expression of *Mycoplasma hyopneumoniae* P52 protein and preparation of the rabbit antisera against P52 protein [J]. *Chinese Journal of Preventive Veterinary Medicine*, 2009, 31(7):572-576
- [14] Jody W, Cheryl J, Stuart JC, et al. Mhp493 (P216) is a proteolytically processed, cilium and heparin binding protein of *Mycoplasma hyopneumoniae* [J]. *Molecular Microbiology*, 2009, 71(3):566-582
- [15] Lisa MS, Ania TD, Cheryl J, et al. A Processed Multidomain *Mycoplasma hyopneumoniae* adhesin binds Fibronectin, Plasminogen, and Swine Respiratory Cilia [J]. *The American Society for Biochemistry and Molecular Biology*, 2010, 285:33971-33978
- [16] Ania TD, Cheryl J, Minion FC, et al. Repeat regions R1 and R2 in the P97 paralogue Mhp271 of *Mycoplasma hyopneumoniae* bind heparin, fibronectin and porcine cilia [J]. *Molecular Microbiology*, 2010, 78(2):444-458
- [17] Lisa MS, Linda F, Ania TD, et al. Mhp107 is a member of the multifunctional adhesin family of *Mycoplasma hyopneumoniae* [J]. *The American Society for Biochemistry and Molecular Biology*, 2011, 286:10097-10104
- [18] Vijaykumar P, Vincent AF. α -Enolase, a Novel Strong Plasmin(ogen) Binding Protein on the Surface of Pathogenic Streptococci [J]. *The American Society for Biochemistry and Molecular Biology*, 1998, 273(33):14503-14515
- [19] Giron JA, Lange M, Baseman J. Adherence, fibronectin binding, and induction of cytoskeleton reorganization in cultured human cells by *Mycoplasma penetrans* [J]. *Infection and Immunity*, 1996, 64(1):197-208
- [20] Katarzyna D, Joseph MP, Claudia FD. Fibronectin binding protein and host cell tyrosine kinase are required for internalization of *Staphylococcus aureus* by epithelial cells [J]. *Infection and Immunity*, 1999, 67(9):4673-4678
- [21] Gerlinde LS, Alexandre P, Matthias M. Proteins complexed to the P1 adhesin of *Mycoplasma pneumoniae* [J]. *Microbiology*, 2000, 146:741-747
- [22] Dallo SF, Chavoya A, Baseman JB. Characterization of the gene for a 30-kilodalton adhesion-related protein of *Mycoplasma pneumoniae* [J]. *Infection and Immunity*, 1990, 58(12):4163-4165
- [23] Plummer JM, Lazzell A, Baseman JB. Shared epitopes between *Mycoplasma pneumoniae* major adhesin protein P1 and a 140-kilodalton protein of *Mycoplasma genitalium* [J]. *Infection and Immunity*, 1987, 55(1):49-56
- [24] Raul B, Oscar QP, Mario FN. *Mycoplasma genitalium* P140 and P110 cytoadhesins are reciprocally stabilized and required for cell adhesion and terminal-organelle development [J]. *Journal of Bacteriology*, 2006, 188(24):8627-8637
- [25] Markham PF, Kanci A, Whithear KG, et al. A novel mechanism for control of antigenic variation in the haemagglutinin gene family of *Mycoplasma synoviae* [J]. *Molecular Microbiology*, 2000, 35(4):911-923
- [26] 刘劫, 吴移谋. 支原体基因组学研究进展[J]. *中国人兽共患病学报*, 2006, 22(11):1073-1077
- Liu Jie, Wu Yi-mou. *Mycoplasma* genomics research progress [J]. *Chinese Journal of Zoonoses*, 2006, 22(11):1073-1077
- [27] Liu Wei, Feng Zhi-xin, Fang Liu-rong, et al. Complete Genome Sequence of *Mycoplasma hyopneumoniae* Strain 168 [J]. *Journal of Bacteriology*, 2011, 193(4):1016-1017
- [28] 刘茂军, 邵国青, 张映. 猪肺炎支原体不同毒力株黏附因子基因序列分析[J]. *江苏农业学报*, 2010, 26(3):522-525
- Liu Mao-jun, Shao Guo-qing, Zhang Ying. Sequence analysis of adhesin gene of different *Mycoplasma hyopneumoniae* strains [J]. *Jiangsu Journal of Agricultural Sciences*, 2010, 26(3):522-525

(上接第 983 页)

- [42] Adesokan AA, Roberts VA, Lee KW, et al. Prediction of HIV-1 integrase/viral DNA interactions in the catalytic domain by fast molecular docking [J]. *Journal of Medicinal Chemistry*, 2004, 47(4):821-828
- [43] Fan L, Roberts VA. Complex of linker histone H5 with the nucleosome and its implications for chromatin packing [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2006, 103(22):8384-8389
- [44] Aloy P, Moont G, Gabb HA, et al. Modelling repressor proteins docking to DNA [J]. *Proteins: Structure, Function, and Bioinformatics*, 1998, 33(4):535-549
- [45] Filikov AV, Mohan V, Vickers TA, et al. Identification of ligands for RNA targets via structure-based virtual screening: HIV-1 TAR [J]. *Journal of Computer-Aided Molecular Design*, 2000, 14(6):593-610
- [46] Daldrop P, Reyes FE, Robinson DA, et al. Novel Ligands for a Purine Riboswitch Discovered by RNA-Ligand Docking [J]. *Chemistry & Biology*, 2011, 18(3):324-335
- [47] Pinto IG, Guilbert C, Ulyanov NB, et al. Discovery of Ligands for a Novel Target, the Human Telomerase RNA, Based on Flexible-Target Virtual Screening and NMR [J]. *Journal of Medicinal Chemistry*, 2008, 51(22):7205-7215
- [48] Holt PA, Chaires JB, Trent JO. Molecular docking of intercalators and groove-binders to nucleic acids using Autodock and Surflex [J]. *Journal of Chemical Information and Modeling*, 2008, 48(8):1602-1615
- [49] Detering C, Varani G. Validation of automated docking programs for docking and database screening against RNA drug targets [J]. *Journal of Medicinal Chemistry*, 2004, 47(17):4188-4201
- [50] Barbault F, Ren B, Rebehmed J, et al. Flexible computational docking studies of new aminoglycosides targeting RNA 16S bacterial ribosome site [J]. *European Journal of Medicinal Chemistry*, 2008, 43(8):1648-1656.
- [51] Janin J. Protein-protein docking tested in blind predictions: the CAPRI experiment [J]. *Molecular Biosystems*, 2010, 6(12):2351-2362
- [52] Janin J, Henrick K, Moult J, et al. CAPRI: A Critical Assessment of PRedicted Interactions [J]. *Proteins-Structure Function and Bioinformatics*, 2003, 52(1):2-9