

第二十四章 药物分子设计概论

Outline of Molecular Drug Design

第一节 导 论

(Introduction)

一、药物的属性 (Attributes of Drug)

药物应具有的基本性质是安全性、有效性、稳定性和质量可控性，这是相互关联和不可分割的四种基本属性。

药物的安全性 (safety) 指在用药的剂量下，药物不呈现或只有轻微的可接受的不良反应或副作用，是用药的前提。药物作为外源性物质，往往分布于体内多种器官或组织中，并与不同的靶标发生相互作用。人们力图将药物产生的无疑效果降低到最低程度，使其尽可能宽的安全范围。当然，药物的安全性也是相对的，对于治疗不同疾病的药物安全性有不同的要求，例如计划生育用药和小儿用药的安全性要求很高，而对威胁生命的肿瘤和艾滋病的药物，可以容许一定程度的不良反应。

药物的有效性 (efficacy) 是指药物的治疗效果，即药物发挥的效力，是用药的目的所在。药物的有效性体现在安全剂量下，分布到靶组织处的药物有足够的浓度，并在相当的时程内，发挥药理作用。所以，药物的有效性体现在一定的时间内有合理的吸收、分布、代谢和排泄 (药代动力学, pharmacokinetics) 性质和对靶标的特异性作用及足够的强度 (药效学, pharmacodynamics)。

药物的稳定性 (stability) 包括两个方面，物理和化学稳定性以及代谢稳定性。前者系指药物原料及其制剂在给定的时间和条件下，保持物理形态和化学结构的稳定不变，以确保药物的治疗效果；代谢稳定性表示药物在体内的结构不变性，较少或不被机体的酶系代谢转化，因而以原形分子发挥药效。稳定性是药物安全性和有效性的保证。

药物的可控性 (quality control) 系指通过物理、化学或生物学的方法确保药物的质量和有效成分的含量，也是药物有效性的保证。

优良的药物应有特异的药理学效应、合理的药代动力学行为、没有或具有可以接受的不良反应、稳定的化学和代谢性质以及良好的物理化学性质。这些属性的物质基础是药物的化学结构，也就是说，在药物的分子结构里，“凝集”了上述诸多的性质。化学结构在很大程度上决定了药效、药代、毒理和物化性质。

显然，在新药创制的过程中，构建药物的化学结构是源头性的重要阶段，药物分子设计 (molecular drug design) 是指通过科学的构思和理性的规划，构建具有治疗价值的新化学实体 (new chemical entities, NCE) 的分子操作。

一、创制新药的四要素 (Four Elements of Drug Innovation)

构成新药研究的主要因素包括以下 4 个方面：①药物作用靶标的确定；②活性评价系统的确定；③先导化合物 (lead compound, 先导物) 的发现；④先导化合物的优化。这四大支柱是不可割裂、互动配合的。就学科而言，前两方面属药理学范畴，后两者为药物化学内容。

（一）靶标的确定

确定治疗的疾病目标和药物作用的环节和靶标，是研发新药的出发点，因为它既体现了研发的目标，也决定了以后各个研发操作环节的特点。选择病种和靶标，除旨在创制非盈利和社会公益性药物外，应以市场的需求、容量和投入回报为导向。新颖的靶标意味着新的作用环节和机制，因而在一定市场内有独占性和市场竞争力，但研发原创性药物的风险较大。

（二）模型的建立

靶标选定之后，要建立生物学模型，以筛选和评价化合物的活性。模型可由不同层次，例如体外的分子水平（酶或受体）试验、细胞培养、离体器官以及整个动物的体内试验，这些模型均应反映出是针对所选定的靶标的作用。模型应有准确性和重现性，要制定出筛选标准。如果化合物符合规定的标准，则研究项目继续进行；否则，的发现新药研究的物质准备应尽早终止研究。

（三）先导物的发现

新药研究的物质准备是化合物的搜集和制备，通过活性筛选，对达到所设定活性指标的化合物，定位入选的苗头化合物（hit）。苗头化合物系指对特定靶标或作用环节显示一定作用的新结构类型的化合物。然而并非所有的苗头化合物都有研发前途，需经结构的修饰或衍变达到先导化合物标准，即所谓的苗头化合物到先导化合物（hit-to-lead）的过渡。先导化合物不仅对某个靶标或模型呈现一定强度和选择性活性，而且其物化性质、药代性质和安全性等满足一定的要求。也就是说，应具有类药性（drug-likeness）和可开发性。先导物可以从多方面的来源得到：天然活性产物、基于受体结构的理性设计（structure-based drug design）、基于相似原理的分子设计（similarity-based drug design）以及随机筛选等。

（四）先导物的优化

先导化合物一般不能直接成为药物，可能由于作用强度或特异性不高、药代动力学性质不适宜、安全性未达到要求，或物化性质的不适宜等，需要对其结构做化学修改。通过优化结构，使上述诸多性质达到最佳配置，这就是先导化合物的优化（lead optimization）。优化的目的就是确定候选药物（drug candidate），进入开发阶段。先导物优化的特征是建立结构与活性的关系，简称构效关系（structure-activity relationship, SAR）。这里所说的活性是泛指生物活性，包括了药效、药代、毒性等。构效关系的研究可以是定性的评价，也可以是定量的研究，后者又分为二维和三维定量构效关系研究（2D-,3D-QSAR）。构效关系的研究不是理论问题，而是实验性的总结和概括，但构效关系的方法学和参数的研究，有许多理论问题。通过构效分析，进一步指导设计与合成，再经体内外活性评价，循环反馈，最终确定候选药物。

第二节 分子的多样性—先导化合物的发现

（Molecular Diversity-Lead Discovery）

一、天然产物具有多样性，是先导化合物的重要来源

（一）青蒿素

古书记载我国西北部盛产的黄花蒿有清热解毒作用，基于此我国科学家从中发现了青蒿素（artemisinin）具有强效抗疟活性。青蒿素对于产生耐药作用的恶性疟疾原虫感染的患者，有明显的治疗作用。青蒿素属于倍半萜类化合物，分子结构中含有过氧键，是呈现抗疟作用的药效基团，如果将过氧键还原成醚键，则完全丧失抗疟活性。然而，由于青蒿素的水溶性低，体内吸收效果不好，以其为先导化合物进行结构优化，经过一系列的研究，得到了如双氢青蒿素、蒿甲醚和青蒿

琥酯等优良药物。

（二）红霉素类

红霉素（erythromycin）为含有十四元环的内酯类抗生素，与其他十六元环内酯共同构成大环内酯类（macrolides）抗感染药物。环上通常连有两个或多个糖基。红霉素的强效抗菌作用，是通过抑制氨基酸酰化的 tRNA 的易位，干扰核糖体的程序化合成蛋白。然而红霉素有明显地刺激胃肠道副作用，研究表明红霉素在胃酸的催化下，6 位羟基与 9 位羰基形成半缩醛，继之与 12 位羟基生成缩酮，缩酮是产生胃肠道刺激和丧失抗菌作用的主要原因。为了避免缩酮的形成，将 6 位羟基甲基化，得到克拉霉素。临床应用表明克拉霉素降低了胃肠道的刺激作用，且因脂溶性增加，提高了口服生物利用度，化学稳定性好，因而改善了抗菌作用。另两个大环内酯阿奇霉素和泰利霉素的抗菌作用更强。3 位的克拉定糖并非抗菌活性所必需，将 3 位糖基水解并氧化成 3-酮，构成新一类红霉素。

（三）洛伐他汀

在生物的进化过程中，某些低等生物的次级代谢产物与哺乳动物的正常代谢产物有许多相似之处。例如，自霉菌发酵液中分离得到的洛伐他汀具有降低血液中胆固醇作用。洛伐他汀分子中含有二羟基戊酸片断，与胆固醇在人体内生物合成的前体 3-羟基-3-甲基戊二酸有相似的结构，后者在 HMG-辅酶 A 还原酶的催化下，被还原成甲羟戊酸，这是体内合成胆固醇的级联反应的限速步骤。洛伐他汀与酶的结合作用 K_i 值为 1.0 nmol/L ，低于底物 3-羟基-3-甲基戊二酸的 K_m 值（ $10 \mu\text{mol/L}$ ）104 倍，因而形成强效特异抑制作用。洛伐他汀分子中的六氢萘片断是维持与酶发生紧密结合的疏水性基团，对酶抑制起重要作用。二羟基戊酸以开环的形式与酶结合，内酯本身没有活性。六氢萘环对抑制过程起辅助作用，可用其他结构替换，例如氟伐他汀、阿托伐他汀和瑞舒伐他汀等分别用吡啶环、吡咯环和嘧啶环替代六氢萘环，仍有很强的抑制活性，但共同特点是有一定的疏水性。

（四）石杉碱

石杉科植物千层塔含有菲啶类生物碱石杉碱甲和乙，具有强效抑制乙酰胆碱酯酶的作用，可提高学习记忆和认知能力。临床用于治疗阿尔茨海默病、健忘症和重症肌无力。为提高石杉碱甲向中枢的分布，增加其脂溶性，将伯胺基制成芳香醛的西佛碱前药 ZT-1，有望成为新一代的抗老年痴呆药物。

（五）考布他汀 (combretastatin) A-4

考布他汀 A-4 是从植物树干中分离的顺式二苯乙烯化合物，有强效的抗白血病作用，对其结构进行修饰，得到一系列活性较高的化合物：

其中 combretastatin A-4 phosphate 是将酚羟基磷酸化的前药，提高水溶解性，用于治疗肝癌的临床研究。

combretastatin A-4 和秋水仙碱 (colchicine) 以及鬼臼毒素 (podophyllotoxin) 的抗癌作用机理都是抑制微管的聚合作用，干扰癌细胞的有丝分裂。它们的化学结构具有某种相似性。

二、组合化学 (Combinatorial Chemistry)

（一）基本原理

先导物的发现在很大程度上取决于化合物的合成和生物评价的速度，合成的数量越多，速度越快，发现活性化合物的概率越大，组合化学是在这个背景的要求下诞生的。组合化学是指由一类组建模块制备一组化合物的组合过程，其原理和方法涵盖的领域很广，如固相合成、固相支持剂、液相平行合成等。

高通量组合合成曾广泛用于先导物的发现，由此延伸出许多新的技术和方法，如自动合成仪、在线纯化、数据处理和高通量筛选等。既往的经验表明，大规模组合库获得的先导物效率低，现在趋向于制备小规模集中库，并常常与**基于受体结构的设计**和**分子模拟相结合**，提高发现先导物的概率。集中库的设计多用于先导化合物的优化。

小规模平行合成是将组合的理念用于普通的有机合成，可以简化并加快合成操作，尤其是在先导物优化和类似物的合成中可提高合成效率。

(二) 固相组合合成 (solid-phase combinatorial synthesis)

固相组合合成源于 Merrifield 的固相多肽合成，现已延伸到各类有机化合物的合成。其原理是将反应物经共价键连接在固相载体（树脂）上，化学反应在树脂上发生，所以固相反应的各个步骤都是在非均相环境中进行的。反应完成后，用化学方法将产物从树脂上裂解下来。固相合成的优点在于，反应底物和中间体始终连接于固体树脂上，溶液中可用过量的反应试剂，以加速反应并有助于反应的完全。当完成反应后，过量试剂与副产物可经简便的过滤或洗涤方法与连接在树脂上的主产物分开，所以固相组合合成本身包含着产物的分离过程。

用于固相合成的树脂通常是含有功能基的高分子材料，最常用的是苯乙烯-二乙烯基苯的共聚物，共聚物上含有氯甲基的树脂 (Merrifield 树脂, Merrifield resin)，或含有羟甲苯氧基的树脂 (Wang 树脂, Wang-type resin)，以及 Rink 树脂和 Ellman 树脂等。这类交联的聚苯乙烯树脂具有良好的化学稳定性，对强酸、强碱、强氧化剂、高温或低温等激烈条件均有较高的耐受性，因而可容许发生多种有机反应。此外，树脂随着功能基化的不同以及共聚交联、孔隙度、载量和树脂珠大小的不同，性质会有很大变化，如亲水性、化学稳定性和膨润性等，不同的规格供不同的底物和反应做选择。

固相合成包含两类反应：一类是底物与树脂的连接和产物自树脂上裂解的反应，另一类是底物结构的变换。这两种不同的操作应互不干扰各自独立进行，所以，应根据不同的底物和反应条件来选择适宜的连接基和树脂。

(三) 液相组合合成 (solution-phase combinatorial synthesis)

液相组合合成与上述的固相合成相反，是将反应试剂共价键连接于树脂上，呈不溶解状态，而反应底物和中间体处于溶解状态。因而纯化过程虽然也是过滤，但是反应产物在溶液之中。由于常用试剂均可结合于树脂上，因而容易实现商品化，例如树脂结合的三苯膦、氢化催化剂、卤化试剂、氧化剂和还原剂等，所以在合成设计上类似于经典的有机合成。

另一种液相组合合成是将传统的有机合成与应用树脂载体的固相合成相结合的方法，即使用可溶性载体代替固相载体，选择合适溶剂使反应在均相液态进行；分离纯化时换溶剂使载体带着产物呈固相析出。由于固相合成是在非均相状态进行的，在平行合成时会产生非线性的动力学行为、化学反应的非均衡分布以及纯度等问题。可溶性载体方法的特征是应用可溶性的聚乙醇单甲醚 (MeO-PEG) 作为反应底物的载体，该聚合物既保障了被连接的底物在反应体系中的可溶性，也是所参与反应的化合物的保护基。聚乙醇单甲醚具有螺旋状结构，因而容易结晶化，所以在组合合成的分离纯化步骤只要保持聚乙醇单甲醚结构就容易在合适的溶剂中结晶纯化，而且含有聚乙醇单甲醚的分子有利于在水相或有机相中溶解。

(四) 研究实例

1. 二肽酶 IV 抑制剂脂肪腈化合物库

二肽酶 IV (dipeptidyl peptidase IV, DPP-IV) 是与 II 型糖尿病相关的靶标, 为丝氨酸蛋白酶, 在体内可促使胰岛素激素-胰高血糖素样肽-1 失活, 引起糖尿病。DPP-IV 特异性裂解 GLP-1 的 N 端二肽残基: Xaa-Pro 或 Xaa-Ala (Xaa 代表一个氨基酸)。选择性地抑制 DPP-IV 酶, 可延长 GLP-1 刺激胰岛素分泌的持续时间, 抑制胰高血糖素的释放, 减缓胃的排空, 因而降低血糖水平。

模拟 Xaa-Pro 残基, 得到环己基甘氨酸-2-氰基吡咯烷, N 端的伯胺结合于酶的 P2 域, 对 DPP-IV 的活性 $IC_{50} = 3 \text{ nmol/L}$, 进而认为 N-取代的甘氨酸氰基吡咯烷活性更强。

三、随机筛选 (Random Screening)

对于全新的靶标发现苗头化合物的最常用方法是高通量筛选化合物库。为了提高苗头化合物的质量和入选率, 化合物的结构多样性、类药性和数量是重要的前提。结构多样性可提高发现苗头化合物的概率; 类药性保障苗头化合物过渡到先导物的可行性; 没有相当大数量的化合物库, 命中率得不到保障。

高通量筛选的样品数量一般很大, 为数十万甚至上百万个。为了从这样多的化合物中确定苗头化合物, 显然只能用离体的酶或受体或细胞进行筛选。为提高速率, 一般采用单点一次性试验, 所以高通量筛选发生假阳性和假阴性 (漏筛) 率较高。

回顾和评价过去用 HTS 方法发现苗头化合物和先导化合物, 发现 HTS 的效率不高, 虽然化合物具有结构多样性, 但由于相当多的化合物无类药性或没有类先导物性, 例如相对分子质量较高、疏水性较强或其他不宜作先导物的性质, 所近年来在新药研究中又提出了配体效率的概念, 作为表征苗头化合物和先导物质量的参数, 并且发展了基于片断的药物分子设计的方法, 被筛选的多为相对分子质量较低的化合物。

四、虚拟筛选 (Virtual Screening)

从组合化学设计合成的化合物库或商业提供的化合物库中, 发现苗头化合物或先导化合物, 不必 (有时也不可能) 全部付诸活性筛选。可用计算机筛选方法对成药的各种因素加以判断, 剔除那些不能成为苗头的化合物。例如对化合物的类药性进行筛选, 将一定范围内的相对分子质量、分配系数、氢键数、极性表面积等数值作为选择标准, 排除范围以外的难以发展成药的化合物。进而用分子对接方法筛选与靶标的活性中心匹配良好的化合物, 或通过药效团搜寻的方法找出可能有活性的分子, 这种方法称为虚拟筛选, 或称为 *in silico* 筛选。采用虚拟筛选操作是基于知识的应用, 犹如用一系列基于知识的“滤片”对化合物库“筛选”, 以“浓缩”出能够满足预定标准的化合物。这样, 将药物筛选的离体和在体的筛选向虚拟方向推移, 成为 *in silico-in vitro-in vivo* 模式。这些虚拟属性的滤片包括类药性、药代动力学性质、毒性、知识产权问题以及与受体的互补性或配体的相似性等, 是通过数据库搜寻和计算化学实现的。图 24-4 是虚拟筛选的示意图。

(一) 类药性

药与非药的结构特征可用统计学分析划分, Lipinski 归纳的“类药 5 规则”, 概括了类药的最低标准, 即相对分子质量在 500 以下, 氢键的给体不超过 5 个, 氢键的受体不超过 10 个; 计算的分配系数 clg P 值不超过 5。上述原则只限于化合物经被动扩散机理的吸收。在虚拟库筛选中, 凡超出上述原则的化合物, 皆在排除之列。

化合物的柔性不宜过强。例如, 分子中含有连续 5 个单键的片断 (烷基除外),

柔性过强，会存在许多种构象，因而有不同分布的药效团。其有利之处是，有可能在活性筛选时，呈现为“苗头”；但更为不利的是，由于有多种药效团，会有多种生物活性，选择性因此不高，也因此所需的某种活性构象的“浓度”低而活性不高。

化合物不得含有重金属和反应活性基团。含有反应活性的功能基的化合物，例如酰卤、酸酐、醛基、卤代烷烃、含杂原子的三元环、氰酸酯、氮-卤化物等可与蛋白质分子中的亲核性基团反应，发生不可逆性和非特异性结合，产生毒性作用。因而除某些抗肿瘤药物和酶不可逆抑制剂外，应避免上述基团。

（二）药代动力学性质

药物的吸收、分布、代谢和排泄对药效起重要作用，根据统计学分析，临床试验被终止淘汰的候选药物 40% 是由于药代动力学不合理造成的。所以，在虚拟库筛选中，发现苗头的阶段要考虑药物的吸收、分布、代谢、排泄等性质，以提高筛选的质量。

决定药物能够穿越细胞膜并在胞浆中转运的性质是分子的化学结构，表现在相对分子质量、解离常数、亲脂性、极性表面积以及形成氢键的能力等。

药物的代谢转化主要在肝中发生。将重要的细胞色素 P-450 如 2D6 和 3A4 催化中心的三维结构作为模板，预测未知化合物的代谢命运。通过分析化合物的三维结构与半衰期的相关性，可以来预测未知物的代谢模式。

（三）毒性的预测

基于已有化合物的毒性和结构特征，经判别分析和多重回归分析得到的模型，可用来预测未知物的毒性。用神经网络算法对化合物的三维结构特征与LD₅₀的相关性分析，也可预测新化合物的毒性。预测毒性可使用基于知识的专家系统，所谓专家系统，是基于实验得到的毒性数据并由这些数据建立起规则（即知识），帮助用户对化合物的毒性做出合理的预测。

（四）基于受体结构的分子设计

在受体结构已知的情况下，可根据其结合部位的三维结构信息，用分子对接方法筛选库中化合物，对互补性好，评分高的化合物，可预测为高亲和力的分子。凡达到设定评分标准的，可作为虚拟筛选的苗头化合物。应当指出，虚拟筛选出的苗头化合物需经试验确定是否是真正的苗头。

（五）基于药效团的数据库搜寻

在受体结构信息未知的情况下，可根据一组活性化合物所具有的共同结构特征，即药效团对化合物库进行相似性对比，凡具有药效团特征的化合物可作为虚拟筛选的入选分子。

第三节 分子的互补性—先导化合物的发现

（Molecular Complementarity —Lead Discovery）

一、分子识别（Molecular Recognition）

配体与受体的相互作用，抗原与抗体的特异性结合，DNA双螺旋结构的形成与复制，DNA对 mRNA转录和 mRNA对蛋白质合成的指导作用，以及药物与靶标的相互作用等，都是以分子间相互作用为基础，主导这些分子间相互作用的本质是分子识别。

分子识别是由于两个分子中的某些特定的原子或基团性质的互补性和空间的适配所驱动的，这种特异性识别的本质是双方的互补性。药物与受体的分子识别和相互作用，大都形成非共价键作用，这与维持机体的生物大分子的空间结构和键合力是相似的，例如维持DNA的双螺旋结构，蛋白质和多肽的α螺旋或β片

层的二级和三级结构等以氢键、静电引力和疏水作用等弱作用力所支配。酶与底物的初始识别和结合也是非共价键作用。

配体-受体结合的特异性，还表现在这些非共价键合的生成不需要越过较高的能垒，因而在动力学上是有利的。相互作用的解离常数 K_d 与热力学的标准自由能 (ΔG^\ominus)、焓 (ΔH^\ominus) 和熵 (ΔS^\ominus) 有如下的关系：

$$\Delta G^\ominus = -RT \ln K_d \quad (24-1)$$

$$\Delta G^\ominus = \Delta H^\ominus - T \Delta S^\ominus \quad (24-2)$$

式 (24-1) 中， R 为摩尔气体常数， T 为热力学温度。由测得的解离常数 K_d 可按照该公式计算复合物的结合自由能 ΔG^\ominus 。为了求解 ΔH^\ominus 和 ΔS^\ominus 值，要测定 K_d 与温度之间的函数关系。如果 ΔH^\ominus 和 ΔS^\ominus 与温度无相关性，则将 K_d 与 $1/T$ 作图可求出 ΔH^\ominus 和 ΔS^\ominus 。也可以用热量计直接测定 ΔH^\ominus 。

支配药物-受体的分子识别和结合的作用可分为两个方面：焓作用和熵作用。焓因素包括静电作用（包括离子-离子、离子-偶极、偶极-偶极、电荷转移作用）、氢键作用和立体作用；熵因素包括疏水作用、转动熵和平动熵以及构象熵等。

药物和受体分子的构象变化，对分子识别和相互作用有很大影响，在诱导契合 (induced fit) 的过程下，药物与受体结合所采取的构象 (药效构象) 并非最低能量构象，所需的能量可由相互作用能补偿，系统的总能量下降。

二、基于受体结构的药物分子设计 (Structure-based Drug Design)

确定并了解受体分子的三维结构，特别是受体与配体分子形成复合物的三维结构，是基于受体结构的分子设计 (structure-based drug design) 的基础。受体一般是蛋白质，作为生物聚合物具有复杂的三维空间结构，并处于不断的变化和运动中。测定受体蛋白的结构式生物化学和结构生物学的研究内容，不仅要鉴定构成蛋白质的氨基酸种类和序列，而且要确定其空间结构。为此，可用 X 射线晶体衍射方法测定蛋白质结晶的衍射图，得到各个原子和基团的坐标，即在空间的相互位置；也可用核磁共振方法测定蛋白质与配体在溶液状态下的空间相对位置；还可用同源蛋白法预测蛋白质的三维结构。计算化学与分子图形学在研究和显示药物与受体分子的三维结构以及揭示它们之间的相互作用有重要的价值。

表皮细胞生长因子受体 (EGFR) 是一个跨膜糖蛋白，是控制细胞生长的重要蛋白质。它的过高表达可造成细胞过度生长，在多种肿瘤细胞都有 EGFR 过高表达，因此和癌症的发生有密切关系。该受体蛋白在膜内具有催化功能的结构域，可使蛋白中的酪氨酸残基磷酸化，所以称为 EGFR 激酶。抑制该激酶活性，是研制抗肿瘤药物的一个途径。EGFR 抑制剂厄洛替尼的设计模拟酶的活性部位 ATP 的作用，4-苯氨基喹唑啉的 1 位 N 与 Met769 的 -NH 形成氢键，距离 1.7 Å (在表征原子距离时多用 Å 为单位)；3 位 N 能够和结构水形成氢键，距离 1.8 Å；该结构水的氧又与 Thr766 的羟基氢形成氢键，距离 2.2 Å；4-取代的苯氨基进入疏水性氨基酸残基构成的疏水腔。图 24-5 是 erlotinib 与 EGFR 复合物的晶体结构图。另一个重要特征是喹唑啉环的 6 位和 7 位的侧链未与酶的表面接触，伸向水相介质，因而可引入亲水性侧链，是调节化合物的物化性质乃至药代性质的重要区域。EGFR 的另外两个抑制剂吉非替尼和拉帕替尼也是基于上述的结合方式成功设计的。

引起艾滋病的人免疫缺陷病毒的蛋白酶是治疗艾滋病的重要靶标，HIV 蛋白酶与该病毒繁殖和成熟过程有密切关系，所以抑制 HIV 蛋白酶是研制艾滋病药物的一个途径。HIV 蛋白酶的三维结构已经解析，在水解蛋白肽键时，酶分子中的

两个天冬氨酸残基 (Asp25 和 Asp25') 与被水解肽键的羰基氧原子形成氢键, 两个异亮氨酸残基 (Ile50 和 Ile50') 与水分子形成氢键, 水又与肽键的羰基形成氢键, 构成了 HIV蛋白酶履行水解功能时的结构特征, 如图 24-6 (a) 所示。

根据酶在催化水解过程的上述过渡态结构, 如果设计的化合物能与这些氨基酸残基相结合, 就能够阻止该酶的催化功能。图 24-6 (b) 中取代的环脲分子有两个羟基与天冬氨酸结合, 脲基的氧原子履行结构水分子的氧原子的功能, 与异亮氨酸结合, 苯环增加了与酶分子疏水腔的结合。由于该化合物具备了与 HIV蛋白酶结合的**必要的原子和基团**, 呈现有强效的抑制作用。有趣的是, 根据自由能变化与解离常数的关系式 $\Delta G = -2.30 RT \lg K$, 计算出结合能为 56 kJ/mol, 这与基于酶分子与该抑制剂结合时功能基的贡献计算得到的结合能 (56 kJ/mol) 非常接近, 表明这种设计的理论计算与实验数据是相当吻合的。

三、反义寡核苷酸 (Anti-sense Oligonucleotide)

基于 DNA 或信使 RNA 的结构设计寡核苷酸是药物分子设计的另一种方法。DNA 双螺旋的互补性结合是由核酸碱基腺嘌呤 (A) 与胸腺嘧啶 (T)、鸟嘌呤 (G) 与胞嘧啶 (C) 之间的氢键所维持。DNA 具有两个最重要的功能, 即在细胞分裂时每股螺旋复制出自身的互补链, 生成子代的 DNA。另一个重要功能是转录出 mRNA, 后者具有指导翻译生成蛋白质的功能。图 24-7 (a) 表示 DNA 的重要功能; 图 24-7 (b) 为反义寡核苷酸与 mRNA 结合的模式图。

如果设计能够与 DNA 或 mRNA 发生特异性结合，分别阻断核酸的转录或翻译功能，则会阻止与病理相关的核酸或蛋白质（例如酶或受体）的生物合成。这种可与 DNA 或 mRNA 结合的互补链成为反义寡核苷酸。反义寡核苷酸作为药物应认为是从遗传信息的复制、转录或翻译的根本干预病理过程，理论上应是“治本”的药物。

许多疾病的发病原因是基因的缺陷或在基因转录或翻译过程中的失常，如癌症和病毒性疾病。癌症可由多种因素引起，但主要是因细胞内致癌因素激活了癌基因。逆转录病毒的遗传物质掺入到宿主的 DNA 中，导致艾滋病病毒性疾病的发生。所以，建造直接针对细胞内 DNA 或 RNA 的分子，原则上应当能够清楚病因，阻止或逆转这些疾病。核酸作为靶标的优点是将基因的表达阻止在早期阶段，而且核酸的结构比蛋白质结构清楚，设计核酸的拮抗剂比设计蛋白质拮抗剂理论上要容易些，但在技术上有许多难点。

反义寡脱氧核苷酸所带的碱基与 DNA 或 mRNA 片段呈互补时，可结合形成 DNA 寡核苷酸或 mRNA 寡核苷酸杂交链，选择性地抑制基因的表达。与 DNA 结合时，是同大沟的 Hoogsteen 碱基对互补结合成三螺旋，从而抑制核酸的转录；也可与 mRNA 做互补性结合，抑制 mRNA 对蛋白质表达。然而 DNA 靶标不像 mRNA 那样简单，因而反义寡核苷酸与双链核酸的交联比与单链核酸困难，因此干预 mRNA 的翻译比干预 DNA 的转录更容易，这是因为外加入的反义寡核苷酸一般浓集在胞浆中，较易与处在胞浆中的 mRNA 结合。

反义寡核苷酸作为药物应满足一下的标准：容易合成并大量制备；在体内稳定，能够耐受核酸酶的水解；能够进入并停留在靶组织处，进入细胞内是重要的；应与靶细胞内特定的核酸发生作用，但不与其他核酸反应。

反义寡核苷酸的分子大小是设计的重要环节。为了只对一个特定基因的识别和选择性结合，反义寡核苷酸要有一定的长度。统计学表明，每 17 个寡核苷酸（17-mer）序列只在人基因组中出现一次，所以合成的反义链长至少含有 15 个（G-C 含量高时）到大约 20 个（A-T 含量高时）核苷酸。增加核苷酸数会提高选择性，因而降低了毒副作用。但超过 25 个碱基的寡核苷酸难以透过细胞膜，不宜作为反义寡核苷酸。

核酸酶存在于细胞核内或核外，功能是裂解核酸中的磷酸二酯键，使核酸降解。核酸酶可降解外切位点（5 或 3-外切核酸酶）或降解内切位点（内切核酸酶）。它的存在限制了正常的寡核苷酸作为“反义”的应用，因为它们会迅速被降解。为了提高对酶降解的抗性，增加稳定性，可对核酸做结构修饰，制成核苷酸模拟物，这是药物化学要解决的问题，例如改变磷酸二酯基、脱氧核糖、碱基和末端的结合物，如图 24-8 所示。

(oblimersen) 是 18 链节的硫代寡核苷酸，临床治疗乳腺癌和直肠癌。

另一个 20 链节的硫代寡核苷酸阿福韦生 (afovirsen)，是已上市的抗艾滋病病毒药物。

5-TCTCCCAGCGTGCGCCAT-3

第四节 分子的相似性-先导化合物的发现和优化

(Lead Discovery and Optimization)

应用相似原理进行药物分子设计的内容很多，也因此药物化学中形成许多重要的原理。相似性原理及其应用，大都是从小分子化合物的结构特征出发，引申出新的结构类型或新的化合物。这些方法包括基于内源性配体结构的分子设计、肽模拟物、生物电子等排、过渡态类似物、同系物、插烯物、剖裂物、拼合物，基于合成中间体的设计和基于代谢转化的设计等。

一、基于内源性配体分子的药物设计 ()

在正常状态下，内源性物质如激素或神经递质的生理功能是通过与体内受体作用而引发的。当内源性物质过多或过低的产生时，造成功能的失衡和病理状态。治疗或纠正这种异常状态，可以通过调节（提高或降低）这些内源性物质的作用，如受体激动剂和拮抗剂（受体调节剂）或酶抑制剂等。当受体或酶的化学结构尚不清楚的情况下，从内源性分子的化学结构出发进行分子设计即基于配体的药物设计是一个重要途径。

最常用的策略是模拟内源性分子的化学结构，通过改变分子的某些原子、基团或片断，或者模仿分子的形状和性质，使仍保持对受体分子识别能力，以激活受体的作用（激动剂），或阻断受体的功能（拮抗剂），或者虽保持了酶的亲和力，却不能使酶履行催化功能而被抑制。

（一）氟尿嘧啶

尿嘧啶是构成核酸的一个碱基，当 5 位氢原子被氟代替时，就成为抗代谢物氟尿嘧啶，是一个临床广泛应用的抗癌药。氟与氢虽然电负性不同，但原子尺寸相近，氟尿嘧啶在体内被三磷酸化后，仍可被胸苷酸合成酶识别而结合，并且结合的能力比正常底物三磷酸尿嘧啶强数千倍。氟尿嘧啶“蒙骗”了癌细胞，致其于死地。

（二）甲氨蝶呤

叶酸是细胞代谢的一种辅酶，在二氢叶酸还原酶催化下，生成四氢叶酸，是转移一碳基团的重要辅酶，参与核酸碱基和某些氨基酸的生物合成。甲氨蝶呤是模拟叶酸的抗代谢物，是将氨基替换了叶酸中的羟基，甲基替换了氢原子，甲氨

但不能催化还原反应，导致癌细胞因合成叶酸所需的碱基匮乏而死亡。

二、肽模拟物（ ）

生物活性肽如激素、神经递质、细胞因子等多种生物活性，与相应受体结合，引发特异的生理作用。然而肽作为药物有许多缺点，例如口服生物利用度低，难以穿越血脑屏障，在体内易被水解失活。此外，注射异源性肽会产生过敏反应。

肽类化合物一般含有多个单键，为柔性分子，在溶液中以多个低能量构象群存在，同一个分子的不同构象，分子形状不同，可被不同的受体识别，因而产生不同的生理效应。一种内源性肽可履行多种功能，这是生物进化的结果。然而在病理下某种功能可能亢进或低下，在调整这种不正常状态时，活性肽的多功能性会产生副作用。

为克服这些缺点，对肽结构作化学修饰或改造，成为肽模拟物。肽模拟物是一类能够模拟肽分子与受体或酶的相互作用，成为可激活或阻断某种生物活性的非肽、类肽或拟肽。其结构特征是，用非肽或拟肽作骨架，在适宜的位置连接药效团，并使药效团的空间分布相似于肽的活性构象的药效团分析。

常用的方法是通过氨基酸的交换，限制肽的构象，例如 α -碳的氢原子被烷基取代， α -氨基的烷基化，引入 α ， β -不饱和双键，引入或拼合脂环，L-构型氨基酸变成 D-构型氨基酸， β -氨基酸替换 α -氨基酸等。将肽链的骨架改变，可避免被水解酶识别，因而增加了代谢稳定性。

糖蛋白 IIb/IIIa 受体拮抗剂的合成是模拟肽的典型实例。血小板的黏着和聚集是血栓形成的病理原因，许多内源性物质可激活血小板，糖蛋白 IIb/IIIa 是血小板膜受体，受体被激活后与纤维蛋白原结合，导致血小板聚集。糖蛋白 IIb/IIIa 受体拮抗剂是防治动脉血栓的药物。

研究发现，三肽 Arg-Gly-Asp (RGD) 是糖蛋白 IIb/IIIa 受体识别的重要片段，其构象限制性三肽仍是受体的识别基团，例如包含有 RGD 片段的环肽 DMP728 是强效拮抗剂。

以 DMP728 为模板设计非肽类拮抗剂，构效关系表明，分子内的甘氨酸残基采取伸展型构象，从碱性的胍基到羧基的空间距离为 $15\sim 17\text{\AA}$ ，而且 Gly 的羰基模拟物芳基或酰胺仍是良好的结合位点，能够满足上述条件的分子并有适宜的空间配置时，可成为糖蛋白 IIb/IIIa 受体的非肽类拮抗剂。用异恶唑作为支撑药效团的骨架，得到罗昔非班 (roxifiban)，对受体结合的 $IC_{50} = 0.25\text{ mol/L}$ ，可口服治疗血栓病。罗昔非班结构中的脒基模拟 RGD 的精氨酸的胍基，羧酸甲酯相当于天冬氨酸的游离羧基，脒基到羧基的空间距离为 $15\sim 17\text{\AA}$ ，苯基异恶唑啉为支撑骨架。

生长抑素 (somatostatin) 为含有 14 个氨基酸和二硫键的环肽，可抑制胰岛素、胰高血糖素和胃泌素的分泌。生长抑素受体激动剂奥曲肽 (octreotide) 保持了生长抑素的主要结构，作为激动剂临床用于治疗消化道肿瘤。用 3-脱氧葡萄糖作为骨架，模拟生长抑素识别受体的 β 转折 Phe7-Trp8-Lys9-Thr10 片段，

吡啶以及氨基，该糖模拟物对生长抑素受体的激动活性为 $50 = 3 \text{ mol/L}$ 。

三、生物电子等排置换 (Replacement of Bioisosteres)

电子等排(isosterism)系指化学上不同的物质，由于外层电子有相同的数目和排布，导致物理性质具有相似性。这个概念几经扩展，发展成生物电子等排原理，是指一组化合物具有相似的原子、基团或片断的价电子数目和排布，可产生相似、相近或相关的生物活性。根据这个广义的概念，在药物化学上一价的原子或基团可相互置换，例如卤素、羟基、甲氧基、烷基和氢原子之间；羧基、磺酰氨基、磺酸基、氧羧基、磷酸基、磷酰氨基、四氮唑基等互为等排体；二价连接基如醚键、硫醚、仲（或叔）氨基之间，羰基、亚砷基、酰胺片断、磺胺片断等为等排体；芳环与芳杂环之间也可做生物电子等排置换。生物电子等排体一般与同一受体部位结合，作用机制相同。表 24-1 列出了重要一价基团的电子等排体。

例如，抗消化道溃疡药 H^+ , K^+ -ATP 酶抑制剂奥美拉唑 (omeprazole)、兰索拉唑 (lansoprazole) 和泮托拉唑 (pantoprazole) 等，是变换环上的一价基团，成为电子等排体。

抗炎药物 COX-2 抑制剂塞来昔布 (celecoxib)、罗非昔布 (rofecoxib) 和伐地考昔 (valdecoxib)，主要是交换中间的杂环骨架。

, SERMs)

用于治疗骨质疏松和绝经后乳腺癌的化疗辅助治疗,例如他莫昔芬(tamoxifen)、雷洛昔芬(raloxifen)和巴多昔芬(bazedoxifene)是环系的等排互换。

抗消化道溃疡药物 H₂ 受体拮抗剂的研制在药物化学上有突出的地位,是以组胺为出发点,通过对咪唑环和碱性侧链的变化调整激动-拮抗作用研制成功的典范。同时也是成功地运用电子等排原理创制出一系列抗溃疡药物的范例。西咪替丁(cimetidine)、雷尼替丁(ranitidine)、法莫替丁(famotidine)和尼扎替丁(nizatidine)均为上市的 H₂ 受体拮抗剂,它们之间芳杂环和碱性侧链的不同,都替换了成功地等排置换。

依据相似性原理做生物电子等排交换,要点是不能改变药的效团及其在空间的正确排布,原子、基团或片断的交换不影响药物与受体结合的模式,当然会改变作用强度以及药学和药代动力学性质。例如羟基、烷氧基、氨基(氮原子上含有或不含氢)与其他一价原子或基团的变换,会由于氢键的有无和偶极作用的大小而影响与受体的结合,溶解性和分配性也会有所变化。

四、 决策法 (Topliss Decision Tree)

以线性自由能相关原理为基础的 Topliss 决策法,本质上是一价生物电子等排基团的变换,只是在结构的相似性演变时,考察多种因素对活性的影响,作为设计新化合物的选择依据。该方法是 Topliss 发明的,其原理是通过对合成的第 2 个化合物与第 1 个化合物的活性强弱的比较,决定第 3 个化合物的设计,在受体结构未知的情况下,这种“试探性”的方法是仔细分析取代基的电性、疏水性和立体性的综合性质进行决策的,根据这些基团性质的相似程度,绘制成演变的树状操作图,故称 Topliss 决策树。Topliss 决策树分芳香族化合物和脂肪族化合物两类,图 24-9 是芳香族化合物优化操作的决策树。

初始物是未取代的 H,并测定了活性。由于许多药物与生物系统的作用依赖于疏水性,随疏水性增高活性加大,故首先选定具有疏水性+ 效应的 4-Cl,4-Cl 也容易合成。测定 4-Cl 活性。若强于 (M) 母体物 H,可能是由于 效应或吸电子+ σ 或+ π + σ 综合效应,因而再引入一个氯原子,即下一步合成 3,4-二氯化合物。

若 3,4-二氯化合物活性强于 4-Cl,提示活性的增加为+ π 和 (或)+ σ 综合效应,因而再引入一个氯原子,即下一步合成 3,4-三氯化合物。

若 3,4-二氯化合物活性强于 4-Cl,提示活性的增加为+ π 和 (或)+ σ 提高了

或 $+\sigma$ 更高的化合物, 即

$3,4\text{-Cl}_2$ 或 $3,4\text{-(CF}_3)_2$ 。这两个化合物的 ϵ 和 ϵ' 都增大了。若活性比 $3,4\text{-Cl}_2$ 高, 则合成 $3\text{-CF}_3\text{-4-NO}_2$ 。

若 3-Cl 的活性等于或低于 4-Cl , 可能是由于 3 位取代产生了不利的空间因素, 也可能是 $+\pi$ 或 $+\sigma$ 超过了最适值。为此, 合成 4-CF_3 , 它具有较高的 σ 值, 亲脂性也强于 4-Cl , 但弱于 $3,4\text{-Cl}_2$, 也无 3 位空间障碍。除 4-CF_3 外, 还可选择 4-Br 或 4-I 。 $3,4\text{-Cl}_2$ 的活性降低也可能是 3 位的立体障碍不利于同受体结合, 可合成 $2,4\text{-Cl}_2$, 它的 π 和 σ 值与 $3,4$ 大体相同, 但空出了 3 位。最后是 4-NO_2 , 硝基的电性比疏水性显著。如果 4-NO_2 的活性增高, 是因为 $+\sigma$ 值强于 3-CF_3 , 而 π 值减低有利于活性。

活性评价结果若 4-Cl 化合物的活性与母体物 H 相同, 可能是适宜的 $+\pi$ 效应和不适宜的 $+\sigma$ (即 $+\sigma$ 增大活性降低)。如果是这样, 则应合成 4-CH_3 , 甲基具有 $+\pi\text{-}\sigma$ 效应。若 4-CH_3 增高了活性, 继续加大 $+\pi\text{-}\sigma$ 效应, 合成 4-t-Bu 。若 4 位叔丁基因立体因素不利于活性, 可合成 $3,4\text{-(CH}_3)_2$ 。最后还可合成 4-异丙氧基、4-苄氧基或 4-苯氧基, 这些都有较高的 $+\pi$ 和低 σ 值。

如果 4-CH_3 活性等于或低于 4-Cl , 这可能是由于对位取代基不利的立体效应所致, 也可能是 $-\pi$ 效应。但 $-\pi$ 效应是随着疏水性增加活性降低, 这种现象比较少见, 因而推测 4-CH_3 未增加活性可能是 4 位取代基的立体障碍。为此, 下一步应合成 3-Cl 。 3-Cl 与 H 相比, 有较高的 $+\pi+\sigma$ 但 4 位未被占据。若 3-Cl 的活性增高, 下一步按照 4-CF_3 路径, 以同样的理位。由向下合成, 只是都须在 3 位。但若 3-Cl 的活性未提高, 可归因于 $+\pi\text{-}\sigma$ 所致。活性的提高可能需要较高的 $+\pi$ 值和较低的 σ 值, 因而应合成 3-CH_3 , 它是 $+\pi\text{-}\sigma$ 基团。如果 3-CH_3 活性没有变化, 则应合成邻位取代基, 如 2-Cl 、 2-CH_3 或 2-OCH_3 。

若 4-Cl 、 4-CH_3 、 3-Cl 、 3-CH_3 和邻位取代基与 4-Cl 相比均不能增高活性, 则 $+\sigma$ 和较小的 π 值可能有利于活性, 故宜合成 4-NO_2 , 它是 $-\pi+\sigma$ 基团, 通过与 $+\pi\text{-}\sigma$ 的取代相比, 判断究竟怎样的综合效应会使活性提高。除 4-NO_2 , 还可合成 4-CN , 4-COCH_3 , $4\text{-SO}_2\text{CH}_3$ 、 4-CONH_2 和 $4\text{-SO}_2\text{NH}_2$ 。若 $-\pi+\sigma$ 基团仍使活性增高, 表明优选的 $-\pi+\sigma$ 起作用, 下一步宜合成 4-F 。此时与母体物 H 相比, 4-F 的 π 和 σ 值变化不大, 若母体物本身已是适宜的 π 和 σ 值, 且为避免在体内被代谢转化成 4-OH , 应合成 4-F 。

如果 3-Cl 活性低于 4-CH_3 , 表明 $-\sigma$ 效应起主导作用, 即随着给电子作用的增强 (σ 值降低), 活性增高, 下一步应合成 $3\text{-N(CH}_3)_2$ 、 3-NH_2 或 3-CH_3 。循此方向下一步应合成 2 位取代基。

若 4-Cl 比 H 的活性低, 可能是由于 4 位取代基的立体障碍或是增高的 $+\pi$ 或 $+\sigma$ 值不利于活性。假定此时 $-\sigma$ 效应起主导作用 (即随着 σ 值降低而活性增高), 则于 4 位引入 OCH_3 。若 4-OCH_3 活性强于起始物 H , 应合成 σ 值更低的 $4\text{-N(CH}_3)_2$, 并最终合成 3-CH_3 、 $4\text{-N(CH}_3)_2$ 。

若二甲氨基化合物的活性未见变化甚至降低, 可能是 π 效应过高所致, 因此宜合成 4-NH_2 或 4-OH 。如果由于二甲氨基的碱性被质子化成为不利的因素, 则应合成 3-CH_3 、 4-OCH_3 或 $4\text{-OCH(CH}_3)_2$ 。

如果 4-OCH_3 的活性相同或低于 4-Cl , 表明 4 位取代作用有不利的影晌, 应合成 3-Cl , 从而回到前述的路径。

这种设计策略尤其适用于化合物合成较难, 活性评价反馈较快的先导物的优化。

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/055111032103011222>