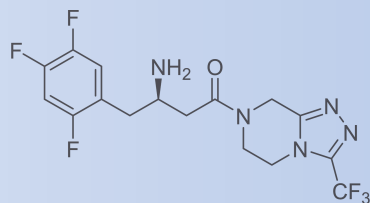


第一章



西格列汀

抗 2 型糖尿病药物西格列汀的 药物化学与工艺研究

徐锦优 梁贵柏

本章以西格列汀研发过程中的重要节点为背景，介绍肠促胰岛素在调控血糖中的作用和DPP-4抑制剂的药物化学研究，包括肠促胰岛素效应（incretin effect）、先导化合物（lead compound）的确认和优化（identification and optimization）、药代动力学与药效学（pharmacokinetics and pharmacodynamics）的评估、新型抑制剂的设计（novel inhibitor design）等内容。最后，本章简要介绍了西格列汀化学合成的工艺研究。

第一节 抗 2 型糖尿病药物西格列汀简介

调控血糖依赖性胰岛素分泌（glucose-dependent insulin secretion, GDIS）又称高血糖诱发性胰岛素分泌（glucose stimulated insulin secretion, GSIS），是调控血糖的新兴策略，近年来受到学术界和制药业的高度重视，其治疗2型糖尿病的相关研究及新药研发均取得了突破性的进展。格列汀类（gliptin）药物属于二肽酰基酶-4（dipeptidyl peptidase-4, DPP-4）抑制剂，主要通过减缓肠促胰岛素（incretin）的重要成员——胰高血糖素样肽-1（glucagon-like peptide-1, GLP-1）的降解，促进餐后高血糖时胰岛素的合成和分泌，从而实现调控血糖的目的。与已有的磺胺酰脲类（sulphonylurea）等糖尿病药物相比，该类药物可以大幅降低诱发低血糖的危险。

2006年10月，美国食品药品监督管理局（Food and Drug Administration, FDA）正式批准由美国默沙东制药公司研发的高效、高选择性的格列汀类药物——西格列汀（sitagliptin）作为治疗2型糖尿病的新药上市（结构式见图1-1-1），为有效控制和治疗2型糖尿病开拓了全新的路径，给全球大约2亿4千万糖尿病患者带来了新的希望。西格列汀上市后年销售峰值超过70亿美元，是一款名副其实的重磅药物。

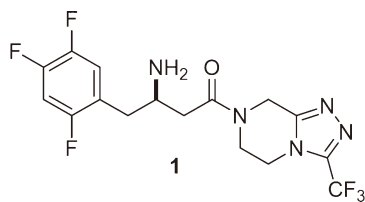


图 1-1-1 西格列汀（化合物 1）

第二节 肠促胰岛素激素及其降解酶二肽酰肽酶-4（DPP-4）

肠促胰岛素激素是从肠道黏膜中提取的对血糖具有调节作用的一组激素，其中两个主要成员是GLP-1和葡萄糖依赖性胰岛素释放肽（glucose-dependent insulintropic peptide, GIP），两者的分泌、功能与降解如图1-2-1所示。

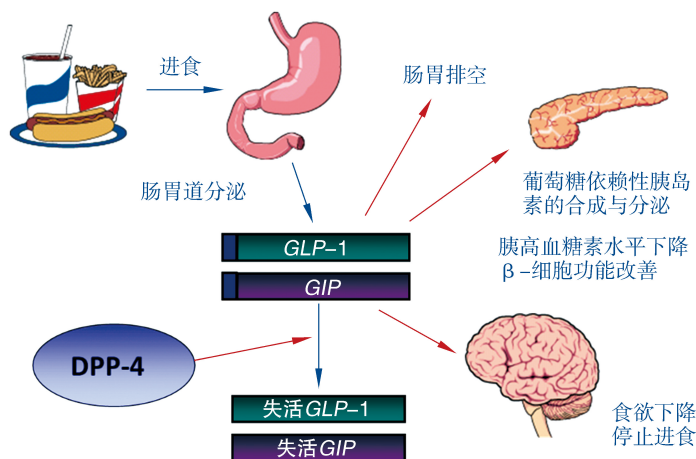


图 1-2-1 肠促胰岛素激素的分泌、功能及其降解

肠促胰岛素激素产生后便可激活它们各自的受体。胰脏β-细胞上的GIP和GLP-1受体被激活后，细胞内环腺苷酸（cAMP）的浓度就会迅速升高，从而导致胰岛素的分泌^[1]。同时，胰岛素的生物合成（biosynthesis）和β-细胞的分裂均增加^[2]。此外，活化的GLP-1受体还有助于抑制胰高血糖素（glucagon）的分泌，使人体减少食物摄入量（food intake）并减缓肠胃排空（gastric emptying^[3]）。由于GIP和GLP-1的合成及分泌只在进食后血糖升高才会发生，继而其受体被活化，所以合成与分泌胰岛素的过程也只会发生在高血糖的情况下才会发生，这被称为“血糖依赖性胰岛素分泌”（glucose-dependent insulin biosynthesis and secretion）。而正常的GLP-1水平不会影响胰高血糖素的分泌，因此，抑制胰高血糖素的分泌同样只发生在餐后的高血糖阶段，几乎不会有低血糖的风险^[4]。

正常人空腹时，GLP-1含量一般在5～10 pmol/L范围，进食后会迅速提高至15～50 pmol/L。但是，人体内循环的GIP和GLP-1不会维持在高水平，很快会在多肽酶的作用下被降解而失去活性成为失活GLP-1（inactive GLP-1）及失活GIP（inactive GIP），其中最主要的降解酶是DPP-4。

20世纪60年代，人们首先从大鼠的肝脏里发现了DPP-4^[5]。随后的研究发现，DPP-4在人体内的各类器官中广泛存在，肾脏、肠道和骨髓含量较高，肝脏、胰脏、胎盘、胸腺和淋巴等处也有存在，其在血浆中存在可溶形式，提示科学家可以针对DPP-4开发降糖药物。至20世纪90年代后期，有关DPP-4对体内能量平衡（energy homeostasis）的作用已经引起了学术界和医药工业界的广泛关注。

DPP-4是细胞表面的丝氨酸蛋白酶，属于脯氨酰寡肽酶家族，对脯氨酸（proline）和丙氨酸（alanine）特异性敏感，可从多肽的N末端切割X-脯氨酸或X-丙氨酸二肽残基（X代表蛋白或者多肽中的天然氨基酸残基）。DPP-4可降解多种蛋白及多肽底物，包括生长因子、趋化因子、神经肽和血管活性肽等。DPP-4的晶体结构发表于2003年^[6]，它是一种可与细胞膜结合的同二聚酶（homodimer），每个单体的分子量为110~150 kDa。目前已知在人体内，GIP和GLP-1是DPP-4的天然底物。

由于DPP-4的降解作用，肠促胰岛素在生物体（包括人体）内的半衰期（half-life）通常很短，如GIP的半衰期仅约10 min，而GLP-1的半衰期仅约1 min，这严重限制了肠促胰岛素在临床上的直接应用^[7]。实验结果表明，由静脉或皮下注入的GLP-1经DPP-4降解后，只剩下10%~20%的活性物质，因此无法降低糖尿病患者体内过高浓度的血糖（主要是葡萄糖）水平^[5]。由此，科学家提出设想：通过有效抑制DPP-4的活性，延长GLP-1在人体内的滞留时间，提高GLP-1在人体内的浓度，进而恢复2型糖尿病患者的肠促胰岛素效应，增加胰岛素的分泌，增强 β -细胞功能（improved β -cell function），可达到控制血糖的效果和目的。经过不断探索，1996年科学家在动物体内实验中首次让这一设想实现，确认DPP-4是一个潜在的抗2型糖尿病新药物靶标（drug target）。此后，各大制药公司便开始纷纷研发DPP-4抑制剂，开启了治疗2型糖尿病的新纪元^[8]。

第三节 基于性质的药物设计

将一个小分子优化成为一个安全有效的药物是一个漫长而又艰难的多参数优化过程（multi-parameter optimization process）。常见需要优化的药物成药性性质包括：生物活性（potency）、靶标的选择性（selectivity）、化合物的溶解性（solubility）、细胞的渗透性（cellular permeability）、代谢稳定性（metabolic stability）、口服生物利用度（oral bioavailability）、与血浆蛋白的结合（plasma protein binding）能力、细胞色素P450（cytochrome P450, CYP）的抑制与诱导能力、hERG等离子通道的抑制，以及化合物的杂泛性（promiscuity）引起的毒性等。

药物化学是科学和艺术的综合，也是一门半经验式的科学，每一名优秀的药物化学工作者都必须经过长期的实践才能够成为真正意义上的药物化学家^[9]。药物化学家在优化先导化合物的生物活性、选择性、药代动力学（pharmacokinetics, PK）和药效学（pharmacodynamics, PD）等性质的同时，要重点关注及考虑避免可能出现的毒性（toxicity）和潜在的包括靶标（on-target）以及非靶标（off-target）等相关的不良反应。因此，先导化合物的优化是一个在活性、口服生物利用度和各种毒副作用之间取得最佳平衡的持续过程。然而，药物化学家在优化化合物的某个性质的同时，经常会忽略结构的变化对其他性质的影响，致使在解决一个问题后又引入新问题。如此，分子结构愈加复杂，分子量和亲脂性相应增加，导致化合物的类药性（drug-like property）降低。

药物研发的每一个循环均包括假设（hypothesis）、合成（synthesis）、体内外生物测试（test *in vitro* and *in vivo*）和分析（analysis）等过程，过程漫长又耗费资金。有经



验的药物设计者 (drug desinger) 会在开始设计分子时尽可能综合考虑及讨论可能出现的各种问题以及变通方法, 明确合成每个化合物的目的。通常, 在先导化合物的确认 (identification) 阶段, 需要合成一定数量的、结构多样性的化合物, 以供选择和确认构-效关系 (structure-activity relationship, SAR), 但进入先导化合物的优化 (optimization) 阶段, 当需要对分子的性质进行微调 (fine-tuning) 时, 则要考虑结构的细微变化对药物整体性质的影响, 尽量避免引入新问题。例如, 为了解决药物候选物的脱靶 (off-target) 活性, 药物化学家经常会在分子上引入极性基团, 这便可能影响化合物的渗透性, 也容易致使分子成为P-糖蛋白 (P-glycoprotein, P-gp) 等外排转运体 (efflux transporter) 底物被转运出细胞以外, 从而影响药物的吸收^[12]。

20世纪70年代初, 随着蛋白质晶体学和磁共振波谱学的快速发展, 测试蛋白质与配体络合物的三维结构成为常规技术, 基于结构的设计 (structure-based drug design) 和计算机辅助药物设计 (computer-aided drug design) 得到大力推进, 翻开了药物设计的新篇章。由于早期缺乏高质量的实验数据, 以及对影响化合物吸收 (absorption)、分布 (distribution)、代谢 (metabolism)、排泄 (excretion) 以及毒性 (toxicity) (简称ADMET) 等因素的认知有限, 当时的计算机辅助药物设计主要用来帮助药物化学家优化化合物的活性。

直到20世纪90年代, 随着各个相关学科的发展, 特别是计算机化学 (computer chemistry)、生理学 (physiology)、生物化学 (biochemistry)、药物代谢学 (drug metabolism)、药代动力学 (pharmacokinetics)、药物转运学 (drug transportaion) 以及不同生物体内药物行为学 (drug behaviors *in vivo*) 等学科的发展, 使药物设计跨入新时代, 逐渐发展出了各种合理药物设计 (rational drug design) 模型和技术。

西格列汀 (sitagliptin) 的研发过程便是药物分子性质影响成药性的成功案例。项目团队成员 (包括本章作者) 采用了基于性质的药物设计方法。基于性质的药物设计 (property-based drug design) 是针对先导化合物或者候选药物的结构进行药代动力学性质的设计与优化, 以实现药物具有良好口服吸收 (A)、定向分布 (D)、可控代谢 (M)、优化消除 (E) 和降低毒副作用 (T) 的目的^[10]。目前, 越来越多的制药公司要求药物化学家在合成化合物前要先对其ADMET性质进行评估, 力争将失败因素控制在先导化合物的确认及优化阶段, 尽量避免发生在药物开发的后期阶段, 如临床试验研究阶段。

第四节 常见影响药物 ADME/T 性质的因素

影响分子类药性 (drug-like property) 的因素包括分子量 (molecular weight)、溶解性 (solubility)、亲脂性 (lipophilicity)、氢键配体数 (number of hydrogen bond donor)、氢键受体数 (number of hydrogen bond acceptor)、可旋转键数目 (number of rotatable bond)、极性表面积 (polar surface area, PSA) 和芳基的数目 (number of aromatic ring) (表1-4-1)。这些因素会直接影响药物的ADMET^[11], 要求药物设计者对影响分子类药性特征的因素、药物的生理过程及靶标如蛋白质等生物大分子的生物活性特



点等均具备深刻的理解和认识。

表 1-4-1 常见影响药物 ADMET 的分子的理化性质

分子的理化性质	理想值
分子量 (Da)	< 500
亲脂性	clogP 在 2 ~ 3 之间
极性表面积	非 CNS 药物 < 120; CNS 药物在 60 ~ 70 之间比较理想 (如当一分子的极性表面积 > 1.2nm ² 时, 其在细胞的穿透性就会变差; 需穿越血 - 脑脊液屏障的药物, 其极性表面积在 0.6 ~ 0.7 nm ² 比较理想)
氢键受体数	< 10 个
氢键供体数	< 5 个
分子内可旋转键数目	< 10 个
芳基的数目	< 2 个

极性表面积 (polar surface area, PSA), 化合物内极性分子的总表面积, 多为氧原子及氮原子, 也包括与其相连的氢原子

分子量影响药物的水溶性、渗透性、代谢和毒性, 是设计药物分子需重点考虑的因素之一。1997年, 美国辉瑞公司的药物化学家利平斯基 (Lipinski) 通过将已上市及在Ⅱ/Ⅲ期临床研究阶段失败的口服药物的特点进行研究与对比, 发表了著名的口服小分子药物利平斯基五规则 (Lipinski rule of five)^[13,14]。口服小分子药物被定义为: 分子量 < 500 Da; 氢键供体 < 5; 氢键受体 < 10; 脂水分配系数计算值 (clogP) < 5。同时, 该规则认为任一化合物如果不能满足其中至少两条, 便难于成为口服药。尽管如此, 随着现代技术的不断发展, 以及科学对肠道吸收机制的理解加深, 如今已有一些口服药物突破了该规则。

分子亲脂性对药物的所有性质, 如水溶性、肠黏膜的吸收、细胞的渗透性、分子与血浆蛋白结合的能力、肝脏和其他器官的代谢以及毒性等都均具有显著影响。通常, 化合物的 logP 在 2 ~ 3 最为理想, 其在溶解性、渗透性和毒性之间可达到良好的平衡。logP 过低, 化合物的渗透性较差, 而 logP 过高 (logP > 5) 则水溶性较差, 化合物更容易被细胞色素酶代谢, 导致高肝清除率 (high hepatic clearance), 从而影响药物的吸收和口服生物利用度。此外, logP 过高会增加药物与血浆蛋白结合的能力, 从而减少血液中自由药物的浓度 (free drug concentration), 造成临床上药代动力学与药效动力学的脱节 (PK/PD disconnection)。logP 过高还可能导致高杂泛性, 增加药物的脱靶活性和不良反应的风险。

药物的口服生物利用度是药物设计者面临的重大挑战。口服生物利用度主要与药物的吸收和代谢相关。常见影响药物吸收的因素包括溶解性、渗透性、小肠首过代谢 (intestine first pass metabolism) 效应、肝脏首过代谢 (liver first pass metabolism) 效应, 以及与药物是否为 P-糖蛋白等外排转运体的底物有关。药物的渗透性与分子量、亲脂性、分子内可旋转键数目、形状以及极性表面积等因素有关 (表 1-4-1)。如前所述, 药物的亲脂性 (logP) 控制在 2 ~ 3 可具备比较理想的渗透性和溶解性。近期研究表明, 相较于亲脂性, 分子的极性表面积能够更准确地反映分子的渗透能力, 已经成为常用的一个参数。当药物的分子内可旋转键数目 < 10, 极性表面积 < 120 Å² 时, 药物通常具有良好的渗透性



和口服生物利用度。而对于要进入大脑的药物，极性表面积一般 $<90 \text{ \AA}^2$ ，在 $60 \sim 70 \text{ \AA}^2$ 范围内比较理想^[15]。

为使药物的口服吸收率最大化，药物分子除应具有较好的渗透性外，还应该避免被P-糖蛋白等外排转运体排出细胞外，同时也需要避免被肠道中的细胞色素酶快速氧化代谢。位于细胞膜上的转运体是体内重要的功能性膜蛋白，在肠道、肝脏、肾脏，以及血-脑脊液屏障等其他生物膜上均有广泛的分布，可以多方位影响药物的吸收、分布以及排泄等过程。介导药物外排的转运体主要包括P-糖蛋白、多药耐药相关蛋白（multidrug resistance-associated proteins, MRPs）、乳腺癌耐药蛋白（breast cancer resistance protein, BCRP）等。人体小肠有效吸收的表面积为 200 m^2 左右，为防止肠道内有毒的物质进入体内，肠道表面分布有众多的外排转运体，包括P-糖蛋白和BCRP，它们和细胞色素等代谢酶一同起到限制药物渗透进入血液循环系统的作用^[16]。

药物分子的血浆蛋白结合能力也是评估药代动力学和毒性的一项重要指标。化合物的高血浆蛋白结合（high plasma protein binding）是西格列汀研发过程中遇到的一个重要问题。由于只有自由的、与血浆蛋白未结合的药物分子才会和药物靶标发生直接的相互作用，因此，高血浆蛋白结合的药物分子不仅会影响药物的分布、代谢和排泄，还会影响药物的药效。由于药物与血浆蛋白结合的程度会受患者的年龄、疾病状态以及同时服用其他药物的影响，FDA现已要求药物研发者必须针对具有高血浆蛋白结合的化合物进行额外的安全性研究^[17]。

临床严重的不良反应往往是导致药物研发失败的最关键的因素。在西格列汀的研发过程中，化合物的脱靶效应（off-target effect）和毒性致使项目团队放弃了许多具有良好生物活性和药代动力学性质的化合物。药物常见的毒性包括细胞毒性（cytotoxicity）、肝毒性（hepatotoxicity）、基因毒性（genotoxicity）、致癌毒性（carcinogenicity）、心脏毒性（cardiac toxicity）等。对于药物设计者来讲，在设计阶段就必须尽可能避免产生这些毒性的因素^[18]。这些毒副作用的来源不仅包括候选药物分子本身，也包括其代谢产物可能产生的毒副作用以及药物-药物相互作用（drug-drug interaction, DDI）引起的毒副作用。

人体内细胞色素P450酶超家族是代谢药物的主要酶，是一类主要存在于肝脏及肠道中的单加氧酶，多位于细胞内质网上，能催化多种内、外源物质的（包括大多数临床药物）氧化代谢过程。抑制或者诱导细胞色素P450酶会影响药物、体内激素和内源毒素，以及所有可能进入人体的物质（包括各种有毒物质）的代谢，所以维持正常的细胞色素P450酶的功能是机体健康的一个重要保证。许多慢性病或多重疾病患者需要服用多种药物或者需要长期联合用药，而抑制或者诱导细胞色素P450酶的药物可能会影响其他药物的药代动力学，从而影响血浆中药物的浓度，导致药物-药物之间产生相互作用。因此，防止药物可能引起的DDI也是药物设计者必须考虑的另一个重要内容^[19]。

在DPP-4抑制剂研发项目中，药物化学家遇到的最大挑战是解决化合物的hERG活性问题。hERG（human ether-à-go-go-related gene）在人的心肌、脑、肝脏、脾脏等组织中均有表达，其中心肌组织表达量最高。当hERG钾离子通道介导的电流通过细胞膜的能力受到抑制时，会导致QT间期延长综合征这一潜在的致命性疾病发生，造成严重的心律失常。目前，检测化合物对hERG钾离子通道的作用已成为FDA要求新药报批的必备资料，



以此规避药物心脏毒性的风险^[20]。

如果一个化合物具有较好的活性、选择性、药代动力学和药效动力学性质，在其正式被选为临床前候选物（pre-clinical candidate, PCC）之前，还需要做大量的脱靶活性和毒性测试。制药公司希望在设计阶段对化合物产生脱靶活性和毒性的可能性进行评估，以避免浪费已经投入的大量人力、物力和财力。杂泛性是这几年被不少公司用来判断药物的选择性和安全性的一个重要概念。杂泛性是指一个化合物可与多个靶标包括各种未知靶标发生作用，产生治疗作用和（或）不良反应。在分析了大量已知化合物与脱靶活性和毒性的关系后，研究者发现化合物的 $\text{clog}P$ 越大，杂泛性越高，即亲脂性分子倾向于与多个靶点结合^[21]。

第五节 早期 DPP-4 抑制剂的药效及毒性研究

早期临床试验结果显示，使用DPP-4抑制剂确实能降低糖尿病患者进食后血糖的升高幅度，提供了疗效的概念证明（proof-of-concept, POC）。但与此同时，早期DPP-4抑制剂显示出一定的毒副作用。

一、早期DPP-4抑制剂的药效研究

默沙东于1999年开始研发DPP-4抑制剂项目，在此之前已有两个公司的DPP-4抑制剂候选物进入临床试验阶段。诺华公司的DPP728（化合物2，图1-5-1A）和前东德Probiobdrug公司的P32/98（化合物3，图1-5-1B、图1-5-1C）的I期临床试验结果均显示，健康志愿者在第一次服药后均无明显的不良反应，血浆中活性GLP-1含量也均有所增加，用餐或直接口服葡萄糖后血糖的升高幅度也明显降低，与动物实验结果类似。

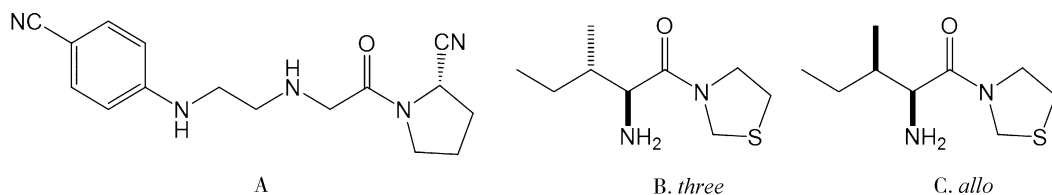


图 1-5-1 早期 DPP-4 抑制剂

A: 化合物2; B: 化合物3 (three); C: 化合物3的异构体4 (allo)

2002年，诺华公司报道了DPP-4抑制剂II期临床试验的初步结果。在93个2型糖尿病患者参加的为期4周的试验中，与服用安慰剂的对照组相比，用药患者用餐或直接口服葡萄糖后的血糖升高幅度、空腹血糖水平和24 h的血糖水平均明显下降，这样的结果首次提供了小分子DPP-4抑制剂可作为抗2型糖尿病药物的直接证据^[22]。这是一个重要的里程碑式的事件，称为“临床的概念证明（clinical proof of concept, cPOC）”。但由于化合物2在大鼠中的半衰期只有15 min，在人体中的半衰期也小于1 h，诺华公司最终决定终止了化合物2进一步的临床试验研究。

二、早期DPP-4抑制剂的毒性研究

2000年，默沙东从Probiodrug公司购买了DPP-4抑制剂化合物3和其异构体4的研发权。在临床前安全性评价化合物3和其异构体4的试验中，项目团队发现实验大鼠发生贫血、血小板减少等不良症状，最终可导致大鼠死亡；在犬的试验中也出现急性胃肠道出血的明显副作用^[23]。对于动物实验中发现的毒性，项目团队及时进行了正确的分析，这是最终成功开发DPP-4抑制剂成为药物的一个关键点。由此可见，分析和研究明确观察到的动物毒副作用是否与抑制DPP-4本身相关[称为机制相关毒性（mechanism-based toxicity）]是至关重要的。如果毒性与抑制DPP-4相关，那么DPP-4就不可能是成药靶标（druggable target），DPP-4抑制剂也不可能进一步开发成为新药。

为明确化合物3和其异构体4引起的毒副作用的来源，项目团队对其脱靶活性进行了深入研究。在药物化学研究策略中，脱靶活性是通过反筛选（counter screen）来实现的，一般分为特异性反筛选（specific counter screen）和覆盖式反筛选（global counter screen）。特异性反筛选通常只针对与靶标非常相近的蛋白质家族成员或同家族的亚型进行评估，而覆盖式反筛选则是测定该化合物对各种已知蛋白质家族成员的作用。因为后者数量众多、实验耗资巨大，一般只用于对少数结构上具有代表性的化合物进行评价，首选临床研究的候选物。在MDS Pharma Services（一家为制药公司提供覆盖式反筛选的CRO公司）的全面反筛选中，化合物3和其异构体4的杂泛性很低，未发现任何 $IC_{50} < 40 \mu\text{mol/L}$ 的靶标外活性。但是在进行特异性反筛选时，化合物3和其异构体（化合物4）对DPP-8和DPP-9显示出较强的抑制活性（图1-5-2A、图1-5-2B）。

根据化合物3（*threo* isomer）和其异构体（化合物4）（*allo* isomer）的动物毒性试验结果，研究者推测化合物的毒副作用应该与抑制DPP-4无关。无论在体内还是体外实验中，化合物3和其异构体（化合物4）对DPP-4的抑制活性几乎完全相同，但它们在动物试验中显示的毒性却差别很大。从动物给药剂量及测得的血药浓度来推算，*allo*-异构体（化合物4）的毒性比化合物3的毒性强约10倍，这与它们分别对DPP-4的抑制活性剂量并不相关，因此基本排除了机制相关毒性的可能性^[23]。分子的毒性（molecule-based toxicity）很可能来源于早期DPP-4抑制剂对靶标外（off-target）酰胺酶家族中其他成员的作用，如对DPP-8/9的非特异性抑制作用。

为进一步证实上述结果，项目团队挑选大鼠中药代动力学参数基本相同、但活性选择性不同的化合物5和化合物6 [化合物5（图1-5-2C）选择性抑制DPP-8/9，化合物6（图1-5-2D）选择性抑制DPP-7（又称QPP，quiescent cell proline peptidase）] 在大鼠中进行了为期两周的比较毒理试验。结果显示：对DPP-8/9专一性抑制的化合物5在大鼠中产生与*allo*-异构体（化合物4）类似的毒性，但对QPP专一性抑制化合物6除在很高剂量时稍有异常之外，未观察到明显毒性。此外，DPP-4特异性抑制剂也未发现任何明显的病变^[23]。这一结果非常有力地支持了项目团队的推测，也突显出DPP抑制剂研发过程中化合物选择性的重要性。



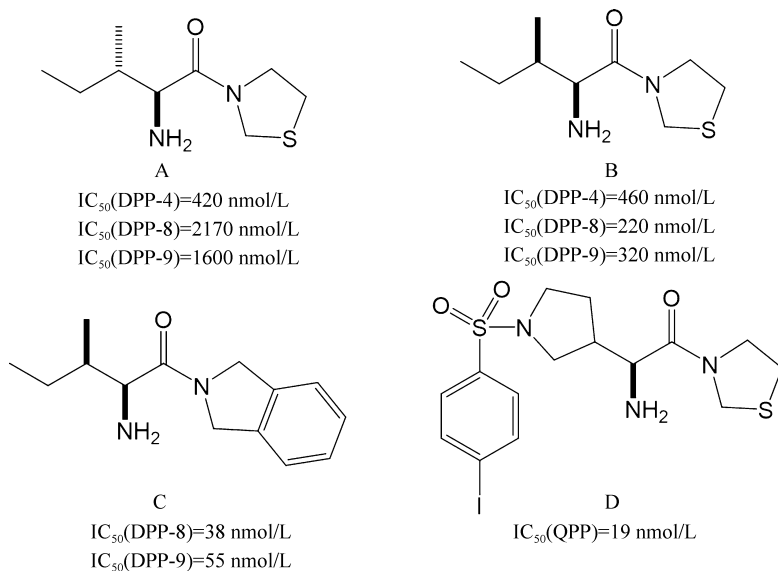


图 1-5-2 早期 DPP-4 抑制剂的选择性

A: 化合物3; B: 化合物3的异构体(化合物4); C: 化合物5; D: 化合物6; QPP: quiescent cell proline peptidase

三、DPP-4抑制剂优化策略的确定

默沙东项目团队最早设定的目标是找到一个同类最优(best-in-class, BIC)的药物。成为同类最优的药物,除了具有很好的活性、药代动力学和药效动力学性质以外,还必须具有很好的选择性和安全性,实际上仅具有最优生物活性的化合物很少成为同类最优的药物。因此,项目团队认为,安全有效的DPP-4抑制剂除了应该具备高DPP-4抑制活性之外,还必须对肽酶家族的其他成员有超过1000倍以上的选择性,尤其是对DPP-8和DPP-9的选择性最为重要。同时,由于糖尿病是慢性疾病,用药周期很长,药物不能对其他具有重要生物功能的蛋白质家族产生作用,如各类离子通道(ion channel)对心脏的正常工作至关重要,不能被扰乱。此外,细胞色素酶是非常重要的代谢酶家族,化合物对细胞色素酶的抑制或诱导都有可能紊乱药物的正常代谢,从而导致DDI,继而产生严重后果。因此,必须将对离子通道、细胞色素酶有作用的分子筛除掉。

第六节 先导化合物的确认

一、什么是理想的先导化合物?

在药物化学研究中,正确选择和设计先导化合物是成功开发药物最关键的第一步。在优化先导化合物成为药物的过程中,往往需要同时优化7~9个不同的参数,所以化合物的合成路线越简单越好。在优化活性和渗透性时,化合物的分子量和亲脂性会增加许多,所以先导化合物的分子量和亲脂性(通过clogP衡量)越小越好。

二、 α -氨基酸衍生的DPP-4抑制剂

在项目团队尚未发现抑制DPP8/9会造成动物多器官毒性以前，早期的DPP-4抑制剂先导化合物（图1-6-1A）都是由 α -氨基酸7衍生而来。化合物8便是一种理想的先导化合物，它合成简单，柔性小，具有非常好的药效学和药代动力学性质及安全性数据，但进一步测试发现化合物8（图1-6-1B）对DPP-8和DPP-9缺乏足够的选择性^[24]。为了解决选择性的问题，项目团队对化合物8进行了较大改造，发现许多活性很好的化合物（如化合物9和化合物10，图1-6-1C，图1-6-1D），但它们对DPP-8和DPP-9仍缺乏足够的选择性。此时，通过高通量筛选技术，项目团队发现了全新的 β -氨基酸衍生物化合物，对DPP-8和DPP-9有一定的选择性，所以项目团队决定暂停对 α -氨基酸衍生物的研究，转而优化 β -氨基酸衍生物系列。由此可见，储备结构多样性的先导化合物，可以帮助项目团队及时地调整优化方向。

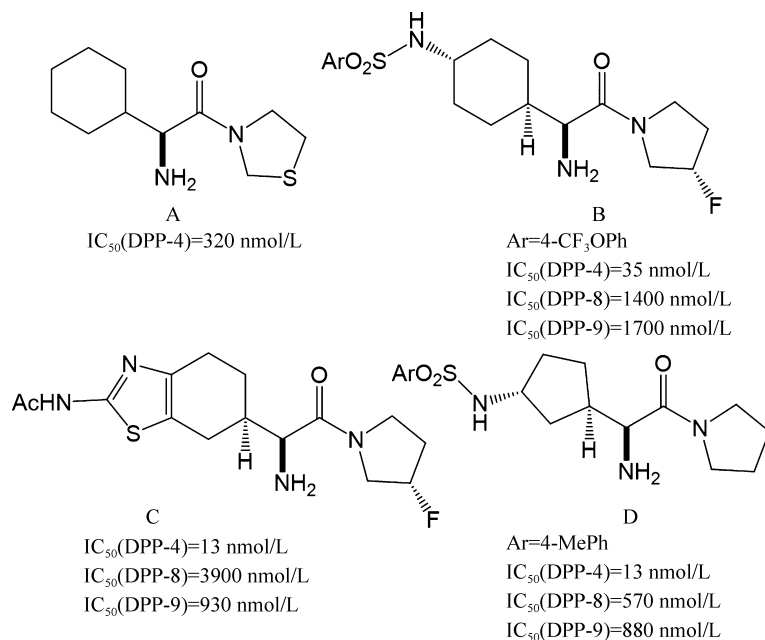


图 1-6-1 早期 α -氨基酸衍生的二肽酰肽酶-4 抑制剂

A: 化合物7; B: 化合物8; C: 化合物9; D: 化合物10

药物研发是个系统工程，候选物可能在成药之前的任何一个环节出现问题，制药公司通常会另外选择一个高质量、结构全新的备选化合物（backup compound）作为应急方案。在发现西格列汀（ β -氨基酸衍生物）以后（开展临床研究之前），项目团队重新研究了 α -氨基酸衍生的DPP-4抑制剂，观察到化合物3和其异构体4的唯一差别是 β -位置的立体异构不同，两者对DPP-8的抑制活性竟相差约10倍。项目团队在化合物3的基础上进一步优化以提高对DPP-8/9的选择性，合成了化合物11（图1-6-2A）。生物活性测试结果表明，化合物11对DPP-8和DPP-9的选择性非常好，但是活性不够理想。经过分析以往 α -氨基酸衍生的DPP-4抑制剂的数据，项目团队设想可通过在苯环上引入不同取代基团来提高该系列化合物的生物活性。比如，在化合物11（ $IC_{50}=1400\text{ nmol/L}$ ）苯基的4-位

引入4-氟苯基,得到化合物12 ($IC_{50} = 64 \text{ nmol/L}$, 图1-6-2B), 其活性提高了20倍, 同时保持对DPP-8和DPP-9较高的选择性, 其药物代谢与药代动力学 (drug metabolism and pharmacokinetics, DMPK) 性质也很好。遗憾的是, 进一步研究发现化合物12具有很高的血浆蛋白结合能力, 对hERG的选择性也较差。继续对化合物12的性质进行微调 (fine-tuning), 通过降低化合物的亲脂性和引进极性基团, 得到了吡啶酮 (pyridone) 化合物13 (图1-6-2C) ^[25,26], 以及进一步优化的三唑 (triazole) 化合物14 (图1-6-2D)。化合物14不仅活性非常好, 也具有很好的选择性和ADMET性质, 最终被选择为西格列汀的备选化合物 (backup compound) ^[27]。从发现化合物11到发现化合物14, 项目团队只合成了300余个化合物, 证明了基于性质的药物设计具有非常高效的策略优势。

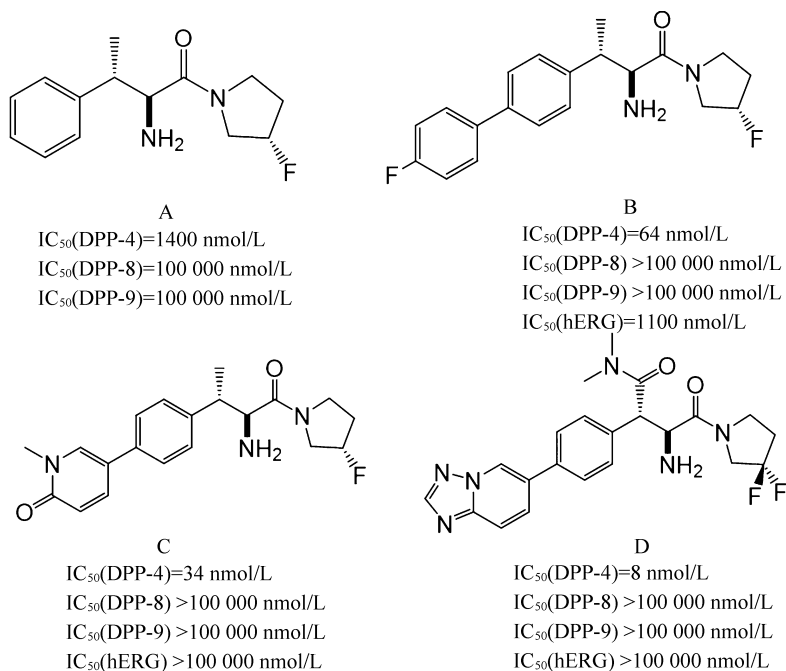


图 1-6-2 α -氨基酸衍生的 DPP-4 抑制剂的优化

A: 化合物11; B: 化合物12; C: 化合物13; D: 化合物14

三、 β -氨基酸系列先导化合物的确认

随着高通量筛选技术的发展, 项目团队发现了全新的 β -氨基酸衍生物化合物15 (图1-6-3A) 和化合物16 (图1-6-3B), 两者对DPP-8和DPP-9有一定的选择性。从先导化合物分子量的角度来分析, 化合物15和16非常不理想。项目团队很快发现这个系列化合物的口服生物利用度非常差。对化合物15进行简化, 发现化合物17 (图1-6-3C) 的活性和化合物15相当, 但分子量 ($MW = 250 \text{ Da}$) 降低较多, 为进一步 β -氨基酸衍生物优化奠定了良好基础。对苯基上的取代基进行深入研究, 项目团队发现了2,4,5-三氟代衍生物化合物18 (图1-6-3D) ($IC_{50} = 119 \text{ nmol/L}$) ^[28]。

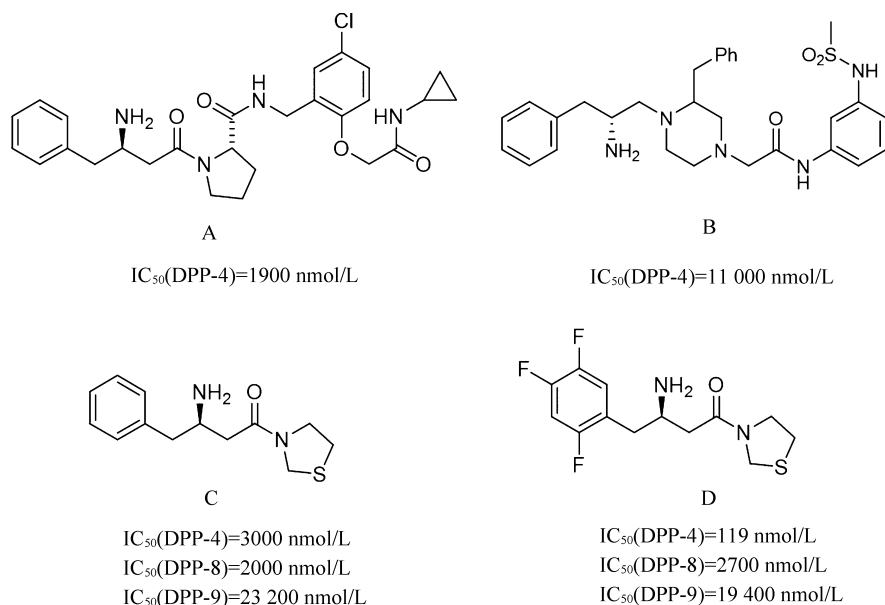


图 1-6-3 高通量筛选中标物的发现与优化

A: 化合物15; B: 化合物16; C: 化合物17; D: 化合物18

化合物18的活性虽然不错,但它对DPP-8和DPP-9的选择性仍旧不够理想。进一步的构-效关系研究发现了新的先导化合物19(图1-6-4)^[29]。化合物19具备了分子量小($MW = 355 \text{ Da}$)、活性高($IC_{50} = 140 \text{ nmol/L}$)、选择性好和容易修饰等特点,为药物的进一步优化提供了必要条件。

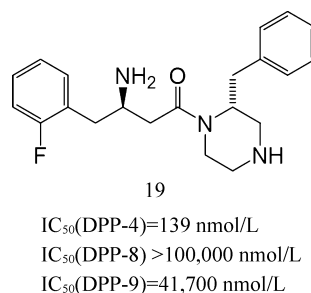


图 1-6-4 先导化合物

第七节 β -氨基酸系列先导化合物的优化

先导化合物与进入临床研究的候选物往往相差甚远。以化合物19为例,研究者发现,在动物体内,2-氟代苯环会形成谷胱甘肽(*glutathione*)的加合物。另外,哌嗪环(*piperazine*)因为体内发生氧化而缺乏代谢稳定性(*metabolic stability*),导致产生口服生物利用度低等不理想的药代动力学性质。

为了消除哌嗪环在体内易被氧化的现象,提高DPP-4抑制剂代谢的稳定性,同时保持对DPP-8/9和hERG的选择性,研发团队进行了一系列化学设计研究,其中最成功的是并环策略化合物21~化合物23(图1-7-1B~D)^[30,31]。

在随后合成的双环哌嗪衍生物里,项目团队发现了许多具有高活性、高选择性的DPP-4抑制剂,比如三唑哌嗪的并环化合物21~化合物23。化合物21($IC_{50} = 460 \text{ nmol/L}$)与非并环哌嗪类似物化合物20(图1-7-1A)($IC_{50} = 3100 \text{ nmol/L}$)相比,活性得到了显著提高。在药物代谢研究中,乙基取代的衍生物22($IC_{50} = 230 \text{ nmol/L}$)证实了三唑哌嗪的并环可以提高体内氧化代谢的稳定性,但其口服生物利用度仍然不理想^[31]。化合物22的渗

透性非常差, 由于提高分子的亲脂性可以提高化合物的渗透性, 三氟甲基衍生物化合物23 ($IC_{50} = 130 \text{ nmol/L}$) 在提高DPP-4抑制活性的同时, 显示出很好的代谢稳定性和口服生物利用度 [$F(\text{rat}) = 44\%$]。构-效关系研究还发现, “左手边”苯环上氟取代可以进一步提高化合物的活性。将两者结合在一起, 得到了如图1-7-2所示的系列化合物。其中2,5-氟代衍生物化合物24 ($IC_{50} = 27 \text{ nmol/L}$, 图1-7-2A) 大鼠的口服生物利用度达到50%, 2,4,5-三氟代衍生物1 ($IC_{50} = 18 \text{ nmol/L}$) 大鼠的口服生物利用度更是提高到76%。化合物24的衍生物1 (西格列汀) 最终被开发成为同类最优新药。

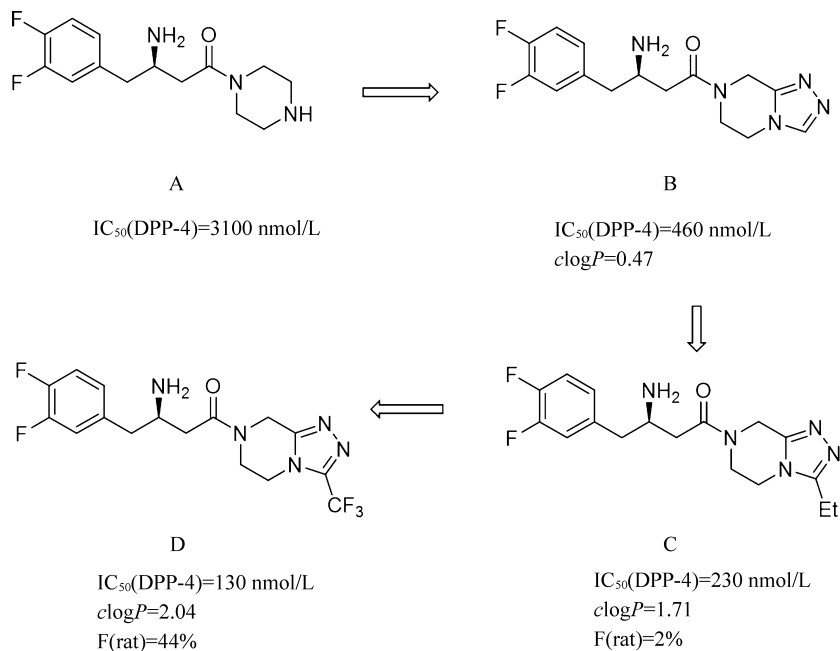


图 1-7-1 优化化合物代谢, 渗透性和生物利用度的策略

A: 化合物20; B: 化合物21; C: 化合物22; D: 化合物23

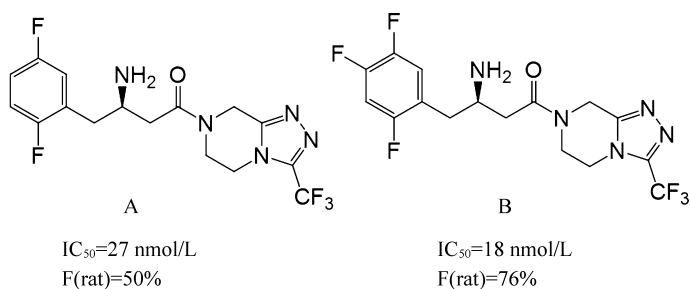


图 1-7-2 临床前候选化合物 24 和 JANUVIA™ (sitagliptin, 西格列汀)

A: 化合物24; B: 化合物24衍生物1 (sitagliptin, 西格列汀)

第八节 新药西格列汀的临床研究

由先导化合物19衍生而来的一系列以 β -氨基酸为基础的DPP-4抑制剂普遍具有良好的选择性，它们对于DPP-8和DPP-9的抑制活性一般较弱（ $IC_{50} > 45 \mu\text{mol/L}$ ）。例如，化合物24和其衍生物（化合物1）对DPP-8和DPP-9选择性均超过1000倍，达到了预期的指标。

在其他反筛选方面，化合物24和其衍生物（化合物1）也都具有很好的选择性。它们对于代谢酶CYP450均无明显抑制或诱导作用，对各类离子通道也均无明显阻碍作用，也未发现值得注意或者值得跟踪的体外脱靶活性。

在小鼠进行的药效实验中，研发者发现化合物24对小鼠DPP-4的抑制活性略有降低（ $IC_{50} = 100 \text{ nmol/L}$ ），但在3 mg/kg口服剂量下即可达到高幅度降低血糖的最佳疗效，其血药浓度可达到700 nmol/L。对于由高脂食物和链脲左菌素（streptozotocin, HFD/STZ）诱发的糖尿病小鼠，在长期服用化合物24后，根据其糖化血红蛋白（glycosylated hemoglobin, HbA1c）的含量可以确认其血糖水平的降低与化合物24的用量相关^[32]。病理切片研究结果显示，这些长期服药的糖尿病小鼠胰岛内的 β -细胞有所增加，为抑制DPP-4有助于恢复胰脏和 β -细胞的功能和胰岛的正常形态提供了依据。

针对化合物24和其衍生物1为期两周的安全性评价（safety evaluation or safety assessment）的研究结果显示，在实验条件下，它们对大鼠和实验犬均无可观测到的毒性。但是在犬的心血管实验模型试验中，化合物24在剂量为10 mg/kg（血浆浓度为46 $\mu\text{mol/L}$ ）时引起了血压和心率降低，以及心电图P波和R波的间隔增加等不良反应。这些现象在剂量为1.0 mg/kg（血浆浓度为6.5 $\mu\text{mol/L}$ ）时则未出现，这在安全性评价实验中被称为“无毒性剂量”（no effect level, NOEL）。在相同的心血管模型试验中，衍生物1则未出现不良反应，其安全性的无毒性剂量为10 mg/kg（血浆浓度为59 $\mu\text{mol/L}$ ），因而安全系数更高。进一步的安全性评价也证实了化合物1是更适合作为临床研究的候选药物（clinical candidate）。

随后的临床试验结果也证实，化合物1（临床代号MK-0431）对DPP-4的抑制与剂量相关。当剂量达到100 mg以上时，MK-0431对DPP-4的抑制活性可以在24 h内保持在80%以上^[33]，这也是在临床前动物实验中实现最大药效所需要的抑制水平。MK-0431在人体内的半衰期可达到8~14 h，适合于每天一次服药（QD dosing），人的口服生物利用度高达87%^[34]。服药后的糖尿病患者再摄入葡萄糖时，其血糖的升高幅度降低，同时伴随着升高的生物标志物有：活性GLP-1、GIP、胰岛素和C-肽，降低的生物标志物是胰高血糖素^[35]。在一项为期24周的单独用药的Ⅲ期临床试验中，与安慰剂相比，糖尿病患者的HbA1c含量从基线的8.0%下降了0.6%~0.8%^[36]。在HbA1c基线更高（>9%）的糖尿病患者人群中，MK-0431引起的HbA1c下降更为显著，可达1.52%。用药患者的空腹血糖和餐后血糖水平均有所下降，还有迹象表明患者的 β -细胞功能有所恢复。临床试验中未发现因服药而引起的低血糖和体重增加等不良反应。

除单独使用外，临床试验结果还表明，MK-0431（后被命名为西格列汀）可以有效地



与其他降血糖药物，（如匹格列酮（pioglitazone）^[37]或二甲双胍（metformin）^[38]）联合用药。在一项为期52周的比较试验里，对服用二甲双胍（ ≥ 1500 mg）的糖尿病患者分别用西格列汀（100 mg）和磺胺酰脲类药物格列吡嗪（glipizide, 5 ~ 20 mg）进行联合治疗^[39]。整个糖尿病患者组的HbA1c下降了0.67%，其中HbA1c总量下降到7%以下的糖尿病患者在西格列汀组中占63%，在格列吡嗪组中占59%，基本相同。但是格列吡嗪组中有32%的患者发生了低血糖，而西格列汀组中只有4.9%发生低血糖。另外，格列吡嗪组患者平均体重增加了1.5 kg，西格列汀组患者平均体重下降了1.1 kg。

2006年10月，FDA正式批准了由美国默沙东制药公司研发的第一个格列汀（-gliptin，DPP-4抑制剂药名的指定词根）类药物——西格列汀作为新药上市，为2型糖尿病的治疗提供了一个全新、安全和有效的治疗手段。在中国，西格列汀被批准用于配合饮食控制和运动，改善2型糖尿病患者血糖控制的单药治疗。另外西格列汀和二甲双胍固定剂量的复方制剂也已经在美国和欧盟被批准上市。

第九节 西格列汀的工艺化学

一、什么是工艺化学？

具体来说，小分子药物合成的工艺化学（process chemistry）就是实用有机合成，涉及从实验室规模的化学合成到中试生产、再到商业生产3个阶段。

工艺研究部门的任务及责任主要包括两部分：①设计和开发可放大量生产的化学工艺，这些工艺应满足合成线路短、原子效率（atom economy）高，成本效益高，技术可靠且对环境影响（environmental impacts）低等重要标准；②制备符合标准的药物活性成分（也称为原料药，active pharmaceutical ingredient, API），在候选药物的整个临床开发阶段中能够支持药物研发的顺利进行，并对上市以后的药物生产提供技术支持^[40]。

除此之外，工艺开发部还需要提供全面的文档记录，详细描述可重复的生产方法。在许多制药公司中，工艺部门还负有对药物商业流程的整个生命周期的管理责任（life cycle management）。

新化合物的合成都是首先由药物化学部门完成的。为了提高效率，同时满足早期筛选的需求，首次合成的新化合物一般都是毫克级，用于体外的生物化学实验、初步筛选与特异性反筛选，以及药代动力学实验等初步评价。随着筛选及优化的推进，通常需要药物化学部门进一步提供1 ~ 100 g的高纯度化合物，进行各种动物模型试验和早期毒理评估等。

通过这些筛选和评估之后，被选出的化合物就会进入新药临床试验申请（investigational new drug, IND）启动研究（enabling studies）阶段，需要的化合物量会跃增至千克级水平。在绝大多数情况下，药物化学部门为早期筛选建立起来的合成方法已经不能满足千克级的生产需要，所以工艺化学部在进入IND之前就会正式启动该项目的工艺研究，以实现从药物化学到工艺化学的无缝衔接，保证项目的顺利进行。

近几十年来，规模化生产小分子药物合成工艺的设计和开发受到了极大关注，已完全



整合到制药研发部门的价值链中，被视为制药行业的核心能力之一，需要很高的科学素养和实验技术能力。

西格列汀的合成，以药物化学部的首次合成为基础，经历工艺部门的迭代升级，实现了大规模、可持续的稳定量产，两次获得美国“总统绿色化学挑战奖”（Presidential Green Chemistry Challenge Award），堪称药物工艺化学的经典。

二、化学反应的效率

说到化学反应的效率，大家一般首先想到的是产率，这是从反应物到产物的转化率，是衡量化学反应效率的重要参数之一，可以通过改变温度、压力、溶剂和反应试剂等各种条件来进行逐个优化，得到理想的产率。除了实际获得的产率之外，衡量化学反应效率还有另一个更为基础的指标，就是原子经济（atom economy），这个概念是美国著名有机合成化学家特罗斯特（Trost）在1991年提出的^[41]。特罗斯特教授认为，对于任何一个化学反应，有多少反应物最终可以被包括在预设的产物中，可以称其为原子经济。而最高程度的原子经济方法是“将两个或多个反应物与仅需催化量的任何其他试剂的简单组合而构成。”他因此提出了如下的原子经济效率的计算公式：

$$\text{原子经济效率}(\%) = \frac{\text{预设产物的分子量}}{\text{所需当量的反应物与试剂的分子量总和}} \times 100\%$$

这个参数不仅考虑了从反应物到产物的转化，同时还体现出反应废料的产生。根据这个计算公式，如维蒂希反应（Wittig reaction），因为等当量的辅基三苯基膦并不能被包括在产物分子里，所以原子效率很低，不适合用于大规模化学工艺流程。相反，狄尔斯-阿尔德环加成反应（Diels-Alder reaction）仅需催化剂就可以将两个反应物结合在一起生成产物，无其他副产物，具有很高的原子效率。有兴趣的同学可以用这个公式去计算所学过的化学反应的原子效率，进而思考如何才能优化得到最佳反应条件，在提高产率的同时减少副产物和废料。

通过简单的计算不难看出，催化反应与等当量的辅基反应相比在原子经济上有很大的优势。西格列汀的工艺研究便是从开始应用手性辅基（chiral auxiliary）反应，到最后发展成为高效率的催化反应工艺，符合绿色、可持续的现代工艺化学理念。

从有机合成的角度分析，西格列汀的分子结构并不复杂，只有一个手性碳原子，如何建立这个手性中心自然成为工艺研究的焦点。限于篇幅，这一节重点讨论手性β-氨基酸的合成工艺，杂环部分的合成工艺不包括在内，有兴趣的读者可以自己查阅文献^[42]。

三、西格列汀的药物化学合成：手性辅基的应用

西格列汀的药物化学合成最初采用了手性辅基的方法。这是一项成熟的不对称合成技术，应用非常广泛，适用于引进多种不同的化学官能团^[43]。在药物化学的早期开发阶段，时间紧、所需的化合物量相对较少，而且手性辅基的诱导效应预见性好，线路开发相对快捷，经常会成为首选方法。

目前可供选择的手性辅基有许多种，最常见的有8-苯基薄荷醇、噁唑烷酮、硫酸伪麻黄碱、叔丁基亚磺酰胺等，可用于多种不同的化学反应，包括羟醛缩合反应、烷基化反



应、狄尔斯-阿尔德环加成反应等常用的碳碳键反应，以及连接各种杂原子等。

α -氨基酸的手性合成方法可有多种选择，已被广泛研究。从 α -氨基酸到 β -氨基酸的转化也有多种成熟的方法，所以默沙东的药物化学团队首先尝试了手性辅基的重氮化反应以及烷基化反应，均获得成功（图1-9-1）^[31]。在小规模合成中这两条线路的实验结果大致相同。

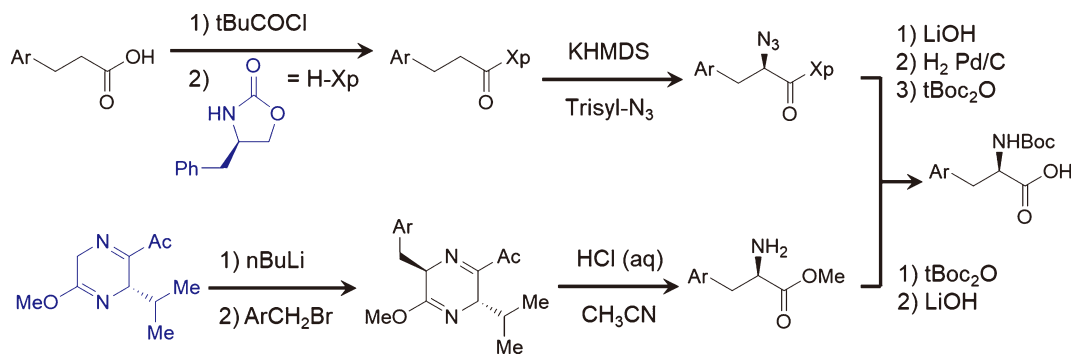


图 1-9-1 西格列汀药物化学合成线路中手性辅基的应用

从 α -氨基酸到 β -氨基酸的转化可通过Wolff重排反应实现（图1-9-2）。该反应条件温和、产率高，但是应用剧毒易爆的重氮甲烷，也不适合于放大。

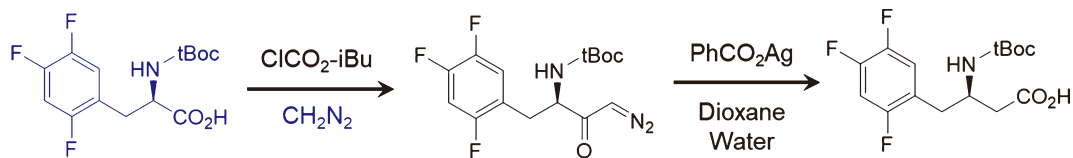


图 1-9-2 西格列汀药物化学合成线路中从 α -氨基酸到 β -氨基酸的转化

手性辅基的成功应用为药物化学研究和早期的动物实验及时提供了足够的西格列汀样品，保证了项目的快速推进。但是，手性辅基也具有一定的局限性。除了上述提到的剧毒易爆重氮试剂之外，连接和去除手性辅基增加了两个线性的合成步骤，不但效率有所降低，而且相关的化学溶剂和试剂的用量、能源消耗都会显著增加，液体和固体废料的处理和排放都会受到影响。从原子经济学的角度看，一个当量的手性辅基不能被包含在产物结构中，完全成了废料，即使可以回收也会增加工序和成本。

四、西格列汀工艺化学：不对称催化氢化的应用

与药物化学部交接之后，工艺化学部认真评估了前期的合成路线，认为应用该路线进行千克级的放量生产不仅步骤长、效率低、环境影响大，而且存在巨大的安全隐患。工艺部门决定放弃手性辅基路线，采用不对称催化氢化反应来合成 β -氨基酸的手性中心。

手性催化氢化反应在有机合成中同样具有非常广泛的应用^[44]。研究手性催化氢化反应的领军人物是日本化学家野依良治（Ryoji Noyori）和美国化学家威廉·诺尔斯（William Knowles），他们于2001年共同获得了诺贝尔化学奖。

目前可供选择的手性氢化催化剂一般由铂族的过渡金属与手性配体组成。常见的过渡金属有钌(Ru)、铑(Rh)、钯(Pd)、铱(Ir)等,最常见的手性配体是烷基膦配体(phosphine ligands),可用于多种不同的催化氢化反应,包括碳碳双键和三键的加氢、羰基的还原,以及烯胺的加氢等常用的还原反应。

(一) 第一代西格列汀工艺化学: 手性氢化工艺合成

第一代工艺化学选择了羰基的不对称还原作为核心反应(图1-9-3)。以廉价的羧酸为起始原料,经丙二酸酯反应延展碳链形成手性氢化的底物 β -羰基酯。随后,研究者通过系统地筛选选择和优化手性氢化反应的催化剂(金属与配体)和反应条件,最后锁定了钌与手性联萘的络合物,催化剂的负载量 $<0.1\text{ mol}\%$,产物经水解后可以分离得到83%产率和94%对映体过量(ce)的羧酸^[42]。

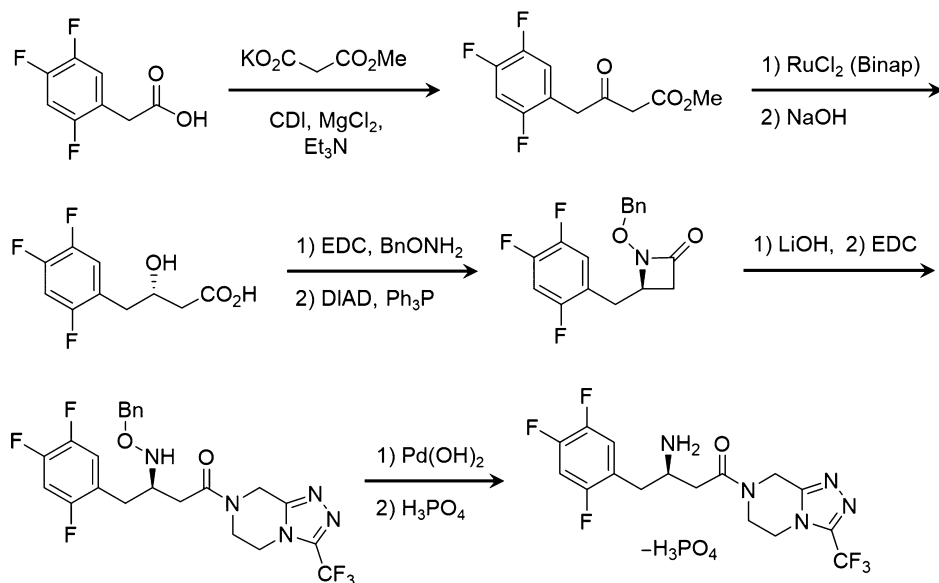


图 1-9-3 第一代西格列汀工艺合成线路

经过标准的EDC酰胺耦联反应之后,通过光延反应(Mitsunobu reaction)形成了四元环内酰胺,并同时实现了手性中心的构型翻转,从而完成了手性 β -氨基酸的不对称合成。整条合成路线经过了9个线性步骤,终产率为52%,平均每步反应产率为93%,被成功用于100千克级的西格列汀合成。

从有机合成的原子经济学角度看,这条工艺路线采用的反应,如光延反应、EDC酰胺耦联反应等,都必须加入等当量的三苯基膦和偶氮二甲酸二异丙酯(DIAD——光延反应的关键中间体)、EDC(酰胺耦联反应)等反应试剂,不能被包括在产物的结构中,所以效率不高,会产生大量废料,对环境的影响也很大,不利于更大规模的生产。

(二) 第二代西格列汀工艺化学: 手性氢化工艺合成

为了进一步提高合成效率,降低对环境的影响,默沙东工艺部继续优化了前体合成的绿色工艺,同时探索出首例无氮取代基的烯胺的不对称催化氢化反应,成功地开发出以烯胺不对称催化氢化为核心的第二代西格列汀生产工艺。

在新的前体合成工艺中,米氏酸的应用实现了从羧酸的延展、到预制杂环胺耦合、再到烯胺转化的“一锅”法合成,大大简化了工艺流程。反应副产物主要是二氧化碳、丙酮、叔丁醇等水溶性小分子,去除了大量对环境有害的化学试剂(图1-9-4)^[45]。

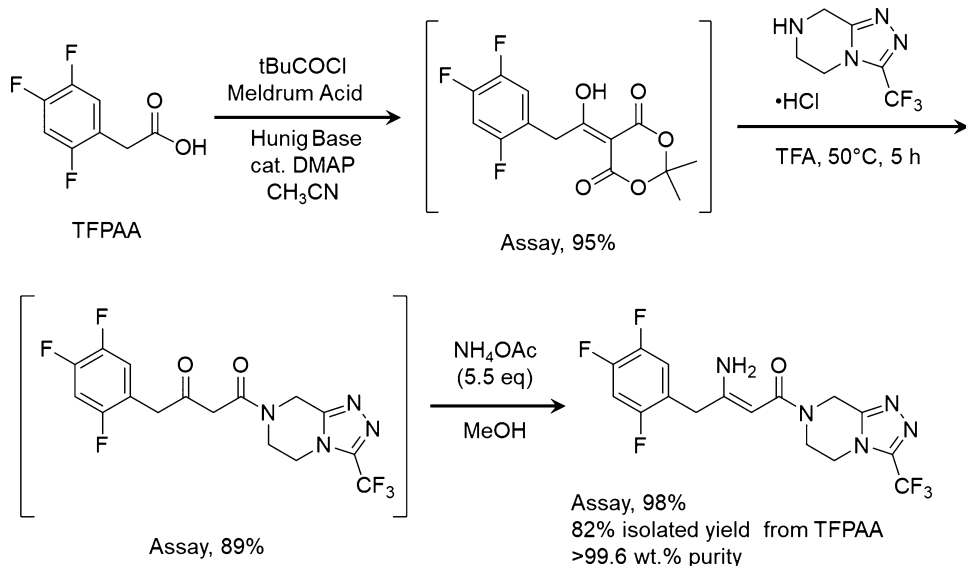


图 1-9-4 第 2 代西格列汀工艺合成线路: 氢化前体的“一锅”法合成

默沙东工艺团队经过系统筛选后,首次发现了以金属铈为核心的高效不对称烯胺催化氢化体系,后经优化首次用于工业生产中(图1-9-5)^[44]。

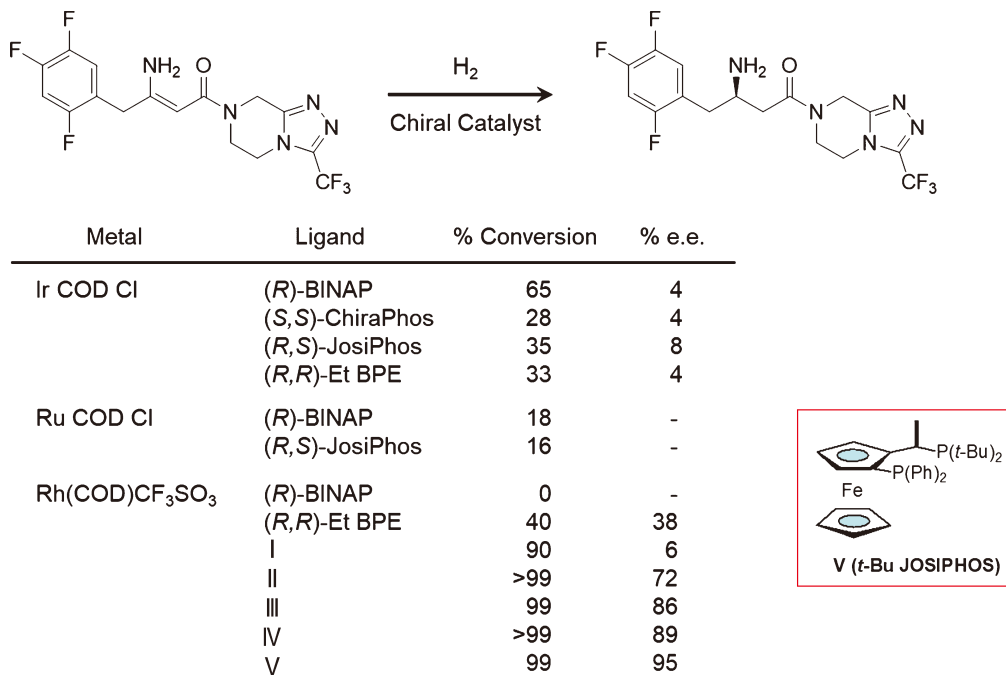


图 1-9-5 第 2 代西格列汀工艺合成线路的关键步骤: 筛选不对称氢化催化剂

从生产成本和收益角度来看,这项新工艺可以回收超过95%的贵金属催化剂铑,大大降低了催化剂的成本;同时新工艺只有3个操作步骤,成本下降70%,总体收益增加近50%。与此同时,以不对称催化氢化为核心的生产工艺使生产西格列汀原料药的工业废料也降低了80%,因此该法荣获了2006年美国总统绿色化学挑战奖。据估计,在西格列汀的整个商品周期里,这一改进使默沙东的生产部门减少了化工废料约15万吨,这是一个惊人的数字。

(三) 第三代西格列汀工艺化学: 酶催化的应用

上述工艺流程中,烯胺催化氢化反应的不对称选择性仍仅为95%,必须经过一步游离碱的重结晶才能达到符合要求的99.7%的原料药标准;同时,催化氢化反应必须在15~20个大气压条件下进行,需要特殊的耐高压反应设备;另外,回收贵金属催化剂铑,以及清除产物中残留的痕量金属铑,也都需要增加操作步骤和生产成本。

为了进一步提高不对称选择性、简化工艺,工艺部门继续成功开发了转氨酶催化工艺。新工艺取代了贵金属催化氢化反应,不再需要任何金属(铑和铁)试剂和高压反应釜设备,进一步大幅降低了资源消耗(图1-9-6)。



图 1-9-6 第三代西格列汀工艺合成路线: 转氨酶还原胺化?

酶催化工艺是新型生物化学工艺,近年来得到快速发展,其中突破性的“酶定向进化”(directed evolution of enzymes)技术最为关键^[46]。2018年,对此项研究做出重要贡献的法兰西斯·阿诺德(Frances Arnold)荣获了诺贝尔化学奖。

开发酶催化工艺的关键是发现或发明适用的酶。对于西格列汀的合成工艺来说,就是找到或开发出一种高转化率、高选择性并适用于大规模工业化生产的转氨酶。

西格列汀是一个完全由人工设计合成的化合物。经过筛选,最初能收集到的转氨酶都未检测到烯氨转换的活性。基于“酶定向进化”技术,采用“诱导”的方式,从化学结构上的相似处着手,研发者合成了一些反应过渡态中间体,使原始转氨酶通过一步一步地“学习”,逐步进化出一个可以检测到转化率的转氨酶,但当时这个酶的转化率仅有0.5%(图1-9-7)。

研发者再接再厉,以原代转氨酶为起点,再次采用定点突变技术对该转氨酶进行随机人工突变,筛选出转化率更好的“突变酶”,使转氨酶完成了一个轮次的“进化”。第二个轮次则以此新“进化”的转氨酶为起点,再进行随机的定点突变和筛选,直到获得满意的结果。在西格列汀合成工艺中烯氨转换酶的筛选及优化过程中,仅用了11个“突变-筛选”轮次便将转氨酶从转化率0.5%提高到95%以上,“进化”之后的最佳转氨酶的催化效率比原代转氨酶提高了惊人的25 000倍。更重要的是,在整个转氨酶的进化过程中,不对称选择性的比率始终保持在99.9%以上(图1-9-8)^[47]。

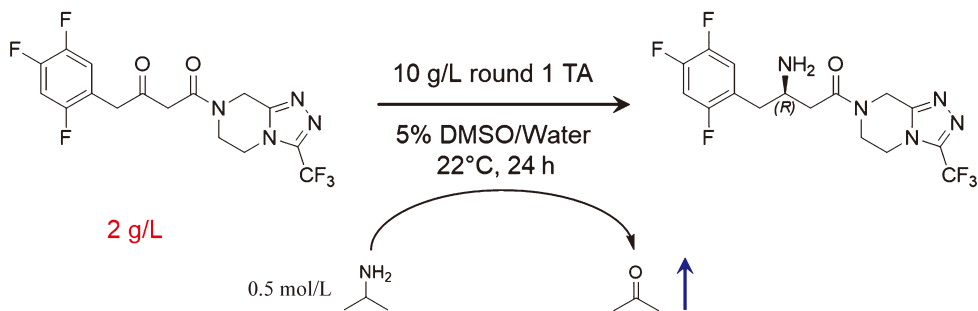
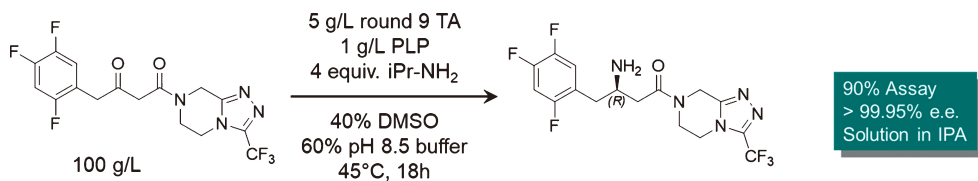


图 1-9-7 第3代西格列汀工艺合成路线: 转化率为0.5%的原代转氨酶



终止反应: 89% ~ 97% sitagliptin, 3% ~ 10% ketoamide, 1% ~ 4% dimer

图 1-9-8 第3代西格列汀工艺合成路线: 进化之后的高效转氨酶

转氨酶催化新工艺将现有设备的生产力提高了56%，总产率提高了10%~13%，整体废物产生量减少了19%，该工艺路线于2010年再次获得美国总统绿色化学挑战奖。

五、绿色化学大有可为

绿色化学 (green chemistry) 或可持续化学 (sustainable chemistry) 的定义是“减少或消除有害物质使用的化学工艺流程”。近年来，随着全球工业化进程的加速以及气候变化与资源损耗的日趋严重，人类对绿色化学工艺和消除有害化工产品的期望值也越来越高，绿色化学在工艺研发中的重要性必将受到越来越多的关注^[48]，绿色化学对未来人类的生活环境也必将起到极为重要的作用。

本章以抗糖尿病药物西格列汀工业制备的工艺绿色化学为结束内容，希望强调可持续化学在药物化学中的重要性，并可以用原子经济学的观点去分析化学反应对人类生活环境的影响。当然，我们也希望西格列汀的工艺研究历史能激发更多的同学们和读者贡献自己的智慧、青春和经验，将目光投向更加可持续的化学工艺。最后，借用一句诺贝尔奖获得者野依良治教授的话鼓励有志的同学们：“绿色化学不仅仅是一个流行的口头禅，而是人类生存之关键。”



数字资源