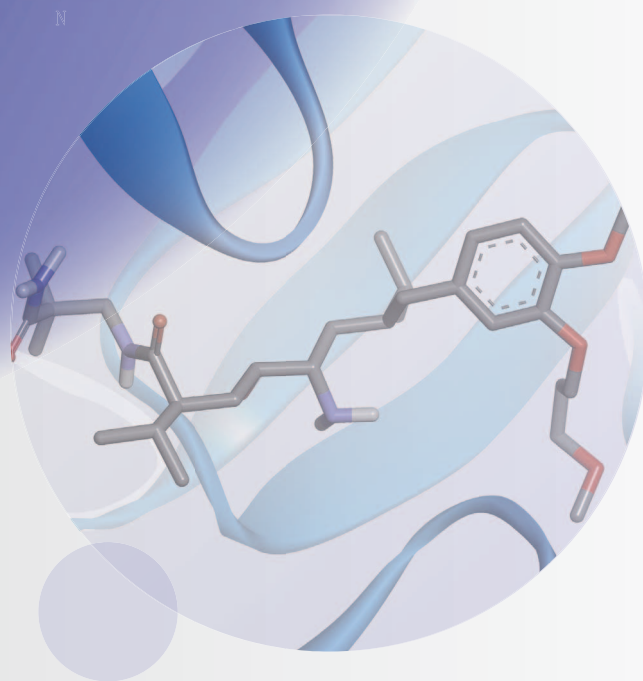


# Discovery Studio

面向生命科学领域的  
综合分子建模和模拟平台



AI  
Driven  
Innovation  
& Quality

## Discovery Studio

面向生命科学领域的综合分子建模和模拟平台

# 目录

---

---

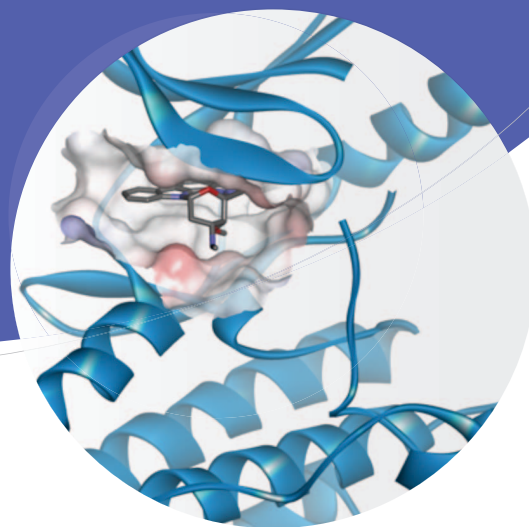
Discovery Studio 软件介绍	1
Discovery Studio 功能模块简介	3
基本界面和显示模块	3
蛋白质建模及模拟模块	4
分子力学、分子动力学、分子力学 / 量子力学模拟模块	5
基于结构的药物发现和设计模块	6
基于片段的药物设计模块	6
基于药效团的药物发现和设计模块	7
基于小分子的药物发现和设计模块	8
生物大分子 X-ray 晶体结构解析模块	9
Discovery Studio 的扩展性	10
Discovery Studio 分子模拟应用案例	13
生命科学领域的解决方案	14
药物化学领域的解决方案	18

---

# Discovery Studio

## 面向生命科学领域的综合分子建模和模拟平台

Discovery Studio (简称 DS)，是基于 Pipeline Pilot 构建的面向生命科学领域的综合分子建模和模拟平台。它服务于生命科学领域的实验生物学家、药物化学家、结构生物学家、计算生物学家和计算化学家，应用于蛋白质结构功能研究，以及药物发现。为科学家提供易用的蛋白质模拟、药物的设计与优化工具。通过高质量的图形界面、经多年验证的科学算法以及集成的环境，DS 将实验数据的保存、管理与专业水准的建模、模拟工具集成在一起，为研究队伍的合作与信息共享提供平台。



### | 应用领域丰富广泛

Discovery Studio 以生物大分子、配体小分子的设计与模拟为核心，研究领域包括且不限于：疾病的发病机理研究、新药发现和设计、生物信息学、结构生物学、酶学、免疫学、病毒学、蛋白质工程、肿瘤研究、食品科学、环境毒理研究等。

### | 使用方式灵活方便

用户可以通过“客户端 (client) —服务器 (Server)”模式使用 Discovery Studio 软件。具有高质量图形界面的可视化客户端，和基于计算的服务器端，为用户提供了数据、工作流程和计算资源的共享和合作。

### | 可扩展性卓越强大

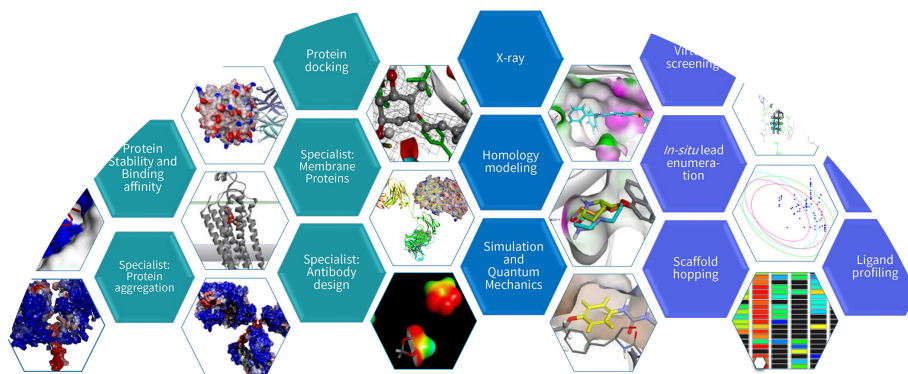
基于领先的科学信息处理平台 Pipeline Pilot (简称 PP) 建立起来的 Discovery Studio 软件让数据的共享和交流变得更为方便和简洁。用户可以依托 Pipeline Pilot 的功能流程定制，制作符合自主需求的科研流程，并加入 DS 中使用。同时，Pipeline Pilot 这个开放平台技术还为使用者整合自己的或第三方的软件工具提供了接口。研究者也可以直接使用 Perl 脚本语言来调用 Discovery Studio 的相关功能。

### | 科学算法准确可信

Discovery Studio 所提供的分子设计与模拟方法中，每个方法所蕴含的科学算法本身都经过详实的验证，并发表在高水平期刊上。

## Discovery Studio

面向生命科学领域的综合分子建模和模拟平台



# Discovery Studio

主要模拟功能包括

### | 生物大分子

蛋白 / 核酸的序列分析、大分子三维结构建模、蛋白结构性质计算、蛋白理性设计、虚拟氨基酸突变、二硫键预测、蛋白聚集效应预测、抗体的设计与优化、抗体人源化、抗体亲和力成熟、蛋白 - 蛋白相互作用预测、X-ray 晶体结构解析、分子动力学模拟和 QM/MM 等。

### | 配体小分子

化合物虚拟筛选、全新药物设计、片段生长、骨架跃迁、受体 - 配体作用机制解释、分子对接、基于片段的药物设计与改造、结合能自由能计算、FEP 自由能微扰、多种药效团模型构建、化合物数据库的构建和筛选、药效团数据库的构建、基于药效团数据库的反向找靶、QSAR/QSPR、构效关系分析 (MMP)、类药性筛选、化合物构象搜索和分析、化合物 ADMET 性质预测等。

AI  
Driven  
Innovation  
& Quality



## Discovery Studio

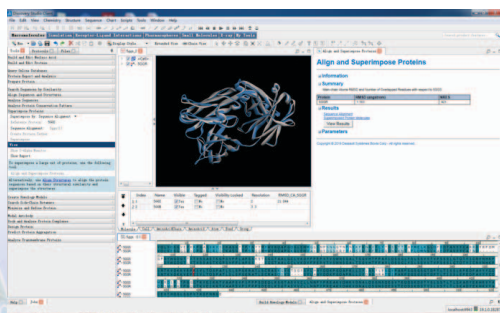
面向生命科学领域的综合分子建模和模拟平台



# Discovery Studio

## 功能模块简介

## 基本界面和显示模块



### | Discovery Studio Standalone

可视化界面，是利用 Discovery Studio 软件进行分子设计和模拟的基础，支持“服务器 (server) - 客户端 (client)”安装在同一台机器上的运行模式。与 Pipeline Pilot Server 结合在一起，提供化学 / 生物学数据显示、动态变化展示、三维作图、动画的输出等显示功能，及许多其它模拟分析功能。

### | Discovery Studio Client

“服务器 (server) - 客户端 (client)”运行模式中，软件的可视化客户端。用于访问和使用服务器端的 Discovery Studio 软件。必须与服务器端配置的 Pipeline Pilot Server 协同工作，为用户提供数据、工作流程和计算资源的共享。

### | Pipeline Pilot Server

“服务器 (server) - 客户端 (client)”运行模式中，软件的计算服务器端。

### | Discovery Studio Visualizer

免费的分子结构显示和分析软件，包含 Discovery Studio Visualizer Client 的部分显示功能，但无法与计算服务器 Pipeline Pilot 连接。此产品通过注册下载获得。

### | ActiveX Control

免费的分子三维结构显示插件。适用于 Windows 环境，该插件可以帮助用户在 IE、PowerPoint 中直接显示和分析分子的三维结构。此产品通过注册下载获得。

# 蛋白质模建及模拟模块

## | DS MODELER

只要给出未知结构序列与另一个已知三维结构且具有一定同源性的蛋白，利用工业标准的快速同源模建方法，便可自动并快速预测蛋白结构。可预测的蛋白结构类型包括酶、抗体、跨膜蛋白。对于靶标发现及蛋白家族功能结构分析，MODELER 是一个理想的解决方案。结合 Sequence Analysis 模块，可以全自动完成抗体 Fab 区、Fv 区、全长抗体以及双特异性抗体结构的模建。同时，MODELER 程序支持并行计算。

## | DS Protein Refine

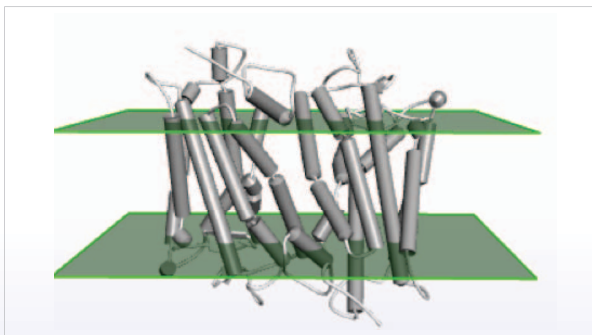
利用 CHARMm 对模建好的蛋白质进行侧链和 loop 区的优化，提高蛋白同源模型的准确性。

## | DS Protein Families

分析蛋白家族蛋白序列的保守模式以及在三维结构上保守残基的位置，可以在分子水平上更好地了解蛋白功能机制。

## | DS Protein Health

蛋白质三维结构合理性评价工具。Protein Health 模块运用 Profiles 3D 方法使用户可以方便快速地找到蛋白质结构中不合理的区域。



## | DS Sequence Analysis

Sequence Analysis 可以让用户使用 BLAST 及 PSI-BLAST 算法来搜索本地数据库或者 NCBI 网上数据库，确定蛋白序列的同源区域。Discovery Studio 同时还支持抗体序列的注释功能、种系基因 V 基因和 J 基因的识别功能，以及基于蛋白序列进行蛋白翻译后修饰位点、抗原线性表位的预测功能和蛋白分子量、等电点、净电荷等生物物理学性质的计算功能。结合 MODELER 和 CHARMm 模块，还可实现抗体人源化，可自动预测并提供异源抗体序列中可人源化突变的残基位点信息及后续人源抗体序列中回复突变信息，尽可能减少抗体的异源性并且同时尽可能确保其结构稳定性以及抗原结合的特异性。

## | DS Protein Docking

快速而准确的蛋白 - 蛋白（核酸 / 多肽）复合物结构预测工具。Protein Docking 采用经典可靠的 ZDOCK 算法预测蛋白 - 蛋白复合物结构，共包括 ZDOCK、ZRANK、RDOCK 三个计算模块。ZDOCK 用于蛋白 - 蛋白刚性对接，ZRANK 用于对 ZDOCK 得到的结构进行基于能量的再次打分，RDOCK 则基于 CHARMm 能量最小化方法优化 ZDOCK 得到的结构。

## | DS Biopolymer

DS Biopolymer 可用于搭建和修饰蛋白质和小肽分子、将蛋白质复合物分成不同的组分、快速产生肽类或蛋白质分子的性质分析报告、计算蛋白的溶剂化能以及蛋白氨基酸残基 pKa 值、计算蛋白偶极矩随 pH 的变化、评估蛋白质的溶解性和黏度。

## | DS Protein Aggregation

预测抗体、蛋白、多肽等生物类型的药物在制造、生产过程的聚集效应，从而确定其稳定性。通过预测影响蛋白自聚集的氨基酸位点，来进行氨基酸突变，从而实现蛋白设计与优化。

# 分子力学、分子动力学、分子力学 / 量子力学模拟模块

## | DS Dynamics

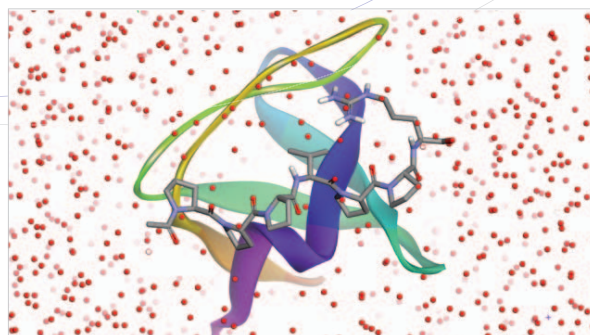
工业标准的分子力学及动力学程序。利用经典而强大的 CHARMM 计算引擎及 CHARMM 系列力场，进行分子动力学模拟，动态研究蛋白、核酸、糖类、脂质、多肽、小分子及相应的复合物等多种分子在不同环境和状态下的热力学及动力学特性。同时，Discovery Studio 还整合了 NAMD 分子动力学模拟引擎，计算速度更快。目前最新版 Discovery Studio 2019 支持 GPU 加速的分子动力学模拟，是 NAMD 使用 8 个 CPU 核进行并行计算速度的 9 倍。为了研究时间尺度更广的生物学过程，DS CHARMM 还支持拉伸分子动力学模拟，可动态研究配体的去结合过程、通道蛋白中离子 / 配体的拉离过程以及蛋白的去折叠过程及构象的变化，预测受体 - 配体结合自由能。

## | DS Analysis

分子动力学结果分析和显示模块。能够分析和显示分子动力学轨迹文件，包括对轨迹进行主成分分析、计算径向分布函数、原子间距离和角度等的变化分析、RMSD 和 RMSF 的分析等。可以将分子轨迹文件以动态、曲线图和表格的方式表现出来。

## | DS CHARMM

基于 DS CHARMM 可以采用虚拟的氨基酸突变方法（如虚拟的丙氨酸扫描、饱和突变等）分析氨基酸突变对蛋白质热稳性、蛋白复合物亲和力的影响，从而快速高效地进行蛋白质设计与优化。基于 CHARMM PBEQ 还可以计算分子表面的静电势分布。在 Discovery Studio 中 CHARMM 也可应用于基于结构的药物设计模块。基于 CHARMM 的对接算法 CDocker 可以精确预测蛋白 - 配体相互作用关系。同时，通过熵的计算可以进一步提高 MM-GBSA 和 MM-PBSA 受体 - 配体结合自由能计算的准确性。支持 FEP 自由能微扰方法精确计算一系列同系配体同受体的相对结合自由能。支持 CHARMM、CHARMM Polar H、charmm36、charmm27、charmm22、charmm19 力场，可模拟多种 Poisson-Boltzmann (PB) 和 Generalized Born (GB) 类型的隐形溶剂模型。支持并行计算。



## | DS CHARMM Lite

用 CHARMM Lite 预测配体与受体亲和力，可得到配体更加精确的打分及排序。所有的工作都可以在后台并行计算。

## | DS CFF (高级 II 类力场)

第二类全原子力场，可以应用于蛋白、核酸、糖类、脂质、多肽等大部分生物分子及小分子结构的模拟研究。

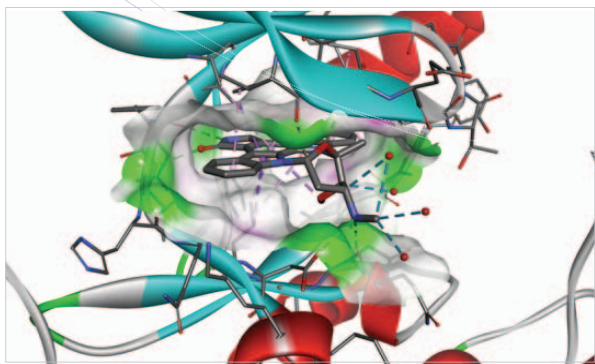
## | DS MMFF (Merck Molecular Force Field)

来自于从头计算和实验数据的第二类力场，专为准确处理构象能和非键相互作用而设计，用于配体或有机小分子的模拟研究。

## | DS QUANTUMm

使用量子力学 / 分子力学 (QM/MM) 方法可以从电子水平精确优化靶标蛋白 - 配体复合物中，结合位点处的构象和能量，计算精度高于单纯的力学优化。其中量化方法采用基于密度泛函理论的 DMOL3 程序，力学优化采用 CHARMM。需要用户定义相应的 QM 区和 MM 区。

## 基于结构的药物发现和设计模块



### | DS Flexible Docking

受体柔性对接工具（“Induced-Fit”对接工具），可以实现配体与受体双柔性对接，模拟配体与受体结合时的“诱导—契合”效应。Flexible Docking 最大的优势在于准确，可以精细地研究配体 - 受体的相互作用信息，适合于作用机理研究。

### | DS LibDock

快速的分子对接工具，适用于对大规模数据库进行快速精确的虚拟筛选。速度快，可以并行运算。

### | DS CDOCKER

是基于 CHARMM 的对接程序，采用 soft-core potentials 以及 optional grid representation 将配体分子与受体活性位点进行对接。采用模拟退火的方法将小分子各个构象在受体活性位点区域进行优化，从而使对接结果更加准确。CDOCKER 在对接过程中还可考虑药效团的限制，提高对接效率和准确度。

### | DS GOLD interface

DS 提供了 GOLD 程序的接口，并提供了设置 interaction filter 的功能。需要用户有 GOLD 软件才能使用。

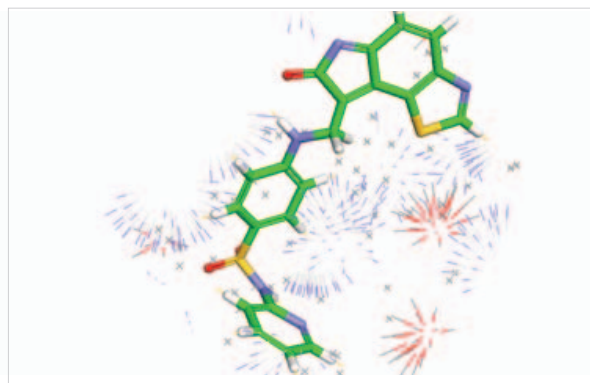
## 基于片段的药物设计模块

### | DS Ludi

全新药物分子设计工具。使用 Ludi 来发现新的具有潜在活性的化合物，可节省研究者大量的时间。Ludi 强大的设计工具允许使用者在实验分析之前模拟筛选，并允许对已有的化合物进行改造。Ludi 易于操作，它包含有 drug-like 片段库，同时也允许用户将自己的分子片段加入到片段库中。

### | DS De Novo Evolution

在 Ludi 算法基础上发展而来的基于受体的小分子药物从头设计方法。在给定母核结构的情况下，DS De Novo Evolution 可以自动地为研究人员设计潜在的与受体有高亲和力的全新小分子化合物，大大缩短了全新小分子药物发现及改造的周期，是设计 Me-Better 类药物分子的有力工具。



### | DS MCSS

基于片段的小分子药物从头设计工具。可以将多个分子片段同时对接进入蛋白质的结合位点，寻找片段的最佳结合位置区域，是强有力的全新药物设计工具，也可以用来表征和分析蛋白质结合位点处的特性。



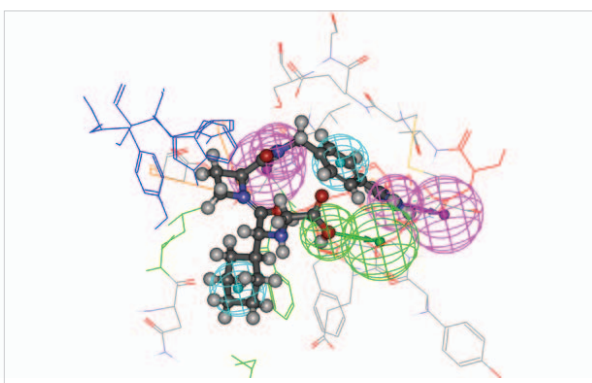
## | DS Grow Scaffold

基于化学反应生长片段工具。能够选取配体分子中某个原子或者某一个基团作为反应位点，选取经典的化学反应进行片段生长，如：酰胺合成反应、醚合成反应等。可选择性的采用 MM-GBMV/SW 模型优化配体分子和 / 或选取的部分蛋白分子侧链。最后根据蛋白口袋匹配特性枚举化合物并进行帕累托优化。通过这种方法设计得到的分子更易于实验合成。

## | DS Replace Fragment

基于电子等排体替换的骨架跃迁工具。可以在有受体蛋白结构的指导下进行或者脱离蛋白结构的需要单独进行。最后根据蛋白口袋匹配特性枚举化合物并进行帕累托优化。

# 基于药效团的药物发现和设计模块



## | DS Catalyst Conformation

多样、完全且快速的构象模型生成工具，可以选择采用 Polling 或 CAESAR 算法，也可采用系统搜索、随机方法或 Boltzmann 随机方法生成构象，生成的构象可用于产生药效团模型或进行数据库搜索。

## | DS Catalyst Hypothesis

对一组化合物进行基于特性结构的比对并自动生成药效团模型的工具。这些特性结构包括亲疏水性基团、氢键给体 / 受体、正 / 负电荷基团和芳香基团等。可选择性的生成排除体积，提高药效团模

型的预测能力；或根据已知具体的活性数据，构建 3D-QSAR 药效团，用于预测新化合物活性。

## | DS Catalyst Shape

用一个特定化合物的形状作模板，确定可能与此模板具有相似形状的化合物。该模块是对基于化学特征药效团搜寻方法的一个有效补充。

## | DS Catalyst Score

提供了分子和药效团之间比较、叠合的工具，快速地评价或排序实验得到的或数据库中的化合物。

## | DS Catalyst Structure Based Pharmacophore (SBP)

基于受体结构产生精准的药效团模型。可以使分子不仅满足和受体结合位点化学特征上的互补，还可以在空间上很好地结合在活性位点空腔处。

## | DS Catalyst DB Build

建立及管理化合物的三维结构数据库，与 DS Catalyst DB Search 可配合使用，用于基于药效团模型的化合物数据库虚拟筛选。

## | DS Catalyst DB Search

基于药效团进行数据库搜索的工具，可以帮助科研人员快速查找到可能的先导化合物。通过设置包含了结合特征、形状以及特性约束等信息的药效团作为数据库的检索条件，对商业或内部数据库中进行查询。可快速筛选成千上万条分子结构记录。DS Catalyst DB Search 在检索时还考虑了分子三维结构的柔性。

## | DS De Novo Ligand Builder

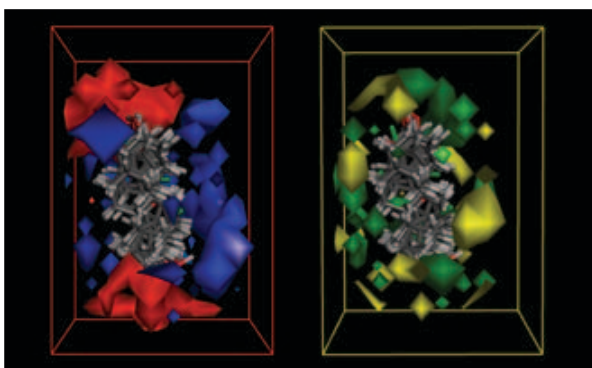
药效团引导的全新分子构建工具。它采用一种新颖而独特的基于片段的药物设计方法（FBDD），片

段放置的位置受药效团支配。这样产生的化合物不仅与蛋白活性位点互补，而且化合物之间的互补性有利于产生新的先导化合物。这个有力的工具能用于迅速产生许多包含与特定靶标结合必需特征的化合物。

## | PharmaDB

药效团数据库，包含基于 scPDB（2012）中 7028 个复合物晶体结构构建的 117423 个药效团模型，并根据不同的靶标类型进行了分类，是目前市场上最大的受体 - 配体复合物药效团数据库。结合 PharmaDB，Discovery Studio 可以快速有效地进行反向找靶、中草药有效成分的确定以及化合物毒副作用评价。

# 基于小分子的药物发现和设计模块



## | DS QSAR, DS QSAR+

定量构效关系研究工具。可以计算接近千种与生物活性或 ADME/T 性质相关的描述符，包括分子拓扑描述符、分子指纹在内的一系列基本性质。同时可以调用半经验量子力学程序 VAMP 计算与电子相关的描述符。还提供了多种统计工具，如 Bayesian 模型、多元线性回归、偏最小二乘法等，用于对各种复杂数据进行建模和数据挖掘。

## | DS MMP

能够利用已有的实验数据快速获取活性悬崖（activity cliff），用来理解结构已知的化合物相关基团之间活性差异，从而设计结构新颖的化合物。

## | DS DMOL3 Descriptor

使用基于泛函密度理论的量子力学程序 DMOL3 计算分子中与电子相关的描述符。

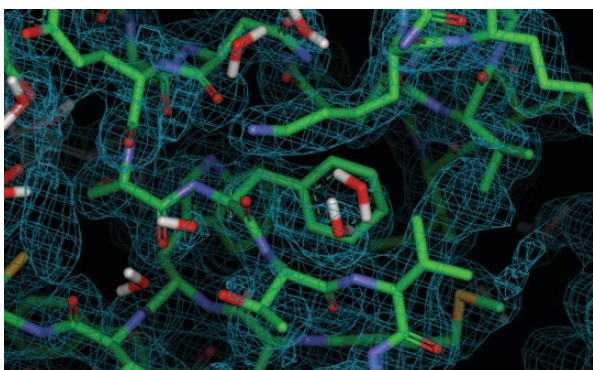
## | DS Library Design

大规模组合化学库的构建和分析工具。Library Design 模块提供了一整套专为化学库设计而定制分子相似性和多样性聚类分析方法。用户可以快速地构建各种组合化学库，应用强大的统计学分析和显示工具对组合化学库的各种性质进行分析和比较，在化学库设计过程中使用帕累托优化方法优化多种性质等，以达到用户按实际需要优化化合物库进行整理和改造的目的。

## | DS ADME & TOPKAT Toxicity Bundle

提供了从化学结构预测吸收、分布、代谢、排泄 (ADME) 和毒理性质的工具。DS ADMET 现在包含了血浆蛋白结合、肝毒性、人细胞色素 P450 2D6 酶结合、被动的肠内吸收性、血脑屏障穿透性、25°C 下的水溶性等 6 个模型。TOPKAT 模块提供了化合物毒理性质预测工具。对各类有机化合物的诸如急性毒性、慢性毒性、诱变性、再生性等毒理性质进行预测，同时还可以考虑化合物对环境毒性的影响。目前，DS 中所有 TOPKAT 毒性模型均以 QMRF 格式提交至欧洲委员会联合研究中心且获批准并发布，遵循 QSAR 模型验证的 OECD 原则，保障了其规范性和可靠性。研究者也可以将自己的毒性数据与已有的毒性数据进行整合，构建新的 TOPKAT 毒性模型。

# 生物大分子 X-ray 晶体结构解析模块



## | DS X-RAY

X-ray 科学功能模块为从事 X 射线晶体学家和基于结构的药物设计科学家提供了从 X 射线晶体学实验数据到结构构建和精修的全套工具和 workflows。基于成熟的 CNX 软件包和 DS 的流程搭建平台，可以很方便地通过工具面板、工作流、命令行等多种方式实现结构的检测和精修，包括产生电子密度图、去除水分子、验证配体的位置和相互作用等操作。此外，还可以进行蛋白 - 配体复合物结构的高通量预测计算。通过配体在蛋白中的放置、结构优化、溶剂化等一系列自动流程化的运作来实现蛋白 - 配体复合物结构的高通量计算。



# Discovery Studio 的扩展性

## 生命科学领域——DS 与 PP 联用自动化蛋白质同源建模流程

蛋白的空间结构对其功能的理解至关重要。现阶段，由于实验技术条件的限制以及自然界中蛋白质的种类数量繁多，人们还无法解析所有蛋白质的晶体结构。采用同源建模方法，人们可以直接从蛋白质的氨基酸序列出发，预测其高级结构。

一般来说，同源建模的主要步骤包括：模板识别、序列比对、模型搭建、模型评估及模型优化。使用 Discovery Studio，用户虽然可以方便的实现上述蛋白质同源建模过程，但仍然需要使用者多次调用 Discovery Studio 的相关运算功能（protocol），这降低了研究者的工作效率。

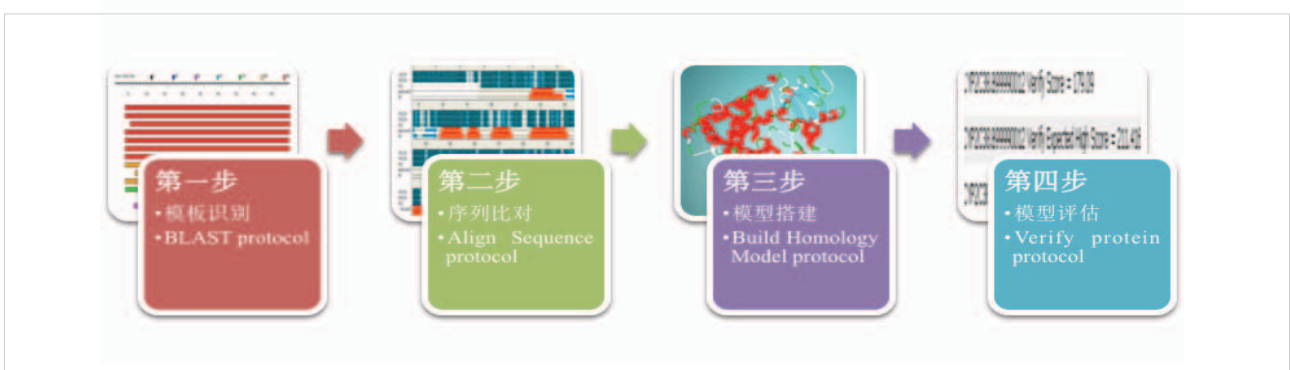


图 1 DS 中同源建模过程

使用 Pipeline Pilot 软件，研究者可以将 Discovery Studio 中同源建模中各个步骤所需的功能（protocol）组合在一起，实现蛋白质同源建模的完全自动化流程。上述自动建模功能又能重新导入回 Discovery Studio，通过 DS 直接使用。

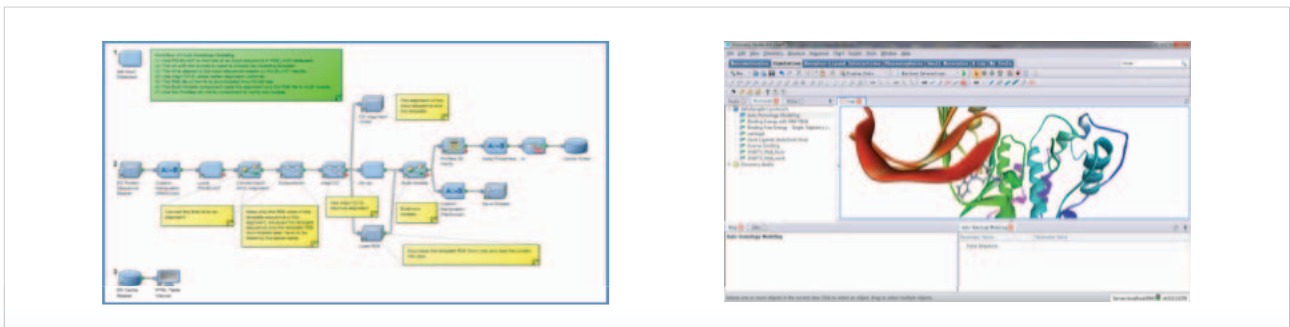


图 2 使用 DS+PP 构建的蛋白同源建模自动化流程重新导入 DS 中使用

## 药物设计领域——DS 与 PP 联用自动化分子对接流程

分子对接模拟方法常常被用来寻找受体与配体间的结合模式。一般，分子对接的主要步骤按模拟流程包括：

- (1) 准备受体结构、准备配体结构；
- (2) 定义配体结合位点；
- (3) 分子对接预测两者间的结合构象；
- (4) 打分评价对接结果；
- (5) 后处理过程；

在 Discovery Studio 中，研究者可以通过多次调用相关功能实现上述分子对接过程，但是要记住每个步骤中所涉及的软件操作，对于初学者来说很难完成。使用 Pipeline Pilot 软件，研究者可以将 Discovery Studio 中分子对接各步骤所需的功能（protocol）组合在一起，完全实现分子对接的自动化过程。上述自动分子对接功能又能重新导入回 Discovery Studio，通过 DS 直接使用。

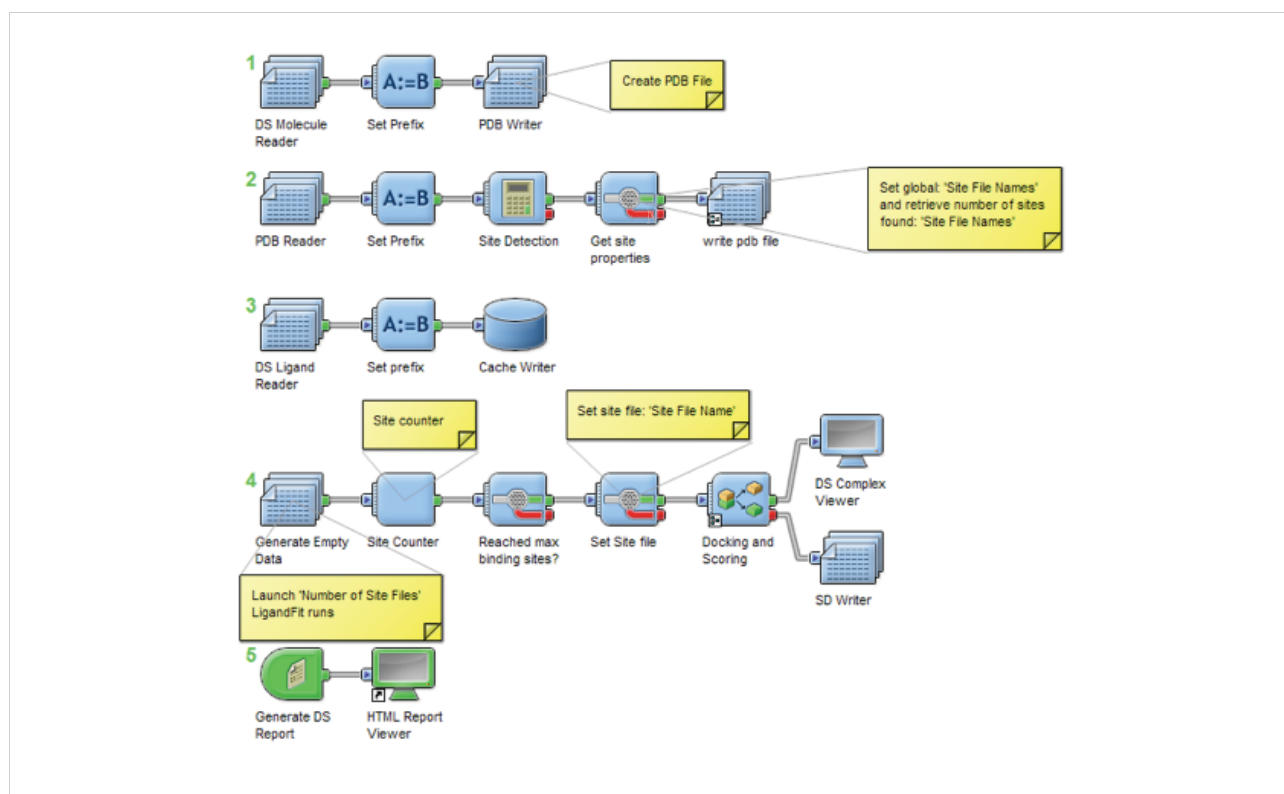


图 1 分子对接的完全自动化流程（使用 DS+PP 构建，分子对接引擎采用 DS 中的对接程序）

## DS 与 PP 联用实现全自动基于结构的全新药物生长策略

Discovery Studio 的计算功能不仅能够通过 Pipeline Pilot 进行自动化定制，研究者还能够使用 PP 将 DS 中的功能按照研究者设计的计算模拟思路进行模拟流程的定制化工作，实现新的分子模拟功能。研究人员使用 Pipeline Pilot 中既有的 Component，构建了一个基于靶标三维结构的全自动全新药物生长 protocol。使用该 Protocol，您只需要输入靶标蛋白的三维结构以及需要保留的化合物骨架等信息，该 Protocol 就可以从您所给定的骨架出发，自动为您生长出与靶标具有最佳结合能力的新化合物来，并采用 GBSA 的方法对新化合物的结合能力进行精确的评价。

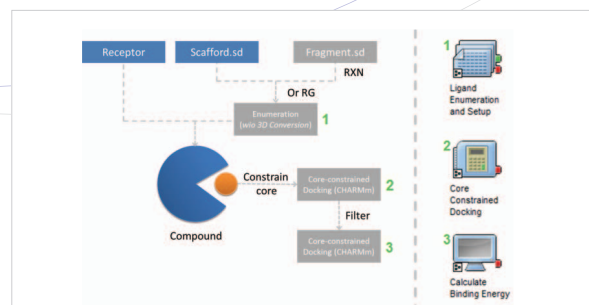


图 1 基于受体结构的药物全新生长策略流程示意图

## 整合打分函数 ID-score

使用 Pipeline Pilot，研究者可以对第三方程序进行整合，并将整合后的程序放置在 Discovery Studio 中使用。这极大的扩展了 DS 的科学计算功能。

Li 等近期开发了一种全新的经验打分函数 ID-score。该打分函数综合考虑了 9 类影响蛋白质与配体相互作用的描述符，包含 2278 个蛋白-配体复合物结构作为训练集，并采用支持向量机回归算法 (SVR) 训练得到了最终的预测模型。基准测试集和大规模的独立测试集的实验结果显示，ID-score 的准确性要优于各种常用的打分函数，而且 ID-score 打分函数对各种生物靶标的预测准确性基本一致，方法评价结果还显示该打分函数对原子类型的高灵敏性，能够很好的预测配体类似物的结合能力。

Li GB, Yang LL, Wang WJ, Li LL, Yang SY. J. Chem. Inf. Model. 2013, 53 (3),592-600.

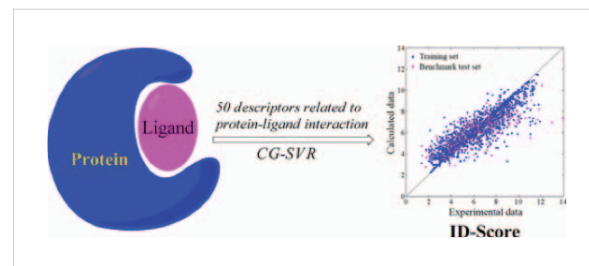


图 1 ID-score 打分函数简介

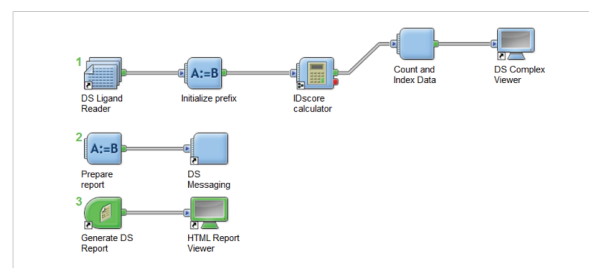


图 2 Pipeline Pilot 整合 ID-score 的计算流程

备注：使用 Pipeline Pilot 整合第三方程序，Pipeline Pilot 本身并不提供第三方程序的使用许可，需要使用者自己提供该程序的使用许可。



图 3 Pipeline Pilot 整合 ID-score 打分函数并在 Discovery Studio 中使用

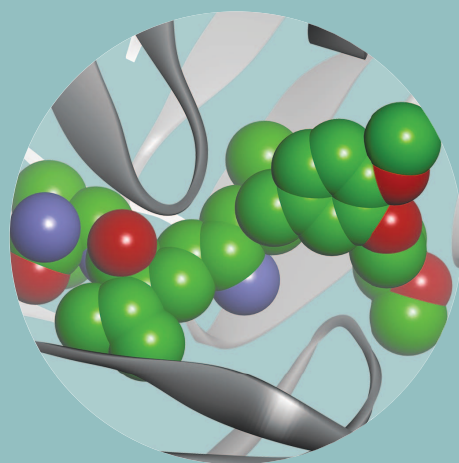
Discovery Studio

面向生命科学领域的综合分子建模和模拟平台



# Discovery Studio

分子模拟应用案例



# Discovery Studio

## 分子模拟应用案例

## 生命科学领域的解决方案

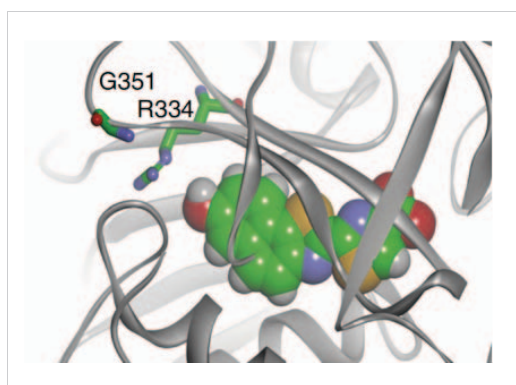


图1 荧光素酶与底物的复合物结构

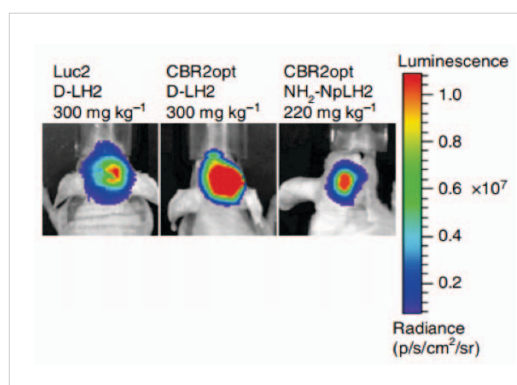


图2 优化后的突变体 (CBR2opt) 与底物在小鼠脑部荧光成像图

## 荧光素酶及底物的设计和改造

【关键词】酶设计与改造、同源建模、分子动力学

【研究背景】萤火虫荧光素酶 (Luc2) 及其底物 D-荧光素 (D-LH2) 介导的生物发光成像方法, 已广泛应用于基因表达检测和临床前小鼠模型的药效检测。Luc2/D-LH2 酶体系可以在 37 摄氏度产生接近 600nm 的峰值, 足以穿透浅层组织如皮肤。但是在深度组织如肺、脑和骨头中, 由于血红蛋白和黑色素等吸收作用, 其灵敏度就无法达到要求。为了改善深度组织成像的分辨率, 其中一种方法就是将生物发光辐射波长提高到近红外区域 (650-900nm)。有研究表明, 通过对 Luc2 突变或增加 D-LH2 的芳香性从而增加酶和底物的  $\pi$  共轭都有利于产生近红外的生物发光。

【设计思路】在此基础之上, 文章作者使用多种实验方法对现有的萤火虫荧光素酶及其底物分别进行了设计和改造。首先将底物 D-LH2 增加了萘环结构, 设计了 NH<sub>2</sub>-NpLH2 和 OH-NpLH2 两种底物, 发现新改造的底物与萤火虫荧光素酶 (CBR) 结合较好, 但荧光效果较弱。因此, 研究结合分子模拟方法使用 Discovery Studio 中 MODELER 构建了基于 CBR 和 OH-NpLH2 复合物三维结构, 并基于该三维结构使用 DS 中分子动力学程序 Standard Dynamics Cascade 程序对新构建的复合物结构进行了优化, 最终确定了稳定的复合物结构, 并分析了两者之间的相互作用。研究表明, 荧光素酶有效的光辐射波长与氢键相互作用相关。根据分析 CBR 和 OH-NpLH2 的氢键相互作用, 选择了 CBR 中关键的两个氨基酸 ARG334 和 GLY351 进行突变, 最终得到 CBR2opt 结构。最后使用实验方法评估了该类突变体的生物发光成像效果。结果表明, 该突变体复合物可以产生近红外辐射, 信号更加稳定, 对于老鼠体内深度组织多光谱成像达到了前所未有的灵敏度和准确度。

Nat Commun. 2018 Jan 9;9(1):132.



# Discovery Studio

## 分子模拟应用案例

## 生命科学领域的解决方案

### 人 / 鼠 PD-1 交叉反应抗体的发现和结合机制的研究

**【关键词】** 抗体设计、抗体同源模建、蛋白 - 蛋白对接

**【研究背景】** 程序性死亡受体 1 (PD-1) 是一种重要的免疫抑制受体分子, 以 PD-1 为靶点的免疫调节在抗肿瘤、抗感染、抗自身免疫性疾病及器官移植存活等方面均有重要的意义。PD-1 抗体是当前备受瞩目, 广为关注的一类肿瘤疗法。考虑到 PD-1 抗体在治疗各种癌症中的重要性和广泛应用, 文章作者希望开发同人 / 鼠 PD-1 有交叉反应的抗体以加速 PD-1 抗体的临床前研究。目前批准用于临床或正在研发中的 PD-1 抗体都只靶向人 hPD-1, 而并不显示与鼠 mPD-1 有交叉反应。因此, 本文开发并首次报道了一种具有人 / 鼠 PD-1 交叉反应的功能性抗体, 并将该抗体的表位与仅同 hPD-1 结合的抗体表位进行了比较。

**【设计思路】** 文章作者首先用 hPD-1 和 mPD-1 免疫大鼠, 通过筛选大量的大鼠杂交瘤克隆发现了既能靶向 hPD-1 又能靶向 mPD-1 的抗体 R9 和只靶向 hPD-1 的抗体 R11。为了快速识别两个抗体的表位, 本研究采用了丙氨酸扫描, 并基于 ELISA 和表面等离子共振技术。同时结合分子模拟的方法, 采用 **Discovery Studio (DS)** 中 **MODELER** 程序构建出了 R11 和 R9 抗体结构模型, 并在此基础上使用 **DS\_ZDOCK/RDOCK** 蛋白 - 蛋白对接 / 优化程序, 将这两个抗体与 hPD-1 进行了对接, 预测了两种抗体不同的结合模式, 研究了他们的结合表位, 从而精细解释了具有人 / 鼠交叉反应抗体的交叉反应性机制。

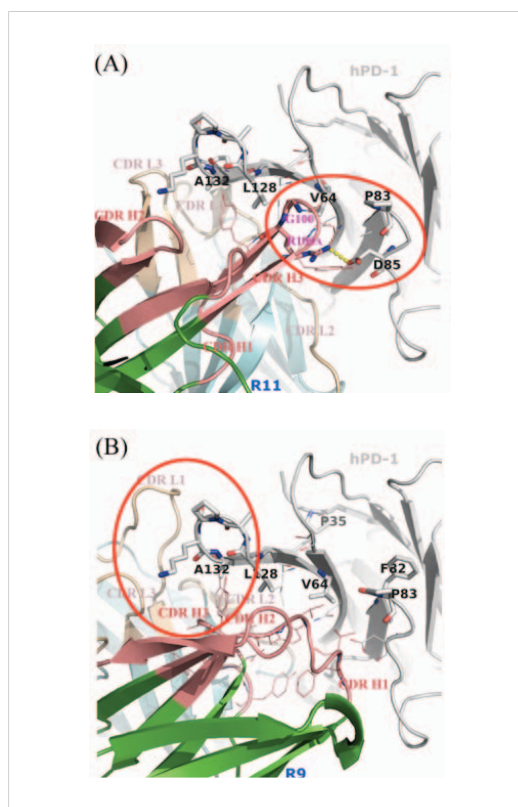


图 1 抗体 R11 (A)、R9 (B) 与 hPD-1 结合模式的预测

# Discovery Studio

## 分子模拟应用案例

## 生命科学领域的解决方案

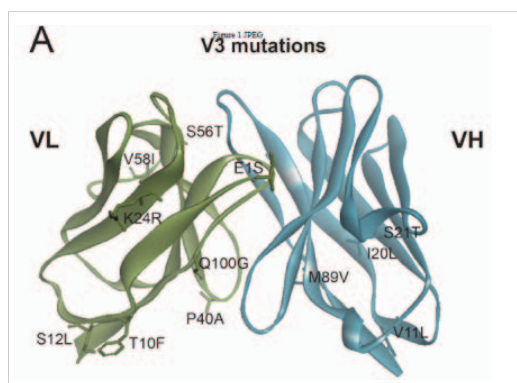


图 1 基于分子模拟方法预测的 hu3F8 突变位点示意图

表 1 优化后 V3-Ile 抗体的结合动力学结果

Table 5   Analysis of binding kinetics measured by surface plasmon resonance.			
Construct	$K_{on}$ ( $S^{-1}M^{-1}$ )	$K_{off}$ ( $S^{-1}$ )	$K_D$ (nM)
hu3F8	$1.15 \times 10^5$	$1.04 \times 10^{-3}$	9.1
V3	$1.09 \times 10^5$	$1.25 \times 10^{-3}$	11.5
V3-Ile	$1.28 \times 10^5$	$0.48 \times 10^{-3}$	3.7
V5	$1.73 \times 10^5$	$3.30 \times 10^{-3}$	19.1

## 抗二唾液酸神经节苷脂 (GD2) 单抗的设计

**【关键词】** 单抗设计、抗体人源化、虚拟氨基酸突变、亲和力 / 稳定性计算

**【研究背景】** 糖基异常化一直以来被认为是癌症的特点之一，GD2（二唾液酸神经节苷脂）是一种特异的糖脂抗原，在小儿和成人肿瘤发生时高度表达，已经成为肿瘤免疫疗法的重要靶标。3F8 是在神经母细胞瘤患者中测试的第一种抗 GD2 单克隆抗体，是对 GD2 具有中等亲和力的鼠源 IgG3。本文基于分子模拟技术对单克隆抗体 3F8 进行了设计和优化。

**【设计思路】** 文章作者首先解析了鼠源 3F8 晶体结构 (PDB: 3VFG)，确定了 3F8 中有 12 个氨基酸位点对于抗体 - 抗原相互作用非常关键 (图 1)，进而采用 **DS\_Calculate Mutation Energy (Binding)** 针对这 12 个位点进行了虚拟的饱和突变。基于虚拟突变结果，本文发现一个单点突变位点 (重链: G54I) 可以提高 3F8-GD2 亲和力并将其工程化至人源化 3F8 (hu3F8) 单抗中，发现在神经母细胞瘤和黑素瘤细胞系中可以有效提高抗体的 ADCC 效应 6-9 倍。本文在此基础上，基于鼠源 3F8 以及人源化 3F8 抗体同源模型 hu3F8 结构 (**DS\_MODELER**、**CHARMm**) 和 **DS\_Calculate Mutation Energy (Stability)** 虚拟氨基酸突变计算模拟方法预测了 12 个突变位点以提高 hu3F8 的抗体热稳性，从而优化了 hu3F8 获得 V3。同前期研究结果 HC:G54I 相结合，本文最终构建得到 V3-Ile 抗体。经试验验证，V3-Ile 抗体表现出更高的热稳性、抗原结合力以及将近提高 6 倍的肿瘤细胞杀伤力 (表 1)，经后续进一步体内验证可进入临床开发用以治疗 GD2 阳性肿瘤。

Front Immunol. 2014 Aug 14;5:372.



# Discovery Studio

## 分子模拟应用案例

## 生命科学领域的解决方案

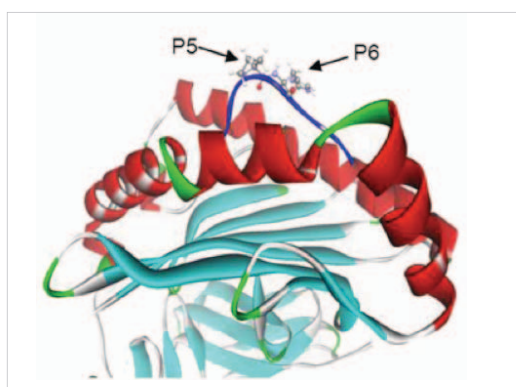


图 1 构建的变构肽配体 - 受体复合物结构

表 1 改造后的变构肽配体和受体复合物的 MM-GBSA 结合能和 RMSDC- $\alpha$  的计算结果

APL candidates	$\Delta\Delta G$ (kcal/mol)	RMSDC- $\alpha$ (Å)	APL candidates	$\Delta\Delta G$ (kcal/mol)	RMSDC- $\alpha$ (Å)
H6F	-7.747	0.38	H6E	-1.243	0.667
H6G	2.980	0.567	H6Q	-0.845	0.512
H6L	11.087	0.523	H6T	23.057	0.934
H6R	2.545	1.432	H6Y	-3.372	0.603

### 基于胰岛素肽的设计和筛选

**【关键词】** 多肽药物设计、loop 区优化、分子动力学、结合自由能计算

**【研究背景】** I 型糖尿病在人类和非肥胖糖尿病小鼠中均是一种自身免疫系统缺陷的疾病，有研究表明其主要由于 CD8<sup>+</sup> T 细胞介导的  $\beta$  细胞损伤，使之不能正常分泌胰岛素。变构肽配体 (Altered peptide ligands, APLs) 可以有效抑制抗原-特异性 T 细胞失效，凋亡或转移。因此开发有效的 APLs 用于 I 型糖尿病的治疗很有必要。本文通过分子模拟技术对变构肽配体进行了设计和优化。

**【设计思路】** 文章作者基于已知的变构肽配体与受体的复合物结构，首先通过 **Discovery Studio** 突变了复合物结构中的关键氨基酸残基，确定了变构肽的关键改造位置，并设计了 8 个变构肽配体。然后结合 **DS\_Loop refinement** 和分子动力学程序 **Standard Dynamics Cascade** 对 8 个变构肽配体的复合物进行了结构模拟和优化，并基于 **MM-PBSA** 方法预测了它们与受体之间的结合能，根据预测结果 (表 1) 选择其中四个肽进行了合成和生物活性测试，其中变构肽 H6F 活性较好。针对 H6F 进行了后续的试验研究，试验结果表明变构肽配体 H6F 有希望成为一个治疗 I 类糖尿病的药物。

Cell Mol Immunol. 2018 Jun 28.[Epub ahead of print]

# Discovery Studio

## 分子模拟应用案例

### 药物化学领域的解决方案

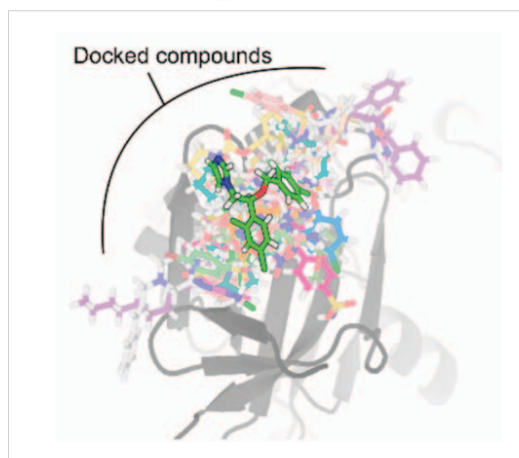


图1 排名前20的化合物对接结果

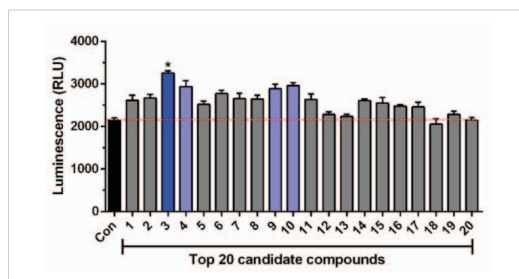


图2 排名前20的化合物 ULK1 激酶活性测试结果

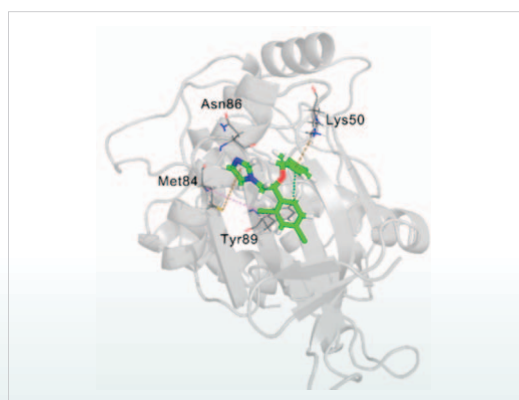


图3 化合物3与ULK1激酶的结合模式

## 新型靶向 ULK1 激酶小分子激动剂的设计

【关键词】 药物设计、虚拟筛选、分子对接

【研究背景】 自噬是一种细胞自我降解的过程，以清除受损或多余的蛋白质和细胞器。自噬作用在机体的生理和病理过程中都可见到，与多种恶性肿瘤、神经退行性疾病等密切相关。哺乳动物基因编码酵母自噬基因 *Atg1* 的同源物 ULK1 蛋白激酶 (UNC-51-like kinase) 是自噬过程中唯一的丝氨酸 - 苏氨酸激酶和启动酶，被认为可能是帕金森病的靶点，因此靶向自噬启动因子 ULK1 小分子激动剂的发现将会为帕金森病的治疗提供可能。

【设计思路】 文章作者首先分析了 ULK1 激动剂结合位点，利用 **Discovery Studio**，通过基于结构的药物设计方法分子对接技术 **DS\_Libdock** 对 ZINC 数据库进行了虚拟筛选，根据对接打分选择排名前 50 的化合物使用 **DS\_CDOCKER** 做进一步筛选，再选择 CDOCKER 打分排名前 20 的化合物进行活性测试发现，其中 15 个化合物具有一定活性，化合物 3 的活性最好 ( $EC_{50} = 261.1 \text{ nM}$ )。通过对化合物 3 和 ULK1 蛋白激酶的非键相互作用分析，作者对化合物 3 做了后续结构优化，最终得到化合物 33i 可作为 ULK1 的强效激动剂，其  $EC_{50}$  值为  $24.14 \text{ nM}$ 。最后作者结合实验方法评估了化合物 33i 的生物活性及其调控自噬的能力，实验结果证明了 33i 靶向 ULK1 的治疗潜力，可作为未来治疗帕金森病的候选药物。

J Med Chem. 2018 Apr 12;61(7):2776-2792.

# Discovery Studio

## 分子模拟应用案例

## 药物化学领域的解决方案

### STAT3 抑制剂的设计和研发

【关键词】 药物设计、片段生长、分子对接

【研究背景】 信号传导子及转录激活子 3 (STAT3) 是一种信号传导蛋白和转录激活物，将细胞外的信号传递到细胞核内而发挥作用，参与多个生理过程，包括细胞的生长、增殖、分化等，因而成为一个潜在的治疗靶标。近来有研究表明紫草素对 STAT3 具有一定的抑制活性，特别在乳腺癌模型中表现出较明显的药理作用。目前已知的 STAT3 抑制剂都具有一些相似的萘醌结构骨架，那么如何基于这些骨架用分子模拟的方法设计出新颖有效的 STAT3 抑制剂就是本文待解决的问题。

【设计思路】 文章作者利用 **Discovery Studio** 分子模拟平台，首先使用 **CDOCKER** 将紫草素对接到 STAT3 的结合位点内，分析了它们之间的相互作用，基于该相互作用使用 **Grow Scaffold**，以紫草素支链上的羟基基团为起点，设计了多个结构多样的紫草素衍生物。然后使用 **CDOCKER** 将设计的紫草素衍生物再次对接到 STAT3 的结合位点内，选择打分较高的分子进行合成和生物活性的评估，最后发现新设计的 PMMB-187 分子对 MDA-MB-231 细胞的抑制活性较好，具有潜在的抗乳腺癌活性。

Biochem Pharmacol. 2017 Dec 15;146:74-86.

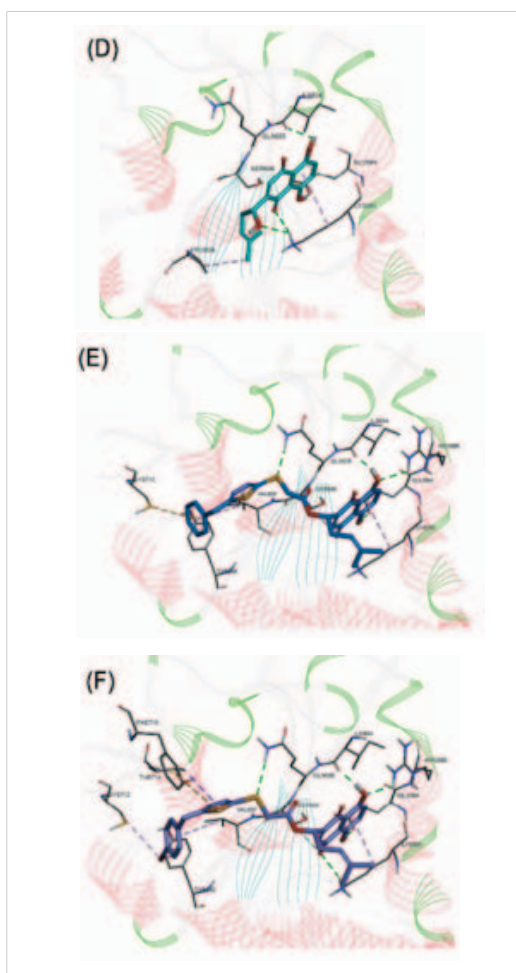


图 1 紫草素 (D) 及其衍生物 PMMB-175 (E) 和 PMMB-187 (F) 与 STAT3 的相互作用图

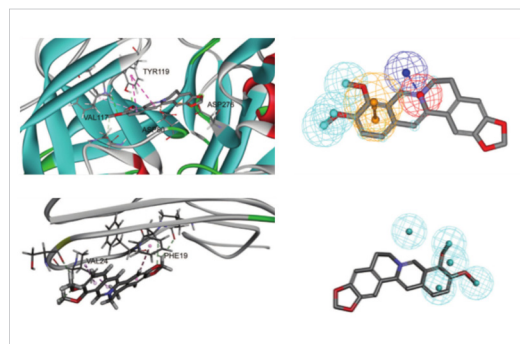
# Discovery Studio

## 分子模拟应用案例

## 药物化学领域的解决方案

表 1 筛选 PharmaDB 的结果

TABLE 1   Highly potential targets of berberine.				
Target name	Gene name	PDB ID	Fit value	Reference
TGF- $\beta$ 1 receptor	TGFB1	3TZM	0.936114	Ogunjimi et al., 2012
A $\beta$ <sub>1-42</sub>	APP	2NAO	0.906099	Wali et al., 2016
PksA	afIC	3HRR	0.818081	Crawford et al., 2009
BACE-1	BACE1	5ENM	0.788156	Wu et al., 2016
MEK-1	MAP2K1	3EQB	0.744816	Warmus et al., 2008
BRAF	BRAF	4E26	0.680164	Clin et al., 2012
CYP11A1	CYP11A1	3MZS	0.677786	Mast et al., 2011
CK2	CSNK2B	2OXD	0.636842	Battistutta et al., 2007
PDE10A	PDE10A	2O8H	0.596286	Chappie et al., 2007
MD-2	LY96	2E59	0.553963	Ohto et al., 2007
AMPK	PRKAA2	4ZHX	0.535163	Langendorf et al., 2016
LXR	NR1H2	3KFC	0.531021	Bernotas et al., 2010
HSP90	HSP90AA1	2XDU	0.503232	Murray et al., 2010

图 1 黄连素与 BACE1、A $\beta$ <sub>1-42</sub> 的作用模式和药效团模型

## 基于多靶标结合模式的黄连素多元药理学研究

**【关键词】**多元药理学、老药新用、药效团模型、药效团数据库

**【研究背景】**多元药理研究越来越多的应用到药物研发中，一个药物分子可能存在多个蛋白靶标，为新药研发提供了新的方向。黄连素作为一种治疗腹泻的常用药，被报道出具有治疗感染，癌症，糖尿病等的功效。

**【设计思路】**本文通过 Discovery Studio 中基于受体 - 配体构建药效团模型方法 (Receptor-Ligand Pharmacophore Generation)，构建了黄连素与多靶标的药效团模型，并通过此模型和药效团数据库 PharmaDB，筛选出可以与黄连素发生相互作用的阿尔兹海默症相关的蛋白靶标，然后进行实验验证，得出黄连素可以作为治疗阿尔兹海默症的一种可能性药物。这种多靶标结合模式研究药物的方法可以很好的帮助我们理解疾病中药物作用。

Front Pharmacol. 2018 Jul 24;9:801.

# Discovery Studio

## 分子模拟应用案例

### 药物化学领域的解决方案

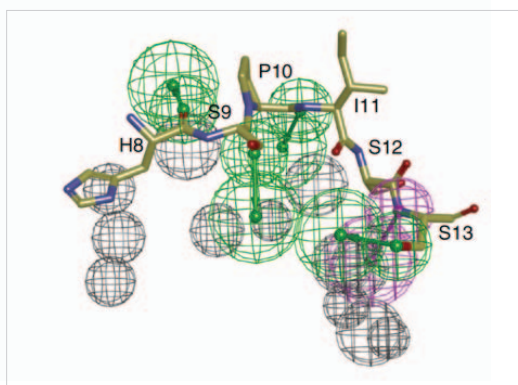


图1 筛选使用的药效团模型

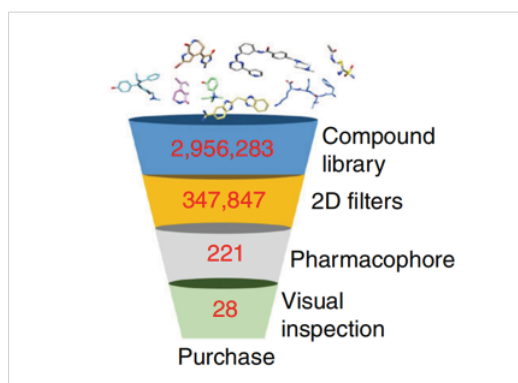


图2 虚拟筛选流程

## 基于阻碍 FL-FLT3 相互作用的抑制剂筛选

【关键词】 药物设计、虚拟筛选、药效团筛选

【研究背景】 外周神经性疼痛 (PNP) 是一种比较难治疗的慢性疾病。背脊髓相关背根神经节中存在的体感神经元致敏性是一个关键的生理病理过程。研究发现，造血细胞在神经损伤位置可产生细胞因子 FL，即 FMS 样酪氨酸激酶 3 受体 (FLT3) 的配体。通过坐骨神经内注射 FL 激活 FLT3，产生疼痛的超敏感性，进而激活了外周神经性疼痛相关的基因表达，产生短期或长期的敏感神经元的敏感性。本研究通过一系列实验，基于不同的老鼠模型发现背神经节中的 FLT3 对于外周神经性疼痛的产生和持续起关键作用。同时发现阻碍 FL-FLT3 之间相互作用的抑制剂可以减缓外周神经性疼痛，该方法有望成为治疗外周神经性疼痛的一种新疗法。

【设计思路】 为了寻求该类抑制剂，文章作者对商业可获得的类药化合物库共 290 万个化合物，使用 **Pipeline Pilot** 按照氢键受体大于等于 1，氢键供体大于等于 4，可旋转键数目小于等于 10，芳香环数目大于等于 1，极性表面积小于等于 90Å，预测的水溶性大于等于 50μM 等规则进行了第一轮筛选，并通过结构预处理，得到 343,847 个具有三维结构的异构体构象。然后使用 **Discovery Studio** 中的 **Receptor-Ligand Pharmacophore Generation** 工具产生基于 FL-FLT3 晶体复合物结构的药效团模型，并利用 **DS\_Screen Library** 模块对之前 343,847 个化合物进行了第二轮筛选。以 Fitvalue 值在 3 以上为标准，得到 285 个结构不同的化合物，最后进行聚类分析和相互作用观察，得到 28 个化合物，经购买和活性测试发现化合物 BDT001 是一种全新的 FLT3 抑制剂，其 IC<sub>50</sub> 为 18 μM。

Nat Commun. 2018 Mar 12;9(1):1042.



## 关于我们 about us

创腾科技专注于生命科学和材料科学领域信息化的开拓与创新。通过 AI 及移动互联技术，我们为广大用户提供：计算模拟与数据建模、科研创新及质量合规等全方位的信息解决方案，全面提升企业的研发效能和数字化转型价值。在中国，已有千余家生命科学和材料科学领域的机构选择了创腾科技的产品和服务，包括国内知名的制药企业、新药研发合同服务企业、石化企业以及高校、科研院所。

## 合作伙伴

Discovery Studio, Pipeline Pilot 是创腾科技合作伙伴达索系统的产品，达索系统授权创腾科技有限公司在中国提供上述产品的销售，培训和技术支持。

AI Driven  
Innovation  
& Quality

创腾科技 | 北京 · 上海 · 苏州 · 广州 · 香港

北京市中关村科学院南路2号融科资讯中心  
C座南楼1512室 (100190)  
电话: 010 82676188  
传真: 010 82677178

上海市浦东新区张江达尔文路88号  
半岛科技园11栋4楼 (201203)  
电话: 021 51821768  
传真: 021 51821858

江苏省苏州市工业园区东长路88号  
2.5产业园A2栋301室 (215028)  
电话: 0512 67509707  
传真: 0512 67509707

广州市天河区黄埔大道西33号  
三新大厦16-E房 (510620)  
电话: 020 88527961



[www.neotrident.com](http://www.neotrident.com) [market@neotrident.com](mailto:market@neotrident.com)