



(12) 发明专利

(10) 授权公告号 CN 115551497 B

(45) 授权公告日 2024. 11. 08

(21) 申请号 202180033571.3

(22) 申请日 2021.03.05

(65) 同一申请的已公布的文献号
申请公布号 CN 115551497 A

(43) 申请公布日 2022.12.30

(30) 优先权数据
62/987914 2020.03.11 US

(85) PCT国际申请进入国家阶段日
2022.11.07

(86) PCT国际申请的申请数据
PCT/US2021/020992 2021.03.05

(87) PCT国际申请的公布数据
W02021/183370 EN 2021.09.16

(73) 专利权人 普渡研究基金会
地址 美国印第安纳州

(72) 发明人 G·乔普拉 P·维耶瓦尔登纳
K·P·杰塔瓦 J·法恩

(74) 专利代理机构 中国专利代理(香港)有限公司 72001
专利代理人 王颖煜 彭昶

(51) Int.Cl.
A61K 31/192 (2006.01)
A61K 31/216 (2006.01)
A61K 31/343 (2006.01)

(56) 对比文件
CN 105705489 A, 2016.06.22
审查员 孟晋东

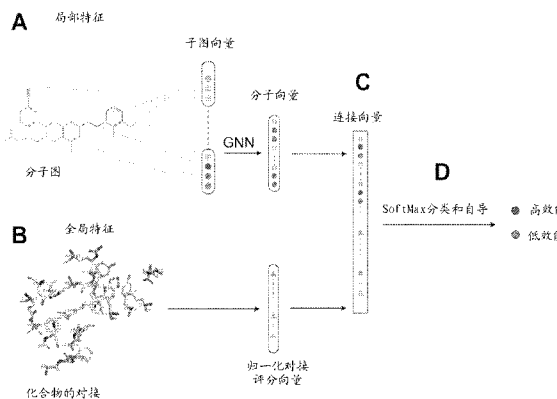
权利要求书1页 说明书40页 附图11页

(54) 发明名称

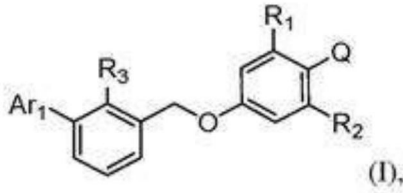
具有免疫调节活性的化合物及其治疗用途

(57) 摘要

本发明总体上涉及用于治疗用途的新化合物。特别地,所公开的具有免疫调节活性的系列化合物可用于治疗免疫系统的功能障碍和各种癌症。用于通过单独地或与其它治疗剂一起施用治疗有效量的这样的化合物来治疗具有免疫疾病和/或癌症的患者的药物组合物物质和方法在本公开内容的范围内。



1. 具有式 (I) 的化合物或其药学上可接受的盐:



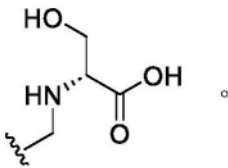
其中

Ar₁是2,3-二氢苯并[b][1,4]-二氧杂环己烯基;

R₁和R₂独立地是甲氧基;

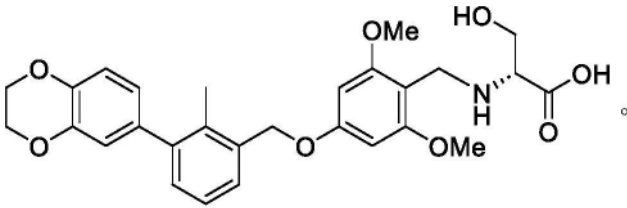
R₃是C₁₋₆烷基;且

Q是:



2. 根据权利要求1所述的化合物或其药学上可接受的盐,其中R₃是CH₃。

3. 根据权利要求1所述的化合物或其药学上可接受的盐,其中所述化合物是



4. 药物组合物,其包含权利要求1~3中任一项所述的一种或多种化合物或其药学上可接受的盐以及一种或多种赋形剂或载体。

5. 根据权利要求1~3中任一项所述的化合物或其药学上可接受的盐,其中所述化合物或其药学上可接受的盐是免疫调节剂。

6. 根据权利要求1~3中任一项所述的化合物或其药学上可接受的盐,其中所述化合物或其药学上可接受的盐是PD-1和PDL-1信号传递途径的抑制剂。

7. 根据权利要求1~3中任一项所述的化合物或其药学上可接受的盐,其中所述化合物或其药学上可接受的盐用于治疗癌症。

8. 药物组合物在制备用于治疗癌症的药物中的用途,所述药物组合物包含权利要求1~3中任一项所述的一种或多种化合物或其药学上可接受的盐和一种或多种载体或赋形剂。

9. 药物组合物在制备用于治疗癌症的药物中的用途,所述药物组合物包含权利要求1~3中任一项所述的化合物或其药学上可接受的盐结合一种或多种具有相同或不同作用模式的其它化合物和一种或多种载体或赋形剂。

10. 根据权利要求9所述的用途,其中所述癌症是去势抗性的前列腺癌。

11. 用作癌症药物的药物组合物,其包含一种或多种根据权利要求1~3中任一项所述的化合物或其药学上可接受的盐以及一种或多种赋形剂或载体。

具有免疫调节活性的化合物及其治疗用途

[0001] 政府支持条款

[0002] 本发明在由国立卫生研究院(National Institutes of Health)给予的CA023168下的政府支持下完成。美国政府具有本发明的某些权利。

[0003] 对相关申请的交叉参考

[0004] 本专利申请与2020年3月11日提交的系列号为62/987,914的美国临时申请有关并要求其优先权权益,其内容特此通过引用整体并入本文。

技术领域

[0005] 本发明总体上涉及具有免疫调节活性的化合物及其治疗用途。本文还描述了这样的化合物的药物组合物和通过单独地、与其它治疗剂一起或在药物组合物中施用治疗有效量的这样的化合物来治疗癌症患者的方法。

[0006] 背景

[0007] 本节介绍了可能有助于促进本公开内容的更好理解的方面。因此,这些陈述应当从这个角度来阅读,而不应被理解为承认什么是或不是现有技术。

[0008] 程序性细胞死亡蛋白1(PD-1)是一种与新癌症治疗剂的创建有关的免疫检查点受体¹。T-细胞受体和主要组织相容性复合物(MHC)之间的长期相互作用导致活化的T-细胞表面上的PD-1的上调²。活化的T细胞产生细胞因子,诸如干扰素- γ ,其又造成肿瘤细胞在其细胞表面上表达程序性死亡配体1(PD-L1)²。肿瘤通过利用PD-1和配体PD-L1之间的相互作用逃避免疫系统的作用,导致更低的效应物T细胞功能和存活,从而导致肿瘤微环境中的抑制性免疫应答²。PD-1/PD-L1相互作用的抑制可以增强抗肿瘤免疫,并且已经进行大量工作来开发单克隆抗体作为PD-1/PD-L1相互作用的抑制剂^{3,4}。例如,派姆单抗和西米普利单抗和纳武单抗是三种FDA批准的抗-PD-1抗体⁴。小分子抑制剂的发现将比抗体具有优势,诸如起效快、体内施用简单、能够穿透细胞膜并与细胞表面受体的细胞质结构域相互作用⁵。几年来,在设计PD-1/PD-L1抑制剂方面已经取得了重大进展^{6,7}。具体而言,Bristol-Myers Squibb(BMS)发现了一组基于肽模拟物分子和非肽小分子的有效PD-1/PD-L1小分子抑制剂^{6,7}。特别地,BMS揭示了含有2-甲基-3-联苯基-甲醇骨架的化学库。后来,Holak等人研究了BMS分子与PD-L1的相互作用,表明BMS分子诱导PD-L1二聚化,并且还报告了具有二聚体PD-L1的化合物的晶体结构^{8,9}。基于这些发现,我们设想开发一个机器学习(ML)框架用于选择和试验新的PD-1/PD-L1抑制剂。对于更有效的癌症治疗,尤其是具有免疫调节活性的PD-1/PD-L1抑制剂,存在未得到满足的医疗需求。

[0009] 附图简述

[0010] 本发明的上述和其它目的、特征和优点将在结合以下描述和附图时变得更加明显,其中:

[0011] 图1. EGNN模型利用了局部(A)和全局(B)特征的组合。使用GNN将权重分配给分子各子图,从分子的分子图计算局部特征。全局特征是用于代表化合物和蛋白之间相互作用的对接评分的集合。将这两个特征组合起来创建一个拼接向量(concatenated vector)

(C),其穿过SoftMax层并自举以将分子分类为对PD-1/PD-L1相互作用具有“低的”或“高的”效能。

[0012] 图2A,上图:在BMS和Incyte专利中的训练数据的分类;左下:BMS专利的主要PD-L1抑制剂骨架。R基团可以是CN、Cl、Br或CH₃;和右下:Incyte专利的主要PD-L1抑制剂骨架。这里A和B表示子骨架。

[0013] 图2B和2C分别显示了BMS和Incyte化合物的成对Tanimoto相似性评分的热图。

[0014] 图3A-3C. 浅粉色链代表PD-1蛋白,且浅青色链代表在PD-1/PD-L1复合物晶体结构(PDB ID: 4ZQK)中的PD-L1蛋白。小麦色链代表PD-L1链A,且蓝白色代表在PD-L1同源二聚体晶体结构(PDB ID: 5N2F)中的PD-L1链B。(图3A)重叠的且对齐的PD-1/PD-L1 (4ZQK)和PD-L1二聚体(5N2F)晶体结构。(图3B)重叠的且对齐的两个晶体结构,具有PD-L1二聚体(5N2F)的确定结合位点(灰色网格)。(图3C) PD-L1二聚体(5N2F)晶体结构,在其结合位点(灰色网格)有小分子(配体ID: 8HW)。

[0015] 图4A显示了用于模型的训练-验证和试验方案;

[0016] 图4B描绘了具有不同训练-验证和试验集的EGNN和GNN的Cohen氏κ评分;

[0017] 图4C显示了具有不同训练-验证和试验集的EGNN和GNN模型的F1评分;且图4D显示了使用Incyte数据训练的EGNN、GNN、RF和SVM模型的BMS和Incyte化合物精确召回曲线之间的成对Tanimoto相似性评分的热图。

[0018] 图5A表明,EGNN通过BMS化合物4a或BMS-1和BMS-1002的骨架跳跃预测了一种新的PD-1/PD-L1抑制剂化合物4b。4b的蓝色部分是从BMS-1002添加,且粉红色部分是从4a(BMS-1)添加。

[0019] 图5B显示了化合物4b在PD-L1同源二聚体晶体结构(PDB ID: 5N2F)中的顶部对接位姿的位置。插图显示了化合物4b的疏水隧道。

[0020] 图5C显示了化合物4b在PD-L1同源二聚体中的顶部对接位姿相互作用的化学相互作用。蓝色和粉红色部分显示为4b的键(sticks)。化合物与残基_AThr20和_AAla121之间的黄色线代表氢键。酪氨酸_ATyr56的芳族环的取向提示与化合物4b中的2,3-二氢-1,4-苯并二氧杂环己烯蓝色芳族环的似乎合理的π-π相互作用。

[0021] 图5D显示了4a(BMS-1对照化合物,红色)和新化合物4b(蓝色)的IC₅₀值对比。为每种试验浓度显示了测定的阳性(PC-DMSO,紫色)和阴性对照(NC-DMSO,绿色)的DMSO对照。

[0022] 详细描述

[0023] 为了促进对本公开内容的原理的理解的目的,现在将参考附图中所示的实施方案,并将使用特定语言对其进行描述。尽管如此,应当理解,无意由此限制本公开内容的范围。

[0024] 本文中使用的以下术语和短语应当具有下述的含义。除非另外定义,否则在本文中使用的所有技术和科学术语具有与本领域普通技术人员通常理解相同的含义。

[0025] 在本公开内容中,术语“约”可以允许值或范围的一定程度的变化性,例如,在规定值或规定的范围限度的10%内、5%内或1%内。在本公开内容中,术语“实质上”可以允许值或范围的一定程度的变化性,例如,在规定值或规定的范围限度的90%以内、95%以内或99%以内。

[0026] 在本文件中,除非上下文另外清楚地指明,否则术语“一个”、“一种”或“所述”用于

包括一个或超过一个。除非另外指出,否则术语“或”用于指非排他性的“或”。另外,应当理解,本文采用的措辞或术语,并且在没有另外定义的情况下,仅用于描述的目的而不是限制的目的。章节标题的任何使用旨在辅助文件的阅读,且不应被解释为限制性的。此外,与章节标题相关的信息可能出现在该特定章节之内或之外。此外,在本文件中提及的所有出版物、专利和专利文献通过引用整体并入本文,如同个别地通过引用并入。如果本文件与如此通过引用并入的那些文件之间的用法不一致,则所并入的参考文献中的用法应视为对本文件的用法的补充;对于不可调和的不一致,以本文件中的用法为准。

[0027] “卤素”指F、Cl、Br或I。“卤素取代”或“卤代”取代指用F、Cl、Br或I替代一个或多个氢原子。

[0028] 本文中使用的术语“烷基”是指碳原子的饱和单价链,其可以任选地是支链的。应当理解,在包括烷基的实施方案中,那些实施方案的示例性变体包括低级烷基,诸如 C_1 - C_6 烷基、甲基、乙基、丙基、3-甲基戊基等。

[0029] 本文中使用的术语“烯基”是指包括至少一个双键的碳原子的不饱和单价链,其可以任选地是支链的。应当理解,在包括烯基的实施方案中,那些实施方案的示例性变体包括低级烯基,诸如 C_2 - C_6 、 C_2 - C_4 烯基等。

[0030] 本文中使用的术语“炔基”是指包括至少一个三键的碳原子的不饱和单价链,其可以任选地是支链的。应当理解,在包括炔基的实施方案中,那些实施方案的示例性变体包括低级炔基,诸如 C_2 - C_6 、 C_2 - C_4 炔基等。

[0031] 本文中使用的术语“环烷基”是指碳原子的单价链,其一部分形成环。应当理解,在包括环烷基的实施方案中,那些实施方案的示例性变体包括低级环烷基,诸如 C_3 - C_8 环烷基、环丙基、环己基、3-乙基环戊基等。

[0032] 本文中使用的术语“环烯基”是指碳原子的不饱和单价链,其一部分形成环。应当理解,在包括环烯基的实施方案中,那些实施方案的示例性变体包括低级环烯基,诸如 C_3 - C_8 、 C_3 - C_6 环烯基。

[0033] 本文中使用的术语“亚烷基”是指碳原子的饱和二价链,其可以任选地是支链的。应当理解,在包括亚烷基的实施方案中,那些实施方案的示例性变体包括低级亚烷基,诸如 C_2 - C_4 亚烷基、亚甲基、亚乙基、亚丙基、3-甲基亚戊基等。

[0034] 应当理解,烷基、环烷基、烯基、环烯基、亚烷基和杂环中的每一个可以任选地被独立地选择的基团取代,所述独立地选择的基团是诸如烷基、卤代烷基、羟基烷基、氨基烷基、羧酸及其衍生物,包括酯、酰胺和亚硝酸酯、羟基、烷氧基、酰氧基、氨基、烷基和二烷基氨基、酰基氨基、硫代等和它们的组合。

[0035] 本文中使用的术语“杂环的”或“杂环”是指碳和杂原子的单价链,其中所述杂原子选自氮、氧和硫,且其一部分、至少一个杂原子形成环。术语“杂环”可以包括“芳族杂环”和“非芳族杂环”二者。杂环包括4-7元单环和8-12元二环,诸如咪唑基、噁唑基、噁唑基、噁唑基、噁唑基、二硫杂环己烷基、二氧杂环己烷基、异噁唑基、异噁唑基、三唑基、呋喃基、四氢呋喃基、二氢呋喃基、吡喃基、四唑基、吡唑基、吡嗪基、哒嗪基、咪唑基、吡啶基、吡咯基、二氢吡咯基、吡咯烷基、哌啶基、哌嗪基、嘧啶基、吗啉基、四氢噻吩基、噻吩基、氮杂环丁基、氧杂环丁基、硫杂丙环基、环氧乙烷基、氮杂环丙基等。“杂环”可以任选地在能够携带氢原子的任何一个或多个位置处被取代。

[0036] 本文中使用的术语“芳基”包括单环和多环芳族碳环基团,其中每个可以任选地被取代。术语“任选地被取代的芳基”是指碳原子的芳族单环或多环,诸如苯基、萘基等,其可以任选地被一个或多个独立地选择的取代基取代,诸如卤素、羟基、氨基、烷基或烷氧基、烷基磺酰基、氰基、硝基等。

[0037] 术语“杂芳基”或“芳族杂环”包括被取代的或未被取代的芳族单环结构,优选5-至7-元环,更优选5-至6-元环,其环结构包括至少一个杂原子,优选一至四个杂原子,更优选一或两个杂原子。术语“杂芳基”也可以包括具有一个或两个环的环系统,其中至少一个环是杂芳族的,例如,其它环状环可以是环烷基、环烯基、环炔基、芳族碳环、杂芳基和/或杂环。杂芳基包括、但不限于吡啶基、N-氧代吡啶基、嘧啶基、吡嗪基、哒嗪基、三嗪基、呋喃基、噻啉基、异噻啉基、噻吩基、咪唑基、噻唑基、吡啶基、吡咯基、噁唑基、苯并呋喃基、苯并噻吩基、苯并噻唑基、异噁唑基、吡唑基、三唑基、四唑基、吡唑基、1,2,4-噻二唑基、异噻唑基、苯并噻吩基、嘌呤基、呋唑基、苯并咪唑基、吡啶基等。在某些实施方案中,所述杂芳基基团具有从1至约20个碳原子,且在其它实施方案中,具有从约3至约20个碳原子。在某些实施方案中,所述杂芳基基团含有3至约14个、3至约7个、或5至6个成环原子。在某些实施方案中,所述杂芳基基团具有1至约4个、1至约3个或1至2个杂原子。

[0038] 在某些实施方案中,“杂环烷基”是指非芳族杂环,其中所述成环原子中的一个或多个是杂原子诸如O、N或S原子。杂环烷基基团可以包括单环或多环(例如,具有2、3或4个稠合环)环系统以及螺环。实例杂环烷基基团包括吗啉代、硫代吗啉代、哌嗪基、四氢呋喃基、四氢噻吩基、2,3-二氢苯并呋喃基、1,3-苯并二氧杂环戊烯、苯并-1,4-二氧杂环己烷、哌啶基、吡咯烷基、异噁唑烷基、异噻唑烷基、吡唑烷基、噁唑烷基、噻唑烷基、咪唑烷基等。在杂环烷基的定义中还包括具有与非芳族杂环稠合(即具有共同键)的一个或多个芳族环的部分,例如杂环的邻苯二甲酰亚氨基、萘二甲酰亚氨基和苯并衍生物。具有一个或多个稠合芳族环的杂环烷基基团可以通过芳族或非芳族部分连接。

[0039] 本文中使用的术语“任选地被取代的”或“任选取代基”意指,所讨论的基团是未被取代的或被指定的取代基中的一个或多个取代。当所讨论的基团被超过一个取代基取代时,所述取代基可以相同或不同。此外,当使用术语“独立地”、“独立地是”和“独立地选自”时,意指所讨论的基团可以相同或不同。某些本文定义的术语可以在结构中出现超过一次,且在这样的出现后,每个术语将彼此独立地进行定义。

[0040] 术语“患者”包括人和非人动物诸如伴侣动物(狗和猫等)和家畜动物。家畜动物是为食物生产而饲养的动物。要治疗的患者优选是哺乳动物,特别是人类。

[0041] 术语“药学上可接受的载体”是本领域公认的,并且是指在携带或运输任何主题组合物或其组分中涉及的药学上可接受的材料、组合物或媒介物,诸如液体或固体填充剂、稀释剂、赋形剂、溶剂或包封材料。在与主题组合物及其组分相容且对患者无害的意义上,每种载体必须是“可接受的”。可以充当药学上可接受的载体的材料的一些例子包括:(1)糖类,诸如乳糖、葡萄糖和蔗糖;(2)淀粉,诸如玉米淀粉和马铃薯淀粉;(3)纤维素及其衍生物,诸如羧甲基纤维素钠、乙基纤维素和醋酸纤维素;(4)粉状的黄蓍胶;(5)麦芽;(6)明胶;(7)滑石粉;(8)赋形剂,诸如可可脂和栓剂蜡;(9)油,诸如花生油、棉籽油、红花油、芝麻油、橄榄油、玉米油和大豆油;(10)二醇类,诸如丙二醇;(11)多元醇,诸如甘油、山梨醇、甘露醇和聚乙二醇;(12)酯类,诸如油酸乙酯和月桂酸乙酯;(13)琼脂;(14)缓冲剂,诸如氢氧

化镁和氢氧化铝；(15) 海藻酸；(16) 无热原水；(17) 等渗盐水；(18) 林格氏溶液；(19) 乙醇；(20) 磷酸盐缓冲溶液；和(21) 在药物制剂中采用的其它无毒的相容物质。

[0042] 本文中使用的术语“施用”包括将本文所述的化合物和组合物引入患者的所有方式，包括、但不限于口服(po)、静脉内(iv)、肌肉内(im)、皮下(sc)、透皮、吸入、含服、眼部、舌下、阴道、直肠等。可以以含有常规无毒的药学上可接受的载体、佐剂和媒介物的单位剂型和/或制剂施用本文所述的化合物和组合物。

[0043] 固体药物形式可以包含惰性组分和载体物质，诸如碳酸钙、磷酸钙、磷酸钠、乳糖、淀粉、甘露醇、海藻酸盐、明胶、瓜尔胶、硬脂酸镁、硬脂酸铝、甲基纤维素、滑石粉、高度分散的硅酸、硅油、较高分子量脂肪酸(诸如硬脂酸)、明胶、琼脂或植物或动物脂肪和油或固体高分子量聚合物(诸如聚乙二醇)；如果需要的话，适合口服施用的制剂可以包含额外的调味剂和/或甜味剂。

[0044] 液体药物形式可以被灭菌，和/或在适当的情况下包含辅助物质，诸如防腐剂、稳定剂、润湿剂、穿透剂、乳化剂、铺展剂、增溶剂、盐、用于调节渗透压或用于缓冲的糖或糖醇、和/或粘度调节剂。这样的添加剂的例子是酒石酸盐和柠檬酸盐缓冲剂、乙醇和掩蔽剂(诸如乙二胺四乙酸及其无毒盐)。高分子量聚合物诸如液体聚氧化乙烯、微晶纤维素、羧甲基纤维素、聚乙烯吡咯烷酮、葡聚糖或明胶适用于调节粘度。固体载体物质的例子是淀粉、乳糖、甘露醇、甲基纤维素、滑石粉、高度分散的硅酸、高分子量脂肪酸(诸如硬脂酸)、明胶、琼脂、磷酸钙、硬脂酸镁、动物和植物脂肪和固体高分子量聚合物，诸如聚乙二醇。

[0045] 用于胃肠外或局部应用的油性混悬剂可以是植物油、合成油或半合成油，诸如在每种情况下在脂肪酸链中具有8至22个碳原子的液体脂肪酸酯，例如棕榈酸、月桂酸、十三烷酸、十七烷酸、硬脂酸、花生酸、肉豆蔻酸、山嵛酸、十五烷酸、亚油酸、反油酸、巴西烯酸、芥酸或油酸，其用具有1至6个碳原子的一元至三元醇酯化，所述醇诸如甲醇、乙醇、丙醇、丁醇、戊醇或其异构体、乙二醇或甘油。这样的脂肪酸酯的例子是商购可得的miglyols、肉豆蔻酸异丙酯、棕榈酸异丙酯、硬脂酸异丙酯、PEG 6-癸酸、饱和脂肪醇的辛酸/癸酸酯、聚氧乙烯甘油三油酸酯、油酸乙酯、蜡状脂肪酸酯诸如人工鸭尾腺脂肪、椰子脂肪酸异丙酯、油酸油醇酯、油酸癸酯、乳酸乙酯、邻苯二甲酸二丁酯、己二酸二异丙酯、多元醇脂肪酸酯等。不同粘度的硅油或脂肪醇诸如异十三烷醇、2-辛基十二烷醇、鲸蜡基硬脂醇或油醇或脂肪酸诸如油酸也是合适的。此外，还可能使用植物油，诸如蓖麻油、扁桃仁油、橄榄油、芝麻油、棉籽油、花生油、大豆油等。

[0046] 合适的溶剂、胶凝剂和增溶剂是水或水混溶性溶剂。合适的物质的例子是醇类，诸如乙醇或异丙醇、苯甲醇、2-辛基十二烷醇、聚乙二醇类、邻苯二甲酸酯类、己二酸酯类、丙二醇、甘油、二丙二醇或三丙二醇、蜡类、甲基溶纤剂、溶纤剂、酯类、吗啉类、二氧杂环己烷、二甲亚砷、二甲基甲酰胺、四氢呋喃、环己酮等。

[0047] 胶凝剂和成膜剂的混合物也是完全可能的。在这种情况下，特别使用离子大分子诸如羧甲基纤维素钠、聚丙烯酸、聚甲基丙烯酸及其盐、支链淀粉半乙醇酸钠、海藻酸或丙二醇海藻酸盐如钠盐、阿拉伯树胶、黄原胶、瓜尔胶或角叉菜胶。以下可以用作额外的制剂助剂：甘油、不同粘度的石蜡、三乙醇胺、胶原、尿囊素和novantisolic acid。制剂还可能需要使用表面活性剂、乳化剂或润湿剂，例如月桂基硫酸钠、脂肪醇醚硫酸盐、N-月桂基-亚氨基二丙酸二钠、聚乙氧基化的蓖麻油或脱水山梨糖醇单油酸酯、脱水山梨糖醇单硬脂酸酯、

聚山梨酯(例如吐温)、鲸蜡醇、卵磷脂、单硬脂酸甘油酯、聚氧乙烯硬脂酸酯、烷基苯酚聚乙二醇醚、鲸蜡基三甲基氯化铵或单-/二烷基聚乙二醇醚正磷酸单乙醇胺盐。用于稳定乳剂的稳定剂诸如蒙脱石或胶体硅酸,或防止活性物质分解的稳定剂诸如抗氧化剂,例如生育酚或丁基羟基茴香醚,或防腐剂,诸如对羟基苯甲酸酯,同样可以用于制备期望的制剂。

[0048] 用于胃肠外施用的制品可以以单独的剂量单位形式存在,诸如安瓿或管形瓶。优选使用活性化合物的溶液,优选水溶液,特别是等渗溶液以及混悬剂。这些注射形式可以制成作为即用型制品得到或仅在使用前如下直接制备:将活性化合物(例如冻干粉剂,在适当的情况下,含有其它固体载体物质)与期望的溶剂或助悬剂混合。

[0049] 鼻内制品可以作为水性或油性溶液或作为水性或油性混悬剂存在。它们也可以作为在使用前使用合适的溶剂或助悬剂制备的冻干粉剂存在。

[0050] 可吸入制品可以作为粉末、溶液或混悬剂存在。优选地,可吸入制品是粉末形式,例如作为活性成分与合适的制剂助剂诸如乳糖的混合物。

[0051] 在常规的抗微生物和无菌条件下生产、分装和密封制品。

[0052] 如上面指出的,本发明的化合物可以作为与其它活性剂的联合疗法施用,所述其它活性剂例如可用于治疗癌症(例如,前列腺癌、卵巢癌、肺癌或乳腺癌)的治疗活性化合物。对于联合疗法,可以将活性成分配制成为在单一剂量形式中含有几种活性成分的组合物和/或成为在单独剂量形式中含有单独活性成分的试剂盒。在联合疗法中使用的活性成分可以共同施用或单独施用。

[0053] 应当理解,本文所述的化合物和组合物的每日总用量可以由主治医师在合理的医学判断范围内决定。任何特定患者的具体治疗有效剂量水平将取决于多种因素,包括所治疗的障碍和障碍的严重程度;所用的具体化合物的活性;所用的具体组合物;患者的年龄、体重、一般健康、性别和饮食;所用的具体化合物的施用时间和排泄速率,治疗的持续时间,与所用具体化合物联合或同时使用的药物;以及研究人员、兽医、医生或其它具有普通技能的临床医师众所周知的类似因素。

[0054] 取决于施用途径,本文考虑了宽范围的可允许剂量,包括落在从约1 μ g/kg至约1g/kg范围内的剂量。剂量可以是单一的或分开的,并且可以根据多种给药方案施用,包括每天1次(q.d.)、每天2次(b.i.d.)、每天3次(t.i.d.)或甚至每隔一天一次、每周一次、每个月一次等。在每种情况下,本文所述的治疗有效量对应于施用的实例,或者可替换地对应于每日、每周或每月总剂量。

[0055] 本文中使用的术语“治疗有效量”是指在组织系统、动物或人中引起生物学或医学应答的活性化合物或药学试剂的量,所述应答是研究人员、兽医、医生或其它临床医师所正在寻求的,包括减轻正在治疗的疾病或障碍的症状。在一个方面,所述治疗有效量是可以适用于任何医学治疗的合理收益/风险比治疗或减轻疾病或疾病症状的量。

[0056] 本文中使用的术语“治疗有效量”是指要施用给患者的量,并且可以基于体表面积、患者体重和/或患者病况。另外,应当理解,如Freireich, E. J., 等人,癌症 Chemother. Rep. 1966, 50 (4), 219(其公开内容通过引用并入本文)所述,为人类确定的剂量与为动物(包括试验动物)确定的那些剂量之间存在相互关系(示例性地基于每平方米体表面的毫克数)。体表面积可以从患者身高和体重近似确定(参见,例如,Scientific Tables, Geigy Pharmaceuticals, Ardley, New York, 第537-538页(1970))。本文所述

化合物的治疗有效量可以定义为可用于抑制恶性细胞或癌细胞群体的生长(或将其杀死)的任何量,所述群体诸如可以在需要从这样的癌症或恶性肿瘤缓解的患者中发现。通常,这样的有效量范围为从约5 mg/kg至约500 mg/kg、从约5 mg/kg至约250 mg/kg和/或从约5 mg/kg至约150 mg/kg的化合物/患者体重。应当理解,有效剂量还可以随施用途径、任选的赋形剂使用以及化合物与其它常规和非常规治疗性治疗(包括其它抗肿瘤剂、辐射疗法等)的共同使用的可能性而变化。

[0057] 本文中使用的术语“患者”包括人类和非人动物诸如伴侣动物(狗、猫等)和家畜动物。家畜动物是为食物生产而饲养的动物。要治疗的患者优选是哺乳动物,特别是人类。

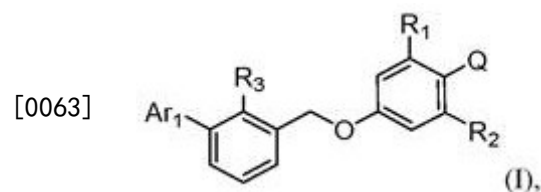
[0058] PD-1程序性细胞死亡-1;PD-L1程序性死亡-配体-1;GNN图神经网络;EGNN能量图神经网络;SVM支持向量机;RF随机森林;FDA,美国食品和药品管理局;HTRF,均相时间分辨荧光;PPh₃,三苯基膦;DIAD,偶氮二甲酸二异丙酯;THF,四氢呋喃;Me₄Si四甲基硅烷;ML,机器学习;AUROC,接受者操作特征曲线下面积。

[0059] 本发明总体上涉及用于治疗用途的新化合物。具体而言,本公开内容涉及可用于治疗各种癌症的具有免疫调节活性的新颖化合物。

[0060] 本文还描述了这样的化合物的药物组合物和通过单独地、与其它治疗剂一起或在药物组合物中施用治疗有效量的这样的化合物来治疗癌症患者的方法。

[0061] 在某些示例性实施方案中,本发明涉及一种药物组合物,其包含一种或多种如在本文中公开的化合物或其药学上可接受的盐、以及一种或多种稀释剂、赋形剂或载体。

[0062] 在某些示例性实施方案中,本发明涉及具有式(I)的化合物或其药学上可接受的盐:



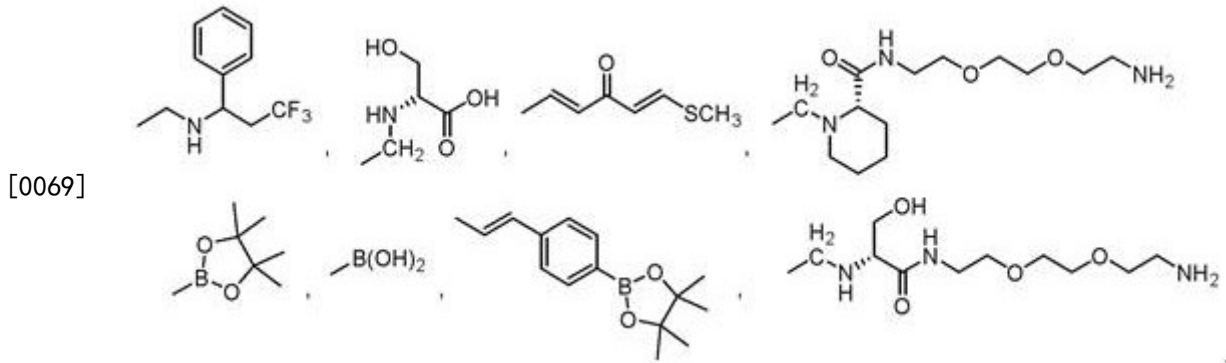
[0064] 其中

[0065] Ar₁是任选地被取代的芳基或杂芳基;

[0066] R₁和R₂独立地是氢、卤素、叠氮基、硝基、氰基、烷基、烯基、炔基、烷基炔基、烷氧基、羟基烷基、氨基烷基、烷硫基(thiolalkyl)、巯基烷基、杂烷基、杂烯基、杂炔基、杂环基、环烷基、环烯基、环杂烷基、环杂烯基、酰基、芳基、杂芳基、芳基烷基、芳基烯基或芳基炔基,它们中的每一个任选地被取代;

[0067] R₃是卤素、叠氮基、硝基、氰基、烷基、烯基、炔基、烷基炔基、烷氧基、羟基烷基、氨基烷基、烷硫基、巯基烷基、杂烷基、杂烯基、杂炔基、杂环基、环烷基、环烯基、环杂烷基、环杂烯基、酰基、芳基、杂芳基、芳基烷基、芳基烯基或芳基炔基,它们中的每一个任选地被取代;且

[0068] Q是:



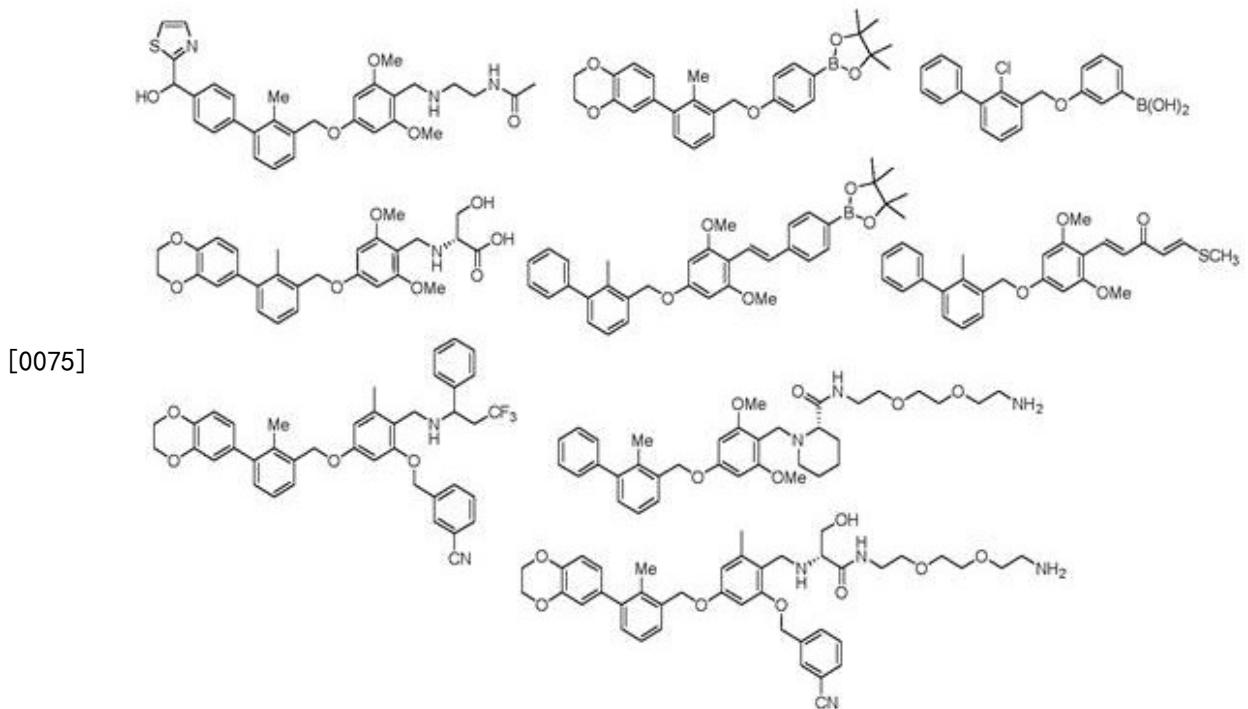
[0070] 或一个或几个氨基酸残基。

[0071] 在某些示例性实施方案中,本发明涉及如在本文中公开的具有式(I)的化合物,其中 Ar_1 是苯基、2,3-二氢苯并[b][1,4]-二氧杂环己烯或苯基(噻唑-2-基)甲醇。

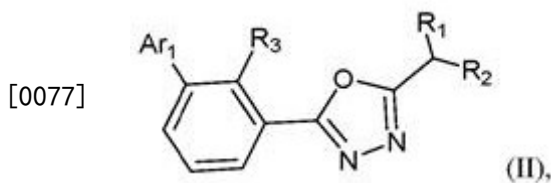
[0072] 在某些示例性实施方案中,本发明涉及如在本文中公开的具有式(I)的化合物,其中 R_1 和 R_2 独立地是氢、甲基、羟基、甲氧基或 $-OCH_2Ar$ 。

[0073] 在某些示例性实施方案中,本发明涉及如在本文中公开的具有式(I)的化合物,其中 R_3 是 CH_3 、CN或Cl。

[0074] 在某些示例性实施方案中,本发明涉及如在本文中公开的具有式(I)的化合物,其中所述化合物包括



[0076] 在某些其它示例性实施方案中,本发明涉及具有式(II)的化合物或其药学上可接受的盐:



[0078] 其中

[0079] R_1 是芳基、被取代的芳基、杂芳基；

[0080] R_2 是伯胺或仲胺,其含有烷基、烯基、炔基、烷基炔基、烷氧基、羟基烷基、氨基烷基、烷硫基、巯基烷基、杂烷基、杂烯基、杂炔基、杂环基、环烷基、环烯基、环杂烷基、环杂烯基、酰基、芳基、杂芳基、芳基烷基、芳基烯基或芳基炔基,它们中的每一个任选地被取代；

[0081] R_3 是卤素、叠氮基、硝基、氰基、烷基、烯基、炔基、烷基炔基、烷氧基、羟基烷基、氨基烷基、烷硫基、巯基烷基、杂烷基、杂烯基、杂炔基、杂环基、环烷基、环烯基、环杂烷基、环杂烯基、酰基、芳基、杂芳基、芳基烷基、芳基烯基或芳基炔基,它们中的每一个任选地被取代；且

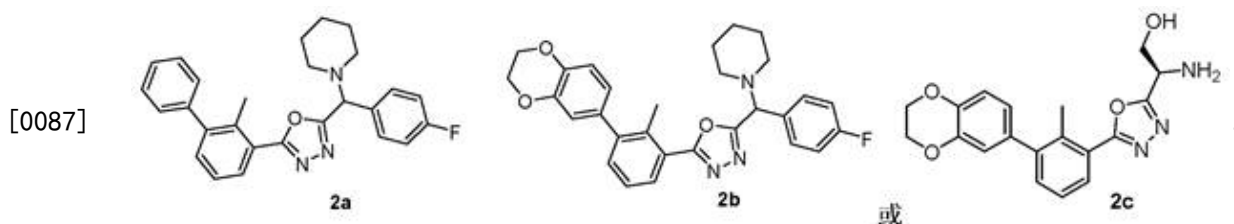
[0082] Ar_1 是芳基、杂芳基、芳基烷基、芳基烯基或芳基炔基,其中每个任选地被取代。

[0083] 在某些示例性实施方案中,本发明涉及如在本文中公开的具有式(II)的化合物,其中 Ar_1 是苯基、2,3-二氢苯并[b][1,4]-二氧杂环己烯或苯基(噻唑-2-基)甲醇。

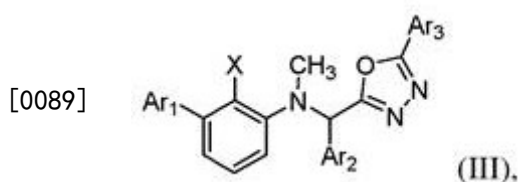
[0084] 在某些示例性实施方案中,本发明涉及如在本文中公开的具有式(II)的化合物,其中 R_1 和 R_2 独立地是哌啶、吡咯烷、苯基、4-卤代苯基(卤代= 氟代、溴代、碘代)和/或作为单个氨基酸或氨基酸组合的一个或多个氨基酸残基。

[0085] 在某些示例性实施方案中,本发明涉及如在本文中公开的具有式(II)的化合物,其中 R_3 是甲基、CN或卤素。

[0086] 在某些示例性实施方案中,本发明涉及如在本文中公开的具有式(II)的化合物,其中所述化合物是



[0088] 在某些示例性实施方案中,本发明涉及具有式(III)的化合物或其药学上可接受的盐:



[0090] 其中

[0091] Ar_1 是苯基、2,3-二氢苯并[b][1,4]-二氧杂环己烯或苯基(噻唑-2-基)甲醇；

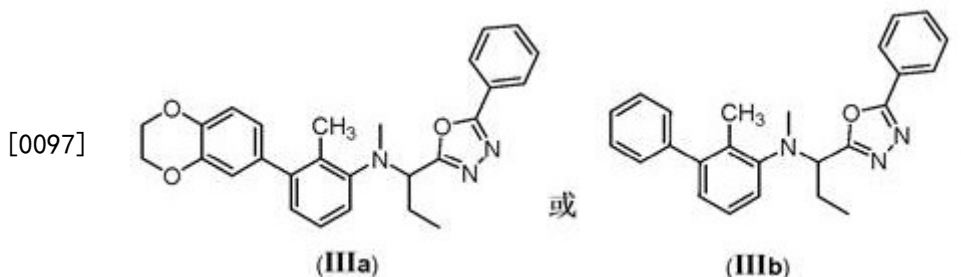
[0092] Ar_2 是哌啶或吡咯烷；

[0093] Ar_3 是苯基、4-卤代苯基(卤代= 氟代、溴代、碘代)；且

[0094] X独立地是卤素、叠氮基、硝基、氰基、烷基、烯基、炔基、烷基炔基、烷氧基、羟基烷基、氨基烷基、烷硫基、巯基烷基、杂烷基、杂烯基、杂炔基、杂环基、环烷基、环烯基、环杂烷基、环杂烯基、酰基、芳基、杂芳基、芳基烷基、芳基烯基或芳基炔基,其中每个任选地被取代。

[0095] 在某些示例性实施方案中,本发明涉及如在本文中公开的具有式(III)的化合物,其中X是甲基、氰基或氯。

[0096] 在某些示例性实施方案中,本发明涉及如在本文中公开的具有式(III)的化合物,其中所述化合物是



[0098] 在某些示例性实施方案中,本发明涉及一种药物组合物,其包含一种或多种具有式(I)、(II)、(III)的化合物或其药学上可接受的盐、以及一种或多种稀释剂、赋形剂或载体。

[0099] 在某些示例性实施方案中,本发明涉及一种或多种具有式(I)、(II)、(III)的化合物或其药学上可接受的盐、以及一种或多种稀释剂、赋形剂或载体,其中所述化合物是免疫调节剂。

[0100] 在某些示例性实施方案中,本发明涉及一种或多种具有式(I)、(II)、(III)的化合物或其药学上可接受的盐、以及一种或多种稀释剂、赋形剂或载体,其中所述化合物是PD-1和PDL-1信号传递途径的抑制剂。

[0101] 在某些示例性实施方案中,本发明涉及一种或多种具有式(I)、(II)、(III)的化合物或其药学上可接受的盐、以及一种或多种稀释剂、赋形剂或载体,其中所述化合物用于治疗癌症。

[0102] 在某些示例性实施方案中,本发明涉及一种用于治疗具有癌症的患者的方法,包括以下步骤:给需要从所述癌症缓解的患者施用治疗有效量的一种或多种具有式(I)、(II)、(III)的化合物或其药学上可接受的盐、和一种或多种载体、稀释剂或赋形剂。

[0103] 在某些示例性实施方案中,本发明涉及一种用于治疗具有癌症的患者的方法,包括以下步骤:给需要从所述癌症缓解的患者施用治疗有效量的一种或多种具有式(I)、(II)、(III)的化合物或其药学上可接受的盐、和一种或多种载体、稀释剂或赋形剂,其中所述化合物是免疫调节剂。

[0104] 在某些示例性实施方案中,本发明涉及一种用于治疗具有癌症的患者的方法,包括以下步骤:给需要从所述癌症缓解的患者施用治疗有效量的一种或多种具有式(I)、(II)、(III)的化合物或其药学上可接受的盐、以及一种或多种具有相同或不同作用模式的其它化合物、和一种或多种载体、稀释剂或赋形剂,其中所述化合物是免疫调节剂。

[0105] 在某些示例性实施方案中,本发明涉及一种用于治疗具有癌症的患者的方法,包括以下步骤:给需要从所述癌症缓解的患者施用治疗有效量的一种或多种具有式(I)、(II)、(III)的化合物或其药学上可接受的盐、和一种或多种载体、稀释剂或赋形剂,其中所述癌症是去势抗性的前列腺癌。

[0106] 在某些示例性实施方案中,本发明涉及一种用于治疗具有癌症的患者的方法,包括以下步骤:给需要从所述癌症缓解的患者施用治疗有效量的一种或多种具有式(I)、

(II)、(III)的化合物或其药学上可接受的盐、和一种或多种载体、稀释剂或赋形剂,其中所述化合物是免疫调节剂。

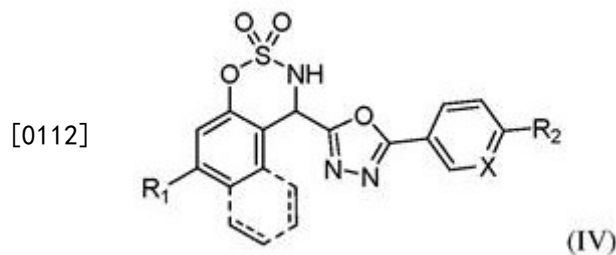
[0107] 在某些示例性实施方案中,本发明涉及用作癌症药物的药物组合物,其包含一种或多种具有式(I)、(II)、(III)的化合物或其药学上可接受的盐、以及一种或多种稀释剂、赋形剂或载体。

[0108] 在某些示例性实施方案中,本发明涉及一种药物缀合物,其包含一种或多种具有式(I)、(II)、(III)的化合物或其药学上可接受的盐、和一种或多种载体、稀释剂或赋形剂。

[0109] 在某些示例性实施方案中,本发明涉及一种药物缀合物,其包含一种或多种具有式(I)、(II)、(III)的化合物或其药学上可接受的盐、和一种或多种载体、稀释剂或赋形剂,其中所述缀合物赋予细胞类型或组织类型靶向或所述缀合物靶向协同那些化合物的作用的另一种途径。

[0110] 在某些示例性实施方案中,本发明涉及一种用于治疗癌症患者的方法,包括以下步骤:给需要从所述癌症缓解的患者施用治疗有效量的一种或多种化合物、以及一种或多种载体、稀释剂或赋形剂,其中所述化合物具有式(I)、(II)或(III)。

[0111] 在某些示例性实施方案中,本发明涉及具有式(IV)的化合物或其药学上可接受的盐:



[0113] 其中

[0114]  代表单键或双键;

[0115]  代表任意的环状环;

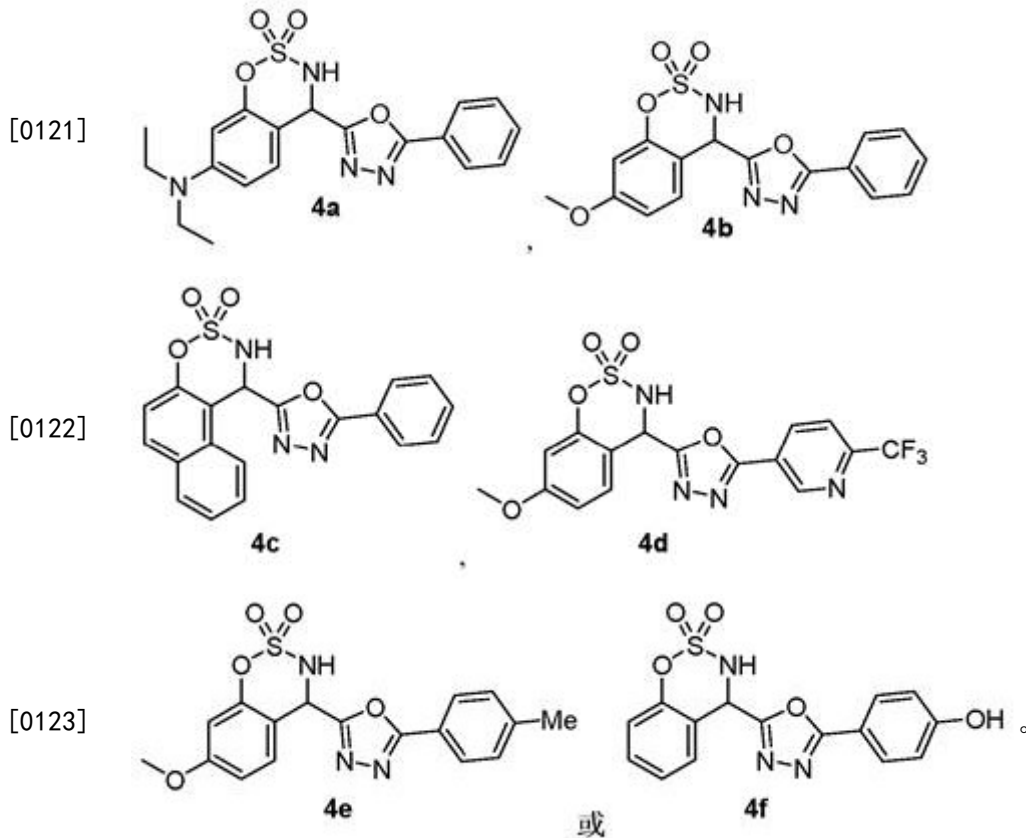
[0116] R_1 是氢、卤素、叠氮基、硝基、氰基、烷基、烯基、炔基、烷基炔基、烷氧基、羟基烷基、氨基烷基、烷硫基、巯基烷基、杂烷基、杂烯基、杂炔基、杂环基、环烷基、环烯基、环杂烷基、环杂烯基、酰基、芳基、杂芳基、芳基烷基、芳基烯基或芳基炔基,它们中的每一个任选地被取代;

[0117] R_2 是氢、卤素、叠氮基、硝基、氰基、烷基、烯基、炔基、烷基炔基、烷氧基、羟基烷基、氨基烷基、烷硫基、巯基烷基、杂烷基、杂烯基、杂炔基、杂环基、环烷基、环烯基、环杂烷基、环杂烯基、酰基、芳基、杂芳基、芳基烷基、芳基烯基或芳基炔基,它们中的每一个任选地被取代;

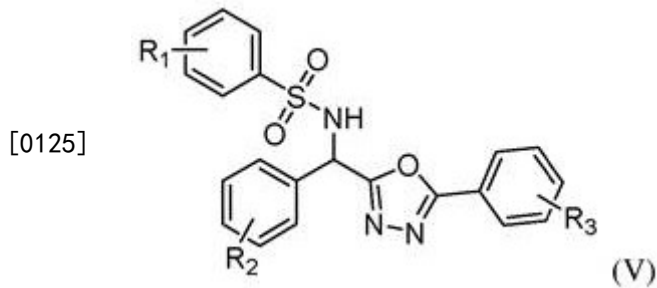
[0118] 且

[0119] X是碳或氮。

[0120] 在某些示例性实施方案中,本发明涉及如在本文中公开的具有式(IV)的化合物,其中所述化合物是



[0124] 在某些示例性实施方案中,本发明涉及具有式(V)的化合物或其药学上可接受的盐:



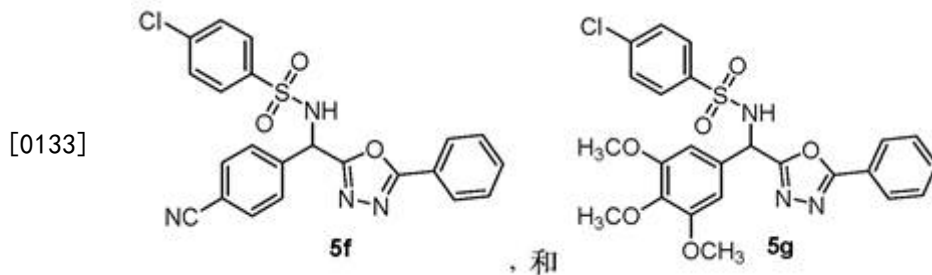
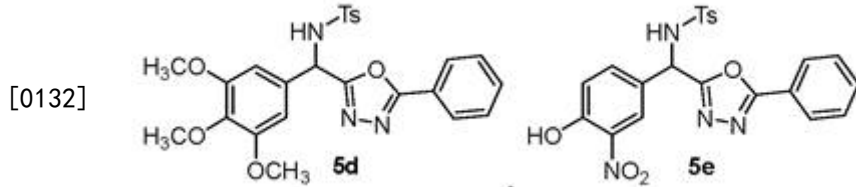
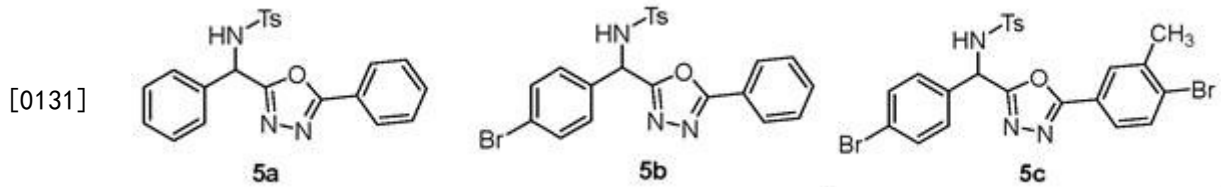
[0126] 其中

[0127] R_1 、 R_2 和 R_3 独立地代表五个取代基,所述取代基选自氢、卤素、叠氮基、硝基、氰基、烷基、烯基、炔基、烷基炔基、烷氧基、羟基烷基、氨基烷基、烷硫基、巯基烷基、杂烷基、杂烯基、杂炔基、杂环基、环烷基、环烯基、环杂烷基、环杂烯基、酰基、芳基、杂芳基、芳基烷基、芳基烯基或芳基炔基,它们中的每一个任选地被取代;或任何两个邻近的取代基连接在一起形成环状或杂环部分。

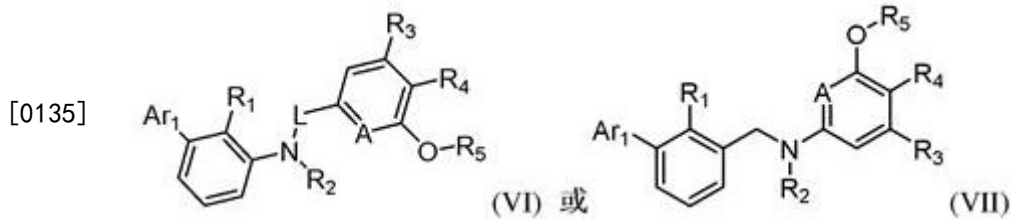
[0128] 在某些示例性实施方案中,本发明涉及如在本文中公开的具有式(V)的化合物,其中 $R_1 = CH_3$ 。

[0129] 在某些示例性实施方案中,本发明涉及如在本文中公开的具有式(V)的化合物,其中 $R_1 = Cl$ 。

[0130] 在某些示例性实施方案中,本发明涉及如在本文中公开的具有式(V)的化合物,其中所述化合物是



[0134] 在某些示例性实施方案中,本发明涉及具有式VI或VII的化合物或其药学上可接受的盐:



[0136] 其中

[0137] A是碳或氮;

[0138] L是 $(\text{CH}_2)_n$ 、 $-\text{SO}$ 、 $-\text{SO}_2$ 、 $-\text{CO}$ 、 $-\text{CO}(\text{CH}_2)_0$,其中n是0、1、2;

[0139] Ar_1 是芳基、杂芳基、芳基烷基、芳基烯基或芳基炔基,它们中的每一个任选地被取代;

[0140] R_1 是卤素、叠氨基、硝基、氰基、烷基、烯基、炔基、烷基炔基、烷氧基、羟基烷基、氨基烷基、烷硫基、巯基烷基、杂烷基、杂烯基、杂炔基、杂环基、环烷基、环烯基、环杂烷基、环杂烯基、酰基、芳基、杂芳基、芳基烷基、芳基烯基或芳基炔基,它们中的每一个任选地被取代;且

[0141] R_2 是H、甲基、乙基或任何烷基;

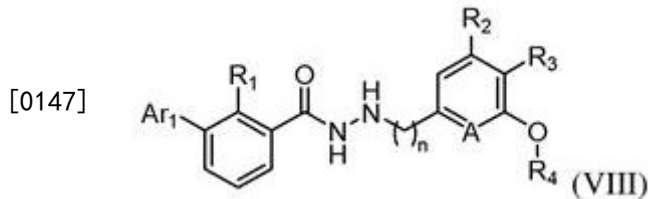
[0142] R_3 是卤素、 $-\text{OR}$ 、 $-\text{NO}_2$ 、氰基、 $-\text{NR}^a\text{R}^b$ 、 $-\text{N}_3$ 、 $-\text{S}(\text{O})_2\text{R}^a$ 、 $-\text{C}$ (烷基)、 $-\text{C}$ (环烷基)、 $-\text{C}$ (炔基)、 $-\text{C}$ (卤代烷基)、酰基、芳基、杂芳基、芳基烷基、芳基烯基或芳基炔基,它们中的每一个任选地被取代,其中 R 、 R^a 和 R^b 独立地是烷基;

[0143] R_4 是卤素、氨基酸、饱和的或不饱和的芳族或杂芳族环、碳水化合物衍生物;或 $-\text{CH}_2)_m\text{NR}^a\text{R}^b$,其中m=0、1、2;其中 R^a 和 R^b 独立地是烷基;

[0144] 或 R_3 和 R_4 连接在一起以形成环状或杂环部分;且

[0145] R_5 是卤素、-OR、-NO₂、氰基、-NR^aR^b、-N₃、-S(O)₂R^a、-C(烷基)、-C(环烷基)、C(炔基)、C(卤代烷基)、酰基、芳基、杂芳基、芳基烷基、芳基烯基或芳基炔基,它们中的每一个任选地被取代,其中R、R^a和R^b独立地是烷基。

[0146] 在某些示例性实施方案中,本发明涉及具有式VIII的化合物或其药学上可接受的盐:



[0148] 其中

[0149] n是0、1、2;

[0150] A是碳或氮;

[0151] R₁是卤素、叠氮基、硝基、氰基、烷基、烯基、炔基、烷基炔基、烷氧基、羟基烷基、氨基烷基、烷硫基、巯基烷基、杂烷基、杂烯基、杂炔基、杂环基、环烷基、环烯基、环杂烷基、环杂烯基、酰基、芳基、杂芳基、芳基烷基、芳基烯基或芳基炔基,它们中的每一个任选地被取代;且

[0152] R₂独立地是卤素、-OR、-NO₂、氰基、-NR^aR^b、-N₃、-S(O)₂R^a、-C(烷基)、-C(环烷基)、C(炔基)、C(卤代烷基)、酰基、芳基、杂芳基、芳基烷基、芳基烯基或芳基炔基,它们中的每一个任选地被取代,其中R、R^a和R^b独立地是烷基,其中R、R^a和R^b独立地是烷基;

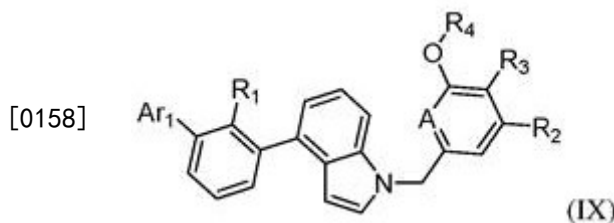
[0153] 或R₂和R₃连接在一起以形成环状或杂环部分;

[0154] R₃独立地是-(CH₂)_mNR^aR^b:其中m=0至2、卤素、任何氨基酸、任何饱和的或不饱和的芳族或杂芳族环或碳水化合物衍生物,其中R^a和R^b独立地是烷基;

[0155] R₄独立地是卤素、-OR、-NO₂、氰基、-NR^aR^b、-N₃、-S(O)₂R^a、-C(烷基)、-C(环烷基)、C(炔基)、C(卤代烷基)、酰基、芳基、杂芳基、芳基烷基、芳基烯基或芳基炔基,它们中的每一个任选地被取代,其中R、R^a和R^b独立地是烷基,其中R、R^a和R^b独立地是烷基;且

[0156] Ar₁是芳基、杂芳基、芳基烷基、芳基烯基或芳基炔基,其中每个任选地被取代。

[0157] 在某些示例性实施方案中,本发明涉及具有式IX的化合物或其药学上可接受的盐:



[0159] 其中

[0160] n是0、1、2;

[0161] R₁是卤素、叠氮基、硝基、氰基、烷基、烯基、炔基、烷基炔基、烷氧基、羟基烷基、氨基烷基、烷硫基、巯基烷基、杂烷基、杂烯基、杂炔基、杂环基、环烷基、环烯基、环杂烷基、环杂烯基、酰基、芳基、杂芳基、芳基烷基、芳基烯基或芳基炔基,它们中的每一个任选地被取

代;且

[0162] R_2 独立地是卤素、-OR、-NO₂、氰基、-NR^aR^b、-N₃、-S(O)₂R^a、-C(烷基)、-C(环烷基)、C(炔基)、C(卤代烷基)、酰基、芳基、杂芳基、芳基烷基、芳基烯基或芳基炔基,它们中的每一个任选地被取代,其中R、R^a和R^b独立地是烷基;

[0163] R_3 独立地是-(CH₂)_mNR^aR^b:其中m=0至2、卤素、任何氨基酸、任何饱和的或不饱和的芳族或杂芳族环或碳水化合物衍生物,其中R^a和R^b独立地是烷基;

[0164] R_4 独立地是卤素、-OR、-NO₂、氰基、-NR^aR^b、-N₃、-S(O)₂R^a、-C(烷基)、-C(环烷基)、C(炔基)、C(卤代烷基)、酰基、芳基、杂芳基、芳基烷基、芳基烯基或芳基炔基,它们中的每一个任选地被取代,其中R、R^a和R^b独立地是烷基;

[0165] 且

[0166] Ar₁是芳基、杂芳基、芳基烷基、芳基烯基或芳基炔基,其中每个任选地被取代。

[0167] 在某些示例性实施方案中,本发明涉及一种药物组合物,其包含一种或多种如在本文中公开的具有式(IV)~(IX)的化合物或其药学上可接受的盐以及一种或多种稀释剂、赋形剂或载体。

[0168] 在某些示例性实施方案中,本发明涉及一种药物组合物,其包含一种或多种如在本文中公开的具有式(IV)~(IX)的化合物或其药学上可接受的盐以及一种或多种稀释剂、赋形剂或载体,其中所述化合物是免疫调节剂。

[0169] 在某些示例性实施方案中,本发明涉及一种药物组合物,其包含一种或多种如在本文中公开的具有式(IV)~(IX)的化合物或其药学上可接受的盐以及一种或多种稀释剂、赋形剂或载体,其中所述化合物是PD-1和PDL-1信号传递途径的抑制剂。

[0170] 在某些示例性实施方案中,本发明涉及一种用于治疗癌症患者的方法,所述方法包括以下步骤:施用治疗有效量的一种或多种如在本文中公开的具有式(IV)~(IX)的化合物或其药学上可接受的盐以及一种或多种稀释剂、赋形剂或载体。

[0171] 在某些示例性实施方案中,本发明涉及一种药物组合物,其包含一种或多种如在本文中公开的具有式(IV)~(IX)的化合物或其药学上可接受的盐以及一种或多种稀释剂、赋形剂或载体,其中所述化合物是PD-1和PDL-1信号传递途径的抑制剂。

[0172] 在某些示例性实施方案中,本发明涉及一种用于治疗癌症患者的方法,所述方法包括以下步骤:给需要从所述癌症缓解的癌症患者施用治疗有效量的一种或多种如在本文中公开的具有式(IV)~(IX)的化合物以及一种或多种载体、稀释剂或赋形剂。

[0173] 在某些示例性实施方案中,本发明涉及一种用于治疗癌症患者的方法,所述方法包括以下步骤:给需要从所述癌症缓解的癌症患者施用治疗有效量的一种或多种如本文中公开的具有式(IV)~(IX)的化合物以及一种或多种具有相同或不同作用模式的其它化合物和一种或多种载体、稀释剂或赋形剂。

[0174] 在某些示例性实施方案中,本发明涉及一种用于治疗癌症患者的方法,所述方法包括以下步骤:给需要从所述癌症缓解的癌症患者施用治疗有效量的一种或多种如在本文中公开的具有式(IV)~(IX)的化合物和一种或多种载体、稀释剂或赋形剂,其中所述化合物是PD-1和PDL-1信号传递途径的抑制剂。

[0175] 在某些示例性实施方案中,本发明涉及一种用于治疗癌症患者的方法,所述方法包括以下步骤:给需要从所述癌症缓解的癌症患者施用治疗有效量的一种或多种如在本文

中公开的具有式 (IV)~(IX) 的化合物和一种或多种载体、稀释剂或赋形剂,其中所述癌症是去势抗性的前列腺癌。

[0176] 在某些示例性实施方案中,本发明涉及用作癌症药物的药物组合物,其包含一种或多种如在本文中公开的具有式 (I)~(IX) 的化合物和一种或多种载体、稀释剂或赋形剂。

[0177] 程序性细胞死亡蛋白1/程序性死亡-配体1 (PD-1/PD-L1) 相互作用是癌细胞用来增强免疫抑制的免疫检查点。与抗体相比,迫切需要开发起效快、有成本效益且易于生物利用的小分子药物。不幸的是,以盲目的方式合成和验证抑制PD-1/PD-L1相互作用的大型小分子库以既费时又昂贵。为了改进该药物发现流程,我们已经开发了一种机器学习方法,其针对专利数据进行了训练以鉴定、合成和验证PD-1/PD-L1小分子抑制剂。我们的模型包含两个特征:作为全局特征的代表结合能量 (E) 的对接评分,和代表局部特征的通过分子拓扑结构的图神经网络 (GNN) 的子图特征。这种能量-图神经网络 (EGNN) 模型优于传统的机器学习方法以及简单的GNN (对于保留试验集,具有0.9524的F1评分和0.8861的Cohen氏 κ 评分),从而提示,小分子的拓扑结构、结合袋中的结构相互作用和训练数据的化学多样性都是增强模型性能的重要考虑因素。使用自举的EGNN模型选择化合物进行合成并对预测的抑制PD-1/PD-L1相互作用的高和低效力进行实验验证。有效抑制剂 (4-((3-(2,3-二氢苯并[b][1,4]二氧杂环己烯-6-基)-2-甲基苄基)氧基)-2,6-二甲氧基苄基)-D-丝氨酸是两种已知生物活性骨架的杂交体,且具有339.9 nM的IC₅₀值,其相对优于已知的生物活性化合物。我们得出的结论是,我们的自举的EGNN模型将可用于鉴定通过骨架跳跃(一种众所周知的药物化学技术)设计的靶标特异性的高效分子。

[0178] 传统上,小分子抑制剂的开发需要对多种类药化合物的大型库进行高通量筛选¹⁰或药物化学家对具有弱受体活性的骨架的迭代以增强效能¹¹。这个过程(i)耗时;(ii)需要昂贵的仪器和机器人;(iii)基于反复实验法;和(iv)快速鉴定几种新骨架的效率非常低¹²。此外,已经开发了使用对接方法的虚拟筛选来改进该过程,但收效甚微¹³。此外,ML架构诸如支持向量机(SVM)¹⁴⁻¹⁶、随机森林(RF)¹⁷⁻¹⁹、图卷积网络²⁰和图神经网络(GNN)^{21,22}已经用于药物设计和预测药物-靶标相互作用^{23,24}。近来,利用蛋白结合位点中的图特征的组合的新架构已经在计算结合亲和力和确定化合物是否将结合靶标方面显示出巨大希望^{20,22}。

[0179] 还已经提出了几种新的基于神经网络的架构,其有望鉴定有效骨架,但许多尚未经过实验性试验^{15,16,25-28},并且挖掘和表征蛋白晶体学数据的能力的发展希望会推动这些模型的创建²⁹。近来,已经表明,通过GNN并入的分子子图特征和由它们的序列编码的蛋白特征可以组合以预测化合物是否可以靶向给定的蛋白²⁴。受这项工作启发并基于我们对开发药物设计和免疫学方法的兴趣²⁹⁻³⁶,我们已经开发了一种新的机器学习模型来预测化合物是否可以抑制PD-1/PD-L1相互作用。我们的方法用代表结合自由能的对接评分替代蛋白序列特征,并且由于结合袋中小分子的这种全局能量相互作用,我们已经将此模型称为“能量图神经网络”(EGNN)。使用CANDOCK³¹计算三维原子相互作用能量评分(图1B),并使用端对端训练方法与局部分子图特征(图1A)组合(图1C-D)。在这项工作中,我们使用该EGNN模型来选择合成设计,并实验性地试验来自这些预测的精选化合物列表,以使用均相时间分辨荧光(HTRF)测定前瞻性地鉴定有效的PD-L1小分子抑制剂。我们还试验了负向预测,表明所述模型可以用于选择作为PD-1/PD-L1抑制剂的有效先导物。

[0180] 用于训练EGNN模型的专利数据

[0181] 我们使用来自四项专利的762种化合物的PD-1/PD-L1小分子抑制数据来训练我们的模型：BMS的WO 2015/034820 A1⁷和WO 2015/160641 A2 (674种化合物)⁶，以及Incyte Corporation的WO 2018/119263 A1³⁷和US 2018/0273519 A1³⁸ (88种化合物)。在专利中，使用均相时间分辨荧光 (HTRF) 结合测定显示针对PD-1/PD-L1相互作用的活性。但是，所述专利并未列出所有化合物的单独的IC₅₀值，而是提供了不同分子的抑制范围。因此，我们为两个数据集训练了带有截止值的二元分类器，以将分子视为“高效”或“低效” (图1)。如果在专利中报道的分子的IC₅₀小于或等于100 nM，则将其视为“高效”分子，否则将其视为“低效”分子。选择该阈值是因为它是四项专利中唯一的共同阈值 (表S4)。应当指出，此处不应考虑IC₅₀的实际值，因为我们的多次重复实验不能获得专利中某些分子的准确报告结果 (参见表2中化合物4a的IC₅₀值，在WO 2015/034820 A1专利7中注解了6-100 nM的BMS-1)。因此，我们考虑与WO 2015/034820 A1专利7中的BMS对照分子 (化合物4a/BMS-1) 的上限相比基于我们的实验IC₅₀值的正向预测 (高效)。带有BMS或Incyte注解的762种小分子的训练数据集显示在支持信息文件 (Supporting Information File) (TrainingData.xlsx) 中。

[0182] 我们选择BMS和Incyte专利以将分子的化学多样性包括在训练数据集中。图2A显示了BMS和Incyte专利中低效和高效分子和一般骨架的分布。BMS专利分别具有372种高效化合物和302种低效化合物，而Incyte专利分别具有47种高效化合物和41种低效化合物。BMS专利骨架含有 (2-甲基-3-联苯基) 甲醇的417种衍生物和 [3- (2,3-二氢-1,4-苯并二氧杂环己烯-6-基)-2-甲基苯基] 甲醇的257种衍生物，如图2A (左下) 所示，R基团为CN、Cl、Br和CH₃。另一方面，Incyte专利骨架具有不同的子骨架，在图2A中表示为A和B (右下)。对于Incyte骨架，X表示N或C-R基团 (R: 烷基基团)。这些骨架提示，Incyte化合物的化学多样性高于BMS化合物的化学多样性，因为在Incyte专利中的化合物的一般结构对于子骨架和原子而言具有更多的结构多样性。我们使用BMS和Incyte化合物的成对Tanimoto相似性评分验证了这一观察结果，分别如图2B和2C中的热图所示。使用半径为2、位长为1024的Morgan指纹来计算成对Tanimoto相似性。在BMS热图中的高红色区域指示，分子对在结构上彼此相似。在Incyte热图中的低红色区域表明，它在分子结构方面具有更多的化学多样性。此外，发现所有BMS化合物的平均成对Tanimoto相似性评分对于所有Incyte化合物分别为0.4434和0.3920，这证实了与BMS化合物相比在Incyte化合物中更高的化学多样性。

[0183] PD-L1同源二聚体和PD-1/PD-L1晶体结构揭示了对接的结合位点

[0184] 以前已经表明，BMS化合物通过诱导PD-L1的二聚化来抑制PD-1/PD-L1相互作用^{8,9}。因此，选择了PD-L1同源二聚体晶体结构 (PDB ID: 5N2F) 来对接本手稿中的所有化合物。还使用PyMol软件包⁴⁰使用PD-1/PD-L1晶体结构 (PDB ID: 4ZQK) 检查同源二聚体晶体结构 (5N2F) 中的PD-L1的结合位点位置是否重叠和相互对齐 (图3A)。在图3B中，显示了PD-L1同源二聚体在重叠的且对齐的晶体结构上的选定结合位点以指示PD-L1与小分子的同源二聚体的形成阻断了PD-1/PD-L1相互作用。在选定的结合位点中的PD-1/PD-L1相互作用的已知抑制剂 (配体ID: 8HW)⁸ (图3C) 提示，对应于PD-L1同源二聚体的选定结合位点与开发PD-1/PD-L1抑制剂有关。因此，PD-L1同源二聚体的对接相互作用将与鉴定PD-1/PD-L1抑制剂有关。此外，没有进行与PD-1/PD-L1的直接对接，因为在PD-1和PD-L1之间的结合位点充满了来自两种蛋白的相互作用的氨基酸残基。因此，没有空间将小化合物与PD-1/PD-L1复合物对接。

[0185] 使用CANDOCK³¹生成小分子与PD-L1同源二聚体的对接构象(细节参见关于在EGNN中使用对接和能量向量(E)生成能量特征的实验部分)。在开发机器学习方法之前,我们还评估了仅使用训练集中化合物的对接评分对CANDOCK中的96个潜在能量评分函数³¹中的每一个进行高效对(vs)低效分子分类的能力。将Cohen氏 κ 评分用于选择最佳评分函数来区分两个类别(表S1)。评分函数径向累积完成15(radial cumulative complete 15) (RCC15)获得了0.41447的最高Cohen氏 κ 评分。但是,RCC15评分不能清楚地区分训练数据中所有高效和低效分类分子(参见图S1中的小提琴图)。仅使用一个评分函数不足以捕捉小分子的PD-1/PD-L1抑制的不同状态。因此,我们使用每个类别的最顶级评分函数开发了EGNN模型,所述评分函数表现出正的Cohen氏 κ 值(表S1)来代表全局特征(细节参见关于在EGNN中使用对接和能量向量(E)生成能量特征的实验部分)。这包括RCR15 ($\kappa=0.37746$)和RCC15 ($\kappa=0.41447$)评分函数。 κ 评分在0.21 - 0.40之间的模型被认为是相当好的一致性模型,且如果 κ 评分是在0.41 - 0.60之间,则所述模型被认为是中等的一致性模型⁴²。

[0186] 具有超参数优化的EGNN模型优于GNN和其它基线模型

[0187] 在实验部分中给出了EGNN模型的详细描述,包括与对接组合的分子GNN的组合。图1显示,EGNN模型是表示为GNN的小分子的局部特征(参见EGNN中的分子图的图神经网络)与表示为对接评分的蛋白-配体相互作用的全局特征的组合(参见章节在EGNN中使用对接和能量向量(E)生成能量特征)。EGNN使用88种小分子进行训练,所述小分子对PD-1/PD-L1的抑制而言具有高效和低效,来自两个Incyte专利(参见用于训练EGNN模型的专利数据)。我们计算了相对于不同超参数的epoch数目的平均F1评分的变化(经五个交叉验证倍数)(图S2)。选择最佳超参数以避免EGNN的过拟合和欠拟合,包括:隐藏分子向量的维度(dim) = 10,子图半径= 2,和隐藏层数= 1(参见关于EGNN训练和超参数优化的实验部分)。

[0188] 将EGNN和GNN模型使用不同的训练集进行训练,以检查化学多样性对高效和低效分子的分类的模型性能的影响。分别和组合使用两个数据集(BMS和Incyte)来训练EGNN模型并确定预测PD-1/PD-L1抑制剂的最佳数据集。使用两种不同的方法将数据集拆分为训练-验证集和试验集(4:1):(1)对混洗的数据使用随机拆分器;和(2)通过DeepChem库使用骨架拆分方法⁴³。然后用五倍交叉验证进行训练,并使用试验集来评价模型的预测能力。在这里,测量了Cohen氏 κ 、F1评分和接受者操作特征曲线下面积(AUROC)以对比仅用BMS数据、仅用Incyte数据和BMS-Incyte组合数据训练的三个模型。此外,作为一个单独的实验,针对仅在BMS数据上训练的EGNN和GNN模型获得所有测量,同时为Incyte数据预测,反之亦然。

[0189] 图4A显示了如何在训练-验证-试验集方案中拆分数据集。将初始数据集基于骨架拆分或随机混洗拆分以4:1比例分成两组。然后将80%的数据集用作训练和验证数据集,而将20%的数据集用作保留试验集以评价模型性能。

[0190] 图4A显示了使用BMS化合物、Incyte化合物以及这些集合的并集训练的两个模型的平均F1评分(经五倍交叉验证)。使用Incyte数据训练的EGNN和GNN模型的平均F1评分分别为0.956 (± 0.051)和0.678 (± 0.157) (图4A)。该结果提示,使用包含多种化学骨架(图2C)的Incyte数据的EGNN训练模型比使用相同数据集训练的GNN表现好得多。但是,当仅用化学多样性低于Incyte的BMS化合物进行相同试验时,平均F1评分对于两种模型而言是相当的,EGNN模型为0.992 (± 0.007),GNN模型为0.948 (± 0.022)。这表明,与训练数据中较大的化学多样性相比,GNN模型在训练数据中较小化学多样性的情况下表现良好。但是,

EGNN模型在两个数据集都表现良好,表明它是优于GNN的模型。

[0191] 具有骨架拆分的GNN模型似乎产生了与EGNN相当的结果(图S4)。但是,这是在意料之中,因为图神经网络在训练中使用了二维分子框架/拓扑结构。当化合物的框架分布相似于在训练-验证和试验集内时,GNN表现良好。但是,我们的意图是开发一个模型,其可以用于筛选大型化合物库,所述库不需要与训练集(即Incyte或BMS)共享相同的骨架分布。因此,我们选择了带有改组的随机拆分器来创建用于性能评价的试验集以开发更通用的模型。

[0192] 对于用BMS化合物、Incyte化合物和这些集的并集训练的两种模型,不同试验集(保留试验集是基于随机拆分)的Cohen氏 κ 评分如图4B所示。使用Incyte数据训练并在保留试验集上试验的EGNN和GNN模型的 κ 评分分别为0.8861和0.4304(图4B)。该结果表明,使用包含多种化学骨架(图2C)的Incyte数据训练的EGNN模型比使用相同数据集训练的GNN表现好得多。但是,当仅用化学多样性低于Incyte的BMS化合物进行相同的试验时,两种模型的Cohen氏 κ 评分是相当的,EGNN模型为0.6416,且GNN模型为0.7164。这表明,与较大的化学多样性相比,GNN模型在训练和试验数据中的化学多样性较小时表现良好。两种模型对于组合数据集也显示出相当的性能。当组合BMS和Incyte数据集二者时,EGNN模型的保留试验集的 κ 评分为0.6072,而GNN模型为0.6729。对三种不同训练集对比,对F1评分观察到了类似的趋势(图4C)。这些结果表明,EGNN模型在化学多样性数据集(诸如Incyte数据)方面优于GNN模型。我们认为这是由于与GNN模型训练数据中小分子的仅“局部”结构特征相比,增加了通过PD-L1同源二聚体的对接评分捕获的“全局”能量特征作为EGNN中的训练数据。

[0193] 我们还研究了在一个化合物集上训练的EGNN和GNN模型预测另一个化合物集内PD-1/PD-L1的高效和低效抑制剂的能力。这些结果表示在图4B和4C(分别为 κ 评分和F1评分)中,用不同的条形图案代表不同的试验集。Incyte和BMS化合物之间的Tanimoto相似性也显示在图4D热图中。0.3044的平均成对Tanimoto相似性评分表明,这两个数据集内的化合物彼此非常不同。当将EGNN和GNN模型在BMS化合物上进行训练并用于预测Incyte化合物时,分别观察到EGNN的0.1505和GNN的0.1200的Cohen氏 κ 评分,且F1评分为0.3810和0.2264。另一方面,当将它们使用Incyte数据训练并用于预测BMS化合物时,两种模型的F1和 κ 评分均有改善(κ 评分为EGNN = 0.3852,GNN = 0.3196和F1评分为EGNN = 0.7400,GNN = 0.6958)。这些结果显示,在用Incyte数据进行训练并在BMS上进行试验时,EGNN和GNN模型的F1评分和Cohen氏 κ 评分都有显著改善。但是,AUROC评分无法正确区分这些模型(图3)。这是在意料之中,因为AUROC不是评价在偏斜/不平衡数据集上训练的模型的好量度,并且它可能阻碍模型的差性能⁴⁴。此外,这些结果提示,EGNN模型在两种情况下都优于GNN模型(细节参见表S2)。这突出了训练数据中化学多样性的重要性,即使在训练集和试验集之间没有太多的化合物相似性。因此,不适合使用单独的BMS或组合的BMS和Incyte数据来训练最终模型对未知分子进行预测。因此,我们仅选择Incyte数据集来训练EGNN模型,这显著改善了EGNN模型性能(表S2)。

[0194] 此外,仅通过对比Tanimoto 2D相似性,我们已经检查了将训练集化合物分类为低效和高效类别的能力。显示低效、高效和所有化合物的Tanimoto 2D相似性评分的分布的小提琴图如图S5所示。这清楚地表明,无论是高效类别还是低效类别的化合物都显示出具有低成对相似性评分的高概率,并且即使考虑到所有化合物也是如此。因此,仅考虑拓扑相似

性来选择有效的PD-1/PD-L1抑制剂是不够的。

[0195] 最后,使用它们的试验集性能,我们将使用Incyte训练数据训练的交叉验证的EGNN模型与GNN、支持向量机(SVM)和随机森林(RF)基线模型进行对比。还在局部和全局特征方面训练了SVM和RF模型二者。使用Weisfeiler-Lehman算法⁴⁵从半径为2的分子图中提取的指纹被用作类似于EGNN和GNN模型的“局部”特征。在这里,我们将零填充到最大指纹长度以保持相同的指纹尺寸。通过CANDOCK获得的相同预选对接评分(RCR15和RCC15)被用作“全局”能量特征。对于所有四个模型,获得的AUROC、AUPRC、精确度、调用、F1评分和Cohen氏 κ 值列在表1中。使用具有“线性”内核的scikit-学习库⁴⁶中的“svm”包训练SVM模型,并使用具有500棵树的scikit-学习库⁴⁶中的“随机森林分类器”训练RF。将scikit-学习包⁴⁶中的“metrics”模块用于统计AUROC、精确度、调用、F1评分和Cohen氏 κ 。使用R编程语言中的“precrec”库⁴⁷获得模型和AUPRC值的精确召回曲线。EGNN模型优于所有其它模型,AUROC、AUPRC、精确度、调用、F1评分和Cohen氏 κ 的值分别为0.9250、0.9212、0.9091、1.0000、0.9524和0.8861(表1)。对比这四个模型的精确召回曲线(图4E)也证实了EGNN模型优于所有其它三个模型。总之,用Incyte数据集,在EGNN中组合的局部和全局特征提供了最佳性能。

[0196] 表1. 与其它基线模型(诸如,随机森林、SVM和GNN模型)相比,用于PD-1/PD-L1抑制剂预测的EGNN的AUROC、AUPRC、精确度、调用、F1评分和Cohen氏 κ 。所有模型都在Incyte数据集上进行了训练,并基于相同的保留试验集进行了评价。

量度	模型			
	随机森林	SVM	GNN	EGNN
AUROC	0.8125	0.7750	0.8625	0.9250
AUPRC	0.7419	0.8266	0.8688	0.9212
精确度	0.7692	0.8000	0.7273	0.9091
调用	1.0000	0.8000	0.8000	1.0000
F1 评分	0.8696	0.8000	0.7619	0.9524
Cohen 氏 κ	0.6494	0.5500	0.4304	0.8861

[0198] 用于PD-1/PD-L1抑制的EGNN预测的合成选择和验证

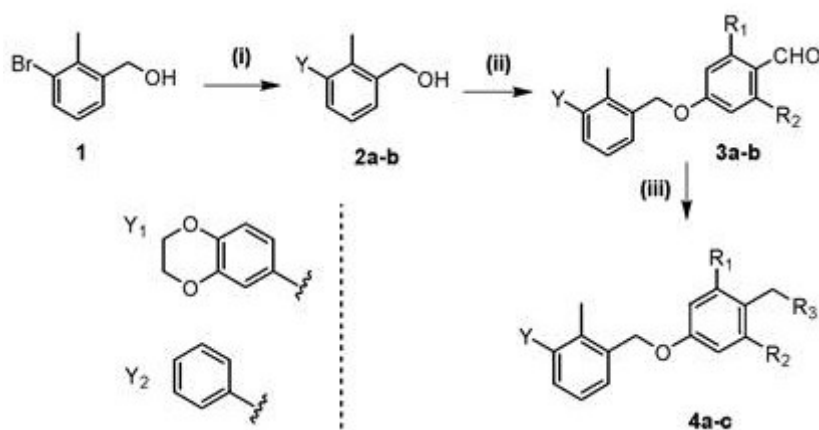
[0199] 将使用最适超参数和Incyte数据集训练的EGNN模型用于获得对小分子设计的内部数据库的预测。我们开发了自举的EGNN模型以使用1000个EGNN模型预测对PD-1/PD-L1抑制具有高效和低效的化合物(参见自举EGNN模型部分)。自举是一种重要的统计技术,其可以用于根据多个模型之间的一致性选择可靠的分子进行合成和实验验证。自举的EGNN模型鉴定出作为PD-1/PD-L1抑制剂的高效和低效小分子,它们被合成并然后用HTRF结合测定进行实验验证(总结参见表2)。具体来说,我们根据自举的EGNN SoftMax平均评分和标准偏差(参见表2中的EGNN SoftMax评分)选择了4种预测对PD-1/PD-L1抑制为高效或低效的分子进行试验。

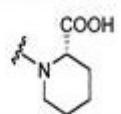
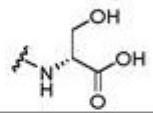
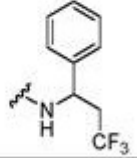
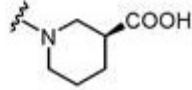
[0200] 在EGNN自举的预测中,我们已经选择了1种高效分子(化合物4b)和3种具有不同骨架的低效分子(化合物4c、4d和4e)用于进一步试验。我们已经定义了一个名为“计数(counts)”的新参数,它记录了100个模型中为感兴趣的分子给出0.5或更高的SoftMax评分

的模型的数目。具体而言,预测化合物4b是一种高效PD-1/PD-L1抑制剂,计数为99,且平均SoftMax评分为0.7771 (± 0.1193)。相比之下,化合物4c的结果为仅69的计数和0.5786 (± 0.1406)的平均SoftMax值,化合物4d的结果为仅5的计数和0.1821 (± 0.1514)的平均SoftMax值,且化合物4e的结果为62的计数和0.5280 (± 0.1259)的平均SoftMax值,表明低效预测。我们还合成了BMS骨架(化合物4a,一种已知的PD-1/PD-L1抑制剂),用作我们的HTRF实验的阳性对照。化合物结构显示在方案1和2中(关于程序和表征参见实验部分)。预测的高效分子(化合物4b)是(4-((3-(2,3-二氢苯并[b][1,4]二氧杂环己烯-6-基)-2-甲基苄基)氧基)-2,6-二甲氧基苄基)-D-丝氨酸,两个BMS分子的杂交体,即分别含有(2-甲基-3-联苯基)甲醇和[3-(2,3-二氢-1,4-苯并二氧杂环己烯-6-基)-2-甲基苄基]甲醇的4a (BMS-1)和BMS-1002 (图5A),并表明了EGNN模型进行骨架跳跃的能力。

[0201] 使用Incyte数据训练的EGNN仅具有三种含有[3-(2,3-二氢-1,4-苯并二氧杂环己烯-6-基)-2-甲基苄基]甲醇骨架的化合物。作为一个单独的实验,我们已经删除了这些并为我们的合成库进行了预测,并且EGNN模型仍然能够预测化合物4b (一种基于[3-(2,3-二氢-1,4-苯并二氧杂环己烯-6-基)-2-甲基苄基]甲醇骨架的化合物)作为高效化合物,具有 0.8285 ± 0.1396 的SoftMax评分和971的计数。该结果表明,EGNN模型可以鉴定具有[3-(2,3-二氢-1,4-苯并二氧杂环己烯-6-基)-2-甲基苄基]甲醇骨架的高效PD-1/PD-L1抑制剂,即使没有在类似的骨架上学习。

[0202] 方案1. 代表性的合成方案^a



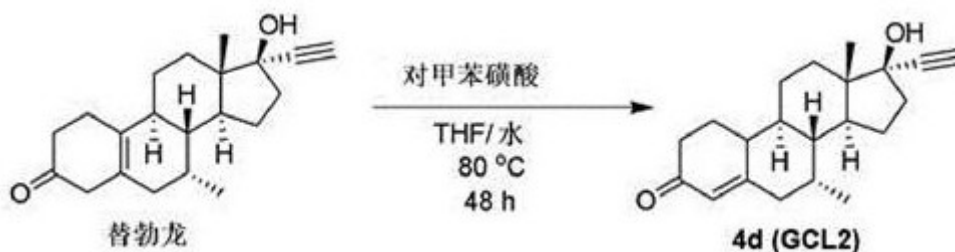
化合物	Y	R ₁	R ₂	R ₃
4a (KPGC01S94)	Y ₂	OMe	OMe	
4b (KPGC01S32)	Y ₁	OMe	OMe	
4c (KPGC01S138)	Y ₂	OMe	OMe	
4e (KPGC01S42)	Y ₂	OMe	OMe	

[0204]

[0205] ^a反应条件: (i) BH₃·THF复合物(1.0 M在THF中),无水THF,0℃至室温,2天; (ii) PPh₃,DIAD,0℃至室温,20 h,无水THF; (iii) 胺组分,NaBH₃CN,催化量AcOH,DMF,80℃或室温,1h或3 h或过夜。

[0206] 方案2. 4d (GCL2)的合成

[0207]



[0208] 化合物4b的PD-L1同源二聚体(PDB ID: 5N2F)中的顶部对接位姿以与先前关于共晶结构所示类似的方式相互作用⁸(图5B)。具体而言,对于化合物4b,2,3-二氢-1,4-苯并二氧杂环己烯基团促进了PD-L1同源二聚体的链A中氨基酸残基Tyr56 (_ATyr56)的移动。已知的是,如果该2,3-二氢-1,4-苯并二氧杂环己烯基团存在,则该_ATyr56不从一端封闭疏水袋⁸从而在对接构象中创建疏水隧道(图5B插图)而不是疏水裂缝。此外,化合物4b(2,3-二氢-1,4-苯并二氧杂环己烯)的芳族环通过与氨基酸残基_ATyr56的 π - π 堆积作用而被稳定化(图5C)。在结构中的中心甲基苄基环(图5A中4b中的品红色)对2,3-二氢-1,4-苯并二氧杂环己烯环旋转了大约30°,且甲基苄基环的甲基基团指向PD-L1同源二聚体的链B。这种取向导致与同源二聚体的链A和B二者的Met115以及与_BAla121的疏水相互作用。4b化合物的D-丝氨酸末端与_AThr20和_AAla121形成氢键,并且在4b分子的_ATyr123的主链NH和两个甲氧基基团之一中的氧之间形成似乎合理的氢键(图5C)。这些结果表明,可以使PD-L1二聚化的化合物4b的有利相互作用将导致PD-1/PD-L1抑制。

[0209] HTRF测定证实,化合物4b抑制PD-1/PD-L1相互作用的IC₅₀为339.9 nM(细节参见实验部分)(图5D)。这比我们实验室合成和试验的BMS化合物4a(BMS专利WO 2015/034820 A1中的BMS-1分子)的521.5 nM的IC₅₀相对更好。应当指出,BMS-1分子在BMS专利中被表示在

HTRF测定中具有6-100 nM的 IC_{50} ⁷。但是,我们的实验的多次重复并未得到抑制PD-1/PD-L1相互作用的小于100 nM的 IC_{50} 值(参见 IC_{50} 值的计算部分和支持文件HTRF_IC50_Data.xlsx)。如前所述,该结果不影响我们的机器学习方法,因为我们已经根据高效和低效对分子进行分类,而不是估计具体的 IC_{50} 值。我们的工作和专利之间的这种实验结果差异的一种可能解释是,用于进行HTRF测定和 IC_{50} 值计算的方案的差异。出于这个原因,我们已经包括了HTRF测定方案的详细说明、用于计算 IC_{50} 的数据的分析以及供科学界使用的支持数据文件(参见实验部分)。为了试验我们的自举的EGNN模型正确鉴定低效预测的有效性,我们还试验了化合物4c、4d和4e,结果是对PD-1/PD-L1相互作用的无/差的抑制(表2)。每种试验化合物的 IC_{50} 图(图S6)以及¹³C和¹H NMR谱作为支持信息予以提供。这些4a-4e之间的成对Tanimoto相似性评分(表S3)显示EGNN模型鉴定高效和低效抑制剂的能力,而不管结构相似性如何。化合物4e与对照BMS化合物(4e)显示出高相似性,Tanimoto相似性评分为0.8018。但是,所述模型将它识别为一种低效分子,且实际试验显示它是一种差的PD1/PD-L1抑制剂, IC_{50} 为1261 nM。另一方面,所述模型将化合物4b鉴定为高效PD1/PD-L1抑制剂,并且HTRF测定证实它具有339.9 nM的非常好的 IC_{50} 。但是,它与对照化合物(4a)的成对Tanimoto相似性评分仅为0.5074。总之,这些结果表明,自举的EGNN模型可以用于选择分子进行合成和PD-1/PD-L1抑制的实验验证,并且它可以鉴定在结构上与对照化合物4a相似的低效分子。

[0210] 表2. 预测的活性和无活性化合物的 IC_{50} 值与EGNN SoftMax评分

化合物	IC_{50} 值	SoftMax 评分	预测
4a (KPGC01S94)	521.5 nM	对照*	对照*
4b(KPGC01S32)	339.9 nM	0.7771 +/- 0.1193	高效
4c (KPGC01S138)	无抑制	0.5786 +/- 0.1406	低效
4d (GCL.2)	无抑制	0.1821 +/- 0.1514	低效
4e (KPGC01S42)	1261 nM (差抑制)	0.5280 +/- 0.1259	低效

[0212] *表示BMS活性化合物。

[0213] 讨论和结论

[0214] 癌症免疫疗法标志着治疗癌症的重要步骤,并且PD-1/PD-L1免疫检查点抑制剂的开发已经成为治疗多种肿瘤的一个重要研究领域。目前,靶向PD-1 (派姆单抗、纳武单抗和西米普利单抗)和PD-L1 (阿特朱单抗、度伐单抗和阿维鲁单抗)的六种治疗性抗体已获得美国FDA批准。近年来,已经开发了几种新的小分子PD-1/PD-L1抑制剂⁴³,以及人PD-1/PD-L1复合物和抑制性配体的共晶的结构测定⁴⁴⁻⁴⁶。本领域仍然非常活跃地在寻找新的小分子来抑制这个重要的检查点,并且我们希望通过使用新的基于结构的ML方法来提高这种搜索的速度,所述方法已经经过广泛的基准测试和前瞻性试验。

[0215] 我们已经开发了一种新的ML方法EGNN,其基于组合小分子拓扑结构的局部特征和在结合袋内相互作用的小分子的全局特征作为能量评分来选择、合成和实验验证PD-1/PD-L1相互作用的有效抑制剂。具体来说,EGNN优于传统的ML架构,诸如,包括局部和全局特征

二者的RF、SVM,以及仅使用小分子拓扑结构的局部特征的GNN模型。当以来自BMS和Incyte专利数据的已知PD-1/PD-L1抑制剂为基准时,我们得出的结论是,小分子的拓扑结构、结合袋中的结构相互作用以及训练数据的化学多样性都是增强模型性能的重要考虑因素。

[0216] 我们使用自举的EGNN模型(基于1000个EGNN模型)来预测和自信地选择新分子用于化学合成并随后使用HTRF PD-1/PD-L1抑制测定进行抑制试验。我们相信,自举是一种重要的统计技术,其在药物设计中与ML方法一起使用以自信地选择分子进行实验验证。预测的高效分子(4-((3-(2,3-二氢苯并[b][1,4]二氧杂环己烯-6-基)-2-甲基苄基)氧基)-2,6-二甲氧基苄基)-D-丝氨酸是两种BMS高效分子骨架的杂交体,并具有抑制PD-1/PD-L1相互作用的339.9 nM的IC₅₀值,这表明EGNN模型进行骨架跳跃以鉴定新的抑制剂的能力。具有不同骨架的低效分子的准确选择表明我们的自举模型在选择化合物进行合成方面的实际效用,这是基于ML的药物设计领域的一个难题。

[0217] 我们的EGNN方法可以通过添加更多化学多样性数据来进一步开发,并将强化迭代学习与在每个步骤中进行的实验相结合,以开发抑制PD-1/PD-L1相互作用的结构多样的小分子库以指导构效关系。鉴于机器学习模型的一般性质和易于使用的对接方法,这种方案可以适用于通过靶向其它免疫检查点来鉴定小分子免疫调节剂,以及通常用于包括局部和全局特征进行基于靶标的药物设计。

[0218] 实验部分

[0219] 用于试验预测化合物的抑制的均相时间分辨荧光(HTRF)测定

[0220] 使用来自Cisbio US, Inc.的PD1/PD-L1 HTRF测定试剂盒试验了4种高效和低效预测化合物对PD-1/PD-L1相互作用的抑制。对于每种预测化合物(4b、4c、4d和4e)和BMS对照化合物(4a),如在试剂盒中所述使用测定方案。简而言之,将2 μ L的化合物、4 μ L的25 nM Tag1-PD-L1蛋白溶液和4 μ L的250 nM Tag2-PD1蛋白添加到Cisbio氏HTRF 96-孔低体积白板中。然后,将板在室温温育15分钟。接下来,加入10 μ L预混合的抗-标签检测试剂(5 μ L的1X抗-Tag1-Eu³⁺和5 μ L的1X抗-Tag2-XL665),并将密封的板在室温温育2小时。最后,除去板密封物,并使用HTRF[®]兼容读数器进行测量。该方案使用12种不同浓度的每种化合物,其中最大和最小测定浓度分别为10,000 nM和0.001 nM。对高效预测化合物4b(36个数据点)和阳性对照化合物4a(48个数据点)进行了不同浓度的几个重复。将由 $\Delta F/\Delta F_{max}$ (使用HTRF比率665 nm/620 nm计算)相对于log[浓度]表示的归一化信号的拟合曲线用于确定化合物的50%抑制浓度(IC₅₀) (参见下一部分IC₅₀值的计算)。

[0221] 为了计算 $\Delta F/\Delta F_{max}$,首先如下计算HTRF比率:

$$[0222] \quad \text{HTRF 比率} = \frac{\text{信号 } 665 \text{ nm}}{\text{信号 } 620 \text{ nm}} \times 10000。$$

[0223] 10000因子的乘法因子不用于处理在计算期间提高数据准确性的十进制值。如下通过从每个化合物(样品)HTRF比率中减去背景HTRF比率(在我们工作中的阴性DMSO对照),计算指示破坏PD-1/PD-L1相互作用的化合物的“特定信号”的 ΔR 比率:

$$[0224] \quad \Delta R = \text{HTRF 比率 (样品)} - \text{HTRF 比率 (背景)}。$$

[0225] 接下来,进行数据归一化以最小化在不同日期、不同读板器仪器上或由不同个体进行测定时的值变化。相对于背景HTRF比率进行归一化,并如下计算:

$$[0226] \quad \Delta F = \frac{\text{HTRF比率(样品)} - \text{HTRF比率(背景)}}{\text{HTRF比率(背景)}} \times 100\%。$$

[0227] 最后,计算 $\Delta F/\Delta F_{max}$ 比率以实现多个实验之间的值的对比。

$$[0228] \quad \Delta F/\Delta F_{max} = \frac{\Delta F(\text{样品})}{\Delta F_{max}}$$

[0229] 其中取 ΔF_{max} 作为测定中阳性DMSO对照的 ΔF 。

[0230] IC_{50} 值的计算

[0231] 通过使用用于Windows的GraphPad Prism软件8.3.0版(GraphPad Software, La Jolla California USA, www.graphpad.com)分析浓度-响应曲线的对数以拟合具有四参数逻辑(4PL)回归的S型曲线,确定PD-1/PD-L1抑制的 IC_{50} 值。 IC_{50} 值提供在表2中。以下方程式定义了回归曲线。

$$[0232] \quad Y = Bottom + \frac{(Top - Bottom)}{(1 + 10^{((LogIC50 - X) \times HillSlope)})}$$

[0233] 其中 X = 浓度的对数, Y = 应答 $\Delta F/\Delta F_{max}$, Top 和 $Bottom$ = 与 Y 相同单位的平台, $LogIC50$ = 与 X 相同的对数单位, $HillSlope$ = 斜率因子或Hill斜率,无单位。使用以上方程式,计算 $LogIC50$ 以得到每种化合物的 IC_{50} 值。包含所有重复的HTRF_IC50_Data.xlsx数据文件作为支持文件提供,用在GraphPad Prism软件中以计算 IC_{50} 值。

[0234] EGNN模型的机器学习架构

[0235] 使用PyTorch⁴⁷开发EGNN模型。用于实现机器学习模型和结果的所有脚本都提供在GitHub上,参见<https://github.com/chopralab/egnn>。图1显示了EGNN机器学习架构的概述。我们为Tsubaki和同事的分子图²⁴实现了图神经网络。简而言之,使用ChemAxon MolConverter⁴⁸软件将分子结构转换为SMILES字符串。然后使用RDKit⁴⁹软件包和Weisfeiler-Lehman算法提取分子的r半径子图的图(图1A)。以下部分包括EGNN架构的详细信息。

[0236] EGNN中的分子图的图神经网络

[0237] 以下方程式和符号以及分子GNN的详细信息已在此处从原始作品²⁴中复制出来,并进行了少量修改以进行澄清。小写粗体字母(例如 $\mathbf{v} \in \mathbb{R}^d$)表示向量,大写粗体字母(例如 $\mathbf{M} \in \mathbb{R}^{m \times n}$)表示矩阵,且斜体非粗体字母(例如 $S, G, v,$ 和 e)表示标量、集合、图、顶点和边。GNN将分子图转换为具有两个基于神经网络的函数的低维实值向量 $\mathbf{y} \in \mathbb{R}^d$;转换和输出²¹。在图 G 中,每个顶点(v)通过转换函数考虑其相邻顶点和边的信息进行更新。这些顶点已通过输出函数映射进实值向量 $\mathbf{y} \in \mathbb{R}^d$ 。两个函数都是可微的。GNN模型的所有输入特征和权重都在交叉熵损失函数的帮助下使用反向传播进行更新。

[0238] 一个图可以定义为 $G = (V, E)$,在这里; V 和 E 分别是顶点和边的集合。当应用于化学时,原子可以定义为顶点,化学键可以定义为边。首先,所有原子和化学键都将嵌入为具有基于其不同类型的d维的实值向量。由于小分子中原子(例如:C、N、O等)和键(例如:单键、

双键、三键等)的多样性是有限的,因此学习参数的数目是有限的。因此,使用了一种称为r-半径子图⁵⁰的策略来避免这种限制。

[0239] r-半径子图

[0240] 在定义的 r 半径内的所有原子的集合,一个原子 i 可以表示为 $N(i,r)$ 。当 $r=0$, $N(i,r) = \{i\}$ 时,它是分子中所有原子的集合。第 i 个顶点(v_i)的r-半径子图定义如下;

[0241] $v_i^{(r)} = (V_i^{(r)}, E_i^{(r)})$, 其中,

[0242] $V_i^{(r)} = \{v_j | j \in N(i,r)\}$

[0243] $E_i^{(r)} = \{e_{mn} \in E | (m,n) \in N(i,r) \times N(i,r-1)\}$

[0244] 第 i 个和第 j 个原子(e_{ij})之间的边的r-半径子图定义如下;

[0245] $e_{ij}^{(r)} = (V_i^{(r-1)} \cup V_j^{(r-1)}, E_i^{(r-1)} \cap E_j^{(r-1)})$

[0246] 根据类型将随机初始化的嵌入(图1)分配给每个r-半径边 $e_{ij}^{(r)}$ 和顶点($v_i^{(r)}$)。反向传播已用于训练这些随机嵌入。

[0247] 顶点转换函数

[0248] 说 $v_i^{(t)} \in \mathbb{R}^d$ 是在时间步长 t 处给定分子图 G 的第 i 个顶点的嵌入向量。然后更新的 $v_i^{(t+1)} \in \mathbb{R}^d$ 向量可以写成如下;

[0249]
$$v_i^{(t+1)} = \sigma \left(v_i^{(t)} + \sum_{j \in N(i)} h_{ij}^{(t)} \right)$$

[0250] 其中, $N(i)$ 表示相邻原子的集合, σ 是被定义为 $\sigma(x) = \frac{1}{1+e^{-x}}$ 的S型函数,且 $h_{ij}^{(t)} \in \mathbb{R}^d$ 是定义邻域的隐藏向量,并可以如下计算;

[0251]
$$h_{ij}^{(t)} = f \left(W_{neighbor} \begin{bmatrix} v_j^{(t)} \\ e_{ij}^{(t)} \end{bmatrix} + b_{neighbor} \right)$$

[0252] 其中, f 是细化线性单元(ReLU),一种非线性活化函数诸如 $f(x) = \max(0,x)$ 。 $W_{neighbor} \in \mathbb{R}^{d \times 2d}$ 和 $b_{neighbor} \in \mathbb{R}^{d \times 1}$ 分别是权重矩阵和偏置向量。在时间步长 t 以后分子图的第 i 个和第 j 个原子(顶点)之间的向量定义为 $e_{ij}^{(t)}$ 。

[0253] 边转换函数

[0254] 如前所述,在训练过程中使用边转换函数更新每个嵌入的边向量 $e_{ij}^{(t)}$ 。

[0255]
$$e_{ij}^{(t+1)} = \sigma \left(e_{ij}^{(t)} + g_{ij}^{(t)} \right)$$

[0256]
$$g_{ij}^{(t)} = f \left(W_{edge} \left(v_i^{(t)} + v_j^{(t)} \right) + b_{edge} \right)$$

[0257] 其中, $W_{edge} \in \mathbb{R}^{d \times d}$ 、 $b_{edge} \in \mathbb{R}^{d \times 1}$ 分别是权重矩阵和偏置向量。此外,添加了 $v_i^{(t)}$

和 $\mathbf{v}_j^{(t)}$, 因为分子图中的边没有方向。

[0258] 分子GNN的分子向量输出

[0259] 转换函数生成一组更新的原子(顶点)向量 $(V = \{\mathbf{v}_1^{(t)}, \mathbf{v}_2^{(t)}, \dots, \mathbf{v}_{|V|}^{(t)}\})$ 。然后输出函数使用这组原子向量获得独特分子向量 $\mathbf{y}_{molecule} \in \mathbb{R}^d$ (图1A), 其定义如下:

$$[0260] \quad \mathbf{y}_{\text{分子}} = \frac{1}{|V|} \sum_{i=1}^{|V|} \mathbf{v}_i^{(t)}$$

[0261] 其中, 完整分子图中的顶点总数由 $|V|$ 表示。

[0262] 在EGNN中使用对接和能量向量(E)生成能量特征

[0263] 首先, 使用MarvinSketch⁵¹软件仔细绘制所有报告的分子。然后, 所有绘制的分子都在3D中清理并转换为sybyl.mol2文件, 所述文件用于与我们内部的CANDOCK³¹软件包(版本0.6.0)对接, 使用默认参数(具有20,000 max_num_possible)和20% 顶级种子百分比(图1B)。CANDOCK源代码可在GitHub上在<https://github.com/chopralab/candock/releases/tag/v0.6.0>得到。用PD-L1同源二聚体晶体结构(PDB ID: 5N2F)完成对接。我们根据蛋白结构中晶体配体的坐标(配体ID: 8HW)选择结合位点。然后, 将radial-mean-reduced-6 (RMR6)³¹用作“选择器(Selector)”参数用于对接以选择顶部位姿³¹。接着, 选择每种对接化合物的顶部位姿, 并使用CANDOCK³¹软件中所有可用的96个不同潜在能量函数重新计算其对接评分。每种分子的所有96个CANDOCK对接能量评分都针对每个潜在能量函数进行归一化, 以用作EGNN模型中的向量;

$$[0264] \quad \hat{S}_{i,j} = \left(\frac{S_{i,j} - \min(S_j)}{\max(S_j) - \min(S_j)} \right)$$

[0265] 其中, $i: 1 \rightarrow n$ 和 $j: 1 \rightarrow m$ 。在这里, n 是潜在能量评分函数的数目, 和 m 是数据集内分子的数目。 $\hat{S}_{i,j}$ 是第 j 个对接分子的用第 i 个潜在能量函数的能量评分的归一化对接能量值。类似地, $S_{i,j}$ 是归一化前的对接能量评分。并且, $\max(S_j)$ 和 $\min(S_j)$ 是所有对接分子的第 j 个评分函数内的最大和最小能量值。然后使用scikit-学习包⁴⁶中的Cohen_κ评分工具为所有训练集数据的每个评分函数计算Cohen氏κ评分。选择了所有给出正Cohen氏κ评分的评分函数, 并为EGNN模型选择了每个类中的顶部。因此, EGNN模型中每种分子的归一化对接评分向量使用RCR15和RCC15归一化的潜在能量评分函数表示为 $\mathbf{y}_{energy} \in \mathbb{R}^2$ (图1B)。

[0266] EGNN的输出

[0267] 如图1C所示, 归一化的对接能量评分向量(\mathbf{y}_{energy})与GNN的分子向量输出($\mathbf{y}_{molecule}$)连接。然后, 如下将连接的长向量 $(\mathbf{y}_{molecule} \oplus \mathbf{y}_{energy}) \in \mathbb{R}^{(d+96)}$ 用于训练以得到输出向量 $\mathbf{x}_{output} \in \mathbb{R}^2$;

$$[0268] \quad \mathbf{x}_{output} = \mathbf{W}_{output} (\mathbf{y}_{molecule} \oplus \mathbf{y}_{energy}) + \mathbf{b}_{output}$$

[0269] 其中 \oplus 表示连接, $\mathbf{W}_{output} \in \mathbb{R}^{2 \times (d+96)}$ 表示权重矩阵, 且 $\mathbf{b}_{output} \in \mathbb{R}^2$ 表示偏置

向量。然后,将SoftMax分类器(图1D)添加到 $\mathbf{x}_{output} = [y_0, y_1]$ 向量的顶部以获得高或低效率的概率。

$$[0270] \quad p_t = \frac{e^{y_t}}{\sum_i e^{y_i}}$$

[0271] 其中, $t \in \{0,1\}$; 0指示低效且1指示高效,且 p_t 是给定 y_t 的概率。

[0272] 自举EGNN模型

[0273] 该模型使用随机数来初始化边和顶点向量。因此,最终模型使用自举来获得预测。训练了一千个具有不同随机种子的不同模型,并获得了内部分子设计试验集的预测。平均SoftMax评分用作自举模型的最终预测结果。最后,合成试验集中的分子根据平均SoftMax评分分类为高效或低效。如果它大于或等于0.5,则认为它是高效,否则认为它是低效(图1)。因此,将EGNN模型使用给定的SMILES字符串、由CANDOCK³¹生成的RCR15和RCC15评分的向量以及它们与PD-L1蛋白的高效或低效状态进行反向传播训练。训练过的模型可以用于预测给定分子是对PD-L1蛋白的高效或低效分子的概率。

[0274] EGNN训练和超参数优化

[0275] 该模型将给定分子的SMILES字符串和对接能量评分字符串作为输入。模型的超参数在用于预测之前进行了优化。通过考虑五倍交叉验证的F1评分,优化了GNN隐藏向量的维度(dim)、GNN的隐藏层数和子图半径。为GNN隐藏分子向量输出的维度使用了三个值(即dim = 5、10和15)。数字1、2和3用于检查GNN中隐藏层的最佳数目。最后,从半径= 1、2和3中选出模型的最佳子图半径。

[0276] F1评分和Cohen氏 κ 的计算

[0277] 以下术语用于计算Cohen氏 κ 和F1评分。预测为高效且实验报告为高效的化合物的数目被认为是真阳性(TP)。预测为高效但在专利中实验报告为低效的化合物的数目被认为是假阳性(FP)。将真阴性(TN)定义为预测为低效且实验也报告为低效的化合物的数目。然后将假阴性(FN)定义为预测为低效但实验报告为高效的化合物的数目。F1评分定义如下。

$$[0278] \quad F1 \text{ 评分} = \frac{2 \cdot \text{精确度} \cdot \text{调用}}{\text{精确度} + \text{调用}}$$

[0279] 其中精确度和调用定义如下。

$$[0280] \quad \text{精确度} = \frac{TP}{TP + FP}$$

$$[0281] \quad \text{调用} = \frac{TP}{TP + FN}$$

[0282] Cohen氏 κ :

$$[0283] \quad \text{Cohen氏 } \kappa = \frac{P_o - P_e}{1 - P_e}$$

[0284] P_o = 在评级者之间的相对观察一致

$$[0285] \quad P_o = \frac{TP + TN}{TP + TN + FP + FN}$$

[0286] P_e =随机一致的概率

$$[0287] \quad P_e = \frac{[TP + FN \quad FP + TN] \begin{bmatrix} TP + FP \\ FN + TN \end{bmatrix}}{(TP + TN + FP + FN)^2}$$

[0288] 合成

[0289] 除非另外指出,否则所有试剂和溶剂购自商业来源并以接收状态使用。在螺旋帽小瓶中进行所有反应。使用 Me_4Si 作为内标使用500 MHz获得质子(^1H)和碳(^{13}C) NMR谱,并以 δ 单位报告。以Hz为单位报告偶合常数(J 值)。使用快速色谱(Teledyne ISCO EZprep)在硅胶上进行柱色谱。使用电喷雾电离(ESI)技术和以TOF质量分析仪获得高分辨率质谱图(HRMS)。有机溶剂和起始材料以接收状态使用。根据报道的程序从化合物1、2a-b、3a-b开始合成BMS化合物4a (BMS-1或KPGC01S94)⁷以及化合物4b-c,且光谱数据与报道的数据一致⁶⁻⁸。

[0290] 化合物4a (BMS-1或KPGC01S94), (S)-1-(2,6-二甲氧基-4-((2-甲基-[1,1'-联苯]-3-基)甲氧基)苄基)哌啶-2-甲酸:将来自方案1的3b (45 mg, 0.125 mmol)、(S)-哌啶-2-甲酸(64.5 mg, 4当量, 0.5 mmol)、氰基硼氢化钠(40.8 mg, 5.2当量, 0.65 mmol)溶解在DMF (1 mL)中,然后加入乙酸(2滴)。将反应混合物在80°C搅拌1小时。通过TLC监测反应。将粗制物通过0-20% DCM:MeOH纯化以提供作为灰白色固体的期望产物(31.5 mg, 53%收率)。 ^1H NMR (500 MHz, $\text{DMSO}-d_6$) δ 7.49 - 7.41 (m, 3H), 7.39 - 7.34 (m, 1H), 7.32 - 7.25 (m, 3H), 7.19 (dd, $J = 7.7, 1.5$ Hz, 1H), 6.41 (s, 2H), 5.17 (s, 2H), 4.08 (s, 2H), 3.78 (s, 7H), 3.11 (t, $J = 5.5, 5.5$ Hz, 1H), 3.08 - 2.99 (m, 1H), 2.60 (dd, $J = 13.5, 6.7$ Hz, 1H), 2.20 (s, 3H), 1.80 (q, $J = 6.0, 5.9, 5.9$ Hz, 2H), 1.55 (q, $J = 6.7, 6.1, 6.1$ Hz, 2H), 1.37 (ddt, $J = 18.4, 12.8, 6.5, 6.5$ Hz, 3H); ^{13}C NMR (126 MHz, $\text{DMSO}-d_6$) δ 172.0, 161.47, 160.32, 142.70, 141.85, 135.83, 134.59, 130.25, 129.63, 128.86, 128.72, 127.44, 126.04, 92.05, 69.23, 64.35, 56.42, 48.70, 46.21, 31.16, 26.11, 22.13, 21.27, 16.41。

[0291] 化合物4b (KPGC01S32), (4-((3-(2,3-二氢苯并[b][1,4]二氧杂环己烯-6-基)-2-甲基苄基)氧基)-2,6-二甲氧基苄基)-D-丝氨酸:将来自方案1的3a (35.6 mg, 0.104 mmol)、D-丝氨酸(32.8 mg, 3当量)、氰基硼氢化钠(19.6 mg, 3当量)溶解在DMF (1 mL)中,然后加入乙酸(0.104 mmol, 1当量, 2滴)。将反应混合物在室温搅拌过夜。通过TLC监测反应。将粗制物通过0-20% MeOH:DCM纯化以提供作为灰白色固体的期望产物(42%收率)。 ^1H NMR (500 MHz, $\text{DMSO}-d_6$) δ 7.42 (dd, $J = 7.6, 1.5$ Hz, 1H), 7.22 (t, $J = 7.6, 7.6$ Hz, 1H), 7.15 (dd, $J = 7.6, 1.5$ Hz, 1H), 6.90 (d, $J = 8.2$ Hz, 1H), 6.76 (d, $J = 2.1$ Hz, 1H), 6.73 (dd, $J = 8.2, 2.1$ Hz, 1H), 6.37 (s, 2H), 5.13 (s, 2H), 4.26 (s, 4H), 3.86 (s, 2H), 3.77 (s, 6H), 3.58 (dt, $J = 8.5, 3.4, 3.4$ Hz, 2H), 2.94 (t, $J = 6.0, 6.0$ Hz, 1H, NH), 2.20 (s, 3H); ^{13}C NMR (126 MHz, $\text{DMSO}-d_6$) δ 172.85, 160.75, 159.69, 159.53, 143.42, 142.96, 142.12, 135.85,

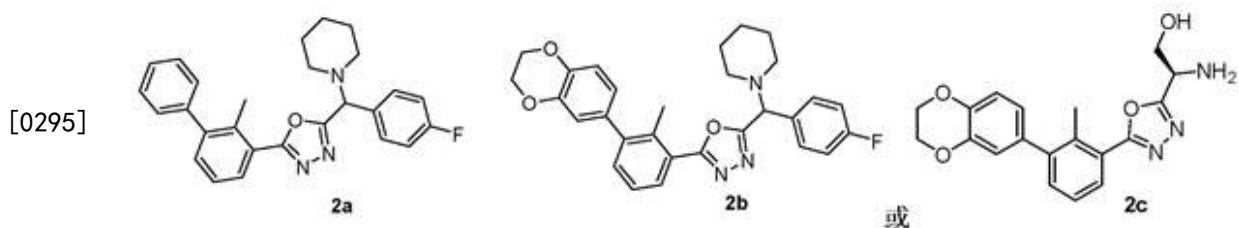
134.96, 134.65, 130.20, 128.56, 125.93, 122.59, 118.17, 117.26, 104.65, 92.06, 69.18, 64.57, 62.51, 61.34, 56.43, 56.32, 16.41; HRMS (ESI): 对于 $C_{28}H_{32}NO_8$ $[M + H]^+$, 实测 510.2132 m/z; 计算质量, 510.2128。

[0292] 化合物4c (KPGC01S138), *N*-(2,6-二甲氧基-4-((2-甲基-[1,1'-联苯]-3-基)甲氧基)苄基)-3,3,3-三氟-1-苯基丙烷-1-胺: 将来自方案1的3b (8 mg, 0.022 mmol)、3,3,3-三氟-1-苯基丙烷-1-胺 (16.7 mg, 0.088 mmol, 4当量)、氰基硼氢化钠 (7.2 mg, 0.114 mmol, 5.2当量) 溶解在 DMF (0.5 mL) 中, 然后加入乙酸 (1滴)。将反应混合物在 80°C 搅拌 3 小时。通过 TLC 监测反应。将粗制物通过 0-20% MeOH:DCM 纯化以提供作为油状产物的期望产物 (68% 收率)。 1H NMR (500 MHz, $CDCl_3$) δ 7.45 - 7.38 (m, 5H), 7.38 - 7.30 (m, 6H), 7.30 - 7.27 (m, 1H), 7.26 (d, $J = 5.4$ Hz, 1H), 6.22 (s, 2H), 5.08 (s, 2H), 3.96 (t, $J = 6.5$, 6.5 Hz, 1H), 3.75 (s, 6H), 3.73 (d, $J = 6.5$ Hz, 2H), 2.46 (s, 1H), 2.27 (s, 3H), 1.29 - 1.24 (m, 1H); ^{13}C NMR (126 MHz, $CDCl_3$) δ 160.23, 159.41, 143.06, 141.94, 135.1, 134.49, 132.81, 130.34, 129.40, 128.49, 128.32, 128.11, 127.91, 127.62, 127.18, 126.90, 125.65, 91.17, 69.34, 56.21, 55.53, 39.11, 16.23; HRMS (ESI): 对于 $C_{32}H_{33}F_3NO_3$ $[M + H]^+$, 实测 536.2419 m/z; 计算质量, 536.2413。

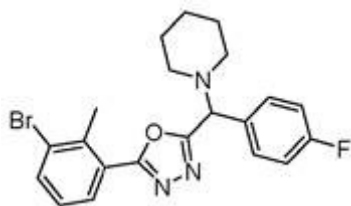
[0293] 化合物4d (GCL.2), (7*R*, 8*R*, 9*S*, 13*S*, 14*S*, 17*R*)-17-乙炔基-17-羟基-7,13-二甲基-1,2,6,7,8,9,10,11,12,13,14,15,16,17-十四氢-3*H*-环戊二烯并[a]菲-3-酮: 将替勃龙 (156 mg, 0.5 mmol) 放入含有 10 mL 的 THF 的圆底烧瓶中并加入 100 μ L 水。接着, 将对甲苯磺酸 (85 mg, 0.5 mmol) 加入它, 并将混合物在 80°C 回流 48 小时, 并通过 TLC 监测反应的进展。然后将有机溶剂蒸发至干燥以得到粗产物, 将其通过快速柱色谱纯化, 使用 20% 的在石油醚中的乙酸乙酯溶剂混合物作为洗脱液, 以产生灰白色固体纯化合物 GCL2 (53% 收率)。 1H NMR (500 MHz, MeOD) δ 5.80 (t, $J = 2.1$, 2.1 Hz, 1H), 2.88 (s, 1H), 2.56 (ddt, $J = 14.1$, 5.4, 1.6, 1.6 Hz, 1H), 2.42 - 2.28 (m, 4H), 2.27 - 2.19 (m, 1H), 2.18 - 2.13 (m, 1H), 2.06 - 1.90 (m, 3H), 1.77 - 1.53 (m, 6H), 1.44 - 1.25 (m, 2H), 1.14 (qd, $J = 11.0$, 11.0, 10.9, 4.2 Hz, 1H), 0.91 (d, $J = 0.7$ Hz, 3H), 0.79 (d, $J = 7.1$ Hz, 3H); ^{13}C NMR (126 MHz, MeOD) δ 201.06, 167.84, 125.32, 87.28, 78.82, 73.45, 48.12, 47.95, 47.78, 47.61, 47.44, 47.27, 47.10, 46.67, 45.82, 43.05, 42.76, 42.18, 38.28, 36.02, 32.25, 30.66, 26.44, 26.41, 21.71, 11.79, 11.77。

[0294] 化合物4e (KPGC01S42), (*S*)-1-(2,6-二甲氧基-4-((2-甲基-[1,1'-联苯]-3-基)甲氧基)苄基)哌啶-3-甲酸: 将来自方案1的3b (24 mg, 0.0066 mmol)、(D)-3-哌啶甲酸 (34.2 mg, 0.0265 mmol, 4当量)、氰基硼氢化钠 (5.2 mg, 0.0343 mmol, 5.2当量) 溶解在 DMF (1 mL) 中, 然后加入乙酸 (1滴)。将反应混合物在室温搅拌 14 小时。通过 TLC (硅胶, 5% DCM:MeOH) 监测反应。将粗制物使用 0-20% DCM:MeOH 通过快速色谱纯化, 以提供作为油状产物的期望产物 (43% 收率)。 1H NMR (500 MHz, DMSO): δ 7.49 - 7.41 (m, 3H), 7.39 - 7.34 (m, 1H), 7.32 - 7.25 (m, 3H), 7.19 (d, $J = 7.7$ Hz, 1H), 6.37 (s, 2H), 5.15 (s, 2H), 3.75 (s, 6H), 3.53 (s, 2H), 2.38 - 2.30 (m, 2H), 2.20 (s, 3H), 1.97 (s, 4H), 1.16 (s, 2H); ^{13}C NMR (126 MHz, DMSO): δ 170.81, 160.50, 160.06,

142.67, 141.87, 135.99, 134.56, 130.18, 129.63, 128.81, 128.71, 127.42, 126.01, 92.00, 69.09, 60.22, 56.27, 55.37, 55.04, 52.65, 49.05, 31.14, 26.68, 21.21, 16.40, 14.54; LCMS/MS (ESI): 对于 $C_{29}H_{33}NO_5$ $[M + H]^+$, 实测 476.3 m/z; 计算质量, 476.24。



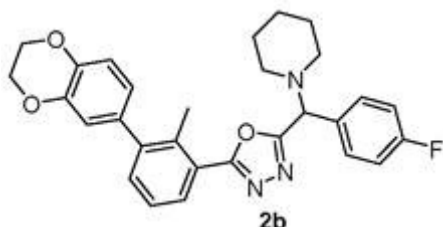
[0296] 2a的合成:



[0297] 步骤-1: 2-(3-溴-2-甲基苯基)-5-((4-氟苯基)(哌啶-1-基)甲基)-1,3,4-噁二唑: 在具有磁力搅拌棒的清洁的干燥的螺旋帽小瓶中, 将哌啶(1当量)、4-氟苯甲醛(1当量)、N-(异氰基亚氨基)三苯基膦(1当量)的混合物溶解在DCM (5 mL/mmol)中。在室温将3-溴-2-甲基-苯甲酸(1当量)在DCM中的溶液缓慢地加入反应混合物中, 并在50-60°C搅拌2小时。在真空中除去溶剂, 并将粗制物通过快速色谱纯化, 使用己烷:乙酸乙酯(0-60%)作为洗脱液, 以得到浅黄色油状产物。

[0298] 步骤-2: 2-((4-氟苯基)(哌啶-1-基)甲基)-5-(2-甲基-[1,1'-联苯]-3-基)-1,3,4-噁二唑(2a): 在具有磁力搅拌棒的清洁的干燥的螺旋帽小瓶中, 放入2-(3-溴-2-甲基苯基)-5-((4-氟苯基)(哌啶-1-基)甲基)-1,3,4-噁二唑(1当量)、苯基硼酸(2当量)和 $PdCl_2(dppf)_2 \cdot CH_2Cl_2$ (3 mol%)的混合物并用氩气净化一次。加入甲苯(4.5 mL)以及乙醇(1.5 mL)并将反应混合物用氩气净化。在净化的同时加入1.5 mL 1M $NaHCO_3$ 并将其在80°C搅拌45分钟。通过TLC监测反应进程。在反应结束后, 加入乙酸乙酯(20 mL)并用水(2 x 20 mL)洗涤。在减压下除去有机溶剂, 并将粗制物通过快速色谱(0-60%, 己烷:乙酸乙酯)纯化, 以得到透明的油状产物。

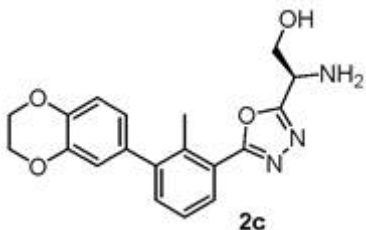
[0299] 2b的合成



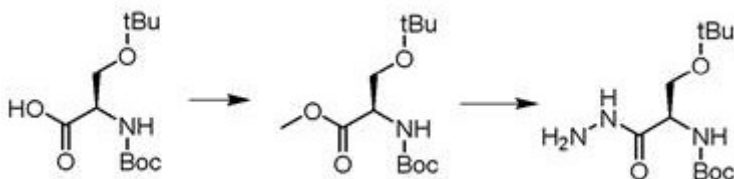
[0300] 在具有磁力搅拌棒的清洁的干燥的螺旋帽小瓶中, 放入2-(3-溴-2-甲基苯基)-5-((4-氟苯基)(哌啶-1-基)甲基)-1,3,4-噁二唑(1当量)、(2,3-二氢苯并[b][1,4]二氧杂环己烯-6-基)硼酸(2当量)和 $PdCl_2(dppf)_2 \cdot CH_2Cl_2$ (3 mol%)的混合物并用氩气净化一次。

加入甲苯 (4.5 mL) 以及乙醇 (1.5 mL) 并将反应混合物用氩气净化。在净化的同时加入 1.5 mL 1M NaHCO₃ 并将其在 80°C 搅拌 45 分钟。通过 TLC 监测反应进程。在反应结束后, 加入乙酸乙酯 (20 mL) 并用水 (2 x 20 mL) 洗涤。在减压下除去有机溶剂, 并将粗制物通过快速色谱 (0-60%, 己烷: 乙酸乙酯) 纯化, 以得到透明的油状产物。

[0301] 2c 的合成



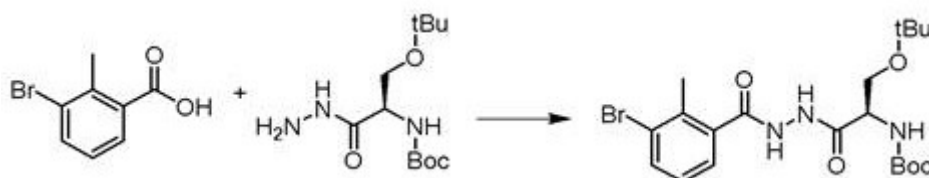
[0302]



[0303] N-(叔丁氧基羰基)-O-(叔丁基)-D-丝氨酸甲酯的制备: 将 Boc-D-丝氨酸 (1 当量)、碘甲烷 (3 当量) 在 DMF 中的混合物在室温搅拌 2 小时。将冰冷的水加入反应混合物中, 并用乙酸乙酯萃取。然后, 将有机层用水洗涤, 并将乙酸乙酯在真空中除去。定量获得油状粗产物, 并将它不经进一步纯化用于下一步。

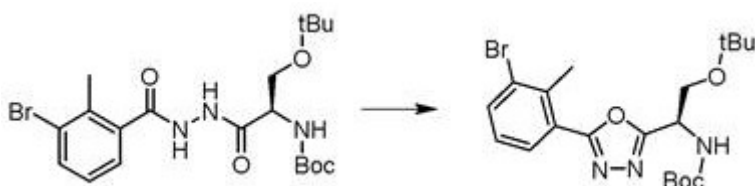
[0304] (R)-(3-(叔丁氧基)-1-胍基-1-氧代丙烷-2-基)氨基甲酸叔丁酯的制备: 将胍一水合物 (2 当量) 缓慢地加入保持在冰浴中的 Boc-D-丝氨酸的甲酯 (1 当量) 在 DCM 中的混合物中。将它在室温搅拌 2 小时, 直到观察到白色悬浮液。将冰冷的水加入反应混合物中, 并用乙酸乙酯萃取。然后, 将有机层用水洗涤并将乙酸乙酯在真空中除去。定量获得白色固体产物, 并将它不经进一步纯化用于下一步。

[0305]

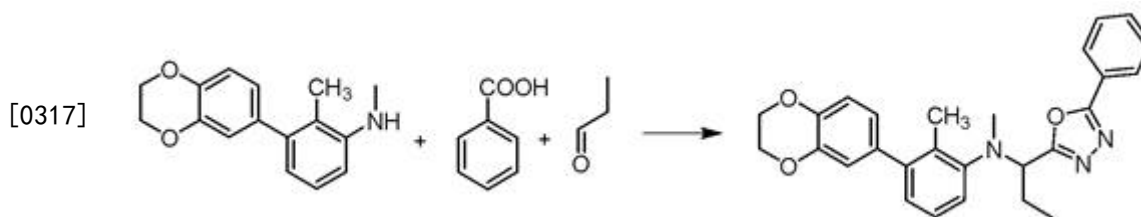


[0306] 将 3-溴-2-甲基苯甲酸 (1 当量)、EDCI (1.5 当量)、DIPEA (5 当量) 在 DMF 中的混合物搅拌 10 分钟, 然后将 (R)-(3-(叔丁氧基)-1-胍基-1-氧代丙烷-2-基)氨基甲酸叔丁酯 (1 当量) 加入反应混合物中。将它在室温搅拌过夜。通过 TLC 监测反应。将粗制物用乙酸乙酯: 冰冷的水萃取, 并将乙酸乙酯层在减压下除去。将产物使用 0-60% 己烷: 乙酸乙酯通过快速柱色谱纯化, 以得到白色固体 (65% 收率)。

[0307]

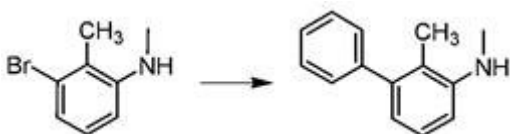


[0308] 将 (R)-(1-(2-(3-溴-2-甲基苯甲酰基)胍基)-3-(叔丁氧基)-1-氧代丙烷-2-基)

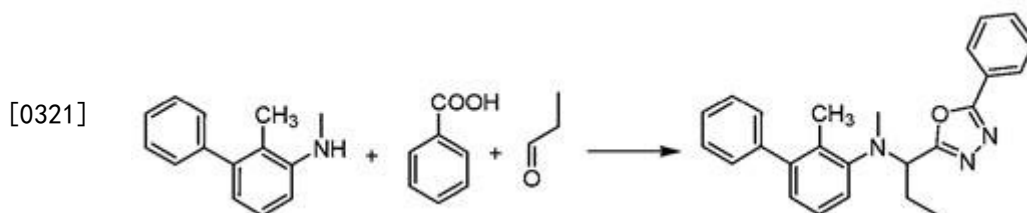


[0318] 步骤-2: 向3-(2,3-二氢苯并[b][1,4]二氧杂环己烯-6-基)-N,2-二甲基苯胺(1当量)、丙醛(1.1当量)、(N-异氰基亚氨基)三苯基膦在DCM中的混合物逐份加入苯甲酸(1.1当量)。将反应混合物在40℃搅拌2-3小时。通过TLC监测反应。将产物使用0-60% 己烷:乙酸乙酯通过快速柱色谱纯化以得到油状产物(54%收率)。

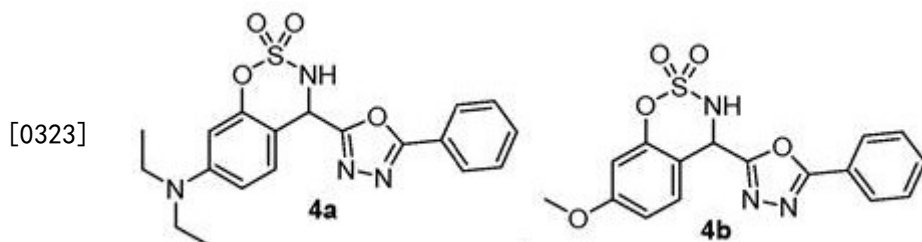
[0319] (IIIb)的合成

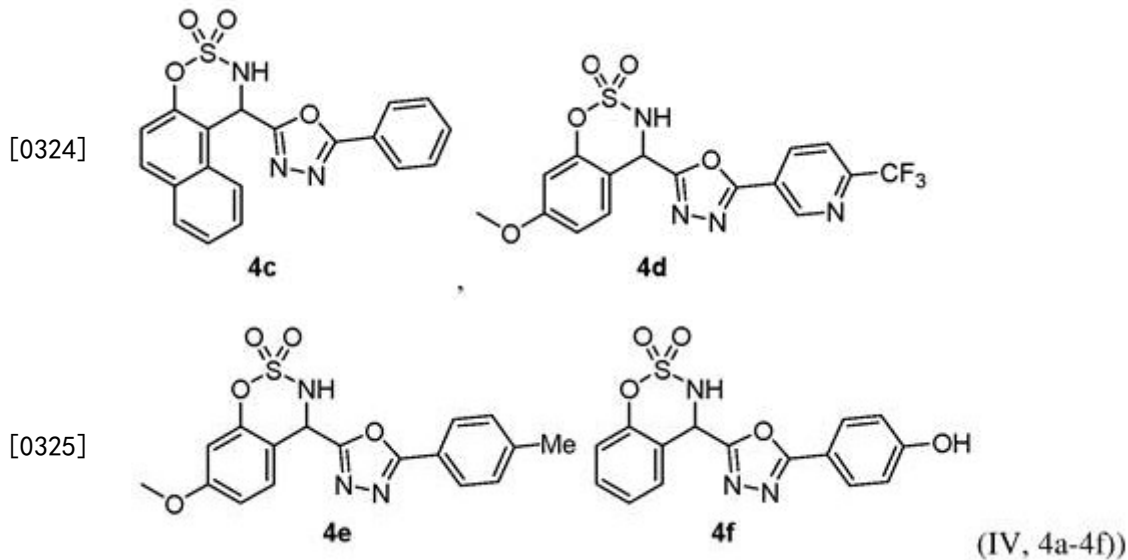


[0320] 步骤-1: 将在甲苯:乙醇(1.5:0.5 mL)中的3-溴-N,2-二甲基苯胺(1当量)、苯基硼酸(1.5当量)、PdCl₂(dppf)·CH₂Cl₂(3 mol%)用氩气净化2次。在惰性气氛下加入1M NaHCO₃(1.5 mL)并将它在80℃搅拌45分钟。通过TLC监测反应。将产物使用0-60% 己烷:乙酸乙酯通过快速柱色谱纯化以得到油状产物(74%收率)。

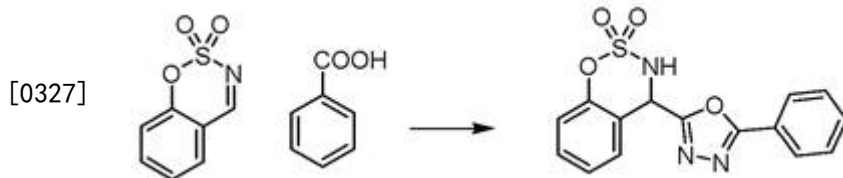


[0322] 步骤-2: 向N,2-二甲基-[1,1'-联苯]-3-胺(1当量)、丙醛(1.1当量)、(N-异氰基亚氨基)三苯基膦在DCM中的混合物逐份加入苯甲酸(1.1当量)。将反应混合物在40℃搅拌2-3小时。通过TLC监测反应。将产物使用0-60% 己烷:乙酸乙酯通过快速柱色谱纯化以得到油状产物(48%收率)。



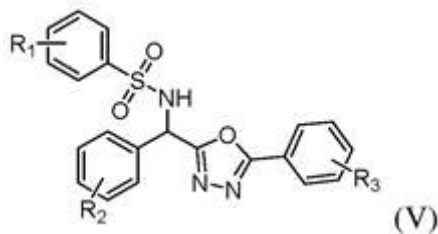


[0326] 代表性的程序：4-(5-苯基-1,3,4-噁二唑-2-基)-3,4-二氢苯并[e][1,2,3]噁嗪2,2-二氧化物[式IV衍生物]：



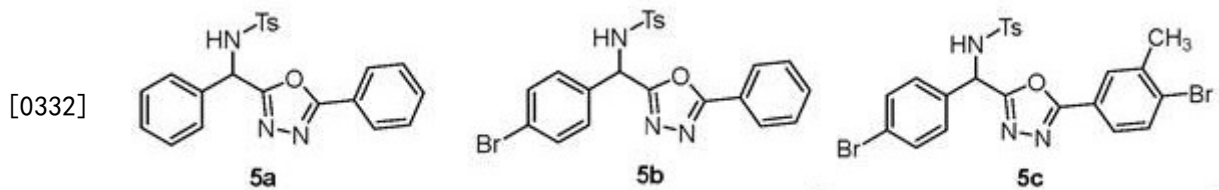
[0328] 在清洁的烘干的螺旋帽小瓶中,放入磺酰基亚胺1a (0.2 mmol)、苯甲酸(1.1当量)、(N-异氰基亚氨基)三苯基膦(1.1当量)并冷却至 -10°C 。接着,加入溶剂 CH_2Cl_2 (3 mL)并在相同温度搅拌5-10分钟。在减压下部分地除去溶剂,并将粗制物直接加载到硅胶筒上并使用EtOAc:己烷作为洗脱液通过快速色谱(Teledyne ISCO)纯化。收集主要级分,并用水:MeCN作为洗脱液通过反相制备型HPLC (Teledyne ISCO)再纯化。

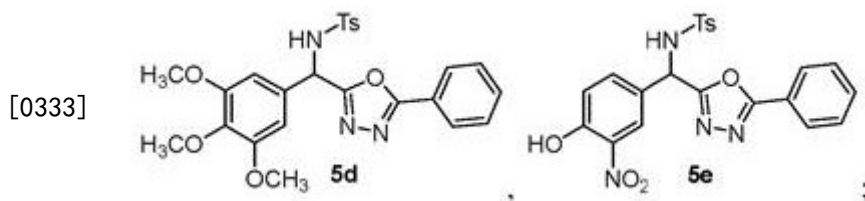
[0329] 具有式V的化合物：



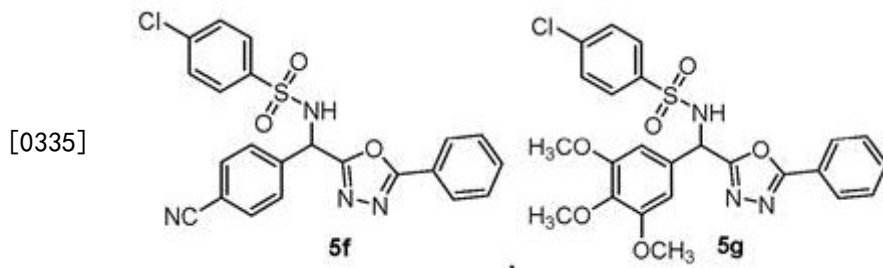
[0330] 根据式(V)的化合物,其中：

[0331] 当 $\text{R}_1 = \text{CH}_3$ 时,所述化合物为





[0334] 当 $R_1 = Cl$ 时,所述化合物为



[0336] 用于合成4-甲基-N-(苯基(5-苯基-1,3,4-噁二唑-2-基)甲基)苯磺酰胺[式V衍生物]的代表性程序:在清洁的烘干的螺旋帽小瓶中,放入N-甲苯磺酰基亚胺(0.2 mmol)、苯甲酸(1.1当量)、(N-异氰基亚氨基)三苯基膦(1.1当量)并冷却至 $-10^{\circ}C$ 。接着,加入溶剂 CH_2Cl_2 (3 mL)并在相同温度搅拌5-10分钟。在减压下部分地除去溶剂,并将粗制物直接加载到硅胶筒上,并使用EtOAc:己烷作为洗脱液通过快速色谱(Teledyne ISCO)纯化以得到固体产物。

[0337] 本领域技术人员会认识到,可以对上述具体实施内容做出众多修改。所述实施内容不应限于所描述的特定限制。其他实施内容也是可能的。

[0338] 尽管已在附图和前述描述中详细说明和描述了本发明,但其应当在性质上视作示例性的而不是限制性的,应当理解,仅显示和描述了某些实施方案,并且落入本发明的精神内的所有变化和修改都期望得到保护。本发明的方法和设备的范围意图由以下权利要求限定。但是,必须理解,在不背离其精神或范围的情况下,可以以不同于具体解释和说明的方式实施本公开内容。

[0339] 参考文献

[0340] (1) Pardoll, D. M. The Blockade of Immune Checkpoints in Cancer Immunotherapy. *Nat. Rev. Cancer* 2012, 12 (4), 252-264.

[0341] (2) Buchbinder, E. I.; Desai, A. CTLA-4 and PD-1 Pathways Similarities, Differences, and Implications of Their Inhibition. *American Journal of Clinical Oncology: Cancer Clinical Trials*. Wolters Kluwer Health 2016, 第98-106页。

[0342] (3) Sasikumar, P. G. N.; Ramachandra, M.; Naremaddepalli, S. S. S. 1, 2, 4-Oxadiazole Derivatives as Immunomodulators. Google Patents 2017年9月。

[0343] (4) Sunshine, J.; Taube, J. M. Pd-1/Pd-L1 Inhibitors. *Curr. Opin. Pharmacol.* 2015, 23, 32-38.

[0344] (5) Imai, K.; Takaoka, A. Comparing Antibody and Small-Molecule Therapies for Cancer. *Nat. Rev. Cancer* 2006, 6 (9), 714-727.

[0345] (6) Chupak, L.; Ding, M.; Martin, S.; Connoly, T.; Xu, N.; Yeung, K.;

Zhu, J.; Langley, D.; 等人. Compounds Useful as Immunomodulators. WO/2015/160641, 2015年10月22日。

[0346] (7) Chupak, L. S.; Zheng, X.; Mingo, P. A. Compounds Useful as Immunomodulators. WO/2015/034820, 2015年12月3日。

[0347] (8) Guzik, K.; Zak, K. M.; Grudnik, P.; Magiera, K.; Musielak, B.; Tö rner, R.; Skalniak, L.; Dö mling, A.; Dubin, G.; Holak, T. A. Small-Molecule Inhibitors of the Programmed Cell Death-1/Programmed Death-Ligand 1 (PD-1/PD-L1) Interaction via Transiently Induced Protein States and Dimerization of PD-L1. *J. Med. Chem.* 2017, 60 (13), 5857-5867.

[0348] (9) Zak, K. M.; Grudnik, P.; Guzik, K.; Zieba, B. J.; Musielak, B.; D ö mling, A.; Dubin, G.; Holak, T. A. Structural Basis for Small Molecule Targeting of the Programmed Death Ligand 1 (PD-L1). *Oncotarget* 2016, 7 (21), 30323-30335.

[0349] (10) Broach, J. R.; Thorner, J. High-Throughput Screening for Drug Discovery. *Nature*. [伦敦: Macmillan Journals], 1869- 1996,第14-16页。

[0350] (11) Artis, D. R.; Lin, J. J.; 等人. Scaffold-Based Discovery of Indeglitazar, a PPAR Pan-Active Anti-Diabetic Agent. *Proc. Natl. Acad. Sci. U. S. A.* 2009, 106 (1), 262-267.

[0351] (12) Böhm, H.-J.; Flohr, A.; Stahl, M. Scaffold Hopping. *Drug Discov. today Technol.* 2004, 1 (3), 217-224.

[0352] (13) Kitchen, D. B.; Decornez, H.; Furr, J. R.; Bajorath, J. Docking and Scoring in Virtual Screening for Drug Discovery: Methods and Applications. *Nat. Rev. Drug Discov.* 2004, 3 (11), 935-949.

[0353] (14) Joachims, T. Text Categorization with Support Vector Machines: Learning with Many Relevant Features. *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* 1998, 1398, 137-142.

[0354] (15) Burbidge, R.; Trotter, M.; Holden, S. Drug Design by Machine Learning: Support Vector Machines for Pharmaceutical Data Analysis. *Comput. Chem.* 2001, 26 (1), 5-14.

[0355] (16) Heikamp, K.; Bajorath, J. Support Vector Machines for Drug Discovery. *Expert Opin. Drug Discov.* 2014.

[0356] (17) Breiman, L. Random Forests. *Mach. Learn.* 2001, 9 (1), 5-32.

[0357] (18) Li, H.; Leung, K. S.; Wong, M. H.; Ballester, P. J. Improving Autodock Vina Using Random Forest: The Growing Accuracy of Binding Affinity Prediction by the Effective Exploitation of Larger Data Sets. *Mol. Inform.* 2015, 34 (2-3), 115-126.

[0358] (19) Zhang, Q. Y.; Aires-de-Sousa, J. Random Forest Prediction of Mutagenicity from Empirical Physicochemical Descriptors. *J. Chem. Inf. Model.*

2007, 47 (1), 1-8。

[0359] (20) Torng, W.; Altman, R. B. Graph Convolutional Neural Networks for Predicting Drug-Target Interactions. *J. Chem. Inf. Model.* 2019。

[0360] (21) Scarselli, F.; Gori, M.; Tsoi, A. C.; Hagenbuchner, M.; Monfardini, G. The Graph Neural Network Model. *IEEE Trans. Neural Networks* 2009, 20 (1), 61-80。

[0361] (22) Lim, J.; Ryu, S.; Park, K.; Choe, Y. J.; Ham, J.; Kim, W. Y. Predicting Drug-Target Interaction Using a Novel Graph Neural Network with 3D Structure-Embedded Graph Representation. *J. Chem. Inf. Model.* 2019, 59 (9), 3981-3988。

[0362] (23) Ma, T.; Zhang, A. *AffinityNet: Semi-Supervised Few-Shot Learning for Disease Type Prediction.*

[0363] (24) Tsubaki, M.; Tomii, K.; Sese, J. Compound-Protein Interaction Prediction with End-to-End Learning of Neural Networks for Graphs and Sequences. *Bioinformatics* 2019, 35 (2), 309-318。

[0364] (25) Gupta, A.; Müller, A. T.; Huisman, B. J. H.; Fuchs, J. A.; Schneider, P.; Schneider, G. Generative Recurrent Networks for De Novo Drug Design. *Mol. Inform.* 2018, 37 (1)。

[0365] (26) Ståhl, N.; Falkman, G.; Boström, J. Deep Reinforcement Learning for Multiparameter Optimization in de Novo Drug Design. *J. Chem. Inf. Model.* 2019。

[0366] (27) Li, Y.; Zhang, L.; Liu, Z. Multi-Objective de Novo Drug Design with Conditional Graph Generative Model. *J. Cheminform.* 2018, 10 (1)。

[0367] (28) Esaki, T.; Ohashi, R.; Watanabe, R.; Natsume-Kitatani, Y.; Kawashima, H.; Nagao, C.; Mizuguchi, K. Computational Model to Predict the Fraction of Unbound Drug in the Brain. *J. Chem. Inf. Model.* 2019。

[0368] (29) Fine, J.; Chopra, G. Lemon: A Framework for Rapidly Mining Structural Information from the Protein Data Bank. *Bioinformatics* 2019。

[0369] (30) Chopra, G.; Samudrala, R. Exploring Polypharmacology in Drug Discovery and Repurposing Using the CANDO Platform. *Curr. Pharm. Des.* 2016, 22 (21), 3109-3123。

[0370] (31) Fine, J.; Konc, J.; Samudrala, R.; Chopra, G. CANDOCK: Chemical Atomic Network Based Hierarchical Flexible Docking Algorithm. *bioRxiv* 2018。

[0371] (32) Kischuk, E.; Majumder, J.; Fine, J. A.; Lantz, T. C.; Dhawan, D.; Knapp, D. W.; Ratliff, T. L.; Chopra, G. Abstract 4709: Cell-Specific Gene Program-Based Small-Molecule Immunomodulators Targeting Solid-Tumor Microenvironments; 2018。

[0372] (33) Esensten, J. H.; Helou, Y. A.; Chopra, G.; Weiss, A.; Bluestone, J. A. CD28 Costimulation: From Mechanism to Therapy. *Immunity.* 2016,第973-988

页。

[0373] (34) Hernandez-Perez, M.; Chopra, G.; Fine, J.; Conteh, A. M.; 和erson, R. M.; Linnemann, A. K.; Benjamin, C.; Nelson, J. B.; Benninger, K. S.; Nadler, J. L.; 等人. Inhibition of 12/15-Lipoxygenase Protects Against β -Cell Oxidative Stress and Glycemic Deterioration in Mouse Models of Type 1 Diabetes. *Diabetes* 2017, 66 (11), 2875-2887.

[0374] (35) Ma, X.; Zhou, J.; Wang, C.; Carter-Cooper, B.; Yang, F.; Larocque, E.; Fine, J.; Tsuji, G.; Chopra, G.; Lapidus, R. G.; 等人. Identification of New FLT3 Inhibitors Potently Inhibit AML Cell Lines via an Azo Click-It/Staple-It Approach. *ACS Med. Chem. Lett.* 2017, 8 (5).

[0375] (36) Fine, J.; Lackner, R.; Samudrala, R.; Chopra, G. Computational Chemoproteomics to Understand the Role of Selected Psychoactives. *Sci. Rep.* 2019.

[0376] (37) Lu, L.; Zhang, F.; Li, J.; Wang, H.; Xiao, K.; Wu, L.; Qian, D.-Q.; Lajkiewicz, N.; Konkol, L.; Li, Z.; 等人. Heterocyclic Compounds Derivatives as PD-L1 Internalization Inducers. WO/2018/119263, 2018年6月28日。

[0377] (38) Xiao, K.; Zhang, F.; Wu, L.; Yao, W. Heterocyclic Compounds as Immunomodulators. US 2018/0273519 A1, 2018年9月27日。

[0378] (39) Willett, P. Similarity-Based Virtual Screening Using 2D Fingerprints. *Drug Discov. Today* 2006, 11 (23-24), 1046-1053.

[0379] (40) DeLano, W. L. *Pymol: An Open-Source Molecular Graphics Tool*; CCP4, 2002;第40卷。

[0380] (41) Max, K.; Weston, S.; Keefer, C.; Engelhardt, A.; Cooper, T.; Mayer, Z.; Kenkel, B.; Team, R. C.; Benesty, M.; Lescarbeau, R.; 等人. Classification and Regression Training. 2016,第198页。

[0381] (42) Robin, X.; Turck, N.; Hainard, A.; Tiberti, N.; Lisacek, F.; Sanchez, J. C.; Müller, M. PROC: An Open-Source Package for R and S+ to Analyze and Compare ROC Curves. *BMC Bioinformatics* 2011, 12 (1), 77.

[0382] (43) Guzik, K.; Tomala, M.; Muszak, D.; Konieczny, M.; Hec, A.; Błaszczewicz, U.; Pustuła, M.; Butera, R.; Dömling, A.; Holak, T. A. Development of the Inhibitors That Target the PD-1/PD-L1 Interaction—A Brief Look at Progress on Small Molecules, Peptides and Macrocycles. *Molecules* 2019, 24 (11), 2071.

[0383] (44) Zak, K. M.; Kitel, R.; Przetocka, S.; Golik, P.; Guzik, K.; Musielak, B.; Dömling, A.; Dubin, G.; Holak, T. A. Structure of the Complex of Human Programmed Death 1, PD-1, and Its Ligand PD-L1. *Structure* 2015.

[0384] (45) Lee, H. T.; Lee, S. H.; Heo, Y. S. Molecular Interactions of Antibody Drugs Targeting PD-1, PD-L1, and CTLA-4 in Immuno-Oncology. *Molecules*. 2019.

- [0385] (46) Zak, K. M.; Grudnik, P.; Magiera, K.; Holak, T. A. Structural Biology of the Immune Checkpoint Receptor PD-1 and Its Ligands PD-L1/PD-L2. *Structure*. 2017.
- [0386] (47) PyTorch Community. Tensors and Dynamic Neural Networks in Python with Strong GPU Acceleration. *Github* 2016, 6.
- [0387] (48) Chemaxon, M. Molecule File Converter, V. 5.10. 1, (C) 1999–2012 ChemAxon Ltd.
- [0388] (49) Landrum, G. A. RDKit: Open-Source Cheminformatics.
- [0389] (50) Costa, F.; De Grave, K. Fast Neighborhood Subgraph Pairwise Distance Kernel. In *ICML 2010 - Proceedings, 27th International Conference on Machine Learning*; Omnipress; Madison, WI, USA, 2010; pp 255-262.
- [0390] (51) Csizmadia, P. MarvinSketch and MarvinView: Molecule Applets for the World Wide Web. In *Proceedings of ECSOC-3, the third international electronic conference on synthetic organic chemistry*, 9月1a30; 2019;第1775页。

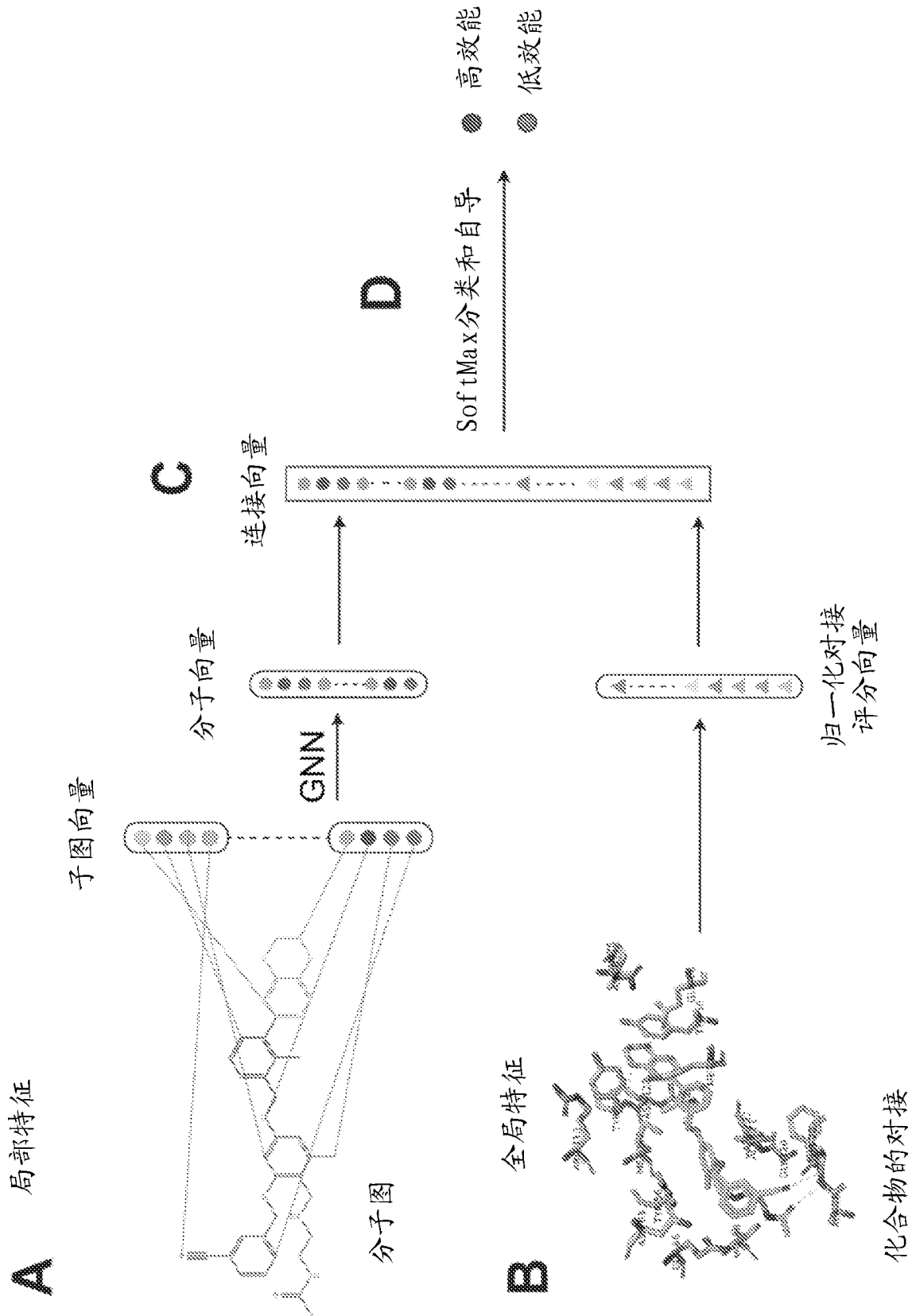
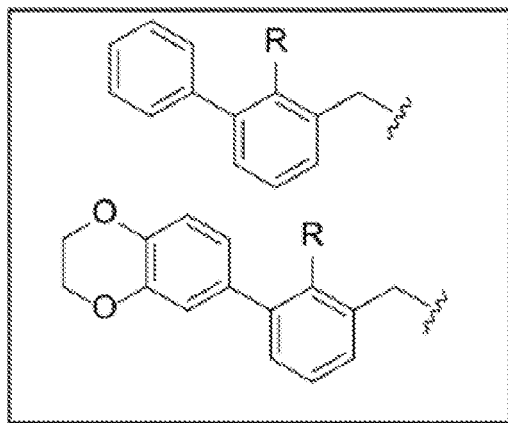
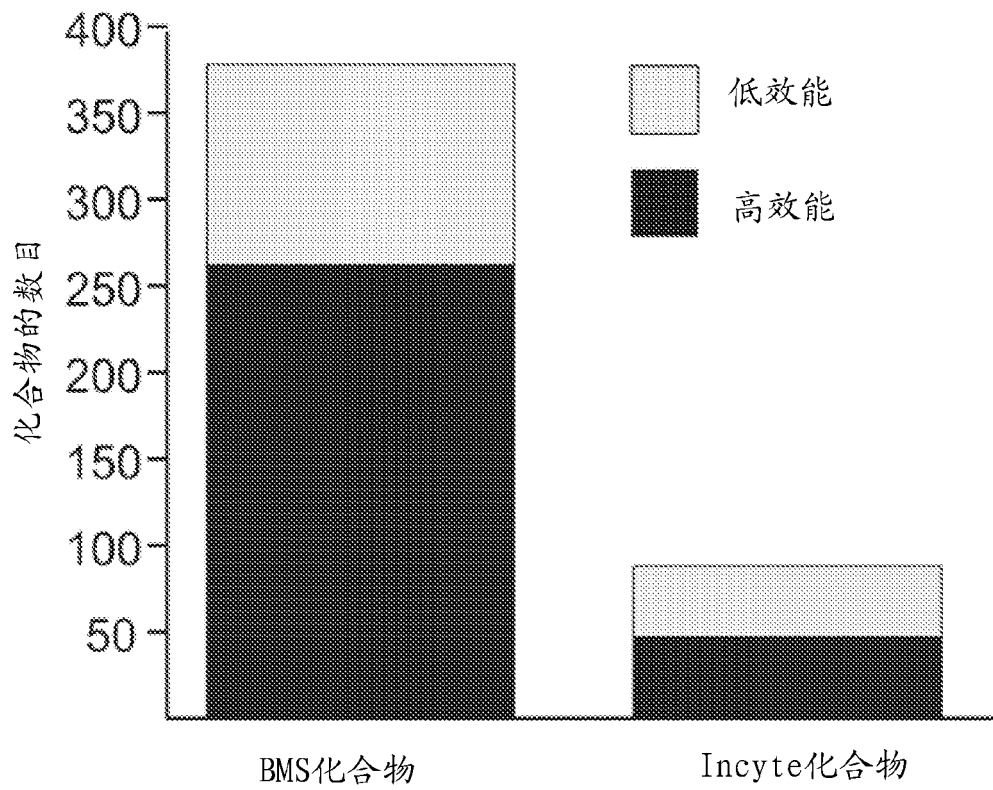
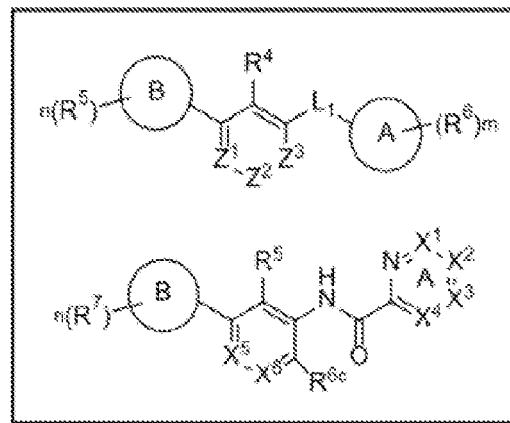


图 1

训练数据的分类



BMS主要骨架



Incyte主要骨架

图 2A

BMS分子的Tanimoto相似性

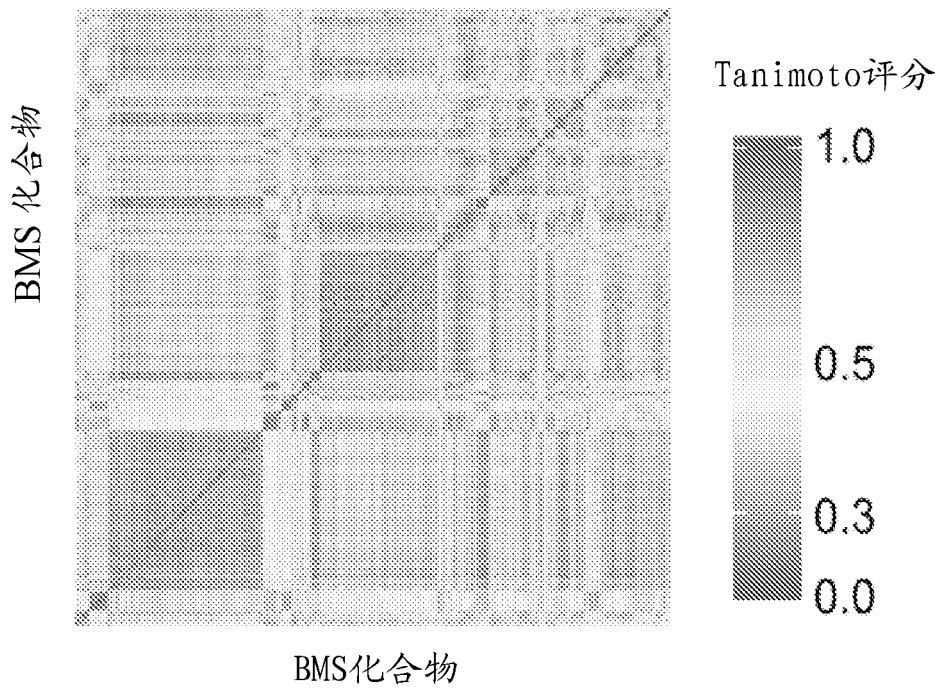


图 2B

Incyte分子的Tanimoto相似性

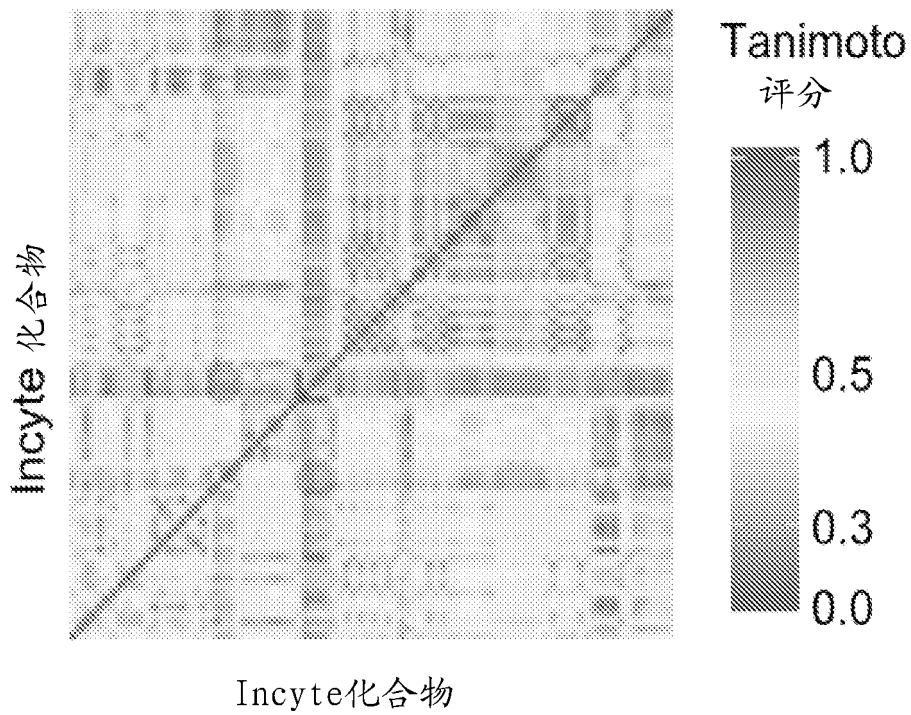


图 2C

对齐的PD-L1同源二聚体和PD-1/PD-L1晶体结构

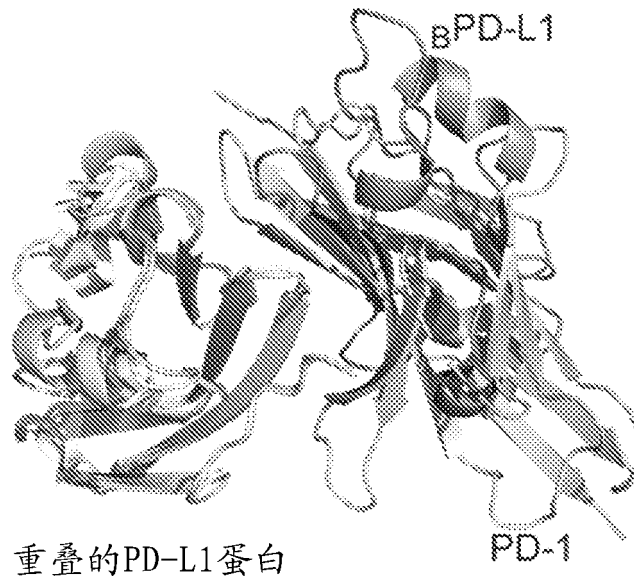


图 3A

PD-L1同源二聚体结合位点
与PD-1/PD-L1相互作用重叠

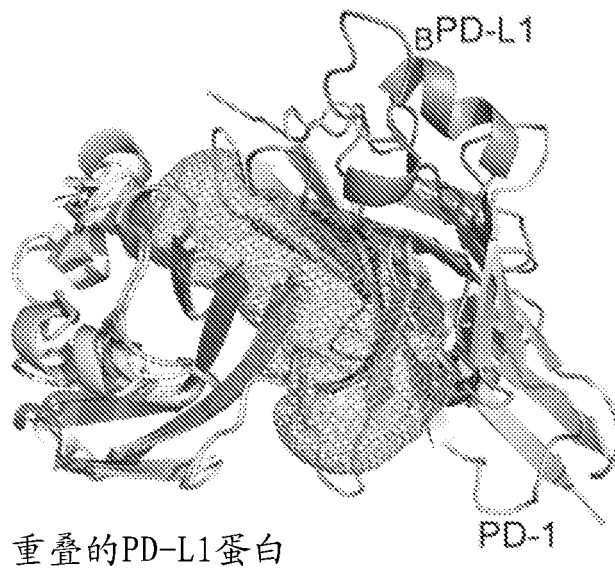


图 3B

PD-L1同源二聚体与其配体8HW

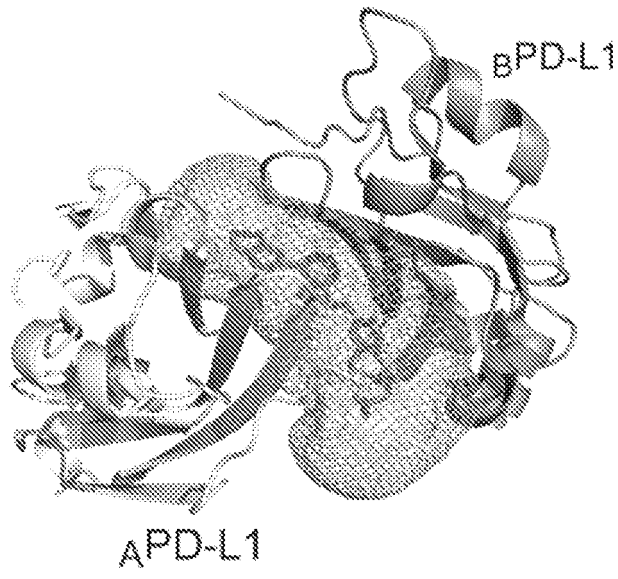


图 3C

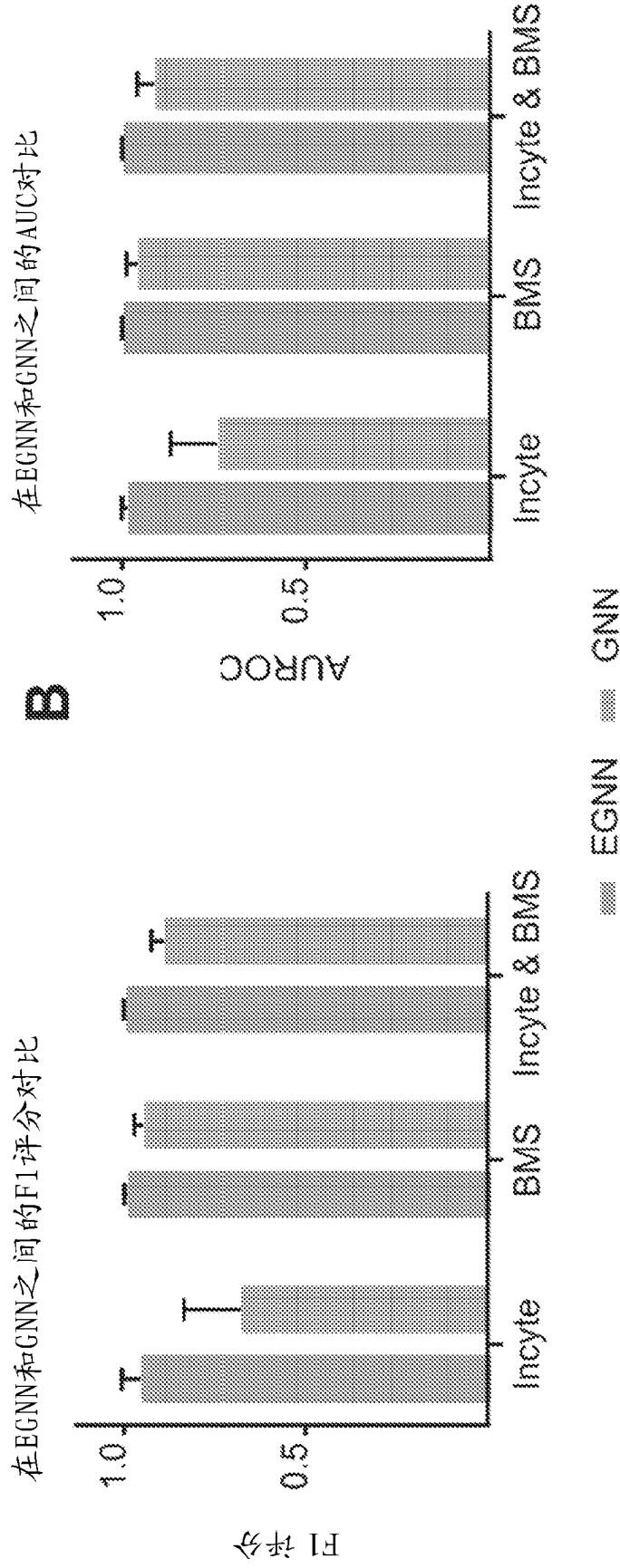


图 4B

图 4A

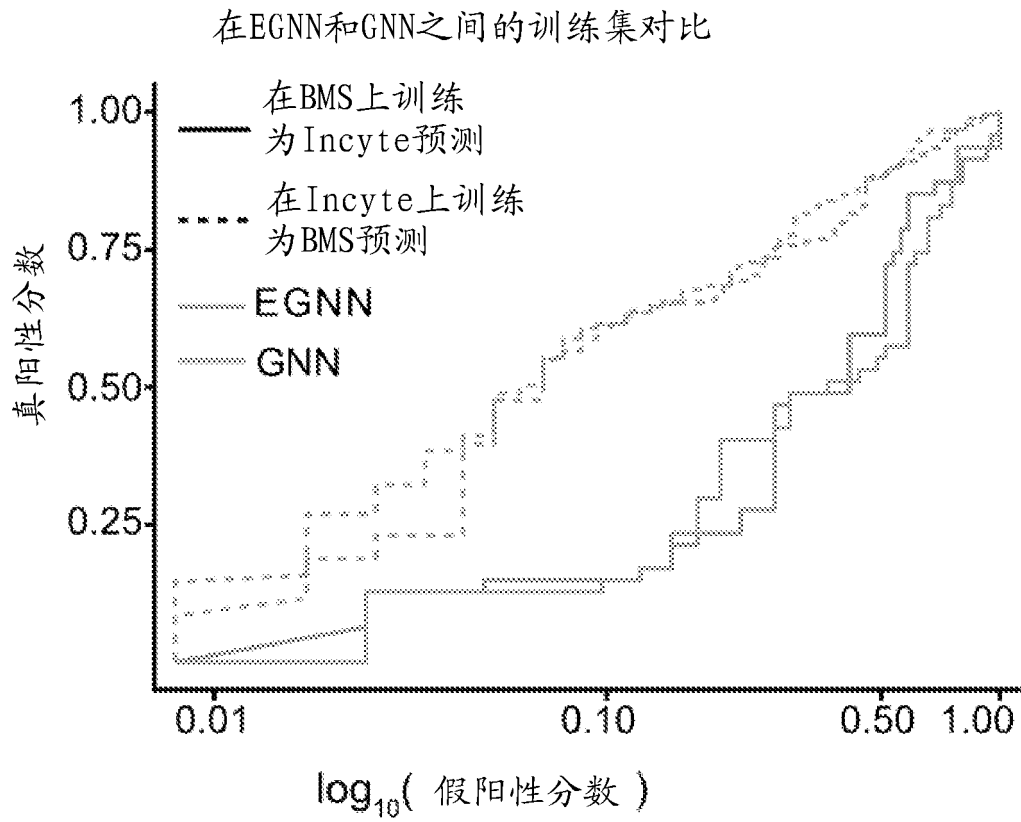


图 4C

EGNN和基线模型的ROC曲线

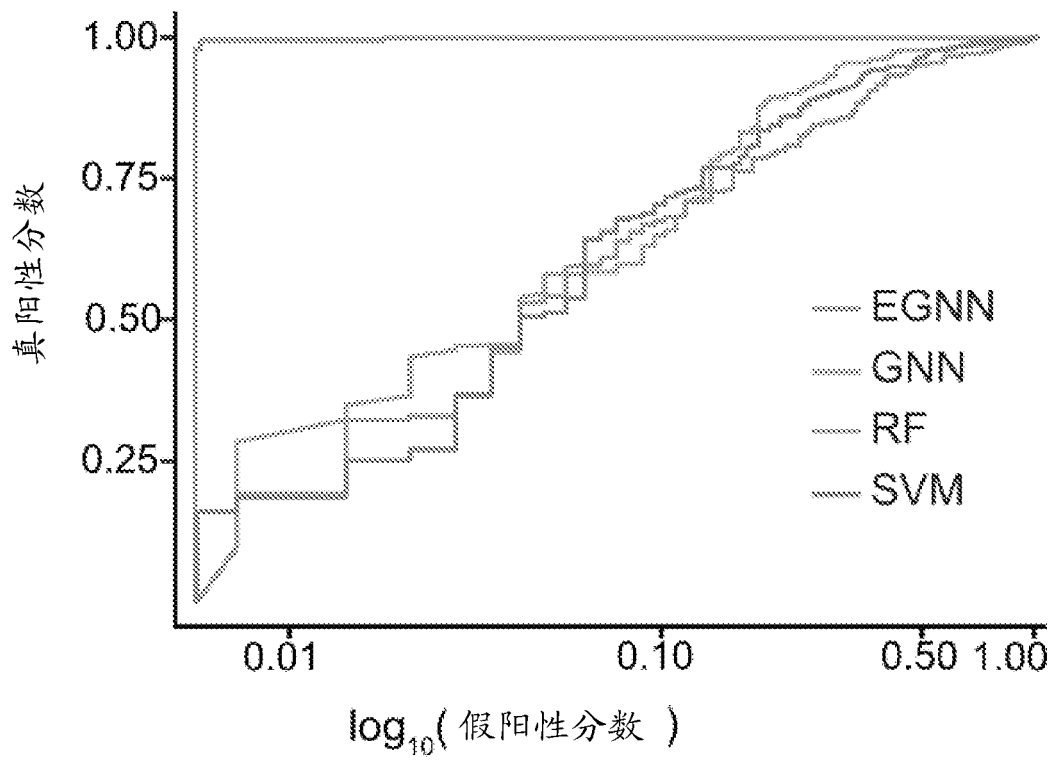


图 4D

来自BMS-1002和BMS-1的4b的合理设计

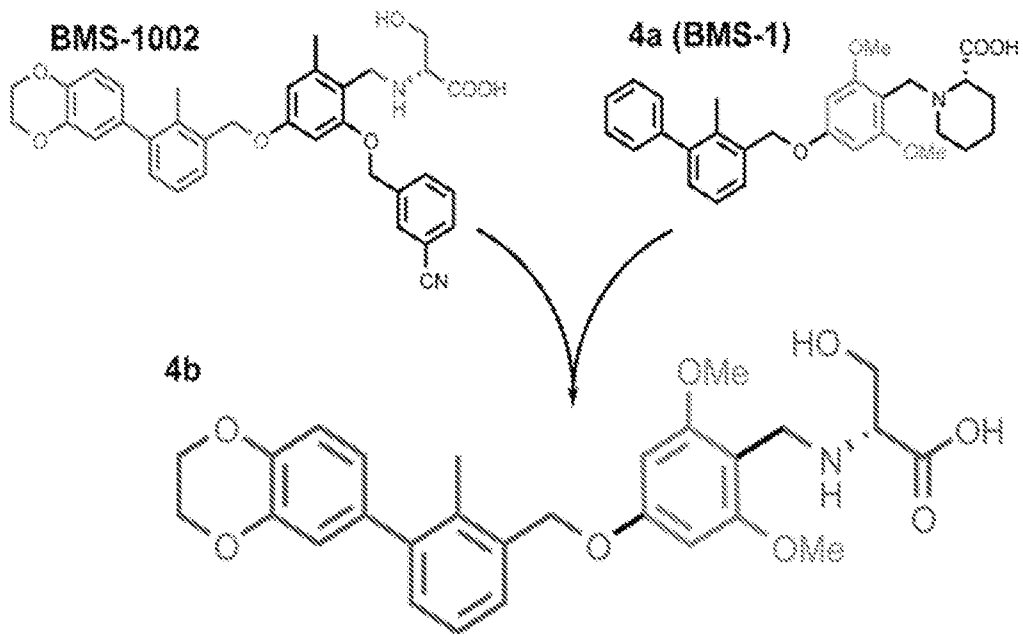


图 5A

PD-L1同源二聚体与4b的结合隧道

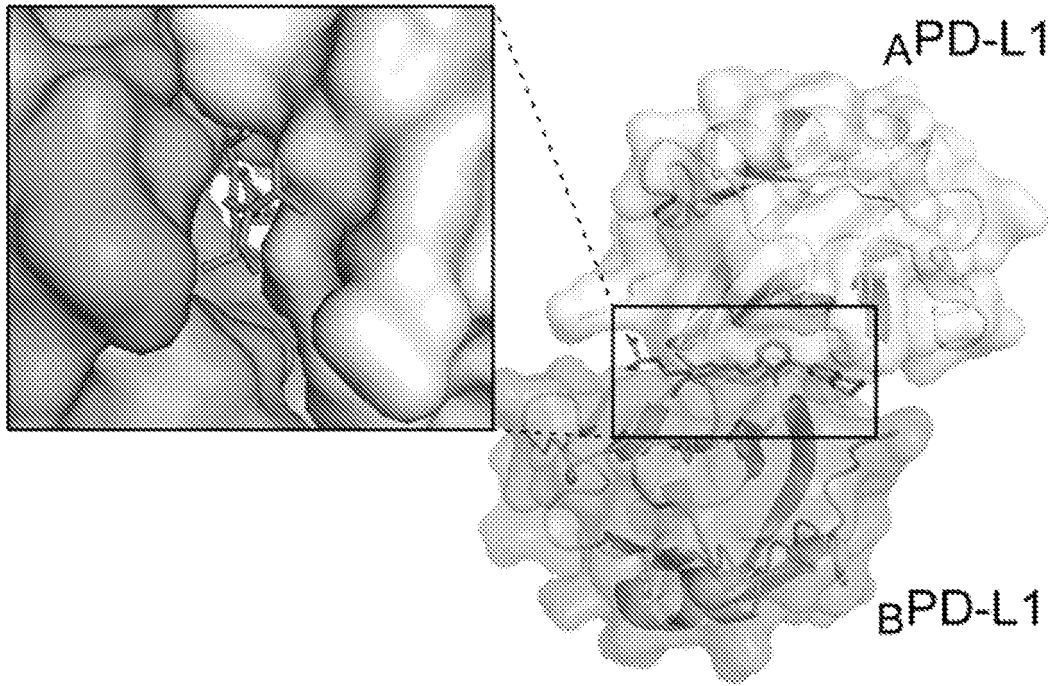


图 5B

4b在PD-L1同源二聚体袋中的对接位姿

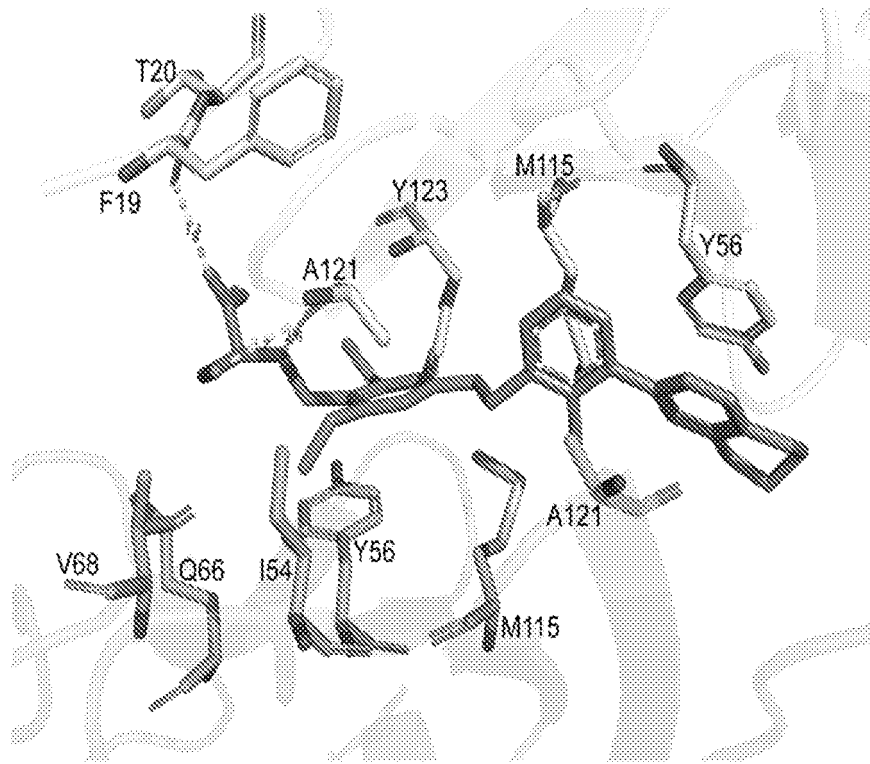


图 5C

4a和4b的PD-1/PD-L1抑制

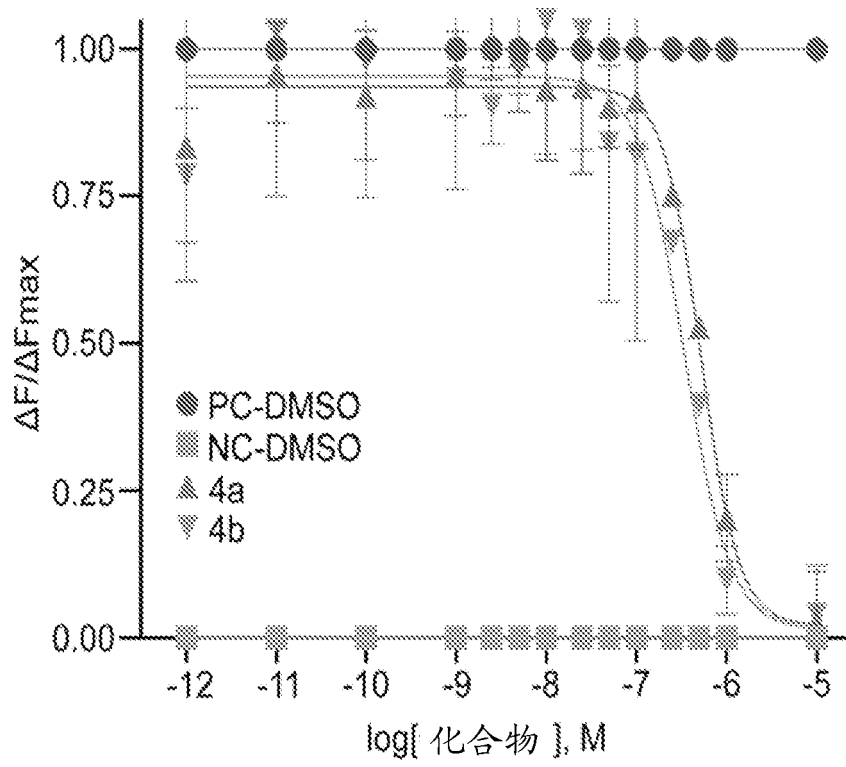


图 5D