新药创制是复杂的智力活动, 涉及科学研究、技术创造、产品开发和医疗效果等多维科技活动。每个药物都有自身的研发轨迹, 而构建化学结构是最重要的环节, 因为它涵盖了药效、药代、安全性和生物药剂学等性质。本栏目以药物化学视角, 对有代表性的药物的成功构建, 加以剖析和解读。

治疗非小细胞肺癌的第三代靶向药物奥希替尼于2009年启动研究, 2015年11月在美国首先上市, 2017年3月CFDA批准在我国上市, 研发和应用都可谓之快速。这是一个有代表性的科技驱动研发的项目, 作为第三代药物虽然有借鉴和跟随的成分, 但饱含了创新之举。由于针对T790M突变的EGFR靶标, 要避开对结构类似的野生型EGFR和胰岛素生长因子受体的作用, 因而是个精准设计和多维研究的工作。从药物化学的视角, 在短时间内反复考量结构细节对活性、选择性、物化性质、脱靶性等的影响, 确定了候选物进入开发阶段, 彰显了高效率。在技术层面上, 结构优化中融汇了多种方法, 如分子模拟、QSAR、结构生物学和动力学同位素效应等。精准医学须要有精准设计与制造的药物支撑, 奥希替尼的研制体现了精准的内涵。 (编者按)

DOI: 10.16438/j.0513-4870.2025-0575

精准设计的第三代肺癌靶向药物奥希替尼

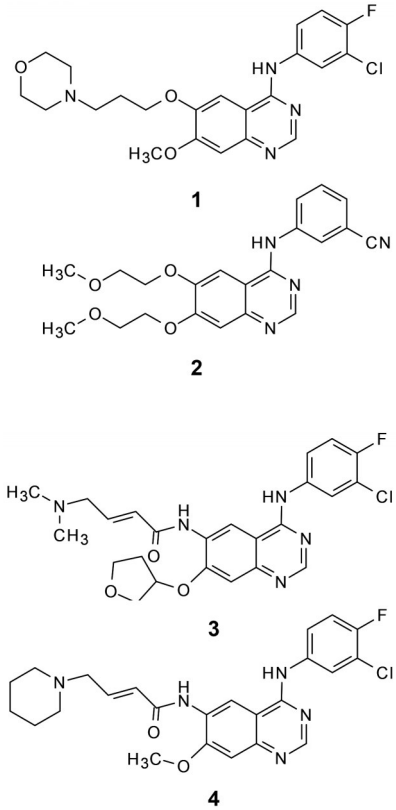
郭宗儒

(中国医学科学院、北京协和医学院药物研究所, 北京 100050)

1 研发背景

1.1 第一代EGFR激酶抑制剂对双位点变异的激酶无效 非小细胞肺癌 (NSCLC) 大约有25%患者的表皮生长因子受体 (EGFR) 酪氨酸激酶发生了变异, 位点是Leu858突变为Arg (L858R), 导致癌细胞对这个信号通路的依赖而发展。对此单变异型激酶, 第一代EGFR酪氨酸抑制剂例如吉非替尼 (**1**, gefitinib) 和厄洛替尼 (**2**, erlotinib) 是敏感的, 成为治疗NSCLC一线治疗药。但用药1年左右有60%患者出现耐药, 是由于激酶的ATP结合位点的门户残基 (gatekeeper) Thr790突变为Met (T790M), **1**和**2**降低了对双位点突变激酶的抑制活性, 治疗效果变差。由于**1**和**2**对野生型的EGFR激酶也有抑制作用, 若加大剂量会因脱靶作用引起不良反应, 例如皮疹和腹泻等, 因而限制了治疗窗口。

1.2 第二代不可逆抑制剂的脱靶作用 第二代“替尼”有一批药物, 特点是不可逆抑制剂, 例如阿法替尼 (**3**, afatinib) 和达可替尼 (**4**, dacomitinib), 结构中含有弱亲电性基团, 与激酶向溶剂的开口处的残基Cys797发生迈克尔加成, 共价键结合导致不可逆抑制。**3**和**4**提高了抑制双突变激酶的活性, 因而对第一代耐药的患者有效。然而由于包括野生型在内的所有EGFR激酶开口处都有Cys797, 降低了选择性作用, 因而呈现不良反应。



1.3 研发目标 为克服两代“替尼”的不足, AstraZeneca公司设定的目标是: ① 对L858R和T790M双突变型

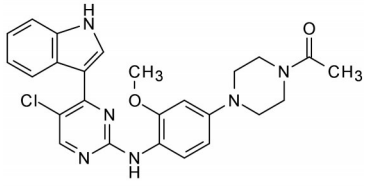
(DM) 的EGFR激酶选择性抑制; ② 不抑制野生型 (WT) 激酶; ③ 不抑制胰岛素生长因子1受体 (IGF1R); ④ 口服吸收。为此, 必须针对突变后生成的Met790的结构特征, 使化合物强效结合ATP结构域, 以超过因T790M变异提高了的与ATP的结合力, 因为后者是不利因素。所以, 目标的本质是选择性: 对双突变激酶高活性, 对野生型激酶和IGF1R无或低活性, 为此, 须精准设计与评价。

2　发现先导化合物

2.1　活性评价　用野生型和L858R/T790M双突变的EGFR激酶作初筛, 再用3种表达不同EGFR激酶的细胞模型评价受试物活性, 即表达野生型激酶的细胞人LoVo细胞 (WT); 缺失外显子19的激活突变的PC9细胞 (AM); 双突变的NCI-H1975细胞 (DM), 以测定磷酸化程度判断终点, IC_{50}表示活性强度。此外, 计算受试物 IC_{50WT}/IC_{50DM} 值以表征选择性, 比值越大选择性越高。

2.2　化合物库的筛选　公司用多年研发酪氨酸激酶抑制剂积累的化合物作酶水平的选择性活性筛选, 发现一些具有嘧啶母核的化合物对双突变EGFR的活性强于野生型, 其中化合物**5**活性较强, $IC_{50\,(DM酶)} = 0.009\ \mu mol\cdot L^{-1}$, 强于野生型88倍, 为可逆性抑制剂。

实验表明, 抑制DM酶活性强度虽然很高, 但对DM细胞的活性降低了90倍 ($IC_{50\,(DM细胞)} = 0.77\ \mu mol\cdot L^{-1}$)。可能的原因是细胞内有高浓度的ATP, 双突变酶对ATP有强效的结合作用, 导致降低了化合物**5**的竞争性抑制, 所以, 对先导物**5**的优化须是提高对DM细胞的抑制活性。



5

3　先导物的优化

3.1　引入共价键结合的基团　为了提高化合物**5**对细胞的活性, 设想分子中加入亲电性基团, 与酶的ATP结合域形成共价结合, 不可逆抑制可提高活性强度。利用酶开口处的残基Cys797可作为共价结合的位点 (第二代抑制剂多结合于该位点)。亲电基团的连接在化合物**5**的位置须经分子模拟确定。

为此, 将**5**对接到EGFR-T790M的晶体结构中, 如图1所示。嘧啶环N1和苯胺的NH同铰链的Met793形成两对氢键 (图中虚线), 氯原子指向门户残基Met790, 苯基哌嗪伸向ATP结合腔的开口处。

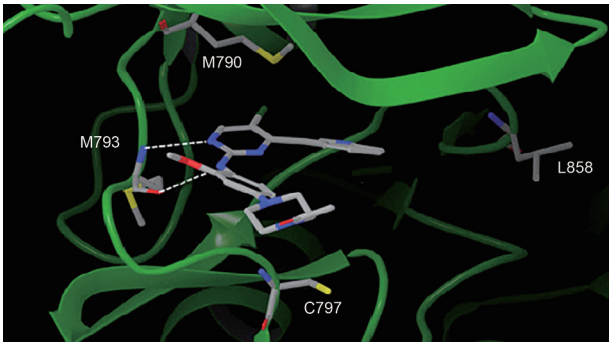


Figure 1　Molecular docking of compound **5** to EGFR-T790M

基于对接模型, 亲电性基团丙烯酰胺 (弱迈克尔加成基) 安排在苯胺的间位, 这是易于同Cys797的巯基发生共价结合的位置。此外, 还去除甲氧基和进入水相的哌嗪环, 以便减小分子尺寸, 为以后添加基团预留出结构空间。合成的化合物**6**在细胞水平上对单突变 (AM) 和双突变 (DM) 的活性都强于野生型 (WT) (表1)。

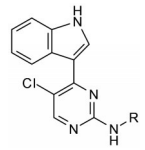
3.2　基团和位置的变换　为了提高活性, 化合物**6**的苯胺环上恢复甲氧基 (**7**), 提高了对DM细胞的活性, 但同时也提高了抑制WT活性, 因而选择性 (WT/DM) 略有降低。变换丙烯酰胺的位置 (**8**)、构象限制 (**9**) 或氰基乙酰胺 (**10**) 都使活性显著降低, 提示间位丙烯酰胺和甲氧基是优化的基团 (表1)。

化合物**6**与EGFR激酶结合的复合物单晶经X-射线衍射图分析, 证明丙烯酰胺确实与Cys797形成了共价键结合 (图省略) (Ward RA, Anderton MJ, Ashton S, et al. Structure- and reactivity-based development of covalent inhibitors of the activating and gatekeeper mutant forms of the epidermal growth factor receptor (EGFR). J Med Chem 2013, 56: 7025−7048)。

3.3　优化物理化学性质　化合物**7**的缺点是脂溶性强, 分布系数 $\log D_{7.4} > 4.3$, 对物化和药代性质不利, 为此, 对**7**的结构作两方面变更: 一是变换吲哚环成极性较强的片段, 另一是在丙烯酰胺链上加入碱性基团。合成的化合物**11**～**13**列于表2。结果表明用吡唑并吡啶置换吲哚环, 虽然对双突变细胞活性略有降低, 但对野生型作用很弱, 因而显著提高了选择性。在丙烯酰胺末端增添二甲胺甲基, 对DM细胞活性也略有减弱, 但降低了脂溶性, 因而亲脂性配体效率 (LLE) 没有下降。化合物**13**显示优良品质。

3.4　嘧啶环上氯原子的变换　前已述及的分子模拟提示, 嘧啶环5位的氯原子接近Met790, 疏水性相互作用有利于同双突变激酶的结合, 但这也因此加大了化合物的亲脂性。优化这一位置的取代基, 用H、F、CH_3和$SOCH_3$作等排置换 (数据从略), 对单突变株活性无

Table 1 Inhibitory activity of compounds **6**−**10** against enzymes and cells



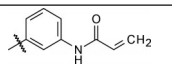
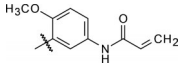
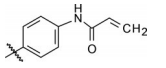
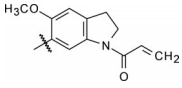
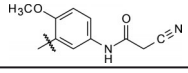
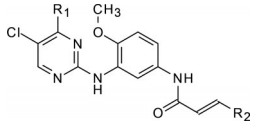
Compd.	R	DM (enz) IC_{50}/μmol·L^{-1}	AM (cell) IC_{50}/μmol·L^{-1}	DM (cell) IC_{50}/μmol·L^{-1}	WT (cell) IC_{50}/μmol·L^{-1}	WT/DM selectivity
6		0.053	0.27	0.081	3.5	43
7		0.006 3	0.029	0.022	0.55	25
8		0.32	19	12.9	> 30	> 1.6
9		0.052	2.8	2.1	> 30	> 15
10		0.065	1.8	2.8	> 30	> 11

Table 2 Physico-chemical property of optimized compounds. * Lipophilic ligand efficiency. LLE = $pIC_{50\ (DM\ cell)}$ − log $D_{7.4}$



Compd.	R_1	R_2	AM cell IC_{50}/μmol·L^{-1}	DM cell IC_{50}/μmol·L^{-1}	WT cell IC_{50}/μmol·L^{-1}	WT/DM selectivity	log $D_{7.4}$	LLE* (DM)
7		H	0.029	0.022	0.55	25	> 4.3	< 3.4
11		H	0.15	0.033	12.8	390	> 4.3	< 3.2
12		CH_2NMe_2	0.085	0.053	1.55	29	4.1	3.2
13		CH_2NMe_2	0.40	0.096	23	240	3.6	3.4

明显变化, 但双突变细胞株活性降低 2～3 倍, 对野生型没有变化。H 置换 Cl 原子导致双向变化: 选择性 (DM/WT) 降低了两倍, 而分布系数 (log $D_{7.4}$) 由 3.6 降至 2.6, LLE 由 3.4 升至 4.0, 因而在活性、选择性和物化性质上各有优势, 这也反映在以后的选择中。在两可中作出判断和取舍需要经验, 也有风险和运气。

3.5 亲电性侧链的反应性研究 丙烯酰胺作为亲电性基团, 可与较强的亲核性基团如巯基发生迈克尔加成反应, 共价结合的强度取决于两个因素: 抑制剂与酶蛋白可逆性结合的平衡常数 (K_i) 和生成共价键的速率常数 (k_{inact}), 所以, 最直接的方法是测定受试物的 k_{inact}/K_i 值。对本项目而言, 由于难度较大不适于化合物的优化筛选, 因而将反应双方都作简化: 用谷胱甘肽作为模板, 与各种简单取代的丙烯酰苯胺作为模型物, 发生的反应如图 2 所示, 用 LC-MS 测定加合产物, 计算反应速率和谷胱甘肽的半衰期, 以此确定苯环上不同取代基 (16 个) 对反应的影响, 并用 Hansch 分析对反应速率与苯环上取代基的电性参数作了相关性分析 (数据从略), 从而确定本项目亲电基团的反应性能窗口 (reactivity window)。

在此基础上合成了 50 多个目标化合物, 考察亲电基团与谷胱甘肽的反应速率 (数据从略), 还用已知不可逆抑制剂阿法替尼作对照, 确定了对谷胱甘肽最低反应性的半衰期小于 400 min。

3.6 嘧啶环上取代基的优化 基于提高反应性能的研究, 在苯胺环的对位引入哌嗪环, 也希望邻位的哌嗪环影响丙烯酰胺的构象和反应能力。由于哌嗪的碱性有利于溶解性, 所以去除了丙烯酰胺上的二甲胺基。合成的化合物是考察嘧啶环 4,5 位取代基对活性的影响, 列于表 3 中。

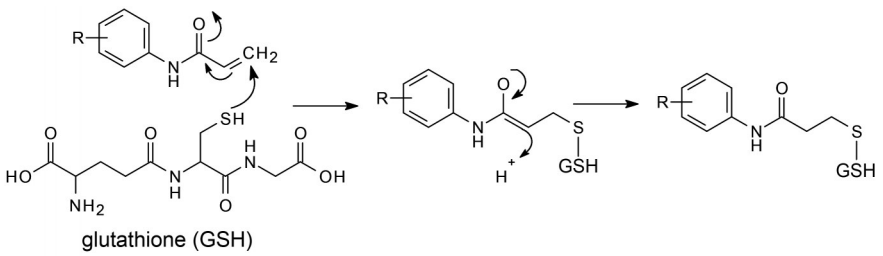
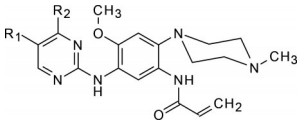


Figure 2　The reaction between glutathione and substituted acryloyl aniline

Table 3　Activity and lipophilicity of piperazine-containing compounds



Compd.	R_1	R_2	AM cell IC_{50}/μmol·L^{-1}	DM cell IC_{50}/μmol·L^{-1}	WT cell IC_{50}/μmol·L^{-1}	WT/DM selectivity	log $D_{7.4}$	LLE (DM)
14	Cl		0.14	0.019	12	620	4.1	4.1
15	H		0.14	0.071	19	270	Not test	−
16	Cl		0.023	0.005 6	0.77	140	3.7	4.6
17	H		0.017	0.068	13	19	3.4	3.7

结果表明, 两个含有吲哚片段的化合物 (**16**, **17**) 虽然对双突变细胞株活性很高, 但也提高了对野生型的活性, 这与前述的化合物相似。吡唑并吡啶杂环似乎仍是优选的, 氯原子有利于提高选择性。

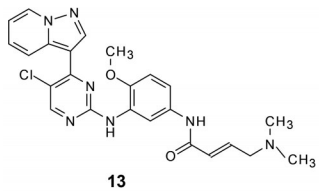
优化至此, 化合物 **13**、**14** 和 **17** 在活性强度和选择性以及物化性质 (综合可用 LLE 表征) 显示优良性质, 数据综合于表 4。**13** 的助溶基团是二甲胺基连接于丙烯酰胺上, **14** 和 **17** 的甲基哌嗪直接连于苯环上, 二者溶解性明显大于 **13**。**14** 尤其突出, 对双突变细胞活性高, LLE 值 4.1, 选择性高于 **13** 和 **17**, 灌胃裸鼠的吸收性和对心脏的安全性也呈现良好的品质。

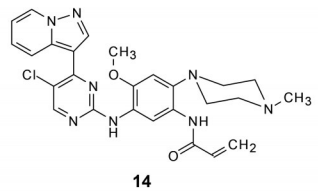
3.7　哌嗪环上的变换　以 **14** 为新的先导物进行优化, 变换的部位是哌嗪片段, 通过外端的氮原子再连接碱性基团 (引出的部分不会影响与活性部位结合)。表 5 列出了化合物的结构和性质。

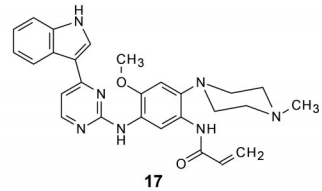
化合物 **18** 的侧链增链为二甲胺甲基, **18** 的活性、选择性和亲脂性配体效率有所降低, **19** 少一个 *N*-甲基, 或氨基换成羟基 **22**, 活性进一步降低。**20** 和 **21** 是一对对映体, *S* 构型强于 *R* 型, 活性与 **18** 相同, 但选择性稍差。连接哌啶 (**23**, **24**) 和哌嗪 (**25**, 成脲基片段) 的活性都显著降低。综上, 哌嗪环上再添加碱性的化合物没有出现优良化合物。

3.8　哌嗪环变换成其他碱性片段　用其他含氮杂环或碱性链替换哌嗪环, 筛选出 **26**～**28**, 活性和物化性质优于上述的化合物, 主要数据列于表 6。其中对血小板样生长因子 1 受体 (IGF1R) 抑制活性代表脱靶效应。

Table 4　*In vitro* and *in vivo* activity of the compounds **13**, **14** and **17**

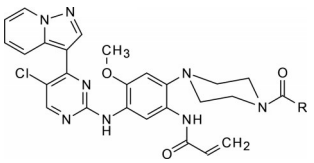

13


14


17

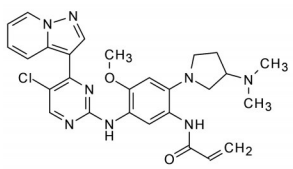
Compd.	AM cell IC_{50}/μmol·L^{-1}	DM cell IC_{50}/μmol·L^{-1}	WT cell IC_{50}/μmol·L^{-1}	log $D_{7.4}$	LLE (DM)	Solubility /μmol·L^{-1}	C_{max}/μmol·L^{-1}; AUC/μmol·L^{-1}·h	hERGs IC_{50}/μmol·L^{-1}
13	0.388	0.091	> 20	3.4	3.6	1	5.8; 67	4.2
14	0.14	0.019	12	3.6	4.1	> 973	2.7; 13.4	7.1
17	0.015	0.063	1.2	3.4	3.8	828	0; 0	9.1

Table 5 Structure, activity and physico-chemical property of the compounds with varied piperazne moiety

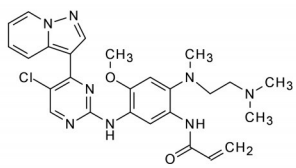


Compd.	R	AM cell IC_{50}/μmol·L^{-1}	DM cell IC_{50}/μmol·L^{-1}	WT cell IC_{50}/μmol·L^{-1}	log $D_{7.4}$	LLE (DM)
18		0.073	0.016	10	2.6	5.2
19		0.156	0.066	7.5	1.8	5.4
20		0.028	0.019	1.6	–	–
21		0.351	0.177	9	2.1	4.7
22		0.179	0.075	11	–	–
23		0.211	0.031	6.06	2.2	5.3
24		0.220	0.057	5.6	2.2	5.0
25		0.247	0.135	10.3	2.0	4.9

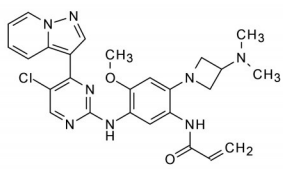
Table 6 Biological activity and physico-chemical property of the compounds **26**–**28**. *Insulin-like epidermal factor 1 receptor-expressing cell



26



27



28

Compd.	AM cell IC_{50}/μmol·L^{-1}	DM cell IC_{50}/μmol·L^{-1}	WT cell IC_{50}/μmol·L^{-1}	log $D_{7.4}$	LLE (DM)	Solubility /μmol·L^{-1}	IGF1R* IC_{50}/μmol·L^{-1}	hERGs IC_{50}/μmol·L^{-1}
26	0.006	0.002	0.357	3.1	5.6	28	0.007	5.6
27	0.002	0.000 6	0.145	2.8	6.4	222	0.006	6.7
28	0.021	0.004	0.938	3.3	5.1	43	0.026	4.0

由于**26**～**28**的活性、选择性和物化性质等都优于既往的化合物, 进行了体内药效学实验, 用EGFR激酶双突变细胞H1975细胞和单突变的PC9细胞接种的两种小鼠模型, 分别灌胃受试物10 mg·kg^{-1}, 观察体内EGFR磷酸化的降低百分率, 结果表明, **26**～**28**都显著降低两种模型的磷酸化程度, 持续时间24 h以上 (数据省略)。

进而用大鼠、小鼠和犬实验了药代动力学性质, 3个化合物在啮齿动物的口服生物利用度是22%～59%, 犬分别为67%、30%、56%, 药代的其他参数也较好。

然而对胰岛素生长因子受体1有较高抑制活性, 可引起高胰岛素血症等不良反应。脱靶作用的原因是IGF1R激酶的ATP结合域的门户氨基酸是甲硫氨酸, 与耐药 (双突变) EGFR激酶的门户氨基酸相同。这个潜在的脱靶作用对治疗剂量是个限制因素。

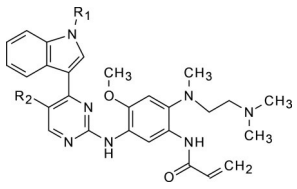
3.9 嘧啶环上取代基的再优化 化合物**26**～**28**呈现良好的药效和药代动力学性质, 提示碱性侧链的重要作用, 使得研发者回顾并重新审视吡唑并吡啶片段的依据, 决定对化合物**13**以及延续使用的这个杂环片段用吲哚置换, 以探索新的碱性链与吲哚的“搭配”是否是优化匹配, 而且吲哚环在合成上有优势。同时还应考察嘧啶环5位的氯原子的作用, 为此设计合成并评价了表7的化合物**29**～**34**。结果表明新的碱性侧链与吲哚片段的组合系列, 可优化出具有良好的活性、选择性和其他性质的化合物。

表中突出的化合物是**33**, 对突变细胞的活性虽然稍弱, 但对EGFR野生型细胞和高表达IGF1R细胞抑制作用很弱, 因而选择性明显高于其他化合物, 显示出有前景的性质。

4 候选化合物的确定和奥希替尼上市

4.1 高选择性化合物的选择 在高活性化合物中确定候选物需要从多维性质的比较中选择。研制者比较了化合物**13**、**26**、**27**、**28**和**33**的活性 (AM、DM和WT细胞)、溶解性、脱靶作用 (IGF1R和hERGs) 等性质 (表8)。

Table 7 Biological activity and physico-chemical property of the compounds with varied substituents on the pyrimidine ring



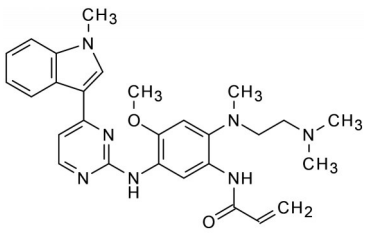
Compd.	R_1	R_2	AM cell IC_{50}/μmol·L^{-1}	DM cell IC_{50}/μmol·L^{-1}	WT cell IC_{50}/μmol·L^{-1}	log $D_{7.4}$	IGF1R IC_{50}/μmol·L^{-1}	Solubility /μmol·L^{-1}	hERGs IC_{50}/μmol·L^{-1}
29	CH_3	CN	0.001	0.000 9	0.046	2.7	0.038	57	4.3
30	CH_3	Cl	0.002	0.002	0.058	3.3	0.040	759	7.4
31	H	Cl	0.000 2	0.000 2	0.011	3.3	0.007	75	14.8
32	CH_3	CH_3	0.001	0.001	0.071	2.8	0.196	576	15.6
33	CH_3	H	0.017	0.015	0.480	3.4	2.900	7	16.3
34	H	H	0.002	0.002	0.033	2.9	0.263	79	17.5

Table 8 Comparison of multiple properties of compounds with high activity

Compd.	AM cell IC_{50}/μmol·L^{-1}	DM cell IC_{50}/μmol·L^{-1}	WT cell IC_{50}/μmol·L^{-1}	Solubility /μmol·L^{-1}	IGF1R IC_{50}/μmol·L^{-1}	hERGs IC_{50}/μmol·L^{-1}	IGF1R/DM
13	0.388	0.091	> 20	1	0.099	4.2	1.1
26	0.006	0.002	0.357	28	0.007	5.6	3.5
27	0.002	0.000 6	0.145	222	0.006	6.7	10
28	0.021	0.004	0.938	43	0.026	4.0	6.5
33	0.017	0.015	0.480	7	2.900	16.3	193

从选择性活性和安全性数据分析, **33**显示出优良性质(Finlay MR, Anderton M, Ashton S, et al. Discovery of a potent and selective EGFR inhibitor (AZD9291) of both sensitizing and T790M resistance mutations that spares the wild type form of the receptor. J Med Chem, 2014, 57: 8249−8267)。

4.2 奥希替尼批准上市 进而通过化合物的动物体内抗耐药瘤株实验、药代动力学实验和对小鼠引起高血糖和高胰岛素血症作用 (数据从略), 结果表明化合物**33**优胜于其他, 确定为候选化合物, 定名奥希替尼(osimertinib), 经临床前和临床研究, 证明对晚期非小细胞肺癌、特别是用一线替尼类药物治疗后疾病恶化的患者有效, 于2015年11月美国FDA批准上市。



Osimertinib

5 后记

奥希替尼是第三代EGFR激酶抑制剂。AstraZeneca公司从2009年启动项目, 2011合成和评价完化合物, 2012年确定了候选物, 2013年2月进入I期临床, 同年10月FDA批准进入快速轨道, 2014年5月进入II期临床, 2015年1月开始III期临床研究, 同年11月批准上市, 7年研发成功 (Yver A. Osimertinib (AZD9291)—a science-driven, collaborative approach to rapid drug design and development. Ann Oncol, 2016, 27: 1165−1170)。

这是一个有代表性的科技驱动研发的项目, 虽然有借鉴和跟随的成分, 但专门对T790M突变的EGFR靶标, 而不抑制野生型EGFR和胰岛素生长因子受体, 则显示出是个精准设计和多维评价的工作。从药物化学的视角, 两年内在不断的修正结构, 反复考量结构细节对活性、选择性、物化性质、脱靶性等的影响, 确定了活性化合物, 是高效率的研究。在技术层面上, 将分子模拟、QSAR、结构生物学和动力学同位素效应 (篇幅所限本文未述及) 等方法融汇到结构优化之中。精准医学需要有精准设计与制造的药物支撑, 奥希替尼的研制体现了这一要求。