Title: Anaplastic lymphoma kinase (ALK) inhibitors in the treatment of ALK-driven lung cancers

Author: <ce:author id="aut0005" author-id="S1043661817300166-f9866cf30d73cde1771a72979b3be151"> Robert Roskoski Jr.

PII: S1043-6618(17)30016-6
DOI: http://dx.doi.org/doi:10.1016/j.phrs.2017.01.007
Reference: YPHRS 3468

To appear in: *Pharmacological Research*

Received date: 5-1-2017
Accepted date: 6-1-2017

Please cite this article as: Roskoski Robert.Anaplastic lymphoma kinase (ALK) inhibitors in the treatment of ALK-driven lung cancers.*Pharmacological Research* http://dx.doi.org/10.1016/j.phrs.2017.01.007

Anaplastic lymphoma kinase (ALK) inhibitors in the treatment of ALK-driven lung cancers

Robert Roskoski Jr.

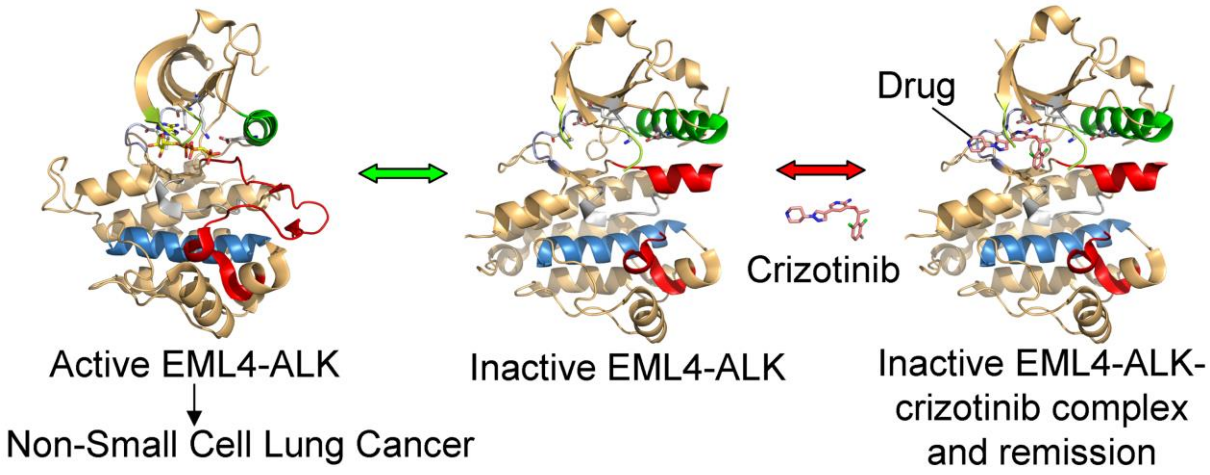
Blue Ridge Institute for Medical Research, 3754 Brevard Road, Suite 116, Box 19, Horse Shoe, North Carolina 28742-8814, United States

Phone: 1-828-891-5637

Fax: 1-828-890-8130

E-mail address: rrj@brimr.org

Graphical abstract



ABSTRACT

Anaplastic lymphoma kinase is expressed in two-thirds of the anaplastic large-cell lymphomas as an NPM-ALK fusion protein. Physiological ALK is a receptor protein-tyrosine kinase within the insulin receptor superfamily of proteins that participates in nervous system development. The EML4-ALK fusion protein and four other ALK-fusion proteins play a fundamental role in the development in about 5% of non-small cell lung cancers. The amino-terminal portions of the ALK fusion proteins result in dimerization and subsequent activation of the ALK protein kinase domain that plays a key role in the pathogenesis of various tumors. Downstream signaling from the ALK fusion protein leads to the activation of the Ras/Raf/MEK/ERK1/2 cell proliferation module and the JAK/STAT cell survival pathways. Moreover, nearly two dozen ALK activating mutations are involved in the pathogenesis of childhood neuroblastomas. The occurrence of oncogenic ALK-fusion proteins, particularly in non-small cell lung cancer, has fostered considerable interest in the development of ALK inhibitors. Crizotinib was the

first such inhibitor approved by the US Food and Drug Administration for the treatment of ALK-positive non-small cell lung cancer in 2011. The median time for the emergence of crizotinib drug resistance is 10.5 months after the initiation of therapy. Such resistance prompted the development of second-generation drugs including ceritinib and alectinib, which are approved for the treatment of non-small cell lung cancer. Unlike the single gatekeeper mutation that occurs in drug-resistant epidermal growth factor receptor in lung cancer, nearly a dozen different mutations in the catalytic domain of ALK fusion proteins have been discovered that result in crizotinib resistance. Crizotinib, ceritinib, and alectinib form a complex within the front cleft between the small and large lobes of an inactive ALK protein-kinase domain with a compact activation segment. These drugs are classified as type I½ B inhibitors because they bind to an inactive enzyme and they do not extend past the gatekeeper into the back pocket of the drug binding site.

Key words: Acquired drug resistance; Catalytic spine; Fusion protein; K/E/D/D; Protein kinase inhibitor classification; Targeted cancer therapy

Chemical compounds studied in this article: Alectinib: (PubMED CID: 49806720); Brigatinib: (PubMED CID: 68165256); Ceritinib: (PubMED CID: 57379345); Crizotinib (PubMED CID: 11626560); Entrectinib: (PubMED CID: 25141092); Lorlatinib: (PubMed CID: 71731823)

Abbreviations: ALK^+-NSCLC, ALK-positive non-small cell lung cancer; AS, activation segment; CS or C-spine, catalytic spine; CL, catalytic loop; EGFR, epidermal growth

factor receptor; GK, gatekeeper; IGF-1R, insulin-like growth factor-1 receptor; InsRK, insulin receptor protein-tyrosine kinase; LTK, leukocyte tyrosine kinase; PKA, protein kinase A; RS or R-spine, regulatory spine; SCLC; small cell lung cancer; Sh1, shell residue 1; VEGFR, vascular endothelial growth factor receptor.

1. Overview of lung cancers

1.1 Incidence

Lung cancer refers to tumors arising from the respiratory epithelium including the bronchi, bronchioles, or alveoli. It is the leading cause of cancer deaths both worldwide and in the United States. Siegel et al. estimated that about 224,000 new cases of lung cancer will be diagnosed in the United States in 2016 (118,000 men and 106,000 women) and 158,000 people will die of the disease (86,000 men and 72,000 women) [1]. Lung cancer deaths account for more that one-quarter of all cancer-related mortalities in the United States. Cancer of the lung occurs most often between the ages of 40 and 70 years of age with a peak incidence in the 50s and 60s. Cigarette smoking accounts for 85–90% all cases of lung cancer and is quantitatively the most important modifiable risk factor. However, 10–15% of lung cancers occur in people who have never smoked. The incidence of lung cancer in heavy smokers (two packs or 40 cigarettes per day for greater than 20 years) is 11%. The rising and falling incidences of lung cancers are correlated with the frequency of cigarette smoking in previous decades. Carcinoid tumors, hamartomas, and inflammatory myofibroblastic tumors represent rare pulmonary tumors not considered in this review, which focuses instead on carcinomas.

The signs and symptoms that prompt lung cancer patients to seek medical help vary with the size and specific location of the tumor [2]. The major chief complaints of patients at their initial clinic visit include cough (75%), weight loss (40%) chest pain (20%), and dyspnea (20%). Other findings include dysphagia (due to esophageal invasion), hemoptysis, hoarseness (due to recurrent laryngeal nerve invasion), post-obstructive pneumonia, and wheezing. Additional symptoms include variable bone pain owing to localized metastatic lesions. Occasionally the patient may be asymptomatic at the time of diagnosis and the finding is based upon an incidental chest X-ray performed for other reasons.

1.2 *Small cell lung cancer*

The main forms of lung cancer include small cell lung cancer (SCLC) [3] and non-small cell lung cancer (NSCLC) [2]. Small cell lung cancer is an aggressive disease that accounts for 10–15% of all lung cancers. Metastases to regional lymph nodes or distant sites is present in at least 90% of patients at the time of diagnosis thereby accounting for the poor clinical outcome. Histologically, small cell carcinomas have scant cytoplasm, finely granular nuclear chromatin (salt and pepper pattern), few nucleoli, and the tumor cells are smaller than three times the diameter of a small resting lymphocyte (< 25 μm). The mitotic rate is high and the cells lack a glandular or squamous organization. The risk of SCLC is correlated with the duration and the intensity of tobacco use. Accordingly, these malignancies rarely occur in never-smokers.

Small cell lung cancers are driven more by mutations and deletions of tumor suppressor genes rather than by alterations in oncogenes [3]. Loss of the functional tumor protein 53 (*TP53*) gene and loss of the retinoblastoma 1 (*RB1*) gene occurs in 75–90%

and nearly 100% of all SCLC cases, respectively. However, increased expression of the protein-tyrosine kinase Kit or its activating ligand (stem cell factor) and BCL-2 anti-apoptotic proteins each occur in about 90% of SCLCs. Poly(ADP-ribose)-1 polymerase (PARP1) overexpression occurs in more that one-half of all SCLC cases. Additionally, amplification of the MYC family of transcription factors is present in about one-fifth of all SCLC patients. Moreover, approximately two-thirds of these patients present with metastases to the liver, adrenal gland, bone, bone marrow, or brain. Symptoms of metastatic disease may include anorexia, bone pain, fatigue, headache, and seizures.

Small cell lung cancer patients rarely survive more than a few months without treatment [3]. In patients without extensive tumor spread (limited disease) at the time of diagnosis, median overall survival is about 17 months and five-year survival is about 12%. Such patients are treated with etoposide, cisplatin, and radiation therapy. In patients with widespread metastases (extensive disease), median overall survival is about 8–13 months and five-year survival is about 1%. These patients are treated with cyclophosphamide, anthracyclines, and vincristine (see Ref. [4] for a description of the mechanism of action of cytotoxic drugs) and radiation [3]. FDA-approved treatments for small cell lung cancer include doxorubicin, etoposide, mechlorethamine, methotrexate, topotecan, and everolimus – an mTOR protein kinase inhibitor (www.cancer.gov/about-cancer/treatment/drugs/lung).

Clinical trials using alisertib (an aurora A protein-serine/threonine kinase inhibitor), BMN-673 (a PARP1 inhibitor), and pembrolizumab or ipilimumab (immune checkpoint inhibitors) for the treatment of SCLC are underway [3]. Other agents undergoing clinical trials include CCT245737 and LY2606368 (checkpoint protein-

serine/threonine kinase 1 antagonists) and the following immune checkpoint inhibitors: (i) utomilumab (targets CD137 of the tumor necrosis receptor family), (ii) PF-04518600 (targets CD134, or OX40), (iii) JNJ-61610588 (targets V-domain Ig suppressor of T cell activation (VISTA) – a transmembrane protein occurring in tumor-infiltrating lymphocytes), (iv) PRD001 (targets PD-1), (v) CA-170 (an oral VISTA inhibitor), (vi) JNJ-64457107 (targets CD40, a co-stimulatory protein found in antigen-presenting cells), and (vii) Poly-ICLC (an immunostimulant) [5]. Targeted small molecule inhibitors in SCLC clinical trials include (i) sitravatinib, or MGCD516 (a multi-kinase inhibitor of PDGFR, VEGFR, and the Trk and Eph families), (ii) AMG 479 (an IGF-1R antagonist), (iii) INCB054828 (a fibroblast growth factor receptor antagonist), (iv) larotrectinib, or LOXO-101 (a Trk1/2/3 antagonist), (v) entrectinib, or RXDX-101 (an ALK, ROS1, and Trk1/2/3 multi-kinase inhibitor), (vi) cideranib, or AZD2171 (a Trk1/2/3 and VEGFR1/2/3 antagonist), (vii) galunisertib, or LY2157299 (a TGFβR1 antagonist), and (viii) ralimetinib, or LY-2228820 (a MAP kinase p38α inhibitor). Olarparib, or AZD2281, is a PARP1 antagonist and EC1456 (a folate-tubulysin conjugate that is taken up by cells expressing the folate receptor and delivers the microtubule inhibitor intracellularly) are two additional drugs that are in SCLC clinical trials [3].

1.3 Non-small cell lung cancer

Although cigarette smoking is the chief causative agent of lung cancers, exposure to asbestos, radon, arsenic, chromium, vinyl chloride, and polycyclic aromatic hydrocarbons are also implicated as risk factors [2]. Adenocarcinomas (40%–50%), squamous cell carcinomas (20–30%), and large cell carcinomas (≈ 10%) are the three main histologic types of NSCLC. Adenocarcinomas are characterized histologically by

gland formation and mucin production (*aden* is Greek for gland); these malignancies usually originate in the lung periphery. *KRAS* mutations occur in 30–35% and *EGFR* mutations occur in 10–15% of people with lung adenocarcinomas. *BRAF* (7–10%) and *MET* (7%) mutations, *EML4-ALK* translocations (4%), *ROS1* fusions (1%), *RET* fusions (1%), and *ERBB2/HER2* mutations (2–4%) represent driver genes that occur in adenocarcinomas. Those non-smokers who acquire lung cancer most commonly have adenocarcinoma. Squamous cell carcinomas are more common in men than women and are characterized microscopically by keratinization with markedly eosinophilic cytoplasm (*squamosa* is Latin for scale like). These carcinomas most often arise in larger bronchi and often metastasize to the central hilar lymph nodes early in the course of the disease, but they generally disseminate outside of the thorax later than other categories of lung cancer. These malignancies display frequent deletions of the *TP53* and *RB1* genes. The cyclin-dependent kinase inhibitor *CDKN2A* gene, which encodes p16/INK4A, is lost in 65% of these tumors. Large-cell lung carcinoma (LCLC) is a variable group of undifferentiated malignancies that originate from transformed lung epithelial cells. Such tumor cells have a moderate amount of cytoplasm and exhibit large nuclei with prominent nucleoli.

Shaw et al. reported that patients with ALK^+-NSCLC are generally younger than the general population with lung cancer [6]. The patients were also former or never cigarette smokers and their tumors were usually adenocarcinomas often with so-called signet ring morphology. ALK^+-NSCLC patients usually lack *KRAS* and *EGFR* mutations.

The mainstay of nonmetastatic lung cancer treatment is surgical removal [2]. Unfortunately, only one-third of the lung cancers are diagnosed before the tumor has

metastasized. Pre-existing medical conditions in patients with early stages (I or II) of lung cancer decrease the number of patients that can undergo surgery (See Ref. [2] for a discussion of the intricate staging classification of lung cancer). Even after a complete surgical resection, the recurrence rates of NSCLC are high, thereby necessitating the use of chemotherapy post-surgery (so-called adjuvant treatment) using cisplatin-vinorelbine. Other agents such as pemetrexed, docetaxel, etoposide, and gemcitabine are also used in various combinations in the treatment of NSCLC. In patients with advanced stages (III or IV) of disease, combined chemoradiation therapy pre-surgery (so-called neoadjuvant treatment) is beneficial. For patients unable to undergo surgery owing to poor pulmonary function or other co-existing morbidities, stereotactic radiation therapy represents a possible option. Although combined chemotherapy and radiotherapy improves the outcome of patients with metastatic lung cancer, the overall five-year survival is less that 15%. A common chemotherapeutic regime for non-squamous cell NSCLC includes carboplatin and paclitaxel. Bevacizumab, an angiogenesis inhibitor that binds to VEGF-A, in combination with paclitaxel and carboplatin, improves the survival of people with advanced, or metastatic, non-squamous cell NSCLC.

FDA-approved treatments for non-small cell lung cancer include the following cytotoxic drugs: carboplatin, gemcitabine, mechlorethamine, methotrexate, paclitaxel, pemetrexed, and vinorelbine (www.cancer.gov/about-cancer/treatment/drugs/lung). Also approved are the following monoclonal antibodies: atezolizumab (an immune checkpoint inhibitor directed against PD-1 – programmed cell death ligand 1), bevacizumab, necitumumab (targets the EGFR receptor and is used in combination with gemcitabine and cisplatin for the treatment of squamous cell NSCLC), nivolumab (an immune

checkpoint inhibitor that binds to PD-1), pembrolizumab (an immune checkpoint inhibitor that targets the PD-1 receptor), and ramucirumab (an angiogenesis inhibitor that targets VEGFR2). Small molecule targeted protein kinase inhibitors that are FDA-approved include: afatinib (an inhibitor of EGFR/ErbB1, ErbB2, and ErbB4), alectinib (an inhibitor of ALK and RET), ceritinib (an inhibitor of ALK, IGF-1R, and ROS1), crizotinib (an inhibitor of ALK, c-MET, ROS1), erlotinib (an inhibitor of EGFR), everolimus (an inhibitor of mTOR), gefitinib (an inhibitor of EGFR), and osimertinib (an inhibitor of EGFR T970M). A variety of regimen combinations that are recommended for the treatment of all types of NSCLC can be found at the National Comprehensive Cancer Network (www.NCCN.org). Because lung cancers represent such a large proportion of all cancers, it is not surprising that there are hundreds of clinical trials focusing on the treatment of these maladies. See Ref. [5] for a review of dozens of cytotoxic, immunological, and targeted therapies in clinical trials for SCLC and NSCLC. For a summary of the clinical trials involving orally effective crizotinib, ceritinib, alectinib, brigatinib, entrectinib, and lorlatinib that target EML4-ALK fusion proteins, see Ref. [7].

2. Role of ALK and ALK-fusion proteins in the pathogenesis of malignancies

2.1 Discovery and architecture of Anaplastic Lymphoma Kinase (ALK)

Anaplastic lymphoma kinase is a member of the insulin receptor protein-tyrosine kinase superfamily. In 1994, two groups characterized a tyrosine-protein kinase in anaplastic large-cell lymphoma (ALCL) cell lines [8,9] where anaplastic denotes dedifferentiated cells with a bizarre morphology. Nearly two-thirds of these anaplastic lymphomas exhibit a chromosomal translocation where the entire nucleophosmin (*NPM*) gene is fused to the 3´ portion of the *ALK* gene, which includes its entire intracellular

component. This cancer-causing ALK protein kinase is thus a chimeric protein resulting from a translocation between chromosomes (2;5)(p23;q35) that encodes an NPM-ALK fusion protein (Fig. 1A). As a consequence, this chromosomal reshuffling results in the ectopic expression driven by the NPM promoter of the NPM-ALK fusion protein kinase that is constitutively activated; the anaplastic lymphoma kinase was named after the disease [8].

Moreover, an EML4 (echinoderm microtubule-associated protein like 4)-ALK oncoprotein was subsequently discovered in non-small cell lung cancers a dozen years later (Fig. 1B) [10,11]. The *ALK* gene and the *EML4* gene, which occur on chromosome 2p, have opposite orientations and are separated by 12 Mb. The *EML4* gene is broken at a position 3.6 kb downstream of exon 13 and is ligated to nucleotides 297 bp upstream of exon 21 of *ALK* giving rise to an EML4-ALK fusion protein. More than two dozen additional oncogenic ALK fusion proteins have been found in a variety of diseases including anaplastic large cell lymphomas, diffuse large B cell lymphomas, inflammatory myofibroblastic tumors, NSCLC, and breast, colorectal, esophageal, and renal cell carcinomas [12,13]. These fusion proteins are derived from chromosomes 1, 2 (where the *ALK* gene is found), 3, 4, 5, 7, 9, 10, 11, 14, 17, 19, 22, and X. In addition to fusion proteins, increased ALK activity owing to point mutations or enzyme overexpression occurs in many childhood neuroblastomas. This neoplasm is the most common cancer in babies and the third most common malignancy in children after leukemia and brain neoplasms.

Physiological ALK is a transmembrane receptor protein-tyrosine kinase that consists of an 18-residue signal peptide, an extracellular portion of 1020 residues, a 21-

residue transmembrane segment, and finally a 561 amino acid intracellular section (UniprotKB ID: Q9UM73) [14]. The intracellular portion consists of a juxtamembrane (JM) segment, a protein kinase domain, and a carboxyterminal tail (Fig. 1C and D). The extracellular region contains an LDLa (low density lipoprotein class A) section sandwiched between two MAM segments followed by a glycine-rich fragment. MAM refers to meprin, A5 protein, and receptor protein-tyrosine phosphatase μ. A single MAM domain consists of about 170 amino acid residues and these MAM modules likely contribute to cell–cell interactions. The LDLa sector is of unknown function and contains two or more disulfide bridges along with a constellation of negatively charged residues. The function of the glycine-rich fragment is uncertain. However, a point mutation of a single glycine residue in the glycine-rich extracellular section in *Drosophila melanogaster* produces ALK inactivation thereby demonstrating the functional importance of this sector, at least in *Drosophila* [15].

Mammalian ALK is hypothesized to play a role in the development of the nervous system based upon the expression of its mRNA during mouse embryogenesis [14,16,17]. Iwahara et al. reported that the expression of ALK mRNA and protein in mice diminishes after birth; the mRNA and protein reach a minimum after three weeks of age and are thereafter maintained at low levels to and through animal adulthood [16]. Morris et al. reported that ALK mRNA is expressed in several adult human tissues including brain, colon, prostate, small intestine, and testis, but not in normal human lymphoid cells, heart, kidney, lung, liver, ovary, pancreas, placenta, skeletal muscle, spleen, or thymus [18]. That ALK expression is not ubiquitously expressed may be advantageous because it decreases the likelihood of toxicity in response to therapies targeting this enzyme.

Because there were no known ligands for the ALK receptor for more than two decades after its discovery, it was considered to be an orphan receptor. However, two small proteins each with a molecular weight of 20 kDa (PFAM150A and B) were discovered to stimulate ALK activity [18,19]. Reshetnyak et al. reported that these ligands, which they called AUG-α and AUG-β (for augmentor), are dual activators of ALK and LTK (leukocyte tyrosine kinase) [19]. LTK is the closest relative of ALK [20]. However, the physiological function of these ligands is presently unclear. Because these ligands are able to superactivate mutant ALK proteins that occur in some neuroblastomas, the development of antibodies that bind to these ligands represents a potential therapeutic approach for children with these tumors.

2.2 Primary structure of the ALK protein-tyrosine kinase domain

Hanks and Hunter investigated the sequences of various groups of 60 protein kinases and partitioned their primary structures into 12 sequential segments (I-VIA,VIB-XI) [21]. The catalytic domain of ALK consists of 287 amino acid residues, which is an average size for a protein kinase. Domain I of ALK contains a glycine-rich loop (GRL) with a GxGxΦG signature (^{1123} GHGAFG^{1128}), where Φ refers to a hydrophobic residue and is phenylalanine in the case of ALK. The glycine-rich loop occurs between the β1- and β2-strands and overlays the ATP/ADP-binding site (Fig. 2A). The glycine-rich loop must be flexible owing to its role in both ATP binding and ADP release and stretches including glycine have this property. Domain II of ALK contains a conserved Ala-Xxx-Lys (^{1148}AVK^{1150}) sequence in the β3-strand and domain III contains a conserved glutamate (E1167) in the αC-helix that forms an electrostatic bond with the conserved β3-lysine in the active protein kinase conformation. Domain VIB of ALK contains a

conserved HRD sequence, which forms part of the catalytic loop (^{1247}HRDIAARN^{1254}). Domain VII contains a ^{1270}DFG^{1272} signature and domain VIII contains a ^{1297}PPE^{1299} sequence, which together represent the beginning and end of the protein kinase activation segment. The activation segment is so named because it adopts different conformations between the more active and less active protein kinase structures. The PPE sequence occurs at the end of the activation segment of several protein kinases, but APE is more common. Domains IX–XI make up the αE–αI helices. The initial X-ray structure of PKA provided a valuable framework for understanding the interrelationship of the 12 domains in mediating catalysis, as described next.

2.3 Secondary and tertiary structure of ALK and the K/E/D/D signature

The elucidation of the X-ray crystal structure of the catalytic subunit of PKA by Susan Taylor's group has enlightened our views on the fundamental biochemistry of the entire protein kinase superfamily (PDB ID: 2CPK) [22,23]. All protein kinases including the ALK have a small amino-terminal and large carboxyterminal lobe [24]. The small lobe contains five conserved β-strands (β1–5) and an important regulatory αC-helix and the large lobe contains four conserved short strands (β6–β9) along with seven helices (αD–αI and αEF). A cleft or pocket between the N- and C-terminal lobes serves as a binding site for ATP. Of the thousands of protein kinase structures that have been reported, all of them possess this fundamental protein kinase fold as first described for the catalytic subunit of PKA [23,24].

Essentially all functional protein kinases contain a K/E/D/D (Lys/Glu/Asp/Asp) signature that plays crucial mechanistic roles in the protein kinase reaction (Table 1) [25]. The lysine and glutamate occur within the amino-terminal lobe and the two aspartate

residues occur within the carboxyterminal lobe. Although both lobes participate in ATP binding, the N-terminal lobe plays a key role in this process. Because there are apparently no X-ray crystallographic structures in the public domain of the active form of ALK, which is a member of the insulin receptor superfamily, we have used the structure of the tris-phosphorylated form of the active insulin receptor protein kinase (InsRK) for comparison (Fig. 2B). The E1047 carboxylate group of the insulin receptor (the E of K/E/D/D) of the αC-helix forms an electrostatic bond with the ε-amino group of K1030 and serves to stabilize its interactions with the ATP phosphates.

The presence of a salt bridge between the β3-lysine and the αC-glutamate is obligatory for the formation of an active protein kinase, which corresponds to an "αC-in" structure (Fig. 2B). The X-ray crystal structure of the ALK-ADP complex fails to show the location of the β3-K1150 beyond the δ-carbon atom so that we cannot tell whether it is within hydrogen-bonding distance with αC-E1047 (Fig. 2A). The failure of these residues to make electrostatic contact constitutes an inactive "αC-out" structure. As illustrated later, some of the inactive forms of ALK bound to drugs do display an "αC-in" conformation. The αC-in structure is necessary, but not sufficient, for the manifestation of catalytic activity. Moreover, Lee et al. reported that the relative interlobe closure between the small and large lobes of this ALK-ADP complex is intermediate between the open catalytically active and the closed dormant conformations observed in the InsRK structures indicating that ALK (PDB ID: 3LCT) is in a dormant state [26].

The large C-terminal lobe participates in protein/peptide substrate binding and plays an essential role in catalysis. Two Mg^{2+} ions are required for each catalytic cycle of several protein kinases [27] and are most likely required for the functioning of ALK. By

inference, D1167 (the DFG-D and the first D of K/E/D/D) binds to Mg^{2+}(1), which in turn binds to the β- and γ-phosphates of ATP. In this active conformation, the DFG-D is directed inward toward the active site. In the inactive form of many protein kinases, the DFG-D points outward (DFG-D out) from the active site. ALK N1254 of the catalytic loop binds Mg^{2+}(2), which in turn binds to the α- and γ-phosphates of ATP. We infer that the activation segment of ALK in its active state forms an open structure that allows protein/peptide binding as depicted for the insulin receptor (Fig. 2B). The activation segment in the less active ALK occurs in a compact and closed structure that blocks protein/peptide binding. The proximal activation loop forms an αAL-helix that is not observed in active protein kinases and this proximal helix abuts with the αC-helix to keep it immobilized and in an inactive configuration (Fig. 2A). Both active and less active protein kinases contain an additional helix (αEF) near the end of the activation segment. The exocyclic 6-amino nitrogen of ATP characteristically forms a hydrogen bond with the carbonyl backbone residue of the first ALK hinge residue (E1197, Fig. 2A) that connects the amino-terminal and carboxyterminal lobes of the protein kinase domain and the *N*1 nitrogen of the adenine base forms a second hydrogen bond with the N–H group of the third hinge residue (M1199, not shown). As noted later, most small-molecule steady-state ATP competitive inhibitors of most protein kinases including ALK make hydrogen bonds with the backbone residues of the connecting hinge.

The activation segment binds the protein-substrate thus playing an important role in the catalytic cycle [28]. The beginning of the segment is located near the N-terminus of the αC-helix and the conserved catalytic loop HRD. The interfaces of these units are linked by hydrophobic interactions. As in the case of most protein kinases [29],

phosphorylation of a residue or residues within the activation segment converts a less active to a more active enzyme [30]. ALK, InsRK, IGF-1RK, LTK, the closest relative of ALK [20]), IRRK (insulin receptor related kinase), and ROS1 contain three phosphorylatable tyrosine residues with a sequence YxxxYY. This segment is Y1278/Y1282/Y1283 in ALK and Y1158/Y1162/Y1163 in InsRK. The following scheme describes the activation of InsRK: phosphorylation of Y1162 is followed by that of Y1158 and then Y1163 [31]. This same ordering was found for IGF-1RK with phosphorylation of the second tyrosine, followed by the first, and then by the third tyrosine in the YxxxYY sequence [32]. The ordering for the activation segment phosphorylation of the NPM-ALK fusion protein is considered next.

Lemmon and Schlessinger reviewed the mechanisms of activation of several receptor tyrosine-protein kinases [33] and this analysis provides us with a possible scheme for ALK activation. One inferred mechanism for the ligand dimer-induced activation of physiological ALK involves the transphosphorylation of the activation loop tyrosines by a partner ALK protein kinase. The newly trisphosphorylated ALK then catalyzes the transphosphorylation of the partner ALK activation loop tyrosines. For the NPM-ALK fusion protein found in anaplastic large cell lymphomas, Tartari et al. reported that Y1278 is phosphorylated first and that of Y1282 and Y1283 occur later (human native ALK residue numbers) [34]. Whether or not this is the order followed in the native ALK receptor or EML4-ALK is unknown and deserves study. In the case of the InsRK, activation loop trisphosphorylation allows the αC-helix to move into its active configuration [26].

A possible mechanism for ligand and dimer-induced activation of ALK involves the phosphorylation of one or more of the juxtamembrane tyrosine residues such as Y1092 or Y1096 (Fig. 1D). These modifications would allow the enzyme to assume its active state and would be followed by the phosphorylation of the activation loop tyrosine residues. Such a mechanism has been described for the activation of the Kit receptor protein-tyrosine kinase [35]. After induced dimerization by the stem-cell factor, or Kit ligand, transphosphorylation of Y568 and Y570 in the juxtamembrane segment initiates Kit activation. These modifications are followed by transphosphorylation of Y823 in the activation loop to yield fully active Kit. Additional experiments are required to determine the ordering of ALK juxtamembrane segment and activation loop phosphorylation. The activation of protein kinases with multiple phosphorylation sites such as the InsRK involves a defined and invariant order. The activation of the various ALK fusion proteins presumably relies upon the dimerization mediated by the N-terminal fusion protein segment followed by the transphosphorylation of one member of the dimer by the other as described for ligand dimer-induced activation [12].

The catalytic loop surrounding the actual site of phosphoryl transfer within the large lobe consists of ^{1247}HRDIAARN^{1254} in ALK. The catalytic aspartate (D1249) in ALK, which is the first D of K/E/D/D, functions as a base and accepts a proton from the tyrosyl –OH group (Fig. 3). The catalytic segment AAR sequence occurs in many receptor protein-tyrosine kinases such as PDGFR, EGFR, and ALK while RAA occurs in many non-receptor protein-tyrosine kinases such as Src [21].

2.4 The hydrophobic spines of inactive ALK and active insulin receptor protein kinase

Kornev et al. examined the structures of about two dozen active and inactive protein kinases and they determined the role of several crucial residues by a local spatial pattern alignment algorithm [36,37]. Their investigation led to the categorization of eight hydrophobic residues as a catalytic or C-spine and four hydrophobic residues that constitute a regulatory or R-spine. These spines consist of amino acid residues occurring in both the amino-terminal and carboxyterminal lobes. The R-spine contains one residue from the αC-helix and another from the activation loop, both of which are key components in determining the more active or less active kinase states. The adenine moiety of ATP constitutes part of the C-spine. The lower portion of the R-spine functions as an anchor for the peptide/protein substrate binding site and the C-spine tethers ATP within the cleft thus enabling catalysis. Moreover, the accurate alignment and arrangement of both spines are required for the production of an active enzyme as described for the cyclin-dependent protein kinases, EGFR, ERK1/2, the Janus kinases, MEK1/2), and Src [4,27,38–42].

The classical protein kinase R-spine consists of the catalytic loop HRD-His, the activation segment DFG-Phe, an amino acid four residues C-terminal to the conserved αC-glutamate, and an amino acid near the beginning of the β4-strand [36,37]. The backbone N–H of the HRD-His is anchored to the hydrophobic αF-helix by a hydrogen bond to an invariant aspartate carboxyl group. Going from the ventral to the dorsal part of the spine, Meharena et al. dubbed the R-spine residues RS0, RS1, RS2, RS3, and RS4 (Fig. 4A and B) [43]. Although the R-spine of inactive enzymes is usually nonlinear or broken, the R-spine of inactive ALK is linear like that of active the InsRK (Fig. 4B and C).

The C-spine of protein kinases is made up of residues from both the N-terminal and C-terminal lobes and it is completed by the adenine base of ATP (Fig. 4A) [43]. The two residues of the small lobe that interact with the adenine component of ATP include a conserved valine at the beginning of the β2-strand (CS7) and a conserved alanine residue from the canonical AxK of the β3-strand (CS8). Furthermore, a lipophilic residue from the β7-strand (CS6) interacts with the adenine portion of ATP. The CS6 residue is sandwiched between two hydrophobic residues (CS4 and CS5) that make hydrophobic contact with the CS3 residue near the beginning of the αD-helix of the large lobe. CS5/6/4 immediately follow the catalytic loop asparagine (HRDxxxxN). Finally, CS3 and CS4 make hydrophobic contacts with the CS1 and CS2 residues of the αF-helix thus forming a completed catalytic spine (Fig. 4A and B) [44]. Significantly, the hydrophobic αF-helix anchors both the C- and R-spines. Furthermore, both spines play an essential role in supporting the protein kinase catalytic residues in an active conformation. CS7 and CS8 and nearby residues in the amino-terminal lobe make up the "ceiling" of the adenine binding pocket and CS5/6/7 and nearby residues make up the "floor" of the ATP-binding pocket.

Using site-directed mutagenesis, Meharena et al. identified three shell (Sh) residues in the catalytic subunit of protein kinase A that stabilize the regulatory spine, which they labeled Sh1, Sh2, and Sh3 [43]. The Sh2 residue corresponds to the gatekeeper residue. The gatekeeper designation indicates the role that this amino acid plays in controlling access to a back cleft of the back pocket [45,46], which is sometimes called hydrophobic pocket II (HPII) [47,48]. In contrast to the identification of the APE, DFG, or HRD signatures, which is based upon their primary structures [21], the spines

were identified by their spatial locations in more active and less active protein kinases [36,37]. Table 2 provides a summary of the spine and shell residues of ALK, InsRK, and PKA. As described later, small molecule protein kinase antagonists often interact with residues that constitute the catalytic spine and sometimes the regulatory spine and shell residues [44].

Inactive ALK does not bear many of the properties associated with dormant protein kinases. For example, the DFG-Asp is pointed inward toward the active site, the configuration of the αC-helix is nearly in its active configuration, and the catalytic and regulatory spines are linear and are neither bent nor broken; these are properties of an active enzyme. However, the activation segment is closed and compact, which is a property of inactive enzymes, in contrast to the activation segments of active enzymes that are open and extended. Moreover, the αAL-helix at the beginning of the activation segment and the N-terminal β-turn within the juxtamembrane segment abut with and immobilize the αC-helix thereby hindering its required movement during catalysis [26]. As noted above, it is unusual to find that the catalytic and regulatory spines of a dormant enzyme such as ALK are nearly superimposable with those of an active enzyme (InsRK) (Fig. 4D). A nine-residue helix at the beginning of the activation loop (αAL) disallows the formation of an open activation segment and results in an inactive enzyme. This compact activation segment blocks the peptide/protein binding site, but it allows ADP/ATP to bind (Fig.2A).

3. Structure of ALK-drug complexes

Crizotinib is a first generation antagonist that is approved for the first-line treatment of ALK^{+}-NSCLC (Table 3). The X-ray crystal structure shows that amino

group of the aminopyrimidine scaffold of crizotinib (Fig. 5A) forms a hydrogen bond with the carbonyl group of the first hinge residue (E1197), the N1 nitrogen of the pyrimidine moiety forms another hydrogen bond with the N–H group of the third hinge residue (M1199) (Fig. 6A). The drug occupies only the front cleft and does not extend past the gatekeeper residue. Additionally, crizotinib makes hydrophobic contact with L1122 within the glycine-rich loop, V1130 (CS7) after the loop, A1148 (CS8), the gatekeeper L1196 (Sh2), the ^{1198}LMA^{1200} triad of the hinge, N1254 within the catalytic loop, and L1256 (CS6) in the large lobe. L1122 and V1130 occur above the plane of the pyrazopyridine scaffold and L1256 occurs below it. The 3-fluorophenyl group of the drug makes van der Waals contact with R1253 of the catalytic loop and with DFG-D1270. The piperidine ring, which is attached to the pyrazol-4-yl moiety, is directed away from the enzyme into the solvent.

Crizotinib targets additional protein kinases including c-MET or the hepatocyte growth factor receptor (HGFR) protein-tyrosine kinase (its original drug target) [51] and ROS1 [52], another member of the insulin receptor superfamily of protein-tyrosine kinases [20]. The ALK-crizotinib complex has DFG-D pointed inward toward the active site, a linear R-spine, and an αC-helix "in" conformation and a disordered activation segment. However, the initial portion of the activation segment possesses an αAL helix indicating that the segment is in a closed or compact conformation and the enzyme is thus in a less active conformation. This corresponds to a type I½ inhibitor and not a type I inhibitor as previously classified [44]. Similarly, ceritinib is a type I½ inhibitor and not a type I inhibitor.

Liao [47] and van Linden et al. [48] divided the gap between the N-terminal and C-terminal lobes into a front cleft or pocket, a gate area, and a back cleft. The back pocket corresponds to the gate area and the back cleft. The front cleft includes the Gly-rich loop, the hinge, the linker connecting the hinge with the αD-helix in the large lobe, and residues within the catalytic loop. The gate area includes residues within the β-3 strand and the initial portion of the activation segment including DFG. They also described several sub-pockets within these three regions. For example, the front cleft contains an adenine pocket (AP) adjacent to FP-I and FP-II subpockets and the gate area contains BP-I-A and BP-I-B subpockets. The back cleft consists of the αC-helix, the αC-β4 loop, portions of the β-4 and β-5 strands, and portions of the αE-helix. Crizotinib binds to the ALK front pocket and the FP-I subpocket. We previously divided type I½ inhibitors into two classes: A and B. The type A inhibitors extend into the back pocket while the type B inhibitors do not. Because crizotinib does not extend into the back pocket it is classified as a type I½ B inhibitor.

van Linden et al. provided a comprehensive summary of ligand and drug binding to more than 1200 human and mouse protein kinase catalytic domains deposited in the protein data bank [48]. Their KLIFS catalogue (kinase–ligand interaction fingerprint and structure) reports on the alignment of 85 protein kinase-ligand binding-site residues, which facilitates the recognition of family specific interaction features as well as the classification of ligands according to their binding properties. These investigators employ a standard residue numbering system that facilitates a comparison among all protein kinases. See Table 2 for the correspondence between the C-spine, Shell, and R-spine residues and the KLIFS residues. Moreover, van Linden et al. have established a valuable

noncommercial and searchable web site that provides updated information on protein kinase-ligand interactions (http://www.vu-compmedchem.nl/).

Crizotinib was the first FDA-approved small molecule antagonist used for the treatment of ALK^+-NSCLC with a response rate of about 60% indicating that primary resistance occurs in 40% of patients [53–56]. Additionally, acquired or secondary resistance occurs in the 60% of those patients that respond with a median duration of about 10.5 months after the initiation of treatment. Acquired resistance in about one-third of the cases is due to mutations or overexpression of EML4-ALK fusion proteins, another third is due to up-regulation of alternative bypass signaling pathways, and the mechanisms for the final third is unknown. The L1196M gatekeeper mutation is the most common ALK mutation conferring crizotinib resistance while other resistance mutations include I1171T, F1174C, G1202R, S1206Y, and G1269A [53]. Such drug-resistant mutations prompted the development of second generation ALK inhibitors such as ceritinib and alectinib.

Ceritinib is a second generation medicinal [57] that is approved by the FDA for the treatment ALK^+-NSCLC in patients who have developed resistance or are intolerant to crizotinib. Ceritinib overcomes ALK-bearing I1171T, L1196M, S1206Y, and G1269A mutations, but not ALK with F1174C or G1202R mutations [7,50,54]. The ceritinib crystal structure depicts a binding pattern similar to that of the ALK–crizotinib complex with a compact activation segment conformation corresponding to a dormant enzyme [56], which indicates that it is a type I½ B inhibitor of ALK. The *N*3 of the pyrimidine (Fig. 5B) forms a hydrogen bond with the N–H of M1199 and the N–H of the 2-aminopyrimidine forms a hydrogen bond with the M1199 carbonyl within the ALK hinge

(Fig. 6B). A water molecule forms hydrogen bonds with β3-K1150 and one of the sulfonyl oxygens (not shown). Ceritinib makes hydrophobic contacts with the β1-strand L1122, H1124 within the G-rich loop, V1130 (CS7) after the G-rich loop, A1148 (CS8), β3-K1150, the gatekeeper L1196 (Sh2), the ^{1198}LMA^{1200} triad of the hinge, and L1256 (CS6). The 5-chloropyridine and sulfonylphenyl drug fragments make hydrophobic contact with L1122, A1148 and H1124 in the ceiling of the cleft. The terminal 2-isopropoxy-3-(piperidin-4-yl)phenyl group lies between the solvent and the ATP-binding cleft. Ceritinib inhibits InsRK, IGF-1RK, and ROS1 protein kinase with an inhibitor profile that partially overlaps that of crizotinib. Ceritinib binds in the ALK front cleft (the adenine pocket and FP-I) without extending to the back cleft [48]. The L1196M gatekeeper mutation of ALK is one of the mechanisms responsible for the acquired resistance to crizotinib and its inhibition by ceritinib thus represents an important FDA-approved second-line therapy for crizotinib-resistant ALK$^+$-NSCLC.

Alectinib is a second generation ALK antagonist that is built upon a 9-ethyl-6, 6-dimethyl-11-oxo benzo[*b*]carbazole scaffold (Fig. 5C) [58]. This drug is effective against the ALK L1196M gatekeeper mutation along with C1156Y and F1174L mutations [7]. The 11-oxo group of the drug forms a hydrogen bond with the N–H group of ALK M1199 of the hinge (Fig. 6C). Alectinib makes hydrophobic contacts with L1122 at the end of the β1-strand, A1148 (CS7), V1180 (Sh1), the gatekeeper L1196 (Sh2), L1182 within the hinge, and L1256 (CS6). L1122 and A1148 occur above the plane of the scaffold while L1256 occurs below the plane. These residues are found within the front pocket and BP-I-B in the gate area [48]. The 4-morpholin-4-piperidine group extends from the enzyme into the solvent. The αC-E1167 forms a hydrogen bond with the β3-

K1150, the R-spine is linear, and DFG-D1270 is pointed toward the active site, properties associated with an active enzyme form. However, the activation segment has an αAL helix indicating that the segment is in the compact conformation of a dormant enzyme. Binding to the front pocket of an inactive enzyme with the DFG-D in conformation indicates that alectinib is a type I½ B inhibitor. Alectinib is a second-line treatment that is FDA-approved for the treatment of crizotinib-resistant ALK^{+}-NSCLC (www.brimr.org/PKI/PKIs.htm).

Brigatinib is an ALK antagonist with a bisanilinopyrimidine scaffold and piperazine and piperidine extensions (Fig. 5D). This drug is effective against most crizotinib-resistant mutants including C1156Y, F1174L, L1196M (the gatekeeper), G1202R, and R1275Q [59]. The *N*3 of the pyrimidine forms a hydrogen bond with the N–H group of M1199 of the hinge while the phenylamino group forms a hydrogen bond with the M1199 carbonyl group (Fig. 6D). The drug makes hydrophobic contacts with L1122 before the glycine-rich loop, V1130 (CS8) after the loop, A1148 (CS7), L1196/1198 of the hinge, L1256/1257 (CS6/4), and V1265 before the activation segment. The 5-chloropyridine component of brigatinib makes hydrophobic contacts with A1148 in the ceiling of the cleft and with L1256 in the floor. The methoxy group binds under hinge residue L1198 and the piperazine-piperidine solubilization moieties extend from the enzyme into the solvent. Owing to an intramolecular hydrogen bond between the scaffold and the C4 aniline N–H group, brigatinib pays a reduced entropic penalty during the formation of the drug-enzyme complex. The drug is found in the adenine pocket and the adjacent FP-I region of the front cleft [48]. The binding pattern of brigatinib with

ALK classifies it as a type I½ B antagonist. This medicinal is in four clinical trials for the treatment of ALK^+-NSCLC, but it has not been FDA-approved.

Entrectinib, which is a brain penetrant type I½ B ALK inhibitor, contains an indazole scaffold (Fig. 5E) [60]. The drug is effective against the ALK crizotinib-resistant L1196M gatekeeper mutation, but not against the G1202R mutant. The N–H group of the indazole forms a hydrogen bond with the carbonyl group of E1197 while its *N*1 forms a hydrogen bond with M1199, both residues of which occur within the hinge. The amino group attached to the 3-position of indazole forms a hydrogen bond with the carbonyl group of M1199. The antagonist makes numerous hydrophobic contacts with ALK including L1122 before the glycine-rich loop, F1127 within the G-rich loop, V1130 (CS8) in the β1-strand following the G-rich loop, A1148 (CS7), V1180 (Sh1), I1183 in the αC-β4 loop, the gatekeeper L1196 (Sh2), L1198 within the hinge, the catalytic loop N1254, and the ^{1255}CLL^{1257} triad (CS5/6/4) after the catalytic loop. The indazole group makes hydrophobic contact with V1130 and A1148 in the ceiling of the cleft and with L1256 in the floor. The 3,5-difluorophenyl group of the medicinal makes hydrophobic contact with F1127 in front of the aromatic group, with D1203 at its bottom, and C1255 and L1256 at its rear (as protein kinases are classically viewed). The drug also makes van der Waals contact with DFG-D1270. The tetrahydropyran and the *N*-methylpiperazinyl solubilizing group extend from the protein into the solvent. The drug occupies the adenine pocket and the adjacent FP-I region of the front cleft [48] and is classified as a type I½ B antagonist. Entrectinib is in four clinical trials against malignancies driven by its target kinases including Trk1/2/3, ROS1, and ALK [7].

Lorlatinib is an orally active brain penetrant cyclic 2-aminopyridine derivative that is a type I½ B ALK inhibitor (Fig. 5F) [61]. This medicinal is an effective antagonist against the more common L1196M and G1269A crizotinib-resistant mutations as well as the less common T1151Ins, L1152R, C1156Y, F1174L, and S1206Y mutants [62]. The X-ray structure shows that the 2-amino group forms a hydrogen bond with Ell97 of the hinge and the *N*1 of the pyridine ring forms a hydrogen bond with the N–H group of M1199. Note that these interactions with the first and third hinge residues correspond to the mode of interaction of the adenine group of ADP with ALK (Fig. 2A). The drug makes hydrophobic contact with L1122 immediately before the glycine-rich loop, V1130 (CS8) near the beginning of the β2-strand, A1148 (CS7), V1180 (Sh1), L1196/1198 within the hinge, N1254 at the end of the catalytic loop, and L1256 (CS6). The pyridine group makes hydrophobic contact with A1148 in the ceiling and L1256 of the floor of the cleft while the pyrazole group makes hydrophobic contact with L1122 in the ceiling. The fluorobenzene component of the drug makes van der Waals contact with R1253 and N1254 and hydrophobic contact with L1256, all in the floor of the cleft. Lorlatinib also makes van der Waals contact with DFG-D1270. The drug occupies the adenine pocket and the adjacent FP-I region of the front cleft [48]. The pyrazole with its methyl and cyano attachments extend from the enzyme into the solvent. Lorlatinib is in three clinical trials in patients with ALK^+ or $ROS1^+$ lung cancer including treatment of related brain metastases. Lung cancer accounts for nearly one-half of all cancer metastases to the brain and lorlatinib was fabricated to traverse the blood brain barrier [61].

4. Mechanisms of ALK-based drug resistance

The replacement of the ALK methionine gatekeeper with leucine does not confer crizotinib resistance by blocking access to the back pocket as seen in the X-ray structure of the drug bound to the L1196M G1269A double mutant (PDB ID: 4ANS). Lovly et al. observed that the EML4-ALK 1196M fusion protein exhibits greater cellular baseline levels of phosphorylation than the wild type protein indicating that this gatekeeper mutation leads to increased protein kinase activity [63]. This observation suggests that the replacement of leucine with methionine stabilizes the ALK active conformation and leads to drug resistance. Recall that the gatekeeper (Sh2) occurs near the top of the regulatory spine, and it is likely that the replacement of leucine with methionine strengthens the R-spine and tips the equilibrium in favor of active ALK. Accordingly, crizotinib resistance results from enzyme activation and not from steric hindrance or to the blockade of the back pocket by the gatekeeper.

Several other ALK crizotinib-resistant mutations have been identified in tumor samples from patients with NSCLC [64]. Val1180 is the Sh1 residue that occurs in the αC-β4 loop (Fig. 7). Its mutation to the larger leucine (V1180L) may hinder drug binding. Asp1203, which is in direct contact with crizotinib, is on the floor of the drug binding site. Although the size difference between aspartate and the mutant asparagine (D1203N) is marginal, models indicate that the asparagine obstructs crizotinib binding. The S1206Y mutation is located on the floor of the crizotinib binding site and the conversion of the small serine to the larger tyrosine may block its binding. Similarly, the G1269A mutation occurs on the floor of the drug binding site immediately before ^{1270}DFG^{1272} and the larger alanine may prevent enzyme-drug complex formation. The I1171T is the RS3 residue that occurs within the αC-helix; perhaps this mutation produces a structural change that

strengthens the R-spine leading to a more active enzyme. However, this suggestion is counter intuitive because threonine is less hydrophobic than isoleucine. The L1152R mutation occurs in the β3-strand, the C1156Y mutation occurs within the β3–αC loop just before the αC-helix, the F1174L mutation occurs within the αC-helix, and the S1206Y and E1210K mutations occur within the αD-helix; in each of these cases the residues are too far to directly block crizotinib binding. The mechanism for conferring resistance to crizotinib is unclear in these situations, but it may result from the destabilization of the inactive enzyme conformation that readily forms a complex with the drug.

Several ceritinib-resistant mutations following the second-line treatment of ALK^+-NSCLC have been discovered [64]. These mutations include C1156Y, F1174L, and L1152R, which were first observed in crizotinib-resistant tumor samples. The G1202R ceritinib-resistant mutation occurs within the hinge and is in close proximity to the drug; the conversion of the smaller to larger residue may directly impede drug binding. The I1171T and V1180L alectinib-resistant mutations were previously found in crizotinib-resistant tissue samples. Moreover, an additional I1171N alectinib-resistant mutation has been reported. Although clinical experience with lorlatinib is limited, an L1198F resistance mutation has been observed. This residue occurs within the hinge and the larger phenylalanine most likely hinders ALK-lorlatinib complex formation. See Ref. [64] for a list of all of the ALK drug resistant mutations.

5. Classification of protein kinase-drug complexes

Historically Dar and Shokat defined three classes of small molecule protein kinase inhibitor: types I, II, and III [65]. The type I inhibitor is a drug that binds within the ATP pocket of the active conformation of the kinase, the type II inhibitor binds to an

inactive (usually DFG-out) conformation of the enzyme while the type III inhibitor is an allosteric or a non-ATP competitive antagonist. Allosteric inhibitors bind to a site distinct from the active site [66] and this refers to ligands that bind outside of the ATP-binding pocket of protein kinases. Zuccotto classified type I½ inhibitors as ligands that bind to inactive protein kinases with the DFG-Asp directed inward toward the active site (in contrast to the DFG-Asp out structure) [67]. Gavrin and Saiah subdivided allosteric antagonists into two classes: III and IV [68]. Accordingly, type III inhibitors bind within the cleft between the amino-terminal and carboxyterminal lobes next to, but separate from, the ATP binding pocket while type IV inhibitors bind outside of the cleft. Lamba and Gosh classified bivalent molecules as those antagonists which span two regions of the protein kinase domain as type V inhibitors [69]. To complete this classification we labeled small molecules that form covalent adducts with the target enzyme as type VI inhibitors [44].

We previously divided the type I½ and type II inhibitors into A and B subtypes [44]. The binding of imatinib to Abl is a prototypical type II inhibitor [67]. This drug binds to Abl with the DFG-D out configuration and extends into the back cleft [44] and we classified drugs that extend into the back cleft as type A inhibitors. In contrast, we classified drugs such as bosutinib that bind to the DFG-D out conformation of Abl while not extending into the back cleft as type B inhibitors. Based upon limited data, the practical consequence of this distinction is that type A antagonists bind to the their target kinase with a long residence time when compared with type B inhibitors.

6. Epilogue

In addition to small molecule inhibitors, two monoclonal antibodies (large molecule inhibitors) that directly or indirectly target VEGFR have been approved for the treatment of NSCLC. Ramucirumab binds to VEGFR2 and prevents its activation by ligands, thus inhibiting angiogenesis [70]. The FDA approved ramucirumab in 2014 in combination with docetaxel for the treatment of metastatic NSCLC with disease progression on platinum-based cytotoxic chemotherapy. In 2006, the FDA approved bevacizumab for use as a first-line treatment of metastatic nonsquamous NSCLC in combination with carboplatin/paclitaxel chemotherapy [71]. This drug binds to VEGF-A and prevents the activation of VEGFR thereby inhibiting angiogenesis [72].

The US FDA has approved 29 small molecule drugs that bind directly to the intracellular protein kinase domain (www.brimr.org/PKI/PKIs.htm) and the year that each of these drugs was approved is indicated in Fig. 8. Note that the FDA approved from 2–5 new drugs in the years from 2011–2015. However, no new small molecule protein kinase antagonists were approved in 2016, but new indications or modifications were reported in 2016 for afatinib, cabozantinib, crizotinib, and palbociclib. Table 4 provides an updated and corrected classification of FDA-approved protein kinase inhibitors based upon the structure of the drug-enzyme complexes as we first reported [44]; the newly added drugs include alectinib, cobimetinib, dabrafenib, ibrutinib, and trametinib. Of the 29 approved drugs, we are lacking X-ray crystal structures of cabozantinib, osimertinib, pazopanib, and regorafenib bound to their drug targets and their consequent structure-based inhibitor classification.

Imatinib was first approved for the treatment of Philadelphia chromosome positive chronic myelogenous leukemia in 2001[25]. This first approved small molecule

antagonist is quite unusual in that it provides a durable response that lasts for more than a decade in the majority of patients. Imatinib has substantially greater beneficial properties when compared with all of the other FDA-approved small molecule protein kinase inhibitors in terms of (i) its effectiveness with greater than 95% complete hematological response against its target disease, i.e., chronic myelogenous leukemia [74] and (ii) the comparative lack of resistance when contrasted with other medications and neoplastic diseases. The development of imatinib resistance owing to mutations in BCR-Abl has lead to the development of second generation drugs that are effective against such mutations including bosutinib, dasatinib, nilotinib, and ponatinib. In contrast to imatinib, resistance to crizotinib occurs in more than 90% of patients within one year as described above and resistance is a grave problem for all targeted and cytotoxic anticancer drugs. As noted by Winer et al. “Biologically, the cancer cell is notoriously wily; each time we throw an obstacle in its path, it finds an alternate route that must then be blocked” [75].

One conundrum in the treatment of ALK^+-NSCLC is whether to begin therapy with crizotinib or with one of the second generation of drugs. The goal is to maximize the duration of therapeutic efficacy. Whether the therapeutic effectiveness is of greater duration with initial crizotinib followed by a second generation ALK inhibitor or whether it is better to begin therapy with a second generation inhibitor remains to be established. Another possibility is to begin the initial treatment of NSCLC with a combination of inhibitors, a strategy that is effective in the treatment of HIV/AIDS [76,77]. Additionally, myeloid-derived suppressor cells (MDSCs) have been implicated in various tumor immune escape mechanisms and targeting these cells alone or in combination with other

treatment modalities has the potential to augment the efficacy of lung cancer treatment protocols [78].

Only a third of the patients with acquired crizotinib-resistance have ALK mutations, another third of patients exhibits the activation of bypass pathways, and the mechanism of resistance is unknown in the final third of patients. Activation of EGFR, MEK1/2, ERK1/2, Src, MET (HGFR), PI3-kinase, or the insulin-like growth factor-1 receptor along with Kit amplification and increased production of stem cell factor (Kit ligand) are among the known activated bypass pathways [53, 63]. Because activation of EGFR appears to be the most important bypass mechanism, concomitant treatment with a combination of ALK and EGFR inhibitors might represent an effective first-line treatment strategy. Although great strides in the treatment of NSCLC have been made in the last decade using targeted protein kinase inhibitors and immune checkpoint inhibitors, there is still a large unmet clinical need in the treatment of such patients.

Conflict of interest

The author is unaware of any affiliations, memberships, or financial holdings that might be perceived as affecting the objectivity of this review.

Acknowledgment

The author thanks Laura M. Roskoski for providing editorial and bibliographic assistance.

References

[1] Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. CA Cancer J Clin 2016;66:7–30.

[2] Bugano Diniz Gomes D. Gold KA, Gibbons DL. Non-small cell lung cancer. in Kantarjian HM, Wolff RA (Eds.), The MD Anderson Manual of Medical Oncology, third ed., McGraw-Hill Education, New York, 2016, 343–376.

[3] Cascone T, Gold KA, Glisson BS. Small cell carcinoma of the lung, in Kantarjian HM, Wolff RA (Eds.), The MD Anderson Manual of Medical Oncology, third ed., McGraw-Hill Education, New York, 2016, 323–342.

[4] Roskoski R Jr. ErbB/HER protein-tyrosine kinases: Structures and small molecule inhibitors. Pharmacol Res 2014;79:34–74.

[5] Dholaria B, Hammond W, Shreders A, Lou Y. Emerging therapeutic agents for lung cancer. J Hematol Oncol 2016;9:138.

[6] Shaw AT, Yeap BY, Mino-Kenudson M, Digumarthy SR, Costa DB, Heist RS, et al. Clinical features and outcome of patients with non-small-cell lung cancer who harbor EML4-ALK. J Clin Oncol 2009;27:4247–53.

[7] Sullivan I, Planchard D. ALK inhibitors in non-small cell lung cancer: the latest evidence and developments. Ther Adv Med Oncol 2016;8:32–47.

[8] Morris SW, Kirstein MN, Valentine MB, Dittmer K, Shapiro DN, Look AT, Saltman DL. Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in non-Hodgkin's lymphoma. Science 1995;267:316–7.

[9] Shiota M, Fujimoto J, Semba T, Satoh H, Yamamoto T, Mori S. Hyperphosphorylation of a novel 80 kDa protein-tyrosine kinase similar to Ltk in a human Ki-1 lymphoma cell line, AMS3. Oncogene1994;9:1567–74.

[10] Soda M, Choi YL, Enomoto M, Takada S, Yamashita Y, Ishikawa S, et al. Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. Nature 2007;448:561–6.

[11] Rikova K, Guo A, Zeng Q, Possemato A, Yu J, Haack H, et al. Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer. Cell 2007;131:1190–203.

[12] Roskoski R Jr. Anaplastic lymphoma kinase (ALK): structure, oncogenic activation, and pharmacological inhibition. Pharmacol Res 2013;68:68–94.

[13] Roskoski R Jr. The preclinical profile of crizotinib for the treatment of non-small-cell lung cancer and other neoplastic disorders. Expert Opin Drug Discov 2013;8:1165–79.

[14] Morris SW, Naeve C, Mathew P, James PL, Kirstein MN, Cui X, Witte DP. ALK, the chromosome 2 gene locus altered by the t(2;5) in non-Hodgkin's lymphoma, encodes a novel neural receptor tyrosine kinase that is highly related to leukocyte tyrosine kinase (LTK). Oncogene 1997;14:2175–88. Erratum in: Oncogene 1997;15:2883.

[15] Lorén CE, Englund C, Grabbe C, Hallberg B, Hunter T, Palmer RH. A crucial role for the Anaplastic lymphoma kinase receptor tyrosine kinase in gut development in *Drosophila melanogaster*. EMBO Reports 2003;4:781–6.

[16] Iwahara T, Fujimoto J, Wen D, Cupples R, Bucay N, Arakawa T, et al. Molecular characterization of ALK, a receptor tyrosine kinase expressed specifically in the nervous system. Oncogene 1997;14:439–49.

[17] Vernersson E, Khoo NK, Henriksson ML, Roos G, Palmer RH, Hallberg B. Characterization of the expression of the ALK receptor tyrosine kinase in mice. Gene Expression Patterns 2006;6:448–61.

[18] Guan J, Umapathy G, Yamazaki Y, Wolfstetter G, Mendoza P, Pfeifer K, et al. FAM150A and FAM150B are activating ligands for anaplastic lymphoma kinase. Elife. 2015;4:e09811.

[19] Reshetnyak AV, Murray PB, Shi X, Mo ES, Mohanty J, Tome F, et al. Augmentor α and β (FAM150) are ligands of the receptor tyrosine kinases ALK and LTK: Hierarchy and specificity of ligand-receptor interactions. Proc Natl Acad Sci U S A 2015;112:15862–7.

[20] Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S. The protein kinase complement of the human genome. Science 2002;298:1912–34.

[21] Hanks SK, Hunter T. Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. FASEB J 1995;9:576–96.

[22] Knighton DR, Zheng JH, Ten Eyck LF, Ashford VA, Xuong NH, Taylor SS, et al. Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. Science 1991;253:407–14.

[23] Knighton DR, Zheng JH, Ten Eyck LF, Xuong NH, Taylor SS, Sowadski JM. Structure of a peptide inhibitor bound to the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. Science 1991;253:414–20.

[24] Taylor SS, Kornev AP. Protein kinases: evolution of dynamic regulatory proteins. Trends Biochem Sci 2011;36:65–77.

[25] Roskoski R Jr. A historical overview of protein kinases and their targeted small molecule inhibitors. Pharmacol Res 2015;100:1–23.

[26] Lee CC, Jia Y, Li N, Sun X, Ng K, Ambing E, et al. Crystal structure of the ALK (anaplastic lymphoma kinase) catalytic domain. Biochem J 2010;430:425–37.

[27] Roskoski R Jr. Src protein-tyrosine kinase structure, mechanism, and small molecule inhibitors. Pharmacol Res 2015;94:9–25.

[28] Taylor SS, Keshwani MM, Steichen JM, Kornev AP. Evolution of the eukaryotic protein kinases as dynamic molecular switches. Philos Trans R Soc Lond B Biol Sci 2012;367:2517–28.

[29] Nolen B, Taylor S, Ghosh G. Regulation of protein kinases; controlling activity through activation segment conformation. Molecular Cell 2004;15:661–75.

[30] Rawlings DJ, Scharenberg AM, Park H, Wahl MI, Lin S, Kato RM, et al. Activation of BTK by a phosphorylation mechanism initiated by SRC family kinases. Science 1996;271:822–5.

[31] Wei L, Hubbard SR, Hendrickson WA, Ellis L. Expression, characterization, and crystallization of the catalytic core of the human insulin receptor protein-tyrosine kinase domain. J Biol Chem 1995;270:8122–30.

[32] Favelyukis S, Till JH, Hubbard SR, Miller WT. Structure and autoregulation of the insulin-like growth factor 1 receptor kinase. Nat Struct Biol 2001;8:1058–63.

[33] Lemmon MA, Schlessinger J. Cell signaling by receptor tyrosine kinases. Cell 2010;141:1117–34.

[34] Tartari CJ, Gunby RH, Coluccia AM, Sottocornola R, Cimbro B, Scapozza L, et al. Characterization of some molecular mechanisms governing autoactivation of the catalytic domain of the anaplastic lymphoma kinase. J Biol Chem 2008;283:3743–50.

[35] Roskoski Jr R. Structure and regulation of Kit protein-tyrosine kinase--the stem cell factor receptor. Biochem Biophys Res Commun 2005;338:1307–15.

[36] Kornev AP, Haste NM, Taylor SS, Ten Eyck LF. Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism. Proc Natl Acad Sci U S A 2006;103:17783–8.

[37] Kornev AP, Taylor SS, Ten Eyck LF. A helix scaffold for the assembly of active protein kinases. Proc Natl Acad Sci U S A 2008;105:14377–82.

[38] Roskoski R Jr. Cyclin-dependent protein kinase inhibitors including palbociclib as anticancer drugs. Pharmacol Res 2016;111:784–803.

[39] Roskoski R Jr. ERK1/2 MAP kinases: structure, function, and regulation. Pharmacol Res 2012;66:105–43.

[40] Roskoski R Jr. Janus kinase (JAK) inhibitors in the treatment of inflammatory and neoplastic diseases. Pharmacol Res 2016;111:784–803.

[41] Roskoski R Jr. MEK1/2 dual-specificity protein kinases: structure and regulation. Biochem Biophys Res Commun 2012;417:5–10.

[42] Roskoski R Jr. Allosteric MEK1/2 inhibitors including cobimetanib and trametinib in the treatment of cutaneous melanomas. Pharmacol Res 2017;117:20–31.

[43] Meharena HS, Chang P, Keshwani MM, Oruganty K, Nene AK, Kannan N, et al. Deciphering the structural basis of eukaryotic protein kinase regulation. PLoS Biol 2013;11:e1001680.

[44] Roskoski R Jr. Classification of small molecule protein kinase inhibitors based upon the structures of their drug-enzyme complexes. Pharmacol Res 2016;103:26–48.

[45] Shah K, Liu Y, Deirmengian C, Shokat KM. Engineering unnatural nucleotide specificity for Rous sarcoma virus tyrosine kinase to uniquely label its direct substrates. Proc Natl Acad Sci U S A 1997;94:3565–70.

[46] Liu Y, Shah K, Yang F, Witucki L, Shokat KM. A molecular gate which controls unnatural ATP analogue recognition by the tyrosine kinase v-Src. Bioorganic Med Chem 1998;6:1219–26.

[47] Liao JJ. Molecular recognition of protein kinase binding pockets for design of potent and selective kinase inhibitors. J Med Chem 2007;50:409–24.

[48] van Linden OP, Kooistra AJ, Leurs R, de Esch IJ, de Graaf C. KLIFS: a knowledge-based structural database to navigate kinase-ligand interaction space. J Med Chem 2014;57:249–77.

[49] Shaw AT, Friboulet L, Leshchiner I, Gainor JF, Bergqvist S, Brooun A, et al. Resensitization to Crizotinib by the Lorlatinib ALK Resistance Mutation L1198F. N Engl J Med 2016;374:54–61.

[50] Fontana D, Ceccon M, Gambacorti-Passerini C, Mologni L. Activity of second-generation ALK inhibitors against crizotinib-resistant mutants in an NPM-ALK model compared to EML4-ALK. Cancer Med 2015;4:953–65.

[51] Cui JJ, Tran-Dubé M, Shen H, Nambu M, Kung PP, Pairish M, et al. Structure based drug design of crizotinib (PF-02341066), a potent and selective dual inhibitor of mesenchymal-epithelial transition factor (c-MET) kinase and anaplastic lymphoma kinase (ALK). Journal of Medicinal Chemistry 2011;54:6342–63.

[52] Bergethon K, Shaw AT, Ou SH, Katayama R, Lovly CM, McDonald NT, et al. ROS1 rearrangements define a unique molecular class of lung cancers. J Clin Oncol 2012;30:863–70.

[53] Katayama R, Shaw AT, Khan TM, Mino-Kenudson M, Solomon BJ, Halmos B, et al. Mechanisms of acquired crizotinib resistance in ALK-rearranged lung cancers. Sci Transl Med 2012;4:120ra17.

[54] Dagogo-Jack I, Shaw AT. Crizotinib resistance: implications for therapeutic strategies. Ann Oncol 2016;27 Suppl 3:iii42–iii50.

[55] Gainor JF, Dardaei L, Yoda S, Friboulet L, Leshchiner I, Katayama R, et al. Cancer Discov 2016;6:1118–1133.

[56] Friboulet L, Li N, Katayama R, Lee CC, Gainor JF, Crystal AS, et al. The ALK inhibitor ceritinib overcomes crizotinib resistance in non-small cell lung cancer. Cancer Discov 2014;4:662–73.

[57] Marsilje TH, Pei W, Chen B, Lu W, Uno T, Jin Y, et al. Synthesis, structure-activity relationships, and in vivo efficacy of the novel potent and selective anaplastic lymphoma kinase (ALK) inhibitor 5-chloro-N2-(2-isopropoxy-5-methyl-4-(piperidin-4-yl)phenyl)-N4-(2-(isopropylsulfonyl)phenyl)pyrimidine-2,4-diamine (LDK378) currently in phase 1 and phase 2 clinical trials. J Med Chem 2013;56:5675–90.

[58] Kinoshita K, Asoh K, Furuichi N, Ito T, Kawada H, Hara S, Ohwada J, et al. Design and synthesis of a highly selective, orally active and potent anaplastic lymphoma kinase inhibitor (CH5424802). Bioorg Med Chem 2012;20:1271–80.

[59] Zhang S, Anjum R, Squillace R, Nadworny S, Zhou T, Keats J, et al. The Potent ALK Inhibitor Brigatinib (AP26113) Overcomes Mechanisms of Resistance to First- and

Second-Generation ALK Inhibitors in Preclinical Models. Clin Cancer Res 2016 [Epub ahead of print].

[60] Menichincheri M, Ardini E, Magnaghi P, Avanzi N, Banfi P, Bossi R, et al. Discovery of entrectinib: a new 3-aminoindazole as a potent anaplastic lymphoma kinase (ALK), c-ros oncogene 1 kinase (ROS1), and pan-tropomyosin receptor kinases (pan-TRKs) inhibitor. J Med Chem. 2016;59:3392–408.

[61] Johnson TW, Richardson PF, Bailey S, Brooun A, Burke BJ, Collins MR, et al. Discovery of (10R)-7-amino-12-fluoro-2,10,16-trimethyl-15-oxo-10,15,16,17-tetrahydro-2H-8,4-(metheno)pyrazolo[4,3-h][2,5,11]-benzoxadiazacyclotetradecine-3-carbonitrile (PF-06463922), a macrocyclic inhibitor of anaplastic lymphoma kinase (ALK) and c-ros oncogene 1 (ROS1) with preclinical brain exposure and broad-spectrum potency against ALK-resistant mutations. J Med Chem 2014;57:4720–44.

[62] Zou HY, Friboulet L, Kodack DP, Engstrom LD, Li Q, West M, et al. PF-06463922, an ALK/ROS1 inhibitor, overcomes resistance to first and second generation ALK inhibitors in preclinical models. Cancer Cell 2015;28:70–81.

[63] Lovly CM, Heuckmann JM, de Stanchina E, Chen H, Thomas RK, Liang C, Pao W. Insights into ALK-driven cancers revealed through development of novel ALK tyrosine kinase inhibitors. Cancer Research 2011;71:4920–31.

[64] Qian M, Zhu B, Wang X, Liebman M. Drug resistance in ALK-positive non-small cell lung cancer patients. Semin Cell Dev Biol 2016, http://dx.doi.org/10.1016/j.semcdb.2016.09.016.

[65] Dar AC, Shokat KM. The evolution of protein kinase inhibitors from antagonists to agonists of cellular signaling. Annu Rev Biochem 2011;80:769–95.

[66] Monod J, Changeux JP, Jacob F. Allosteric proteins and cellular control systems. J Mol Biol 1963;6:306–29.

[67] Zuccotto F, Ardini E, Casale E, Angiolini M. Through the "gatekeeper door": exploiting the active kinase conformation. J Med Chem 2010;53:2681–94.

[68] Gavrin LK, Saiah E. Approaches to discover non-ATP site inhibitors. Med Chem Commun 2013;4:41.

[69] Lamba V, Ghosh I. New directions in targeting protein kinases: focusing upon true allosteric and bivalent inhibitors. Curr Pharm Des 2012;18:2936–45.

[70] Kurzrock R, Stewart DJ. Exploring the benefit/risk associated with antiangiogenic agents for the treatment of non-small cell lung cancer patients. Clin Cancer Res 2016; clincanres.1968.2016. [Epub ahead of print].

[71] Sliwkowski MX, Mellman I. Antibody therapeutics in cancer. Science 2013 341;1192–1198.

[72] Roskoski R Jr. Vascular endothelial growth factor (VEGF) signaling in tumor progression. Crit Rev Oncol Hematol 2007;62:179–213.

[73] Fischer PM. Approved and Experimental Small-Molecule Oncology Kinase Inhibitor Drugs: A Mid-2016 Overview. Med Res Rev 2016. doi: 10.1002/med.21409. [Epub ahead of print].

[74] Kantarjian H, Sawyers C, Hochhaus A, Guilhot F, Schiffer C, Gambacorti-Passerini C, et al. Hematologic and cytogenetic responses to imatinib mesylate in chronic myelogenous leukemia. N Engl J Med 2002;346:645–52. Erratum in: N Engl J Med 2002;346:1923.

[75] Winer E, Gralow J, Diller L, Karlan B, Loehrer P, Pierce L, et al. Clinical cancer advances 2008: major research advances in cancer treatment, prevention, and screening—a report from the American Society of Clinical Oncology. J Clin Oncol 2009;27:812–26. Erratum in: J Clin Oncol 2009;27:3070–1.

[76] Venning GR, Scott GM. New drug treatments for HIV infection and AIDS. A review following the Seventh International Conference on AIDS, Florence 16-21 June 1991. Br J Clin Pharmacol 1992;33:349–56.

[77] Das K, Arnold E. HIV-1 reverse transcriptase and antiviral drug resistance. Part 2. Curr Opin Virol 2013;3:119–28.

[78] Adah D, Hussain M, Qin L, Qin L, Zhang J, Chen X. Implications of MDSCs-targeting in lung cancer chemo-immunotherapeutics. Pharmacol Res 2016;110:25–34.

Figure legends

Fig. 1 (A) Architecture of the anaplastic large cell lymphoma kinase NPM-ALK fusion protein. (B) Structure of the EML4-ALK fusion protein (variant 1) in NSCLC. (C) Overview of the native human ALK receptor tyrosine protein kinase. The extracellular segment of ALK (residues 19–1038) contains an LDLa domain (453–471) sandwiched between two MAM domains (264–427 and 480–626) along with a glycine-rich component. The transmembrane (TM) segment (residues 1039–1059) connects the extracellular segment with the intracellular juxtamembrane (JM) component (1060–1620). The protein-tyrosine kinase domain consists of residues 1116–1392, which is followed by a C-terminal tail. (D) A close up view of the intracellular portion of ALK depicting various tyrosine phosphorylation sites within specific segments.

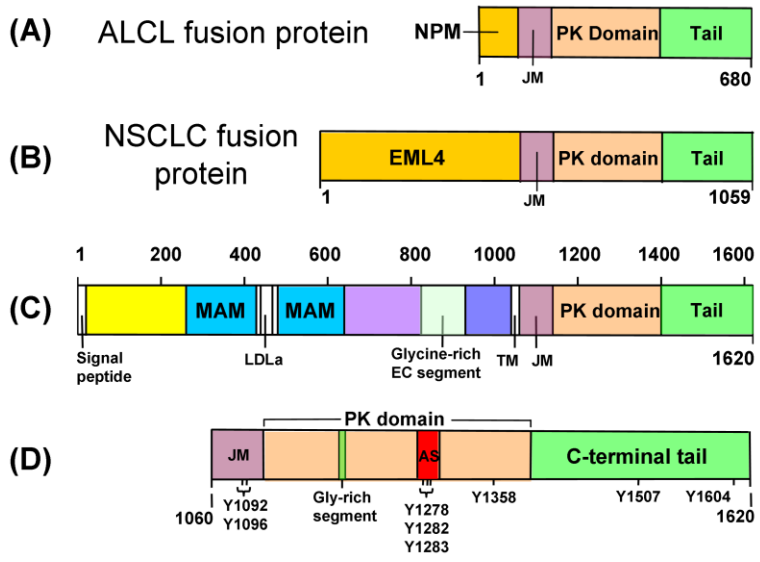
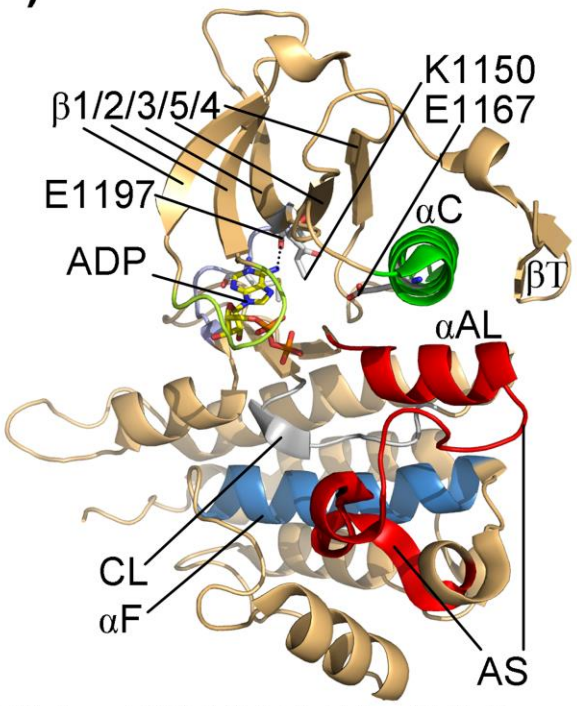


Fig. 2. (A) Ribbon diagram structure of the dormant ALK catalytic domain. (B) Structure of the active InsRK catalytic domain. αAL, the proximal activation loop helix; AS, activation segment; βT, N-terminal β-turn; CL, catalytic loop. Figs. 2, 3, 4 and 6 were prepared using the PyMOL Molecular Graphics System Version 1.5.0.4 Schrödinger, LLC.

(A) ALK PDBID: 3LCT



(B) InsRK PDB ID: 4XVL

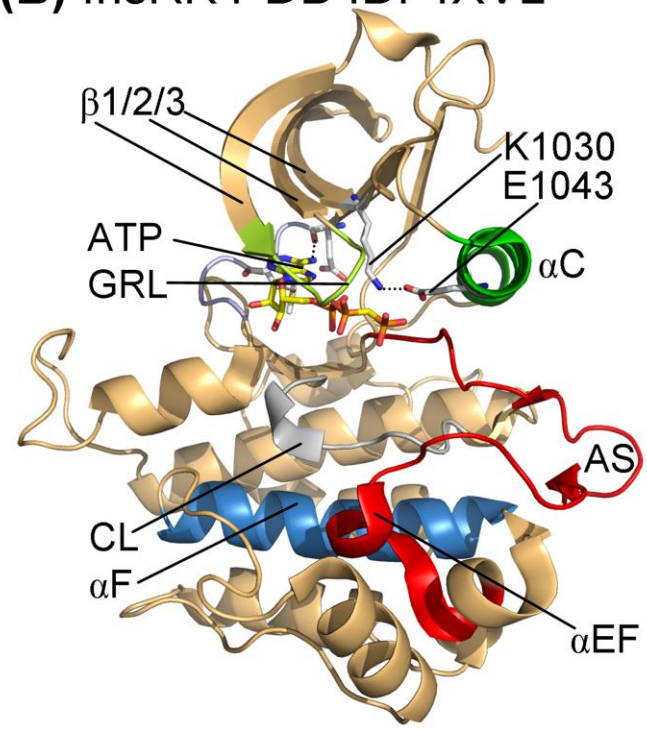


Fig. 3. Inferred mechanism of the ALK reaction based upon the structure of the InsRK as prepared from PDB ID: 4XVL. The chemistry occurs within the blue circle.

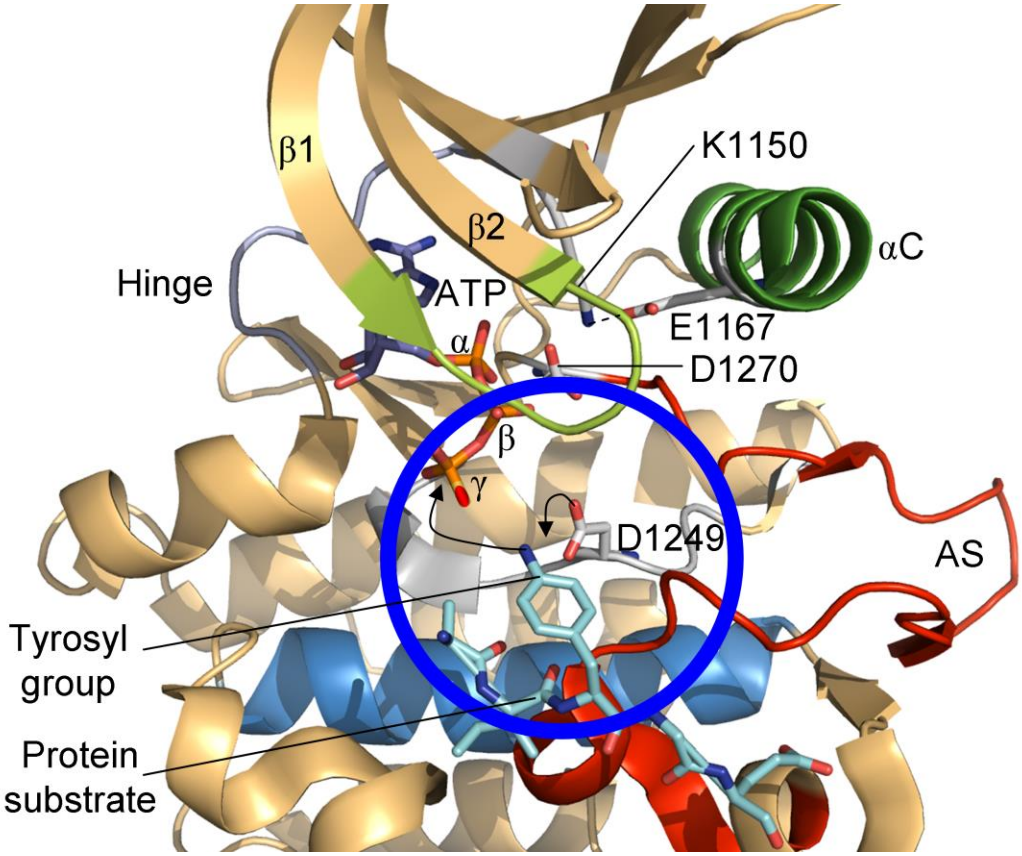


Fig. 4. Catalytic and regulatory spine structures. (A) Catalytic spine (CS), regulatory spine (RS), and shell (Sh) residue labels as observed from the classical frontal view of protein kinases. (B) Spines of dormant human ALK (C) Spines of active human InsRK. (D) Superposition of the dormant ALK spine residues (gray) and active InsRK spine residues (dark blue). Ad, adenine.

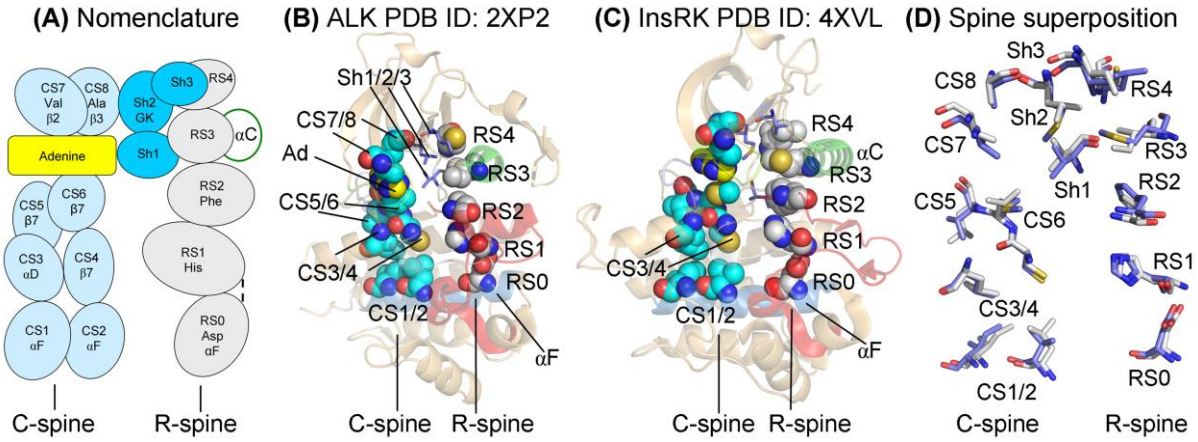


Fig. 5. Chemical structure of selected ALK antagonists.

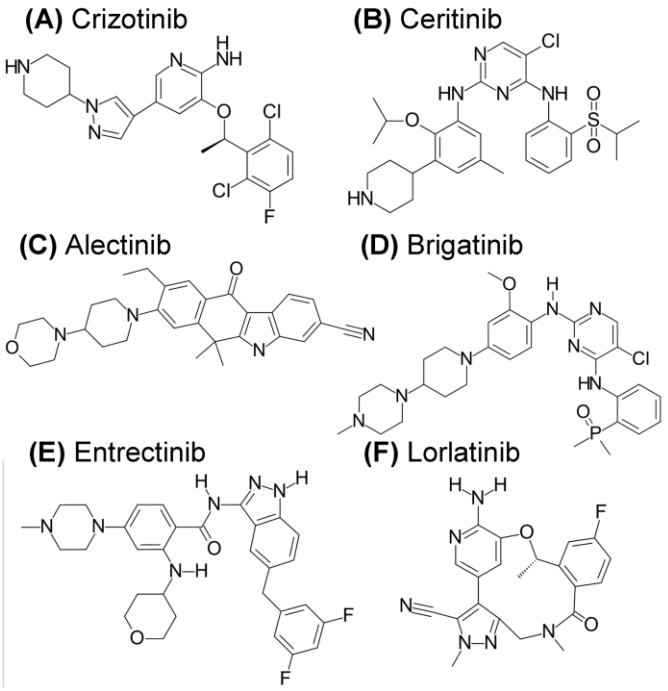


Fig. 6. (A–F) Human ALK-drug complexes. Dashed lines indicate polar bonds. The protein data bank identification numbers are listed. AS, activation segment; CL, catalytic loop; GK, gatekeeper L1196.

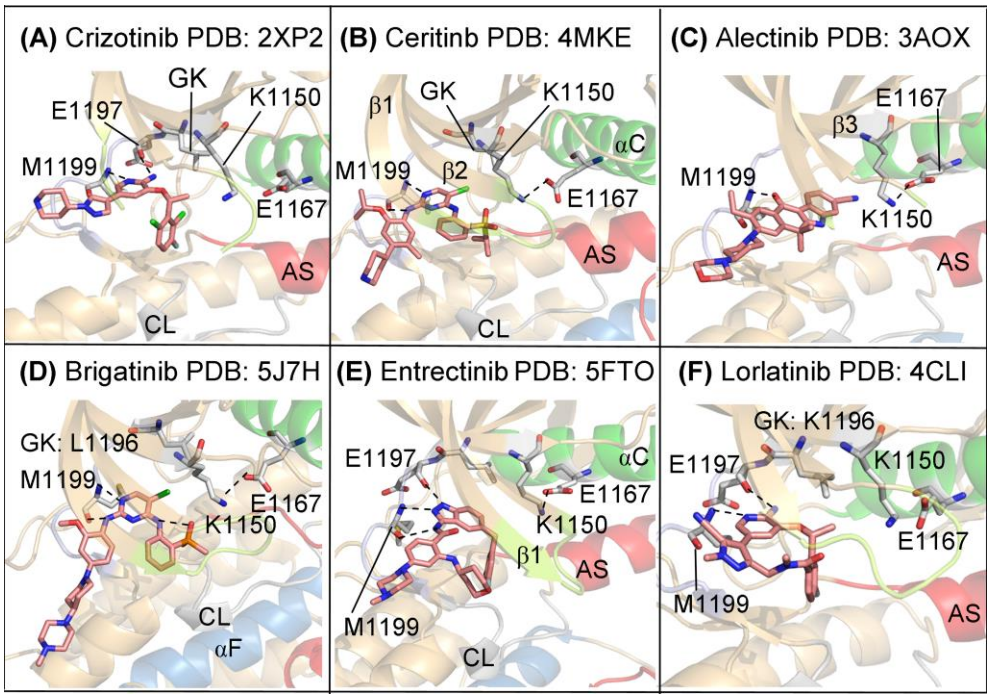


Fig. 7. Location of ALK drug-resistant mutations. AS, activation segment; βT, β-turn; CL, catalytic loop.

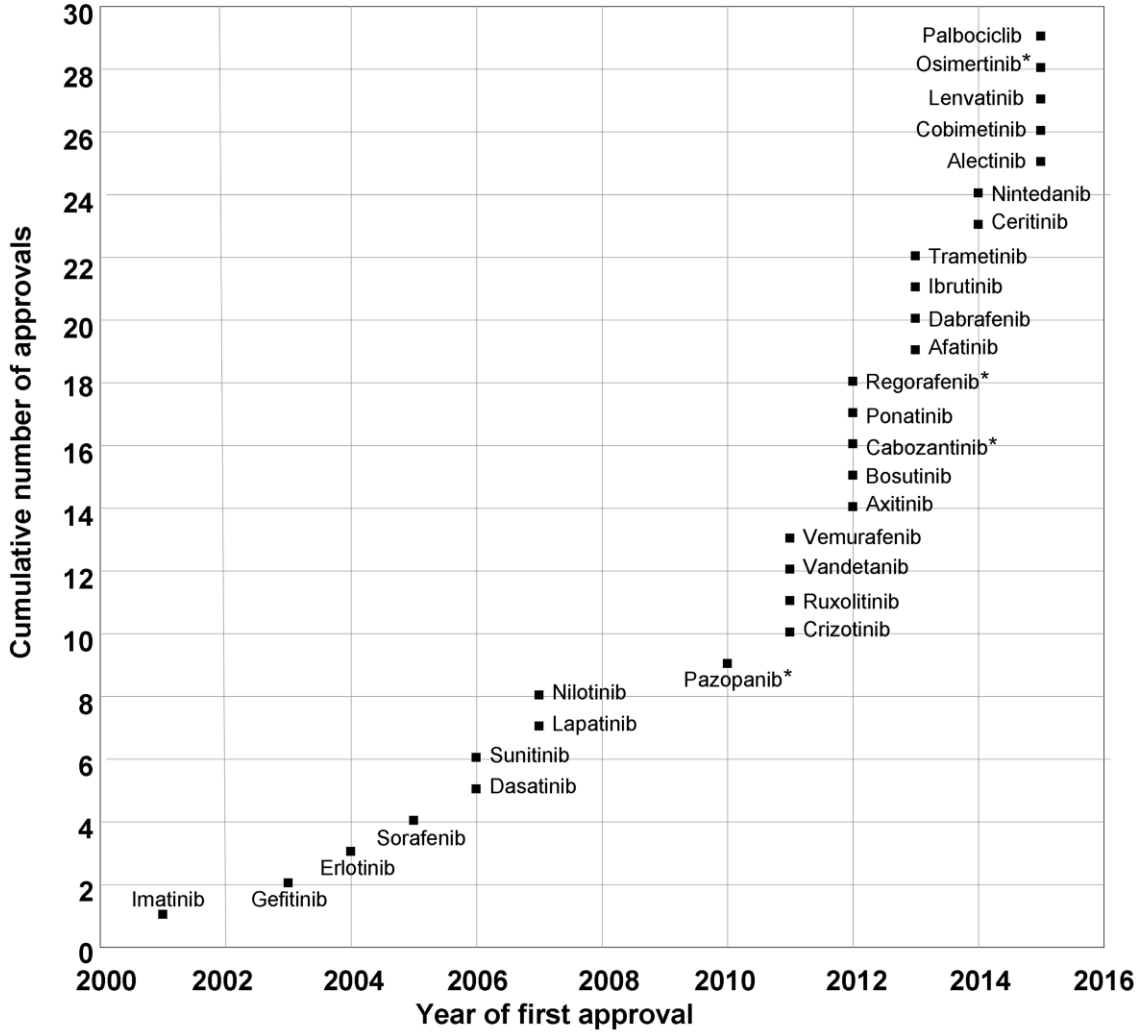


Fig. 8. Year of initial authorization of all FDA-approved small molecule inhibitors that bind directly to the target protein kinase domain. The asterisks indicate medications for which no drug-enzyme X-ray structures are available. Adapted from Ref. [73].

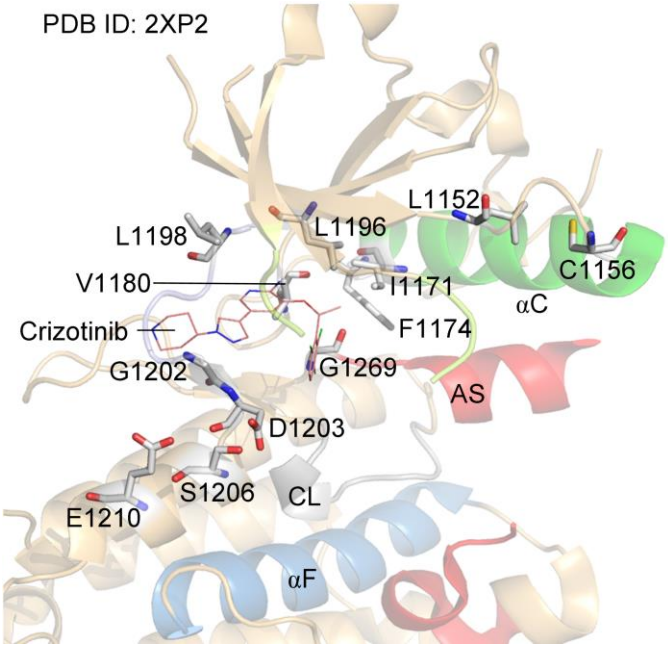


Table 1: Important residues in human ALK

	Comments	Residues
No. of residues		1620
Protein kinase domain		1116-1392
Molecular Wt (kDa)		176.4
N-lobe		
Glycine-rich loop: GxGxΦG	Anchors ATP β-phosphate	^{1123} GHGAFG^{1128}
β3-lysine (K of K/E/D/D)	Forms salt bridges with ATP α- and β-phosphates	K1150
αC-glutamate (E of K/E/D/D)	Forms ion pair with β3-K	E1167
Hinge residues	Connects the N- and C-lobes	^{1197}ELMAGG^{1202}
C-lobe		
Catalytic loop HRD (first D of K/E/D/D)	Catalytic base (abstracts proton)	D1249
Catalytic loop Asn (N)	Chelates Mg^{2+}(2)	N1254
Activation segment	Positions protein substrate	D1270–E1299
AS DFG (second D of K/E/D/D)	Chelates Mg^{2+}(1)	D1270
AS phosphorylation sites	Stabilizes the AS after phosphorylation	Y1278,Y1282,Y1283
End of activation segment		^{1297}PPE^{1299}
UniProtKB accession no.	Q9UM73	

Table 2: Spine and shell residues of human ALK and InsRK and murine PKA

	Symbol	KLIFS No.[a]	ALK	InsRK	PKA[b]
Regulatory spine					
β4-strand (N-lobe)	RS4	38	C1182	L1062	L106
C-helix (N-lobe)	RS3	28	I1171	M1051	L95
Activation loop F of DFG (C-lobe)	RS2	82	F1271	F1151	F185
Catalytic loop His/Tyr (C-lobe)	RS1	68	H1247	H1130	Y164
F-helix (C-lobe)	RS0	None	D1311	D1191	D220
R-shell					
Two residues upstream from the gatekeeper	Sh3	43	I1194	V1074	M118
Gatekeeper, end of β5-strand	Sh2	45	L1196	M1076	M120
αC-β4 loop	Sh1	36	V1180	V1060	V104
Catalytic spine					
β2-strand (N-lobe)	CS8	11	V1130	V1010	V57
β3-AxK motif (N-lobe)	CS7	15	A1148	A1028	A70
β7-strand (C-lobe)	CS6	77	L1256	M1139	L173
β7-strand (C-lobe)	CS5	76	C1255	C1138	I174
β7-strand (C-lobe)	CS4	78	L1257	V1140	L172
D-helix (C-lobe)	CS3	53	L1204	L1084	M128
F-helix (C-lobe)	CS2	None	L1318	V1198	L227
F-helix (C-lobe)	CS1	None	I1322	I1202	M231

[a]KLIFS (kinase–ligand interaction fingerprint and structure) from Ref. [48].

[b]From Ref. [36,37,43].

Table 3: Properties of selected orally effective small molecule ALK inhibitors approved and in clinical trials[a]

Name, code, trade name®	Targets	PubChem CID	Formula	MW (Da)	ALK K_i (nM)	D/A[b]	cLogP[c]	FDA-approved indications
Crizotinib, PF-2341066, Xalkori®	ALK, c-Met, ROS1	11626560	$C_{21}H_{22}Cl_2FN_5O$	450.3	1.6[d]	2/6	3.7	ALK^+ (2011) and $ROS1^+$ (2016) NSCLC
Ceritinib, LDK 378, Zycadia®	ALK, IGF-1R, InsR, ROS1	57379345	$C_{28}H_{36}ClN_5O_3S$	558.1	0.10[d]	3/8	6.4	ALK^+-NSCLC after crizotinib resistance (2014)
Alectinib, CH5424802, Alecensa®	ALK, RET	49806720	$C_{30}H_{34}N_4O_2$	482.6	0.09[d]	1/5	5.2	ALK^+-NSCLC after crizotinib resistance (2015)
Brigatinib, AP26113	ALK, EGFR	68165256	$C_{29}H_{39}ClN_7O_2P$	584.1	0.10[d]	2/9	4.6	None
Entrectinib, RXDX-101	ALK, ROS1, Trk1/2/3	25141092	$C_{31}H_{34}F_2N_6O_2$	560.6	12[e]	3/8	5.7	None
Lorlatinib, PF-06463922	ALK	71731823	$C_{21}H_{19}FN_6O_2$	406.4	0.14[d]	1/7	1.5	None

[a]clinicaltrials.gov
[b]No. of hydrogen bond donors/acceptors.
[c]Calculated log of the partition coefficient as determined by MedChem Designer® v.1.0.1.15.
[d]From supplement to Ref. [49].
[e]From Ref. [50].

Table 4: Drug-protein kinase interactions

Drug-enzyme[a]	PDB ID	DFG-D	AS[b]	αC	R-Spine	GK[c]	Inhibitor class	Pockets and sub-pockets occupied[d]
Bosutinib-Src	4MXO	in	open	in	active	T	I	F, G, BP-I-A/B
Dasatinib-Abl	2GQG	in	open	in	active	T	I	F, FP-I
Erlotinib-EGFR	1M17	in	open	in	active	T	I	F, G, BP-I-A/B
Gefitinib-EGFR	2ITY	in	open	in	active	T	I	F, G, BP-I-A/B
Palbociclib-CDK6	2EUF	in	open	in	active	F	I	F
Ruxolitinib-Src[e]	4U5J	in	open	in	active	T	I	F[g]
Tofacitinib-JAK3	3LXK	in	open	in	active	M	I	F, FP-I/II
Tofacitinib-JAK1	3EYG	in	open	in	active	M	I	F, FP-I/II
Vandetanib-RET	2IVU	in	open	in	active	V	I	F, G, BP-I-A/B
Dabrafenib-B-Raf	5CSW	in	closed	out	RS3→	T	I ½ A	F, G, B, FP-II, BP-I-A, BP-I-B, BP-II-in, BP-II-A-in
Dasatinib-Lyn[f]	2ZVA	in	open	in	RS2/3/4 up	T	I ½ A	F, G, B, BP-I-A/B
Lapatinib-EGFR	1XKK	in	closed	out	RS2/3→	T	I ½ A	F, G, B, BP-I-A/B, II in, IIA in
Lenvatinib-VEGFR	3WZD	in	?	in	RS3/4 up	V	I ½ A	F, G, B, IB, II in
Vemurafenib-B-Raf	3OG7	in	?	out	RS3→	T	I ½ A	F, G, B
Alectinib-ALK	3AOX	in	closed	in	active	L	I ½ B	F, FP-I
Ceritinib-ALK	4MKC	in	closed	in	active	L	I ½ B	F, FP-I
Crizotinib-ALK	2XP2	in	closed	in	active	L	I ½ B	F, FP-I
Crizotinib-Met	2WGJ	in	closed	out	←RS4	L	I ½ B	F, FP-I

Erlotinib-EGFR	4HJO	in	closed	out	RS2/3→	T	I ½ B	F, G, BP-I-A/B
Sunitinib-CDK2	3TI1	in	closed	out	RS3→	F	I ½ B	F
Axitinib-VEGFR	4AG8	out	closed	in	←RS2/4	V	IIA	F, G, B, BP-I-B, II out
Imatinib-Abl[f]	1IEP	out	closed	in	←RS2	T	IIA	F, G, B, BP-I-A/B, II out, IV
Imatinib-Kit	1T46	out	closed	in	←RS2	T	IIA	F, G, B, BP-I-A/B, II out, IV
Nilotinib-Abl	3CS9	out	closed	in	←RS2	T	IIA	F, G, B, BP-I-A/B, II out, III,V
Ponatinib-Abl[f]	3OXZ	out	closed	in	←RS2	T	IIA	F, G, B, BP-I-A/B, II out, III,IV
Ponatinib-Kit	4U0I	out	closed	in	←RS2	T	IIA	F, G, B, BP-I-A/B, II out, III,IV
Ponatinib-B-Raf	1UWH	out	?	in	←RS2/4	T	IIA	F, G, B, BP-I-B, II out, III
Sorafenib-CDK8	3RGF	out	?	in	←RS2/4	F	IIA	F, G, B, BP-I-B, II out, III
Sorafenib-VEGFR	4ASD	out	closed	in	←RS2	V	IIA	F, G, B, BP-I-B, II-out, III
Bosutinib-Abl	3UE4	out	open	in	←RS2	T	IIB	F, G, BP-I-A/B
Nintedanib-VEGFR2	3C7Q	out	closed	in	←RS2	V	IIB	F, G, BP-I-B
Sunitinib-Kit	3G0E	out	closed	in	←RS2	T	IIB	F
Sunitinib-VEGFR	4AGD	out	closed	in	←RS2	V	IIB	F, BP-I-B
Cobimetanib-MEK1	4AN2	in	closed	out	RS3→	M	IV	F,G, B, BP-II-in
Trametinib-MEK1	Model[h]	in	closed	out	RS3→	M	IV	F,G, B, BP-II-in[g]
Afatinib-	4G5J	in	open	in	active	T	VI	F, G, BP-I-

EGFR								A/B
Ibrutinib-BTK	Model[i]	in	closed	in	inactive	T	VI	F, G, B, BP-I-A/B[g]

[a]All human proteins unless otherwise noted.
[b]Activation segment.
[c]Gatekeeper residue.
[d]F, front cleft; G, gate area; B, back cleft; from http://klifs.vu-compmedchem.nl/ unless otherwise noted.
[e]Chicken enzyme.
[f]Mouse enzyme.
[g]Inferred.
[h]Ref. [42].
[i]Ref. [40].