BIOCHEMISTRY
including biophysical chemistry & molecular biology
Accelerated Publication

Biochemistry **2010,** *49,* 6813–6825 6813
DOI: 10.1021/bi1005514

Crystal Structures of Anaplastic Lymphoma Kinase in Complex with ATP Competitive Inhibitors

Roberto T. Bossi,[†,‡] M. Beatrice Saccardo,[†] Elena Ardini,[†] Maria Menichincheri,[†] Luisa Rusconi,[†] Paola Magnaghi,[†] Paolo Orsini,[†] Nilla Avanzi,[†] Andrea Lombardi Borgia,[†] Marcella Nesi,[†] Tiziano Bandiera,[†,§] Gianpaolo Fogliatto,[†] and Jay A. Bertrand*[,†]

[†]*Nerviano Medical Sciences S.r.l., Viale Pasteur 10, 20014 Nerviano (MI), Italy.* [‡]*Current address: Tecan Italia S.r.l., Via Brescia 39, 20063 Cernusco sul Naviglio (MI), Italy.* [§]*Current address: Drug Discovery and Development, Italian Institute of Technology, Via Morego 30, I-16163 Genova, Italy.*

Received April 12, 2010; Revised Manuscript Received July 4, 2010

ABSTRACT: Anaplastic lymphoma kinase (ALK) is a receptor tyrosine kinase involved in the development of several human cancers and, as a result, is a recognized target for the development of small-molecule inhibitors for the treatment of ALK-positive malignancies. Here, we present the crystal structures of the unphosphorylated human ALK kinase domain in complex with the ATP competitive ligands PHA-E429 and NVP-TAE684. Analysis of these structures provides valuable information concerning the specific characteristics of the ALK active site as well as giving indications about how to obtain selective ALK inhibitors. In addition, the ALK-KD−PHA-E429 structure led to the identification of a potential regulatory mechanism involving a link made between a short helical segment immediately following the DFG motif and an N-terminal two-stranded β-sheet. Finally, mapping of the activating mutations associated with neuroblastoma onto our structures may explain the roles these residues have in the activation process.

Anaplastic lymphoma kinase (ALK)[1] is a receptor tyrosine kinase in the insulin receptor superfamily whose physiological expression is limited to neuronal cells, especially during development. In mice and humans, the full-length ALK genes encode proteins of 1620 and 1619 residues, respectively, that contain an extracellular domain, a hydrophobic transmembrane domain, and an intracellular kinase domain (*1*, *2*). The ALK kinase domain (ALK-KD) contains three tyrosines in the activation loop (YxxxYY) that represent the major autophosphorylation sites that are associated with enzyme activation (*3*). To date, no structures of the ALK-KD are available, although a homology model has been reported (*4*). ALK was originally identified in 1994 as the product of a recurring gene fusion in anaplastic large cell lymphomas (ALCLs), a subtype of non-Hodgkin lymphomas characterized by a typical cell morphology and by the cell surface expression of the CD30 antigen (*5*, *6*). A discrete subset of ALCLs carry a balanced chromosomal translocation in which the entire nucleophosmin (NPM) gene on chromosome 5 is fused to the 3′ portion (including the entire kinase domain) of the ALK gene on chromosome 2. This chromosomal rearrangement results in the ectopic expression of the NMP−ALK fusion protein that has a constitutively activated ALK-KD. The NPM−ALK chimeric protein was demonstrated to have a strong oncogenic potential and to be responsible for neoplastic transformation. In addition, the constitutive expression of the human NPM−ALK protein in mouse T-cell lymphocytes is sufficient for the development of lymphoid neoplasia in transgenic animals with a short period of latency (*7*, *8*). Moreover, preclinical experimental data have demonstrated that ALK downregulation by RNA interference or inhibition using specific small molecules impairs cell proliferation of ALK+ ALCL cell lines both in vitro and in vivo (*9*). More recently, an additional chromosomal rearrangement involving the cytoplasmic portion of ALK has been identified in a subset of non-small cell lung cancer (NSCLC) patients (*10*, *11*). As observed for the NPM−ALK protein, this new fusion variant (EML4−ALK) that contains the N-terminal portion of the echinoderm microtubule-associated protein like 4 and the entire intracellular portion of ALK has a constitutively active ALK kinase activity and was demonstrated to be oncogenic. The EML4−ALK rearrangement was originally identified in Chinese and Japanese populations (frequency of 4−6%), but more recently, genetic screening performed on two population-based NSCLC cohorts from Switzerland and the United States confirmed the existence in Western populations (frequency of 3%) (*12*). Furthermore, treatment of EML4−ALK+ NSCLC cell lines with ALK inhibitors demonstrated that these cells are clearly dependent upon ALK for proliferation and survival (*13*).

Recent data reported from four independent groups established the primary role of ALK as a critical oncogene in the pathogenesis of neuroblastoma, an aggressive and often lethal childhood cancer (*14*−*17*). Neuroblastoma is the most common solid tumor of early childhood and originates from neural

*To whom correspondence should be addressed. E-mail: jay.bertrand@nervianoms.com. Phone: +39 0331 58 1395. Fax: +39 0331 58 1360.

[1]Abbreviations: Abl, Abelson tyrosine kinase; ACK1, activated Cdc42-associated kinase 1; ALCL, anaplastic large-cell lymphoma; ALK, anaplastic lymphoma kinase; ALK-CD, anaplastic lymphoma kinase cytoplasmic domain; ALK-KD, anaplastic lymphoma kinase domain; A-loop, activation loop; AMP-PNP, adenosine 5′-(β,γ-imido)-triphosphate; BSA, bovine serum albumin; DFG, Asp-Phe-Gly; Aur-A, Aurora-A; Aur-B, Aurora-B; BRK, breast tumor kinase; DTT, dithiothreitol; FAK, focal adhesion kinase; FGF1R, fibroblast growth factor receptor 1 kinase; FLT3, FMS-like tyrosine kinase-3; GST, glutathione *S*-transferase; IGF-1R, insulin-like growth factor 1 receptor; IRK, insulin receptor kinase; LCK, lymphocyte-specific kinase; NSCLC, non-small cell lung cancer; NPM, nucleophosmin; PLK1, Polo-like kinase 1; Sf9, *Spodoptera frugiperda*; VEGFR3, vascular endothelial growth factor receptor 3.

Published on Web 07/13/2010

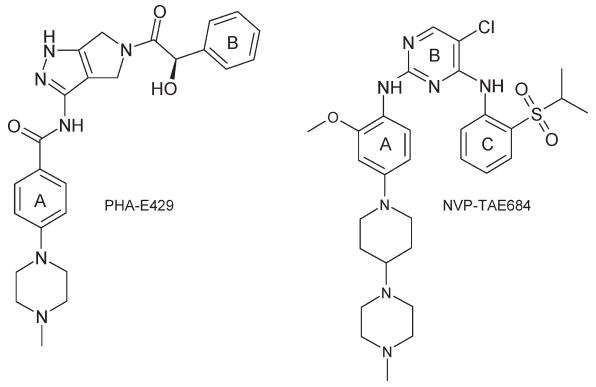

FIGURE 1: Chemical structures of PHA-E429 and TAE684.

crest-derived tissues. Germline mutations in the ALK-KD were found to be the cause of the inherited version of neuroblastoma. In addition, ALK somatic mutations and gene amplification were proven to play a role in more than 30% of sporadic neuroblastoma, the most common form of the disease. The most frequent ALK mutations have been fully characterized, and the resulting data indicate that they are gain-of-function mutations. Data supporting this gain of function include the constitutive autophosphorylation of ALK and the consequent activation of downstream signaling pathways, as well as the ability to transform Ba/F3 cells.

Taken together, these data support the hypothesis that ALK inhibition may represent an effective and innovative therapy for ALCL, NSCLC, and neuroblastoma patients whose tumors harbor ALK genetic alterations. With this in mind, a program was initiated to identify ALK small molecule inhibitors and, in parallel, crystallographic work began on the human ALK kinase domain (ALK-KD) to enable structure-based drug design. Here we report the crystal structures of ALK-KD in complex with the pyrrolo-pyrazole inhibitor PHA-E429 and with the high-affinity ALK inhibitor NVP-TAE684 (Figure 1). These structures provide valuable information concerning the specific characteristics of the ALK active site and give indications of how inhibitors could be modified to improve and/or maintain ALK selectivity.

MATERIALS AND METHODS

Chemical Synthesis. The synthesis of PHA-E429 is reported by Fancelli and co-workers (*18*). TAE684 was synthesized following the procedures reported in International Patent Application WO 2005/016894 (February 24, 2005). Compounds PHA-E589, NMS-E107, and NMS-E828 were synthesized in our laboratories following the procedures reported in the Supporting Information, and high-resolution mass spectrometry and NMR confirmed their structures.

Expression, Purification, and Characterization. The human ALK cytoplasmic domain (ALK-CD), residues 1060−1620, was cloned into vector PVL1393 (Pharmingen), modified by the insertion of glutathione *S*-transferase (GST), a PreScission (GE) protease cleavage site, and a Kozak consensus sequence. *Spodoptera frugiperda* (Sf9) cells were transfected with the pVL-GST-ALK-CD plasmid using the BaculoGold transfection system (BD Biosciences) as described by the manufacturer's protocol. Three rounds of viral amplification were performed to obtain a high-titer viral stock. Infected SF21 cells were harvested 48 h postinfection and lysed by liquid extrusion with a Gaulin homogenizer (Niro Soavi Italy) in 50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 0.2% CHAPS, 20% glycerol, 20 mM DTT, 1 mM Na_3VO_4, and Complete Protease Inhibitor Cocktail (Roche Diagnostics). The lysate was centrifuged at 20000*g* for 30 min, and the supernatant was loaded onto a GSH affinity chromatography column. After being extensively washed, recombinant GST-ALK-CD was eluted with 10 mM glutathione in 100 mM Tris-HCl (pH 8.0) and 10% glycerol. As a final purification step, GST-ALK-CD was loaded onto a Heparin Sepharose FF (GE Life Sciences) column and eluted with 25 mM Tris (pH 7.5), 500 mM NaCl, 2 mM DTT, and 20% glycerol. Fractions containing pure GST-ALK-CD were pooled, dialyzed against 50 mM Tris (pH 7.4), 150 mM NaCl, 2 mM DTT, and 20% glycerol, and stored at −80 °C for subsequent use in biochemical assays.

For human ALK-KD, residues 1094−1407, the cloning and expression were similar to those described above with the following exceptions. SF21 cells were harvested 72 h after infection and lysed in 50 mM Tris-HCl (pH 8.0), 500 mM NaCl, 20 mM DTT, 1 mM Na_3VO_4, and Complete Protease Inhibitor Cocktail (Roche Diagnostics), and the supernatant was loaded onto a GSH affinity chromatography column equilibrated with buffer A [50 mM Bicine (pH 8.4), 150 mM NaCl, and 5 mM DTT]. After being extensively washed, the bound protein was treated overnight with PreScission (GE) protease at 4 °C. The following morning, the cleaved protein was eluted from the column, pooled, and concentrated to reduce the volume. Next, a 2-fold molar excess of ligand (PHA-E429 or TAE684) was added to the protein, and the resulting mixture was loaded onto a GE Superdex 200 gel filtration column equilibrated with buffer A. Fractions were pooled, concentrated to 5−10 mg/mL, and stored at −80 °C for subsequent crystallization experiments. The purified ALK-KD was ~95% pure as judged by SDS−PAGE and was found to be homogeneous and unphosphorylated by electrospray mass spectrometry analysis. Autophosphorylation of ALK-KD after incubation in the presence of magnesium and ATP confirmed that the isolated kinase domain was active as judged by Western blot analysis using anti-phosphotyrosine.

Protein Kinase Activity Assays. Automated kinase assays were conducted as described by Pevarello and co-workers (*19*). Briefly, phosphoryl transfer was quantified after incubation of each specific kinase with ATP/[γ-^{33}P]ATP mix [2Km] and substrate [5Km] in a final volume of 60 μL. After incubation, the reaction was stopped and phosphorylated substrates were separated from nonincorporated radioactive ATP using Dowex resin. In certain cases, the kinases were preactivated by incubation with ATP to linearize the reaction kinetics. PHA-E429 was profiled against a panel of 19 kinases, and TAE684, PHA-E589, NMS-E107, and NMS-E828 were profiled against 45 kinases.

For ALK-CD, the kinase activity was determined at 25 °C by measuring the rate of phosphorylation of the peptide substrate H-ARDIYRASFFRKGGCAMLPVK-CONH2 (American Peptide Co.) using ATP traced with radiolabeled ATP (Redivue [γ-^{33}P]ATP, Amersham Pharmacia Biotech). The reaction buffer consisted of 50 mM Hepes-NaOH (pH 7.5), 3 mM $MgCl_2$, 1 mM $MnCl_2$, 1 mM DTT, 3 μM Na_3VO_4, and 0.2 mg/mL BSA. Prior to the reaction, ALK-CD was preactivated by incubation at 400 nM for 1 h at 30 °C with 50 μM ATP in the reaction buffer. Transphosphorylation rates were measured in 96-well plates through the selective capture of unreacted ATP by a strong anion exchanger resin (Dowex 1X8, formate form, Supelco) according to the procedures reported in U.S. Patent 6,927,037 B2 with some modifications. Briefly, to 60 μL of reaction mixture

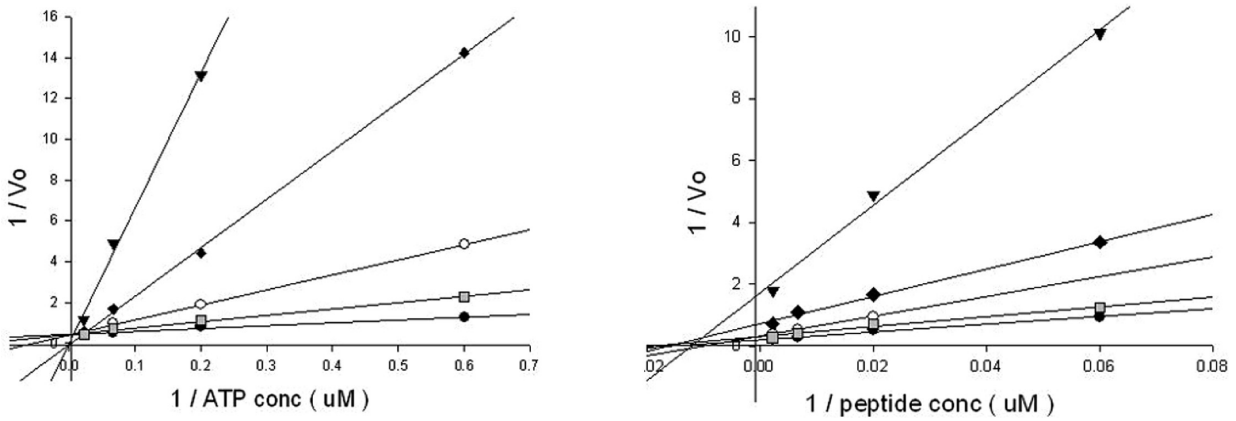

FIGURE 2: Double-reciprocal plots resulting from the kinetic analysis of PHA-E429 vs ATP (right) and PHA-E429 vs peptide (left). PHA-E429 concentrations of 0 (●), 26.7 (□), 80 (○), 240 (◆), and 720 nM (▼) were used.

was added 210 μL of a 1:3 (v/v) suspension of resin equilibrated with 150 mM sodium formate (pH 3.0), which also served to stop the reaction. After the solution had been mixed and the resin allowed to settle for at least 1 h, 30 μL of supernatant was carefully withdrawn using a Biomek FX instrument (Beckman Coulter Inc.) and added to a 96-well plate (Optiplate, Packard) containing 150 μL of MicroscintTM 40 (Perkin-Elmer). The radioactivity left in the supernatant and thus incorporated into the peptide substrate was then measured using a Top Count NXT instrument (Packard). Results for ALK-CD were as follows: K_m for ATP = 4.09 ± 0.38 μM, K_m for peptide = 98.31 ± 8.97 μM, and V_{max} = 6.92 ± 0.26 nM min^{-1} nM^{-1}.

Inhibition constants (K_i) of PHA-E429, TAE684, and staurosporine (*20*) were evaluated using a matrix of varied concentrations of ATP, peptide, and inhibitors, the values ranging from 30 to 900% of their IC_{50} values. The data were fitted using the rapid equilibrium bireactant model in the presence of a competitive inhibitor:

$$v = (V_{max}[A][B])/$$
$$\left(\alpha K_a K_b + \alpha K_a[B] + \alpha K_b[A] + [A][B] + \alpha K_a K_b\frac{[I]}{K_i} + \frac{\alpha K_a[B][I]}{\beta K_i}\right)$$

where V_{max} is the maximum velocity, [A], [B], and [I] are the ATP, peptide, and inhibitor concentrations, respectively, K_a and K_b are the Michaelis constants, α is the ATP and peptide interaction factor, K_i is the inhibition constant, and β is the inhibitor interaction factor. Data from the time course matrix experiments were converted into double-reciprocal plots for visual inspection. Interpretation of the experiments indicates that PHA-E429, TAE684, and staurosporine behave as ATP competitive inhibitors with inhibition constants of 23.68 ± 2.35, 0.65 ± 0.07, and 5.56 ± 0.75 nM, respectively. The double-reciprocal plots from the kinetic analysis of PHA-E429 are shown as an example in Figure 2.

Crystallization and Structure Determination. Unphosphorylated ALK-KD, residues 1094−1407, was used for all crystallographic studies. Both reported structures came from batches of protein that were copurified in the presence of the inhibitors, and the protein was crystallized without any subsequent addition of inhibitor. The protein−inhibitor complex at a concentration of 5−10 mg/mL in buffer A was mixed with a reservoir containing 20% PEG 3350 and 0.15 M DL-malic acid (pH 7.0) (complex with PHA-E429) or 18% PEG 3350 and 0.1 M Tris-HCl (pH 8.5) (complex with TAE684). Crystals grew in 2−3 days at 277 K by the vapor diffusion method. Crystals were cryoprotected before being flash-frozen in liquid nitrogen by being soaked in 20% PEG 3350 and 1 M DL-malic acid (pH 7.0) or 26% PEG 3350, 0.1 M Tris-HCl (pH 8.5), and 20% glycerol. Diffraction data were collected at the European Synchrotron Radiation Facility (Grenoble, France) on beamlines ID23-1 and ID14-3. Indexing, integration, and scaling were performed using MOSFLM and SCALA (*21*). Both ALK-KD structures were determined by molecular replacement using Phaser (*22*). Search models were an in-house determined structure of the triphosphorylated IGF-1R kinase domain for the ALK-KD−PHA-E429 complex and the latter structure for the ALK-KD−TAE684 complex. Model building was done using Coot (*23*), and refinement was done with RefMac (*24*) and CNX (*25*). Crystallographic data are listed in Table 1, and SigmaA weighted difference omit maps of the inhibitors are shown in Figure 3. Structural images have been generated with CCP4mg (*26*) and PyMol (*27*). Superposition of the two ALK-KD structures gave a root-mean-square deviation (rmsd) of 0.50 Å for the 279 structurally equivalent Cα atoms.

RESULTS

Protein Production and Crystallization. Efforts to obtain crystals of the ALK kinase domain began with a multiple-construct strategy using structural information from homologous tyrosine kinases as well as results from limited proteolysis. A large number of these constructs were evaluated in *Escherichia coli* (~40 constructs) and a smaller subset in insect cells (19 constructs). In the end, none of the constructs expressed in *E. coli* and only four of the 19 constructs expressed in insect cells provided soluble and monodisperse protein that was suitable for crystallization trials. Furthermore, even with these four constructs (residues 1077−1405, 1085−1400, 1085−1405, and 1094−1407), the yields after purification were too low to allow adequate sampling of crystallization space. However, these yields could be significantly improved by incorporation of information coming from buffer optimization experiments and by the addition of inhibitors during the final steps of the purification. These efforts eventually led to a reproducible crystallization protocol that allows us to quickly and easily obtain crystals of ALK-KD (construct 1094−1407) in complex with ATP competitive inhibitors. To date, we have determined seven structures of ALK-KD in complex with ATP competitive inhibitors, two of which are reported here.

Overall Structure of the ALK Catalytic Domain. The ALK structure begins with a 13-residue N-terminal segment

containing two antiparallel β-strands ($\beta1'$ and $\beta2'$) that pack against the back of the αC helix (Figure 4). This is then followed by the bilobal protein kinase fold (*28*) which is characteristic of all tyrosine and Ser/Thr kinases. The N-terminal (or small) lobe is composed of a core five-stranded antiparallel β-sheet and a single α-helix (αC). The C-terminal (or large) lobe is predominantly α-helical and contains the regulatory activation loop upon which are located three tyrosine residues whose phosphorylation is associated with activation. The cavity between the N- and C-terminal lobes and the hinge region that interconnects the two defines the binding site for ATP and ATP competitive inhibitors. In the ALK structures, the final model comprises residues 1095−1401 with gaps in different central regions that are presumably due to disorder. In the ALK−PHA-E429 complex, the disordered regions include residues 1123−1129 of the glycine-rich loop, residues 1137−1144 from the loop interconnecting strands $\beta2$ and $\beta3$ of the antiparallel β-sheet, and residues 1279−1288 from the activation loop (A-loop), while in the ALK−TAE684 complex, the gaps are for residues 1137−1144 and 1275−1288.

Table 1: Crystal Structure Data and Refinement Statistics[a]

Data Collection		
ALK-KD complex	PHA-E429	NVP-TAE684
PDB entry	2XBA	2XB7
space group	$P2_12_12_1$	$P2_12_12_1$
cell parameters (Å)		
a	51.884	57.781
b	57.013	57.307
c	105.541	105.157
X-ray source	ID23-1 ESRF	ID14-3 ESRF
resolution (Å)	1.95	2.50
no. of observations		
total	110917	47954
unique	23448	11104
completeness (%)	99.6 (99.8)	98.3 (99.5)
R_{sym}	0.088 (0.448)	0.094 (0.428)
$I/\sigma I$	16.3 (2.9)	14.7 (3.0)
Refinement		
resolution (Å)	1.95	2.50
no. of reflections		
working set (%)	22214 (94.5)	10527 (92.9)
test set (%)	1180 (5.0)	552 (4.9)
R_{cryst}/R_{free}	0.197/0.230	0.198/0.246
rmsd		
bond lengths (Å)	0.012	0.009
bond angles (deg)	1.2	1.2
Ramachandran plot (%)		
most favored region	98.5	96.8
allowed region	1.5	2.9
outlier region	0	0.4
average *B*-factor/no. of atoms		
ligand	25.27/34	43.85/42
water	32.68/136	31.21/42
protein	26.14/2245	33.99/2253
total	26.49/2415	34.09/2361

[a]Numbers in parentheses correspond to values for the highest-resolution shell.

Analysis of the Kinase Conformation. The ALK structures reported here display features that are characteristic of an inactive protein kinase. It is generally accepted that a high level of similarity exists for the active protein kinase structures, while there are significant differences for the inactive structures. Thus, to highlight the inactive features of the ALK-KD structures, a comparison was made with the active structure of triphosphorylated IRK in complex with an ATP analogue and a peptide substrate [IRK3P−AMP-PNP, Protein Data Bank (PDB) entry 1IR3]. IRK was chosen as a reference on the basis of its sequence similarities with ALK-KD [44% identical and 66% homologous (Figure 4)] as well as the fact that it represents a transition state complex for the phosphoryl transfer to a peptide substrate (*29*). For the sake of simplicity, the ALK−PHA-E429 complex was chosen for the comparison with the IRK3P−AMP-PNP complex, but most of the features described apply to both of the ALK structures reported here. As shown in Figure 5, the structures of the IRK3P−AMP-PNP and ALK−PHA-E429 complexes superpose quite well, giving a rmsd of 1.42 Å for the 235 structurally equivalent Cα atoms. Analysis of the superposition reveals that although the two structures are quite similar there are differences in the degrees of lobe closure between the N- and C-terminal domains and in the positions of the αC helices. In effect, the

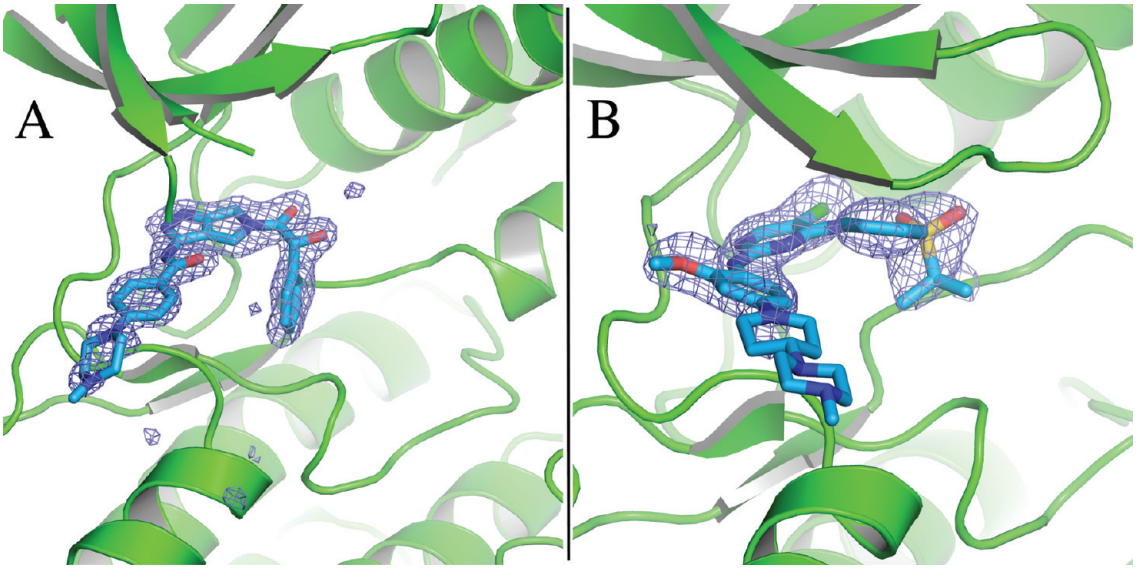

FIGURE 3: SigmaA weighted difference omit maps associated with the inhibitors PHA-E429 (A) and TAE684 (B) bound in the ALK-KD active site.

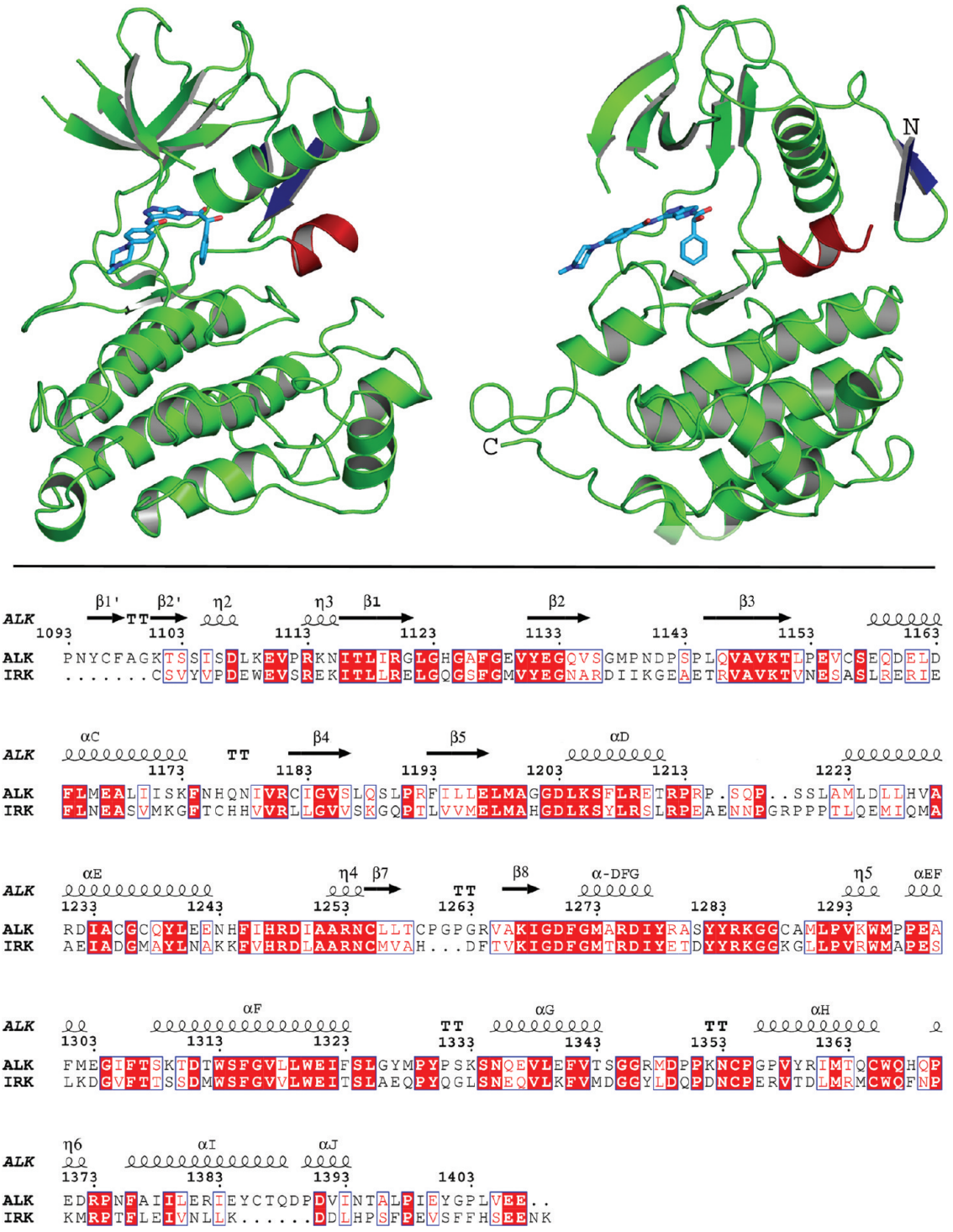

FIGURE 4: Overall architecture of ALK-KD with PHA-E429 bound in the active site (top). The right image has been rotated 60° with respect to the image on the left. Secondary structures are represented as helices (α-helix), arrows (β-strand), and loops (disordered loop). The N-terminal two-stranded antiparallel β-sheet is colored blue, and the short helical segment following the DFG motif is colored red. The inhibitor PHA-E429 is shown as a ball-and-stick representation with light blue carbon atoms. The bottom panel shows a sequence alignment of the kinase domains of ALK and IRK; secondary structure elements corresponding to the ALK-KD structure are shown above the sequence. The alignment was produced using ESPript (*42*).

active IRK structure is more closed and its αC helix has rotated toward the C-terminal lobe. Thus, the ALK-KD structure is inactive on the basis of its degree of lobe closure and the incorrect positioning of the αC helix. However, it is worth noting that the position of the αC helix allows the formation of the conserved hydrogen bond between the catalytic lysine (Lys1150) and the glutamate residue of the αC helix (Glu1167), a feature that is generally representative of an active kinase. In active kinases, the HRD motif is involved in two conserved hydrogen bonds: one between the arginine side chain and the main chain carbonyl of the DFG + 1 residue and the other between the histidine (or tyrosine) side chain and the carbonyl oxygen of the DFG − 1 residue. In the ALK−PHA-E429 structure, the carbonyl oxygen of Met1273 (DFG + 1) does not form a hydrogen bond with the side chain of Arg1248 (HRD motif); instead, it makes an α-helical hydrogen bond with the backbone nitrogen of Ile1277.

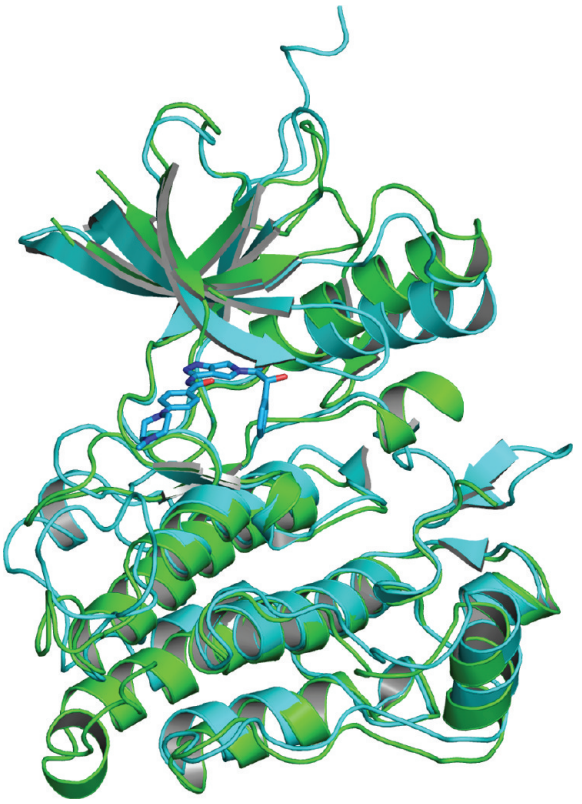

FIGURE 5: Ribbon diagram showing the superposition of ALK with PHA-E429 (green ribbons) and triphosphorylated IRK with an ATP analogue and a peptide substrate (cyan ribbons; PDB entry 1IR3).

Furthermore, in the ALK−PHA-E429 structure, a peptide flip of Gly1269 directs the carbonyl oxygen into rather than away from the ATP pocket, and as a result, the hydrogen bond between Nε2 of His1247 (HRD motif) and the carbonyl oxygen of Gly1269 is lacking. Other features that are apparent from the structural superposition with the IRK3P−AMP-PNP complex are the differences in the A-loop regions. In the IRK structure, the A-loop is completely ordered and the loop extends toward the back of the structure before looping around and coming back toward the front. Two short β-strands (β9 and β10) are present in the IRK A-loop, and they pair with structurally adjacent strands; β9 begins at the DFG + 3 position and pairs with β6, and β10 begins with the adjacent autophosphorylation tyrosines and pairs with β12. In contrast, the A-loop in the ALK−PHA-E429 structure is initiated with a short ordered segment containing the DFG motif and a seven-residue helical segment, followed by the disordered segment (residues 1279−1288), and then becomes ordered again several residues before Pro1292 at the C-terminal end of the A-loop. Both structures are in the "DFG-in" conformation that is typical of an active kinase, and the two structures superpose relatively well for this initial segment of the A-loop (Figure 5). However, immediately after the DFG motif, the two structures deviate significantly in that the IRK structure begins β9 at position DFG + 3 while the ALK structure begins the short helical segment that extends toward the solvent region (hereafter termed the DFG helix). Actually, a steric clash between the DFG + 2 residue (Ala1274) and the side chain of Phe1245 may be the reason why ALK-KD is unable to initiate β9. Clearly, Phe1245 would need to rotate to make space for β9, but in so doing it would sterically clash with the side chain of Phe1098, a residue from β1′ of the two-stranded antiparallel β-sheet that packs against the αC helix. It is worth noting that this side chain position of Phe1245 is conserved in all seven of our ALK-KD structures. Interestingly, the last visible residue of the DFG helix is Tyr1278, one of the three tyrosine residues phophorylated during the activation of the enzyme which leads to the correct formation of the substrate binding site. Moreover, in the ALK−PHA-E429 structure, the hydroxyl group of Tyr1278 forms a hydrogen bond to the backbone nitrogen of Cys1097, another residue from β1′. Presumably, the formation of the helix, the steric clash involving Phe1245, and the hydrogen bond made by Tyr1278 are all part of the mechanism of enzyme regulation because together they impede the formation of the substrate binding site. However, the lack of an apoenzyme ALK structure makes it difficult to confirm this hypothesis since the inhibitors in our ALK structures may impact the conformation of the A-loop. Nonetheless, the DFG helix and the hydrogen bond involving Tyr1278 are consistent features of six of the seven internally determined ALK structures. The one exception is the ALK−TAE684 structure that lacks electron density for the majority of the DFG helix residues.

ALK with PHA-E429. PHA-E429 is a high-affinity Aurora kinase inhibitor coming from a structure-based combinatorial expansion (*18*) with IC_{50} values of 6 and 48 nM for Aur-A and Aur-B, respectively. Kinase selectivity profiling of PHA-E429 reveals that it inhibits four other kinases with IC_{50} values of < 100 nM; these are Abl, FGFR1, LCK, and ALK with IC_{50} values of 14, 18, 61, and 91 nM, respectively. To confirm that PHA-E429 is an ATP competitive inhibitor for ALK-CD, we determined the inhibition constant (K_i). Figure 2 shows the double-reciprocal plots, and interpretation of the data gives an inhibition constant of 23.68 ± 2.35 nM. To understand the potential this scaffold may have for ALK inhibition, we determined the crystal structure of PHA-E429 in complex with ALK-KD. The 1.95 Å structure reveals a donor−acceptor−donor series of hydrogen bonds between the pyrazole amide moiety of PHA-E429 and the ALK hinge residues (Figure 6). The inhibitor's hydroxyl group interacts with two water molecules, one of which bridges with the side chain nitrogen of Lys1150 and one of the side chain oxygens of Asp1270. The central pyrrolopyrazole ring system of PHA-E429 sits above Leu1256, and the attached phenyl ring (ring B) bends down into a cavity formed by the Arg1253-Asn1254-Cys1255-Leu1256 and Gly1269-Asp1270 segments. As mentioned above, the glycine-rich loop is disordered in the ALK−PHA-E429 structure, with Leu1122 being the last ordered residue before the loop and Val1130 the first after the loop. Interestingly, both of these residues make interactions with the inhibitor; Val1130 together with Leu1256 sandwiches the central pyrrolopyrazole ring system, and Leu1122 together with Gly1202 sandwiches ring A. Finally, the methylpiperazine group that is attached to ring A extends toward the solvent.

Analysis of the ALK-KD residues interacting with the inhibitor and the comparison with those of the other kinases that are inhibited by PHA-E429 reveals information that can be valuable for the evaluation of affinity and selectivity. For example, all of the kinases with IC_{50} values of < 100 nM have leucine residues in the position that corresponds to Leu1256 (Table 2), suggesting that leucine has the most complementary shape to that of PHA-E429 and the "fit" may be a necessary requirement for obtaining high affinity. Another observation is that only ALK contains a glycine (Gly1269) in the DFG − 1 position, whereas all the others with IC_{50} values of < 100 nM contain alanine. Thus, one strategy for improving ALK selectivity would be to add meta substituents on ring B that would sterically clash with kinases having an

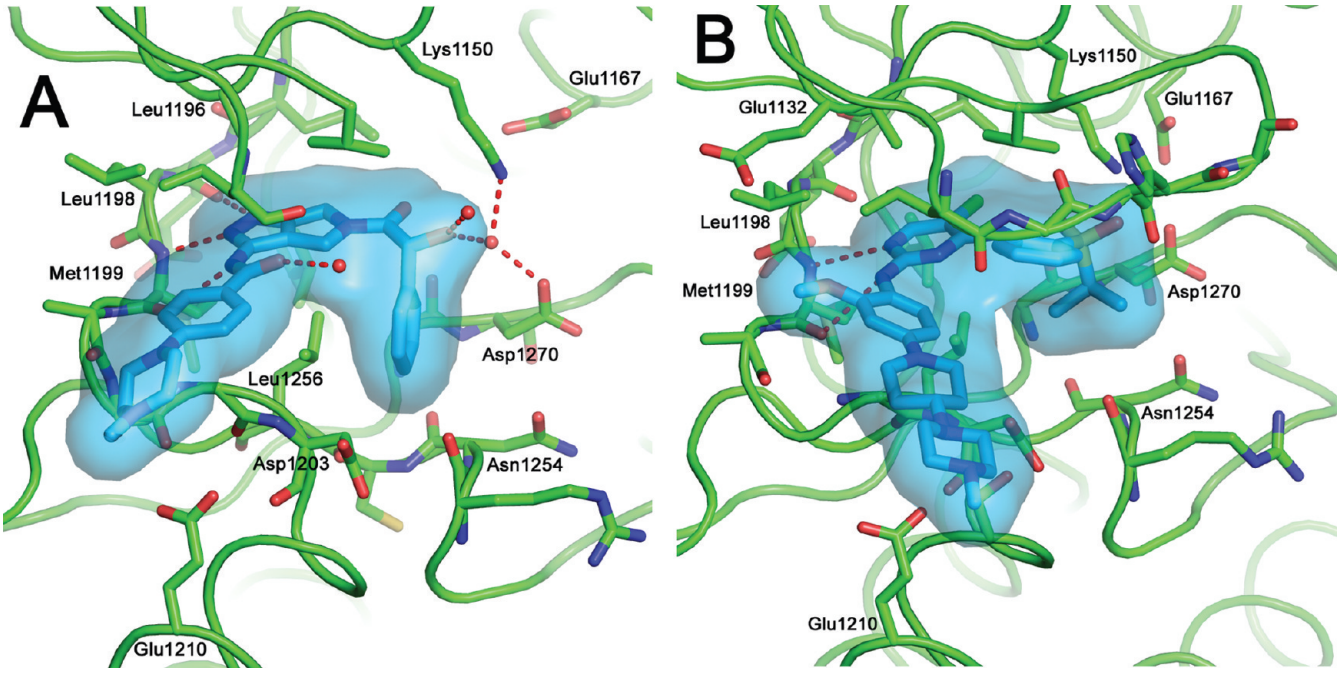

FIGURE 6: View of PHA-E429 (A) or TAE684 (B) bound in the active site of ALK-KD. Inhibitors are shown with blue carbon atoms and a semitransparent molecular surface. Ordered water molecules are depicted as red spheres and hydrogen bonds as dashed lines.

Table 2: Comparison of Select Active Site Residues in ALK and Other Kinases

ALK	L1122	A1126	F1127	V1130	V1180	L1196	L1198	D1203	L1256	G1269
ABL	L	Q	Y	V	V	T	F	N	L	A
ACK	L	S	F	V	I	T	L	S	L	G
Aur-A	L	K	F	V	L	L	Y	T	L	A
Aur-B	L	K	F	V	L	L	Y	E	L	A
BRK	L	Y	F	V	L	T	L	S	L	G
FAK	I	Q	F	V	V	M	L	E	L	G
FGFR1	L	A	F	V	I	V	Y	N	L	A
FLT3	L	A	F	V	V	F	Y	D	L	C
IGF-1R	L	S	F	V	V	M	L	D	M	G
IRK	L	S	F	V	V	M	L	D	M	G
LCK	L	Q	F	V	V	T	Y	S	L	A
VEGFR3	L	A	F	V	V	V	F	N	L	C

alanine in the position corresponding to Gly1269. Then, there are also sequence differences for the residues in the positions corresponding to Val1180 and Leu1196 toward the back of the ATP pocket. So, the addition of substituents to position 6 of the bicyclic tetrahydropyrrolo[3,4-*c*]pyrazole scaffold could modify the selectivity and affinity. Finally, an additional sequence difference between the kinases is observed for the acidic hinge residue that usually interacts with the ribose group of ATP in that only ALK contains an aspartate (Asp1203) in this position. As a result, the ribose pocket could be another zone to target to improve the selectivity versus Aur-A, Abl, FGFR1, Aur-B, and LCK and to potentially increase the affinity of inhibitors for ALK. Because these ideas concerning scaffold modifications are based solely on structural and sequence analysis, the decision was made to test a limited series of derivatives against ALK and the other kinases that are inhibited by PHA-E429 (Figure 7 and Table 3). The analogues of PHA-E429 available in our internal collection all contain gem-dimethyl groups at position 6 of the bicyclic tetrahydropyrrolo[3,4-*c*]pyrazole scaffold and are of racemic mixtures rather than the pure chiral compounds. Nonetheless, they provide valuable information concerning the impact of the modifications on affinity and selectivity. For example, PHA-E589 that contains the gem-dimethyl group on the pyrrolo ring and fluorines in the meta positions of ring B shows a higher affinity for ALK (61 nM) and improved affinity versus Aur-A, Aur-B, and FGFR1. However, the affinity for Abl and LCK remains unmodified compared to that of the parent compound PHA-E429. Next, the addition of an amino group to ring A (NMS-E107) maintains the affinity for ALK (58 nM) but decreases the affinity for Abl and LCK to 50 and 115 nM, respectively. Finally, NMS-E828 that contains a D-prolinamide shows good affinity for ALK (44 nM) and a high level of selectivity versus the other kinases inhibited by PHA-E429, with values ranging from 214 nM to >10 μM. Thus, as shown by NMS-E828, the affinity and selectivity of PHA-E429 could be improved by the introduction of substituents that extend (1) toward the back of the adenine pocket near Val1180 and Leu1196, (2) into the bottom of the cavity formed by Arg1253-Asn1254-Cys1255-Leu1256 and Gly1269-Asp1270 segments, and (3) into the ribose pocket situated above Asp1203 and below the glycine-rich loop residues Leu1122, Gly1123, and His1124. Furthermore, selectivity profiling of NMS-E828 against 45

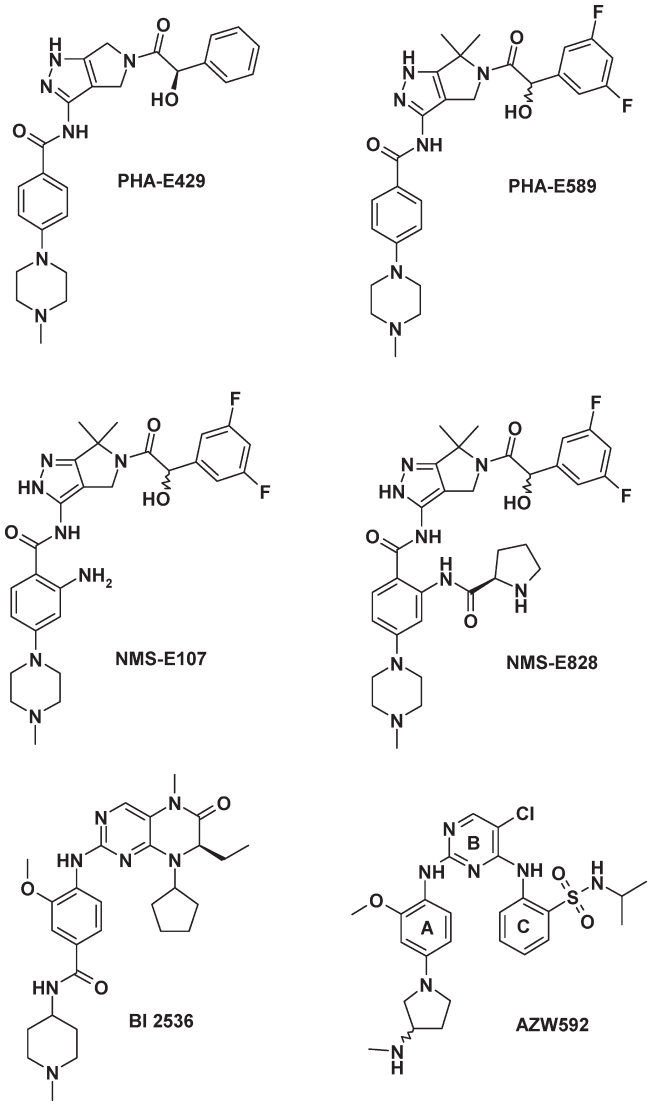

FIGURE 7: Chemical structures of kinase inhibitors PHA-E429, PHA-E589, NMS-E107, NMS-E828, BI 2536 (PLK1), and AZW592 (FAK).

Table 3: IC_{50} Values (in micromolar) for PHA-E429 and Analogues for Kinases Abl, ALK, Aur-A, Aur-B, FGFR1, and LCK

compd	ABL	ALK	AUR-A	AUR-B	FGFR1	LCK
PHA-E429	0.014	0.091	0.007	0.051	0.018	0.061
PHA-E589	0.021	0.061	0.173	0.775	0.045	0.057
NMS-E107	0.050	0.058	0.256	0.765	0.055	0.115
NMS-E828	0.348	0.044	1.341	> 10	0.387	0.214

kinases reveals that ALK is the only kinase with an IC_{50} value of < 100 nM.

ALK with TAE684. TAE684 is a potent and selective inhibitor of ALK that was identified by cellular screening of Ba/F3 NPM-ALK (*30*). TAE684 inhibits the proliferation of Ba/F3 NPM-ALK cells with an IC_{50} of 3 nM, while the survival of the parental Ba/F3 cell is not affected at concentrations up to 1 mM. Profiling of TAE684 on our internal kinase selectivity panel reveals that it inhibits nine kinases with IC_{50} values of < 100 nM; these are FAK, IGF-1R, ALK, Aur-A, IRK, FLT3, VEGFR3, BRK, and ACK1 with IC_{50} values of 12, 16, 17, 19, 27, 33, 33, 40, and 53 nM, respectively. TAE684 is a potent ATP competitive inhibitor of ALK-CD with an inhibition constant (K_i) of 0.65 ± 0.07 nM (data not shown). The 2.5 Å cocrystal structure of TAE684 with ALK-KD reveals how the inhibitor binds in the ATP-binding site (Figure 6). The hinge interactions include hydrogen bonds between the pyrimidine and amino nitrogens of TAE684 and the backbone nitrogen and oxygen of Met1199, respectively. Surprisingly, these are the only direct polar interactions observed between ALK-KD and TAE684. The pyrimidine ring is situated below Ala1148 and above Leu1256, with the chlorine substituent directed toward the back of the pocket. Ring C fills the cavity underneath the glycine-rich loop and makes hydrophobic interactions with Val1130 and the Gly1123-His1124 segment. The isopropylsulfonyl substituent on ring C is oriented so that the sulfonyl oxygens are directed toward Lys1150 and the isopropyl group is directed downward and packs against Asp1270 at the back of the pocket. Although the sulfonyl oxygens are directed toward Lys1150, they do not form polar interactions. Presumably, this is because Lys1150 is already making polar interactions; one is the conserved interaction with Glu1167, and another is with Asp1270 of the DFG motif. Ring A of TAE684 is sandwiched between Leu1122 and Gly1202, and the attached methoxy group is in a cavity formed by the side chain of Glu1132 and the Leu1198-Met1199-Ala1200 hinge segment. Although the electron density is weak for the piperidinyl-1-methylpiperazine group, the assumption is that it packs against the outer side of the αD helix, making Van der Waals interactions with Asp1203, Ser1206, and Glu1210.

As mentioned above, TAE684 shows a high level of affinity for Aur-A, FLT3, and VEGFR3, three kinases that contain aromatic side chains in the hinge position equivalent to Leu1198 (Table 2). This result was somewhat surprising considering that the modeling studies reported by Galkin and co-workers (*30*) predicted a steric clash with the orthomethoxy group attached to the 2-aniline substituent for residues bulkier than Leu at this position. In addition, analysis of the PLK1−BI 2536 crystal structure (PDB entry 2RKU) led Kothe and co-workers (*31*) to assign a similar role in selectivity to the orthomethoxy group in BI 2536, a selective PLK inhibitor currently in clinical trials (Figure 7). However, a superposition of the two structures [ALK-KD−TAE684 and PLK1−BI 2536 (not shown)] reveals that BI 2536 is positioned slightly higher in the active site and, as a result, is more likely to have a steric conflict between residues bulkier than Leu and the orthomethoxy group. This "higher" positioning of BI 2536 is most likely related to the arginine and phenylalanine residues that sit below the inhibitor in the positions corresponding to Gly1201 and Leu1256 in ALK-KD. The positioning of TAE684 in the ALK-KD active site suggests that when residues bulkier than Leu (e.g., Phe or Tyr) are in the hinge position corresponding to Leu1198 the other kinases may be able to avoid the steric clash by positioning the methoxy group underneath the aromatic plane of the side chain. However, this strategy of using the space that exists between the hinge and the aromatic plane of the side chain would work only for small ortho substituents. Therefore, adding larger substituents in ortho could be a valid strategy for improving ALK selectivity against kinases containing aromatic side chains in the position equivalent to Leu1198. Another strategy for improving the selectivity of TAE684 would be to replace the isopropyl group with something that extends deeper into the pocket formed by the Arg1253-Asn1254-Cys1255-Leu1256 and Gly1269-Asp1270 segments. As described above for PHA-E429, a deep extension into this pocket would only be tolerated by kinases containing glycine in the DFG − 1 position (Gly1269 in ALK). This approach should work for

obtaining selectivity versus kinases Aur-A, FLT3, and VEGFR3 that contain Ala, Cys, and Cys, respectively, in the DFG − 1 position (Table 2). Furthermore, this strategy might also work for obtaining selectivity versus IGF-1R and IRK because Leu1256 makes up one side of the cavity and the bulkier residue methionine, the corresponding residue in IGF-1R and IRK, would restrict the size of the cavity. Of the kinases having IC_{50} values of < 100 nM, the only two that have not been addressed by the strategies described above are ACK1 and FAK, two kinases containing leucines in the position corresponding to Leu1256 and glycines in the DFG − 1 position. However, both ACK1 and FAK have gatekeeper residues different from the Leu of ALK-KD, Thr and Met, respectively. Presumably, a substitution of the chlorine on the pyrimidine ring would modify the selectivity for these kinases, but it is not clear if this type of substitution could lead to ALK selectivity. Finally, sequence differences between the three kinases are observed for the acidic hinge residue that usually interacts with the ribose group of ATP; Asp1203 in ALK-KD corresponds to Ser in ACK1 and Glu in FAK. Thus, as specified above for PHA-E429, the ribose pocket could be another zone to target to improve the selectivity versus ACK1 and FAK and to potentially increase the affinity of inhibitors for ALK.

Comparing the ALK-KD−TAE684 and FAK−AZW592 Structures. Structures of the FAK kinase domain in complex with aminopyrimidine inhibitors that are analogues of TAE684 have been determined (*32*). The closest analogue to TAE684 is the compound AZW592 that contains an isopropylsulfonamide group rather than the isopropylsulfonyl on ring C and a methylaminopyrrolidine group rather than the piperidinyl-1-methylpiperazine group on ring A (Figure 7). Superposing the FAK−AZW592 structure (PDB entry 2JKM) onto the ALK-KD−TAE684 structure (1.23 Å rmsd for the 241 structurally equivalent Cα atoms) highlights the similarities and differences in the binding modes for the two inhibitors. Not surprisingly, both kinases make the same type of hinge interactions, and both position ring C underneath the glycine-rich loop. However, one significant difference between the two structures relates to the number of polar interactions that are made by the inhibitors; TAE684 makes only the two hinge interactions with ALK, while AZW592, in addition to the hinge interactions, also makes two additional interactions with the oxygens of the isopropylsulfonamide moiety. Specifically, one oxygen from the AZW592 isopropylsulfonamide moiety forms a hydrogen bond to Nζ from the conserved lysine in FAK and the other hydrogen bonds with a water molecule. Actually, in the FAK−AZW592 structure, a network of five water molecules fills a cavity that exists between the inhibitor and the residues of the DFG motif. In contrast, in the ALK-KD−TAE684 structure, this cavity does not exist and the residues of the DFG motif fill the region occupied by water in the FAK−AZW592 structure. Presumably, some of the differences in the positions of the DFG motifs are due to the peptide flip of Gly1269 in the ALK-KD structure that pushes the glycine up higher in the ATP pocket. Another relevant detail is the fact that the position and conformation of the DFG motif in the FAK−AZW592 structure are slightly distorted with respect to those of an active kinase. In effect, the side chain of the DFG Asp is rotated away from its usual position and the Gly has been shifted up by ~2 Å. It is also worth noting that the organization of the DFG motif in the FAK−AZW592 structure is significantly different from that of the other reported structures of the FAK kinase domain with aminopyrimidine inhibitors (*32*, *33*). Interestingly, in all of the other FAK structures, the residues of the DFG motif adopt a helical conformation that places the DFG + 1 residue (Leu) so that it interacts with the inhibitors in the active site.

Activating Mutations of ALK-KD in Neuroblastoma. In 2008, four separate groups reported the identification of activating mutations of the ALK kinase receptor in neuroblastoma (*14*−*17*), one of the most frequent solid tumors in children arising from the peripheral sympathetic nervous system. Combined, the groups identified 11 ALK-KD residues that were mutated in tumor samples from neuroblastoma patients or neuroblastoma-derived cell lines (Gly1128Ala, Thr1151Met, Met1166Arg, Ile1171Asn, Phe1174Ile, Phe1174Leu, Arg1192Pro, Ala1234Thr, Phe1245Cys, Phe1245Val, Ile1250Thr, Arg1275Gln, and Tyr1278Ser). At the time, efforts were made to localize the mutations using homology models. However, now that the ALK-KD structure is available it is worthwhile to revisit these mutations and contemplate what effects they might have on the activation process. Figure 8 indicates where these residues are located in the ALK-KD structure. The first two mutations, Gly1128Ala and Thr1151Met, would presumably have an impact on the phosphate-binding site because the first is directly on the glycine-rich loop and the second forms a hydrogen bond with Glu1129 of the glycine-rich loop and is adjacent in sequence to the catalytic lysine (Lys1150). In the ALK−PHA-E429 structure, residues Gly1128 and Glu1129 were disordered; however, both are visible in the ALK−TAE684 structure, and the side chain oxygen of Thr1151 forms a hydrogen bond with the carbonyl oxygen of Glu1129. The mutation of Gly1128 to Ala would effectively rigidify the glycine-rich loop, and the mutation of Thr1151 to Met would remove the possibility of forming a side chain hydrogen bond with Glu1129. Although it is not immediately clear how these mutations might contribute to the activation, one possibility is that the mutations might make it easier for ATP to gain access to the active site. Interestingly, three of the mutated residues (Met1166Arg, Arg1275Gln, and Tyr1278Ser) cluster together at the interface between the αC helix and the DFG helix (Figure 8). Both Arg1275 and Tyr1278 are on the upper side of the DFG helix, and Met1166 fills the cavity that exists between the two residues. Thus, the mutation of any of these three residues would most likely destabilize the DFG helix and consequently facilitate the shift of the αC helix. Also, as mentioned above, the hydroxyl group of Tyr1278 forms a hydrogen bond with the backbone nitrogen of Cys1097, so the mutation to Ser would result not only in a loss of hydrophobic interactions but also in the loss of a stabilizing hydrogen bond. Ile1171 is one of the four residues that make up the hydrophobic spine (*34*) of ALK-KD, the others being Cys1182, Phe1271 of the DFG motif, and His1247 of the HRD motif (Figure 8). On the basis of the model proposed by Kornev and co-workers, the dynamic assembly of the hydrophobic spine is the most important feature of kinase activation. Already, in the ALK−PHA-E429 structure, the hydrophobic spine is partially aligned but the packing of the residues is slightly loose, presumably due to the fact that the αC helix has not "flipped" into its active position. The mutation of Ile1171 to Asn might allow for the formation of a hydrogen bond with the backbone carbonyl of Val1180 which would help stabilize the hydrophobic spine and facilitate activation. Phe1174 is at the C-terminal end of the αC helix, where it is the central residue of a hydrophobic cluster that includes Phe1098, Phe1271 (DFG motif), and Phe1245 (Figure 8). As mentioned above, in the ALK−PHA-E429 structure, Phe1271 of the DFG motif is in the DFG-in conformation and Phe1245

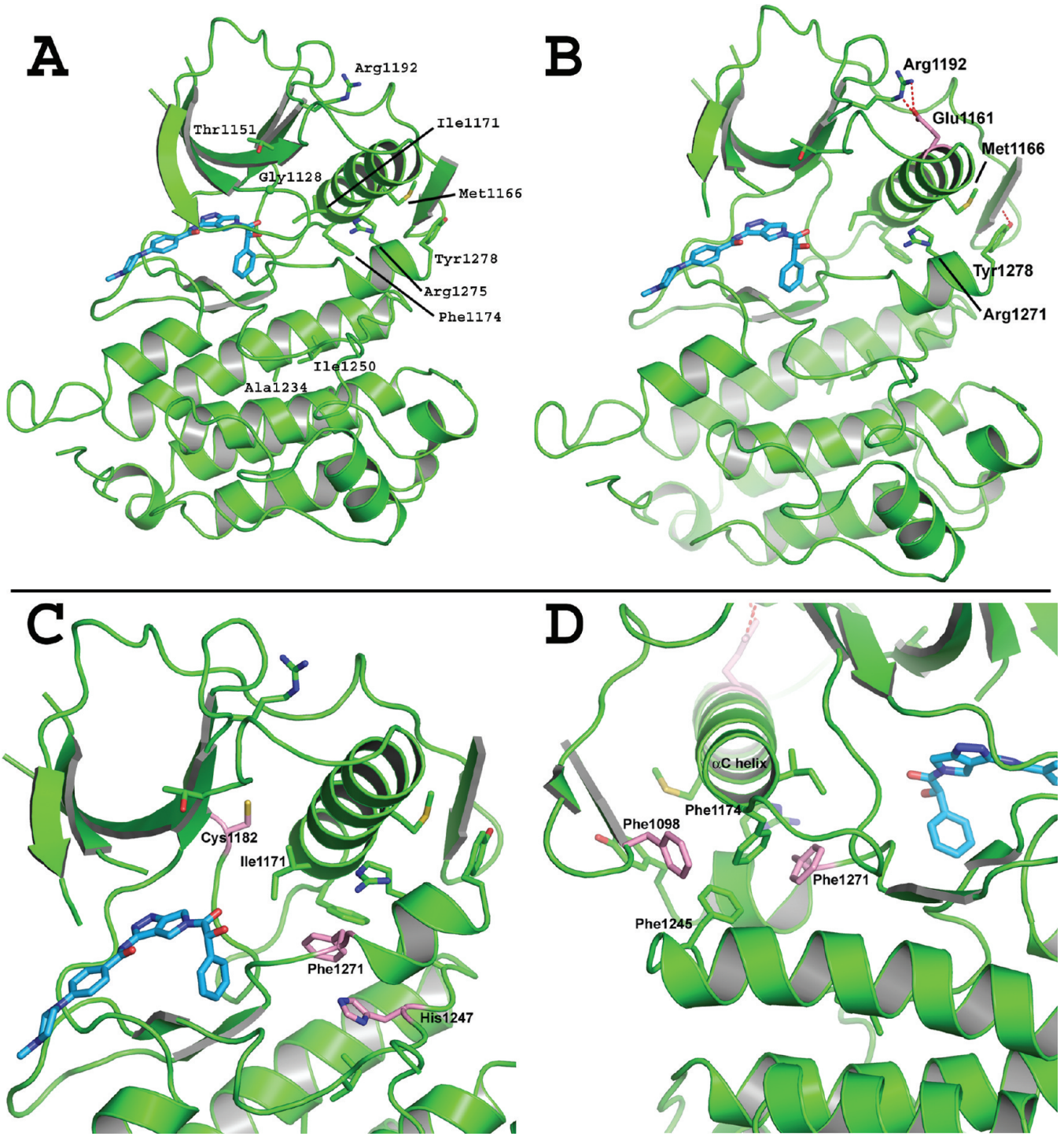

FIGURE 8: (A) Side chain positions for the 11 ALK-KD residues that were mutated in tumor samples from neuroblastoma patients or neuroblastoma-derived cell lines (Gly1128, Thr1151, Met1166, Ile1171, Phe1174, Arg1192, Ala1234, Phe1245, Ile1250, Arg1275, and Tyr1278). The Phe1245 side chain is poorly visible in the structure because it is hidden behind the DFG helix. (B−D) Close-up views of the ALK-KD−PHA-E429 structure highlighting the different mutated residues. Panel B shows residues Met1166, Arg1275, and Tyr1278 that are located at the interface between the αC helix and the DFG helix. Panel C shows the four residues of the hydrophobic spine. Panel D shows the cluster of four phenylalanines. Residues mutated in neuroblastoma are shown with green carbon atoms and other residues of interest with pink carbon atoms.

is in an "inactive" rotomer that is incompatible with the normal position of the β9 strand. To make space for the β9 strand, Phe1245 needs to rotate its side chain to the "active" position and, in so doing, forces Phe1098 to vacate its current position. However, for Phe1245 to rotate it needs to overcome a steric clash with the side chain of Phe1174 that comes from the packing of the hydrophobic cluster. Consequently, mutation of Phe1174 to either Ile or Leu would provide a shorter residue, thus relaxing the steric clash for the side chain rotation of Phe1245. In the same manner, mutation of Phe1245 to Cys or Val would also relieve the steric clash, allowing its side chain to rotate more freely. Arg1192 sits at the N-terminal end of β5 where it makes a salt bridge with Glu1161, a residue at the N-terminal end of the αC helix. Potentially, this salt bridge helps to hold the αC helix away from the C-lobe, and the mutation to Pro removes this interaction and facilitates the shift of the helix. The activating mutation of

Ala1234 to Thr is difficult to understand on the basis of its distance from the other key residues. However, one possibility is that since the location of Ala1234 on the αE helix precedes the loop containing Phe1245, its mutation induces a shift of αE and consequently a shift of Phe1245 away from the other two phenylalanines (Phe1174 and Phe1271), thus allowing it to more easily rotate to the active position. Ile1250 is the residue immediately after the HRD motif, and the side chain of Ile1250 packs against the side chain of His1247. As a result, the mutation of Ile1250 to Thr could influence the activation properties of the HRD which include the role of His1247 in the hydrophobic spine and the formation of a hydrogen bond between Arg1248 and the carbonyl oxygen of residue DFG + 1 that helps to stabilize the N-terminal end of β9.

DISCUSSION

Although protein kinases all adopt strikingly similar structures when they are active, significant differences are observed for the inactive structures which relates to the different mechanisms used to maintain the enzyme in a state of minimal activity (*35*). In most cases, the transition from the inactive to active state involves the coupling between the αC helix and the A-loop. However, in certain cases, the presence or absence of intermolecular or intramolecular interactions may also play a role in the transition. IRK and IGF-1R are examples of the first type because they both use an autoinhibition mechanism to maintain the inactive state by positioning the unphosphorylated A-loop so that it prevents the access of ATP and protein substrates to the active site (*36*, *37*). Src and Hck are examples of the second because the SH3 and SH2 domains make intramolecular interactions that lock the catalytic domain in an inactive conformation (*38*, *39*). On the basis of sequence similarities, we expected ALK-KD to use an autoinhibition mechanism that is similar to those of IRK and IGF-1R. Thus, finding evidence of a drastically different autoinhibition mechanism that uses intramolecular interactions between the N-terminal β-sheet and the DFG helix to lock the structure in an inactive conformation was surprising. The specific details of ALK's autoinhibition include (1) the inactive position of the αC helix which impacts the proper placement of the catalytic residues, (2) the order or disorder in the A-loop which precludes binding of the peptide substrate and also sequesters Tyr1278, making it inaccessible for phosphorylation, and (3) the degree of domain closure between the N- and C-terminal lobes that is not optimal for catalysis. In the ALK-KD−PHA-E429 structure, the DFG helix passes underneath the αC helix and connects to the N-terminal two-stranded antiparallel β-sheet through a hydrogen bond between Tyr1278 and Cys1097. During activation, the αC helix needs to shift ~4 Å toward the C-terminal lobe to properly position the catalytic residues. However, for the shift of the αC helix to occur, the DFG helix would have to unravel itself and, in the process, remove the link to the N-terminal β-sheet that involves Tyr1278. Presumably, the next step would be the phosphorylation of the three tyrosines and the formation of β9 and β10, but to make space for the formation of β9, Phe1245 would have to rotate its side chain to a position that would cause a steric clash with the side chain of Phe1098, a residue from β1′ of the two-stranded antiparallel β-sheet. Thus an initiating step for the formation of β9 would have to be the rotation of Phe1098 or, potentially, the complete removal of the N-terminal antiparallel β-sheet from the surface of ALK-KD. It is not clear how the autophosphorylation of ALK-KD fits with its autoinhibition, but it is interesting to note that in a study of NPM-ALK phosphorylation (*40*) Tyr1278 is reported to be the first site of phosphorylation and Tyr1096, a residue from β1′, is one of eight tyrosines phosphorylated outside of the A-loop. Specifically, Tyr1096 and Phe1098 are the two aromatic residues from the N-terminal β-sheet that packs against the C-terminal end of the αC helix. One explanation is that after activation is initiated the phosphorylation of the two tyrosines (Tyr1096 and Tyr1278) prevents them from returning to their positions in the autoinhibited conformation.

Several of the activating mutations in neuroblastoma influence, either directly or indirectly, the DFG helix and the N-terminal two-stranded antiparallel β-sheet. As a result, they provide independent experimental evidence supporting the autoinhibition mechanism described herein for ALK-KD. As cited above, three of the mutated residues cluster at the interface between the αC helix and DFG helix (Met1166Arg, Arg1275Gln, and Tyr1278Ser), and mutation of these residues presumably destabilizes the DFG helix and, in so doing, promotes the shift of the αC helix. Then there is the hydrophobic cluster involving the four phenylalanines (Phe1098, Phe1174, Phe1245, and Phe1271) that contains two of the residues that are mutated in neuroblastoma (Phe1174Ile, Phe1174Leu, Phe1245Cys, and Phe1245Val). Together, these mutations promote the formation of the β9 strand by removing the steric hindrance resulting from the rotamer position of Phe1245.

In general, ATP competitive kinase inhibitors can be divided into two types which depend on whether they take advantage of the hydrophobic pocket that results from the flip of the DFG motif. Type 1 inhibitors are the more common and recognize the active conformation that is conducive to phospho transfer. In contrast, type 2 inhibitors recognize the inactive DFG-out conformation resulting from the reorganization of the N-terminal portion of the activation loop that exposes the hydrophobic pocket. On the basis of our structural information, it does not appear likely that ALK-KD uses a flipping of the DFG motif. Therefore, we predict that type 2 inhibitors would be ineffective on ALK, leaving only type 1 inhibitors. This prediction correlates with the fact that all seven of our internally determined ALK-KD structures contain type 1 inhibitors and the DFG-in conformation. Nonetheless, it may be possible to exploit features of the ALK-KD inactive conformation for the design of selective inhibitors. Analysis of the ALK-KD structures with PHA-E429 and TAE684 provides valuable information concerning zones that could be exploited for the design of selective and potent ALK inhibitors. One zone that could potentially be exploited is the cavity defined by the Arg1253-Asn1254-Cys1255-Leu1256 and Gly1269-Asp1270 segments. As explained above, Gly1269 is flipped so that the carbonyl oxygen points toward the side chain of the conserved lysine (Lys1150) rather than making a conserved hydrogen bond with the histidine of the HRD motif. This flip of the glycine modifies the shape of the cavity, effectively extending the back of the pocket higher and reducing the access to Val1180 that would be present if the glycine were flipped the other way. Because glycine residues in this position are relatively rare and only 14.5% of kinases contain glycines in this position (*41*), targeting this pocket would be a logical strategy for improving selectivity. Other exploitable zones include the ribose pocket, the cavity near Val1180 and Leu1196 toward the back of the adenine pocket, and the cavity formed by the side chain of Glu1132 and the Leu1198-Met1199-Ala1200 hinge segment. Interestingly, the conversion of PHA-E429 to NMS-E828

exploits three of the four zones described above and shows a significant improvement in ALK selectivity and affinity.

In conclusion, this article reports the first crystal structures of the ALK kinase domain in complex with ATP competitive inhibitors, thus enabling structure-based drug design for this attractive oncology target. We also highlight zones of the active site that could be exploited to obtain selective ALK inhibitors. Furthermore, the ALK-KD−PHA-E429 structure led to the identification of a potential regulatory mechanism involving a link made between the DFG helix and the N-terminal two-stranded β-sheet. Finally, we were able to map the activating mutations associated with neuroblastoma onto our structure and by so doing provide explanations for the roles these residues might have in the activation process. Currently, it is not clear what impact the activating mutations will have on inhibitor affinity, and this will depend on the types of interactions made by the inhibitor and whether the same will be compatible with the constitutively active form of the enzyme.

ACKNOWLEDGMENT

We thank Sandrine Thieffine for the initial clones, Silvia Messali for support with the cells, and Elena Casale, Christian Orrenius, and Fabio Zuccotto for their helpful review of the manuscript.

SUPPORTING INFORMATION AVAILABLE

Experimental procedures describing the synthesis of inhibitors PHA-E589, NMS-E107, and NMS-E828. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Iwahara, T., Fujimoto, J., Wen, D., Cupples, R., Bucay, N., Arakawa, T., Mori, S., Ratzkin, B., and Yamamoto, T. (1997) Molecular characterization of ALK, a receptor tyrosine kinase expressed specifically in the nervous system. *Oncogene 14* (4), 439–449.
2. Morris, S. W., Naeve, C., Mathew, P., James, P. L., Kirstein, M. N., Cui, X., and Witte, D. P. (1997) 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 14* (18), 2175–2188.
3. Donella-Deana, A., Marin, O., Cesaro, L., Gunby, R. H., Ferrarese, A., Coluccia, A. M., Tartari, C. J., Mologni, L., Scapozza, L., Gambacorti-Passerini, C., and Pinna, L. A. (2005) Unique substrate specificity of anaplastic lymphoma kinase (ALK): Development of phosphoacceptor peptides for the assay of ALK activity. *Biochemistry 44* (23), 8533–8542.
4. Gunby, R. H., Ahmed, S., Sottocornola, R., Gasser, M., Redaelli, S., Mologni, L., Tartari, C. J., Belloni, V., Gambacorti-Passerini, C., and Scapozza, L. (2006) Structural insights into the ATP binding pocket of the anaplastic lymphoma kinase by site-directed mutagenesis, inhibitor binding analysis, and homology modeling. *J. Med. Chem. 49* (19), 5759–5768.
5. Morris, S. W., Kirstein, M. N., Valentine, M. B., Dittmer, K. G., Shapiro, D. N., Saltman, D. L., and Look, A. T. (1994) Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in non-Hodgkin's lymphoma. *Science 263* (5151), 1281–1284.
6. Shiota, M., Fujimoto, J., Semba, T., Satoh, H., Yamamoto, T., and Mori, S. (1994) Hyperphosphorylation of a novel 80 kDa protein-tyrosine kinase similar to Ltk in a human Ki-1 lymphoma cell line, AMS3. *Oncogene 9* (6), 1567–1574.
7. Chiarle, R., Gong, J. Z., Guasparri, I., Pesci, A., Cai, J., Liu, J., Simmons, W. J., Dhall, G., Howes, J., Piva, R., and Inghirami, G. (2003) NPM-ALK transgenic mice spontaneously develop T-cell lymphomas and plasma cell tumors. *Blood 101* (5), 1919–1927.
8. Jäger, R., Hahne, J., Jacob, A., Egert, A., Schenkel, J., Wernert, N., Schorle, H., and Wellmann, A. (2005) Mice transgenic for NPM-ALK develop non-Hodgkin lymphomas. *Anticancer Res. 25* (5), 3191–3196.
9. Piva, R., Chiarle, R., Manazza, A. D., Taulli, R., Simmons, W., Ambrogio, C., D'Escamard, V., Pellegrino, E., Ponzetto, C., Palestro, G., and Inghirami, G. (2006) Ablation of oncogenic ALK is a viable therapeutic approach for anaplastic large-cell lymphomas. *Blood 107* (2), 689–697.
10. Soda, M., Choi, Y. L., Enomoto, M., Takada, S., Yamashita, Y., Ishikawa, S., Fujiwara, S., Watanabe, H., Kurashina, K., Hatanaka, H., Bando, M., Ohno, S., Ishikawa, Y., Aburatani, H., Niki, T., Sohara, Y., Sugiyama, Y., and Mano, H. (2007) Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. *Nature 448* (7153), 561–566.
11. Rikova, K., Guo, A., Zeng, Q., Possemato, A., Yu, J., Haack, H., Nardone, J., Lee, K., Reeves, C., Li, Y., Hu, Y., Tan, Z., Stokes, M., Sullivan, L., Mitchell, J., Wetzel, R., Macneill, J., Ren, J. M., Yuan, J., Bakalarski, C. E., Villen, J., Kornhauser, J. M., Smith, B., Li, D., Zhou, X., Gygi, S. P., Gu, T. L., Polakiewicz, R. D., Rush, J., and Comb, M. J. (2007) Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer. *Cell 131* (6), 1190–1203.
12. Perner, S., Wagner, P. L., Demichelis, F., Mehra, R., Lafargue, C. J., Moss, B. J., Arbogast, S., Soltermann, A., Weder, W., Giordano, T. J., Beer, D. G., Rickman, D. S., Chinnaiyan, A. M., Moch, H., and Rubin, M. A. (2008) EML4-ALK fusion lung cancer: A rare acquired event. *Neoplasia 10* (3), 298–302.
13. Koivunen, J. P., Mermel, C., Zejnullahu, K., Murphy, C., Lifshits, E., Holmes, A. J., Choi, H. G., Kim, J., Chiang, D., Thomas, R., Lee, J., Richards, W. G., Sugarbaker, D. J., Ducko, C., Lindeman, N., Marcoux, J. P., Engelman, J. A., Gray, N. S., Lee, C., Meyerson, M., and Jänne, P. A. (2008) EML4-ALK fusion gene and efficacy of an ALK kinase inhibitor in lung cancer. *Clin. Cancer Res. 14* (13), 4275–4283.
14. Chen, Y., Takita, J., Choi, Y. L., Kato, M., Ohira, M., Sanada, M., Wang, L., Soda, M., Kikuchi, A., Igarashi, T., Nakagawara, A., Hayashi, Y., Mano, H., and Ogawa, S. (2008) Oncogenic mutations of ALK kinase in neuroblastoma. *Nature 455* (7215), 971–974.
15. George, R. E., Sanda, T., Hanna, M., Fröhling, S., Luther, W., II, Zhang, J., Ahn, Y., Zhou, W., London, W. B., McGrady, P., Xue, L., Zozulya, S., Gregor, V. E., Webb, T. R., Gray, N. S., Gilliland, D. G., Diller, L., Greulich, H., Morris, S. W., Meyerson, M., and Look, A. T. (2008) Activating mutations in ALK provide a therapeutic target in neuroblastoma. *Nature 455* (7215), 975–978.
16. Janoueix-Lerosey, I., Lequin, D., Brugières, L., Ribeiro, A., de Pontual, L., Combaret, V., Raynal, V., Puisieux, A., Schleiermacher, G., Pierron, G., Valteau-Couanet, D., Frebourg, T., Michon, J., Lyonnet, S., Amiel, J., and Delattre, O. (2008) Somatic and germline activating mutations of the ALK kinase receptor in neuroblastoma. *Nature 455* (7215), 967–970.
17. Mossé, Y. P., Laudenslager, M., Longo, L., Cole, K. A., Wood, A., Attiyeh, E. F., Laquaglia, M. J., Sennett, R., Lynch, J. E., Perri, P., Laureys, G., Speleman, F., Kim, C., Hou, C., Hakonarson, H., Torkamani, A., Schork, N. J., Brodeur, G. M., Tonini, G. P., Rappaport, E., Devoto, M., and Maris, J. M. (2008) Identification of ALK as a major familial neuroblastoma predisposition gene. *Nature 455* (7215), 930–935.
18. Fancelli, D., Moll, J., Varasi, M., Bravo, R., Artico, R., Berta, D., Bindi, S., Cameron, A., Candiani, I., Cappella, P., Carpinelli, P., Croci, W., Forte, B., Giorgini, M. L., Klapwijk, J., Marsiglio, A., Pesenti, E., Rocchetti, M., Roletto, F., Severino, D., Soncini, C., Storici, P., Tonani, R., Zugnoni, P., and Vianello, P. (2006) 1,4,5,6-Tetrahydropyrrolo[3,4-c]pyrazoles: Identification of a potent Aurora kinase inhibitor with a favorable antitumor kinase inhibition profile. *J. Med. Chem. 49* (24), 7247–7251.
19. Pevarello, P., Brasca, M. G., Amici, R., Orsini, P., Traquandi, G., Corti, L., Piutti, C., Sansonna, P., Villa, M., Pierce, B. S., Pulici, M., Giordano, P., Martina, K., Fritzen, E. L., Nugent, R. A., Casale, E., Cameron, A., Ciomei, M., Roletto, F., Isacchi, A., Fogliatto, G., Pesenti, E., Pastori, W., Marsiglio, A., Leach, K. L., Clare, P. M., Fiorentini, F., Varasi, M., Vulpetti, A., and Warpehoski, M. A. (2004) 3-Aminopyrazole inhibitors of CDK2/cyclin A as antitumor agents. 1. Lead finding. *J. Med. Chem. 47* (13), 3367–3380.
20. Tamaoki, T., Nomoto, H., Takahashi, I., Kato, Y., Morimoto, M., and Tomita, F. (1986) Staurosporine, a potent inhibitor of phospholipid/Ca^{++} dependent protein kinase. *Biochem. Biophys. Res. Commun. 135* (2), 397–402.
21. Collaborative Computational Project, Number 4 (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr. D50* (5), 760–763.
22. McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr. 40*, 658–674.

23. Emsley, P., and Cowtan, K. (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D60*, 2126–2132.
24. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of Macromolecular Structures by the Maximum-Likelihood method. *Acta Crystallogr. D53*, 240–255.
25. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography and NMR System: A new software suite for macromolecular structure determination. *Acta Crystallogr. D54*, 905–921.
26. Potterton, L., McNicholas, S., Krissinel, E., Gruber, J., Cowtan, K., Emsley, P., Murshudov, G. N., Cohen, S., Perrakis, A., and Noble, M. (2004) Developments in the CCP4 molecular-graphics project. *Acta Crystallogr. D60*, 2288–2294.
27. Delano, W. L. (2008) The PyMol Molecular Graphics System, DeLano Scientific LLC, Palo Alto, CA.
28. Knighton, D. R., Zheng, J. H., Ten Eyck, L. F., Ashford, V. A., Xuong, N. H., Taylor, S. S., and Sowadski, J. M. (1991) Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science 253* (5018), 407–414.
29. Hubbard, S. R. (1997) Crystal structure of the activated insulin receptor tyrosine kinase in complex with peptide substrate and ATP analog. *EMBO J. 16* (18), 5572–5581.
30. Galkin, A. V., Melnick, J. S., Kim, S., Hood, T. L., Li, N., Li, L., Xia, G., Steensma, R., Chopiuk, G., Jiang, J., Wan, Y., Ding, P., Liu, Y., Sun, F., Schultz, P. G., Gray, N. S., and Warmuth, M. (2007) Identification of NVP-TAE684, a potent, selective, and efficacious inhibitor of NPM-ALK. *Proc. Natl. Acad. Sci. U.S.A. 104* (1), 270–275.
31. Kothe, M., Kohls, D., Low, S., Coli, R., Rennie, G. R., Feru, F., Kuhn, C., and Ding, Y. H. (2007) Selectivity-determining residues in Plk1. *Chem. Biol. Drug Des. 70* (6), 540–546.
32. Lietha, D., and Eck, M. J. (2008) Crystal Structures of the FAK Kinase in Complex with TAE226 and Related Bis-Anilino Pyrimidine Inhibitors Reveal a Helical DFG Conformation. *PLoS One 3* (11), e3800.
33. Roberts, W. G., Ung, E., Whalen, P., Cooper, B., Hulford, C., Autry, C., Richter, D., Emerson, E., Lin, J., Kath, J., Coleman, K., Yao, L., Martinez-Alsina, L., Lorenzen, M., Berliner, M., Luzzio, M., Patel, N., Schmitt, E., LaGreca, S., Jani, J., Wessel, M., Marr, E., Griffor, M., and Vajdos, F. (2008) Antitumor activity and pharmacology of a selective focal adhesion kinase inhibitor, PF-562,271. *Cancer Res. 68* (6), 1935–1944.
34. Kornev, A. P., Haste, N. M., Taylor, S. S., and Ten Eyck, L. F. (2006) Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism. *Proc. Natl. Acad. Sci. U.S.A. 103* (47), 17783–17788.
35. Huse, M., and Kuriyan, J. (2002) The conformational plasticity of protein kinases. *Cell 109* (3), 275–282.
36. Hubbard, S. R., Wei, L., and Hendrickson, W. A. (1994) Crystal structure of the tyrosine kinase domain of the human insulin receptor. *Nature 372*, 746–754.
37. Munshi, S., Kornienko, M., Hall, D. L., Reid, J. C., Waxman, L., Stirdivant, S. M., Darke, P. L., and Kuo, L. C. (2002) Crystal structure of the Apo, unactivated insulin-like growth factor-1 receptor kinase. Implication for inhibitor specificity. *J. Biol. Chem. 277* (41), 38797–38802.
38. Schindler, T., Sicheri, F., Pico, A., Gazit, A., Levitzki, A., and Kuriyan, J. (1999) Crystal structure of Hck in complex with a Src family-selective tyrosine kinase inhibitor. *Mol. Cell 3* (5), 639–648.
39. Xu, W., Doshi, A., Lei, M., Eck, M. J., and Harrison, S. C. (1999) Crystal structures of c-Src reveal features of its autoinhibitory mechanism. *Mol. Cell 3* (5), 629–638.
40. Wang, P., Wu, F., Ma, Y., Li, L., Lai, R., and Young, L. C. (2010) Functional characterization of the kinase activation loop in nucleophosmin (NPM)-anaplastic lymphoma kinase (ALK) using tandem affinity purification and liquid chromatography-mass spectrometry. *J. Biol. Chem. 285* (1), 95–103.
41. Vulpetti, A., and Bosotti, R. (2004) Sequence and structural analysis of kinase ATP pocket residues. *Farmaco 59* (10), 759–765.
42. Gouet, P., Courcelle, E., Stuart, D. I., and Metoz, F. (1999) ESPript: Multiple sequence alignments in PostScript. *Bioinformatics 15*, 305–308.