PAPER

How does conformational flexibility influence key structural features involved in activation of anaplastic lymphoma kinase?†

Tatyana G. Karabencheva,*[a] Christian C. Lee,[b] Gary W. Black,[a] Rossen Donev[c] and Christo Z. Christov*[a]

Anaplastic Lymphoma Kinase (ALK) plays a major role in developing tumor processes and therefore has emerged as a validated therapeutic target. Applying atomistic molecular dynamics simulations on the wild type enzyme and the nine most frequently occurring and clinically important activation mutants we revealed important conformational effects on key interactions responsible for the activation of the enzyme.

Introduction

Anaplastic Lymphoma Kinase (ALK) is a member of the insulin receptor tyrosine kinase superfamily that normally participates in the development of the nervous system in embryos, the expression of which decreases after birth.[1,2] Twenty different ALK-fusion proteins resulting from different chromosomal rearrangements have been shown to be connected with the pathogenesis of diseases such as anaplastic large-cell lymphoma, diffuse large B-cell lymphoma and inflammatory myofibroblastic tumors.[3] A series of germline and somatic single nucleotide polymorphisms in the catalytic domain of ALK were characterized recently from neuroblastoma patient samples and many were shown to be related with deregulation of ALK and its constitutive activation.[4–6] Recently resolved crystal structures of the wild type enzyme,[7] its complexes with adenosine diphosphate (ADP), and staurosporine (STU),[7] adenosine triphosphate (ATP) competitive inhibitors – PHA-E429, NVP-TAE684[8] and of two activation mutants[9] revealed a canonical protein kinase domain fold and topology. The molecule is built of a smaller N-terminal lobe and a larger C-terminal lobe (Fig. 1).[7] The N-terminal lobe contains five stranded antiparallel β-sheets, an important regulatory αC-helix which can be in active and inactive positions and an ATP-binding glycine-rich loop (P-loop). The juxtamembrane segment of the
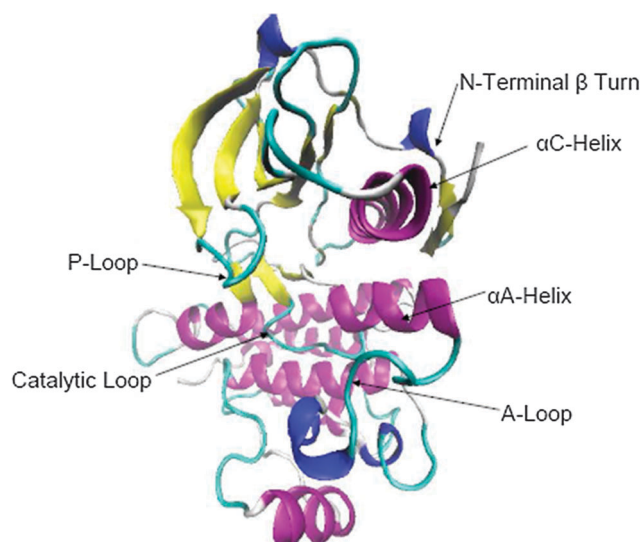

Fig. 1 Structure of ALK (3L9P.pdb[7]). The α-helices are shown in purple, the β-sheets are shown in yellow, and the 3–10 helices are shown in blue, the loops are shown in cyan and the coils in white shown in red.

N-terminus forms a β-turn motif which is unique for ALK. The C-terminal lobe contains eight α-helices (αD–αI) and two β stranded sheets. The activating loop (A-loop) which contains residues 1270–1299 is located here and starts with a DFG-motif, followed by a short αA-helix and ends with a Pro-Pro-Glu motif. The catalytic loop is located between αE and the first strand of the β-sheet in the C-lobe. The αA-helix is located below and orthogonally to the αC-helix and against the N-terminal β-turn by multiple polar and hydrophobic interactions.

The crystal structures display several features of inactive kinase conformation such as the degree of lobe closure between N- and

[a] Department of Applied Sciences, Faculty of Health and Life Sciences, Northumbria University, Newcastle upon Tyne, NE1 8ST, UK. E-mail: christo.christov@northumbria.ac.uk, tatyana.karabencheva-christova@northumbria.ac.uk

[b] The Genomics Institute of the Novartis Research Foundation, San Diego, CA 92121, USA

[c] Biomed Consult Ltd, Swansea SA3 5LF, UK

C-terminal lobes, the position of the αC helix and its orientation with respect to the C-lobe and the location of the distal part of the A-loop which obstructs the peptide binding site. The structure does not possess, however, all the negative regulatory elements that exist in the insulin receptor kinases. The A-loop is in an intermediate position between the inactive and active kinases and its proximal part does not block the ATP-binding site as in other inactive kinase conformations. The DFG-motif exhibits a "DFG-in" conformation as in the case in the active kinases, and the αC-helix allows the proper positioning of Glu1167 with Lys1150 for salt-bridge formation, which is also typical for the active conformation.[7]

ALK activation involves complex changes in several structural features and interactions including but not limited to the mutual orientation of the αC and αA helices, the position of the A-loop, the hydrogen bond between Tyr1278 and Cys1097 from the N-terminal β-turn which locks the inactive conformation and the salt bridge between Glu1167 and Lys1150 which is distinctive for the active conformations.[1,7] Lack of insights into conformational dynamics is a major limitation of X-ray protein crystal structures as they are largely static in nature and reflect some of the details of the experimental methodology, such as ensemble and time averaged properties, imposition of crystal packing effects, and some deviations due to the non-linearity of the registered signals and methods of interpretation.[10] Most importantly, crystal structures could miss some important information about the conformational dynamics and relaxation which are fundamental protein properties.[11,12] The delicate interactions which define such unique intermediate conformations would be influenced by the internal conformational flexibility under native conditions; therefore exploration using molecular dynamics (MD) is important in order to complement the crystallographic structural information. In addition the correlation between the protein dynamics in enzyme catalysis is mater of intensive studies investigations and discussions.[13–16] Recently molecular dynamics study of the on the crizotinib resistance mechanism of C1156Y mutation in ALK was carried on (ref. 17) demonstrating the potential of molecular dynamics to investigate atomistic effects of mutations, however it was carried on only for 15 ns and one mutation only.

The focus of our study is to establish whether the structure of ALK, revealed by crystallographic studies as an intermediate between the active and inactive conformations, is altered by the conformational flexibility and to define the interactions responsible for this. To achieve this insight we carried out a comprehensive set of atomistic molecular dynamics (MD) Simulations followed by detailed analysis of the time evolution and flexibility of multiple interactions representative for both the inactive and active ALK conformations. Here we discuss the influence of the conformational flexibility of the key structural features of the inactive and active conformations whilst the more comprehensive structure/mobility/to activity relationship would be an objective of further project.

Methods

The MD simulations were carried out with the GROMACS code (version 4.3.1)[18] and Gromos43b1 force field for

100 nanoseconds (ns). Twelve MD simulations were carried out for the wild-type apoenzyme ALK, ALK–ADP complex, ALK–staurosporine complex and nine activation mutations – M1166R, I1171N, F1174L, F1245C, I1250T, G1128A, T1151M, R1192P and R1275Q.[4–6] The system was prepared from the following crystal structures of ALK taken from PDB[19] as follows: for the wild type free ALK MD simulation 3L9P.pdb structure was used;[7] for the ALK–ADP complex the 3LCT.pdb structure which contains the ligand ADP was used;[7] for ALK–staurosporine complex, 3LCS.pdb structure which contains the inhibitor staurosporine was used;[7] for the nine simulations of the activation mutant forms of the free ALK the 3L9P.pdb structure was used from which all nine structures of activation mutations were prepared from the crystal structure of the wild-type enzyme using the "What if" web interface (http://swift.cmbi.ru.nl/servers/html/index.html).[20] In all simulations the missing residues in the crystal structures of ALK were added using PyMol.[21] Gromacs utilities for system preparation were used thoroughly. The hydrogen atoms were added according the protonation states of the ionogenic groups and consequently minimized. Gromos43a1 force field was thoroughly applied. The ADP and staurosporine parameters were generated using PRODRG web-based tool.[22] Subsequently the entire system was minimized using the steepest descent algorithm. The system was then solvated using rectangular SPC[23] water box placed 10 Å from the edges of the protein and neutralized. The solvent was then equilibrated for 50 picoseconds (ps) at 310 K using NVT ensemble. Production MD simulations were run for 100 ns in NPT ensemble at 310 K applying Berendsen thermostat[24] treating the electrostatic interactions by particle mesh Ewald method.[25] The simulation step was 2 fs and nonbonding cutoff of 14 Å was applied. The simulation system is presented on Fig. 2.

The quality of the simulations was assessed by Root Mean Square Deviation (RMSD) (Fig. S1–S3A–I, ESI†), root mean
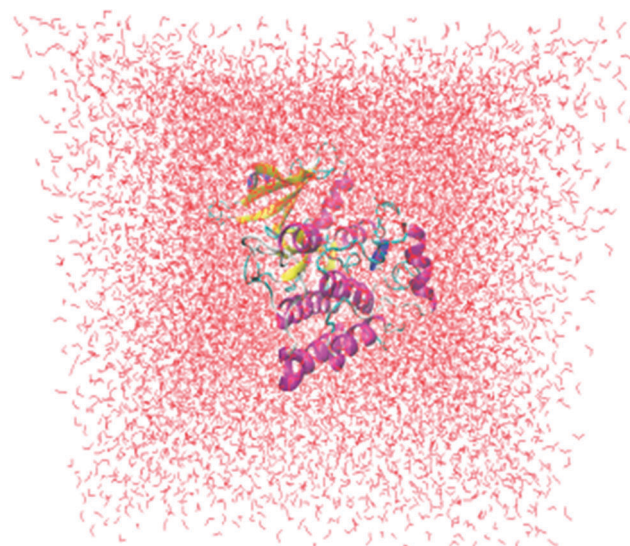

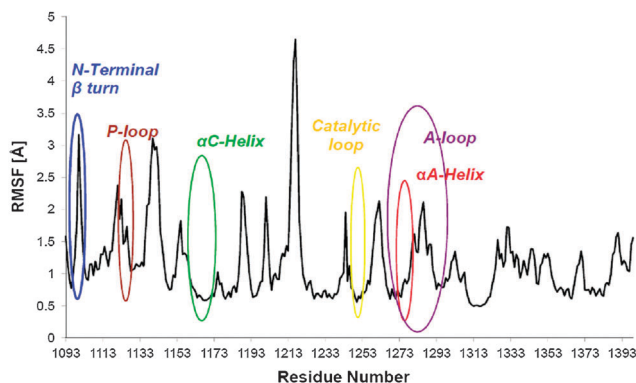Fig. 2 Solvated and neutralized system for MD simulations.

Fig. 3   RMSF values of the wild type free enzyme.

Fig. 4   Interactions between the aminoacid residues in the αC- and αA-helices.

square fluctuation (Fig. 3 in the main text and Fig. S4A–D and S5A–D in ESI†), radius of gyration and total, kinetic and potential energy values long the time of the simulation (the plots are available from the authors by request). All calculations were done for the αC atoms in reference to the initial mini-mized structure. The stability of hydrogen bonds during the simulations was assessed using Gromacs g_hbond utility. After that the average number of the hydrogen bonds between particular couples of residues per MD trajectory was extracted. Salt-bridge distances were calculated using g_salt utility and the averaged value for a particular couple of residues was calculated for each trajectory. Several hydrophobic interactions were calculated by g_dist command. The cluster analysis was done using cutoff distance of 0.25 nm and the central structure of the largest cluster was used for comparison with the mini-mized crystal structure.

Results and discussion

The N-terminal β-turn, followed by the glycine-rich (P-loop) and the A-loop (Fig. 3) exhibit high flexibility in the wild type free enzyme (based on the RMSF values) which suggest that during the process of activation they would undergo sensitive changes. In the simulations of the structures with activation mutants the N-terminal β-turn is characterized with higher flexibility than the wild type enzyme in all mutants with the exception of I1171N and T1151M (Fig. S5 in ESI†). The flexibility of the A-loop is also increased by the activating mutations (with excep-tion of G1128A). The P-loop is characterized with increased flexibility in five mutants (I1250T, R1275Q, M1166R, R1192P and T1151M). The N-terminal β-turn, glycine-rich (P-loop) and the A-loop which are very flexible in the wild type enzyme and as a result of the activating mutations their flexibility is increased even more. The binding of ADP and staurosporine does not influence the flexibility of the catalytic loop, αA- and αC-helices (Fig. S4 in the ESI†) but (especially the binding of staurosporine) leads to increased flexibility of the P-loop and the A-loop (excluding the αA-helix).

The orientation between the αC- and αA-helices which is in is key importance for switching between the inactive and active con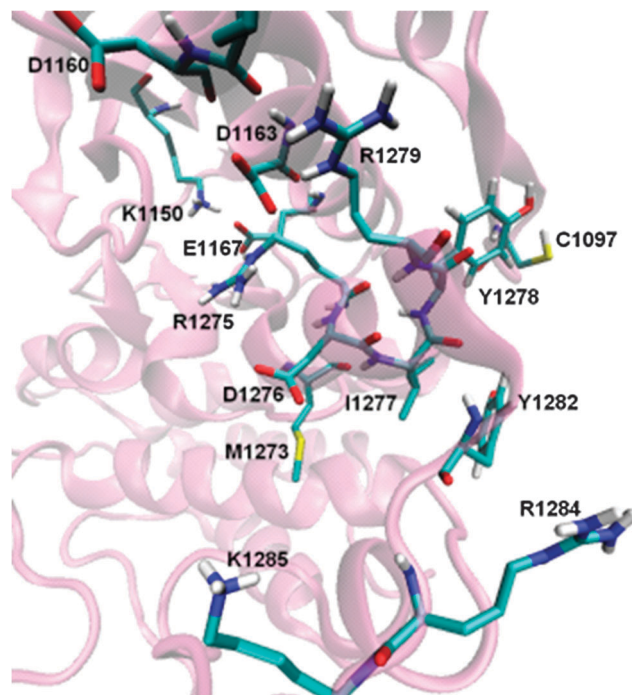formations is stabilized by hydrogen bonds and electrostatic interactions (Fig. 4). An insight about the orientation of the αA and αC helices is provided by analysis of the hydrogen bonding and electrostatic interactions between them (Tables 1 and 2). The hydrogen bond between the side chain of Arg1275 and the carbonyl oxygen of Asp1163 is presented in 60% and 70% from the MD trajectories of T1151M and R1275Q mutants, 40% in the wild type enzyme and the F1245C, I1250T, I1171N mutants and with lower frequency in the G1128A, M1166R and F1174L. Arg1279 interacts with Asp1163, as in the crystal structure, which is indicated by both the hydrogen bonding profile and the salt-bridged distance but this interaction tends to become less strong. In the wild type Arg1275 and Glu1167 still interact electrostatically and by a hydrogen bond, however in most of the mutant forms, this hydrogen bond is presented with lower frequency. The hydrogen bond between Arg1279 and Gln1159 is stable in the MD trajectory of the wild type ALK and is presented with differential stability in the mutants (e.g. it is relatively stable in FF1174L and T1151M but is missing in G1128A and R1275Q).

The side chain of Asp1276 rotates towards the C-lobe during the simulation and the hydrogen bond between Arg1275 and Asp1276, which exists in the crystal structure, is not presented in the simulation of the wild type and mutant forms (the averaged distance between the side chains is higher than in the crystal structure; 4.2 Å compared to 2.0 Å for the wild type and even large differences in the mutant forms). The hydrogen bond between Lys1285 from the A-loop and the αC-helix residue Asp1163, disappears during the simulation of the wild type and all activation mutant enzyme forms and Lys1285 moves away from the αC-helix, due to the flexibility of the A-loop. The repulsive interactions between the side chains of Arg1275 and

Table 1 Averaged number of hydrogen bonds from the MD trajectories of the wild type apoenzyme ALK and all activation mutants

H-bond	Cryst. struct.	Wild type	F1174L	F1245C	G1128A	I1250T	I1171N	M1166R	R1275Q	T1151M
R1275-D1163	1	0.4	0.2	0.4	0.3	0.4	0.4	0.3	0.7	0.6
R1279-D1163	1	1.6	1.4	1.4	0.4	1.1	1.6	1.0	1.2	1.5
R1275-E1167	1	0.8	0.5	0.5	0.6	0.9	0.4	0.4	0	0.5
K1285-D1163	1	0	0	0	0	0	0	0	0	0
D1285-D1160	0	0	0	0	0	0	0	0	0	0
K1285-E1167	0	0	0	0	0	0	0	0	0	0
D1276-R1275	1	0	0	0	0	0	0.3	0	0	0
R1279-Q1159	1	0.7	0.8	0.5	0	0.6	0.5	0.4	0	0.9
Y1278-C1097	1	0.6	0.5	0.2	0	0	0.7	0	0.7	0
Y1282-R1284	1	0.1	0	0	0	0	0	0.1	0.2	0
K1150-E1167	0	0	0.5	0.6	0	0.5	0	0.5	0.3	0.9
R1275-R1248	0	0	0	0	0	0	0	0	0	0
D1276-R1248	0	0	0	0	0	0	0	0	0	0
R1248-I1246	0	0.2	0	0	0	0.1	0	0	0	0.3
M1273-R1248	0	0	0	0	0	0	0	0	0	0
M1273-I1277	1	0.98	0.8	0.9	0.7	0.8	1.0	0.9	1.0	0.8
H1247-G1269	0	0.1	0.3	0	0	0	0	0	0	0

Table 2 Averaged electrostatic interactions in the wild type and the activation mutant forms of the free ALK

Interact	Cryst. struct.	Wild type	F1174L	F1245C	G1128A	I1250T	I1171N	M1166R	R1275Q	T1151M
R1275-D1163	3.3	2.95	3.3	3.1	3.0	2.8	3.3	3.1	5.8	2.7
R1279-D1163	2.1	2.2	2.5	2.4	4.0	2.5	2.2	2.8	2.3	2.2
R1275-E1167	1.9	2.4	2.6	2.7	2.5	2.3	2.7	2.8	6.3	2.9
K1285-D1163	2.6	13.6	11.5	12.2	7.8	11.0	8.6	9.1	11.6	10.0
D1285-D1160	4.4	15.7	13.5	15.4	9.3	12.4	7.9	9.4	13.0	11.6
K1285-E1167	9.2	17.4	16.4	16.7	13.3	14.8	10.3	10.7	11.4	14.3
D1276-R1275	1.99	4.2	4.8	4.3	4.4	4.9	3.2	3.9	5.2	4.5
R1279-Q1159	6.6	6.6	7.4	7.1	9.2	6.95	6.7	7.3	8.96	6.9
Y1278-C1097	6.4	6.4	7.1	7.7	5.8	7.0	6.3	8.9	6.4	8.3
Y1282-R1284	6.8	6.8	6.94	6.7	5.9	6.8	7.9	6.7	5.7	8.8
K1150-E1167	4.7	3.2	2.5	2.5	2.7	2.5	2.9	2.5	2.9	2.2
R1275-R1248	10.5	10.0	12.0	12.1	12.6	11.5	12.1	11.7	10.8	10.6
D1276-R1248	9.7	9.7	9.8	9.7	11.5	9.0	10.2	9.7	9.8	7.9
M1273-I1277	6.1	6.1	6.6	6.5	6.5	6.5	5.95	6.3	5.8	6.4

Arg1279 with the side chain of Lys1285 might also contribute to this effect. The balance between these repulsive interactions and the attractive interactions between the side chains of Lys1285 and Asp1163, Asp1160, Glu1167 and Asp1276, together with the high flexibility of the A-loop, can be a controlling instrument for the positioning of the distal part of the A-loop in more active-like orientation and thus postulated to be part from the activation mechanism. The hydrogen bonding and electrostatic interactions between αC and αA helices become less strong and many of them partly or almost disappear during the simulations of the wild type and mutant forms. This could be an initial step from more long-range process of reorientation of the αC and αA helices as a part from the enzyme activation.

An important interaction that locks the αC helix in inactive conformation is the hydrogen bond between Tyr1278 from the A-loop (the preferred tyrosine residue for the autophosphorylation reaction), and Cys1097, from the N-terminal β-turn. It is stable during 60% of the wild type ALK MD simulation and is also observed with different frequency in F1174F, I1171N, R1275Q and F1245C but missing in F1245C, G1128A, I1250T, M1166R and T1151M trajectories (Table 1 in the main text and Table S1 in the ESI†) which suggest that the activation effect of this mutants might involve change of the orientation between the N-terminal β-turn and αC helix.

Importantly, the hydrogen bond between the second A-loop tyrosine Tyr1282 and Arg1284 also almost disappears in the simulations of the wild type and the mutant forms of the enzyme which also might be part from the initiation events of the activation. This could be facilitated by the flexibility of the A-loop and the stronger interactions between the side chains of Arg1284 and Glu1303 and between the side chains of Tyr1282 and Arg1248. In contrast the hydrogen bond, representative for the inactive conformation between the peptide backbones of DFG + 1 residue Met1273 and Ile1277 that exists in the crystal structure is preserved over the simulation with very high frequency and is also present in all activation mutants. The existence of this interaction in ALK is in contrast to all known active kinases.

Distinctive interaction for the active kinases conformations is the salt bridge between Lys1150 and Glu1167. It is stronger than in the crystal structure during the simulation of the wild type enzyme and even much stronger in all mutant forms (Table 1) which indicate that the structure became more active-like regarding this interaction.

The comparison between the minimized crystal structure and the central structure of the most populated cluster (Fig. 5)
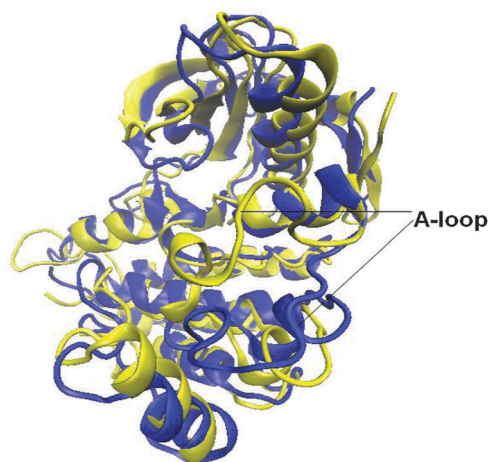
Fig. 5 Superimposed MD-averaged and crystal structures of ALK. The MD structure is in blue and the crystal structure is in yellow.

from the simulation of the wild type enzyme confirms that the A-loop moved more towards the active position, which is more expressed in its distal part during the simulation, and the P-loop moved more towards the catalytic loop, whilst the inter-lobe did not change considerably during the simulation. Two interactions might be related to the positioning of the P-loop closer to the catalytic loop. The salt bridge between Lys1150 and Glu1167, which is characteristic for the active conformation, is strengthen during the MD simulation of the wild type and all mutants. In addition Arg1120 switches between two glutamate residues – Glu1132 and Glu1129.

The results provide important insight revealing that numerous interactions undergo initial changes relevant to the activation, namely: the hydrogen bonds and electrostatic interactions which keep the inactive orientation between αC and αA helices are less strong than in the crystal structure as well as the hydrogen bond between the αC-helix to the N-terminal β-turn; the A-loop is displaced in the more active position due to the repulsive interactions between Lys1285 with Arg1275 and Arg1279. Essentially, the salt-bridge between Glu1167 and Lys1150 which is characteristic for the active conformation is strengthened and might help for displacing of the P-loop towards the catalytic loop. The N-terminal β-turn which is unique for ALK exhibits high flexibility in the wild type ALK and all mutant forms. The activating mutations tend to increase even more the conformational flexibility with most profound effect of the A-loop and N-terminal β-turn and the P-loop confirming their role during the activation process. The complex process of activation might be initiated by several events such as loosening the electrostatic interactions and hydrogen bonding between the αC- and αA-helices, increasing the flexibility of the A-loop and displacing it in a more active position; increasing the flexibility the N-terminal β-turn in combination with the loosening and breaking of the "inactive conformation" hydrogen bond between Tyr1278 and Cys1097, strengthening the "active conformation interaction" between Glu1167 and Lys1150; and displacing of the P-loop towards the

active site. The gained atomistic insight can be used for further atomistic and mechanistic studies and as a background in designing more effective ALK inhibitors as potential drugs.

References

1 R. Roskoski Jr., Anaplastic lymphoma kinase (ALK): Structure, oncogenic activation, and pharmacological inhibition, *Pharmacol. Res.*, 2013, **68**, 68.

2 S. Morris, M. Kirstein, M. Valentine, K. Dittmer, D. Shapiro, D. Saltman and A. Look, Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in non-Hodgkin's lymphoma, *Science*, 1994, **263**, 1281.

3 F. Tabbo, A. Barreca, R. Piva and G. Inghirami, ALK signaling and target therapy in anaplastic large cell lymphoma, *Front. Oncol.*, 2012, **2**, 41.

4 Y. Chen, J. Takita, Y. L. Choi, M. Kato, M. Ohira, M. Sanada, L. Wang, M. Soda, A. Kikuchi, T. Igarashi, A. Nakagawara, Y. Hayashi, H. Mano and S. Ogawa, Oncogenic mutations of ALK kinase in neuroblastoma, *Nature*, 2008, **455**, 971.

5 Y. P. Mosse, M. Laudenslager, L. Longo, K. A. Cole, A. Wood, E. F. Attiyeh, M. J. Laquaglia, R. Sennett, J. E. Lynch, P. Perri, G. Laureys, F. Speleman, C. Kim, C. Hou, H. Hakonarson, A. Torkamani, N. J. Schork, G. M. Brodeur, G. P. Tonini, E. Rappaport, M. Devoto and J. M. Maris, Identification of ALK as a major familial neuroblastoma predisposition gene, *Nature*, 2008, **455**, 930.

6 R. E. George, T. Sanda, M. Hanna, S. Fröhling, W. Luther II, J. Zhang, Y. Ahn, W. Zhou, W. B. London, P. McGrady, L. Xue, S. Zozulya, V. E. Gregor, T. R. Webb, N. S. Gray, D. G. Gilliland, L. Diller, H. Greulich, S. W. Morris, M. Meyerson and A. T. Look, Activating mutations in ALK provide a therapeutic target in neuroblastoma, *Nature*, 2008, **455**, 975.

7 C. C. Lee, Y. Jia, N. Li, X. Sun, K. Ng, E. Ambing, M.-Y. Gao, S. Hua, C. Chen, S. Kim, P.-Y. Michellys, S. A. Lesley, J. L. Harris and G. Spraggon, Crystal structure of the ALK (anaplastic lymphoma kinase) catalytic domain, *Biochem. J.*, 2010, **430**, 425.

8 R. T. Bossi, M. B. Saccardo, E. Ardini, M. Menichincheri, L. Rusconi, P. Magnaghi, P. Orsini, N. Avanzi, A. L. Borgia, M. Nesi, T. Bandiera, G. Fogliatto and J. A. Bertrand, Crystal Structures of Anaplastic Lymphoma Kinase in Complex with ATP Competitive Inhibitors, *Biochemistry*, 2010, **49**, 6813.

9 L. F. Epstein, H. Chen, R. Emkey and D. A. Whittington, he R1275Q Neuroblastoma Mutant and Certain ATP-competitive Inhibitors Stabilize Alternative Activation Loop Conformations of Anaplastic Lymphoma Kinase, *J. Biol. Chem.*, 2012, **287**, 37447.

10 D. Kruschel and B. Zagrovic, Conformational averaging in structural biology: issues, challenges and computational solutions, *Mol. BioSyst.*, 2009, **5**, 1606.

11 S. A. Adcock and J. A. McCammon, Molecular dynamics: Survey of methods for simulating the activity of proteins, *Chem. Rev.*, 2006, **106**, 1589.

12 M. Karplus and J. A. McCammon, Molecular dynamics simulations of biomolecules, *Nat. Struct. Biol.*, 2002, **9**, 646.

13 D. Glowacki, J. N. Harvey and A. Mulholland, Taking Ockham's razor to enzyme dynamics and catalysis, *Nat. Chem.*, 2012, **4**, 169.

14 M. Karplus and J. Kuriyan, Molecular dynamics and protein function, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 6679.

15 C. Z. Christov, A. Lodola, T. G. Karabencheva-Christova, S. Wan, P. V. Coveney and A. J. Mulholland, Conformational Effects on the pro-S Hydrogen Abstraction Reaction in Cyclooxygenase-1: An Integrated QM/MM and MD Study, *Biophys. J.*, 2013, **104**, L5.

16 A. Lodola, J. Sirirak, N. Fey, S. Rivara, M. Mor and A. J. Mulholland, Structural Fluctuations in Enzyme-Catalyzed Reactions: Determinants of Reactivity in Fatty Acid Amide Hydrolase from Multivariate Statistical Analysis of Quantum Mechanics/Molecular Mechanics Paths, *J. Chem. Theory Comput.*, 2010, **6**, 2948.

17 H.-Y. Sun and F.-Q. Ji, A molecular dynamics investigation on the crizotinib resistance mechanism of C1156Y mutation in ALK, *Biochem. Biophys. Res. Commun.*, 2012, **423**, 319.

18 B. Hess, C. Kutzner, D. van der Spoel and E. Lindahl, GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation, *J. Chem. Theory Comput.*, 2008, **4**, 435.

19 H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov and P. E. Bourne, The Protein Data Bank, *Nucleic Acids Res.*, 2000, **28**, 235.

20 R. Rodriguez, G. Chinea, N. Lopez, T. Pons and G. Vriend, Homology modeling, model and software evaluation: three related resources, *Bioinformatics*, 1998, **14**, 523.

21 The PyMOL Molecular Graphics System, Version 1.5.0.3'.

22 A. W. Schuttelkopf and D. M. F. van Aalten, PRODRG: a tool for high-throughput crystallography of protein-ligand complexes, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2004, **60**, 1355.

23 H. J. C. Berendesen, J. P. M. Postma, W. F. v. Gunsteren and J. Hermans, Interaction Models for Water in Relation to Protein Hydration, in *Interaction Models for Water in Relation to Protein Hydration*, ed. B. Pullman, Dordrecht, 1981.

24 H. J. C. Berendsen, J. P. M. Postma, W. F. v. Gunsteren, A. DiNola and J. R. Haak, Molecular dynamics with coupling to an external bath, *J. Chem. Phys.*, 1984, **81**, 3684.

25 U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, A smooth particle mesh Ewald method, *J. Chem. Phys.*, 1995, **103**, 8577.