Indian Journal of Biochemistry & Biophysics
Vol. 52, February 2015, pp. 23-28

Analyzing resistance pattern of non-small cell lung cancer to crizotinib using molecular dynamic approaches

Anish Kumar and K Ramanathan*

Industrial Biotechnology Division, School of Bio Sciences and Technology, VIT University, Vellore-632014, Tamil Nadu, India

Received 12 May 2014 ; revised 21 November 2014

Crizotinib is the potential anticancer drug used for the treatment of non-small cell lung cancer (NSCLC) approved by FDA in 2011. The main target for the crizotinib is anaplastic lymphoma kinase (ALK). Evidences available indicate that double mutant ALK (L1196M and G1269A) confers resistance to crizotinib. However, how mutation confers drug resistance is not well-understood. Hence, in the present study, molecular dynamic (MD) simulation approach was employed to study the impact of crizotinib binding efficacy with ALK structures at a molecular level. Docking results indicated that ALK double mutant (L1196M and G1269A) significantly affected the binding affinity for crizotinib. Furthermore, MD studies revealed that mutant ALK-crizotinib complex showed higher deviation, higher fluctuation and decreased number of intermolecular H-bonds, when compared to the native ALK-crizotinib complex. These results may be immense importance for the molecular level understanding of the crizotinib resistance pattern and also for designing potential drug molecule for the treatment of lung cancer.

Keywords: Non-small cell lung cancer, Crizotinib resistance, Molecular dynamic simulation.

Lung cancer is the prominent cause of cancer deaths in the world[1]. It is divided into two main classes based upon their histology, which are non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC). Most importantly, NSCLC accounts for the 85% of death worldwide. The most common forms of NSCLC are adenocarcinoma (ADC) and squamous cell carcinoma (SCC)[2]. Chromosomal rearrangements in the anaplastic lymphoma kinase (ALK) gene that codes for anaplastic lymphoma kinase have been identified as one of the causes of NSCLC. There are two types of tyrosine kinase, receptor and cytoplasmic tyrosine kinase. The ALK is a cytoplasmic tyrosine kinase, where crizotinib (a potential anticancer drug used in the treatment of NSCLC) is bound[3].

Chromosomal rearrangements involving the ALK gene occur in different malignant conditions, including NSCLC, anaplastic large cell lymphoma (ALCL) and inflammatory myofibroblastic tumor (IMT)[4]. These rearrangements lead to the expression of ALK fusion genes. ALK fusion gene possesses different properties from the two genes that it was originally derived from, can then code for the new ALK fusion protein, which is abnormally and constitutively activated[5]. The new protein contains the tyrosine kinase domain of ALK and the coiled coil domain of EML4. The coiled coil domain of EML4 allows this protein to bind with other ALK fusion proteins and form dimerised and activated fusion proteins[5].

The most prevalent ALK fusion oncogene in NSCLC is the echinoderm microtubule-associated protein-like 4 (EML4)-ALK fusion gene and is present in 4-5% of cases of NSCLC and that too mostly in adenocarcinoma[6]. An inversion in the chromosome 2 brings together the 5' end of the EML4 gene and the 3' end of the ALK gene resulting in the formation of the EML4-ALK fusion gene[7]. The affected persons tend to have typical clinical features like early age of onset, little or absence of any smoking history. Targeted drug therapy is used against NSCLC of which tyrosine kinase inhibitors are amongst the best method in treatment methodology. In particular, crizotinib is one such tyrosine kinase inhibitor which is the first and only

*Corresponding author.
E-mail: kramanathan@vit.ac.in
Tel.: +91 4162202545.
Fax: +91 4162243092.

Abbreviations: ALK, anaplastic lymphoma kinase; ALCL, anaplastic large cell lymphoma; IMT, inflammatory myofibroblastic tumor; MD, molecular dynamics; NSCLC, non-small cell lung carcinoma; PDB, protein data bank; Rg, radius of gyration; RMSD, root mean square deviations; RMSF, root mean square fluctuations; SASA, solvent accessible surface area; SCLC, small cell lung cancer.

drug to have gained FDA approval for the treatment of NSCLC[8]. Although crizotinib has proved itself as an efficient counter to ALK type NSCLC, acquired resistance has made its beneficial effects temporary and has emerged as a major roadblock for crizotinib.

Since no computational study has been focused on this ALK mutation, the actual effect of the mutation on the drug-target binding process remains unexplained. The present study has been aimed to understand the mechanism of drug resistance conferred by ALK mutation using molecular dynamics approaches. The study might be useful to develop new therapeutic strategies to overcome crizotinib resistance in the lung cancer.

Methodology

The native and mutant (L1196M and G1269A) ALK 3D structures were retrieved from the crystal structures of the Brookhaven Protein Data Bank (PDB)[9] for the analysis. The corresponding PDB codes were 2XP2 and 4ANS for the native and mutant structures, respectively[10]. Crizotinib was used as the small molecule in our analysis. PubChem[11] was used to collect the SMILES strings and CORINA program was utilized to construct the 3D structure of the drug molecule[12]. The PubChem database holds over 27 million records of unique chemical structures of compounds (CID) derived from nearly 70 million substance depositions (SID). It contains more than 449,000 bioassay records with over thousands of *in vitro* biochemical and cell-based screening bioassays established, with targeting more than 7000 proteins and genes linking to over 1.8 million of substances. Hence, we used PubChem database for our computational analysis.

Initially, docking experiments were performed using the software Argus Lab 4.0.1. The flexible ligand docking was performed in our analysis[13]. AScore, an empirical scoring function was used for the docking calculations. The following contributions, such as van der Waals interaction between the ligand and protein, hydrophobic effect, hydrogen bonding between the ligand and the protein, deformation effect and effects of translational and rotational entropy loss were taken into account for AScore function[14]. AScore.prm file read from the AScore function was used to calculate of binding energies of the resulting docked structures.

Subsequently, GROMACS Package 4.5.3 implemented with Gromos 43a1 force field was utilized to perform molecular dynamics (MD) calculations of our docked complexes[15,16]. All the heteroatom and the water molecules were removed before performing MD simulations. The following commands were used to run MD simulation for the ALK-crizotinib complex. All commands are shown in italics in the following lines. Lines started with ">" are command lines used for MD. The parameter files such as em.mdp and md.mdp were saved in the directory along with ALK-crizotinib complex file. The total simulation time was set to 5000 ps with integration time step of 2 fs.

The first step was the generation of protein topology file used by the GROMACS. The topology file (.top extension) was generated by the GROMACS pdb2gmx command mentioned below.

> pdb2gmx -f PDBcode.pdb -o conf. pdb

The pdb2gmx command will also add hydrogen's to the protein structure.

Ligand topology (DRGGMX.ITP) and co-ordinate files (DRGGMX.PDB) are generated using PRODRG 2 server[17].

Secondly, the protein-drug complex is solvated in a cubic box of water in order to make the simulation a bit more realistic. The "edit conf" and "gen box" command mentioned below is used to define a box and to add water respectively around the protein-drug complex.

> editconf -f -bt cubic conf.pdb -o box.pdb –d 0.9
> genbox -cp box.pdb -cs spc216 -o water.pdb -p topol.top

This resulted in the addition of 22,269 and 23,506 water molecules to the native and mutant complex structures, respectively. Subsequently, energy minimization was performed to remove the close contacts. The following commands were employed.

>grompp -f em.mdp -c water.pdb -p topol.top -o em.tpr –maxwarn
> mdrun -v -s em.tpr -c em.pdb

In a way, energy minimization shakes the entire system until close contacts are reduced and the system tends towards a lowest energy state called stable condition. In many cases, the protein might carry the charge. So, ions will be added in order to neutralize the system by employing following commands.

>genion -s em.tpr -o ions.pdb –np
>grompp -f em.mdp -c water.pdb -p topol.top -o em.tpr –maxwarn

In the present simulation study, six sodium (6 Na+ ions) counter ions were added to neutralize the total charge of the system. The energy minimization initiated again with the help of neutral structure using following commands.

>grompp -f em.mdp -c ions.pdb -p topol.top -o em.tpr -maxwarn
>mdrun -v -s em.tpr -c em.pdb

Finally, MD simulation was initiated with the aid of following commands.

> grompp -f md.mdp -c em.pdb -p topol.top -o md.tpr -maxwarn
>mdrun -v -s md.tpr -c md.pdb –nice

The progress could be monitored by looking at the step number. The trajectory files were stored for structural analysis at every picosecond. The parameter, such as root mean square deviation (RMSD), root mean square fluctuation (RMSF), intramolecular hydrogen bonds (NHbond), radius of gyration (Rg) and solvent accessibility (SASA) were examined from the MD trajectories by the following program.

For RMSD (Backbone) analysis: *>g_rms –f traj –s md*
For RMSF (Cα residues) analysis: *>g_rmsf –f traj –s md –res*

For H-Bond analysis: *>g_hbond –f traj –s md*
For Rg analysis: *>g_gyrate –f traj –s md*
For SASA analysis: *>g-sas –f traj –s md*

Results and Discussion

Argus Lab 4.0.1 (http://www.ArgusLab.com) was used to determine binding free energies of the native and mutant types of ALK-crizotinib complex. For the native ALK-crizotinib complex, the binding energy was found to be -7.0 kcal/mol, whereas for the mutant ALK-crizotinib complex, the binding energy was found to be -2.5 kcal/mol. There was a significant difference in binding energy of the mutant complex with respect to native complex. This result undoubtedly indicated that the crizotinib was less effective against mutant type (L1196M and G1269A) ALK. The docked complex structures were visualized with the help of PyMOL tool (Fig. 1)[18].

Subsequently, MD simulations were performed for the native and mutant (L1196M and G1269A) ALK-crizotinib complex using GROMACS package 4.5.3. The ligand binding process is a microscopic event that takes place in mere millionths of a second.

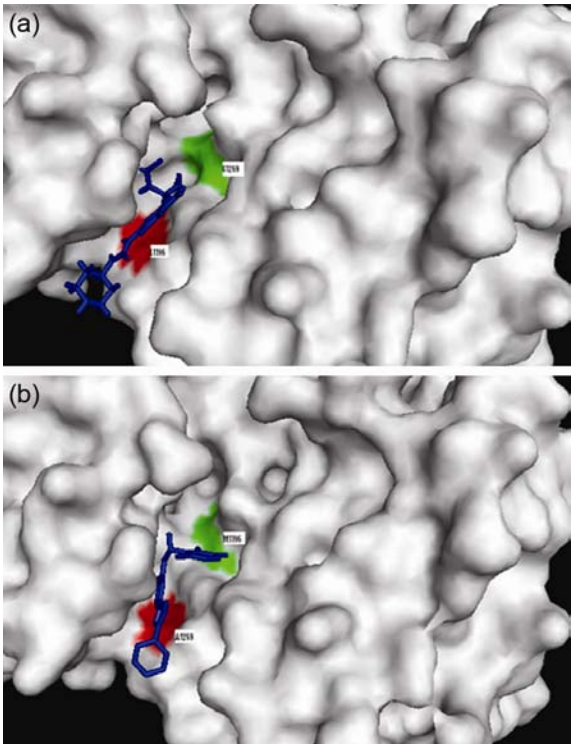

Fig. 1—Docked surface view of crizotinib with native (a) and mutant (b) type ALK using PyMOL

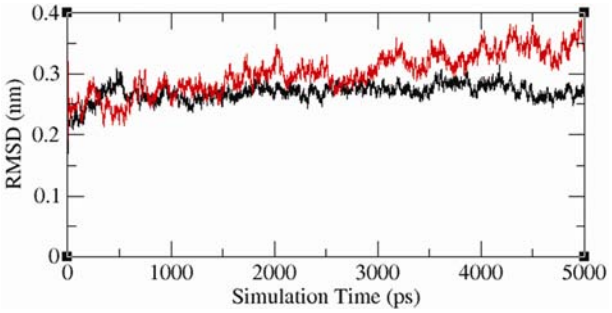

Fig. 2—Root mean square deviations (RMSD) correspond to the native (black) and mutant (red) type ALK-crizotinib complex along the MD simulation at 300 K

The complete understanding of the atomistic energetic and mechanics of binding is not feasible using current experimental techniques. However, MD study is certainly helpful to address the above issues[19].

The parameters, such as RMSD, RMSF, hydrogen bond, radius of gyration and solvent accessibility details were analyzed from the MD trajectory file. RMSD was the measure of the deviation of the mutant structure from their native structure. RMSD analysis could give an idea of how much the three-dimensional structure had deviated over time (Fig. 2). After the relaxation period, native structure acquired ~0.25 nm at 1000 ps during the simulations, while mutant structure acquired ~0.30 nm

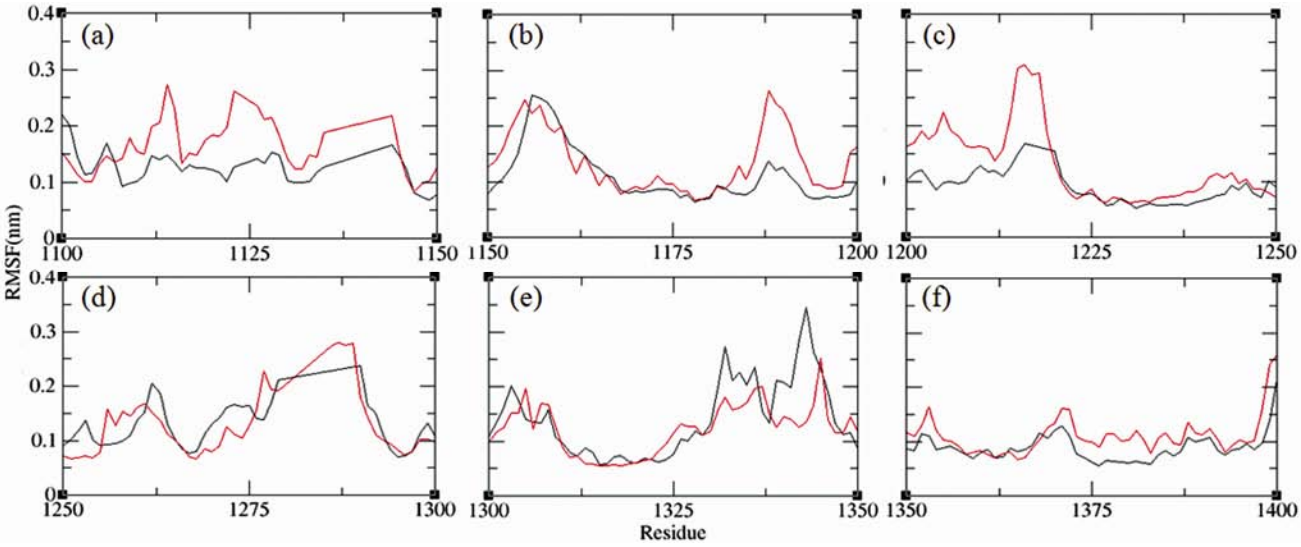

Fig. 3—Root mean square fluctuations (RMSF) of the backbone Cα residues in different regions of our protein sequence [(a) 1100-1150 residues, (b) 1150-1200 residues, (c) 1200-1250 residues, (d) 1250-1300 residues, (e) 1300-1350 residues, and (f) 1350-1400 residues. Colour coding scheme: native, black and mutant, red]

of backbone RMSD at 1000 ps. Between a period of 2000 and 4000 ps, a mutant structure was able to maintain ~0.35 nm of backbone RMSD, while the native structure showed frequent changes (~0.25 to ~0.27 nm) in backbone RMSD. Mutant gained RMSD dominance over native and attained RMSD of ~0.38 nm at 5000 ps. On the other hand, the native showed less deviation than the mutant and achieved ~0.26 nm at 5000 ps. The higher RMSD value of mutant-crizotinib complex clearly indicated that function of the mutant structure was affected. This data suggested that the mutant structure had undergone conformational changes throughout the simulation, whereas such changes were not observed in the native structure.

The flexibility of amino acids is also considered to be a relevant parameter to understand the binding efficiency. The loss of flexibility impairs the binding effect and *vice versa*[20]. Therefore, we examined the flexibility behavior of the amino acid residues (Fig. 3) using RMSF analysis. Although the variation in flexibility of residues was less between the native and mutant structures, there was a significant variation in the flexibility observed in the region 1100–1400. The binding site from 1100 to 1400 belonged to the protein kinase domain in ALK tyrosine kinase receptor. The domain region from 1100-1150 revealed that mutant had fluctuated more than the native structure. The region 1124 was an ATP-binding domain, where mutant structure had a RMSF value of ~0.27 nm, whereas native structure had a RMSF value of ~0.15 nm. The region 1203 acted an inhibitor,

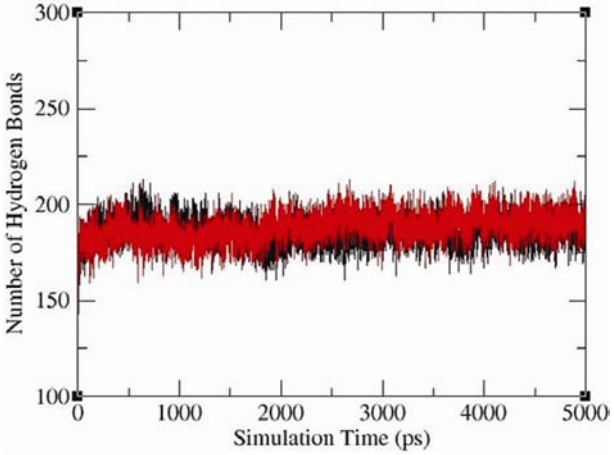

Fig. 4—Intramolecular H bonds observed during the MD simulation [The symbol coding scheme is as follows: native ALK (black) and mutant (L1196M and G1269A) ALK (red)]

where mutant and native structure had a RMSF value of ~0.18 nm and ~0.10 nm, respectively. In the domain region 1120-1380, a mutant structure was able to maintain ~0.26 nm. On the other hand, native structure showed frequent changes (~0.13 to ~0.15 nm) in RMSF value. The RMSF values of ~0.26 nm and ~0.20 nm were observed for mutant and native type ALK, respectively at the end. Overall, the mutant structure fluctuated more than the native structure.

The deviation in flexibility was further validated by the number of intramolecular, as well as intermolecular hydrogen bond (NH bond) analysis (Figs 4 and 5). During the 500 ps, the native structure showed increase in intramolecular H bonds, compared

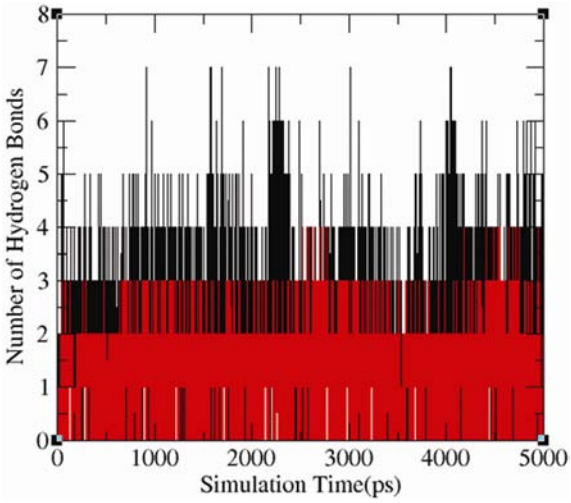

Fig. 5—Intermolecular H bonds observed during the MD simulation [The symbol coding scheme is as follows: native ALK-crizotinib (black) and mutant (L1196M and G1269A)-crizotinib (red) complexes]

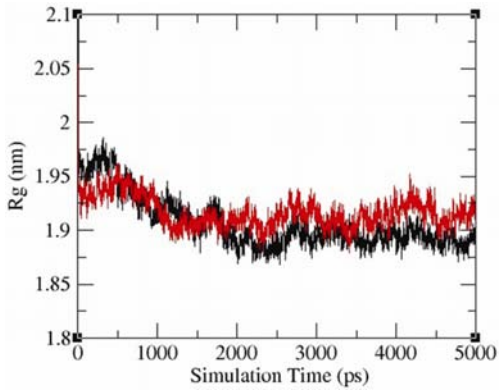

Fig. 6—Radius of gyration correspond to native (black) and mutant (red) types ALK along the MD simulation at 300 K

to the mutant structure. It can be seen in Fig. 4, the mutant structure generated 190 H bonds between periods of 2000 to 4000 ps, while native structure generated 160-170 H bonds between periods of 2000 to 4000 ps. At the end of 5000 ps, the mutant structure generated 200 H bonds, whereas native structure generated 180 H bonds. The frequency of occurrence of intramolecular H bond was slightly higher in the mutant structure than native structure (Fig. 4). This higher number of hydrogen bonds in the mutant structure was the clear evidence of the rigidity of the mutant (L1196M and G1269A) type ALK than native type ALK. Furthermore, intermolecular NH bond analysis (Fig. 5) suggested that the native ALK-crizotinib complex had seven intermolecular hydrogen bonds, whereas three intermolecular

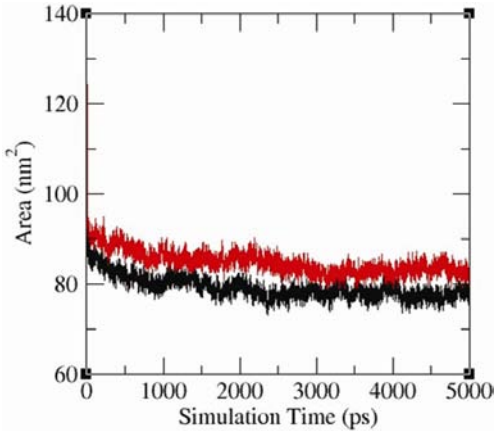

Fig. 7—Solvent-accessible surface area (SASA) corresponds to native (black) and mutant (red) types ALK along the MD simulation at 300 K

hydrogen bonds were observed in the mutant ALK-crizotinib complex throughout the simulation time. These variations within the H bond networks might be mainly responsible for the decrease in binding energy of mutant (L1196M and G1269A) ALK–crizotinib complex.

We also analyzed the radius of gyration (Rg) and solvent accessibility surface area (SASA) for the native and mutant structures of ALK. The Rg is defined as the mass-weight root mean square distance of collection of atoms from their common centre of mass and is a parameter linked to tertiary structural volume of a protein[21]. Thus, Rg helps in understanding the overall dimension of protein and notion of rapid conformational changes in the mutant structure can be observed with Rg. The Rg plot result is shown in Fig. 6. After 2000 ps, there was an abrupt rise in Rg values in mutant structure. At the end of 5000 ps, the mutant structure showed Rg value of 1.94 nm, whereas in native it was found to be 1.90 nm. The changes in Rg of mutant showed that the protein ALK might be undergoing a significant structural transition.

SASA is responsible for bimolecular surface area accessible to solvent molecules. Figure 7 indicated that SASA for mutant structure was higher than the native structure. The rise in SASA values for mutant structure denotes the relative expansion in mutant structure, when compared with native[21]. It was evident that SASA of the mutant structure gradually increased along the trajectory as compared to native structure, indicating that after mutation the overall structure became more unfolded during the simulation. This resulted in the loss of its structural

and functional activity and led to improper binding with crizotinib. These computational evidences indicated that ALK double mutation (L1196M and G1269A) had a significant impact on the ALK structure and its function.

Conclusion

In the present investigation, molecular docking and molecular dynamics approaches were used to gain an insight into crizotinib resistance. Docking analysis clearly indicated that ALK mutation altered the affinity of crizotinib. RMSD data obtained from molecular dynamic simulation revealed structural stability of the native ALK-crizotinib complex. Furthermore, RMSF and the number of hydrogen bonds analysis during the MD simulation conferred stable binding of crizotinib with native type ALK than mutant type ALK. The results obtained from this study might be helpful in the management of crizotinib resistance against ALK mutation.

Acknowledgements

The authors thank the management of Vellore Institute of Technology for all the support, assistance and constant encouragement to carry out this work.

References

1 Siegel R, Naishadham D & Jemal A (2012) *Cancer J Clin* 62, 10-29
2 Skarda J, Marian H & Vítezslav K (2008) *Cancer Ther* 6, 377-388
3 Hallberg B & Palmer R H (2013) *Nature Rev Cancer* 13, 685–700
4 Chiarle R, Voena C, Ambrogio C, Piva R & Inghirami G (2008) *Nat Rev Cancer* 8, 11–23
5 Katayama R, Alice T S, Tahsin M K, Mari M K, Benjamin JS, Balazs H, Nicholas A J, John C W, Alan T Y, Cyril B, Lisa D, Jamal Carlos S, Katherine C, Lecia VS, John I, & Jeffrey A E (2012) *Sci Transl Med* 4(120), ra17
6 Young LC, Manabu S, Yoshihiro Y, Junpei T, Takahiro N, YasushiY, Kengo T, Toru H, Yuichi I, Hidenori H & Hideki K (2010) *N Engl J Med* 363, 1734-1739
7 Shaw AT & Solomon B (2011) *Clin Cancer Res* 17, 2081-2086
8 Ou S-H I, (2011) *Drug Des Devel Ther* 5, 471–485
9 Berman H M, Westbrook J, Feng Z, Gilliland G, Bhat T N, Weissig H, Shindyalov I N & Bourne P E (2000) *Nucleic Acids Res* 28, 235–242
10 Cui J J, Tran-Dube M, Shen H, Nambu M, Kung PP, Pairish M, Jia L, Meng J, Funk L, Botrous I, McTigue M, Grodsky N, Ryan K, Padrique E, Alton G, Timofeevski S, Yamazaki S, Li Q, Zou H, Christensen J, Mroczkowski B, Bender S, Kania R S & Edwards M P (2011) *J Med Chem* 54, 6342-63
11 Feldman J, Snyder K A, Ticoll A, Pintilie G & Hogue C W (2006) *FEBS Lett* 580, 1649-1653
12 Gasteiger J, Rudolph C & Sadowski J (1990) *Tetrahedron Comput Method* 3, 537–547
13 Thompson MA (2004) Poster presentation: molecular docking using arguslab: an efficient shape-based search algorithm and the Ascore scoring function. Fall ACS Meeting, Philadelphia Oda A, Okayasu M, Kamiyama Y, Yoshida T, Takahashi O & Matsuzaki H (2007) *Bull Chem Soc Jpn* 80, 1920–1925
14 Hess B, Kutzner C, Spoel D & Lindahl E (2008) *J Chem Theory Comput* 4, 435–447
15 Spoel D, Lindahl E, Hess B, Groenhof G, Mark A E & Berendsen H J (2005) *Comput Chem* 26, 1701–1718
16 Schuttelkopf AW & Van Aalten D M F (2004) *Acta Crystallogr* 60, 1355-1363
17 Warren L D (2002) *The PyMOL Molecular Graphics System*, DeLano ScientificLLC, San Carlos, CA, USA
18 Durrant J D & McCammon J A (2011) *BMC Biol* 9, 71–79
19 Hinkle A & Tobacman L S (2003) *J Biol Chem* 278, 506–513
20 Kumar A & Purohit R (2013) *PLoS One* 8(5), e64364