Molecular insights of benzodipyrazole as CDK2 inhibitors: combined molecular docking, molecular dynamics, and 3D QSAR studies

Varun Guttikonda¹, Divya Raavi², Sasi Kala Maadwar³, and Deepak Reddy Gade⁴

¹Department of Biochemistry and Chemistry, Lamar University, Beaumont, TX, USA, ²Department of Bioinformatics, Gulbarga University, Gulbarga, India, ³Department of Pharmaceutical Chemistry, College of Pharmacy, Holy Mary Institute of Technology & Science, Hyderabad, India, and ⁴Department of Pharmaceutical Chemistry, Jawaharlal Nehru Technological University-OTRI, Anantapur, AP, India

Abstract

Benzodipyrazoles have been previously evaluated for their in vitro CDK2 inhibitory activity. In the current investigation, we identified a six-feature common pharmacophore model (AADDRR.33) which is predicted to be responsible for CDK2 inhibition. An efficient 3D QSAR (r² = 0.98 and q² = 0.82) model was also constructed by employing PLS regression analysis. From the molecular docking studies, we examined the binding patterns of compound 7aa with the target protein and also calculated the binding energy using MM-GBSA calculations. Three hydrogen bonds with Lys 33, Glu 81, and Leu 83 are conserved even after 1000 ps run in a molecular dynamics simulation. We identified the slight displacement in bond lengths and the conformational changes occurred during the dynamics. The results also elucidated the protein residue–ligand interaction fractions which clearly explained the involvement of non-H-bond interactions.

Keywords

CDKs, benzodipyrazoles, MM-GBSA, molecular dynamics, QSAR

Introduction

Cyclin-dependent kinases (CDKs), of family serine/threonine kinases, regulate the complex processes of the cell division cycle (1,2). Specific CDK–cyclin complexes are responsible for various events known to take place during cell cycle in a sequential and orderly fashion. CDK2 is required to complete G1 and to trigger the S-phase, and CDK4 is required to integrate extracellular signals and directs the cell cycle (3,4). Further explicating, CDKs are inactive as monomers and activation requires binding to the corresponding regulatory proteins and phosphorylation by CDK-activating kinases (CAK) on a specific threonine residue (5). A great deal of effort is now revolved around the CDK2 regulation due to its physiological role in cell division and gene transcription, which rationally precedes cancer development and has benefits as an attractive target for discovery of new antitumor agents (6–8).

From the recent discoveries, benzodipyrazole moieties are affirmed as a potent CDK2 inhibitors based on the pre-evaluated in vitro analysis (9). Aiding the conventional drug discovery, we computationally simulated and profoundly evaluated these compounds for their structural conformations with relevance to the CDK protein by assessing the pharmacophore model and QSAR studies. Molecular docking and dynamics simulations helped in understanding the binding mechanism of a CDK2 inhibitor, which are also supported by MM-GBSA calculations and analysis of several trajectories with respective to protein ligand interactions.

Methodology

Data set and ligand optimization

Precise data set (Table 1) of 60 compounds with benzodipyrazole scaffold was primitive selected (9). Ligands from the data set were structurally produced, using theChem sketcher tool of Schrödinger suite (Schrödinger, New York, NY). These were later refined, initially by optimizing geometrically and energy minimization of the structures 2D to 3D was performed by using the Ligprep tool of Schrodinger suite. This includes generation of possible ionization states at target biological pH through EPiK program, and also conformers and tautomers by applying optimized potentials for liquid simulations (OPLS – 2005) force field. Later these geometrically optimized structures were conditionally chosen to generate pharmacophore hypothesis and built the 3D QSAR model (10). Compound selection and experimental activity (IC₅₀) against CDK2 were all originated from the same laboratory. All the geometrically optimized and energy minimized ligands were arranged as a 3D library for further usage.
Table 1. Structures and 3D QSAR activity data of the selected dataset ligands.

| Compound no. | R1     | R2     | R3     | pIC50  | pIC50 predicted | Pharm set | Data set |
|--------------|--------|--------|--------|--------|----------------|-----------|----------|
| 7a           | –ph-4-So2NH2 | –CONH2 | –CH2–CH2– | 8.398  | 8.65           | Active    | Training |
| 7b           | –ph-4-So2 NMe | –CONH2 | –CH2–CH2– | 7.399  | 7.73           | Inactive  | Test     |
| 7c           | –ph-4-So2 NMe2 | –CONH2 | –CH2–CH2– | 8.000  | 7.97           | Active    | Training |
| 7d           | –ph-4-So2 NHBu | –CONH2 | –CH2–CH2– | 8.699  | 8.76           | Active    | Training |
| 7e           | –ph-4-SO2 NH(CH2–CH2)2N–Me | –CONH2 | –CH2–CH2– | 5.000  | 5.26           | Inactive  | Training |
| 7f           | –ph-4-SO2 Me | –CONH2 | –CH2–CH2– | 7.523  | 7.56           | Active    | Training |
| 7g           | –H      | –CONH2 | –CH2–CH2– | 6.338  | 6.51           | Training  |          |
| 7h           | –ph     | –CONH2 | –CH2–CH2– | 6.824  | 6.49           | Active    | Training |
| 7i           | –ph-4-Me | –CONH2 | –CH2–CH2– | 6.854  | 6.45           | Test      |          |
| 7j           | –ph-4-Ome | –CONH2 | –CH2–CH2– | 6.854  | 6.69           | Active    | Training |
| 7k           | –ph-4-Cl | –CONH2 | –CH2–CH2– | 6.796  | 6.48           | Training  |          |
| 7l           | –ph-4-F  | –CONH2 | –CH2–CH2– | 6.569  | 6.51           | Active    | Test     |
| 7m           | –ph-4-CF3 | –CONH2 | –CH2–CH2– | 7.097  | 7.22           | Active    | Training |
| 7n           | –ph-4-OCF3 | –CONH2 | –CH2–CH2– | 6.293  | 6.84           | Test      |          |
| 7o           | –ph-4-CN  | –CONH2 | –CH2–CH2– | 6.722  | 6.51           | Test      |          |
| 7p           | –ph-(CH2–CH2)O | –CONH2 | –CH2–CH2– | 6.252  | 6.12           | Training  |          |
| 7q           | –ph-(2-imidazolo) | –CONH2 | –CH2–CH2– | 6.443  | 6.39           | Training  |          |
| 7r           | –ph-3-Cl | –CONH2 | –CH2–CH2– | 6.377  | 6.41           | Training  |          |
| 7u           | –2-pyridyl | –CONH2 | –CH2–CH2– | 6.585  | 6.36           | Training  |          |
| 7v           | –3-pyridyl | –CONH2 | –CH2–CH2– | 5.796  | 5.90           | Inactive  | Training |
| 7w           | –Bn     | –CONH2 | –CH2–CH2– | 5.656  | 6.22           | Active    | Training |
| 7x           | –Me     | –CONH2 | –CH2–CH2– | 6.538  | 6.53           | Training  |          |
| 7y           | –CH3    | –CONH2 | –CH2–CH2– | 7.699  | 7.43           | Active    | Training |
| 7z           | –CH2–CH2–OH | –CONH2 | –CH2–CH2– | 5.940  | 6.43           | Inactive  | Training |
| 7aa          | –ph-4-So2 NH2 | –CONH2 | –CH = CH– | 9.523  | 9.39           | Active    | Training |
| 7bb          | –ph-4-So2 NMe | –CONH2 | –CH = CH– | 6.699  | 7.89           | Inactive  | Test     |
| 7cc          | –ph-4-So2 NMe2 | –CONH2 | –CH = CH– | 6.268  | 6.64           | Training  |          |
| 7dd          | –ph-4-So2 NHBu | –CONH2 | –CH = CH– | 6.678  | 6.81           | Training  |          |
| 7ee          | –ph-4-So2 NH(CH2–CH2)2N–Me | –CONH2 | –CH = CH– | 6.824  | 6.60           | Training  |          |
| 7ff          | –ph-4-So2 Me | –CONH2 | –CH = CH– | 9.000  | 8.99           | Active    | Training |
| 7gg          | –H      | –CONH2 | –CH = CH– | 7.108  | 7.06           | Active    | Training |
| 7hh          | –ph     | –CONH2 | –CH = CH– | 7.854  | 7.89           | Active    | Test     |
| 7jj          | –ph-4-Ome | –CONH2 | –CH = CH– | 8.046  | 8.59           | Active    | Training |
| 7kk          | –ph-4-Cl | –CONH2 | –CH = CH– | 8.979  | 7.89           | Active    | Training |
| 7ll          | –ph-4-F  | –CONH2 | –CH = CH– | 7.444  | 7.77           | Active    | Training |
| 7mm          | –ph-4-CF3 | –CONH2 | –CH = CH– | 7.887  | 7.81           | Active    | Training |
| 7nn          | –ph-4-OCF3 | –CONH2 | –CH = CH– | 7.432  | 7.31           | Active    | Training |
| 7oo          | –ph-4-CN  | –CONH2 | –CH = CH– | 8.302  | 8.02           | Active    | Training |
| 7pp          | –ph-(CH2–CH2)O | –CONH2 | –CH = CH– | 7.746  | 7.83           | Active    | Training |
| 7qq          | –ph-(2-imidazolo) | –CONH2 | –CH = CH– | 7.125  | 6.95           | Active    | Training |
| 7rr          | –ph-3-Me | –CONH2 | –CH = CH– | 8.046  | 7.99           | Active    | Training |
| 7ss          | –ph-3-Cl | –CONH2 | –CH = CH– | 7.537  | 7.82           | Active    | Test     |
| 7tt          | –ph-3-F  | –CONH2 | –CH = CH– | 7.553  | 7.75           | Active    | Training |
| 7uv          | –3-pyridyl | –CONH2 | –CH = CH– | 7.208  | 7.16           | Active    | Training |
| 7ww          | –Bn     | –CONH2 | –CH = CH– | 6.854  | 6.88           | Training  |          |
| 7xx          | –Me     | –CONH2 | –CH = CH– | 7.745  | 7.75           | Active    | Training |
| 7yy          | –CH3    | –CONH2 | –CH = CH– | 8.699  | 8.20           | Active    | Test     |
| 7zz          | –CH2–CH2–OH | –CONH2 | –CH = CH– | 7.553  | 8.01           | Active    | Test     |
| 6j           | –ph-4-Ome | –COOEt | –CH2–CH2– | 5.932  | 5.39           | Inactive  | Training |
| 8j           | –ph-4-Ome | –COOH | –CH2–CH2– | 5.678  | 5.62           | Inactive  | Training |
| 10j          | –ph-4-Ome | –CONHNH2 | –CH2–CH2– | 5.678  | 5.77           | Inactive  | Training |
| 11j          | –ph-4-Ome | –CONHOCH2CH = CH2 | –CH2–CH2– | 6.161  | 6.18           | Training  |          |
| 12j          | –ph-4-Ome | –CONHOCH2CH = CH2 | –CH2–CH2– | 5.347  | 5.45           | Inactive  | Training |
| 16j          | –ph-4-Ome | –NH3 | –CH2–CH2– | 6.046  | 6.03           | Test      |          |
| 6d           | –ph-4-So2NHBu | –COOEt | –CH2–CH2– | 5.000  | 4.87           | Inactive  | Training |
| 12d          | –ph-4-So2NHBu | –CONHOCH2CH = CH2 | –CH2–CH2– | 5.545  | 5.45           | Inactive  | Training |
Generation of common pharmacophore hypotheses

By the allocation of activity threshold range, the whole data set was fractioned into active (pIC$_{50} > 7$), inactive (pIC$_{50} < 6$), and intermediate (pIC$_{50} = 6–7$). These ligands were internally validated through PHASE (a fine-grained conformational sampling and a range of scoring technique tool) to identify a common pharmacophore hypothesis, accompanied by a set of aligned conformations with the relativity of bind affinity (11). A 3D pharmacophore model was developed using a set of pharmacophore features to generate sites for all the compounds, concurring with various chemical features that may make easy non-covalent binding among the ligand and its binding pocket. Distinctive six pharmacophore features are hydrogen bond acceptor (A), hydrogen bond donor (D), hydrophobic group (H), negatively ionizable (N), positively ionizable (P), and aromatic ring (R). A maximum of six and a minimum of five sites were selected in order to obtain an efficient pharmacophore model. Hypotheses were generated by a systematic variation of the number of sites ($n_{sites}$) and the number of matching active compounds ($n_{act}$). With $n_{act} = n_{act–tot}$. Initially ($n_{act–tot}$) is the total number of active compounds in the training set, $n_{sites}$ (12). The scoring protocol provides a ranking of different hypotheses to choose the most appropriate for further investigation.

3D QSAR studies

For QSAR development, from the outcome of the scoring protocol, the best hypothesis was chosen (AADDRR.33). Pharmacophore-based QSAR models were generated for hypothesis using 48 training set ligands (80% of dataset were selected randomly) and a grid spacing of 1.0 Å. By implementing regression analysis corresponding with PLS factors, QSAR models containing one to nine PLS factors were generated. These models were validated by predicting the activity of test set ligands. The accuracy of the models increases with increasing number of PLS factors until over-fitting starts to occur (13). The predictive value of the models was evaluated by leave one-out (LOO) and leave-half-out (LHO) cross-validation. The cross-validated coefficient, $R_{cv}^2$, was calculated using the following equation:

$$ r_{cv}^2 = 1 - \frac{\sum (Y_{predicted} - Y_{observed})^2}{\sum (Y_{observed} - Y_{mean})^2} $$

where $Y_{predicted}$, $Y_{observed}$, and $Y_{mean}$ are the predicted, observed, and the mean values of the target property (pIC$_{50}$), respectively. $(Y_{observed} - Y_{mean})^2$ is the predictive residual sum of squares (PRESS). The predictive correlation coefficient ($r_{pred}^2$), based on the molecules of the test set, is defined as

$$ r_{pred}^2 = \frac{SD-PRESS}{SD} $$

where SD is the sum of the squared deviation between the biological activities of the test set and mean activities of the training set molecules, PRESS is the sum of squared deviation between predicted and actual activity values for every molecule in the test set. According to the literature (14), 3D-QSAR models were accepted if

$$ R^2 > 0.6; R_{cv}^2(Q^2) > 0.5 $$

Molecular docking and binding energy calculations

In silico binding interactions of CDK2 and benzodipyrazole derivatives were performed by using the Glide XP algorithm of Schrodinger Suite. Prior to the docking studies, CDK2 protein was retrieved from the PDB website (PDB ID: 2BKZ) and was processed by removing non-protein components and adding missing amino acids and hydrogens to satisfy the valence and be optimized by using the OPLS-2005 force field. A grid was generated by the glide module at the auto-phosphorylation site of kinase domain of CDK2 and compound 7aa was docked into the prepared grid by employing flexible protein and ligand parameters (15). MM/GBSA calculations were performed for docked protein ligand complex to determine the binding energy (16). The total free energy of binding is then expressed in the form mentioned below in the following equation:

$$ \Delta G_{bind} = G_{complex} - (G_{protein} + G_{ligand}) $$

where $\Delta G_{bind}$ is ligand binding energy

In other terms

$$ \Delta G_{bind} = \Delta E_{gas} + \Delta G_{solvation} - T\Delta S $$

where gas phase molecular mechanical energy is $\Delta E_{gas}$ (i.e. Contributions from the van der Waals energy, electrostatic energy, and internal energy), the solvation free energy are $\Delta G_{solvation}$ (polar and non-polar contributions) and $T\Delta S$ is the entropy (17).

Molecular dynamics

Docking complex of the most active compound of the data set molecules (compound 7aa) was selected for molecular dynamic studies using Desmond by D.E. Shaw research (18). The docked complex is also used for trajectory analysis of 1000 ps (equals to 1 ns) molecular dynamics simulations. A solvent system was built around the docked complex with the TIP4PEW water solvent model and neutralizing Sodium ions were added. The system was maintained at 0.15 M salt concentration (Na$^+$ Cl$^-$) which is similar to the physiological conditions. Later the soaked simulated system is introduced into the molecular dynamics workspace. Dynamics simulations were performed at a constant number of atoms, at constant pressure (1.01352 bar) and constant temperature (300 K), i.e., NPT ensemble for 1000 ps. For ensembling, we used the Nose–Hoover chain thermostat with a relaxation time of 1 ps and Martyna–Tobias–Klein barostat with 2 ps relaxation time and isocratic cooling style and Coulombic interactions were cutoff at 9 Å radius (19).

Results and discussion

Pharmacophore model development

A total of 24 common pharmacophore hypotheses with six pharmacophoric features were obtained from the three variant
hypotheses: AADDRR, AAADDR, and ADDRRR. Survival scores of best selected common pharmacophore hypotheses are tabulated in Table 2. Survival scores and activities of AADDRR.33 and AAADDR.46 are almost similar and it might be due to the presence of an H-bond acceptor and loss of an aromatic ring in AAADDR.46. It is very clear that the selected scaffold benzodipyrazole consists of three aromatic rings and it is the reason for selection of AADDRR.33 as best fit CPH and its 3D spatial arrangement with inter-pharmacophoric feature distances are shown in Figure 1. Overall multi-ligand alignment score for a given reference pharmacophore (AADDRR.33) is the average score from the best individual alignments. Alignment of 3D structures over the pharmacophore AADDRR.33 is shown in Figure 2.

**QSAR statistics and model validation**

We built a pharmacophore-based QSAR model by segregating dataset into different training sets and test sets, with varying the grid space, and visualizing the model results. Contour mapping of the reference ligand, interpreted as the dipyrazole units attached to the benzene ring, the carbonyl linkage and alkyl groups attached to the pyrazole nucleus, and the phenyl moiety attached to the nitrogen of pyrazole, are all shown to be favorable conditions for the increase in activity.

From regression analysis of the dataset ligands over AADDRR.33 hypothesis (Table 3) of the total nine factors, PLS-7 was considered to be the best as the regression coefficients $r^2$ was 0.98 (for training set) and $q^2$ was 0.82 (for test set). PLS 8 and 9 have shown to be better regression coefficients but they are nearer to 1 and this condition is called over fitting which should not be considered. Predicted activity of all the dataset ligands obtained from QSAR studies considering PLS 7 is listed in Table 1. Data were plotted by

| ID     | Survival scores | Survival index | Site score | Volume | No. of Matches | Activity |
|--------|-----------------|----------------|------------|--------|----------------|----------|
| AADDRR.33 | 3.884           | 1.667          | 0.99       | 0.896  | 20             | 8.097    |
| AADDRR.46 | 3.821           | 1.757          | 0.94       | 0.885  | 21             | 8.097    |
| ADDRRR.118 | 3.813           | 1.856          | 0.94       | 0.876  | 21             | 7.985    |
| ADDRRR.136 | 2.452           | 1.225          | 0.05       | 0.488  | 22             | 7.746    |
extrapolating and correlating results with the actual (experimental) and predicted activities obtained from the QSAR based on the PLS regression analysis. The efficiency of the system was based on the slope of the line and the alignment of the ligands around the line (the slope of the line passes nearer to the origin of the plot, if the regression coefficient is nearer to one). The plots of actual activity versus activity predicted using PHASE of all dataset ligands and training set ligands are shown in Figure 3.

Molecular docking and binding energy calculations

Least energy conformation state of compound 7aa was considered for elucidation of binding phenomenon with the kinase domain of the CDK2 receptor (PDB ID: 2BKZ). Molecular docking studies confirm that the compound 7aa form three hydrogen bonds each with Lys 33 (2.06 Å), Glu 81 (1.86 Å), and Leu 83 (2.03 Å) which is shown in Figure 4, and compound 7aa has a dock score of −9.988 (XP Gscore). One of the two pyrazole rings is contributing with two H-bonds with Glu 81 (hydrogen attached to nitrogen is acting as an H-bond donor) and Leu 83 (nitrogen is acting as an H-bond acceptor). The third H-bond is present between ester C=O and Lys 33 where oxygen of ester is acting as an H-bond acceptor. Dock scores of all benzodipyrazole derivatives were tabulated in Supplementary Table 1.

To elicit the energy system of docked complex of 2BKZ and compound 7aa, we calculated its binding energy ($\Delta G_{\text{bind}}$) using the MM-GBSA method which is simply a molecular mechanics which is based on the Generalized Born model that is improved with the hydrophobic solvent accessible surface area. GB/SA is one of the most commonly used implicit solvent models. From the calculation, we found that compound 7aa is having $\Delta G_{\text{bind}}$ of −67.09 kcal/Mol. A detailed analysis shows that there is a remarkable contribution of cumbic (−19.12 kcal/Mol) and van der Waal’s (−37.66 kcal/Mol) terms, whereas the contribution of covalent term is negligible (−0.25 kcal/Mol).

### Table 3. Regression scores of AADDRR.33 pharmacophore hypothesis.

| ID   | PLS Factor | $R^2$-squared | $F$     | $P$      | $Q^2$-squared | Pearson-R |
|------|------------|---------------|---------|----------|---------------|-----------|
| AADDRR.33 | 1 | 0.48  | 42.5   | 4.775e−008 | 0.01  | 0.29       |
|       | 2 | 0.70  | 53.6   | 1.23e−012 | 0.49  | 0.71       |
|       | 3 | 0.84  | 74.6   | 2.724e−017 | 0.64  | 0.81       |
|       | 4 | 0.90  | 92.5   | 1533e−020 | 0.69  | 0.84       |
|       | 5 | 0.95  | 152.1  | 9.039e−026 | 0.75  | 0.88       |
|       | 6 | 0.96  | 184.3  | 4.945e−028 | 0.79  | 0.9        |
|       | 7 | 0.98  | 311.5  | 8.193e−033 | 0.82  | 0.91       |
|       | 8 | 0.99  | 414.2  | 3.046e−035 | 0.84  | 0.93       |
|       | 9 | 0.99  | 528.3  | 4.481e−037 | 0.84  | 0.94       |
Figure 3. QSAR plots of data set ligands and training set ligands (inset).

\[ y = 0.94x + 0.42 \quad (R^2 = 0.96) \]

Figure 4. (a) Protein ligand interactions between CDK2 (2BKZ) and compound 7aa elucidates the Hydrogen bonds with distances. (b) 2D interaction diagram of docked complex for conformation identification.

\[ y = 0.98 + 0.13 \quad (R^2 = 0.98) \]
Molecular dynamics simulations were performed to identify the conformational changes of ligand and receptor pocket when exposed to a solvent model (water) and at a constant temperature of 300 K (normal body temperature). We introduced previously docked complex of 2BKZ and compound 7aa into a solvent model and the energy of the whole system was minimized in order to attain stable conformation in insolvent environment. With the energy minimization, compound 7aa changed its conformation and rearranged its H-bonds with receptor pocket and surrounding solvent system (Figure 5). Compound 7aa formed a new H-bond with one of the surrounding water molecules and also the remaining three previously formed H-bonds were rearranged, such as bond distance between ligand and Lys 33 was reduced to 1.544 Å, bond distance between ligand and Glu 81 was reduced to 1.628 Å, and dramatically a slight increase was observed in the bond distance with Leu 83 (2.237 Å). The water molecule that bound to ligand also formed an H-bond with Asp 86 amino acid with larger bond length. From the 2D illustrations of the protein ligand complexes (Figure 5), it is clear that there is a conformational change during energy minimization with solvent model which is absent during docking studies. There is a shift of amino acids or compound structure towards amino acids like Ile 10, Val 18, and Phe 80 after energy minimization.

After 1000 ps run, H-bonds with previous three amino acids remain intact with slight difference in their bond distance, i.e., Lys 33 (1.794 Å), Glu 81 (1.477 Å), and Leu 83 (2.047 Å) which is shown in Figure 6. There was a complete displacement of H-bond with water molecules. Before molecular dynamics Asp 86 was bound to water, the later Asp 86 was replaced with Ile 10. We observed that the total flip of the conformation of ligand after 1000 ps run Ile 10 amino acid following solvation energy minimization was near to the terminal ethoxy group, whereas after molecular dynamics, Ile 10 reached nearer to the ester group and formed an H-bond with water molecules.

During in depth analysis of the bonding pattern between residues and ligand, we observed a significant role of
hydrophobic interactions, and water bridges along with H-bonds. Gila 81 is the only residue showing 100% H-bond interaction fraction with ligand, whereas Leu 83 with 98% while the remaining 2% was contributed by water bridge. Dramatically, during the molecular dynamics simulation, H-bond interaction fraction of Lys 33 was reduced to 36% and also an ionic bond was formed by contributing 55.5% of the total interaction fraction. Leu 134, Tyr 15, Ile 10, Ala 31, Val 18, and Phe 80 formed hydrophobic interactions with ligand in which Leu 143 had the highest hydrophobic interaction fraction (79.5%). An interaction feature of various residues involved in binding is shown in Figure 7. Formation of water bridges is significant during 450–500 ps and 800–950 ps time intervals. Binding pattern of protein residues over a time period of 1000 ps was illustrated in the protein–ligand interaction diagram in Figure 8.

Root mean square fluctuation (RMSF) plots are useful for understanding the local changes in the protein chain during
the molecular dynamic simulations. RMSF of protein side chains and alpha carbons are depicted in Figure 9. With respect to the secondary structure of the protein, the RMSF plot was categorized into three regions as alpha helical (pink), beta-strand (blue), loops in white, and green lines, representing that the amino acids interacted with ligand. A major portion of the interactions was confined between 85 and 150 residues and also all the interactions were formed from the amino acids present in the beta-strands. Very few residues near 50 and 240 in protein backbone Cz exceeded 1.5 Å and most of the remaining amino acids lied below 1 Å. Fluctuations are relatively higher in the case of the side chain structure of protein. Higher fluctuations were observed in loop regions. Monitoring the RMSD of the simulated protein can give insights into its structural conformation throughout the simulation. Small protein changes of the order of 1–3 Å are perfectly acceptable and changes much larger than that indicate that the protein is undergoing a large conformational change during the simulation. Protein RMSD data of C- alphas and backbone structures in Figure 10 revealed that the conformational changes of protein over 1000 ps simulation were acceptable and also most of the protein deviations were confined between 0.9 Å and 1.20 Å. The ligand RMSD over right Y-axis indicates the ligand stability with respect to the protein and its binding pocket. The RMSD plot confirmed that the ligand remained intact.

Figure 8. Protein–ligand interaction diagram of protein residues during 1000 ps time period.

Figure 9. Fluctuations for the C- alphas and side chain of CDK2 protein.
with the binding site of protein as at the end the plot of ligand was no larger than that of the protein (backbone and C-alphas). If the ligand RMSD were found to be significantly larger than that of the protein, ligand will be diffused from the initial binding site.

We also examined the molecular properties like ligand RMSD, molecular surface area (MolSA), solvent accessible surface area (SASA), and polar surface area (PSA) of the ligand during the course of molecular dynamics (Figure 11). RMSD of ligand is almost consistent between 0.7 and 0.8 Å till 450 ps and later fluctuations in RMSD were significant. There is only 8 Å change in the final molecular surface of ligand at the end of 1000 ps. Similarly, initial and final polar surface areas also showed no significant change. However, a remarkable change was observed in the SASA during the simulation. Initial SASA of ligand 7aa was about 48 Å² and, after 1000 ps simulation, it increased to 76 Å². It might be the reason for the formation of an H-bond with a water molecule (water bridge) during the last stage of the simulation. This large increase of the surface area might be the reason for the remarkable change in the conformation of the ligand.

**Conclusion**

In the current investigation, we have elucidated the common pharmacophore model (AADDRR.33) responsible for CDK2 inhibition from a series of Benzodipyrazoles. The identified pharmacophore model consists of two hydrogen bond acceptors, two hydrogen bond donors, and two aromatic rings. An efficient 3D QSAR model was established for the given set and we proposed that this model provides a possible structural basis for further drug investigations for CDK2 inhibition. Molecular docking studies revealed the binding possibilities of benzodipyrazoles in the kinase domain.
Further molecular dynamic simulations clearly elucidated the conformational changes in the protein ligand complex during a certain period of time.

Acknowledgements
The authors would like to thank Dr. Yamini Lingala, Department of Chemistry, Nijam College – Osmania University, Hyderabad, India, and Mr. Ravi Raja Merugu, Department of Bioinformatics, SRM University, Tamil Nadu, India, for their great support and providing the necessary tools to carry out the simulation studies.

Declaration of interest
The authors report no conflicts of interest.

References
1. Malumbres M, Barbacid M. Cell cycle, CDKs and cancer: a changing paradigm. Nat Rev Cancer 2009;9:153–66.
2. Collins I, Garrett MD. Targeting the cell division cycle in cancer: CDK and cell cycle checkpoint kinase inhibitors. Curr Opin Pharmacol 2005;5:366–73.
3. Malumbres M, Barbacid M. Mammalian cyclin dependent kinases. Trends Biochem Sci 2005;30;630–41.
4. Tsutsui T, Hesabi B, Moons DS, et al. Targeted disruption of CDK4 delays cell cycle entry with enhanced p27(Kip1) activity. Mol Cell Biol 1999;19:7011–19.
5. Tetsu O, McCormick F. Proliferation of cancer cells despite CDK2 inhibition. Cancer Cell 2003;3:233–45.
6. Vermeulen K, Van Bockstaele DR, Berneman ZN. The cell cycle: a review of regulation, deregulation and therapeutic targets in cancer. Cell Prolif 2003;36:131–49.
7. de Cárcer G, Pérez de Castro I, Malumbres M. Targeting cell cycle kinases for cancer therapy. Curr Med Chem 2007;14:969–85.
8. Orzáez M, Gottat A, Mondragón L, et al. ATP-noncompetitive inhibitors of CDK-cyclin complexes. ChemMedChem 2009;4:19–24.
9. D’Alessio R, Bargiotti A, Metz S, et al. Benzodipyrrozole: a new class of potent CDK2 inhibitors. Bioorg Med Chem Lett 2005;15:1315–19.
10. Ligprep, version 2.6. New York: Schrödinger, LLC; 2013. (http://www.schrodinger.com/LigPrep/).
11. Phase, Version 3.5. New York: Schrodinger, LLC; 2013. (http://www.schrodinger.com/Phase/).
12. Dixon SL, Smodyrev AM, Knoll EH, et al. PHASE: a novel approach to pharmacophore modeling and 3D database searching. J Comput Aided Mol Des 2006;20:647–71.
13. Rajendra Prasad VV, Deepak Reddy G, Appaji D, et al. Chemosensitizing acridones: in vitro calmodulin dependent cAMP phosphodiesterase inhibition, docking, pharmacophore modeling and 3D QSAR studies. J Mol Graph Model 2013;40:116–24.
14. Golbraikh A, Tropsha A. Beware of q2!. J Mol Graph Model 2002;20:269–76.
15. Friesner RA, Murphy RB, Repasky MP, et al. Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein–ligand complexes. J Med Chem 2006;49:6177–96.
16. Guimaraes CR. MM-GBSA rescoring of docking poses. Methods Mol Biol 2012;819:255–68.
17. Muñoz C, Adasme F, Alzate-Morales JH, et al. Study of differences in the VEGFR2 inhibitory activities between semaxanib and SU5205 using 3D-QSAR, docking, and molecular dynamics simulations. J Mol Graph Model 2012;32:39–48.
18. Desmond molecular dynamics system, version 3.9. New York, NY: D. E. Shaw Research; 2014. (http://www.desshawresearch.com/resources_desmond.html).
19. Shivakumar D, Harder E, Damm W, et al. Improving the prediction of absolute solvation free energies using the next generation OPLS force field. J Chem Theory Comput 2012;8:2553–8.

Supplementary data available online
Supplementary Table 1