QSAR, pharmacophore modeling and molecular docking studies to identify structural alerts for some nitrogen heterocycles as dual inhibitor of telomerase reverse transcriptase and human telomeric G-quadruplex DNA

R. D. Jawarkar¹,4*, R. L. Bakal¹, P. N. Khatale², Israa Lewaa³, Chetan M. Jain¹, Jagdish V. Manwar⁴ and Minal S. Jaiswal⁴

Abstract

Background: Telomerase reverse transcriptase (TERT) and human telomeric G-quadruplex DNA are amongst the favorable target for researchers to discover novel and more effective anticancer agents. To understand and elucidate structure activity relationship and mechanism of inhibition of telomerase reverse transcriptase (TERT) and human telomeric G-quadruplex DNA, a QSAR modeling and molecular docking were conducted.

Results: Two robust QSAR model were obtained which consist of full set QSAR model ($R^2$: 0.8174, $CCC_{tr}$: 0.8995, $Q^2_{loo}$: 0.7881, $Q^2_{LMO}$: 0.7814) and divided set QSAR model ($R^2$: 0.8217, $CCC_{tr}$: 0.9021, $Q^2_{loo}$: 0.7886, $Q^2_{LMO}$: 0.7783, $Q^2_{F1}$: 0.7078, $Q^2_{F2}$: 0.6865, $Q^2_{F3}$: 0.7346) for envisaging the inhibitory activity of telomerase reverse transcriptase (TERT) and human telomeric G-quadruplex DNA. The analysis reveals that carbon atom exactly at 3 bonds from aromatic carbon atom, nitrogen atom exactly at six bonds from planer nitrogen atom, aromatic carbon atom within 2 Å from the center of mass of molecule and occurrence of element hydrogen within 2 Å from donar atom are the key pharmacophoric features important for dual inhibition of TERT and human telomeric G-quadruplex DNA. To validate this analysis, pharmacophore modeling and the molecular docking is performed. Molecular docking analysis support QSAR analysis and revealed that, dual inhibition of TERT and human telomeric DNA is mainly contributed from hydrophobic and hydrogen bonding interactions.

Conclusion: The findings of molecular docking, pharmacophore modelling, and QSAR are all consistent and in strong agreement. The validated QSAR analyses can detect structural alerts, pharmacophore modelling can classify a molecule’s consensus pharmacophore involving hydrophobic and acceptor regions, whereas docking analysis can reveal the mechanism of dual inhibition of telomerase reverse transcriptase (TERT) and human telomeric G-quadruplex DNA. The combination of QSAR, pharmacophore modeling and molecular docking may be useful for the future drug design of dual inhibitors to combat the devastating issue of resistance.

*Correspondence: rahuljawarkar@gmail.com
4 Department of Medicinal Chemistry, Dr. Rajendra Gode College of Pharmacy, University-Mardi Road, Amravati, Maharashtra 444602, India
Full list of author information is available at the end of the article

© The Author(s) 2021. Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.
Background

Transcriptase has emerged as a possible drug target in cancer therapeutics because of following whys and wherefores [1]: Every human being has telomerase, which is active in the early stages of life to preserve telomere duration. This will maintain the chromosomal integrity of recurrently dividing cells, which is supposed to be inactive in most somatic cells and is maintained during adulthood [1]. The telomerase enzyme is found in 80–90% of cancer cells isolated from major human tumors, although it is not found in the adjacent cells of healthy human tissue [2, 3]. Given these evidences and particulars, telomerase has gotten a lot of attention as a potential target for developing new anticancer drugs.

Telomerase is an RNA–protein complex (RNP) that comprises 30 ends of linear chromosomes which give rise to generation of the small telomere-repeat sequence (TTG GGG in ciliates and TTA GGG in humans) by using RNA prototype. The RNA prototype is a component of its essential telomerase RNA (TER) component and its basic telomerase reverse transcriptase (TERT) [4]. Telomerase activity is strongly controlled throughout development and oncogenesis [5]. The reported evidences for telomerase makes it a popular therapeutic target, as well as inhibitory agents with potential for cancer treatment [6]. Many hTERT inhibitors were reported [2, 7], and some of them, comprising BIBR1532 [8–10] showed promising anticancer effects [11–13].

In silico development of new molecules by using a logical structure-based drug design approach that identifies prominent hits and confirms activity for carefully selected hits. As a result, finding such hits using a hybrid ligand- and structure-based drug design approach will aid in the balanced development of more effective telomerase inhibitors [14, 15].

Furthermore, noncoding repetition orders of the guanine-rich DNA, which are important for preventing the cell from recombination and degradation, are found at the telomeric ends of the chromosomes. Telomere are the extremities of eukaryotic chromosomes, are necessary for maintaining genome integrity and appear to play crucial a role in cellular aging and the cancer.

Keywords: QSAR, TERT, Telomeric, G-quadruplex DNA, Molecular docking

Graphical abstract
They consist of tandemly replicated DNA sequences with a G-rich strand directed 5′ to 3′ towards the chromosome’s end [13]. The telomerase activation is associated with the shortening of telomere, tracked by the activation of the DNA destructive responses. This involves the cell cycle arrest, senescence and apoptosis [16–22]. Telomerase is an imminent anticancer drug target which is supported by the evidence that its activity is described and reported in 85–90% of all human tumors, but not in normal cells [23]. The Cell cycle arrest and apoptosis effects have rendered telomerase as a striking target in the field of anticancer drug discovery and novel therapeutics [24–28].

To date, a diverse array of G-quadruplex-stabilizing compounds has been investigated and reported by various researchers, including macrocyclic oxazolo [29–33], anthraquinones [34, 35], acridines [36–40], cationic porphyrins [41–46], bistriazoles [47], perylene [48–50], ethidium derivatives [51, 52], fluorenines [53], pentacyclic acridinium salts [54–56], and fluoroquino phenoazoxines [57–62].

Quantitative structure–activity relationship (QSAR) studies attempt to find new and similar molecules in the broad databases of reported molecules with known established observed activities or properties [63–65]. The discovery of such a statistical correlation opens up the possibility of predicting the activities and properties of new compounds and, as a result, guiding the synthesis of new molecules without having to implement it.

Multidrug resistance (MDR) is one of the major concerns associated during the course of anticancer treatment. On the other hand, the inhibition of a single target repeatedly display momentary effectiveness due to emergence of the drug resistance [66]. Perceiving that cancers are heterogeneous entities, the simultaneous inhibition of multiple targets is needed to obtain the optimal effect. As a result, finding new and safe dual inhibitors is critical to overcoming the resistance issue in cancer treatment. The main emphasis in current study is to build a QSAR model which find the various structural alerts and the features in nitrogen heterocycles containing compound and their correlation with telomerase reverse transcriptase (TERT) and human telomeric G-quadruplex DNA inhibition. Furthermore molecular docking, may perhaps to be used to understand the dual inhibitory mechanism and interactions between the ligands and receptor for predicting their binding affinity. In addition, pharmacophore modeling is used to reveal consensus pharmacophoric features required for the dual inhibition of TERT and human telomeric G-quadruplex DNA.

**Method**

The QSAR experimental methodology consist of selection of dataset, calculation of molecular descriptors, feature selection algorithm, validation process, and correlation in relation to structural landscapes (i.e. OECD guidelines). The main goal of using experimental methodology is to build a QSAR model with a good equilibrium of external predictive capability (quantitative/predictive QSAR) and understanding of QSAR model as well as molecular descriptors in terms of structural alerts responsible for biological activity (qualitative/descriptive QSAR) [67–73].

Thus, classic method has been followed to develop a stable QSAR model for inhibitory activity of nitrogen ring containing heterocycles for TERT and human telomeric G quadruplex DNA. More particulars about the technique followed in the current work are available in the literature [71, 74–78].

**Selection of dataset**

In this study, a dataset of structurally varied 82 nitrogen ring containing heterocycles experimentally tested against Human TERT inhibitory potential has been carefully chosen for QSAR investigation from renown and publicly accessible ChEMBL database (https://www.ebi.ac.uk/chembl/). The dataset includes a wide range of molecules with different substituents, such as acridine rings, triazoles rings, pyrimidine rings, and so on. As a result, in order to create a robust QSAR model, we try to cover as much chemical space as possible. The experimentally reported EC50 values range from 2 to 23,500 nM, which were transformed to pEC50 (−log10EC50) prior to QSAR model building. The SMILES notations, EC50 and pEC50 for selected nitrogen heterocycles as an example only are presented in Table 1. In Table 1, we have depicted five most and least active compounds as representative examples only. In addition, the common scaffolds have been presented in Fig. 1 (Presentation of Serial number, CheMBL id, Smiles and Pec50 value of 82 compounds is given in Additional file 1: Table S5).

The 2D-structures of all 82 compounds were sketch by ACD ChemSketch Freeware (www.acdlabs.com) tracked by conversion to 3D structures using Avogadro ver. 1.02 (https://avogadro.cc/) by means of MMFF94 force field for geometry optimization and partial charge assignment. The resulting parameters were used for geometry optimization: Force Field: MMFF94, Algorithm: Steepest Descent, numeral of steps used for optimization: 1000.

**Calculation and pruning of molecular descriptors**

PyDescriptor were employed for molecular descriptor calculation using 3D-optimized structures of
In the present study, we have implemented objective feature selection (OFS) to minimize the group of molecular descriptors. The OFS method intricated removal of persistent, nearly persistent (95% molecules) and greatly correlated molecular descriptors ($R > 0.90$). OFS lead to condensed bunch of 494 molecular descriptors, which still encompasses comprehensive range of PyDescriptor because of incidence of 1D- to 3D-molecular descriptors. This reduced descriptor pool was further used for the subjective feature selection (SFS) [71, 78–80] (Calculated Py-Descriptor values used to build QSAR model is depicted in Additional file 1: Table S6).

**Subjective feature selection (model building)***

The method of SFS was executed by using QSARINS-2.2.4 with default setting in which a number of descriptors was set to 1000. The genetic algorithm (GA) module available in QSARINS-2.2.4 employ $Q^2$ as a fitting parameter to circumvent over fitting and insertion of redundant variables during model building. A key decision in evolving a successful QSAR model is to stop adding molecular descriptors to the model at appropriate point of time. In the current study, breaking point is obtained by means of graph was drawn amid the number of descriptors intricated in the models and $Q^2$ value. In the conclusion, the number of molecular descriptors conforming to the breaking point was considered optimal for model development. The graph amid numbers of variables used in the models against $Q^2$ value is shown in Fig. 2. From Fig. 2, it is clear that the breaking point links to five variables. Therefore, five descriptor were used to derive robust QSAR model while QSAR models with more than five descriptors were excluded [79, 81]. To gain deep understanding of structural features prominent for Dual inhibition of telomerase reverse transcriptase as well as human telomeric G DNA, we are expected to develop (Descriptive QSAR) statistically satisfactory GA-MLR originated QSAR model [82, 83].

**Validation of QSAR model***

Following the development of a QSAR model, it is extremely crucial and important to validate the model for external predictive potential in order to determine its performance and scope for predicting biological activity in lead/drug optimization during the drug discovery phase [56, 57, 73]. As a result, not only were extensive internal validations and the Y-scrambling technique used, but an external prediction range of 20% molecules was also used to verify the model's statistical robustness. A Williams plot was also created to ascertain the applicability domain of the developed QSAR model. Additionally, the guidelines were used to choose and validate a QSAR model. $R^2_{tr}$ 0.6, $Q^2_{loo}$ 0.5, $Q^2_{LMO}$ 0.6, $R^2 > Q^2$, $R^2_{ex}$ 0.6 RMSE$_{tr}$ < RMSE$_{cv}$, DK 0.05, CCC 0.80, $Q^2$-Fn 0.60, $r^2_m 0.5, (1 - r^2_o/r^2_m) < 0.1 0.9 k 1.1 or (1 - $r^2/p^2_o) < 0.1, 0.9 k' 1.1, r^2_o < 0.2 < 0.3$ with RMSE and MAE as low as possible. The $Q^2_{LMO}$ value stated here is mean value of 2000 repetitions with 30% of the population (molecules) arbitrarily excluded from the training set at each time. The external predictive capacity of model was find out by using external validation parameters, viz., RMSE$_{ex}$ MAE$_{ex}$, $R^2_{ex}$, $Q^2F1$, $Q^2F2$, $Q^2F3$, and CCC$_{ex}$. All QSAR models that do not hallucinate.
Fig. 1 Structural variations in dataset compounds used for QSAR modeling
not achieve the endorsed lower-limit values for above statistical parameters have been directly excluded,

The inter-correlation between descriptors were tested by the QUIK rule ($Q$ under Influence of $K$). QUICK rule was fixed to 0.05 to lessen inter-correlation among descriptors. The reliability of the developed QSAR model was ascertained by Y-randomization set at 2000 iterations to check the fitting of the randomly reordered Y-data. For the randomization of the build QSAR model, the dependent variables (PIC50 value) of the training set have been shuffled and new coefficients of determination were calculated. The significantly low value of the coefficients of determination of the new models specify that the reported model in the present QSAR analysis is not obtained by chancy correlation. The external validation of all the models were verified with the subsequent validation criteria: $r^2_{\text{ext}}$ (external determination coefficient), $Q^2_{\text{F1}}$, $Q^2_{\text{F2}}$, $Q^2_{\text{F3}}$, concordance correlation coefficient (CCC), $CCC_{\text{ext}}$, $r^2_{m}$, and $\Delta r^2_{m}$. The parameter $r^2_{m}$ (overall) penalizes a model for large differences between observed and predicted values of the compounds of the whole set (considering both training and test sets). The $\Delta r^2_{m}$ estimated the indulgent between the values of the predicted and the resultant experimental activity data ($\mu$IC$^{50}$ value). It has been reported that, the observed value for the $\Delta r^2_{m}$ should be preferentially lower than 0.2 provided that, value of $r^2_{m}$ > 0.5.

Additionally, all the QSAR models were evaluated for validation parameters, such as Golbraikh and Tropsha's criteria to justify model reliability and robustness. Generally, good predictive ability of the developed QSAR model depends upon closeness of predicted value against observed (experimental biological activity) value. Even, presence of single outlier diminish the predictive capability of the developed QSAR model. Subsequently, we have tried to highlight the outlier on the basis of those compounds who showed significantly high residual value in GA-MLR QSAR models. Moreover, we have identified the outlier compounds by comparing the predicted value with three standardized residual values. Likewise, structural variation in database compounds was observed by leverage effect in Williams plot. The applicability domain of the developed QSAR model is ascertained by merging the leverage and the standard residuals [68, 69, 72, 74, 84, 85].

**Pharmacophore modeling**

To achieve a consensus pharmacophore model, we have generated lowest energy conformer of the most active compound 82 that was used to align all of the molecules in the dataset. Then, LIQUID 1.0, a free PyMOL plugin, was used to produce consensus pharmacophore model expending default settings. Whereas, in Fig. 12, the pdb files (5cqg) were obtained from publicly accessible database (www.rcsb.org). Afterward, the bound pdb ligands were isolated without any alteration or optimization, that is, the X-ray crystallographic resolved crystal structure of extracted ligand was used as it is to generate the pharmacophore model with LIQUID 1.0 [86] (Fig. 4).

**Docking of the inhibitors**

Molecular docking studies were performed using the NRGSuite software package. The NRGSuite package equipped with FlexAID that contains four primary panels to specify the input target protein and ligand to be docked, configuration of the target and ligand and simulation. All the crystallized water molecules and coordinated molecules were found in the crystal structure of BIBR1532 anchored with the *Tribolium castaneum* catalytic subunit of the RNA prototype the telomerase (tcTERT) were preserved (pdb id-5cqg) during a docking procedure. The binding site residues of tcTERT was ascertain by using site finder function in NRGSuite to determine active site in crystallized ligand BIBR1532.
with the extraction of ligand structure present in pdb for redocking purpose. Site finder option reveals two binding site in chain A and B where crystallized ligand BIBR1532 was found to be bound. As a result, the inhibitor compound 82 and the co-crystallized ligand were docked using NRGSuite’s site finder function. The FlexAID uses genetic algorithm. A number of important parameters, notably the number of chromosomes and generations can be defined in this panel. Additionally, the number of top results that are visualized during the simulation and the frequency (in numbers of generations) to refresh the visualization can be set. Then, the poses obtained during the placement stage were then fine-tuned using the Induced Fit method, which allows for protein versatility during ligand binding and thus improves the interaction prediction accuracy. The top five uppermost scoring poses were then achieve with the GBVI/WSA dG scoring utility. The final performance was analyzed, and docked poses inside the binding site that were not properly oriented (for catalytic site) and whose versatile alignments were compared to the top scoring pose of compound 82 were included. The docking poses of compound 82 were then superimposed on top of the co-crystallized ligand BIBR1532 to determine the most favorable docked conformation for telomerase inhibition.

**Results**

**Full set model**

\[ pEC_{50} = 4.191 \pm 0.596 + 0.407 \pm 0.09 \cdot fdonH2A + -0.251 \pm 0.095 \cdot ringN\_acc\_8A + 0.261 \pm 0.131 \cdot faroCC3B + 0.214 \pm 0.095 \cdot fplaNN6B + 0.263 \pm 0.065 \cdot com_aroC\_2A++ \]

\[ R^2: 0.8174, \quad R^2_{\text{adj}}: 0.8054, \quad R^2 – R^2_{\text{adj}}: 0.0120, \quad \text{LOF}: 0.1499, \quad K_{\text{cv}}: 0.2604, \quad \Delta S_{\text{cv}}: 0.1186, \quad \Delta S_{\text{esc}}: 0.3400, \quad \text{MAE}_{\text{cv}}: 0.2510, \quad \text{RSS}_{\text{cv}}: 9.4780, \quad \text{CCC}_{\text{cv}}: 0.8995, \quad s: 0.3531, \quad F: 68.0476, \quad Q^2_{\text{loo}}: 0.7881, \quad R^2 – Q^2_{\text{loo}}: 0.0293, \quad \text{RMSEE}_{\text{cv}}: 0.3663, \quad \text{MAE}_{\text{esc}}: 0.2702, \quad \text{PRESS}_{\text{cv}}: 11.0009, \quad \text{CCC}_{\text{cv}}: 0.8831, \quad Q^2_{\text{LMO}}: 0.7814, \quad R^2 – Y_{\text{sc}}: 0.0597, \quad Q^2 – Y_{\text{esc}}: 0.0987, \quad \text{RMSEE}_{\text{Y}_{\text{cv}}}: 0.7714. \]

(Depiction of QSAR results along with their experimental and predicted EC50 values for full set model is given in Additional file 1: Table S4).

**Divided set model (80:20)**

\[ pEC_{50} = 5.355 \pm 0.563 + 0.367 \pm 0.1 \cdot fdonH2A + -0.287 \pm 0.098 \cdot ringN\_acc\_8A + 0.312 \pm 0.119 \cdot notringC_aroC\_2B + 0.336 \pm 0.127 \cdot fplaNN6B + 0.274 \pm 0.071 \cdot com_aroC\_2A++ \]
$R^2$: 0.8217, $R^2_{\text{adj}}$: 0.8068, $R^2-R^2_{\text{adj}}$: 0.0149, LOF: 0.1600, $K_{\text{xx}}$: 0.2783, Delta $K$: 0.0936, RMSE$_{\text{tr}}$: 0.3394, MAE$_{\text{tr}}$: 0.2636, RSS$_{\text{tr}}$: 7.6042, CCC$_{\text{tr}}$: 0.9021, $s$: 0.3560, $F$: 55.2934, $Q^2_{\text{loo}}$: 0.7886, $R^2-Q^2_{\text{loo}}$: 0.0331, RMSE$_{\text{cv}}$: 0.3696, MAE$_{\text{cv}}$: 0.2882, PRESS$_{\text{cv}}$: 9.0164, CCC$_{\text{cv}}$: 0.8842, $Q^2_{\text{LMO}}$: 0.7783, $R^2_{Y_{\text{extr}}}$: 0.9777, $Q^2_{Y_{\text{extr}}}$: −0.1253, RMSE$_{\text{AVY}_{\text{extr}}}$: 0.7717, MSE$_{\text{extr}}$: 0.4114, MAE$_{\text{extr}}$: 0.3415, PRESS$_{\text{extr}}$: 2.7438, $R^2_{\text{extr}}$: 0.7020, $Q^2-F1$: 0.7078, $Q^2-F2$: 0.6865, $Q^2-F3$: 0.7346, CCC$_{\text{extr}}$: 0.8354, $r^2$ aver.: 0.5878, $r^2$ delta: 0.0810.

In the present QSAR analysis, compound 82 was detected as X outlier while compound 8, 13 were depicted as high leverage influential (see Fig. 5). Moreover, apart from fitness function $Q^2$, we have displayed another fitness function of concordance correlation coefficient (CCC$_{\text{extr}}$) as one of the external validation parameter. (Depiction of QSAR results for divided set model along with their experimental and predicted $EC_{50}$ values for divided set model is given in Additional file 1: Table S3).

**Discussion**

**faroCC3B**
The descriptor faroCC3B point out frequency of occurrence of carbon atom exactly at 3 bonds from aromatic carbon atom. Since this descriptor has positive coefficient, this means that, increase in the value of this descriptor will increase its $pEC_{50}$ value for the molecules used in present study. In compound 81, C3 and C6 propanamide substituent placed at a topological distance of 3 bonds from C3 and C6 aromatic carbon atoms of acridine ring, amino methyl carbon atom placed at C2 position of pyrrolidine ring is separated by the topological distance of three bonds from C9 aromatic carbon of acridine ring and C3, C4 carbon atoms of pyrrolidine ring is placed at a topological distance of three bonds from C9 aromatic carbon of acridine Ring. Same pattern of descriptor faroCC3B is observed in compound 7 except pyrrolidine ring is replaced by amino cyclopropyl substituent on C9 position of acridine ring. If we compare activity profile of compound 81 and 75, we observed that pyrrolidine substituent is more favorable for anticancer activity rather than amino cyclopropyl substituent. This observation highlight the variation in the activity of compound 81 and 75 in nanomolar range.

To add more, both aliphatic chain and unsaturated centers in molecule significantly contributed to the overall lipophilicity of the molecule. Therefore, lipophilicity is the key feature that govern dual inhibition of TERT and human G DNA. Subsequently, attachment of pyrrolidine ring to the acridine ring through the single bond impart enough flexibility to the compound 81 therefore, sterically lock it into the active conformation within TERT and human telomerase G DNA.

Ranganathan et al. and Khanna et al. have reported that, most of the molecules in the metabolite dataset used studies contains a carbon atoms in the range 35–55.
which is 32% i.e. 5–25 carbon atoms per molecule. The carbon atom content in metabolites has a mean of 33 atoms and maximum up to the 100. On the other hand, drugs molecule have an average of 18 carbon atoms per molecule, with a maximum of 256 and 76% of drugs consist of carbon atoms in the range of 5–25 [85]. In QSAR model descriptor faroCC3B highlight the importance of occurrence of carbon atom exactly at 3 bonds from aromatic carbon atom, further diminishes the anticancer activity of compound 66. For increasing anticancer profile of compound 66, substitution of pyrrolidine ring with an amino methyl substitution at C9 position of aromatic carbon of acridine ring is recommended. This observation supported the fact that, compound 81 and 75 have five such carbon atoms placed at a topological distance of three bonds from aromatic carbon atoms while compound 66, 45 and 8 shows two centers except compound 13 in which carbon with topological distance of three bond from aromatic carbon atom is missing, therefore it is clear that variation in pEC50 value is due to absence of carbon atoms at a topological distance of three bonds from aromatic carbon atoms. Here greater the number of carbon atom at a topological distance of three bonds from aromatic carbon atom, higher will be the anticancer activity of stated compounds under study. Enhance cloud of carbon atoms in the range of 5–25 [85]. In QSAR model descriptor faroCC3B highlight the importance of occurrence of carbon atom exactly at 3 bonds from aromatic carbon atom in dataset molecules (see Fig. 6).

Further, in compound 66, descriptor faroCC3B is observe at C3 and C6 aromatic carbon of acridine ring which reveals that, decrease in the cloud of carbon atoms placed at topological distance of three bonds from the aromatic carbon atom, further diminishes the anticancer activity of compound 66. For increasing anticancer profile of compound 66, substitution of pyrrolidine ring with an amino methyl substitution at C9 position of aromatic carbon of acridine ring is recommended. This observation supported the fact that, compound 81 and 75 have five such carbon atoms placed at a topological distance of three bonds from aromatic carbon atoms while compound 66, 45 and 8 shows two centers except compound 13 in which carbon with topological distance of three bond from aromatic carbon atom is missing, therefore it is clear that variation in pEC50 value is due to absence of carbon atoms at a topological distance of three bonds from aromatic carbon atoms. Here greater the number of carbon atom at a topological distance of three bonds from aromatic carbon atom, higher will be the anticancer activity of stated compounds under study. Enhance cloud of carbon atom augments lipophilicity which in turn indicate maximum hydrophobic interaction with receptor.

fplaNN6B
The descriptor fplaNN6B stand for the frequency of occurrence of nitrogen exactly at six bonds from planer nitrogen atom. The positive coefficient designates that an increase in the number of such Nitrogen atoms may plausibly enhances the anticancer activity (pEC50 value). Ranganathan et al. and Khanna et al. have recognized that drugs molecules clearly possess the maximum number of the nitrogen atoms, followed by toxin molecules and lastly, metabolites [85].

Pennington et al. established the importance of nitrogen in heterocyclic compounds. Pennington et al. reported that the replacement of a CH group with N atom in aromatic and hetero aromatic ring structures can have many beneficial effects on molecular and physicochemical properties and intra and intermolecular interactions that may give rise to improved pharmacological profiles in drug discovery. Moreover, Pennington et al. also investigated that, a N atom in aromatic and heteroaromatic ring systems can influence the number of intra- and intermolecular orbital, steric, electrostatic, and hydrophobic interactions such as lone pair, dipole–dipole, hydrogen bonding, metal coordination, van der Waals, σ-hole, σ**S−X, and π-system interactions, which in turn can translate to modified pharmacological profiles [87].

This observation is reinforced by a simple comparison of the subsequent pair of molecules: comp-81 (pEC50-7.74 nm) with comp-40 (pEC50-6.59 nm), comp-80 (pEC50-7.74 nm) with comp-45 (pEC50-6.67 nm) (see Fig. 2). In case of compound 2, this feature is missing. Therefore, we can say that presence of planer nitrogen is most important for augmenting biological activity performance of molecule (see Fig. 7). In our dataset, compound 62 to 82 contains two planer nitrogens, compound 40, 45, 46, 52, 67, 78, 82 consist of one planer nitrogen while planer nitrogen in absent compounds 1 to 4 and 6 to 32. Here, nitrogen becomes planar when its lone pair becomes involved in pi-bonding. The five-membered rings have significant delocalization of electrons to produce a cloud system similar to that in benzene. As a result, planer nitrogen increases the electron cloud in the molecule, which strengthens electrostatic interactions with the receptor surface by exhibiting pi bonding.

ringN acc_8A
This descriptor depicts occurrence of ring nitrogen within 8 Å from the acceptor atom. In QSAR model, this descriptor has negative correlation with pEC50 value. Therefore, number of ring nitrogen within 8 Å from acceptor atom must be retained, as low as possible to enhance the anticancer activity (pEC50 value). Increase in the value of descriptor ringN acc_8A will further decreases the anticancer activity profile of the compounds in dataset. This is observed when pEC50 of compound 23 compared with compound 41. This could be the possible reason for the difference in the pEC50 value of compound 23 and 41 (see Fig. 8).

In general, it is established that, close combination of ring nitrogen and acceptor is avoided to prevent intramolecular hydrogen bonding in the molecule. Specially, when oxygen and nitrogen are connected by single bond to the neighbor atoms. Thus, the descriptor ringN acc_8A provides a hint to avoid close proximity of ring nitrogen with acceptor atom to avoid the prospect of intramolecular bonding.

com aroC 2A
The descriptor com aroC 2A specifies occurrence of aromatic carbon atom within 2 Å from center of mass of molecule. This descriptor is positively correlated with pEC50. Hence, this value must be kept as high as possible. In case of compound 82 (pEC50=8.69 nm), there are
Aromatic Carbon atom

non-aromatic Carbon atom

Comp-81 (pEC50 = 7.745 nm)

Aromatic Carbon atom

non-aromatic Carbon atom

Comp-75 (pIC50 = 7.30 nm)

Aromatic Carbon atom

non-aromatic Carbon atom

Comp-66 (pIC50 = 7.04 nm)

Aromatic Carbon atom

non-aromatic Carbon atom

Comp-45 (pIC50 = 6.67 nm)

Aromatic Carbon atom

non-aromatic Carbon atom

Comp-8 (pEC50 = 5.237 nm)

Aromatic Carbon atom

non-aromatic Carbon atom

Comp-13 (pIC50 = 5.53 nm)

Fig. 6 2D pictorial depiction of faroCC3B descriptor for the compounds 81, 8, 13, 45, 66 and 75
Fig. 7  2D pictorial depiction of fplaNN6B descriptor for compound 81, 80, 40, 45 and 2
around six carbons present within the radius of 2 Å from the center of mass of the molecule while in compound 19, only three carbons are present within 2 Å from center of mass of molecule. Therefore, it is reasonable to settle that, difference in the activity of stated compound is due to the number of carbons present within 2 Å from the center of mass (see Fig. 9).

The same is true for compound 13, which has five carbons within 2 Å of the center of mass. It is reasonable to conclude that the activity of compounds 82 and 13 differs solely due to the number of carbons within 2 Å of the center of mass, and that this may be the cause of the differences in the activity profiles of both molecules.

This could be the possible reason for the differences in the activity profile of both molecules. In general, the presence of aromatic carbon atom affect overall lipophilicity of the molecule, therefore, it is rational to predict that, carbon atoms present in the vicinity of center of mass plays crucial role in hydrophobic interactions with the receptor.

\textbf{notringC\_aroC\_2B}

The descriptor notringC\_aroC\_2B describes the occurrence of non-ring carbon atom exactly at or within 2 bonds from aromatic carbon atoms, providing different level and type of useful information. Since this descriptor has a negative coefficient in the model, raising its value can result in a lower activity profile. In this case, a compound with a higher number of non-ring carbon atoms exactly at or within two bonds from aromatic carbon atoms might have lower activity than one with fewer of these aromatic carbons. This is observed when comparing compound 37 to 61 and compound 33 to 78 (see Fig. 10). This finding supports the fact of the variation in activity of the stated compounds.

In compound 78, non ring carbon atom containing amide group is present at terminal position, therefore, it may establish that, these substituents occupy lipophilic pocket of the TERT as well human telomeric G DNA. To add more, these substituent varying steric bulk in the receptor pocket, thereby blocking the enzyme. Besides, the presence of aliphatic chain along with unsaturated pyrrolidine imparts good lipophilicity as well as flexibility to the molecule.

\textbf{fdonH2A}

The descriptor fdonH2A indicates frequency of occurrence of element hydrogen within 2 Å from donar atom. It has positive coefficient in the developed models, therefore the number of Hydrogen atoms in the neighborhood of ring Nitrogen atoms is favorable blend to be used for lead/drug optimization. Since Hydrogen is the smallest element, it suggests that there should be minimum bulk in the vicinity of donar atoms. Therefore, in future structural modifications, steric bulk nearer to donar atoms should be circumvented to have better anticancer activity (see Fig. 11).

As the descriptor fdonH2A specifies necessity of higher number of donar feature with presence of hydrogen atom within 2 Å. In case of compound 68, five donar features are present with five hydrogen atoms within 2 Å while four donar are reported in compound 38. As a result, it is reasonable to conclude that a greater number of donars containing hydrogen atoms is needed for greater telomerase inhibition. This may be a plausible explanation for the differences in PEC50 values among the compounds mentioned. When we compared the PEC50 values of compound 68 with 38 and compound 26 with 4, we came to the same conclusion. In case of compound 68, steric bulk increase from amide oxygen due to butyl and hexyl aliphatic side chain. These substituents in turn augment the lipophilicity of the compound along with the selectivity towards receptor.

To add further, in compound 68, donar is capable of getting more surface area for hydrogen bonding within receptor pocket due to aliphatic side chain as compared to bulky aromatic substituent. To add more, the interaction between compound 68 and the receptor is possible due to the flexibility of aliphatic butyl side chain.

In our QSAR analysis, diverse Py molecular descriptors demonstrating dissimilar structural landscapes have provided expressive visions into the whys and wherefores for differences in the anticancer activity of dataset compounds.
Pharmacophore modeling

It is a deep-rooted and successful branch of Computer assisted drug design which is executed to recognize key structural alerts (properties) accountable for binding affinity and overall pharmacological activity of ligand. The consensus pharmacophore model displays two larger hydrophobic regions separated at a distance of 7.2 Å and 4.2 Å from hydrogen bond donor and two hydrogen bond acceptors placed at a distance of 2.8 Å and 2.3 Å from hydrogen bond donor. The pharmacophore modeling emphasized the significance of the hydrophobic nature of central acridine ring and its nearby substituents atoms. The similar observation is also reinforced by the occurrence of the descriptor faroCC3B, notringCaroC2B and comaroC2A in the QSAR model as well as recent crystal structures for BIBR1532 with TERT (see Fig. 12).

Based on a comparison of Pharmacophore model with co-crysallized ligand (pdb-5cqg) with pharmacophore model for Most active compound 82, the consensus pharmacophore model and the pharmacophore model obtained using the X-ray resolved crystal structure of extracted ligands are very close especially with respect to the presence of two large hydrophobic region (green colored) at the both end and one H-bond acceptors in the vicinity of acridine nitrogen (red colored). Thus, QSAR and pharmacophore modeling led to recognition of consensus and matching structural
topographies and justified by recent crystal structure of TERT with BIBR1532. Moreover, compound 82 display hydrogen bonding and hydrophobic interaction with human G DNA, therefore, presence of hydrophobic as well as acceptor feature is crucial for binding as well as inhibition of human G DNA. The similar remark is highlighted by the occurrence of the descriptor faroCC3B, notringC_aroC_2B and com_aroC_2A in the QSAR model.

Molecular docking of compound 82 with TERT
Telomerase enzyme is a ribonucleoprotein (RNP) reverse transcriptase responsible for replicating the ends of chromosomes and sustaining genome authenticity. The TERT structure comprises four separate areas (TRBD, fingers, palm, and thumb) well-arranged into a ring thus, producing large interior binding pocket for RNA prototype and telomeric DNA during the whole process of telomere elongation. At the present, BIBR1532 molecule is in clinical trial and chemically, it is a (2-[(E)-3-naphtalen-2-ylbut-2-enoylamino]-benzoic acid). It is a non-nucleosidic, non-competitive, small-molecule inhibitor of telomerase that is regularly and constantly introduced in studies of telomerase function. (Docking results for Compound 82, BABR1532 and Epirubicin into the TERT is given in Additional file 1: Table S1).

Experimental and simulated annealing study reveals the presence of superficial but well-defined hydrophobic pocket located on the external surface of the thumb area of telomerase and following the TRBD-thumb border.
This pocket is formed as a minor gap, around 10 Å wide-ranging and 8 Å deep, prepared by the collection of the tips and connecting loops of the helices 20, 21, and 22, 23. This cavity form as pocket and, referred as the FVYL motif/pocket, present on the well-preserved hydrophobic residues: F478, V491, Y551, and L554. The FVYL amino acid residues exert extensive hydrophobic interactions to stabilize the placement of the rings and helices near the pocket. Many well-maintained and typically hydrophobic amino acid residues occupying the interior of this pocket, which comprise M482, M483, F494, I497, W498, I550, Y551, and L554.

Examination of the BIBR1532 telomerase amalgamated crystal structure data revealed hydrogen bonding interactions amid Asn 421, Arg 433, Lys 437 of telomerase and the carbonyl and carboxylic acid groups of BIBR1532. From the docking analysis of pose 1 of compound 82display docking score of $-9.125$ kcal/mol and occupied FVYL motif/pocket through pi-alkyl hydrophobic interaction with residue L554 via phenyl ring of 4-amino-phenyl substituent. Acridine nitrogen form key hydrogen bonding interaction with water molecule HOH: B735 (3.40 Å) while pyrrolidine ring carbon exhibit two carbon hydrogen bonding with residue GLY: B283 (2.56 Å, 2.81 Å) (see Fig. 13).

Meanwhile, B:MET 482 residue form pi–sulphur interactions with acridine ring, PHE: 494 execute amide–pi interaction with phenyl ring, ARG486 and ILE550 anchored alkyl hydrophobic interactions with pyrrolidine ring. In addition to this ILE497, Leu554 and ILE550 which form shape of the interior of FVYL motif/pocket and exhibit pi-alkyl hydrophobic contact with amino phenyl substituent. Compound 82 acquired same binding conformations as that of crystallized ligand BIBR1532 (see Fig. 13). As previously mentioned, we chose two binding sites 23 and 47 in the site finder choice in MOE to investigate the binding mode of compound 82 due to crystallized ligand BIBR1532 binding at two separate sites during the docking process. Docking analysis divulges that compound 82 acquired two best docked conformation with the docking score of $-9.125$ kcal/mol and $-9.004$ kcal/mol at first site while third docking conformation was acquired in another binding site and exhibit two hydrogen bonding interaction in which water molecule HOH: 784 (3.40 Å) bind with acridine ring and HOH: 799 (2.58 Å) attached with pyrrolidine nitrogen (see Fig. 14).

Likewise, middle acridine form electrostatic contact with ARG 486 residue through pi–cation interactions and PHE494 involved in three pi–pi stacked hydrophobic interactions with all the three acridine rings. Therefore, it is reasonable to say that acridine ring plays crucial role in enhancing binding affinity against TERT and actively involved in drug receptor interactions. In addition to this, MET482, ILE497, LEU554, ILE550, MET483, ARG486 residue from the interior lining of the FVYL motif/pocket are involved in pi-alkyl hydrophobic interactions with amino phenyl substituent and acridine ring.

Again it is not hard to see that compound 82 acquired different conformations within FVYL motif/pocket and
Fig. 13 Depiction of 2D interaction of pose 1 conformation of compound 82 with TERT and 3D view of superimposed alignment of compound 82 (green) with crystallized ligand BIBR1532 (yellow)
key interaction involves hydrogen bonding and hydrophobic interactions with the involvement of water molecule and hydrophobic residues MET482, ILE497, LEU554, ILE550, MET483, and ARG486. When the docking findings were compared to those of the crystallized ligand BIBR1532, the interaction between compound 82 and amino acid residues was found to be close to that of BIBR1532 (see Fig. 15).

The descriptor com_aroC_2A and faroCC3B point out towards importance of lipophilicity in telomerase inhibition. Here Ligand lipophilicity influences target affinity momentarily as most discovered binding sites shows presence of at least one hydrophobic pocket in a nearby aqueous environs. The hydrophobicity give rise to the interaction between the ligands and the protein binding sites through altering the interactions between the protein and solvating waters, therefore exhibiting more promising hydrophobic interactions for both ligand and protein. Therefore, these descriptors give key information about lipophilicity is the most important factor required for telomerase inhibition and plays crucial role in monitoring the balance of hydrophobic features of molecule. Thus, it is sensible to say that docking outcomes are in complete agreement with descriptor com_aroC_2A and faroCC3B. The Descriptor ringN_acc_8A and fplaNN6B highlight the importance of ring nitrogen and planer nitrogens in QSAR model. Docking results depicted that acridine ring nitrogen (ringN_acc_8A) form pi-alkyl hydrophobic interactions with TERT receptor therefore, it is rational to say that docking results are entirely correlated with QSAR findings. Moreover pyrrolidine ring nitrogen exhibit hydrogen bonding with water molecule which again put forth that, nitrogen atom is essential for TERT inhibition.

**Molecular docking study on human telomeric G-quadruplex DNA**

The human and mammalian telomeric DNAs comprises 5’-TTAGGG-3’ repeating sequences that contain numerous base pairs. Binding of small ligand to human telomeric DNA is documented to be stabilize G4 DNA, impedes in functioning of gene expression/regulation is one of the strategy to develop new anticancer agents.

The established X-ray crystal structure of a human quadruplex G DNA made from four uninterrupted human telomeric DNA which repeats and developed at a K1 concentration that come close to its intracellular concentration. K1 ions were reported in the crystal structure. The folding and occurrence of the DNA in reported (pdb id-1kf1) intramolecular quadruplex, is primarily different from the Reported Na1-containing quadruplex arrangement [88, 89]. All four DNA strands are present in analogous fashion and, shows three linking trinucleotide coils placed on the outer core of the quadruplex and look like as propeller-like arrangement.

Docking studies of compound 82 in complex with human telomeric G-quadruplex DNA shows (dock score = −7.2503 kcal/mol) that planer acridine ring loaded on the G terminal and align in between DG: 8 and DG: 9 where it exhibit pi–pi stacked interaction with DG: 8 (see Fig. 16). Here central cationic acridine ring nitrogen atom covering the central polarized carbonyl channel of negative electrostatic potential that runs through the stack of G quartets and exhibited a contact with potassium ion (K, A: 46) at topological distance of 3.34 Å. At this point, one propanamide substituent at 6 position of acridine ring orient in between DG: 8 and DG: 9 where it exhibit one hydrogen bonding interactions with water molecule HOH: 1050 (2.75 Å), one carbon hydrogen bond with DG: 8 (2.85 Å) while another propanamide substituent at 2 position stacked on DG: 20 where it form a contact of hydrogen bond with DG: 14 (2.76 Å) and carbon hydrogen bonding with DG: 20 (2.57 Å).

The central acridine ring is stabilized by pi–pi stacked interactions with DG: 8 (5.99 Å) and DG: 14 (5.61 Å). Moreover acridine ring exhibit one more pi–pi stacked contact with DG: 14 (5.13 Å) to concrete stabilization of acridine ring and human telomeric G-quadruplex complex (see Fig. 16). The binding site itself is extremely disturbed as it appears exterior to the load of three G quartets which is connected to the channels generated from the phosphodiester backbones. (Docking results for Epirubicin and compound 82 in the human telomeric G DNA is given in Additional file 1: Table S2).
Fig. 15 Depiction of 2D interaction of pose 3 of compound 82 and 3D presentation of superimposed alignment of pose 3 of compound 82 with BIBR1532.
Further amide carbonyl oxygen of propanamide substituent exhibit metal acceptor contact with potassium ion (K: 26) at interatomic distance of 3.21 Å which disclose close contact of 2-propanamide substituent than acridine nitrogen. Third sub-stunt 4-aminophenyl at 9th position of acridine ring align in between DG: 14 and DG: 10 where phenyl ring form pi–alkyl contact with DG: 14 (5.39 Å) while amino substituent exhibit hydrogen bonding again with DG: 14 (2.09 Å). Here it is important to note that amino substituent align very closely near DG: 12 DNA base (see Fig. 17).

Furthermore, we docked Epirubicin against human telomeric G-quadruplex and analyzed its binding orientation and modes of interaction to compare the docking findings of compound 82. The docking analysis reveals binding of Epirubicin with human telomeric G-quadruplex DNA which yielded negative docking score of −6.1933 kcal/mol. Binding at DA: 13, DG: 14, 55, Potassium (K: 26) and 11 site is stabilized by polar and hydrophobic contacts with Epirubicin. Methoxy group substituted on 1 position align in between DG: 15 and DG: 14 where ring A, B and C form pi–pi hydrophobic contact with DG: 14 successively placed at interatomic distance of 3.70, 3.78 and 6.05 Å (see Fig. 18). In addition to this, ring A exhibit carbon hydrogen bond with 11, 5-dione moiety form metal-acceptor contact with potassium (K: 26) at 3.47 Å and 6-hydroxy substituent exert hydrogen bonding interaction with DG: 14 (2.77 Å). Moreover, ethereal linkage oxygen at 10th position exhibit covalent bond with 55, 5-hydroxy substituent form hydrogen bonding contact with 55 (2.77 Å) while 4-amino group display hydrogen bonding contact with DA: 13 at 2.20 Å.

There is distinct difference in binding of compound 82 and Epirubicin. The compound 82 was directed on the G terminal and orient in between DG: 8 and DG: 9 where it exhibit pi–pi stacked interaction with DG: 8 whereas another pyrrolidine end orient on the DG: 20 and slightly inclined near DG: 21.

Moreover, central cationic acridine ring nitrogen exhibited a metal-acceptor contact with potassium ion. In comparison to this, methoxy terminal of Epirubicin align in between DG: 15 and DG: 14 where ring A, B and C form pi–pi hydrophobic contact with DG: 14 with successive placement at interatomic distance of 3.70, 3.78 and 6.05 Å while orientation of pyran ring of Epirubicin spread over DG: 4, DG: 8 and DG: 9 (see Fig. 19).

Conclusions
In our study, telomerase reverse transcriptase (TERT) and human telomeric G4 DNA were used as therapeutic targets to build a QSAR model. We used an 82-compound dataset that includes a variety of structures such as the acridine ring, triazoles ring, and pyrimidine ring. QSARINS 2.2 software is used to create a QSAR model using PyDescriptor and the GA-MLR method. The derived QSAR model showed high external and internal predictive ability. Compound 82 attained diverse conformations inside FVYL motif/pocket of TERT and key interaction involves hydrogen bonding and hydrophobic interactions with the involvement of water molecule and hydrophobic residues MET482, ILE497, LEU554, ILE550, MET483, and ARG486. With the bottomless analysis of docking outcomes compared with crystallized ligand BIBR1532, the interaction amid compound 82 and amino acid residues were similar to those of BIBR1532.

Moreover, molecular docking revealed that compound 82 and Epirubicin bind to G4 DNA’s external grooves and loop, forming pi–pi stacking hydrophobic and hydrogen bonding associations with G DNA bases. In binding interaction of both ligands, DG: 14 and potassium ion are
the common DNA base and ion, with which both ligand exhibit pi–pi stacking hydrophobic and metal acceptor interactions. In terms of orientation, both ligands terminals align by their own way on different grooves and loops which may leads to specificity of interactions with G4 DNA that may in turn responsible for their different binding individualities. The analysis divulges that both compound 82 and Epirubicin act as potential inhibitors of human telomeric DNA where both ligands stabilizes G4 DNA in presence of K\(^+\) ion which was plausibly affect association of telomerase with telomeric DNA therefore exhibiting cell induced apoptosis as an alternate mechanism to damage on binding to quadruplexG4 DNA.

Furthermore, we obtained promising results in this study because compound 82 binds to the FVYL motif/pocket of TERT and adopts the same conformation as the clinical trial agent, BIBR1532. Beside this, acridine ring was established and investigated to be bind with human telomeric G DNA. In light of this rationale, docked analysis of the most active compound 82 in human telomeric G DNA revealed that compound 82 binds to Human telomeric G DNA through hydrogen bonding and hydrophobic interactions. While compound 82 has a different binding conformation than Epirubicin, both exhibit hydrogen bonding and hydrophobic interactions. Based on the docking findings, it is clear that the most active compound 82 binds to both TERT and human telomeric G DNA through similar interactions. As a result, it is reasonable to conclude that compound 82 has potent anticancer activity via dual inhibition of TERT and human telomeric G DNA. The outcome of QSAR and molecular docking study may possibly support many researchers to put forward novel and dual inhibitors of TERT and Human telomeric G DNA with higher anticancer activities. Therefore, they may possibly decrease time, cost and even the accessibility of the laboratories equipped to bring out the synthesis and tests.

Fig. 17 Display of 2D and 3D interaction of compound 82 with human telomeric G-quadruplex DNA

Fig. 18 3D depiction of superimposed structures of compound 82 (green) with Epirubicin (yellow)
Fig. 19 Depiction of Epirubicin and human telomeric G-quadruplex DNA Interactions
Abbreviations
Q SAR: Quantitative structure activity relationship; GA-MLR: Genetic algorithm multiple linear equation; TERT: Telomerase reverse transcriptase; RNP: RNA–protein complex; MDR: Multidrug resistance; OECD: Organization for Economic Corporation and Development; OFS: Objective feature selection; SFS: Subjective feature selection; CCC: Concordance correlation coefficient.

Supplementary Information
The online version contains supplementary material available at https://doi.org/10.1186/s43094-021-00380-7.

Additional file 1: Table S1. Docking results for Compound 82, BABR1532 and Epirubicin into the TERT; Table S2. Docking results for Epirubicin and compound 82 in the human Telomeric G DNA; Table S3. Depiction of QSAR results along with their experimental and predicted EC50 values for divided set model; Table S4. Depiction of QSAR results along with their experimental and predicted EC50 values for full set model; Table S5. Presentation of Serial number, ChemBL id, Smiles and Pec50 value of 82 compounds; Table S6. Calculated Descriptor values used to build QSAR model.

Acknowledgements
We are thankful to Shri Yogendraji R. Gode, President, Dr. Rajendra Gode Institute of Pharmacy, University Mardi Road, Amravati, for providing extended support and Research Amenities during entire course of Research work. Authors of this manuscript are thankful to Dr. Paola Gramatica for providing a copy of QSARINS 2.24 software.

Authors’ contributions
Conceptualization: RDJ and RLB; Formal analysis: CMJ, MSJ and JVM, Data curation: CMJ. Authors of this manuscript are thankful to Dr. Paola Gramatica for providing a copy of QSARINS 2.24 software.

Funding
Not applicable.

Availability of data and materials
Data and material are available on request. A dataset of structurally varied 82 nitrogen ring containing heterocycles experimentally tested against Human Telomeric G DNA and poly(ADP-ribose) polymerase-1 activity sensing based on the high fluorescence selectivity and sensitivity of TOTO-1 towards G bases in single-stranded DNA and poly(ADP-ribose). Chem Sci 10(3):3706–3714. https://doi.org/10.1039/c9sc05770b.

Declarations
Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
Authors declares no competing interest.

Author details
1Department of Medicinal Chemistry, Dr. Rajendra Gode Institute of Pharmacy, University-Mardi Road, Amravati, Maharashtra 444602, India. 2Department of Pharmaceutical Chemistry, SGSPS Institute of Pharmacy, Akola, Maharashtra 444001, India. 3Department of Business Administration, Faculty of Business Administration, Economics and Political Science, British University in Egypt, Elsherouk City, Cairo 11837, Egypt. 4Department of Medicinal Chemistry, Dr. Rajendra Gode College of Pharmacy, University-Mardi Road, Amravati, Maharashtra 444602, India.

Received: 21 May 2021   Accepted: 16 November 2021   Published online: 27 November 2021

References
1. Shay JW, Bacchetti S (1997) A survey of telomerase activity in human cancer. Eur J Cancer 33(5):787–791. https://doi.org/10.1016/S0959-8049(97)00062-2
2. Chen Z, Monia BP, Corey DR (2002) Telomerase inhibition, telomere shortening, and decreased cell proliferation by cell permeable 2′-O-methoxyethyl oligonucleotides. J Med Chem 45(25):5423–5425. https://doi.org/10.1021/jm025563v
3. Bodnar AG (1998) Extension of life-span by introduction of telomerase into normal human cells. Science 279(5349):349–352. https://doi.org/10.1126/science.279.5349.349
4. Blackburn EH, Collins K (2010) Telomerase: an RNP enzyme synthesizes DNA. Cold Spring Harbor Perspect Biol 3(5):a003558-a. https://doi.org/10.1101/cshperspect.a003558
5. Shay JW (2016) Role of telomerases and telomerase in aging and cancer. Cancer Discov 6(6):584–593. https://doi.org/10.1158/2159-8290.Cd-16-0862
6. Mada Y, Masutomi K (2015) Telomerase reverse transcriptase moonlights: therapeutic targets beyond telomerase. Cancer Sci 106(11):1486–1492. https://doi.org/10.1111/cas.12806
7. Ruden M, Puri N (2013) Novel anticancer therapeutics targeting telomerase. Cancer Treatment Rev 39(5):444–456. https://doi.org/10.1016/j.ctrv.2012.05.007
8. Barma DK, Elayadi A, Falck JR, Corey DR (2003) Inhibition of telomerase by BIBR 1532 and related analogues. Biorg Med Chem Lett 13(7):1333–1336. https://doi.org/10.1016/s0960-894x(03)00101-x
9. El-Daly H, Kull M, Zimmermann S, Pantic M, Waller CF, Martens UIM (2005) Selective cytotoxicity and telomere damage in leukemia cells using the telomerase inhibitor BIBR1532. Blood 105(4):1742–1749. https://doi.org/10.1182/blood-2003-12-4322
10. Yang H, Fu F, Li W, Wei W, Zhang Y, Liu S (2019) Telomerase and poly(ADP-ribose) polymerase-1 activity sensing based on the high fluorescence selectivity and sensitivity of TOTO-1 towards G bases in single-stranded DNA and poly(ADP-ribose). Chem Sci 10(3):3706–3714. https://doi.org/10.1039/c9sc05770b
11. Saygin C, Carraway HE (2021) Current and emerging strategies for management of myelodysplastic syndromes. Blood Rev 48:100791. https://doi.org/10.1016/j.blre.2020.100791
12. Hernandez-Meza G, Felden J, Gonzalez-Kozlova EE, Garcia-Lezana T, Peix J, Portela A et al (2021) DNA methylation profiling of human hepatocarcinogenesis. Hepatology 74(1):183–199. https://doi.org/10.1002/hep.31659
13. Carloni L-E, Wechselberger R, De Vlijder T (2021) Characterization of in vitro G-quadruplex formation of imetelstat telomerase inhibitor. Nucleic Acid Ther. https://doi.org/10.1089/nat.2020.0918
14. Cheng T, Li Q, Zhou Z, Wang Y, Bryant SH (2012) Structure-based virtual screening for drug discovery: a problem-centric review. AAPS J 14(1):133–141. https://doi.org/10.1002/ajps.22932
15. Sekeroğlu ZA, Sekeroğlu Y, Kucuk N (2020) Effects of reverse transcriptase inhibitors on proliferation, apoptosis, and migration in breast carcinoma cells. Int J Toxicol 40(1):52–61. https://doi.org/10.1177/1091582019961498
16. Hastie ND, Dempster M, Dunlop MG, Thompson AM, Green DK, Allshire RC (1990) Telomere reduction in human colorectal carcinoma and with ageing. Nature 346(6387):866–868. https://doi.org/10.1038/346664a0
17. Counter CM, Avilion AA, LeFeuvre CE, Stewart NG, Greider CW, Harley CB (1992) Telomere shortening associated with chromosome instability is arrested in immortal cells which express telomerase activity. EMBO J 11(5):1921–1929. https://doi.org/10.1002/j.1460-2075.1992.tb05245.x
18. Kim N, Piatyzek M, Prowse K, Harley C, West M, Ho P (1994) Specific association of human telomerase activity with immortal cells and cancer. Science 266(5193):2011–2015. https://doi.org/10.1126/science.7605428
19. Counter CM, Botelho FM, Wang P, Harley CB, Bacchetti S (1994) Stabilization of short telomeres and telomerase activity accompany immortalization of Epstein–Barr virus-transformed human B lymphocytes. J Virol 68(5):3410–3414. https://doi.org/10.1128/jvi.68.5.3410-3414.1994
20. Karlseder J (1999) p53- and ATM-dependent apoptosis induced by telomeres lacking TRF2. Science 283(5406):1321–1325. https://doi.org/10.1126/science.283.5406.1321
