Design, synthesis, screening and molecular modeling study of a new series of Ros-1 receptor tyrosine kinase inhibitors

Ibrahim M. El-Deeb, b Byung Sun Park, c Su Jin Jung, a Kyung Ho Yoo, a
Chang-Hyun Oh, a Seung Joo Cho, d,e Dong Keun Han, a Jae Yeol Lee, c and So Ha Lee a,*

a Life Sciences Research Division, Korea Institute of Science and Technology, P.O. Box 131, Cheongryang, Seoul 130-650, Republic of Korea
b Department of Biomolecular Science, University of Science and Technology, 113 Gwahangno, Yuseong-gu, Daejeon 305-333, Korea
c Research Institute for Basic Sciences and Department of Chemistry, College of Science, Kyung Hee University, Seoul 130-701, Korea
d Research Center for Resistant Cells, Chosun University, Gwangju 501-759, Republic of Korea
e Department of Cellular and Molecular Medicine, College of Medicine, Chosun University, 375 Seosuk-dong, Dong-gu Gwangju 501-759, Republic of Korea

Abstract – A series of rationally designed Ros-1 tyrosine kinase inhibitors was synthesized and screened. Compound 12b has showed good potency with IC50 value of 209 nM, which is comparable with that of the reference lead compound 1. Molecular modeling studies have been performed, i.e., a homology model for Ros1 was built, and the screened inhibitors were docked into its major identified binding site. The docked poses along with the activity data have revealed a group of the essential features for activity. Overall, simplification of the lead compound 1 into compound 12b kept the activity, while facilitated the synthetic advantages. A molecular interaction model for Ros1 kinase and inhibitors has been proposed.

The traditional cancer treatments using classical cytotoxic drugs cannot distinguish between normal and cancerous cells, leading to the production of serious side effects, and limiting their long term use in most of cases.1,2 These common drawbacks of cytotoxic agents have encouraged scientists to search for more selective and less toxic therapies for different types of cancers. After the new advances in the science of genomics and the characterization of the human genome, the incidence of many types of cancers has been correlated with mutations in genes that confer growth advantage.3-5 Among these genes are those encoding for protein kinases, a group of important players in the process of signal transduction that controls normal cell growth and proliferation.6,7 Mutations at such genes result usually in the production of the encoded kinases in abnormal high levels and/or constitutive uncontrolled activity, leading to uncontrolled cell proliferation and consequently cancer.8-10 Glioblastoma multiforme is the most advanced astrocytic neoplasm, and is one of the most aggressive human cancers with a median survival of less than one year. Despite decades of therapeutic research, effective chemotherapeutic treatment for high grade astrocytomas is not yet available, and patient care ultimately focuses on palliative management.11-13 In a survey of 45 different human tumor cell lines, the tyrosine kinase Ros1 was found to be expressed in glioblastoma-derived cell lines at high levels, while not expressed at all or expressed minimally in the remaining cell lines.14 Ros1 is a receptor tyrosine kinase that is homologous to the Drosophila sevenless tyrosine kinase receptor.15-18 It's encoded by Ros1 gene which is located at the chromosome 6 region 6q16→6q22 which is involved in non-random chromosomal arrangements in specific neoplasias.16 A microdeletion at 6q21 results in the fusion of FIG, a gene coding for a Golgi apparatus-associated protein, to the kinase domain of the proto-oncogene Ros1.17,18 The fused protein product FIG-ROS is a potent oncogene, and its transforming potential resides in its ability to interact with and become localized to the Golgi apparatus.17 The ectopic expression of Ros1 receptor protein has been reported mainly in meningiomas and astrocytomas (25% of low grade and 30% of malignant glioma tumors) suggesting a key role for Ros1 in these CNS malignancies.17,18 Hence the targeting of the tyrosine kinase Ros1 could be a useful strategy for treatment of astrocytic neoplasms.

In a previous work reported by our group,20 a new potent and highly selective Ros1 kinase inhibitor 1 was...
synthesized and evaluated (Figure 1). The synthesized compound 1 was screened over 45 different kinases, and showed a high selectivity towards Ros1 kinase, associated with good potency (IC_{50} = 199 nM). In this study, a series of rationally designed compounds have been synthesized and screened over Ros1 kinase in order to study the structure activity relationship for this new class of inhibitors. A molecular modeling study was also made in order to have a better molecular insight on the mode of enzyme inhibition of these inhibitors, where a homology model for the enzyme Ros1 was built, and the tested compounds were docked into the potential binding site of the built homolog.

The synthesis work started with the preparation of the key ester, methyl 3-methoxy-5-methylbenzoate (5) as illustrated in Scheme 1. In the first step, the sodium salt of ethyl 2-hydroxy-4-oxopent-2-enooate (2) was prepared according to literature procedure\(^2\), through the condensation of diethyl oxalate with acetone in the presence of sodium ethoxide in absolute ethanol. The resulted salt 2 was then cyclized into Claisen furan derivative 3 by heating in 50% acetic acid followed by acidification with sulfuric acid.\(^6\) The resulted Claisen compound underwent rearrangement and aromatization into 3-hydroxy-5-methylbenzoic acid (4) within less than one hour by heating with magnesium oxide in boiling water, followed by acidification with hydrochloric acid to precipitate the product.\(^22\)

Methyl esterification and O-methylation of the resulted phenolic acid 4 were achieved in a single step and in a high yield (94%) to give compound 5 through a little modification of the literature procedure,\(^23\) where the acid 4 was refluxed with excess potassium carbonate and iodomethane in acetone, and in the presence of a catalytic amount of DMAP.

In Scheme 2, the benzoate ester 5 underwent a nucleophilic attack at its carboxylic carbon by the activated methylene group of 2-chloro-4-methylpyrimidine. The activation of this methyl group into an active methylene was achieved by dropwise addition of lithium bis(trimethylsilyl)amide (LHMDS) in dry THF at room temperature. The resulted tautomeric α,β-unsaturated ketone 6 was then converted to the required pyrazole derivative 7 through two successive steps. In the first step, compound 6 was heated with excess N,N-dimethylformamide dimethylacetal for 12 hours, and the resulted product was taken to the next step without further purification, where it was cyclized with hydrazine monohydrate in absolute ethanol into the pyrazole derivative 7. The reaction of the resulted pyrazole 7 with iodoacetonitrile in the presence of excess potassium carbonate produced two different regiosomers; a major isomer 8 with R\(_f\) value of 0.74 (EtOAc), and a minor isomer 9 with R\(_f\) value of 0.84 (EtOAc). A mixture of these two isomers was taken to the next step without separation, where it underwent a Suzuki coupling with a series of arylboronic acids, in the presence of dichlorobis(triphenylphosphine) Pd(II) and potassium carbonate, in a mixed solvent of THF and water in a (4:1) ratio. These series of Suzuki coupling reactions produced two isomers in every reaction 10a-j and 11a-j, which were separated in a pure form by preparative TLC. In all of these reactions, the 1H-pyrazole isomer (isomer 10) was the product with the lower R\(_f\) value while the 2H-pyrazle isomer (isomer 11) was the product with the higher R\(_f\). This was proved by the 2D-NOESY NMR spectrum of compounds 10h and 11h as a model for the whole series. The absence of any cross peak between the acetonitrile –CH\(_2\)- and any of the aromatic protons of 3-methoxy-5-methylphenyl group in the 2D-NOESY NMR spectrum of compound 10h while the presence of such peaks in the 2D-NOESY NMR spectrum of compound 11h has confirmed the assigned structures. The final hydroxyl derivatives 12a-j and 13a-j were obtained by demethylation of the methoxy group of the corresponding methoxy compound using 10 equivalents of borontrifluoride-dimethylsulfide complex in dichloromethane.

In Scheme 3, typical procedures to those used in Scheme 2 have been applied. The only difference was in the starting amine, where in Scheme 3, 2-chloro-5-methylpyridine was used instead of 2-chloro-4-methylpyridine in Scheme 2.
Scheme 2. Reaction conditions and yields: (i) LHMDS, THF, N₂, rt, 18 h, 72%; (ii) DMF-DMA, reflux, 12 h; (iii) hydrazine hydrate, abs. EtOH, rt, 2 h, 81%; (iv) K₂CO₃, iodoacetonitrile, acetone, reflux, 4 h, 92%; (v) arylboronic acid, Pd(PPh₃)₂Cl₂, Na₂CO₃, N₂, THF/H₂O (4:1), 70 °C, 12 h; (vi) BF₃·S(CH₃)₂, dichloromethane, N₂, rt, 24 h.

Scheme 3. Reaction conditions and yields: (i) LHMDS, THF, N₂, rt, 18 h, 55%; (ii) DMF-DMA, reflux, 12 h; (iii) hydrazine hydrate, abs. EtOH, rt, 4 h, 62%; (iv) K₂CO₃, iodoacetonitrile, acetone, reflux, 4 h, 80%; (v) arylboronic acid, Pd(PPh₃)₂Cl₂, Na₂CO₃, N₂, THF/H₂O (4:1), 70 °C, 12 h; (vi) BF₃·S(CH₃)₂, dichloromethane, N₂, rt, 24 h.

Kinase assays were performed at Reaction Biology Corporation using the “HotSpot” assay platform. In the initial screening step, the 14 selected compounds showed in Figure 2 were tested over Ros1 kinase at a single dose of 20 µM, and the reaction was carried out at 10 µM ATP concentration. At this concentration, compound 12b showed a significant inhibition of 96 % for the activity of Ros1 kinase, while the inhibition exerted by all of the other tested compounds was below 30%. Accordingly, compound 12b was further tested in a 10-dose IC₅₀ mode with 3 fold serial dilutions starting at 20 µM. Staurosperine²⁴ was used as a control compound in a 10-dose IC₅₀ mode with 5 fold serial dilutions starting at 20 µM, and the reaction was carried out at 10 µM ATP concentration. The new compound 12b has showed an IC₅₀ value of 209 nM which is comparable with that of the reference compound 1 (Figure 2).

In order to apply a molecular modeling study for the tested derivatives and correlate between their activities
and their binding mode to the enzyme, a homology model for Ros1 kinase was built using MOE software. The last 1347 amino acids of the total 2347 residues of Ros1 which are predicted to contain the kinase domain of the enzyme were used to build the homolog model over the 3D crystal structure of the insulin-like growth factor i receptor protein (PDB code 1P4O).

Table 1. Synthesized derivatives from Scheme 2 and 3.

<table>
<thead>
<tr>
<th>Compound</th>
<th>X</th>
<th>R</th>
<th>Compound</th>
<th>X</th>
<th>R</th>
<th>Compound</th>
<th>X</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>10a</td>
<td>H</td>
<td>H</td>
<td>11i</td>
<td>H</td>
<td>4-N(CH₃)₂</td>
<td>13g</td>
<td>H</td>
<td>3-NH-Ac</td>
</tr>
<tr>
<td>10b</td>
<td>N</td>
<td>H</td>
<td>11j</td>
<td>H</td>
<td>4-OPh</td>
<td>13h</td>
<td>H</td>
<td>4-CN</td>
</tr>
<tr>
<td>10c</td>
<td>H</td>
<td>2-Ac</td>
<td>12a</td>
<td>H</td>
<td>H</td>
<td>13i</td>
<td>H</td>
<td>4-N(CH₃)₂</td>
</tr>
<tr>
<td>10d</td>
<td>H</td>
<td>3-Ac</td>
<td>12b</td>
<td>N</td>
<td>H</td>
<td>13j</td>
<td>H</td>
<td>4-OPh</td>
</tr>
<tr>
<td>10e</td>
<td>H</td>
<td>4-Ac</td>
<td>12c</td>
<td>H</td>
<td>2-Ac</td>
<td>18a</td>
<td>N</td>
<td>H</td>
</tr>
<tr>
<td>10f</td>
<td>H</td>
<td>2-NH-Ac</td>
<td>12d</td>
<td>H</td>
<td>3-Ac</td>
<td>18b</td>
<td>H</td>
<td>2-Ac</td>
</tr>
<tr>
<td>10g</td>
<td>H</td>
<td>3-NH-Ac</td>
<td>12e</td>
<td>H</td>
<td>4-Ac</td>
<td>18c</td>
<td>H</td>
<td>3-Ac</td>
</tr>
<tr>
<td>10h</td>
<td>H</td>
<td>4-CN</td>
<td>12f</td>
<td>H</td>
<td>2-NH-Ac</td>
<td>19a</td>
<td>N</td>
<td>H</td>
</tr>
<tr>
<td>10i</td>
<td>H</td>
<td>4-N(CH₃)₂</td>
<td>12g</td>
<td>H</td>
<td>3-NH-Ac</td>
<td>19b</td>
<td>H</td>
<td>2-Ac</td>
</tr>
<tr>
<td>10j</td>
<td>H</td>
<td>4-OPh</td>
<td>12h</td>
<td>H</td>
<td>4-CN</td>
<td>19c</td>
<td>H</td>
<td>3-Ac</td>
</tr>
<tr>
<td>11a</td>
<td>H</td>
<td>H</td>
<td>12i</td>
<td>H</td>
<td>4-N(CH₃)₂</td>
<td>20a</td>
<td>N</td>
<td>H</td>
</tr>
<tr>
<td>11b</td>
<td>N</td>
<td>H</td>
<td>12j</td>
<td>H</td>
<td>4-OPh</td>
<td>20b</td>
<td>H</td>
<td>2-Ac</td>
</tr>
<tr>
<td>11c</td>
<td>H</td>
<td>2-Ac</td>
<td>13a</td>
<td>H</td>
<td>H</td>
<td>20c</td>
<td>H</td>
<td>3-Ac</td>
</tr>
<tr>
<td>11d</td>
<td>H</td>
<td>3-Ac</td>
<td>13b</td>
<td>N</td>
<td>H</td>
<td>21a</td>
<td>N</td>
<td>H</td>
</tr>
<tr>
<td>11e</td>
<td>H</td>
<td>4-Ac</td>
<td>13c</td>
<td>H</td>
<td>2-Ac</td>
<td>21b</td>
<td>H</td>
<td>2-Ac</td>
</tr>
<tr>
<td>11f</td>
<td>H</td>
<td>2-NH-Ac</td>
<td>13d</td>
<td>H</td>
<td>3-Ac</td>
<td>21c</td>
<td>H</td>
<td>3-Ac</td>
</tr>
<tr>
<td>11g</td>
<td>H</td>
<td>3-NH-Ac</td>
<td>13e</td>
<td>H</td>
<td>4-Ac</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11h</td>
<td>H</td>
<td>4-CN</td>
<td>13f</td>
<td>H</td>
<td>2-NH-Ac</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A good homology between the two proteins was predicted with E value of 1.4e⁻⁵⁷. The sequences of the two proteins were aligned and the built homolog was relaxed applying force filed Amber99 calculations. The site identifier has identified a major binding site made up of 28 residues and having a size of 172 Å. The docking results for the reference compound 1 and compound 12b have revealed that both of the two compounds have the same pose that occupies the same space inside the binding pocket (Figure 3).

Figure 2. % of Ros1 enzyme inhibition exerted by single dose concentration of 20 µM (upper panel). Dose-activity curve for compound 12b on Ros1 kinase (lower panel).

Figure 3. Similar docking poses for the reference compound (yellow) and compound 12b (cyan)
In this binding pose, the compound is fit into the receptor site through four hydrogen bonds. Two hydrogen bonds between the acetonitrile nitrogen and leucine 1105 and alanine 1106, one hydrogen bond between the phenolic OH and leucine 1122, and a fourth hydrogen bond between the pyridyl nitrogen and lysine 1111 (Figure 4).28

Figure 4. Hydrogen bonding interactions with compound 12b.

By investigating the binding interactions of the reference compound 1 (Figure 3), it was found that the pyrimidinyl nitrogens don't share in any hydrogen bonding in the receptor. In addition, the 1-aminopropan-2-ol group doesn't support any significant binding for the compound in the receptor site. These observations were consistent with the screening results for the two active compounds 1 and 12b, where they showed nearly the same IC₅₀ values for enzyme inhibition, and this means that the structural differences between them (the extra pyrimidine nitrogen and the 1-aminopropan-2-ol side chain) has no significant effect on compound activity.

Figure 5. Proposed clash between compound 13g and the receptor surface when occupy the same pose of compound 12b.

The docking results for the remaining screened compounds have showed different binding modes to that of compound 12b, and in the majority of them, the compounds were loosely bound to the receptor site with only one or maximum two hydrogen bonds. It showed also that these derivatives failed to occupy the same binding pose of compound 12b, due to either the lack of hydrogen bonding capability with lysine 1111, or clashing with the receptor surface caused by the excessive bulkiness of the substituent at the phenyl ring (Figure 5).

It became clear from these results that the terminal pyridyl moiety in compounds 1 and 12b is an essential feature for activity and is highly conserved, and that the replacement of this moiety with unsubstituted or substituted phenyl group, results in a great loss of activity. The simple deviation in the position of this terminal pyridyl group in space caused by the shift in its position from the meta-position (compounds in Scheme 2) to the para-position (compounds in Scheme 3) relative to the central pyridine, results in a great loss of activity too, and this means that the orientation of the pyridyl group is also highly conserved (Figure 6).

Figure 6. Binding mode of compound 20a (Scheme 3)

From Figure 5, we can also conclude that the 2H-pyrazole isomers of this series can't fit well in the intended pose at the receptor site, because of the loss of the two hydrogen bonds formed with the acetonitrile nitrogen, in addition to the possible clash between the acetonitrile group at this position and the receptor surface. For a more clear comparison between the different binding modes of all of the screened compounds, the binding complexes were aligned together, and a barcode presentation mode was established to view all the interacting residues against the docked compounds, and the characteristic bindings with each of them. As showed in Figure 7, a unique common binding feature in compounds 1 and 12b is the binding with leucine 1122. This characteristic binding
was not observed at any of the other derivatives. This figure reveals also that these two active compounds (1 and 12b) have typical binding profiles, which accounts for their identical inhibitory activity.

Figure 7. Barcode presentation mode for the characteristic interactions of the screened compounds at the receptor site (the order from top to bottom: 1, 12b, 12a, 12c, 12d, 12e, 12f, 13f, 13g, 12h, 12i, 12j, 20a, 20b, 20c.)

In conclusion, the molecular modeling study has showed that the more simple structure of compound 12b contains three essential features for activity (terminal pyridyl, \(\text{NH}\)-acetonitrile and phenolic hydroxyl), and that this structure is highly conserved in most of its parts. A possible site that is yet viable for more modifications is the methyl group at the 3-hydroxy-5-methylphenyl ring. The extension of structure originating from this point may produce derivatives with extending binding features and possible enhancement in activity.

Acknowledgment

This research was supported by Korea Institute of Science and Technology and Bio-strategy Technology Research Program through Korea Institute of Industrial Technology Evaluation and Planning funded by Ministry of Knowledge Economy. We also appreciate to Dr. Sean W. Deacon and Dr. Haiching Ma from Reaction Biology Corporation for kinase screening.

References and Notes

28. The showed residues' numbers refer to the amino acid sequence in the built model which is 1000 less than the actual number in the whole original protein.