Terminating Oncogenic Signaling: Rational Design of a Molecular Inhibitor distorting and inactivating GTP bound to Oncogenic Mutated RAS Proteins.

Sushant Shekhar

LinkedIn: [www.linkedin.com/in/sushant-shekhar-462808323]

November 2024

1 Abstract

Oncogenic mutations in RAS proteins, particularly in GTP-binding domains, drive uncontrolled cellular proliferation and tumorigenesis and cause mutated Ras driven cancers like pancreatic cancer, colorectal cancer, lung adenocarcinoma, melanoma, thyroid cancer, bladder cancer, liver cancer, and endometrial cancer. These mutations hinder GTP hydrolysis because the mutated Ras becomes insensitive to GAP(GTPase-activating protein), resulting in persistent activation of downstream signaling pathways. Here, we present the design and characterization of a novel drug designed to distort and inactivate GTP bound to mutated RAS proteins. The drug is designed based on the following data and approach:

- 1. Ligands like MG GDP, ACT and GOL have high to good affinity towards mutated Ras (The data has been collected from RCSB Protein Data Bank).
- 2. We have used Tert-butyl phenylsulfate CC(C)(C)OS(=O)(O)OC1 = CC = CC = C1 as one of the fragments of the drug molecule because tert-butyl phenylsulfate (CC(C)(C)OS(=O)(O)OC1=CC=CC=C1) exhibits a high affinity for GTP (Tested using Swissdock), it can disrupt Ras signaling by interfering with the GTP bound to Ras. Ras typically becomes active upon binding GTP, and its activation is essential for transmitting signals that promote cell proliferation. In the case of mutated Ras proteins that lack intrinsic GTPase activity, GTP remains bound to Ras, leading to continuous signaling. By directly interacting with the GTP molecule and mutated Ras, tert-butyl phenylsulfate in

combination with other used fragments can defunctionalize GTP and the whole mutated Ras proteins, thereby inhibiting its activation and terminating the uncontrolled signaling.

- 3. We have introduced additional electron-donating group (EDG) to the molecule. We will add electron-donating group (EDG) to the molecule. Suitable EDGs, such as -NH2 and -OCH3 at meta position to the Oc1cccc1 portion of CC(C)(C)OS(=O)(O)OC1 = CC = CC = C1 molecule to enhance the nucleophilic strength of the tert-Butyl hydrogen sulfate fragment i.e CC(C)(C)OS(=O)(O)O so that CC(C)(C)OS(=O)(O)OC1 = CC = CC = C1 will interact with GTP bound to mutated Ras with better affinity defunctionalizing the GTP.
- 4. We have introduced a suitable protective group for the drug molecule so that the protective group must satisfy the criterias like selective Reactivity(It should be protecting the molecule from reacting with non-mutated RAS or other nearby proteins), structural Stability(preventing the intramolecular reactions that could destabilize the molecule), targeted activation (The protective group should specially allow specially CC(C)(C)OS(=O)(O)OC1 = CC = CC = C1 to interact with GTP bound to mutated RAS proteins). A benzyl-based protective group with additional targeting features and appropriate positioning is used, as it is chemically stable and can incorporate targeting moieties (e.g., ester linkages or functional groups that are cleaved in the presence of specific mutations).

Based on the above analysis we got the fragments for our drug molecule. They are MG GDP, ACT , GOL , tert-butyl phenylsulfate CC(C)(C)OS(=O)(O)OC1=CC=CC=CC, EDGs such as -NH2 and -OCH3 groups and a Benzyl-based protective group. We will assamble these fragments to design the drug in such a way that the drug molecule will target mutated Ras proteins like 7R0M, 7R0N , 7R0Q , 8BE3 , 7VVG , 8VM2 , 6ZIZ , 7T47 selectively only interacting with the GTP (bound with these mutants in active state) causing the GTP and the mutant Ras to loose its functionality and hence terminating the uncontrolloed downstream cell proliferation signaling.

Through computational modeling, we demonstrate the ligand's capability to interact with GTP-RAS altering, distorting and defunctionalizing it, thereby terminating oncogenic signaling pathways. This study provides a mechanistic foundation for targeting mutated RAS proteins using ligand-based approaches, highlighting its potential as a therapeutic strategy against RAS-driven cancers.

2 Designing the Drug Molecule:

Step 1: We will first assemble following three fragments (GDP, Acetate and Glycerol) with our main fragment CC(C)(C)OS(=O)(O)OC1 = CC = CC = C1 using python RDKit library, specifically Chem for molecular manipulation and AllChem for advanced operations.

```
1. c1nc2c(n1[C@H]3[C@@H]([C@H](C@H](O3)CO[P@](=O)(O)OP(=O)(O)O)O)N = C(NC2 = O)N GDP
```

- 2. CC(=O)[O-] Acetate
- 3. C(C(CO)O)O Glycerol

```
19 # Editable molecule
20 base_mol_edit = Chem.RWMol(base_molecule)
21
22
23 # Function to attach a fragment to a molecule
24 def attach_fragment(base_mol, fragment, base_attach_idx,
      frag_attach_idx):
      # Copy atoms from the fragment to the base molecule
25
      frag_atoms = {}
26
      for atom in fragment.GetAtoms():
27
          new_idx = base_mol.AddAtom(atom)
28
          frag_atoms[atom.GetIdx()] = new_idx
29
30
      # Add bonds from the fragment to the base molecule
      for bond in fragment.GetBonds():
32
          base_mol.AddBond(
33
              frag_atoms[bond.GetBeginAtomIdx()],
              frag_atoms[bond.GetEndAtomIdx()],
35
              bond.GetBondType(),
36
          )
37
38
      # Add the bond between the base molecule and the
39
          fragment
      base_mol.AddBond(base_attach_idx, frag_atoms[
          frag_attach_idx], Chem.BondType.SINGLE)
      return base_mol
41
42
43
_{44}| # Attach GDP (fragment 1) to the base molecule
base_attach_idx_1 = 3  # Example: Oxygen atom in the base
      molecule
46 frag_attach_idx_1 = 0 # Example: Specific atom in the
     fragment
47 base_mol_edit = attach_fragment(base_mol_edit,
     frag1_molecule, base_attach_idx_1, frag_attach_idx_1)
48
49 # Attach Acetate (fragment 2) to the updated base molecule
50 base_attach_idx_2 = 5 # Example: Another atom in the
     updated molecule
51 frag_attach_idx_2 = 0 # Example: Specific atom in the
     fragment
52 base_mol_edit = attach_fragment(base_mol_edit,
     frag2_molecule, base_attach_idx_2, frag_attach_idx_2)
_{54}| # Attach Glycerol (fragment 3) to the updated base molecule
base_attach_idx_3 = 7 # Example: Another atom in the
     updated molecule
frag_attach_idx_3 = 0 # Example: Specific atom in the
      fragment
57 base_mol_edit = attach_fragment(base_mol_edit,
```

```
frag3_molecule, base_attach_idx_3, frag_attach_idx_3)

# Finalize the molecule
final_molecule = base_mol_edit.GetMol()

# Generate and print the SMILES of the new molecule
final_smiles = Chem.MolToSmiles(final_molecule)
print("Final molecule SMILES:", final_smiles)
```

We get the following molecule:

$$CC(C)(Cc1nc2c(=O)[nH]c(N)nc2n1[C@@H]1O[C@H](CO[P@](=O)(O)OP(=O)(O)O)[C@@H](O)[C@H]1O)OS(=O)(CC(=O)[O-])(Oc1ccccc1)OC(O)CO)OCO$$

Step2: We will add a protective group to the above molecule which will shield

reactive sites to prevent undesired interaction with non-mutated RAS or surrounding proteins, stabilize the molecule and prevent intramolecular reactions, allow the specific moiety CC(C)(C)OS(=O)(O)OC1 = CC = CC = C1 to selectively interact with GTP bound to mutated RAS proteins.

For this task, we will use a sterically bulky group with selective reactivity, such as a benzyl-based protecting group. The group will selectively expose the functional region upon encountering a specific trigger (e.g., mutated RAS environment).

We will identify attachment points for the protective group. Likely sites include hydroxyl (-OH), phosphate (-PO4), or amine (-NH2) groups. for adding protective group using the following python program.

```
2 from rdkit import Chem
 from rdkit.Chem import AllChem
  # Base molecule
  base_smiles = "CC(C)(Cc1nc2c(=0)[nH]c(N)nc2n1[C@@H]10[C@H](
     CO[P@](=0)(0)OP(=0)(0)O)[C@@H](0)[C@H]10)OS(=0)(CC(=0)[O
     -])(0c1ccccc1)0C(0)C(0)C0"
  # Protective group: Benzyl group with selective activation
  protective_group_smiles = "C1=CC=CC01" # Example of a
     benzyl protecting group
 # Convert SMILES to RDKit molecule
base_molecule = Chem.MolFromSmiles(base_smiles)
13 protective_group = Chem.MolFromSmiles(
     protective_group_smiles)
14
# Editable molecule
16 base_mol_edit = Chem.RWMol(base_molecule)
17
 # Identify a potential attachment point: Look for reactive
     groups (e.g., -OH, -NH2, -PO4)
  attachment_idx = None
  for atom in base_molecule.GetAtoms():
20
      # Check for -OH groups
21
      if atom.GetSymbol() == "0" and any(neighbor.GetSymbol()
         == "H" for neighbor in atom.GetNeighbors()):
```

```
attachment_idx = atom.GetIdx()
23
      # Check for phosphate oxygen atoms
      elif atom.GetSymbol() == "O" and atom.GetDegree() > 1:
26
          attachment_idx = atom.GetIdx()
27
          break
      # Check for -NH2 groups
29
      elif atom.GetSymbol() == "N" and any(neighbor.GetSymbol
30
          () == "H" for neighbor in atom.GetNeighbors()):
          attachment_idx = atom.GetIdx()
31
          break
32
33
  if attachment_idx is None:
34
      raise ValueError("No suitable attachment point found for
35
           the protective group.")
36
37 # Add the protective group to the base molecule
38 # Append the protective group's atoms to the base molecule
39 protective_atoms = {}
40 for atom in protective_group.GetAtoms():
      new_idx = base_mol_edit.AddAtom(atom)
41
      protective_atoms[atom.GetIdx()] = new_idx
42
43
44 # Add bonds within the protective group
  for bond in protective_group.GetBonds():
45
      base_mol_edit.AddBond(
46
          protective_atoms[bond.GetBeginAtomIdx()],
47
          protective_atoms[bond.GetEndAtomIdx()],
48
          bond.GetBondType(),
49
      )
52 # Add a bond between the base molecule and the protective
# Attach at the specified attachment point
54 base_mol_edit.AddBond(attachment_idx, protective_atoms[0],
     Chem.BondType.SINGLE)
56 # Finalize the molecule
final_molecule = base_mol_edit.GetMol()
59 # Generate and print the SMILES of the new molecule
final_smiles = Chem.MolToSmiles(final_molecule)
61 print("Protected molecule SMILES:", final_smiles)
```

We will get the following molecule upon adding the benzyl protected group:

Step3: We will add electron-donating group (EDG) to the molecule. Suitable EDGs, such as -NH2 and -OCH3 at meta position to the Oc1cccc1 portion of the above molecule To enhance the nucleophilic strength of the tert-Butyl hydrogen sulfate fragment i.e CC(C)(C)OS(=O)(O)O.

We will get the following molecule:

Step4: Since MG (PDB ID) has the stongest affinity for all mutant Ras proteins including 7R0M, 7R0N, 7R0Q, 8BE3, 7VVG, 8VM2, 6ZIZ and 7T47 so we will strategically attach MG to the above molecule with the second and third carbons in ribose of the GDP fragment in the above molecule.

We will get the following molecule:

 $\begin{array}{lll} \texttt{COc6ccc} & \texttt{(OS(=0)(CC(=0)[0-])(OC(0)C(0)C0)OC(C)(C)Cc2nc1c(=0)[nH]c(N)nc1n2[C@H]4[C@@H]30[Mg]0[C@@H]3[C@@H](COP(=0)(0)0P(=0)(0)0)[0+]4C5=CC=CC05)c(N)c6 \end{array}$

- 3 Molecular Docking of the ligand with the target GTP linked mutated Ras proteins and calculated affinity data:
- 1. Swissdock Interaction result of Tert-butyl phenylsulfate i.e CC(C)(C)OS(=

$$O(O)OC1 = CC = CC = C1$$
 with GTP

Click here to view the docking result of Tert-butyl phenylsulfate with GTP

Model	Calculated affinity (kcal/mol)
1	-2.201
2	-2.138
3	-2.132
4	-2.130
5	-2.103
6	-2.077
7	-2.070
8	-2.037
9	-2.037
10	-2.033
11	-1.996
12	-1.996
13	-1.996
14	-1.994
15	-1.963
16	-1.959
17	-1.954
18	-1.943
19	-1.942
20	-1.940

2. Swissdock Interaction result of the following drug molecule

with mutated Ras 7R0M (KRasG12C):

Click here to view the docking result of drug molecule with 7R0M

Model	Calculated affinity (kcal/mol)
1	-7.313
2	-6.369
3	-6.368
4	-6.206

3. Swissdock Interaction result of the following drug molecule

with mutated Ras 7R0Q (KRasG12C):

Click here to view the docking result of drug molecule with 7R0Q

Model	Calculated affinity (kcal/mol)
1	-2.316
2	0.097

4. Swissdock Interaction result of the following drug molecule

with mutated Ras 8VM2 (NRAS Q61K):

Click here to view the docking result of drug molecule with 8VM2

Model	Calculated affinity (kcal/mol)
1	-6.203
2	-5.722
3	-5.557
4	-5.226
5	-5.164
6	-5.155
7	-5.061
8	-4.989
9	-4.288
10	-4.017
11	-3.912

The above results shows strong affinity between the drug molecule and the target mutated Ras, though we have used the limited value of exhaustiveness due to the lack of system. We need to work on the mRMSD(MOVING RMSD) value, molecular dynamics simulations, ADMET profiling for pharmacokinetics and toxicity, pharmacophore modeling, virtual screening, and binding free energy calculations. Experimental tests involve in vitro assays like SPR, ELISA, and MTT, cellular tests for viability and apoptosis, Western blotting, flow cytometry, in vivo animal models for efficacy and safety, PK/PD studies, toxicology assessments, electrochemical assays, real-time PCR, and IHC for protein expression. Combining these approaches will ensure comprehensive evaluation, guiding optimization and reducing testing needs, drug optimization, experimental validation etc. To ensure the drug molecule enters the cell, computational and experimental procedures assess its cell permeability. Computational methods will include predicting lipophilicity (LogP) and evaluating interactions with membrane transporters like P-glycoprotein (P-gp). Molecular dynamics simulations can study the drug's interaction with membranes, while ADMET profiling and models like Caco-2 cell lines or PAMPA assess permeability. Experimental tests involve using fluorescently labeled drugs in cell-based assays to measure uptake, as well as mass spectrometry or fluorescence microscopy to track intracellular concentrations. Flow cytometry can quantify cellular uptake, and transporter inhibition studies can identify efflux mechanisms. Cellular viability and behavior tests after drug exposure provide insights into the drug's effectiveness.

So far the drug molecule definitely holds the potential to interact with the mutated Ras proteins and defunctionalize the protein and GTP bound to it

4 Initial Conformational Optimization:

We have introduced several changes for initial conformational optimization of the molecule to enhance its stability and functionality. Hydrogen atoms are added for a more complete representation, while preserving the fundamental molecular structure and checking for any alterations in the total number of atoms or chiral centers. Geometric optimization is performed by generating multiple 3D conformations, applying MMFF (Merck Molecular Force Field) optimization, and calculating key geometric properties such as the radius of gyration, which describes spatial distribution, and RMSD (Root Mean Square Deviation), which measures structural differences. Energy minimization was conducted to find the lowest energy conformer and reduce the overall molecular energy, aiming for a more stable configuration. Additionally, conformational changes were introduced by adjusting bond angles and torsions to minimize steric strain and find a more energetically favorable 3D arrangement. Throughout this process, the primary goal was to identify the most stable 3D configuration of the molecule while maintaining its core structural integrity, with the SMILES string and atomic connectivity preserved, modifying only the 3D spatial arrangement.

```
import rdkit
from rdkit import Chem
from rdkit.Chem import AllChem, Descriptors

def optimize_molecule_conformation(smiles, num_conformers)
```

```
=50, max_iterations=500):
      Perform initial conformational optimization of a
         molecule using RDKit
      Parameters:
      ______
      smiles : str
          SMILES representation of the molecule
13
      num_conformers : int, optional
14
          Number of conformers to generate (default: 50)
      max_iterations : int, optional
16
          Maximum number of optimization iterations (default:
17
              500)
18
      Returns:
19
20
      optimized_mol : rdkit.Chem.Mol
21
          RDKit molecule with optimized 3D coordinates of
              lowest energy conformer
23
      # Convert SMILES to RDKit molecule
24
      mol = Chem.MolFromSmiles(smiles)
25
26
      # Add hydrogens to the molecule
27
      mol = Chem.AddHs(mol)
28
29
      # Embedding parameters
30
      params = AllChem.ETKDGv3()
      params.randomSeed = 42 # Reproducibility
32
      params.maxAttempts = 100 # More attempts to generate
33
          conformers
34
      # Generate multiple conformers
      num_conformers = min(num_conformers, 200) # Limit to
36
          200 to prevent excessive computation
37
      # Embed and optimize multiple conformers
38
      conformer_ids = AllChem.EmbedMultipleConfs(mol, numConfs
39
          =num_conformers, params=params)
40
      # Perform MMFF optimization
41
      res = AllChem.MMFFOptimizeMoleculeConfs(mol, maxIters=
42
         max_iterations)
      # Extract energies
45
      conformer_energies = [e[1] for e in res]
46
      # Find the lowest energy conformer
47
      if conformer_energies:
```

```
lowest_energy_index = conformer_energies.index(min(
49
              conformer_energies))
           # Create a new molecule with only the lowest energy
51
              conformer
           optimized_mol = Chem.Mol(mol)
          optimized_mol.RemoveAllConformers()
          optimized_mol.AddConformer(mol.GetConformer(
54
              conformer_ids[lowest_energy_index]))
           return optimized_mol
57
      return mol
58
59
60
  # Main execution
61
62 def main():
      # SMILES string of the molecule
63
      smiles = "COc6ccc(OS(=0)(CC(=0)[O-])(OC(O)C(O)CO)OC(C)(C)
64
          ) Cc2nc1c(=0)[nH]c(N)nc1n2[C0H]4[C00H]30[Mg]0[C00H]3[
          C@@H](COP(=0)(0)OP(=0)(0)0)[0+]4C5=CC=CC05)c(N)c6"
65
66
      try:
           # Optimize the molecule
           optimized_mol = optimize_molecule_conformation(
68
              smiles)
           # Generate and print the SMILES of the optimized
70
              molecule
           optimized_smiles = Chem.MolToSmiles(optimized_mol)
71
           print("Optimized Molecule SMILES:")
72
73
          print(optimized_smiles)
74
      except Exception as e:
          print(f"An error occurred during optimization: {str(
              e)}")
78
  if __name__ == "__main__":
79
      main()
```

We get the following molecule upon Initial Conformational Optimization:

Optimized Molecule SMILES:

[H]OC([H])([H])C([H])(O[H])C([H])(O[H])OS(=0)(Oc1c([H])c([H])c(OC([H])([H])[H])c([H])c1N([H])[H])(OC(C([H])([H])[H]) (C([H])([H])[H])C([H])([H])c1nc2c(n1[C@]1([H])[O+](C3=C([H])C([H])=C([H])C([H])([H])03)[C@]([H])(C([H])([H])0P(=0) (O[H])OP(=0)(O[H])O[H])[C@@]3([H])O[Mg]O[C@]31[H])N=C(N([H])[H])N([H])C2=O)C([H])([H])C(=O)[O-]

The optimized structural changes are introduced to improve precision and maintain molecular integrity. Hydrogen atoms are explicitly represented with [H] annotations, ensuring an exact hydrogen count for better computational accuracy and precise 3D conformational analysis. Stereochemical centers are preserved, with specific notations like [C@] and [C@@] retained to maintain the molecule's 3D spatial arrangement. The core molecular skeleton remains unchanged, preserving bond connections and atom arrangements, confirming a successful optimization without altering the fundamental structure. Each atom now explicitly displays its hydrogen count, offering granular structural information crucial for advanced molecular modeling. Functional groups, including the sulfone group OS(=O), heterocyclic rings, and the metal coordination site [Mg], are preserved in detail, ensuring the molecule's chemical properties remain intact.

The optimization provides significant computational benefits, including more accurate energy minimization and precise spatial configuration, which are essential for reliable molecular modeling. These enhancements also improve the molecule's suitability for molecular dynamics simulations, enabling more detailed studies of its behavior and interactions under various conditions.

5 Reterosynthetic route for the preparation of drug:

To analyze the retrosynthetic route for this molecule with multiple functional groups, stereochemical centers, and a coordination complex with Magnesium, it is necessary to train a specialized ML model tailored to handle its unique structural and functional challenges. Current models, such as IBM RXN and Sypra, failed to design routes for this drug molecule because they are typically trained on general reaction datasets and lack the depth to address rare functional groups like metal coordination complexes, intricate stereochemistry, and multi-functional group interactions. Our specialized model would focus on these challenges by incorporating reaction data specific to organometallic chemistry, complex stereochemical transformations, and advanced protecting group strategies. Furthermore, it would use tailored heuristics and quantum chemi-

cal insights to predict feasible synthetic pathways, overcoming the limitations of existing models. Additionally, to evaluate the final synthesized compound, it will be necessary to hire Organic Chemists to test the drug's effects in vitro, ensuring its biological activity and safety.

6 Conclusion:

This molecule holds significant potential as a therapeutic agent for cancers driven by various mutated Ras proteins. The molecule's design, with its ability to selectively disrupting Ras signaling by interfering with the GTP bound to Ras., can effectively disrupt the aberrant signaling pathways that promote uncontrolled cell proliferation in Ras-driven tumors. By targeting the GTP-bound state of mutated Ras, this molecule prevents the activation of downstream signaling cascades such as the MAPK/ERK and PI3K/Akt pathways, which are commonly upregulated in Ras-driven cancers. Furthermore, the molecule's specific binding to mutated Ras ensures minimal off-target effects, offering a potential treatment with reduced toxicity compared to conventional therapies. When used as a part of a targeted therapy regimen, this drug has the potential to not only inhibit tumor growth but also provide a more precise and effective treatment option for patients with Ras-driven cancers, offering a promising advancement in cancer therapeutics.

7 Reference:

1. Bugnon M, Röhrig UF, Goullieux M, Perez MAS, Daina A, Michielin O, Zoete V. SwissDock 2024: major enhancements for small-molecule docking with

Attracting Cavities and AutoDock Vina. Nucleic Acids Res. 2024

Eberhardt J, Santos-Martins D, Tillack AF, Forli S.. AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings. J. Chem. Inf. Model., 2021

- 2. Landrum, G. A. (2023). RDKit: Open-source cheminformatics software (Release 2023.03.1). Retrieved from https://www.rdkit.org.
- 3. Murray, C. W., Rees, D. C. (2009). The rise of fragment-based drug discovery. Nature Chemistry, 1(3), 187–192. DOI: 10.1038/nchem.217.
- 4. Berman, H. M., Battistuz, T., Bhat, T. N., Bluhm, W. F., Bourne, P. E., Burkhardt, K., ... Zardecki, C. (2002). The Protein Data Bank. Acta Crystallographica Section D: Biological Crystallography, 58(6), 899–907. DOI: 10.1107/S0907444902003451.