3,5-dinitrophenyl clubbed azoles against latent tuberculosis- a theoretical mechanistic study

N.V.S.Viswanadha Murthy, M1, V.Girija Sastry1, Syed Hussain Basha2

1Department of Pharmaceutical Chemistry, A.U. College of Pharmaceutical Sciences, Andhra University, Visakhapatnam, Andhra Pradesh – 530 003, Andhra Pradesh, India. 2Innovative Informatica Technologies, Telangana, Hyderabad – 500 049, India.

Abstract: In this present study, novel series of dinitrophenyl linked azole compounds were designed keeping in view of the importance of azole molecules in combinations with nitro and amino group by hyphenation of these two pharmacophores as a single molecular scaffold against latent tuberculosis targeting Isocitrate Lyase (ICL) enzyme by using various computational tools. Our docking studies evidenced that all the novel designed compounds have the potential to inhibit the ICL enzyme with a binding energy in a range of -4.6 to -8.9 Kcal/mol. Moreover, all the designed compounds were predicted to be having promising ADMET parameters and found to be well in compliance with Lipinski’s rule of five along with no toxicity profile. Among all the thirty five compounds tested, compound 3F is the best lead like molecule with -8.9 kcal/mol of binding energy compared to Rifampicin and 3-nitropropionate which has shown -8.1 kcal/mol and -4.6 Kcal/mol. Molecular dynamic simulation studies for compound 3F in complex with Isocitrate Lyase has elucidated several interesting molecular level protein-ligand interactions with some of the important amino acid residues present at the active binding site of ICL enzyme compared to Rifampicin and 3-nitropropionate. Conclusively, novel designed compound 3F of the present study have shown promising tuberculosis inhibition potential worth considering for further evaluations.

Keywords: Mycobacterium tuberculosis, Isocitrate Lyase, docking, ADMET, MD simulations.

I. INTRODUCTION

An infection with Mycobacterium tuberculosis (M.Tb) could result in engaged balanced “symbiosis” for survival, by which 10% of infected people will develop active disease (TB) while the rest shows no symptoms to carry the long-lived M.Tb. commonly known as latent tuberculosis infection (LTBI), which acts as enormous reservoir for potential reactivation of TB, which is a prevalent barrier for treatment and control accompanied with drug resistant M.Tb. Isocitrate Lyase (ICL) enzyme is an key enzyme in glyoxylate shunt catalyses Isocitrate to succinate and glyoxylate. An erasure or inhibition of ICL impairs the endurance of M. tuberculosis in activated macrophages and also reported to eliminate it from the lungs of chronically infected mice. Several studies reveal that ICL can acts as a promising drug target for latent tuberculosis due to its nonexistence in mammals [1-2]. However, due to the ability of latent bacilli to exhibit different metabolic pathways, majority of current drugs including Delamanid and Bedaquiline are inadequate when applied for long term regimens due to their resistance mechanism and toxicities [3-6]. Hence, there is a need to design a novel entity targeting both active and latent tuberculosis. Dinitro phenyl moiety- a vital and fortunate pharmacophore with significant effect on latent tuberculosis [7] via encompassing inhibition with crucial enzymes i.e., Isocitrate Lyase (ICL) in glyoxylate shunt for sustaining M.Tb in chronic phase and Isocitrate dehydrogenase (IDH) involved in citric acid cycle (TCA) is well established in literature [8-9]. Recently, several nitro group-containing molecules shown to be exhibiting potential inhibition of mycobacterial decaprenyl-phosphoryl-β-D-ribofuranose 2’-oxidase (DprE1), which involves in the biosynthesis...
of arabinan polymers in mycobacterium [10]. After extensive comparison between the isocitrate lyase enzyme inhibitors in M.Tb and Non-M.Tb species, it is been observed that molecules containing nitro and amino group showed significant ICL enzyme inhibition.

On the other hand, azoles–nitrogen containing five membered heterocyclic compounds posses interesting pharmacokinetics / dynamics properties. It’s simple synthesis through Click chemistry [11] and ability to inhibit sterol synthesis via sterol demethylase inhibition in eukaryotic organism [12-14] with demonstrated better potential to inhibit latent bacilli in mice compared to rifampicin [15] pulls an attention towards this pharmacophore.

After extensive literature search, it was observed that, till date not enough efforts have been put towards hyphenation of these two pharmacophores as a single molecular scaffold, which creates a huge interest in crafting novel inhibitors for both latent and active tuberculosis targeting Isocitrate Lyase (ICL) enzyme. In this scenario, in this present study, a novel series of dinitrophenyl linked azole compounds were designed (Fig. 1) keeping in view of the importance of azole molecules in combinations with nitro and amino group by hyphenation of these two pharmacophores as a single molecular scaffold against latent tuberculosis targeting Isocitrate Lyase (ICL) enzyme by using various computational tools.

II. RESULTS AND DISCUSSION

a. Docking of the compounds with Isocitrate Lyase

We have performed the docking studies for the present studied thirty five compounds (supplementary table 1) with the Isocitrate Lyase protein in order to know the binding energy involved in this complex formation and to know the molecular interactions responsible for this target specific inhibition. Docking results are tabulated in supplementary table 2. All the thirty five compounds studied in this present work have shown to be successfully docking inside the active site of Isocitrate Lyase with a binding energy in a range of -4.6 to -8.9 Kcal/mol. As per the molecular docking results, it was revealed that Compound 3F has the best estimated -8.9 Kcal/mol of binding energy for the Isocitrate Lyase electrostatic interactions with Leu60, Gly340, Glu308, Met309, Leu310 and Asp307 along with Vander waals interactions with Val26, Ala276, Asp280, Ala279, Lys24, Asp25 and Trp57 residues (Fig. 2a). While, Rifampicin – well known antibiotic used for several bacterial infections including tuberculosis, which we have taken as a control for our tested compounds has shown a binding energy of -8.1 Kcal/mol by forming a hydrogen bonding with GLN79 along with Electrostatic interactions with Tyr388, Arg386, Ala83, Arg82 and Gly387. Whereas, residues Ala353, Leu376, Glu380, Arg379 and Ala383 were observed to be forming Vander waals interaction with the compound (Fig. 2b).

Whereas, 3-nitropropionate – a potential ICL inhibitor, has shown a binding energy of -4.6 Kcal/mol by forming a direct hydrogen bond with Ile150 along with electrostatic interactions with Ala177, Gly175, Gly52, Leu147, Leu48, Asn125, Asn145 and Ala148 residues. Val55, Pro149 and Val176 were observed to be forming Vander waals interaction with the compound (Fig. 2b).

![Figure 1: Rational in Design of Proposed Hybrids](image)

Figure 1: Rational in Design of Proposed Hybrids: (A) Nitro group from M.Tb ICL inhibitors [6] (B) Dinitro phenyl moiety from ICL inhibitors and DprE1 inhibitors [8-10] (C) Triazole ring from marketed azoles i.e., Fluconazole (D) Oxadiazole ring and Dinitro phenyl from potent molecule against latent M.Tb [7] on molecular hybridization lead to proposed hybrids (E).

a. Prediction of pharmacological properties

Pharmacological properties of the present studied compounds have been evaluated using Osiris Property Explorer online server according to Lipinski’s Rule of Five [16] for Oral Bioavailability (supplementary table 3). For the tested compounds it was inferred that most of the present studied compounds are compiling according to Lipinski’s rule of five for good bioavailability, however 3d, 3g, 4c, 6b-f, 6j-k, 7c, 7e-f along with Rifampicin were found to be not compiling with the Lipinski’s rule of five for good oral bio-availability. The result of toxicity analysis of all the analyzed compounds is shown in (supplementary table 4). “None” means low toxic tendency, “medium” means the midcore and “high” means high tendency of toxicity. Supplementary table 4 shows that all the analyzed thirty five compounds has low or no toxic tendency however compound 3-nitropropionate has predicted to be having high toxicity notably.
Based on the promising pharmacological properties, toxicity profile and better binding efficiency we have taken compound 3F for further studies using molecular dynamic simulations in order to reveal its mode of action and to further probe the molecular interactions responsible for this strong inhibition compared to Rifampicin.

b. MD simulations of Isocitrate Lyase protein in complex with compound 3F compared to Rifampicin and 3-nitropropionate

In order to analyze the conformational changes in the ICL enzyme upon binding of compound along with the aim of revealing their underlying molecular interactions at atomic level of Isocitrate Lyase in complex with our best predicted drug like compound (compound 3F) among the thirty five present investigated compounds compared with Rifampicin and 3-nitropropionate. Before starting the analysis part of the MD simulations trajectories, firstly we have analyzed the simulation quality parameters, which is an important issue to deal with to conform that the molecular dynamic simulations were carried out under our given temperature, pressure and simulation box volume conditions throughout the simulated time. 10 nano second of simulation timescale used in this present study is of enough time for the side chain rearrangements in native as well as protein–ligand complexes in order to facilitate the most stable binding conformation [17]. For the molecular dynamic simulations, input parameters given such that the temperature is at 300 k; pressure is at 0 atmospheric and volume of the simulations water box is kept at 550000 Å³ throughout the simulated time. As evident with Fig. 3, the given input simulation parameters were maintained thoroughly throughout the simulated time of 10 ns each, thus confirming the quality of the simulations carried out.

c. Simulation event analysis:

After confirming the simulation quality parameters, we have studied the Root mean square deviation (RMSD) for the protein backbone (Fig. 4), RMSD of the Ligand (Fig. 5), Radius of gyration (ROG) and intra molecular hydrogen bonds (Fig. 6), Root mean square fluctuations (RMSF) of the protein individual residues along with its influence on its total secondary structure elements (SSE) (Fig. 7), contributions with
time dependent function of MD simulations in order to understand the stability and conformational changes of the Isocitrate Lyase protein in its apo state and in complex with compound 3F compared to Rifampicin and 3-nitropropionate at each frame of the MD simulations trajectory.

Figure 4: RMSD graphs of Isocitrate Lyase protein’s backbone in its apo state (red), in complex with Rifampicin (blue), in complex with compound 3F (green) and in complex with 3-nitropropionate (black). Red and blue color ribbons represent the pre and post simulation protein superimposition respectively with their calculated averaged RMSD.

Protein’s backbone RMSD was shown to be fluctuating between 1.5 and 5.0 Å, with an average of 1.71, 1.97, 0.87 and 0.63Å in its apo state, in complex with Rifampicin, in complex with compound 3F and in complex with 3-nitropropionate respectively as shown in figure 5. Moreover, movements of the secondary structural elements (SSE) which are responsible for these protein backbone RMSD as pointed out with small arrows in the figure 5 also shows that majority of helices and loops movements observed in case of apo protein, in complex of Rifampicin and in complex with 3-nitropropionate has been much minimized in case of ICL complex with compound 3F. Much reduced RMSD for ICL protein backbone in complex with compound 3F thus stabilizing the overall protein and SSE is an indication towards its better inhibitory potential compared to Rifampicin and 3-nitropropionate, as the flexibility and movements of a protein is what determines it’s functionality.

From the MD simulation trajectory for ligands in complex with Isocitrate lyase, we have sampled two best snapshots which represents the least and highest ligand RMSDs to study the change in the molecular interactions. Initially, Rifampicin was found to be forming a hydrogen bond with Arg386 along with Vander waals interactions with Tyr388, Gly387, Thr389, Ala83, Leu85, Gly84, Ala349, Tyr356, Met76, Ala353, Ala383, Gln80, Glu380, Leu376, Ser357, Gln79, Leu69, Ile346 along with Electrostatic interactions with Arg82 and Arg379 residues at around 0.8Å of its RMSD.

When these interactions were compared with Rifampicin at its highest RMSD of 1.5Å, molecular interactions profile was observed to be changed significantly with a hydrogen bond with GLN79, Pi-pi stacking with Tyr388, Vander waals interactions with Pro316, Gln80, Ala349, Met76, His393, ly350, Ser357, Tyr356, Ala353, His352, Leu69, Ile346, Asn67, Gly387, Leu85, Ala83 and electrostatic interactions with Arg379, Arg386 and ARG82 residues.

In case of 3-nitroproionate, it was found to be forming a hydrogen bonds with Arg51 and Ile150 along with Vander Waals interactions with Leu48, Ala177, Ala49, Val176, Ala148, Pro149, Leu128, Leu147 and Val55 initially. When the last MD simulation snapshot was compared with this snapshot an extra hydrogen bond with Asn125 was revealed to be playing an important role in stabilizing this compound.

While compound 3F, Initially, it was found to be forming a hydrogen bond with Asp307, Pi-pi stacking with Trp57, Vander Waals interactions with Trp23, Pro21, His61, Ala279, Tyr305, Asp25, Val26, Gln308, Pro277, Ala276, Val224, Met309, Leu281, Leu60, Asp280 and Electrostatic with Arg22 and Lys24 residues at its highest RMSD peak of around 1.5Å observed at 3.8 nanoseconds into simulation trajectory. Leu281, Leu60, Val224, Leu56, Phe343, Trp57, Met309 and Electrostatic interactions with Asp25, Arg213 and Asp280 residues. Strong salt bridges formed by Lys24, Lys342 and Asp307 residues were most likely to be the major factor towards forming this much stabilized protein-ligand complex (Fig. 5).

In contrary to Rifampicin, compound 3F was found to be quite stabilizing as the simulation progress.
setting at well below 0.5Å by forming Pi-Pi stacking with Arg22; strong Salt bridges with Lys24, Lys342, Asp307; Vande Waals interactions with Pro21, Trp23, Val26, Gln308, Ala279, Protein structure compactness can be predicted using radius of gyration (ROG) [18]. We have analyzed the Isocitrate Lyase enzymes ROG in its apo state, in complex with Rifampicin, in complex with 3-nitropropionate and in complex with compound 3F in order to estimate the effect of these compounds on the overall compactness and stability of the protein and found out that the ROG of the protein is maintaining an average of 23.618, 23.374, 22.948 and 23.312Å respectively. From the figure 8 it has been calculated that the Isocitrate Lyase protein structure has an average intra molecular hydrogen bonds of 321, 324, 324 and 335 in ICL in its apo state, in complex with Rifampicin, 3-nitropropionate and in complex with compound 3F respectively. ROG and Total number of intra molecular hydrogen bonds are valuable parameters towards understanding the underlying forces contributing for the stability of the protein structure; Lower ROG and higher rigidity in the protein. This data clearly evidences that our proposed compound 3F has the potential to inhibit the Isocitrate Lyase protein by causing contraction in the proteins overall conformation resulting in increased intra molecular hydrogen bonds contributing for the proteins rigidity. Increased rigidity in the protein is a direct indication of reduced activity of the protein functionality due to minimized protein flexibility.

Each residue fluctuations (RMSF) were calculated and averaged for the entire 10ns simulated timescale in order to identify the higher flexible regions in the protein. When the RMSF graph of Isocitrate Lyase in complex with compound 3F was analyzed and compared with apo state protein and protein in complex with Rifampicin, high peaks of fluctuations were observed at 230-250 and 350-380 residue positions.

Secondary structural elements which were responsible for this high RMSF peaks has been visualized, however no significant loss of SSE percentage was observed due to these high fluctuations in the protein structure. Helices at 350-380 residues were thought to be responsible for ligand recognition and complex formation (Fig. 9). To further understand the underlying energy landscapes, we have performed a separate metadynamic simulation with selected variables at the tip of the compound and the core of the macromolecule following the default protocol as.

Figure 5: RMSD graphs of the Ligands Rifampicin (blue) and compound 3F (green) in complex with Isocitrate Lyase enzyme. Upper panel represents the 2D molecular interactions of Rifampicin first MD snapshot with least RMSD (left) and last MD snapshot with Maximum RMSD (right). Lower panel represents the 2D molecular interactions of compound 3F at the highest RMSD (left) and least RMSD (right). Blue and red color ball and stick representations are the superimposition of the least and highest RMSD orientations of the ligands.
explained briefly elsewhere [19]. As per the meta-
dynamic results, it was revealed that the free energy for
the protein-ligand complex has observed to be at its
minima at -10.38 Kcal/mol for compound 3F compared
with -10.13 & -8.33 Kcal/mol for Rifampicin and 3-
nitropropionate respectively (Supplementary figure 1).
The order of Compound 3F > Rifampicin > 3-
nitropropionate has been observed to the well in
agreement with our docking binding energy order.

d. Molecular interactions of Isocitrate
Lyase-compound 3F complex during
MD simulations:

Detailed molecular interactions between
Isocitrate Lyase and compound 3F compared with
text control Rifampicin were studied using Desmond’s
simulation interaction diagram program. There were
about 27 contacts found in between Isocitrate Lyase and
Rifampicin in total among which hydrogen bonds were
observed with H-Bonds: Gln79, Ala349, Arg386,
Tyr388; Hydrophobic interactions were observed with
Asn67, Leu69, Met76, Arg82, Ala83, Leu85, Pro316,
Leu321, Gln333, Ile346, Ala83, Ala349, Ala353,
Tyr388, Leu376, Arg379, Tyr388, Thr389, Lys392, His393; an
ionic interaction with Arg386 and several water bridges
with Asn67, Asn75, Met76, Ala83, Ala349, His352,
Ala353, Ser357, Asp360, Glu380, Arg386, Gly387 and
Tyr388 residues were noted. While, 3-nitropropionate
exhibited a total of 14 contacts among which three were
direct hydrogen bonds with Arg51, Asn125 and Ile150,
and rest of the contacts were water bridging bonds with
Thr47, Leu48, Val112, Arg122, Asn125, Gln129,
Asn145, Leu147, Ala148, Ala174, Gly175 and Ala177
residues. Whereas, in case of compound 3F a total of 23
contacts were found, hydrogen bonds were observed
with Lys24, Asp307, Lys342; Hydrophobic interactions
with Pro21, Val26, Trp57, Leu60, His61, Leu63,
Val224, Leu281, Pro306, Met309, Ala337, Gly340,
Phe343; Ionic and salt bridges with Lys24, Asp307,
Lys342; Water bridges with Lys24, Asp25, Leu60,
His61, Ala276, Ala279, Asp280, Lys302, Pro306,
Asp307, Gln308, Met309, Ala337 and Lys342 residues
were observed (figure 10). Although the number of
contacts count was less in case of compound 3F, strength
of the interactions during the simulation was way more
than the interactions strength observed in case of
Rifampicin and 3-nitropropionate. Especially, salt bridge
bond with Lys342 was observed to be playing a crucial
role in stabilizing the protein – ligand complex in case of
compound 3F.
III. CONCLUSION

Our *In silico* studies provided a rationalization to the ability of present studied novel compounds as a valuable small inhibitor compound with high binding affinity towards Isocitrate Lyase consolidating their complex’s thermodynamic stability for plausible antituberculosis activity. Moreover our hypothesis further substantiated by the predicted high negative binding energy values of these compounds that they have the plausibility to inhibit Isocitrate Lyase. Further, de novo simulations for 10 nanoseconds revealed ligand interactions with the residues of Isocitrate Lyase, all or some of which fall under catalytic active site important residues for its structurally stability and/or functionality. Helices at 350-380 residue position of the protein were thought to be responsible for ligand recognition and complex formation. Salt bridges formed by compound 3F with residues Lys24, Lys342, Asp307 were thought to be crucial role players in its better binding efficiency.
prepared to Rifampicin and 3-nitropropionate. Knowledge gained through this study would be of high value towards enhancing the discovery of Isocitrate Lyase target specific drug molecules. The present study also evidences that Isocitrate Lyase is a potential drug target for the present investigated compounds for their anti-tuberculosis property. These promising binding energies, along with better ADMET properties of compound 3F, compared with the standard Rifampicin substantiates the need of further evaluating this compounds ability to inhibit tuberculosis.

### IV. COMPUTATIONAL METHODS

#### a. Softwares and programs

Schrodinger’s maestro visualization program v9.6 [20] is utilized to visualize the receptors, ligand structures, hydrogen bonding network, to calculate length of the bonds and to render images. Chemsketch was used to draw the ligand compounds. Autodock 4.0 [21] is the preliminary docking program used in this work for the semi-flexible protein ligand docking studies. Preparation of the ligands and protein receptors in pdbqt file; determination of the size of the grid box size was done using AutoDock Tools version 1.5.6. Molinspiration, Osiris property explorer program was used to study the ADMET properties of the compounds. Schrodinger’s Desmond module v3.6 [22] was used for molecular dynamic simulation studies. Preparation of protein receptor and Ligand: The Isocitrate Lyase crystal structure [PDB: 1F61] was imported from the Protein Data Bank (PDB) [23]. Many missing atoms were identified in the crystal structure, which were supplemented by the repair commands of AutoDock. Co-crystallized water molecules present in the protein structure has been removed; missing H-atoms were added in order to clean the structure before docking studies. This formatted clean, energy minimized structure so obtained was used for the semi-flexible dockings. The ligand molecules were drawn using chemsketch software. Accelrys Discovery Studio [24] was used to energy minimize the ligand compounds and receptor using Steepest Descent and Conjugate Gradient methods with CHARMM force field [25].

#### b. Semi-flexible docking

Autodock v4.0 is used to identify binding pose with associated energy along with the IC50 value prediction of the compounds with drug target Isocitrate Lyase. Default parameters of the software program have been applied similar to the protocol followed elsewhere [26-29]. Briefly, Lamarckian Genetic Algorithm (LGA) [30] with default atomic salvation parameters 126 Å (x, y, and z) grid box for scoring energy was set centered at X = 48.461; Y = 67.792 and Z = 17.982 with 0.375 angstroms grid points spacing. Care was given during the grid box preparation to ensure that the active site of Isocitrate Lyase was surrounded by the 3D grid box centered at its active ligand binding site location.

#### c. Pharmacological properties of the compounds

Osiris Property Explorer online server (www.organicchemistry.org/prog/peo/) was used to check the pharmacological properties of the designed small ligand molecules. 2D descriptor values H-bond acceptors, donors, logP, molecular weight etc., were analyzed along with toxicity and drug score.

#### d. MD Simulations in Water

In order to study the thermodynamic stability of the protein-ligand complex, “Desmond v3.6 Package” was used. OPLS 2005 force field parameters have been applied to simulation TIP3P water models [31-32]. Periodic boundary conditions were used to determine the specific size and shape of the water box buffered at 10 Å distances and box volume was calculated as 550000 Å³. Appropriate counter ions placed randomly in the solvated system were added to neutralize the system electrically. Default energy minimization protocol has been employed [33], to relax the system before starting the actual production run in the NPT ensemble [34] with 1 atmospheric pressure and 300k temperature using Nose-Hoover temperature coupling and isotropic scaling [35].

Author’s contribution: NVSVMM and VGS conceived the idea. NVSVMM and SHB performed the experiments. NVSVMM, VGS and SHB interpreted the data and wrote the manuscript.

**SUPPORTING INFORMATION**

**Supplementary figure 1:** Metadynamic energy landscape of compound 3F compared with Rifampicin and 3-nitropropionate respectively (available at online version of article).

**Supplementary table 1:** Structures of novel designed 3, 5-dinitrophenyl clubbed azoles used in the present study (available at online version of article).

**Supplementary table 2:** Docking results of the 35 novel designed 3,5-dinitrophenyl clubbed azole compounds targeting ICL drug target (available at online version of article).
Supplementary table 3: Molecular descriptor properties of novel designed 3,5-dinitrophenyl clubbed azole compounds (available at online version of article).

Supplementary table 4: Toxicology descriptor properties of novel designed 3,5-dinitrophenyl clubbed azole compounds (available at online version of article).

REFERENCES

29. Samad, Firoz Abdul, et al. "A comprehensive In Silico analysis on the structural and functional impact of SNPs in the congenital 3,5-dinitrophenyl clubbed azoles against latent tuberculosis


3,5-dinitrophenyl clubbed azoles against latent tuberculosis