Design, Synthesis, and Development of 4-[(7-Chloroquinoline-4-yl)amino]phenol as a Potential SARS-CoV-2 Mpro Inhibitor

James Guevara-Pulido,* Ronald A Jiménez, Sandra J. Morantes, Deissy N. Jaramillo, and Paola Acosta-Guzmán
SUPPORTING INFORMATION

Preparation of nanoparticles 2
RP-HPLC method for estimation of drug release 3
IN VITRO cytotoxicity assay 3
Ligand-Based virtual screening (LBVS) 3
Structure-Based virtual screening (SBVS) 6
Descriptor vs descriptor correlation 7
Descriptor correlation vs IC\textsubscript{50} 7
Experimental IC\textsubscript{50} VS Predicted IC\textsubscript{50} 8
R\textsuperscript{2} According to hidden layer nodes 9
Characterization and NMR Spectra 10
HPLC of 4 - ((7-chloroquinoline-4-yl) amino) phenol 12
References 13
PREPARATION OF NANOPARTICLES

Materials
Polycaprolactone (PCL) average Mn 80,000, Dichloromethane (DCM) y Polyvinyl alcohol (PVA) Mw 89,000-98,000 99+% hydrolyzed were purchased from Merck KGaA (Darmstadt, Germany).

Preparation of Unloaded Nanoparticles

A PCL-based nanoparticulate system was chosen for drug incorporation. Unloaded PCL nanoparticles were prepared using the emulsion-solvent evaporation technique.[1] Briefly, for the organic phase 60 mg of PCL were dissolved in 10 mL of DCM using a magnetic stirrer (800 rpm for 10 min), while an aqueous solution of 1% PVA (w/w) was used as the water phase. Then, the organic phase was slowly added into the water phase while stirring at 600 rpm, and the resulting mixture was homogenized via sonication (Fisher Scientific at 600 Hz with a 25% amplitude for 2 min) followed by overall stirring (3h) to remove solvent (DCM). After that, the resulting nanoparticle suspension was centrifuged at 4500 rpm for 45 min using the Thermo Scientific SORVALL ST40R at 40°C. The resulting pellet was washed three times using deionized water, then the pellet was reconstituted in 5 mL of deionized water and frozen at -20°C for 12 h. Finally, the reconstituted pellet was freeze-dried using the VirTis AdVantage Plus Freeze Dryer at 120 mmHg for 48 h, and then stored ensuring it was protected from light and humidity.

Preparation of Loaded Nanoparticles

Method 1: loaded nanoparticles were prepared by absorption. First, 50 mg of unloaded PCL nanoparticles were added to 5 mL of a 10% w/w aqueous solution of the chosen candidate, and the resulting mixture was stirred at 100 rpm for 12 h. Then, nanoparticles were collected by centrifugation at 4500 rpm for 45 min (Thermo Scientific SORVALL ST40R) at 4°C. The resulting pellet was washed thrice with deionized water, reconstituted, freeze-dried, and stored ensuring it was protected from light and humidity.

Method 2: loaded nanoparticles were also prepared following the steps described in the section “Preparation of Unloaded Nanoparticles,” adding 30 mg of the chosen candidate to the organic phase prior to its incorporation into the water phase.

Characterization of Nanoparticles

Nanoparticle size and zeta potential (ζ) were determined by dynamic light scattering and electrophoretic light scattering respectively employing a Malvern Zetasizer (Malvern Instruments, UK) and using 10% (w/w) aqueous suspensions of the nanoparticles previously prepared.

Entrapment Efficiency

Three 50 mg samples of loaded nanoparticles were degraded using 5 mL of DCM at 37°C. The samples were then centrifuged at 3000 rpm for 10 min, and a sample of 100 µL of each supernatant was collected. The candidate was quantified by UV/VIS spectrophotometry at 560 nm employing a Dymax® spectrophotometer. Concentrations of the candidate were determined from the corresponding absorbance readings, by using a previously obtained calibration curve (0µg – 100µg).

In vitro release

Three 50 mg samples were resuspended in 1.5 mL of PBS (pH 7.4), and then incubated in 2 mL Eppendorf tubes at 37°C under orbital stirring (100 rpm). For 8 days, samples were centrifuged after
6 h, 12 h and every 24 h at 3000 rpm for 5 min. After every centrifugation, samples (100 µL) of supernatant were collected and replaced with fresh portions of PBS (100 µL). Quantification of the candidate was carried out by UV/VIS spectrophotometry at 560 nm employing a Dymax® spectrophotometer. Concentrations of the candidate were determined from the corresponding absorbance readings, by using a previously obtained calibration curve (0µg – 100µg).

RP-HPLC METHOD FOR ESTIMATION OF DRUG RELEASE

The estimation of 4-((7-chloroquinolin-4-yl)amino)phenol release was carried out using an amino Phenomenex column and water as the mobile phase (100%). UV detection was set at 350 nm and column temperature was maintained at 40°C. The retention time of 4-((7-chloroquinolin-4-yl)amino)phenol was 7.9 min. For the analysis, 20 mg of the compound were dissolved in 10 mL of deionized water, and a 1 mL aliquot was taken and diluted to a total volume of 10 mL. Then, from the previous solution, five aliquots (50 mL, 100 mL, 200 mL, 300 mL, and 400 mL) were taken and diluted to a total volume of 10 mL, to obtain five solutions with the following concentrations respectively: 3.25 nM, 6.5 nM, 13.0 nM, 19.5 nM and 26.1 nM.

IN VITRO CYTOTOXICITY ASSAY

The cytotoxic potential of the free candidate (4-((7-chloroquinolin-4-yl)amino)phenol) and of the loaded and unloaded nanoparticles was evaluated in vitro using murine fibroblast (L929) and human lung adenocarcinoma (A549) cell lines via the resazurin reduction assay.[2,3] Cells were cultured in DMEM medium supplemented with heat-inactivated 10% (v/v) FCS (fetal calf serum; Gibco BRL), streptomycin (10 μg/ml) and penicillin (100 U/ml) and were maintained at 37°C in a humidified atmosphere with 5% CO₂. Both the free candidate and chloroquine (control) were dissolved in DMSO at a stock concentration of 1mM, while the nanoparticles were suspended in growth medium at a concentration of 1000 µg/mL. When cells reached their logarithmic growth phase, cells were seeded in 96-well plates at 5000 cells per well, incubated for 24 h and treated for 24, 48 and 72 h with 10 serial dilutions 1:3, starting with 500 µM of both compounds, and 500 µg/mL of the nanoparticles. Final concentration of DMSO in each well was kept below 0.1%. However, DMSO at 25% was used as the control for cell death. After treatment, the medium was replaced with 100 µL of resazurin (44 µM), and after incubation (4 h), the fluorescence emitted by viable and/or metabolically active cells was quantified at excitation and emission wavelengths of 535 and 595 nm respectively using a TECAN GENios spectrofluorometer. Assays were carried out thrice in three different weeks, and results were reported in viability percentages after comparison with the untreated control.

LIGAND-BASED VIRTUAL SCREENING (LBVS)

Development of artificial neural network (ANN)

An artificial neural network (ANN) with backpropagation was built, where each input node received a particular molecular descriptor, and the output node generated the predicted response variable (IC₅₀) (Figure 1). The number of nodes in the hidden layer was modified after each computation, taking into consideration the leave-one-out cross-validation method and the determination coefficient (R²) to determine the best-fitting model. Thus, the number of hidden nodes that yielded the best-fitting model (R² closest to 1) was chosen for the final prediction.

The next section will describe in greater detail the structure of the proposed network, the validation method and the validation metrics used.

Neural Network Structure

The proposed neural network (Figure 1) features an initializing algorithm followed by feed-forward and backpropagation algorithms.[4]
Figure 1. Structure of Neural Network

Step 1: Initialization:
- Define weights ($\theta$) and bias nodes ($X_0, A_0$)
- Initialize weights using a random number, different than zero.
- Set bias nodes equal to 1

Step 2: Feed-forward:
- Calculate hidden nodes values
  
  \[ A^{(2)}_1 = g(\theta^{(1)}_{10} X_0 + \theta^{(1)}_{11} X_1 + \theta^{(1)}_{12} X_2) \]  
  \[ A^{(2)}_2 = g(\theta^{(1)}_{20} X_0 + \theta^{(1)}_{21} X_1 + \theta^{(1)}_{22} X_2) \]  

- Activate hidden nodes using the Sigmoid function
  
  \[ g(z) = \frac{1}{1 + e^{-z}} \]

Where: $z = \theta_0 X_0 + \theta_1 X_1 + \cdots + \theta_n X_n$

- Calculate output node values
  
  \[ A^{(3)}_1 = g(\theta^{(2)}_{10} A^{(2)}_0 + \theta^{(2)}_{11} A^{(2)}_1 + \theta^{(2)}_{12} A^{(2)}_2) \]

- Calculate cost function to determine adjustment of output according to the descriptors (error) and include the regularization parameter ($\lambda$) to reduce overfitting.

\[
J(\theta) = -\frac{1}{m} \sum_{i=1}^{m} y^{(i)} \log h_\theta(x^{(i)}) + (1 - y^{(i)}) \log (1 - h_\theta(x^{(i)})) + \frac{\lambda}{2m} \sum_{j=1}^{n} \theta_j^2
\]

Step 3: Backpropagation:
- Adjust weights ($\theta$) based on error measured between output of the network and input by applying the gradient descent algorithm to the cost function. This step was carried out using equations 1-6 to 1-9, based on Figure 2.
Figure 2. Backpropagation Example.

\[
\delta_{j}^{(3)} = A_{j}^{(3)} - Y_{j} \quad (1-6)
\]

\[
\delta_{j}^{(2)} = \theta^{(2)} \delta_{j}^{(3)} \cdot g'(\xi^{(2)}) \quad (1-7)
\]

Where:

- \( \delta_{j}^{(L)} = \text{error of node } j \text{ in } L \text{ layer} \)
- \( Y_{j} = \text{value of node } j \text{ in output layer} \)
- \( \xi^{(L)} = \begin{bmatrix} z_{1}^{(L)} & \ldots & z_{j}^{(L)} \end{bmatrix} \)

- Determine the gradient of the cost function

\[
\frac{\partial}{\partial \theta_{ij}^{(L)}} J(\theta) = \frac{1}{m} \Delta_{ij}^{(L)} \text{ if } j = 0 \quad (1-8)
\]

\[
\frac{\partial}{\partial \theta_{ij}^{(L)}} J(\theta) = \frac{1}{m} \left[ \Delta_{ij}^{(L)} + \lambda \theta_{ij}^{(L)} \right] \text{ if } j \neq 0 \quad (1-9)
\]

Where:

\[
\Delta_{ij}^{(L)} := \Delta_{ij}^{(L)} + A_{j}^{(L)} \delta_{i}^{(L+1)} \text{ (initialize } \Delta_{ij}^{(L)} = 0) \]

Validation Method

An exhaustive cross-validation was carried out by the leave-one-out method, in which one instance (molecular descriptors and experimental IC\textsubscript{50} of a single molecule) is used as a test set, while all other instances are used as a training set. This process is applied to each molecule of the set\textsuperscript{[5]} thus, the weight of each cross-validation iteration was determined, and the average of those weights was calculated to predict IC\textsubscript{50} values.

Now, the leave-one-out method was selected for validation because the error of estimation didn’t vary depending on the data used for the test set and validation, which indicated that the error of estimation was more stable in contrast to other cross-validation methods like k-fold or Montecarlo.\textsuperscript{[5]} Although
this method entails a greater computational cost than the other cross-validation methods mentioned, it is commonly used for small data sets such as the one used in this study (13 molecules).\[^5\]

**Validation Metrics**

The determination coefficient ($R^2$) was used to compare predicted and experimental IC$_{50}$ values and evaluate the goodness of fit of the model (equation 1-10). $R^2$ was used because it’s straightforward and it’s the most commonly used validation metrics for evaluating the goodness-of-fit of a model.\[^6\]

$$R^2 = \frac{\sum_{i=1}^{n}(\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2} \quad (1-10)$$

Where:

\[ \hat{Y} = \text{predicted IC}_{50} \]

\[ Y = \text{experimental IC}_{50} \]

\[ \bar{Y} = \text{Average of experimental IC}_{50} \text{ set} \]

**Prediction of IC$_{50}$**

Once the structure of the ANN was validated, 13 chloroquine derivatives with experimental IC$_{50}$ values were optimized in Avogadro\[^7\] using the MMFF94s force field with four steps per update. Next, the molecular descriptors of the derivatives were calculated with the software PaDEL,\[^8\] and a Pearson correlation among descriptors was carried out in MATLAB\[^9\] to identify those descriptors that presented a correlation value nearest to zero; descriptors with correlation values greater than 0.6 were dismissed. Then, a Pearson correlation between descriptors and experimental IC$_{50}$ values was performed to identify those descriptors that presented a correlation value closest to ±1. Subsequently, the chosen descriptors (ATSC4p, ZMIC0, MLFER_A) and the experimental IC$_{50}$ values were entered into the ANN, and the number of hidden nodes was defined according to the number that yielded the best $R^2$. In this case, 1600 nodes were necessary to yield the best determination coefficient ($R^2=0.899$).

Afterward, the molecular descriptors (ATSC4p, ZMIC0, MLFER_A) of 8 new potential SARS-Cov-2 Mpro inhibitors were calculated, and their IC$_{50}$ values were predicted (Table 1) with the previously described ANN.

**STRUCTURE-BASED VIRTUAL SCREENING (SBVS)**

After the IC$_{50}$ values of the 8 new compounds were predicted, SBVS was carried out to evaluate each compound’s binding affinity to Mpro bound to a non-covalent inhibitor (PDB 6W79).\[^10\] First, the non-covalent inhibitor was removed, and the unbound Mpro was optimized according to the AutoDockTools protocol.\[^11,12\] Then, the 13 chloroquine derivatives, and the 8 new potential SARS-Cov-2 Mpro inhibitors were docked with Mpro. Grid box size was adjusted to 34 x 36 x 28 points with a grid spacing of 0.375 Å, while the central position of the grid box was set at 16.278, -27.000 and 22.278. Calculations were done in triplicate and the average energy of the pose with the lowest RMSD value is shown in table 1.
**4-((7-chloroquinolin-4-yl)amino)phenol:** To a 100 mL round flask, were added 4,7-dichloroquinoline (495 mg, 2.5 mmol, 1 equiv.), aminophenol (273 mg, 2.5 mmol, 1 equiv.), hydrochloric acid (123 µL, 3.75 mmol, 1.5 equiv.) and 96% ethanol (20 mL), the mixture was stirred at reflux during 5 hour. Finally, 5 mL of an aqueous solution of sodium hydroxy (0.1M) was added to precipitate the free amine. The crude was purified using flash chromatography with AcCN/AcOEt (2:1) as eluent. The product was obtained as a yellow solid (400 mg, 1.48 mmol, 59%) mp 275–277 °C (Hexane).\(^{1}\)H NMR (400 MHz, Methanol-\(d_4\)) \(\delta\): 8.42 – 8.26 (m, 2H), 7.86 (d, \(J = 2.2\) Hz, 1H), 7.52 (d, \(J = 9.1\) Hz, 1H), 7.22 (d, \(J = 8.2\) Hz, 2H), 6.93 (d, \(J = 8.2\) Hz, 2H), 6.65 (d, \(J = 5.6\) Hz, 1H).\(^{13}\)C NMR (100 MHz, Methanol-\(d_4\)) \(\delta\): 155.8, 151.4, 151.0, 149.0, 135.0, 130.8, 126.7, 126.6, 125.1, 123.8, 117.8, 116.2, 100.5. HRMS calcd for C\(_{15}\)H\(_{12}\)ClN\(_2\)O (M + H) 271.0560, found 271.0559

**Reference:** Molecules 2013, 18, 10648-10670.
Analysis Report

<Sample Information>
Sample Name: 20-300nm
Sample ID: patron
Data Filename: FYF27.lcd
Method Filename: COVID.lcm
Batch Filename:
Vial #: 1-25
Injection Volume: 10 uL
Date Acquired: 26/01/2021 11:30:41
Acquired by: System Administrator
Date Processed: 26/01/2021 11:48:10
Processed by: System Administrator
Sample Type: Unknown

<Chromatogram>
Detector A Channel 1 300nm

mV

Detector A Channel 2 254nm

mV

<Peak Table>
Detector A Channel 1 300nm

C:\Users\Usuario\Desktop\INQAIFYF27.lcd
References

[1] R. Jenjob, T. Phakkeeree, F. Seidi, M. Theerasilp, D. Crespy, *Macromol. Biosci.* 2019, 19, 1900063.

[2] J. O’Brien, I. Wilson, T. Orton, F. Pognan, *Eur. J. Chem.* 2000, 267, 5421–5426.

[3] B. Drasler, P. Sayre, K. G. Steinhäuser, A. Petri-Fink, B. Rothen-Rutishauser, *NanoImpact* 2017, 8, 99–116.

[4] E. Beqari, “A Very Basic Introduction to Feed-Forward Neural Networks,” can be found under https://dzone.com/articles/the-very-basic-introduction-to-feed-forward-neural, 2018.

[5] M. A. Little, G. Varoquaux, S. Saeb, L. Lonini, A. Jayaraman, D. C. Mohr, K. P. Kording, *Gigascience* 2017, 6, 1–6.

[6] K. Roy, R. N. Das, P. Ambure, R. B. Aher, *Chemom. Intell. Lab. Syst.* 2016, 152, 18–33.

[7] M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek, G. R. Hutchison, *J. Cheminform.* 2012, 4, 17.

[8] C. W. Yap, *J. Comput. Chem* 2011, 32, 1466–1474.

[9] The MathWorks Inc., MATLAB, 2010.

[10] A. D. Mesecar, “6W63: Structure of COVID-19 main protease bound to potent broad-spectrum non-covalent inhibitor X77,” DOI 10.2210/pdb6w63/pdb can be found under https://www.rcsb.org/structure/6W63, 2020.

[11] R. Huey, G. M. Morris, S. Forli, *The Scripps Research Institute* 2012, 32.

[12] O. Trott, A. J. Olson, *J. Comput. Chem* 2010, 31, 455–461.