

Evaluation of predictive machine learning models for drug repurposing against delta variant of SARS-CoV-2 spike protein

Sudipta Dash^{1*}, Ishani Ishani², Dibyajit Lahiri³ & Moupriya Nag³

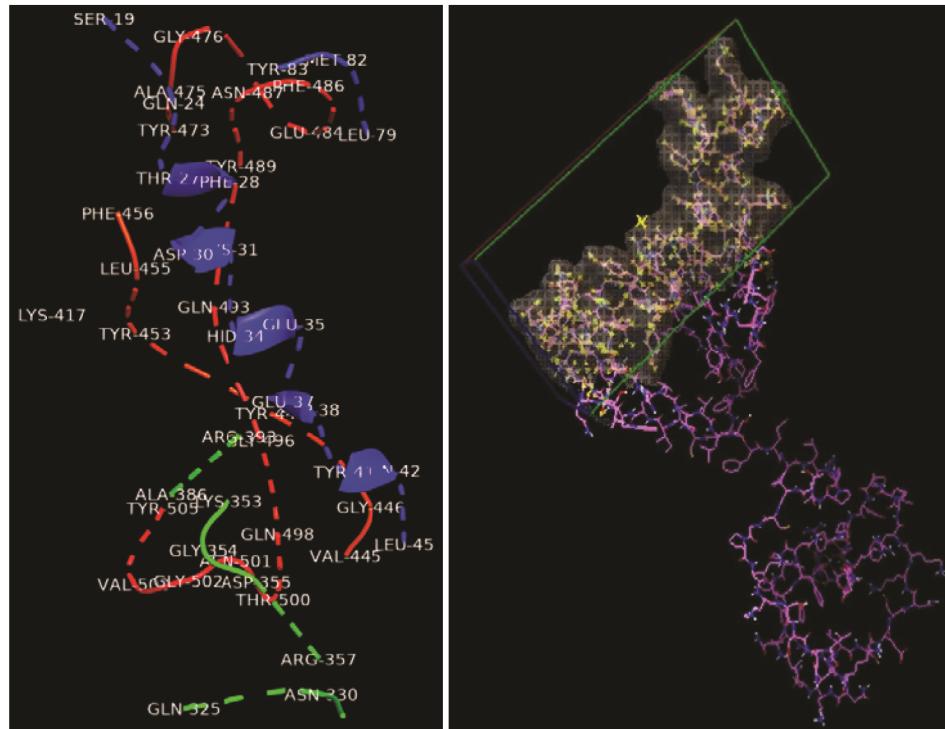
¹Department of Biotechnology, Indian Institute of Technology Kharagpur, Kharagpur-721 302, West Bengal, India

²Department of Life Sciences, School of Natural Sciences, Shiv Nadar University, Greater Noida-201 314, Uttar Pradesh, India

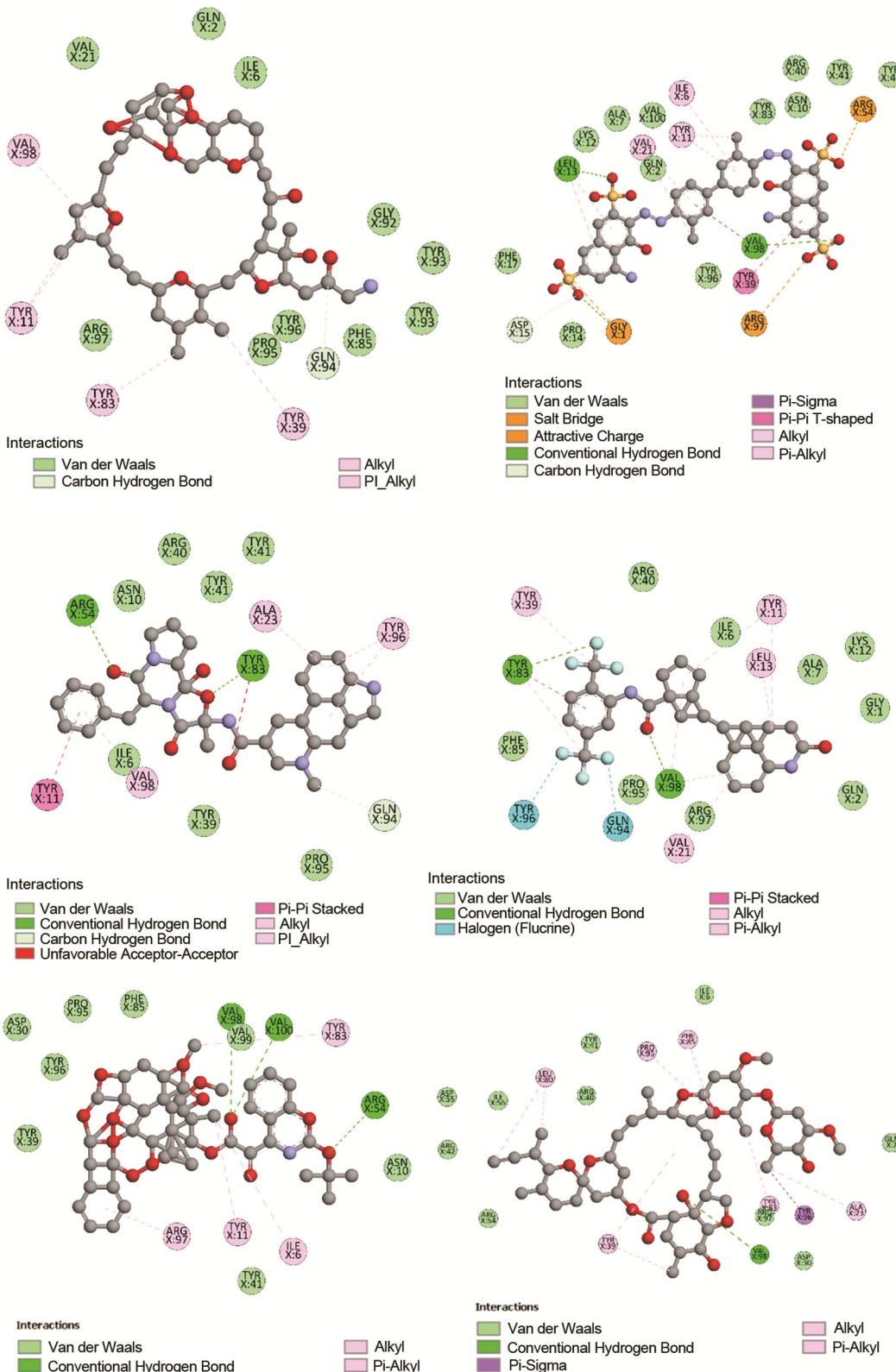
³Department of Biotechnology, University of Engineering and Management, Kolkata-700 156, West Bengal, India

Received 05 November 2021; revised 26 August 2022

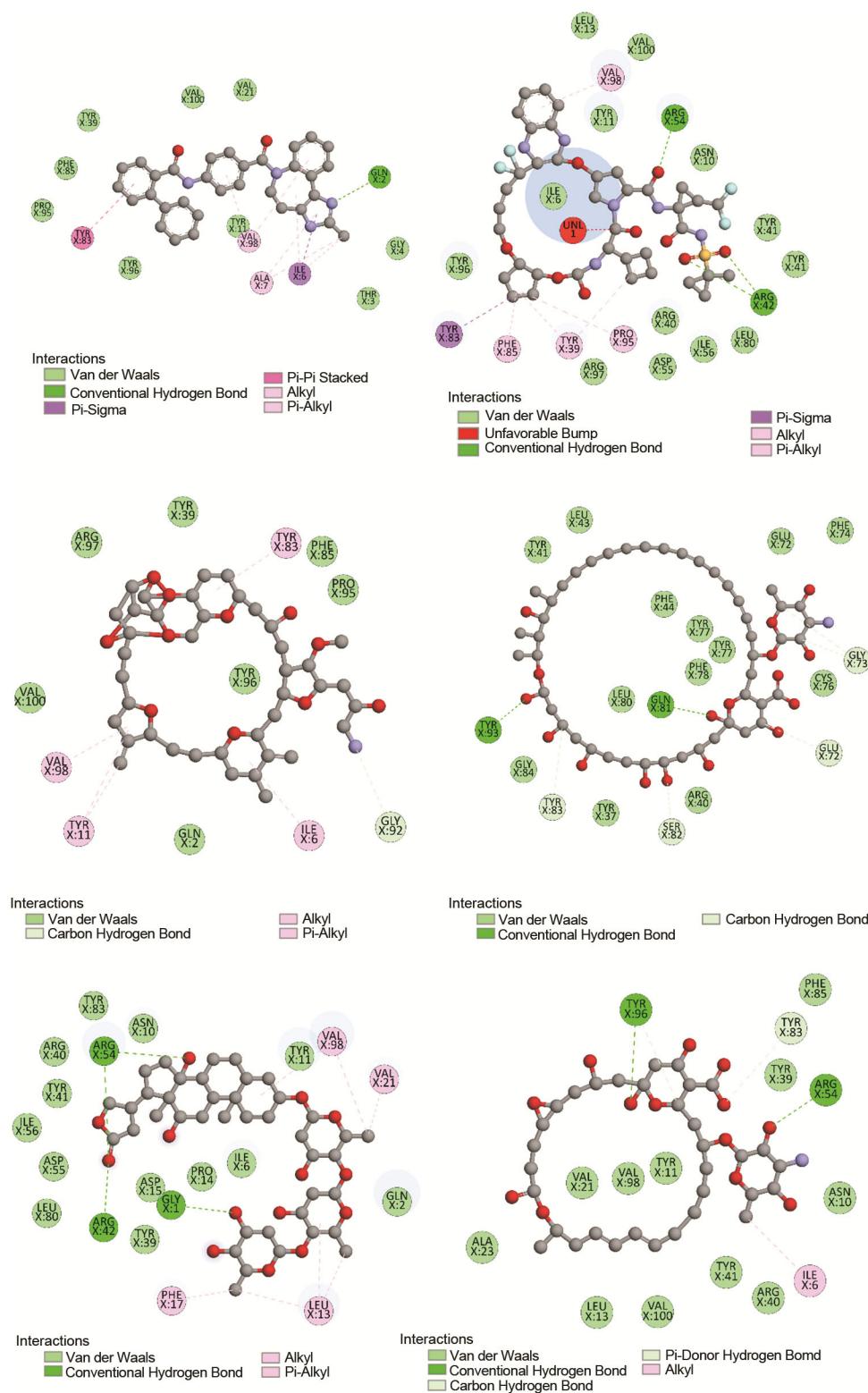
Supplementary Data



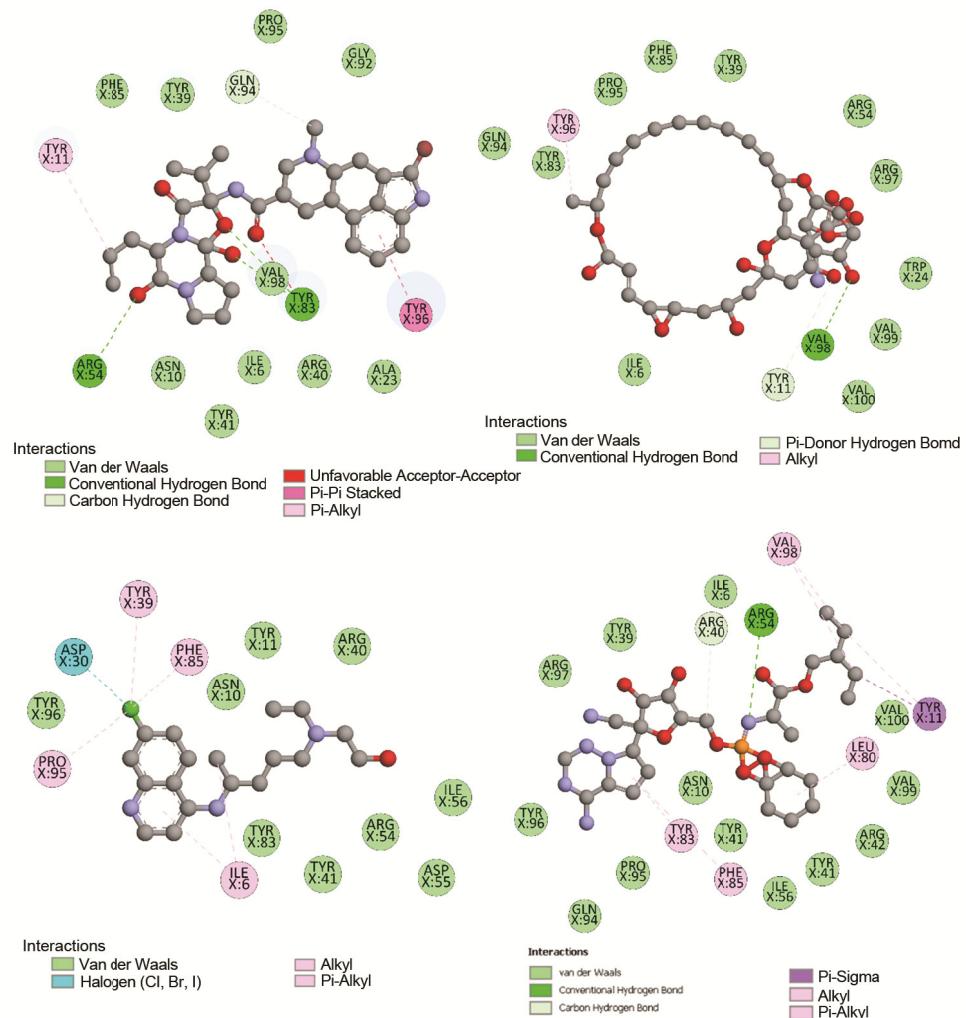
Suppl. Fig. 1 — (A) Predicted active site residues; and (B) Docking grid generation by Autodock Vina on the modelled spike protein. Docking Grid Dimensions: Grid box = $56 \times 60 \times 112$ with a spacing of 0.375. The binding pocket was set at $x = 78.999$, $y = -17.378$, and $z = -13.748$



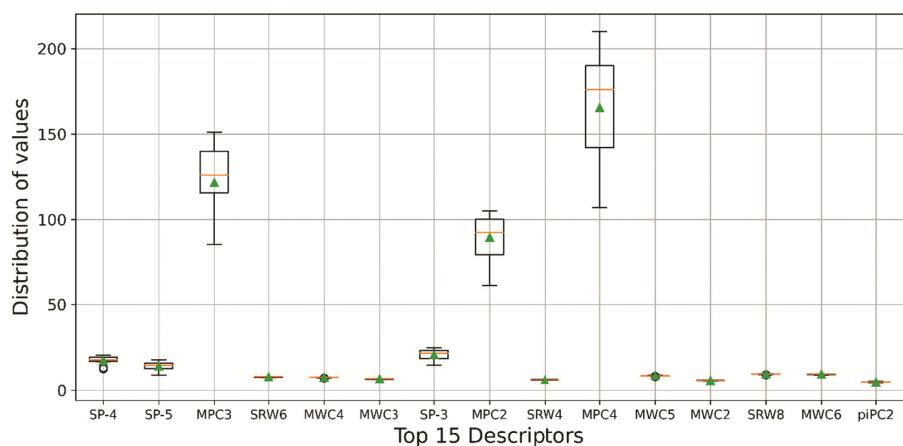
(Contd.).



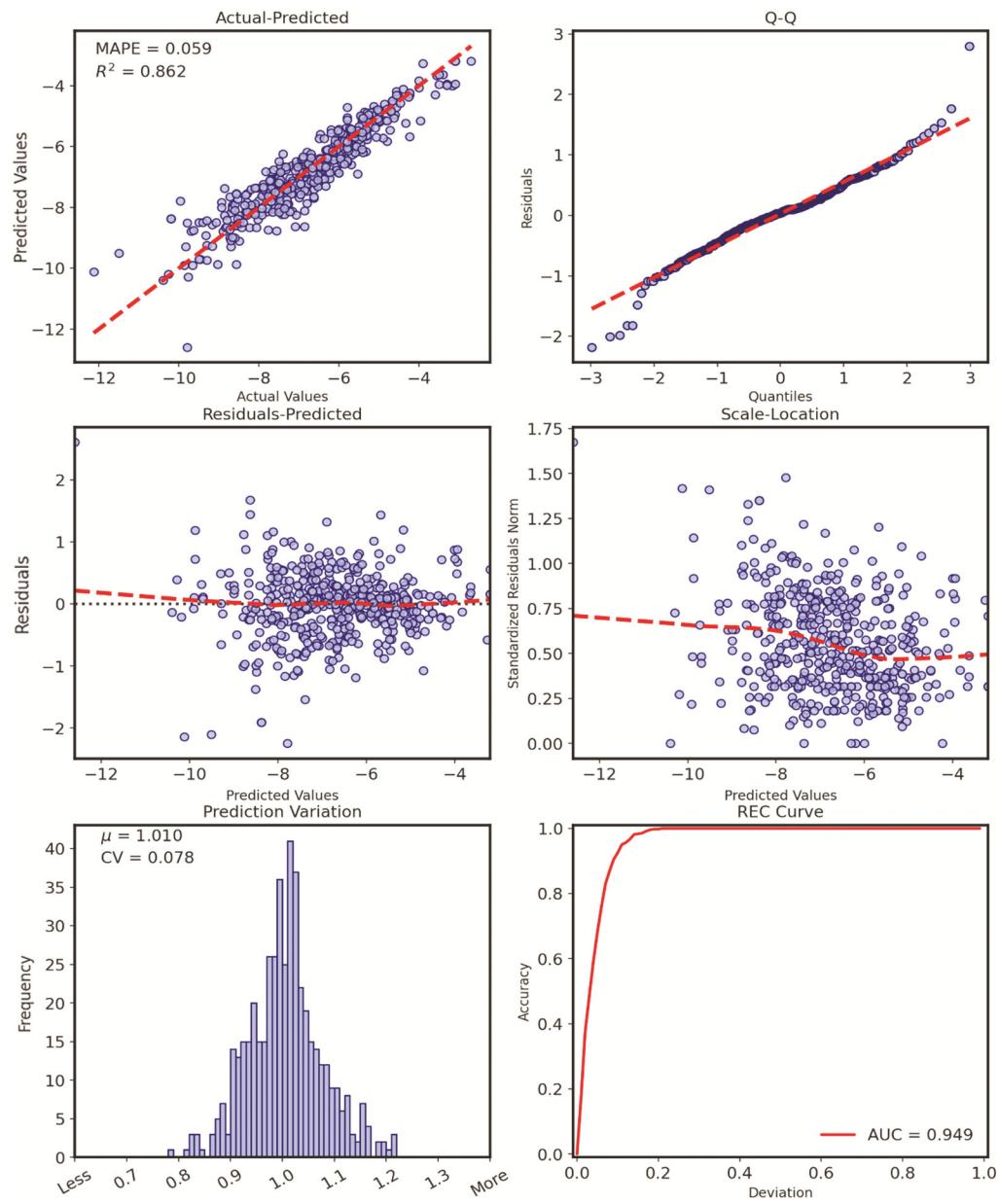
(Contd).



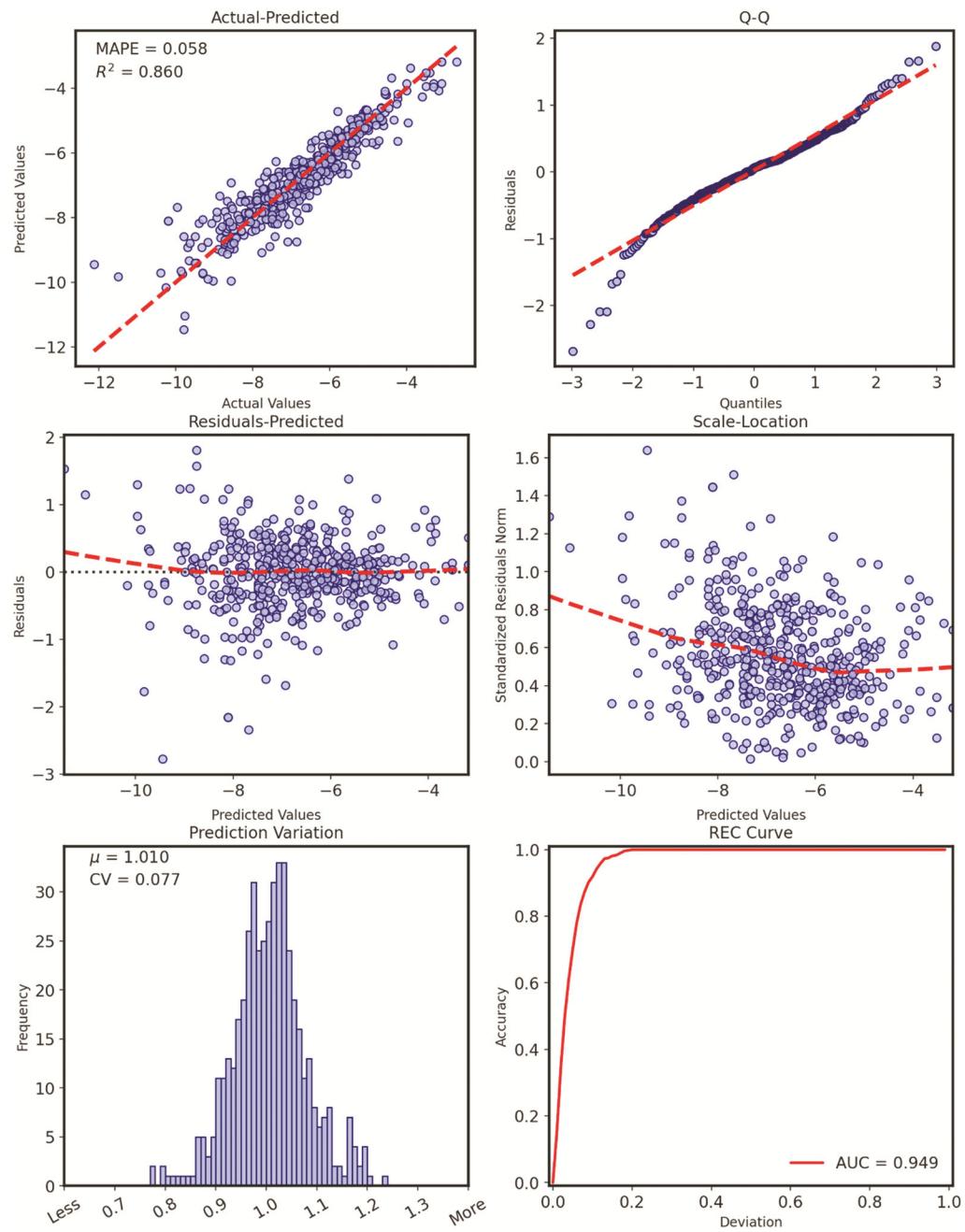
Suppl. Fig. 2 (A-P) — 2D interaction diagram of top 7-20 potential drugs that can inhibit the binding of spike protein and hACE2 protein (A) ZINC000169621219, (B) ZINC000169289767, (C) ZINC000003978005, (D) ZINC000003932831, (E) ZINC000085536932, (F) ZINC000253630390, (G) ZINC000012503187, (H) ZINC000164528615, (I) ZINC000169344691, (J) ZINC000253387843, (K) ZINC000242548690, (L) ZINC000169621220, (M) ZINC000053683151, (N) ZINC000008220909 (O) Hydroxychloroquine (P) Remdesivir



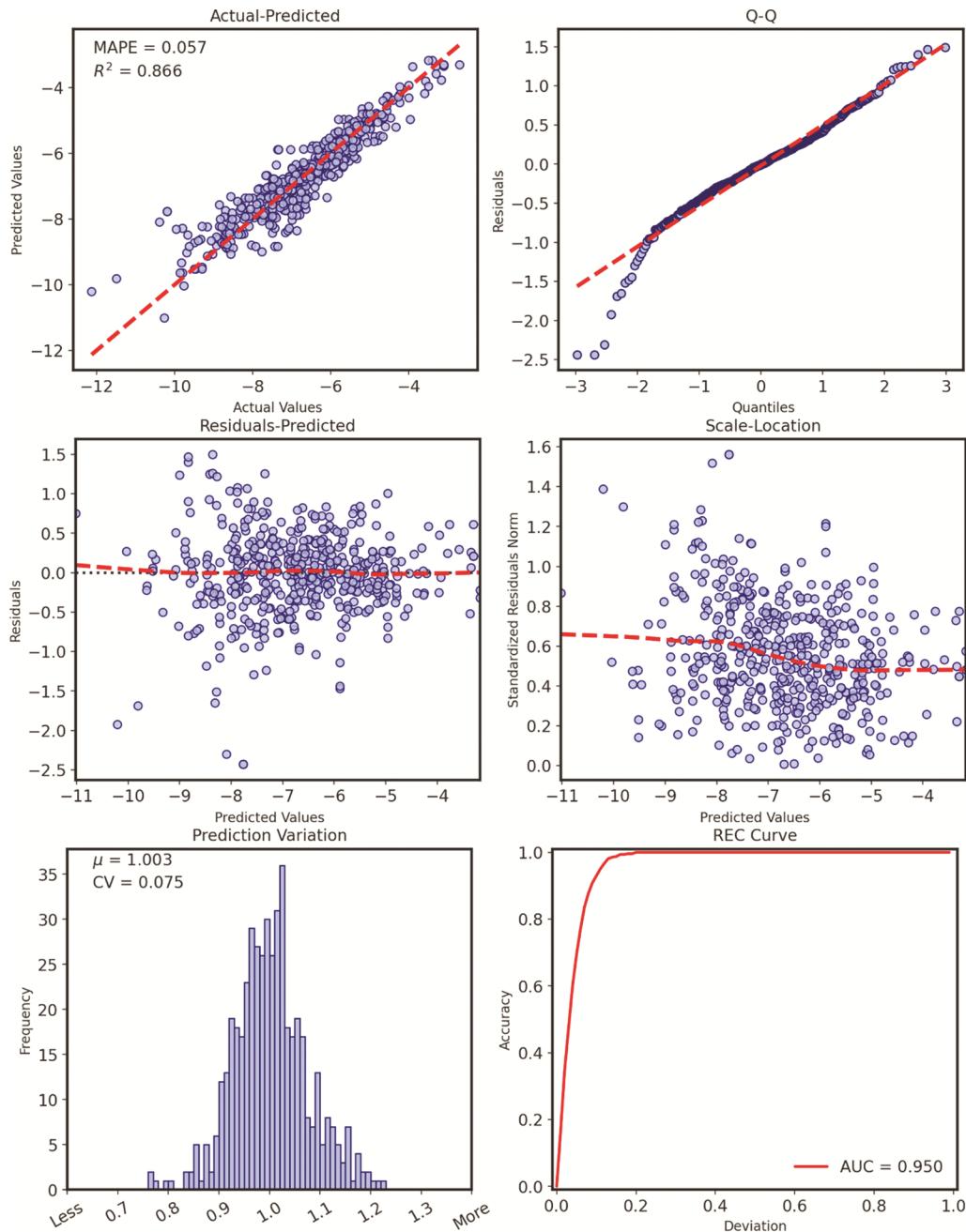
Suppl. Fig. 3 — Graph of the top 15 molecular descriptors and their range of values for the top 20 potential drugs/inhibitors



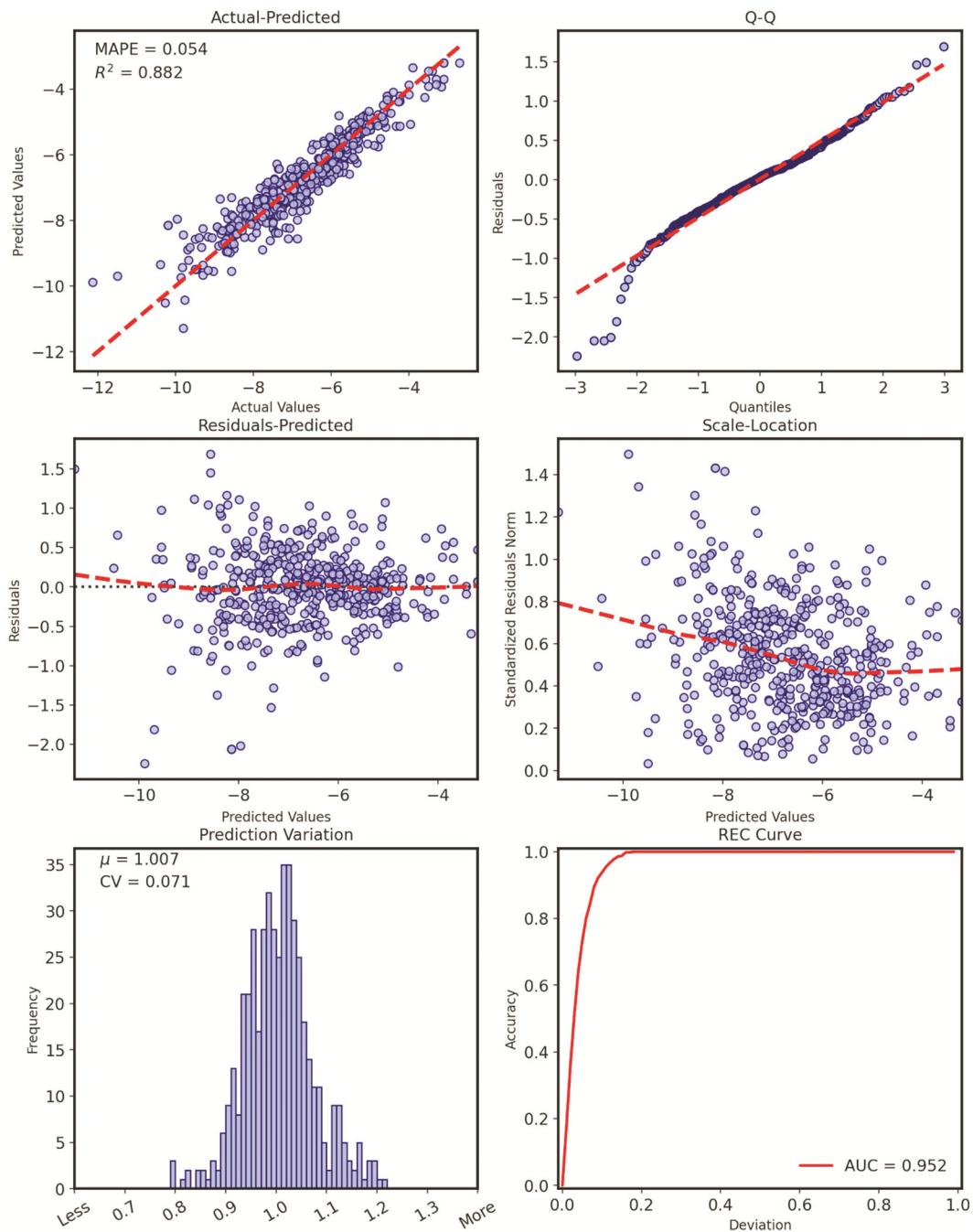
Suppl. Fig. 4 — SlickML evaluation results for K-Nearest Neighbor algorithm



Suppl. Fig. 5 — SlickML evaluation results for Random Forest algorithm



Suppl. Fig. 6 — SlickML evaluation results for Support Vector Machine algorithm



Suppl. Fig. 7 — SlickML evaluation results for Ensemble Stacking method