Supplementary Information - Automated Discovery of Noncovalent Inhibitors of SARS-CoV-2 Main Protease by Consensus Deep Docking of 40 Billion Small Molecules

Francesco Gentile[a], Michael Fernandez[a], Fuqiang Ban[a], Anh-Tien Ton[a], Hazem Mslati[a], Carl F. Perez[a], Eric Leblanc[a], Jean Charle Yaacoub[a], James Gleave[a], Abraham Stern[b], Bill Wong[c], François Jean[d], Natalie Strynadka[e], and Artem Cherkasov[a]*

[a]*Vancouver Prostate Centre, Department of Urologic Sciences, The University of British Columbia, Vancouver, BC, Canada*

[b]*NVIDIA Corporation, Santa Clara, CA, United States*

[c]*Dell Canada, North York, ON, Canada*

[d]*Department of Microbiology and Immunology, The University of British Columbia, Vancouver, BC, Canada*

[e]*Department of Biochemistry and Molecular Biology, The University of British Columbia, Vancouver, BC, Canada*

**Corresponding author: Vancouver Prostate Centre, 2660 Oak Street, Vancouver, British Columbia, V6H3Z6, Canada, acherkasov@prostatecentre.com*

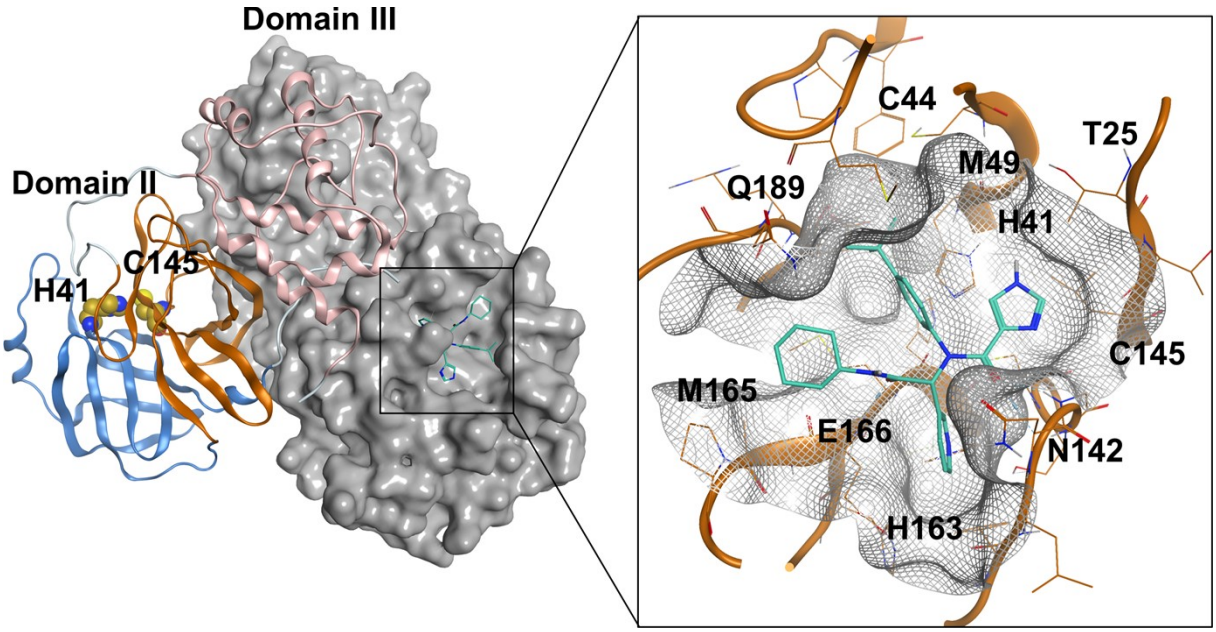

Figure S1. Dimeric structure of SARS-CoV-2 Mpro bound to X77 (from PDB:6W63), with a focus view of the active site.

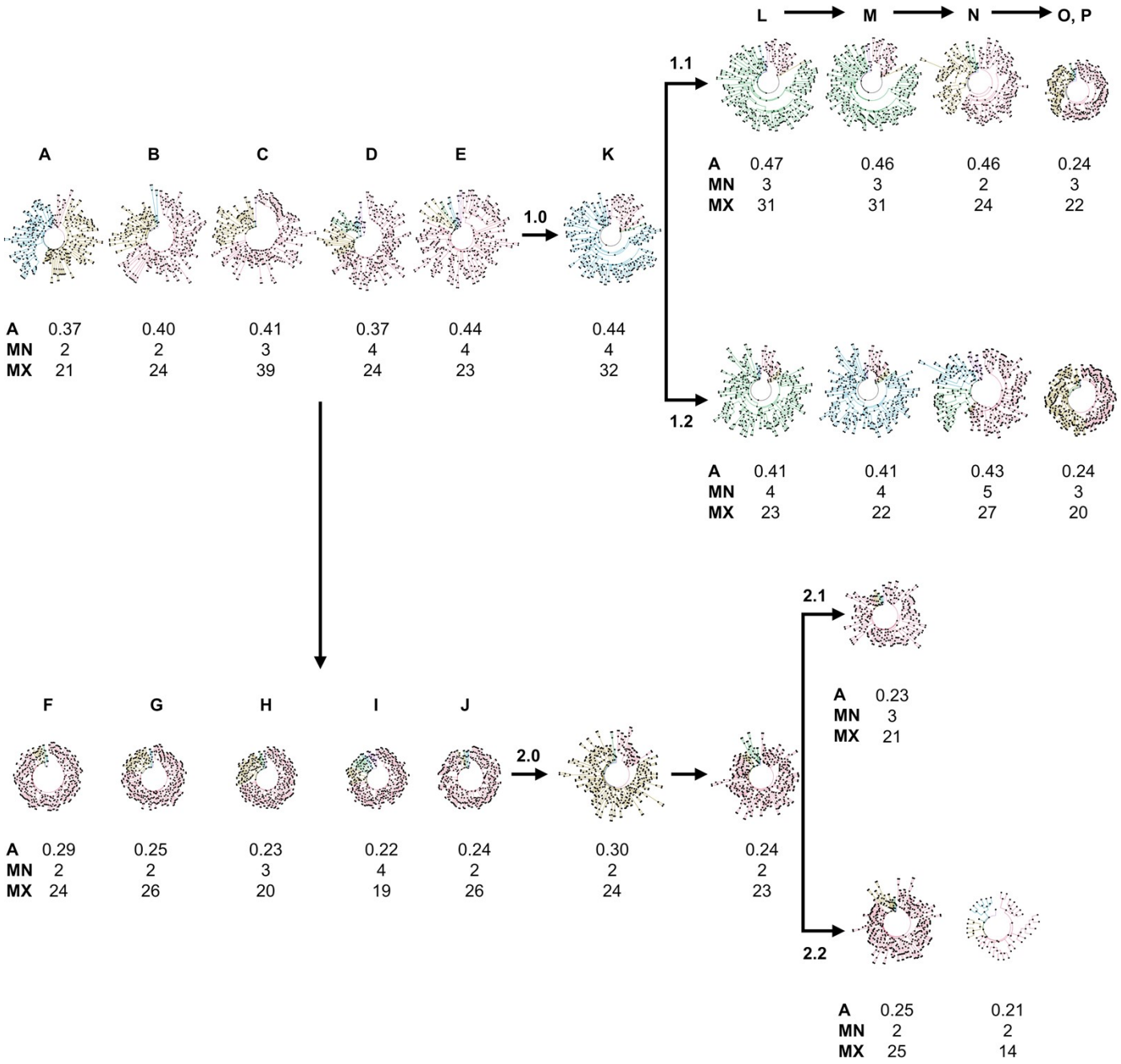

Figure S2. Tanimoto-based hierarchical agglomerative clustering of top-100 compounds returned by each selection strategy. Below each circular dendrogram, the corresponding agglomerative coefficient (A), minimum (MN) and maximum (MX) number of nodes are reported.

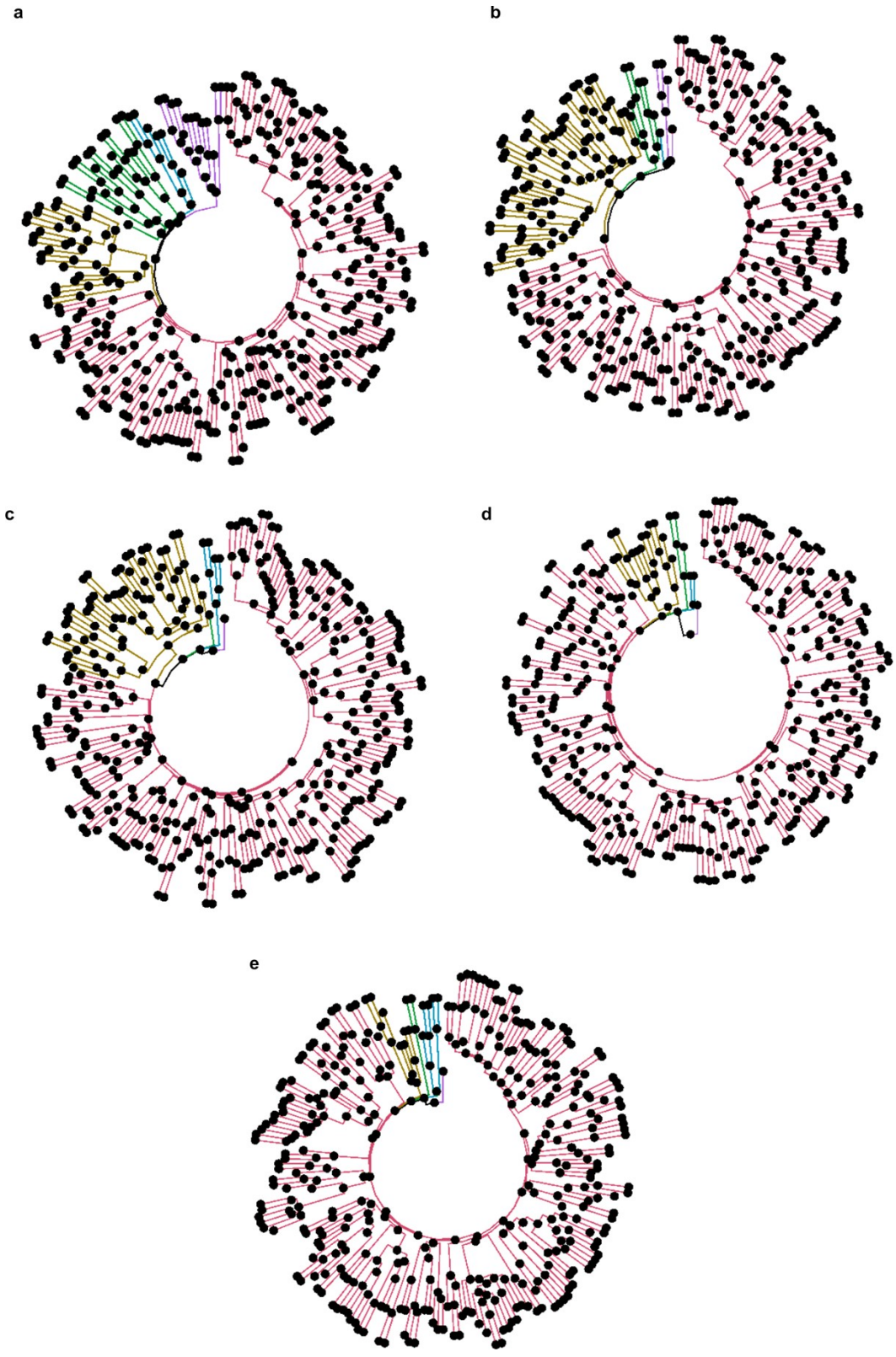
a
b
c
d
e

Figure S3. Tanimoto similarity trees of top-100 compounds from individual docking programs after chemical diversity selection, where the top-five clusters are highlighted by different colors in each dendrogram. **a**, Autodock GPU, **b**, Glide SP, **c**, FRED, **d**, ICM and **e**, QuickVina2.

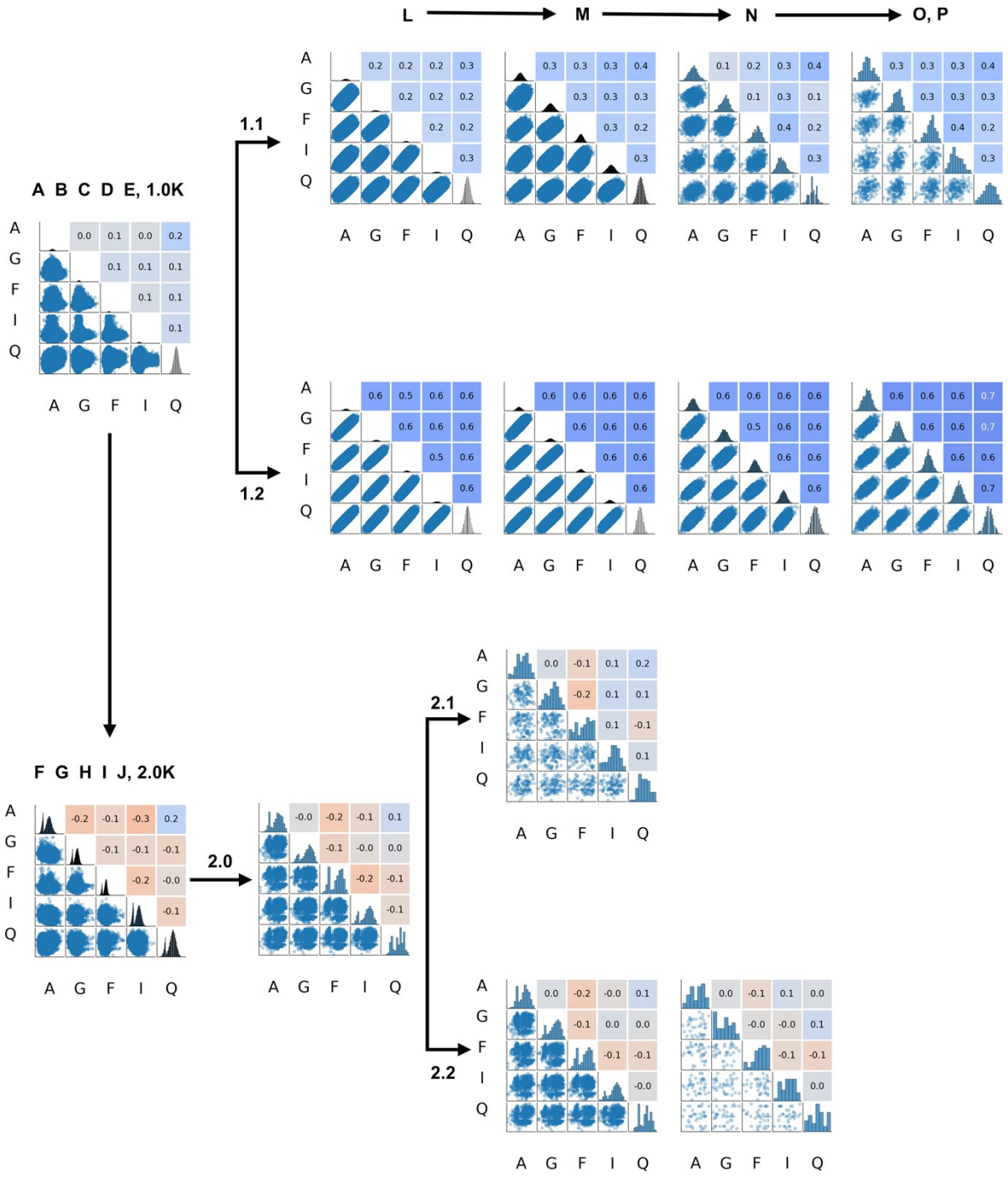

Figure S4. Correlation of docking scores for different hit selection strategies. Score distributions, correlation plots and Pearson coefficients between individual docking scores for the twenty-six filtering strategies.

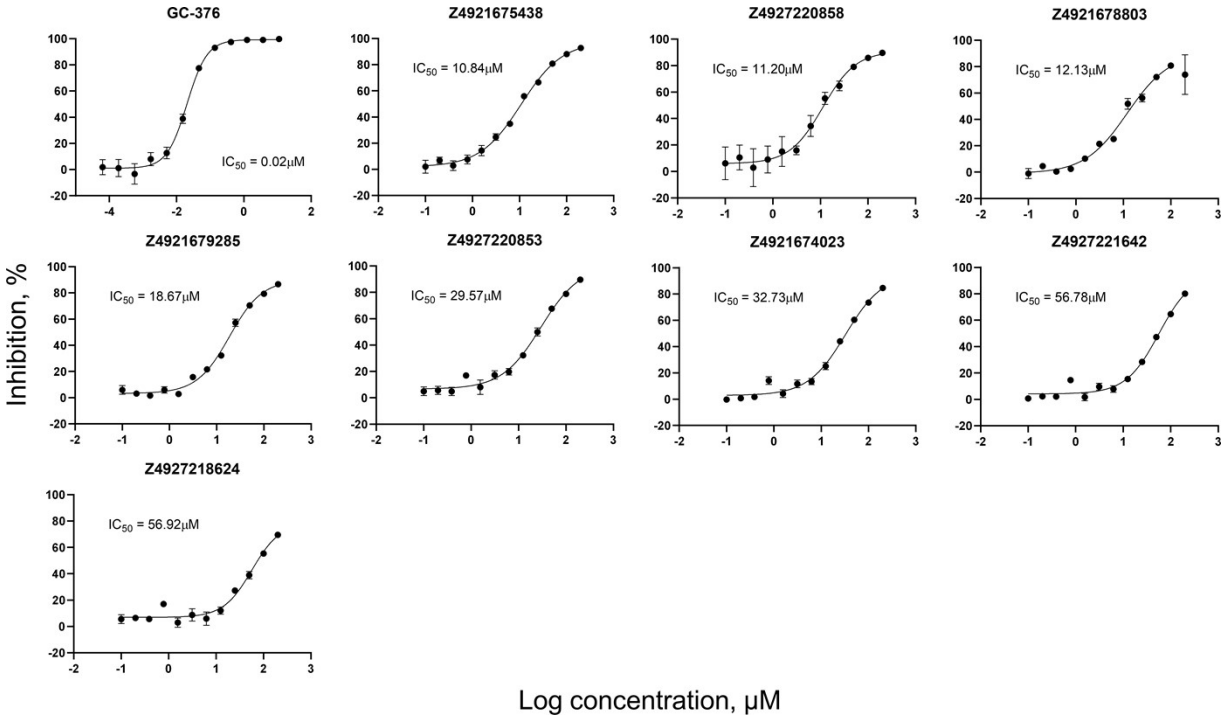

Figure S5. Dose-response curves and measured IC_{50} for control compound (GC-376) and eight identified inhibitors with IC_{50} values below 100 μM.