
Figures and figure supplements

NICEdrug.ch, a workflow for rational drug design and systems-level analysis of drug metabolism

Homa MohammadiPeyhani et al

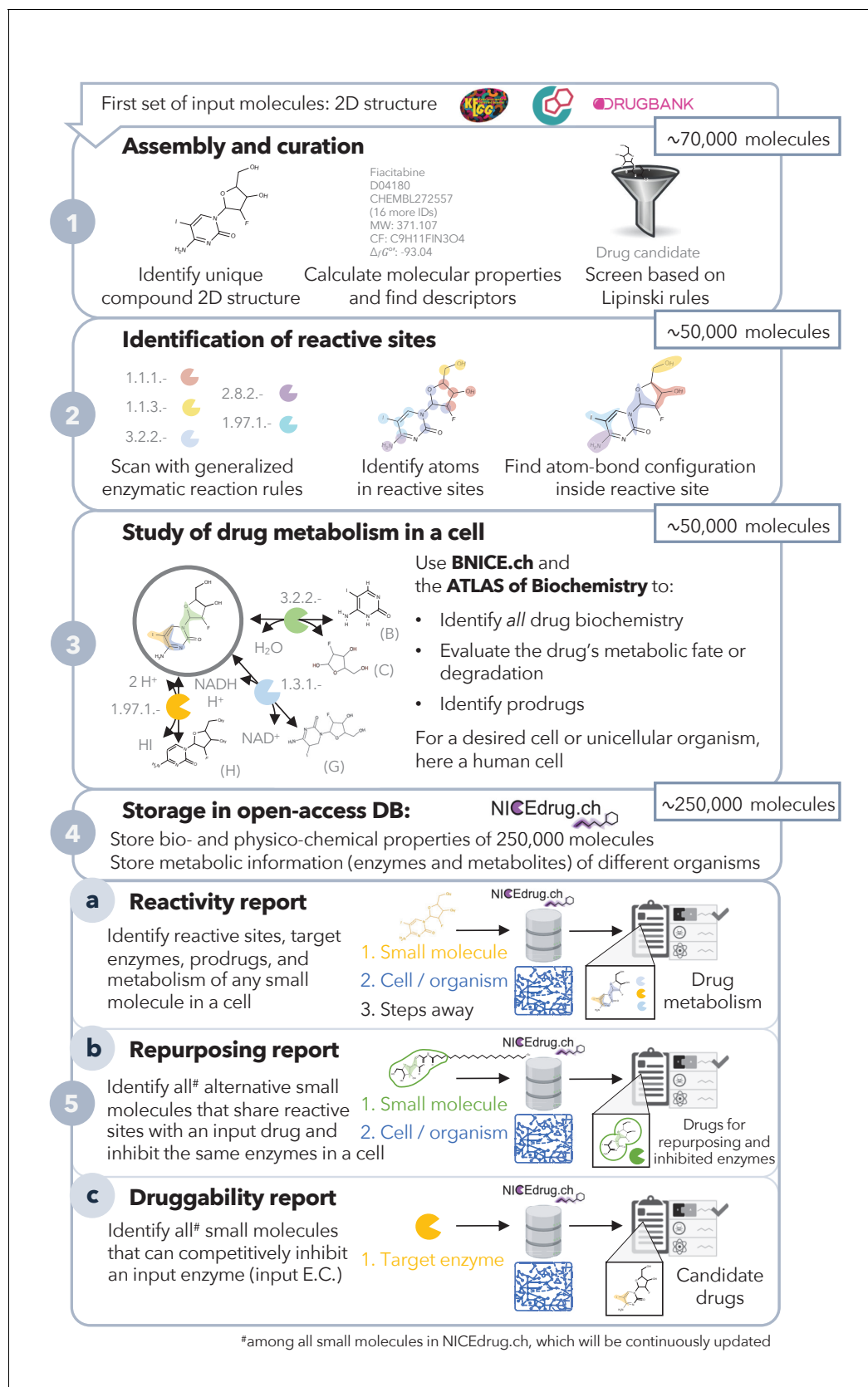


Figure 1. Pipeline to construct and use the NICEdrug.ch database. NICEdrug.ch (1) curates available information and calculates the properties of an input compound; (2) identifies the reactive sites of that compound; (3) explores the hypothetical metabolism of the compound in a cell; (4) stores all

Figure 1 continued on next page

Figure 1 continued

functional, reactive, bio-, and physico-chemical properties in open-source database; and (5) allows generation of reports to evaluate (5a) reactivity of a small molecule, (5b) drug repurposing, and (5c) druggability of an enzymatic target. See also **Figure 1—figure supplement 1**; **Figure 1—figure supplement 2**; **Figure 1—figure supplement 3**, and **Supplementary file 1**.

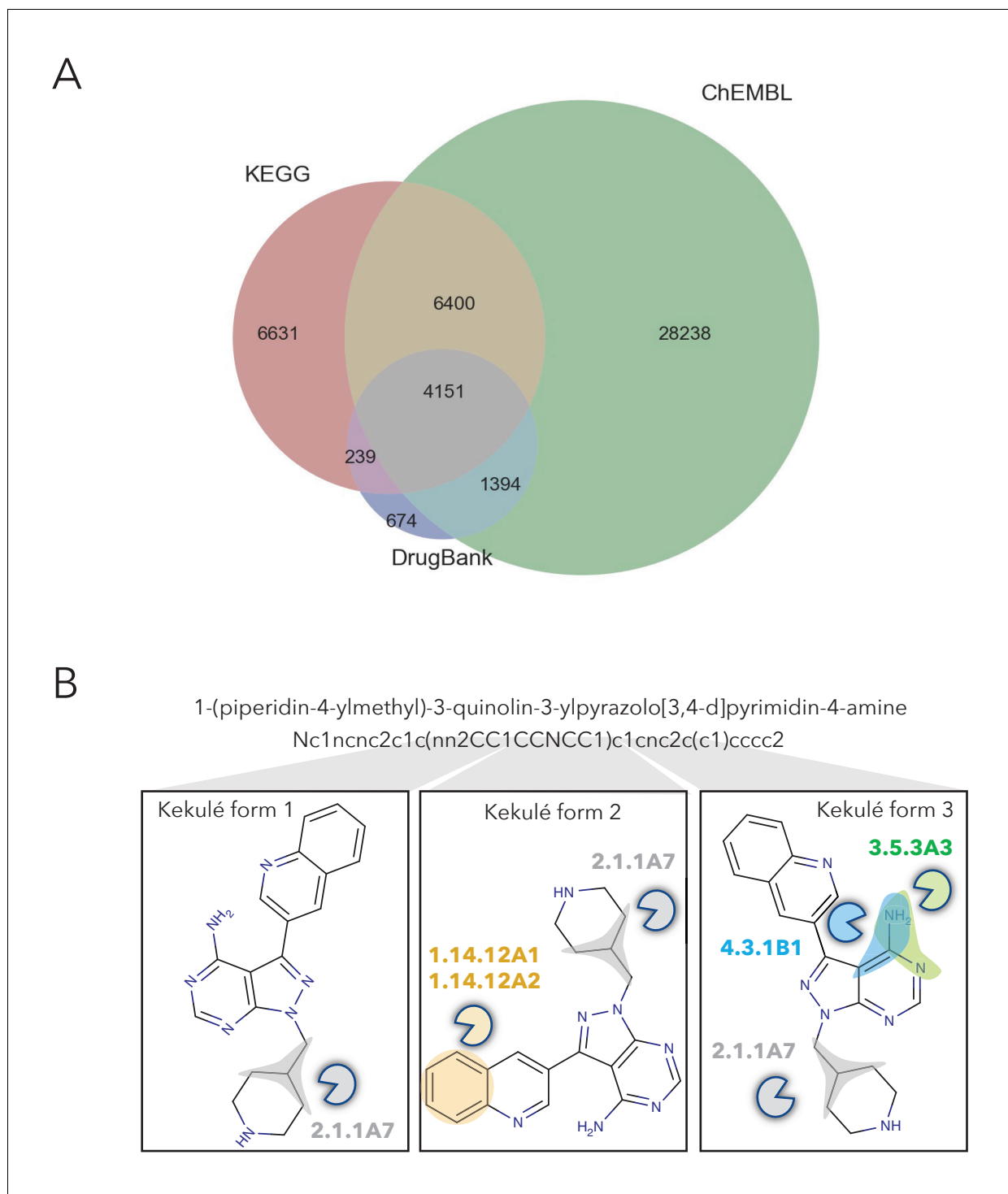


Figure 1—figure supplement 1. Overview of number of molecules in NICEdrug.ch and their structural curation. **(A)** Venn diagram showing the number of compounds in NICEdrug.ch and their source database: KEGG, DrugBank, ChEMBL NTD, and ChEMBL. **(B)** Representation on how different kekulé forms affect the identification of reactive sites and prediction of biological activity for an example molecule.

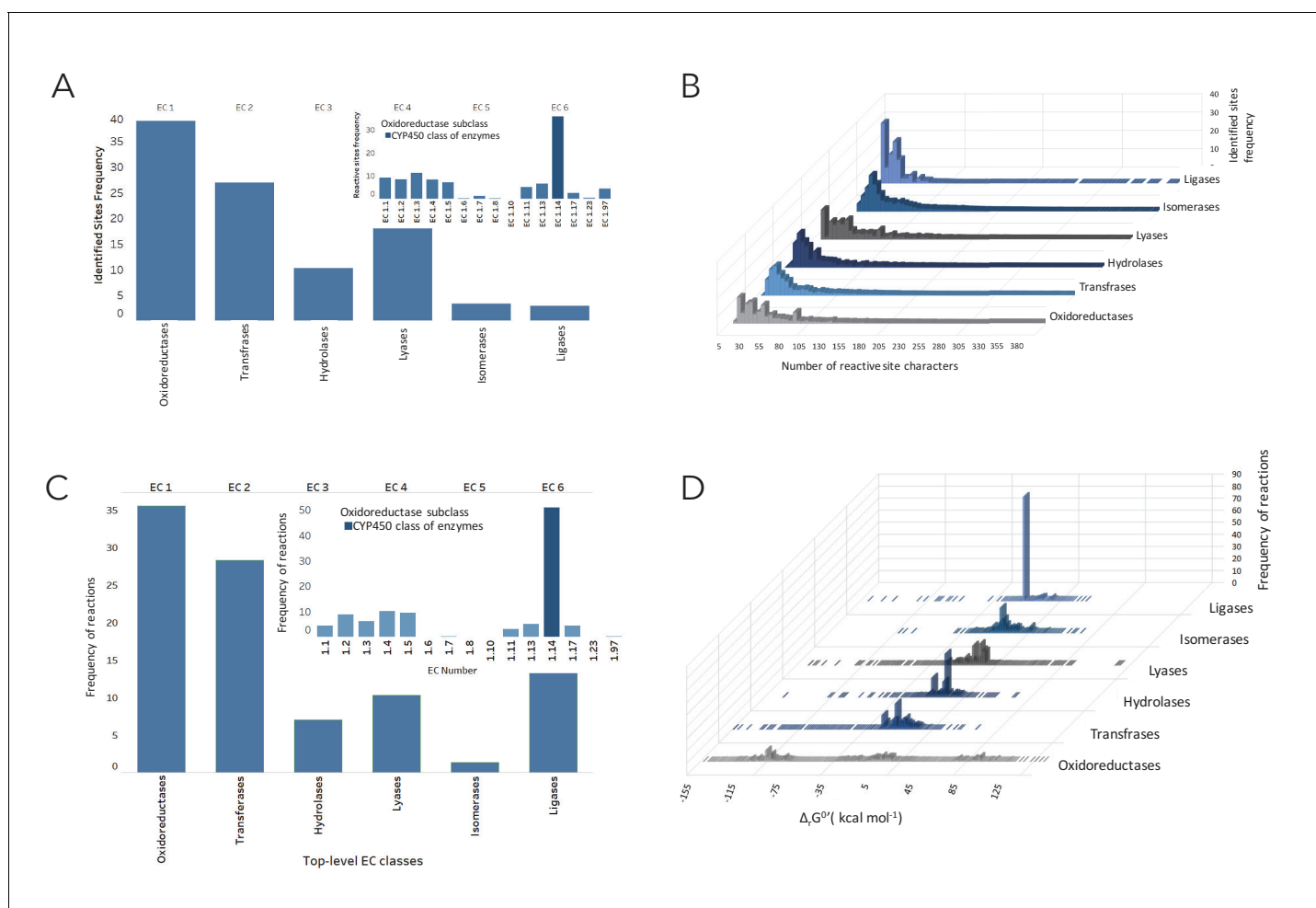


Figure 1—figure supplement 2. Distribution of reactive sites and metabolic reactions as of EC numbers linked to all molecules in NICEdrug.ch. (A) Distribution of reactive sites identified in all molecules of NICEdrug.ch among classes of EC numbers. (B) Specificity of reactive sites identified in drugs based on length and types of participating atoms. (C) Distribution of drug metabolic reactions based on class of EC number. (D) Distribution of Gibbs free energy for the drug metabolic reactions, which are the reactions linked to all molecules of NICEdrug.ch.

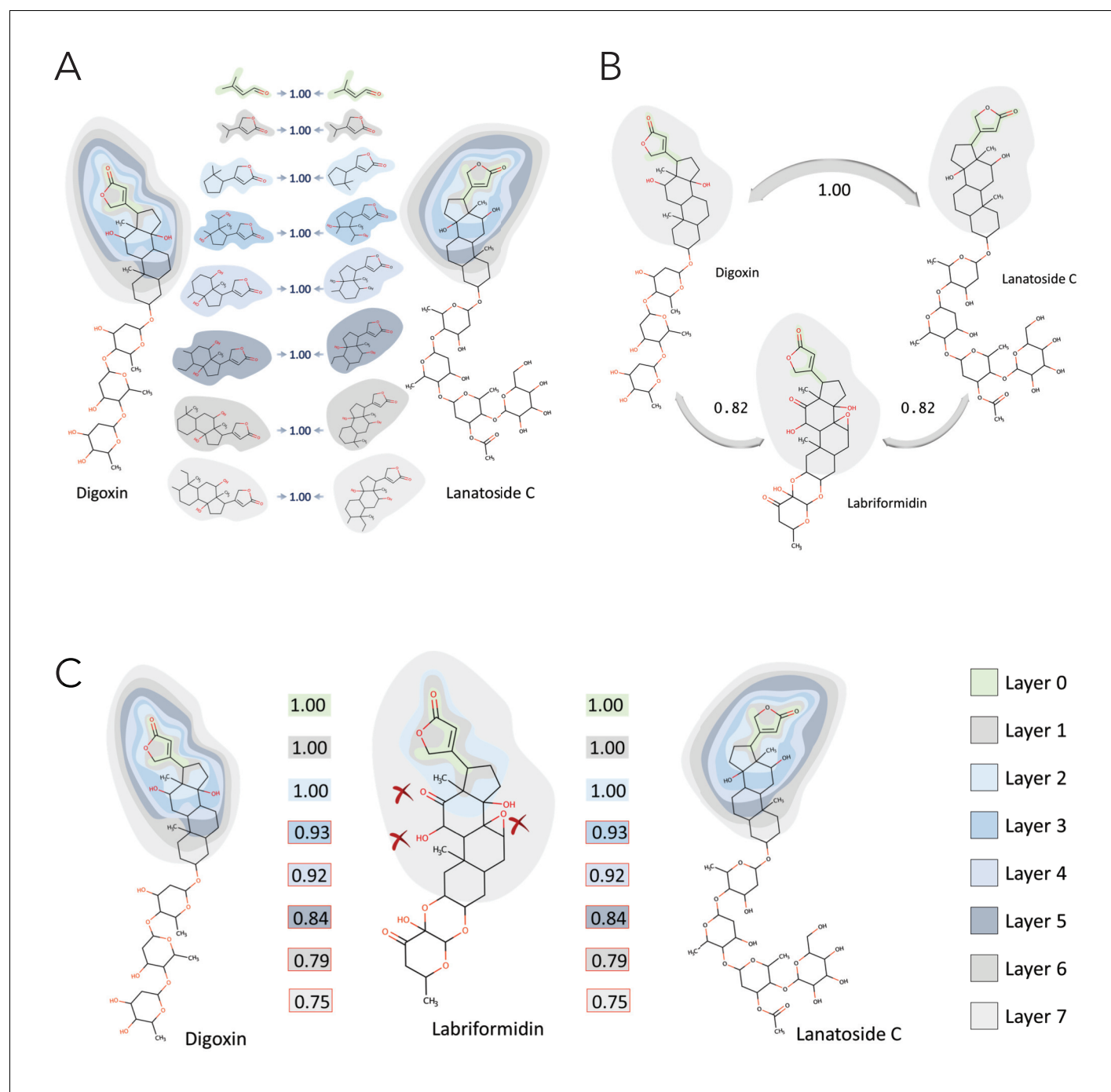


Figure 1—figure supplement 3. Description of NICEdrug score. Example of NICEdrug score calculation. The NICEdrug score takes into account the structure of a molecule's reactive site and its seven-atom-away neighborhood for similarity evaluation, analogous to BridgIT.

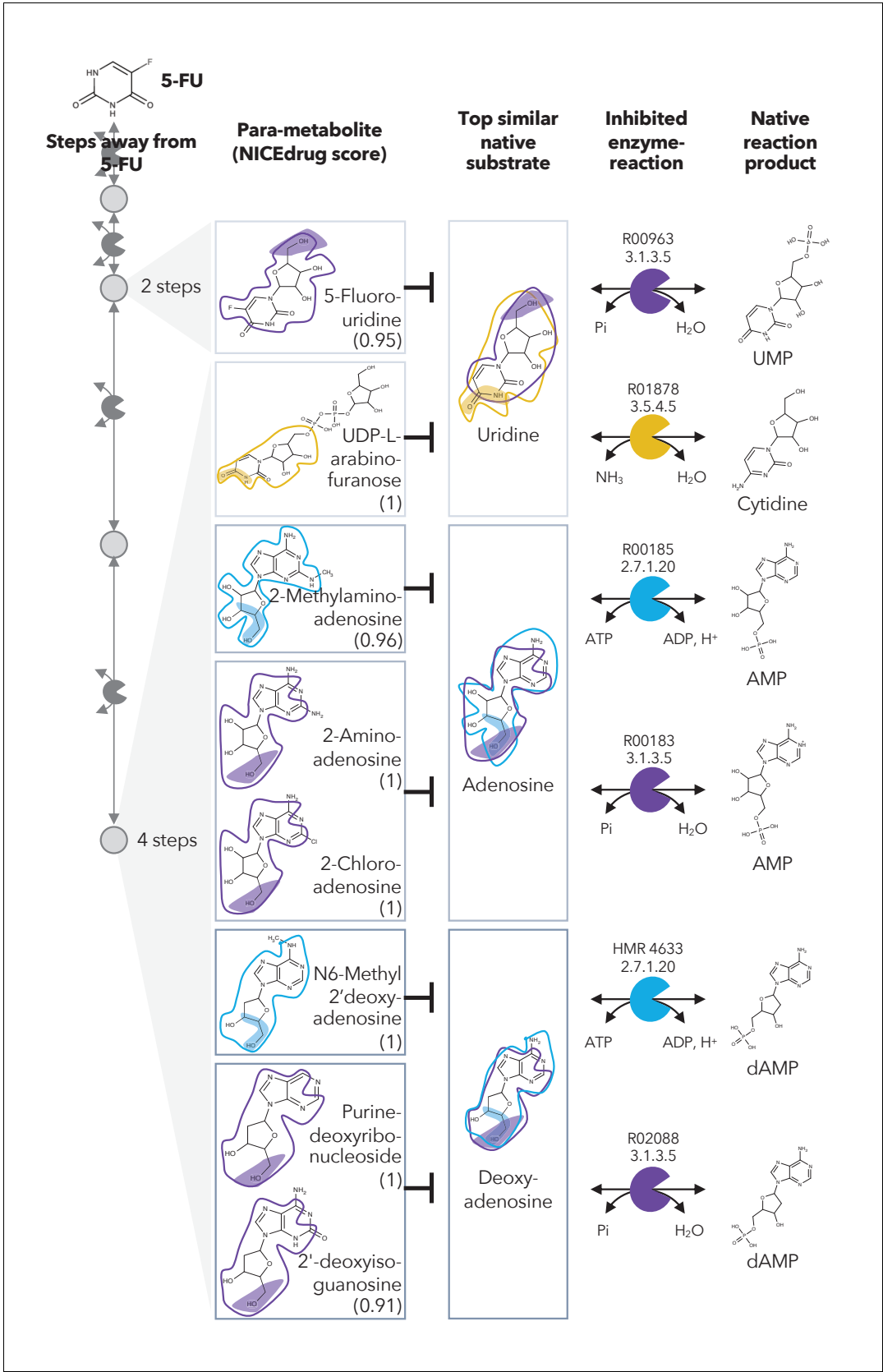


Figure 2. Similarity in reactive site and neighborhood defines para-metabolites in 5-FU metabolism and inhibited human metabolic enzymes. Eight para-metabolites in the 5-FU metabolic neighborhood (represented as defined in 'Materials and methods'). We show the most similar native human metabolites, inhibited enzymes, and native products of the reactions. See also **Supplementary file 3**.

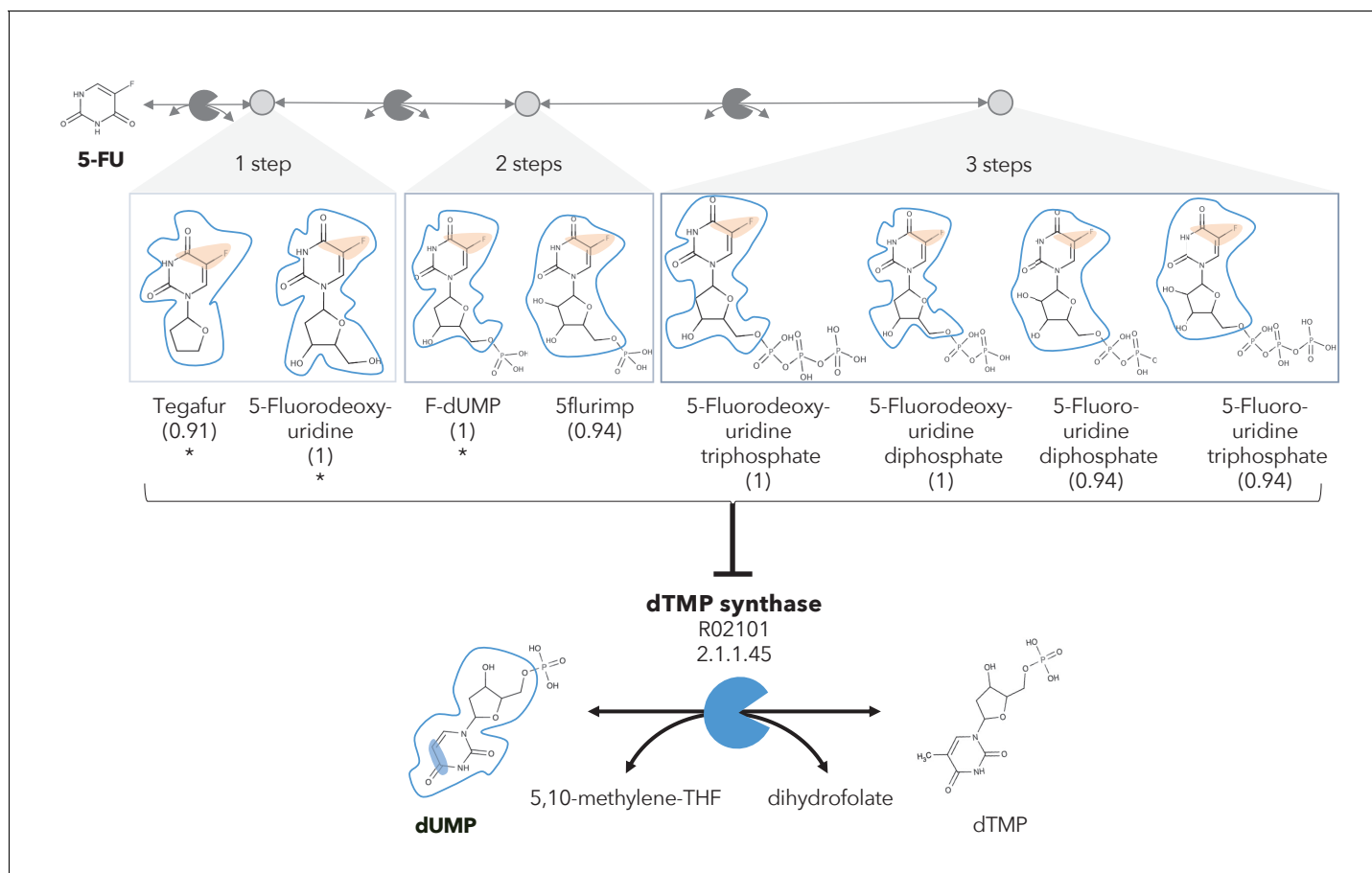


Figure 3. A different reactive site but similar neighborhood defines top anti-metabolites in 5-FU metabolism and inhibited human metabolic enzyme. Eight anti-metabolites of dUMP in the 5-FU metabolic neighborhood (represented as defined in 'Materials and methods'). Note that the reactive site of the anti-metabolites is different than the one of the native human metabolite, but the neighborhood is highly similar, which determines the high NICEdrug score (value in parenthesis). We show the inhibited human enzyme (dTMP synthase) and reaction, and its native product. See also *Supplementary file 3*

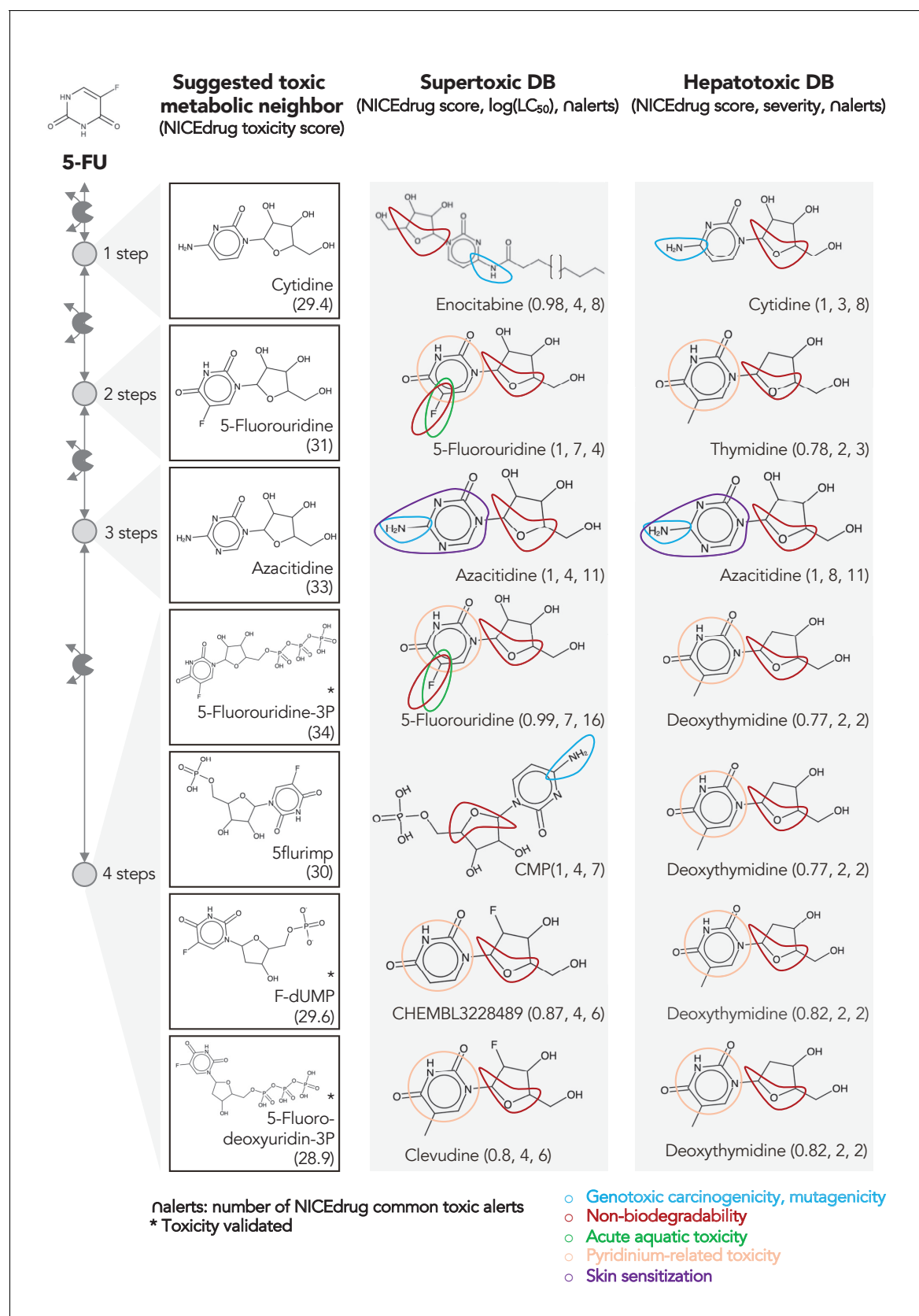


Figure 4. Comparing downstream products to known toxic molecules and analyzing their common structural toxic alerts explains metabolic toxicity of 5-FU. Example of six suggested toxic molecules in the 5-FU metabolic neighborhood (represented as defined in 'Materials and methods'). We show Figure 4 continued on next page

Figure 4 continued

toxic compounds from the supertoxic and hepatotoxic databases that lead to the highest NICEdrug toxicity score (number under toxic intermediate name, 'Materials and methods'). We highlight functional groups linked to five NICEdrug toxic alerts (legend bottom right). See also **Supplementary file 3**.

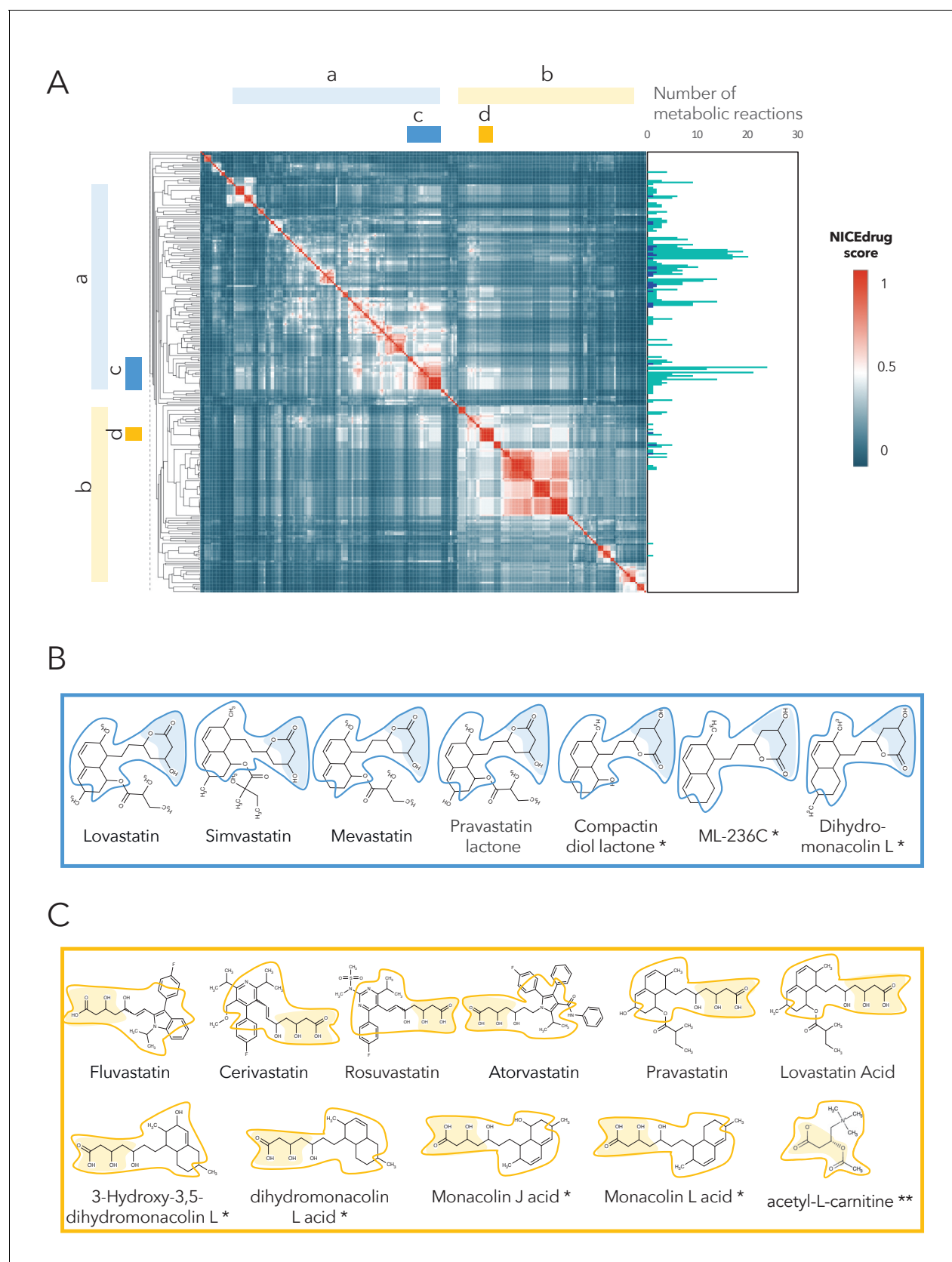


Figure 5. Clustering of molecules with statin reactive sites based on NICEdrug score suggests drugs for repurposing. (A) Pairwise NICEdrug score between all molecules with statin reactive sites (heat map) and number of metabolic reactions in which they participate (right). We highlight clusters of

Figure 5 continued on next page

Figure 5 continued

statins of type 1 (cluster a) and type 2 (cluster b), and clusters of most similar molecules to type one statins (cluster c) and type two statins (cluster d). Within the metabolic reactions, we indicate the total number of reactions (dark color) and the number of reactions that involve the statin reactive site (light color). (B) Examples of statins and Mevastatin analogues of type one from cluster c (blue) and of type two from cluster d (gold). We left the known statins unmarked, which are appropriately clustered together based on the NICEdrug score, and we mark with * new molecules that cluster with statins and that NICEdrug.ch suggests could be repurposed to act as statins. Reactive sites in type one statins and type two statins are colored in blue and orange, respectively. The reactive site neighborhood as considered in the NICEdrug score is also marked. See also; **Figure 5—figure supplement 1** , and **Supplementary file 4**.

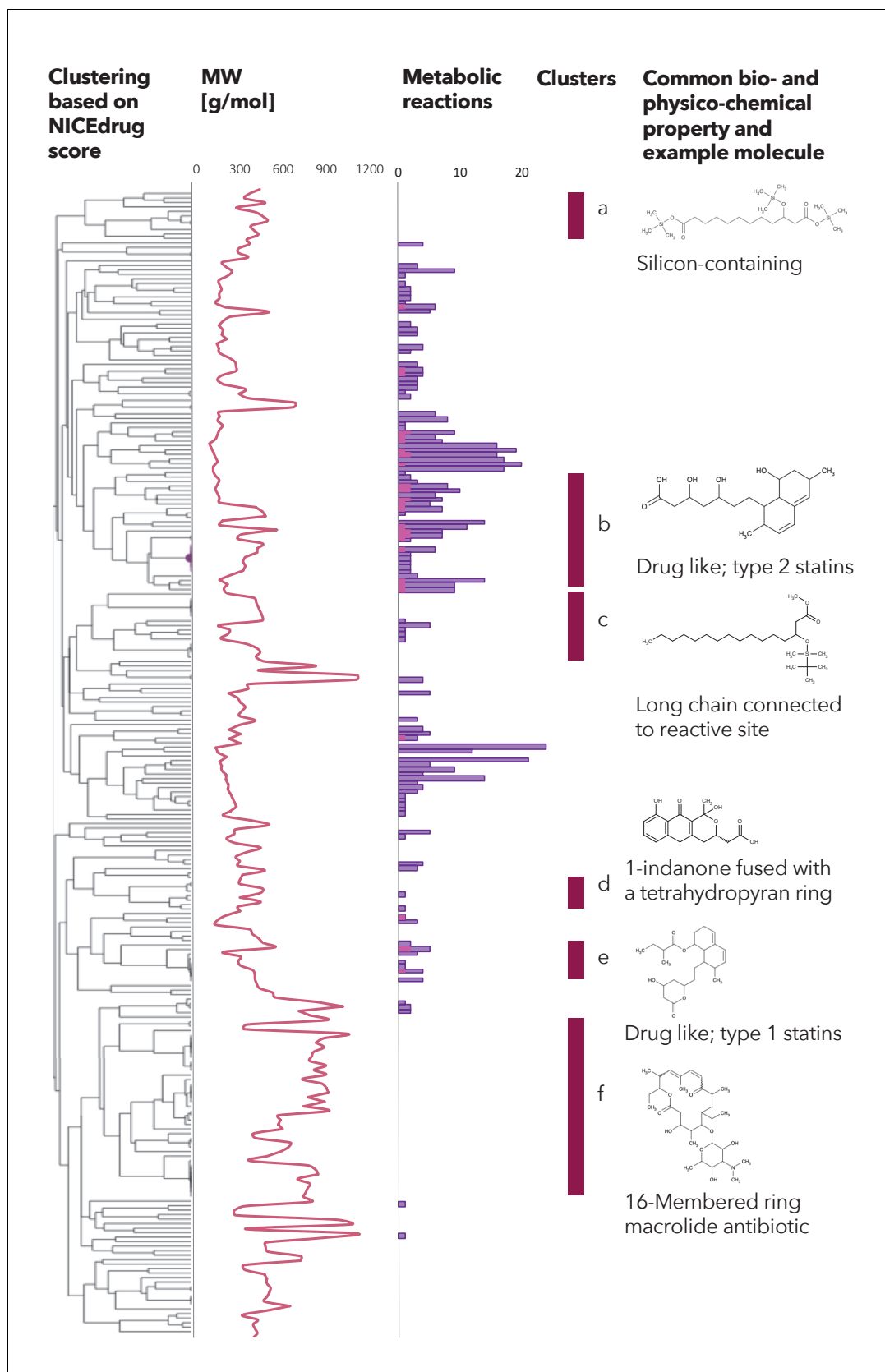


Figure 5—figure supplement 1. Clustering based on NICEdrug score, molecular weight, and reactivity of statin like molecules. Hierarchical clustering based on the NICEdrug score of all molecules in NICEdrug.ch that contain statin reactive site (left). We report the molecules' molecular weight (middle) Figure 5—figure supplement 1 continued on next page

Figure 5—figure supplement 1 continued

left) and number of drug metabolic reactions or reactions in which these drugs participate (middle). The molecular weight seems to be inversely correlated with the number of drug metabolic reactions. We highlight six clusters of drugs (a–f, middle right) and an example representative molecule (left). Interestingly, these clusters also group molecules based on bio- or physico-chemical properties: ‘cluster a’ involves a range of silicon-containing chemical molecules, ‘cluster b’ are drug-like molecules of type two statins, ‘cluster c’ includes chemical molecules with a long chain connected to the reactive site, ‘cluster d’ involves molecules with 1-indanone fused with a tetrahydropyran ring, ‘cluster e’ comprises drug-like molecules of type one statins, and ‘cluster f’ are 16-membered ring macrolide antibiotics.

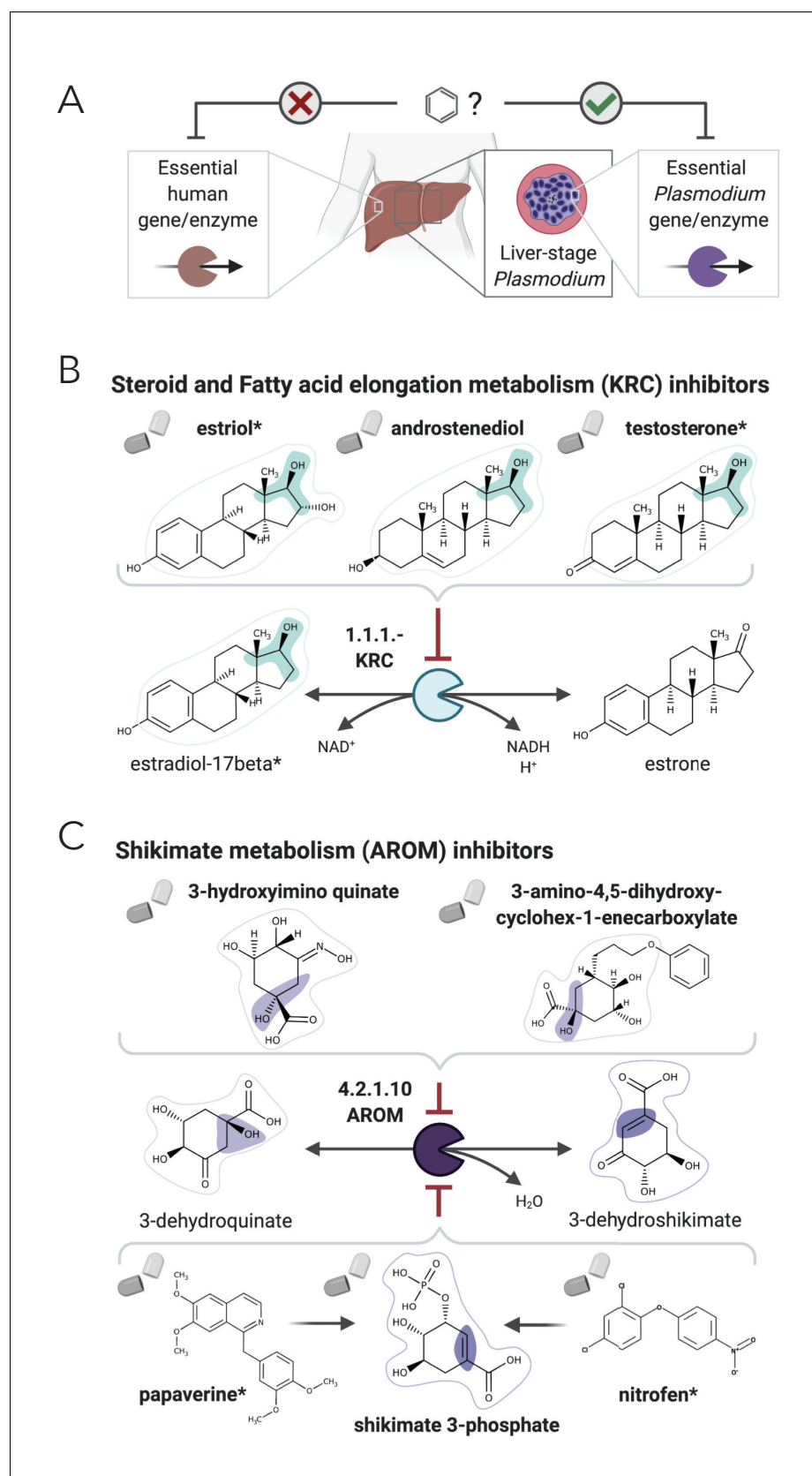


Figure 6. NICEdrug.ch suggests shikimate 3-phosphate as a top candidate to target liver-stage malaria and minimize side effects in host human cells. (A) Schema of ideal scenario to target malaria, wherein a drug efficiently

Figure 6 continued on next page

Figure 6 continued

inhibits an essential enzyme for malaria parasite survival and does not inhibit essential enzymes in the host human cell to prevent side effects. **(B)** Shikimate 3-phosphate inhibits enzymes in the *Plasmodium* shikimate metabolism, which is essential for liver-stage development of the parasite. Shikimate 3-phosphate does not inhibit any enzyme in the human host cell since it is not a native human metabolite, and it does not show similarity to any native human metabolite. **(C)** Mechanistic details of inhibition of aroC by shikimate 3-phosphate and other NICEdrug candidates. See also **Supplementary file 5; Supplementary file 6**.

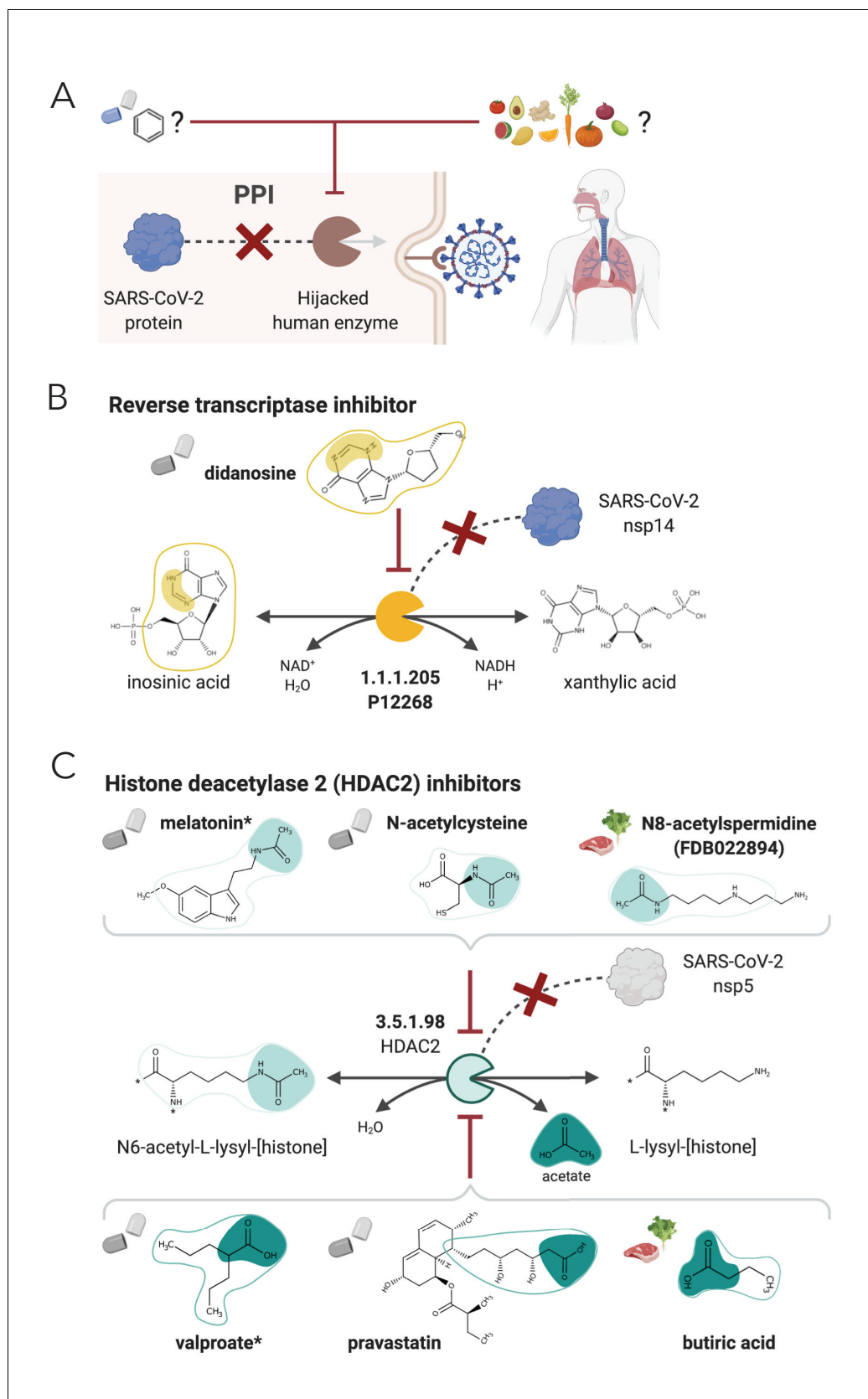


Figure 7. NICEdrug.ch strategy to fight COVID-19, and NICEdrug.ch candidate inhibitors of SARS-CoV-2 host factors: reverse transcriptase and HDAC2. (A) Schema of NICEdrug.ch strategy to target COVID-19, wherein a drug (top-left) or molecules in food (top-right) efficiently inhibit a human

Figure 7 continued on next page

Figure 7 continued

enzyme hijacked by SARS-CoV-2. Inhibition of this host factor reduces or abolishes protein–protein interactions (PPI) with a viral protein and prevents SARS-CoV-2 proliferation. **(B)** Inhibition of the reverse transcriptase (EC: 1.1.1.205 or P12268) and the PPI with SARS-CoV-nsp14 by didanosine based on NICEdrug.ch. **(C)** Inhibition of the HDAC2 (EC: 3.5.1.98) and the PPI with SARS-CoV-nsp5 by molecules containing acetyl moiety (like melatonin, N-acetylcysteine, and N8-acetylspermidine), and molecules containing carboxylate moiety (like valproate, stains, and butyrate) based on NICEdrug.ch. See also **Figure 7—figure supplement 1**; **Figure 7—figure supplement 2**, **Supplementary file 7**; **Supplementary file 8**.

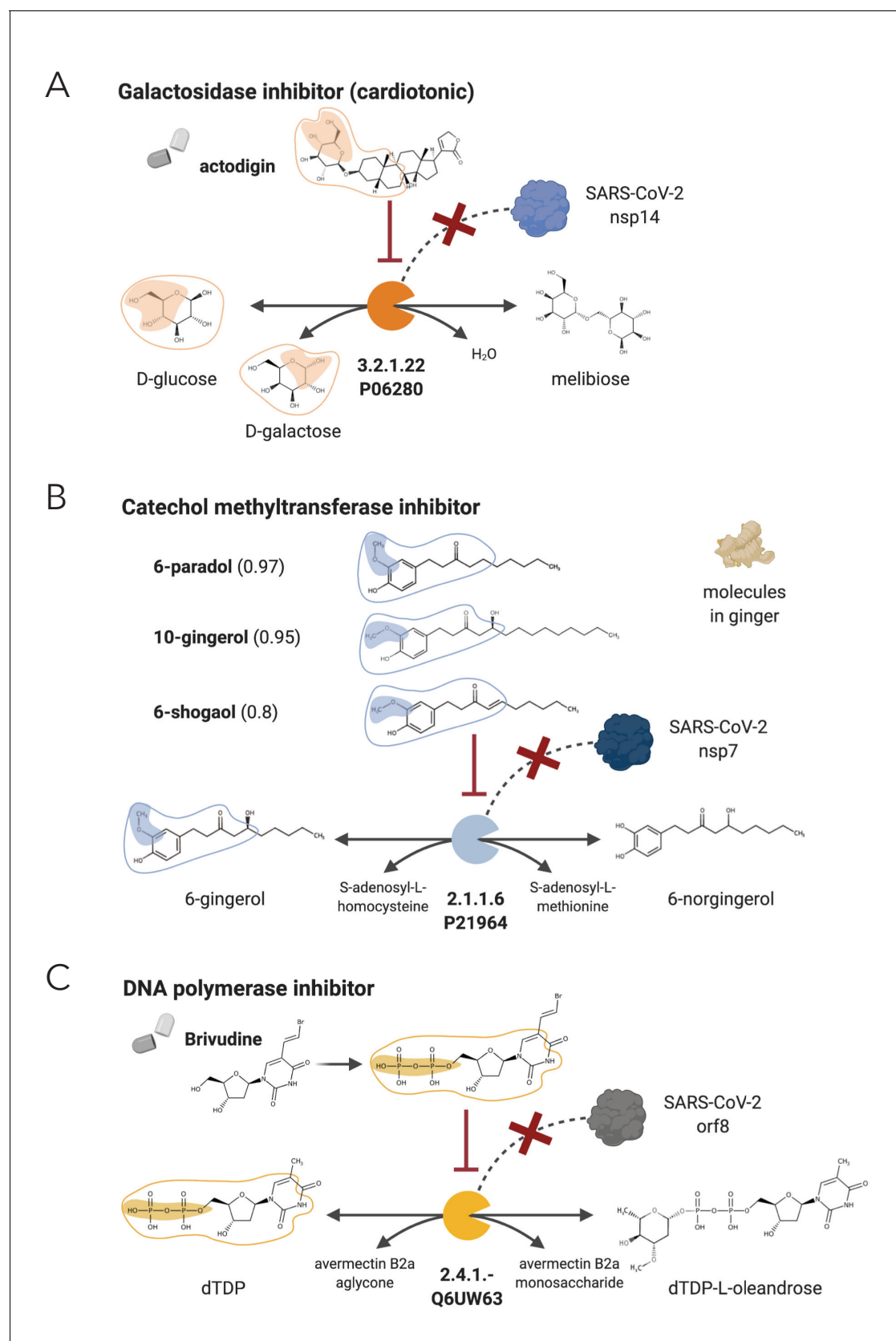


Figure 7—figure supplement 1. NICEdrug candidate inhibitors of SARS-CoV-2 host factors: galactosidase, catechol methyltransferase, and DNA polymerase. (A) Inhibition of the galactosidase (EC: 3.2.1.22 or P06280) and the PPI with SARS-CoV-2 nsp14 by actodigin based on NICEdrug.ch. (B) Figure 7—figure supplement 1 continued on next page

Figure 7—figure supplement 1 continued

Inhibition of the catechol methyltransferase (EC: 2.1.1.6 or P21964) and the PPI with SARS-CoV-2 nsp7 by 6-paradol, 10-gingerol, and 6-shogaol, which are molecules in ginger, based on NICEdrug.ch. (C) Inhibition of the DNA polymerase (EC: 2.4.1.-) and the PPI with SARS-CoV-2 nsp8 by brivudine based on NICEdrug.ch.

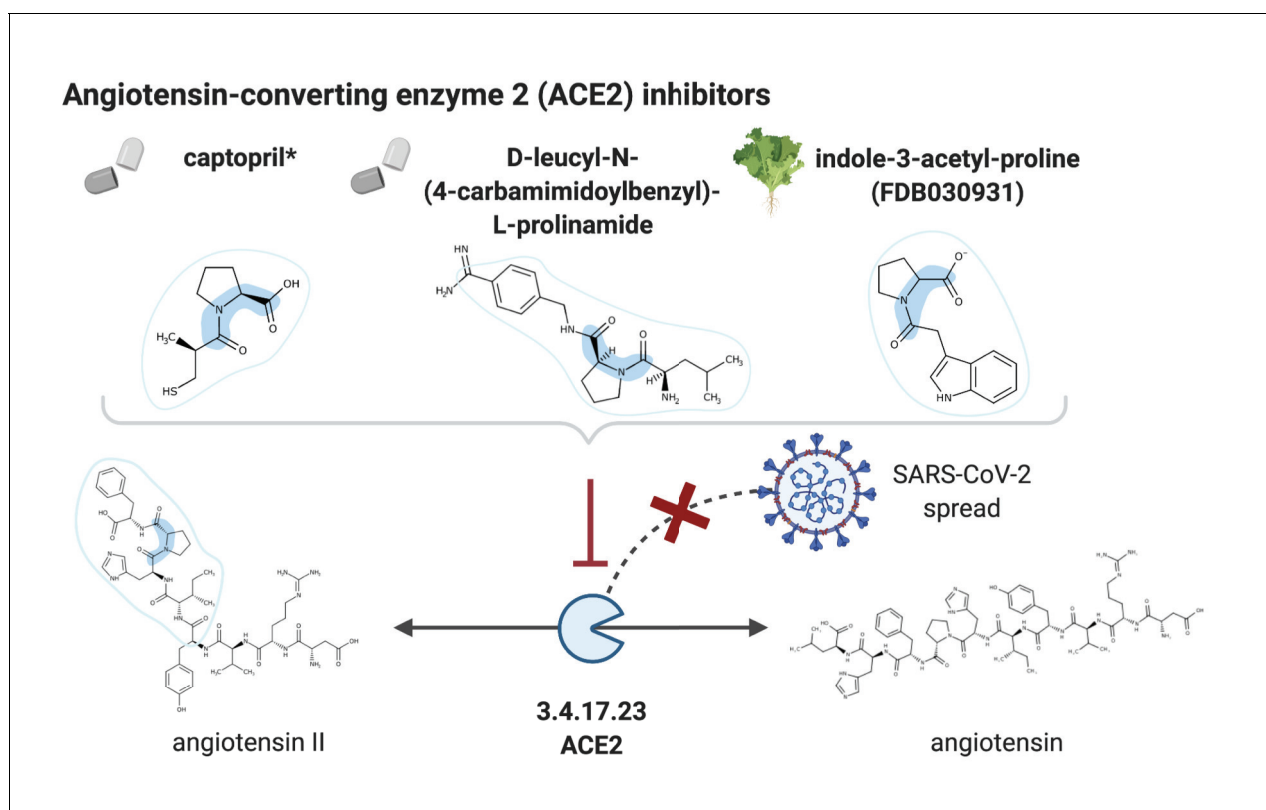


Figure 7—figure supplement 2. NICEdrug candidate inhibitors of ACE2. Inhibition of the ACE2 (EC: 3.4.17.23), a putative host factor of SARS-CoV-2, by the known inhibitor captopril, and NICEdrug candidates D-leucyl-N-(4-carbamimidoylbenzyl)-L-prolinamide and indole-3-acetyl-proline.

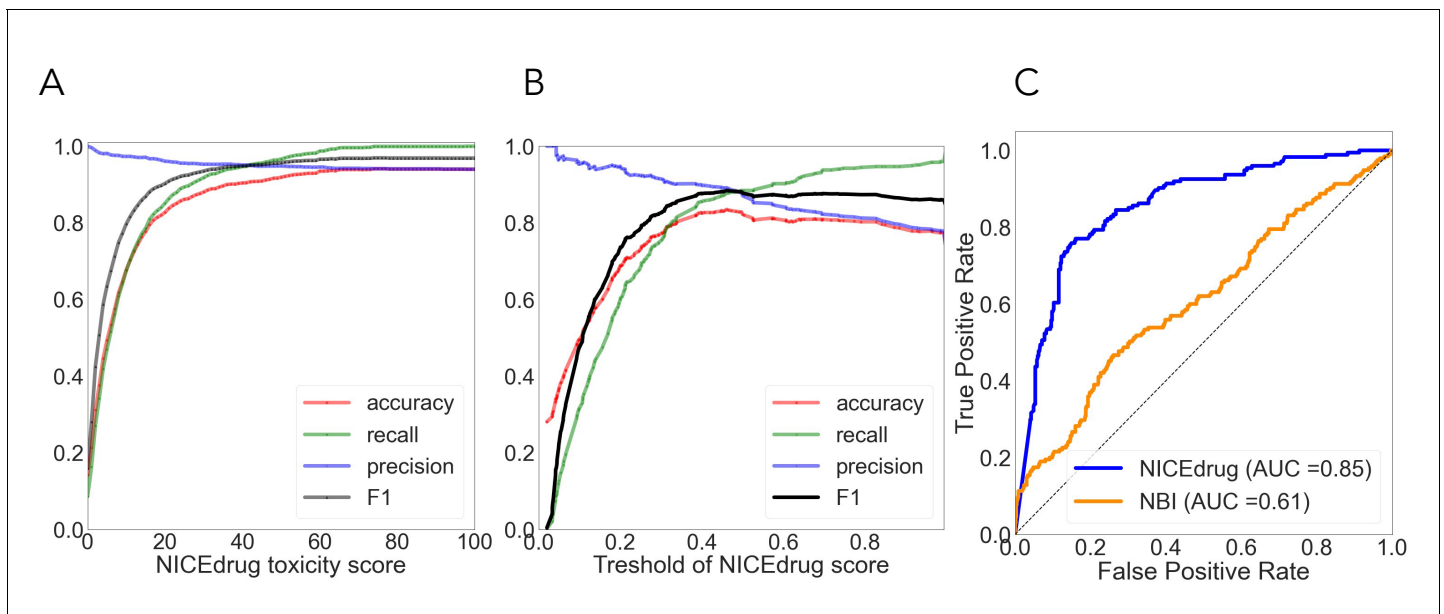


Figure 8. Quantitative comparison of drug toxicity (A) and drug–enzyme pairs (B, C). (A) Evaluation of NICEdrug.ch toxicity predictions on the test dataset of 1777 drug candidates from PubChem. Y-axis represents the value of a metric, namely accuracy (red), precision (green), recall (blue), and F1 (yellow) based on putative threshold for NICEdrug toxicity score. X-axis represents the predicted NICEdrug.ch toxicity score. (B) Evaluation of overall performance of NICEdrug.ch and NBI tool in predicting interaction of enzymes and drug pairs in terms of ROC curves and value of AUC. (C) Quality of enzymes and drug pairs predictions using NICEdrug.ch in terms of statistical measures including: precision, recall, F1, and accuracy.