JOURNAL OF
EMERGING INVESTIGATORS

Article

In silico screening of DEAB analogues as ALDH1 isoenzymes inhibitors in cancer treatment

Kennard Liong[1]*, Gia Bao Le[1]*, William Tan[2]* and Carl Sangree[1]

[1] Northfield Mount Hermon School, Gill, Massachusetts

[2] Jakarta Intercultural School, Jakarta, Indonesia

* These authors contributed equally to this work.

SUMMARY

Inhibiting aldehyde dehydrogenase (ALDH) isoenzymes show promise in cancer treatment, especially as various cancer types exhibit elevated activities of ALDH1 isoenzymes. Several inhibitory drugs, including diethylaminobenzaldehyde (DEAB), have been studied, and demonstrate high efficacy in combating tumor proliferation and enhancing chemotherapy. Even then, significant drawbacks such as low solubility, cytotoxicity, and inconsistent potency have resulted in DEAB's poor efficacy as cancer therapeutics. In this study, we performed in silico screening tests to analyze 55 DEAB analogs – classified into three different scaffolds – to identify new compounds with enhanced ligand binding affinity and reduced toxicity. Given that the ALDH1 active site is predominantly composed of nonpolar residues, we hypothesized that analogs that have low polarity would have lower (better) free energies of binding. Through molecular docking, we found that analogs that are highly hydrophobic with more aromatic rings, higher molecular weights, and high molar refractivity (indicating high polarizability) had significantly stronger binding affinities. Among the analogs screened, analogs 16 and 17 had the best docking scores among ligands that met the drug-like criteria for bioavailability. They had high logPs, indicating high hydrophobicity. The absorption, distribution, metabolism, excretion, and toxicity (ADMET) test also deemed both ligands non-toxic toward human cell culture. However, other analogs with lower logP exhibited high (worse) docking scores among ligands that met the drug-like criteria for bioavailability. With these results highlighting promising ALDH1 isoenzyme inhibitors, a proper in vitro procedure testing ligands 16 and 17 is essential to validate the inhibitory efficacy and cytotoxicity of these two potential therapeutics.

INTRODUCTION

Aldehyde dehydrogenases (ALDHs) are a family of enzymes with at least 19 different isoenzymes that participate in the detoxification process of human alcohol metabolism through nicotinamide-adenine dinucleotide phosphate (NADP)+, a cofactor involved in the oxidation of aldehydes into carboxylic acids (1). ALDHs operate in a homodimer or homotetramer structure, especially among the ALDH1 and ALDH3 classes; each monomer has catalytic, cofactor binding, and oligomerization structural domains (1). In ALDH1A1, for reference, the Cys303 residue within the active site acts as a nucleophile, attacking the carbonyl carbon of the aldehyde to form a thiohemiacetal intermediate (1). The aldehyde's oxygen forms partial bonds with the amide groups on Asn169 and Cys303 (1). A transient intermediate is formed from the reduction of NADP+ through hydride transfer with the intermediate, before being hydrolyzed by a water molecule and deprotonated by Glu269 to release the carboxylic acid and free enzyme (1) (**Figure 1A-D**). The ALDH active site is commonly targeted by inhibitor drugs (1).

Because of ALDHs' pivotal roles in detoxification pathways, dysregulation of ALDHs is often linked with many severe diseases, such as cancer, cardiovascular, and neurological diseases (2). Some ALDH1 isozymes, in particular, are involved in differentiation and transcriptional regulation within stem cells (3). It is suggested that ALDHs can regulate cell function through pathways like the biosynthesis of retinoic acid, the scavenging of reactive oxygen species, and the breakdown of toxic aldehydes (4). Moreover, ALDH1

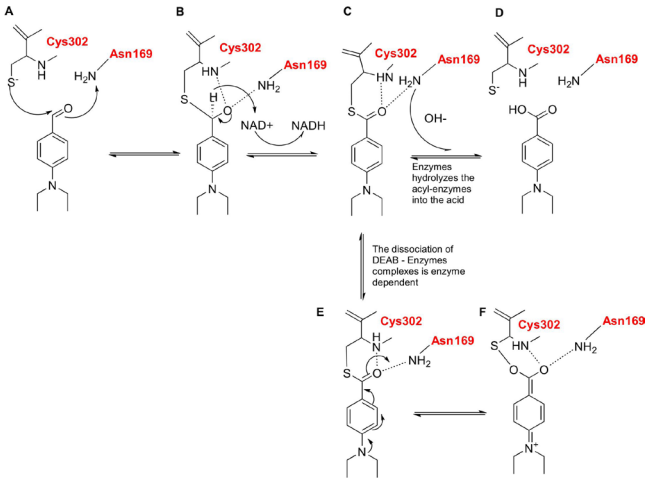

Figure 1: Mechanism of DEAB as an irreversible inhibitor of ALDH1A2. The residue names are bolded and colored red. A) The Cys302 (or Cys303 for ALDH1A1 and ALDH1A3) residue performs a nucleophilic attack on the carbonyl carbon of the aldehyde group on DEAB. B) A thiohemiacetal intermediate is formed; its hydrogen reduces NAD(P)+ to NAD(P)H, which is normally used for enzyme regeneration. C) The aldehyde's oxygen forms partial bonds with amide groups on Asn169 and Cys302 or Cys303, forming an acyl-enzyme intermediate. D) Normally, hydrolysis by a water molecule deprotonated by Glu269 releases the carboxylic acid and free enzyme. E, F) With DEAB, possible resonance structures may stabilize the DEAB-ALDH complex, inhibiting the enzyme.

https://doi.org/10.59720/24-343

isozymes have been identified as markers in stem cells and cancer stem-like cells (CSCs) found in various cancers; CSCs exhibiting high ALDH1 activity demonstrate similar characteristics to human stem cells, including self-renewal and tumor-initiating capacities, contributing to metastasis (1, 3, 4).

Further studies conducted by researchers have observed the overexpression of ALDHs in lung, invasive cervical, prostate, breast, colorectal, gastric, esophageal, head and neck, and other types of cancers that exhibit adherent tumors (5-12). Moreover, elevated ALDH activity is correlated with a worse prognosis, and ALDHs even have metabolic roles that confer resistance to chemotherapy via metabolic inactivation, potentially contributing to relapses among patients (1). Targeting and inhibiting ALDH1 proteins in tumor stem cells could, therefore, be a way to enhance drug sensitivity and inhibit differentiation and tumor proliferation among patients.

Diethylaminobenzaldehyde (DEAB) is a compound that can act as a competitive inhibitor on human ALDH isoenzymes in CSCs (13). For its ability, DEAB is commonly used in negative control experiments in assays testing for ALDH activity (14). Previous studies have examined enzyme kinetics and employed quadrupole time-of-flight mass spectroscopy to analyze the inhibitory effects of DEAB on ALDH isozymes (14). DEAB was identified as a competitive, reversible inhibitor for the ALDH1A1, ALDH1A3, and ALDH1B1 isozymes, and an irreversible, covalent inhibitor for ALDH1A2 due to the compound's exceptionally low turnover (14). The proposed mechanism of DEAB action is similar to that of its usual aldehyde substrate up to the acyl-enzyme intermediate (14). From here, the acyl-enzyme intermediate may reach a stable state due to resonance of the thioester intermediate, granted by a lone electron pair from the para-diethylamino substituent (14). As extra electrons occupy the carbonyl oxygen, it becomes more negatively charged, forming a stable intermediate with the amide groups supplied by Asn169 and Cys303 that is unlikely to be hydrolyzed (14) (**Figure 1E, F**). Depending on the isozyme, the prominence of this resonance structure may differ, leading to different inhibition strengths (14).

In an earlier study, researchers found interesting structure-activity relationship (SAR) trends of DEAB analogs when testing these compounds *in vivo* (7). Particularly, they observed an enhanced inhibitory effect with analogs bearing a methyl group at the meta position and constrained heterocyclic rings at the para position with respect to the aldehyde functional group (7). For instance, 3-methyl-4-(1-piperidinyl) benzaldehyde has an IC50 of 0.29 μM toward ALDH1A3's activity, while 3-methyl-4-(1-pyrrolidinyl) benzaldehyde and DEAB have IC50s of 1.15 μM and 10.4 μM against the same protein's activity, respectively (7) (**Figure 2**). Besides that, when looking at analogs with halogenated R-groups, 4-diethylamino-chlorobenzaldehyde and 4-diethylamino-bromobenzaldehyde had IC50s of 0.55 μM and 0.63 μM respectively (7) (**Figure 2**). The similar inhibitory effect of these analogs implied that differences in atomic volume and electronegativity between halogens were not major factors affecting analogs' inhibitory effects against ALDH1A3 (7).

For their selectivity to specific isozymes, DEAB and its analogs provide an opportunity for further development of novel ALDH1 inhibitors for targeted CSC therapy. Herein, our goal was to identify and analyze the correlation between ALDH1 binding and chemical properties of DEAB's analogs via *in silico* virtual screening methods, which aim to provide essential SAR insights when developing ALDH1 inhibitors. We hypothesized that analogs incorporating nonpolar functional groups would be the most efficient in binding to ALDH1 enzymes' binding pocket, based on the pervasive distribution of hydrophobic residues in the binding pocket of ALDH1 enzymes.

In accordance with our hypothesis, we observed that analogs that were larger and heavier – namely those with higher logP values, molecular weights, and molar refractivities – exhibited lower (better) average binding scores to the ALDH1 active sites than smaller analogs. Given that analogs derived from scaffold 2 illustrated high inhibitory potency against ALDH1A1 and ALDH1A3, future studies should focus on analyzing the inhibitory mechanisms of scaffold 2-derived analogs – particularly analogs 16 and 17 – against ALDHs. Additionally, *in vivo* experiments testing these analogs' toxicities are required to substantiate the potential of DEAB analogs as therapeutics.

RESULTS

Docking DEAB analogs against ALDH1 isozymes

We began our study by compiling a library of 55 DEAB analogs, sourced from the PubChem database and previously reported in the literature (7, 16). We then conducted high-throughput virtual screening through molecular docking of all 55 compounds to the ALDH1A1, ALDH1A2, ALDH1A3, and ALDH1B1 isozymes, as the ALDH1 isozymes are known markers for CSCs (1, 3, 4). The grid box and docking parameters used contained the catalytic site of ALDH, as DEAB and its analogs have been postulated to act upon

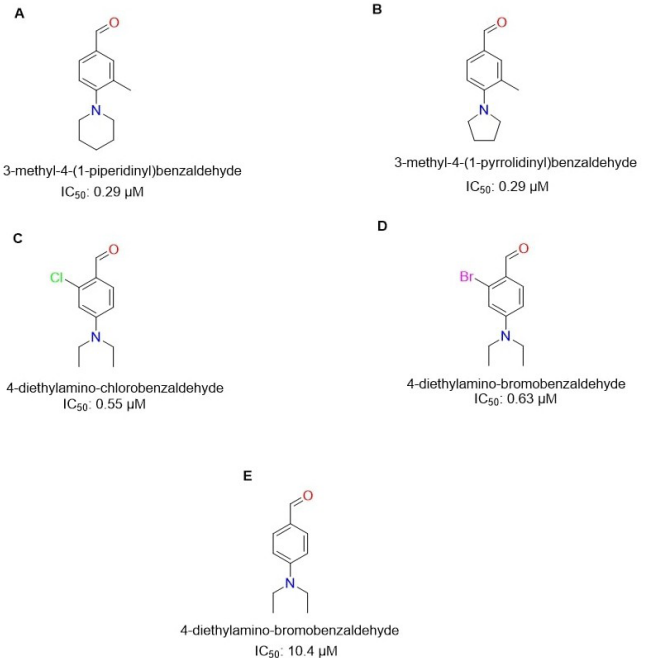

Figure 2: DEAB and a few analogs with IC_{50} values against ALDH1A3 listed. A, B) Analogs with a heterocyclic ring replacing the aliphatic moiety at the para position and methyl group at the meta position. C, D) Analogs with halogenated methyl group at ortho position; E) Structure of unmodified DEAB.

https://doi.org/10.59720/24-343

the active Cys303 residue in ALDH1A1 (with slight residue number differences across isozymes, but still a part of the same structural domain). For each enzyme, the free energies of binding were as follows: -9.3 to -5.5 kcal/mol (median = -6.5 kcal/mol) for ALDH1A1; -8.4 to -4 kcal/mol (median = -6.1 kcal/mol) for ALDH1A3; -8.4 to -5.6 kcal/mol (median = -6.5 kcal/mol) for ALDH1A3; and -7.8 to -5.4 kcal/mol (median = -6.3 kcal/mol) for ALDH1B1 (**Figure 3**). As DEAB and its analogs have been observed to have more potent antiproliferative activity against ALDH1A1 and ALDH1A3, and their free energies of binding had a lower range and median, we focused on these two isozymes for the following SAR analysis.

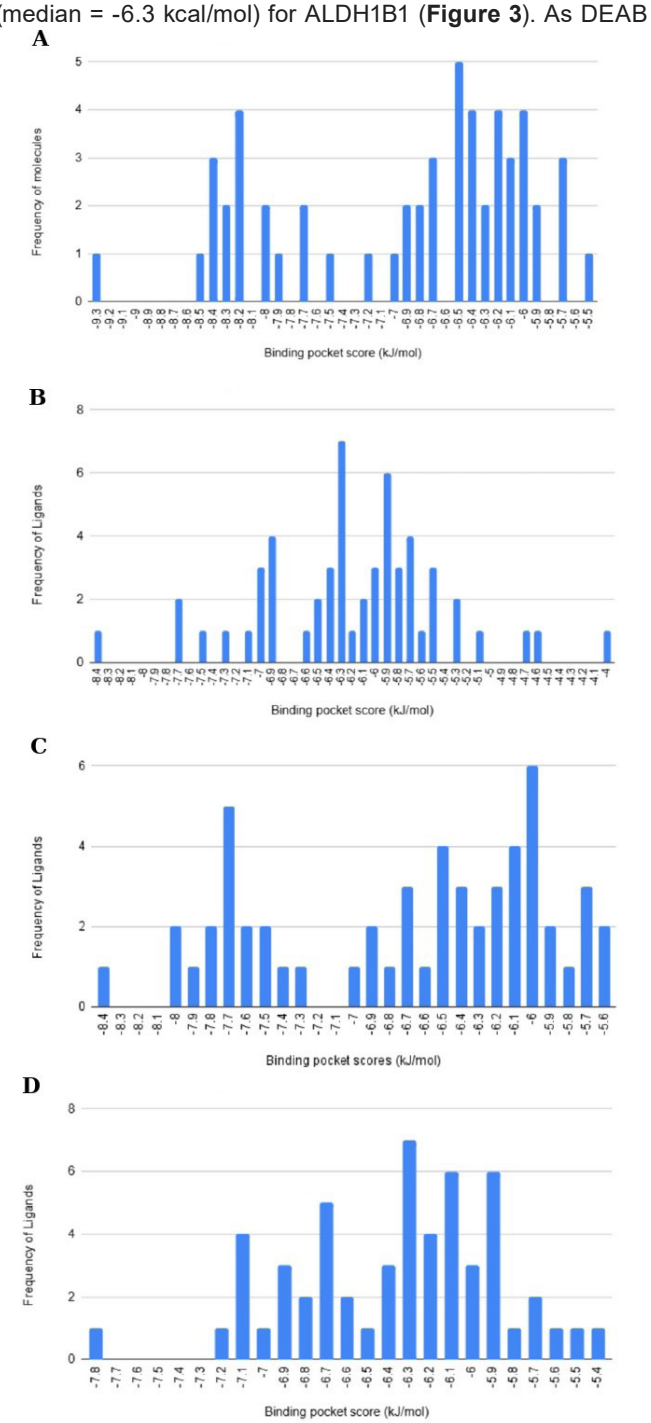

Figure 3: Distribution of free energies of binding across the ALDH1 isozymes. Binding pocket scores were arranged from lower (left) to higher (right) values along the x-axis. The frequency of each docking score was plotted in the y-axis. A-D) Distribution of free binding energies against ALDH1A1, ALDH1A2, ALDH1A3, and ALDH1B1, respectively.

Scaffold and R-group analysis

The 55 ligands tested could be grouped into three scaffolds, based on differences in their nitrogen-containing substituents (scaffold 1), substituents that differ on the ortho position (scaffold 2), or both (scaffold 3) (**Figure 4**). For ALDH1A1 and ALDH1A3, scaffold 2 had a lower average free energy of binding compared to scaffolds 1 and 3; this difference was statistically significant for ALDH1A1, where scaffold 2 had an average free energy of binding of -7.2 kcal/mol, and scaffolds 1 and 3 had -6.3 kcal/mol (two-sample one-tailed *t*-test, $p = 0.0038$) and -6.4 kcal/mol (two-sample one-tailed *t*-test, $p = 0.0086$) (**Figure 5A, B**). On the other hand, for ALDH1A2 and ALDH1B1, the differences in average free binding energies were deemed statistically insignificant; for ALDH1A2, scaffold 1 and 2 had a *p*-value of 0.243 (two-sample one-tailed *t*-test), and scaffold 2 and 3 had a *p*-value of 0.241 (two-sample one-tailed *t*-test) (**Figure 5C**). Similarly, for ALDH1A3, the *p*-values were 0.312 between scaffolds 1 and 2 (two-sample one-tailed *t*-test), and 0.467 between scaffolds 2 and 3 (two-sample one-tailed *t*-test) (**Figure 5D**).

Due to their higher average free energies of binding, the R-groups of DEAB analogs under scaffold 2 were further analyzed. Among scaffold 2 analogs with the top half of binding scores, 17 out of 18 compounds had an aromatic ring on the R-group, with none in the bottom half having one. The compounds in the bottom half predominantly contained ether groups as well, with halogen or alkyl chain R-groups performing worse. Upon analyzing their chemical properties, R-groups in scaffold 2 with higher molar refractivities, molecular weights, and Wildman-Crippen log partition coefficient (logP) values – an indication of drug hydrophobicity – outperformed those with lower values in binding affinity to ALDH1A1 and ALDH1A3, although the relationship was weaker for the latter protein (**Figure 6A-C, G-I**). In addition, higher values of molar refractivity, molecular weights, and logP were shown to have an inversely proportional relationship with binding affinity of DEAB analogs (Pearson's correlation, $R^2 = 0.712$, 0.631, 0.587 (ALDH1A1) and 0.605, 0.497, 0.475 (ALDH1A3) respectively, $p < 0.001$) (**Figure 6A-C, G-I**). A higher number of rotatable bonds and aromatic rings also correlated to a lower free energy of binding (Pearson's correlation, $R^2 = 0.37$,

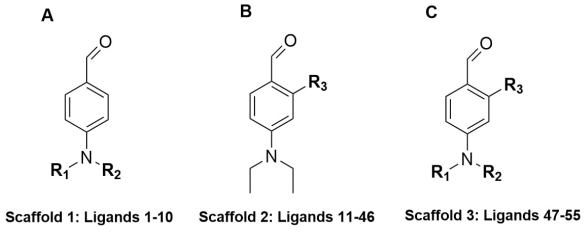

Figure 4: Common scaffolds across the DEAB analogs used. A) Scaffold 1 (Ligand 1 - 10) features modification at R_1 and R_2 positions. B) Scaffold 2 (Ligand 11 - 46) featured modifications at R_3 position. C) Scaffold 3 (Ligand 47 - 55) featured modifications at R_1, R_2, and R_3 positions.

https://doi.org/10.59720/24-343

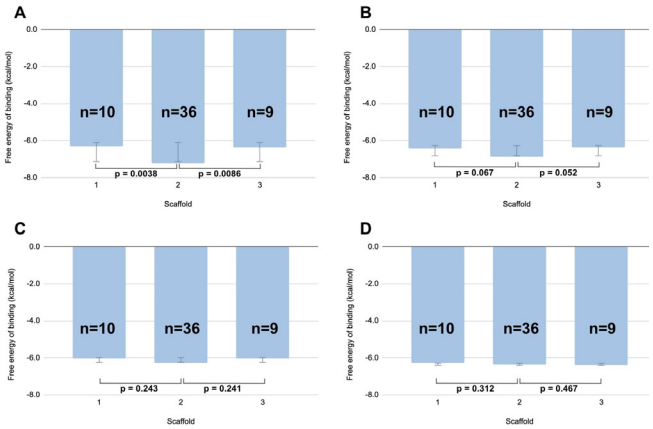

Figure 5: Mean free energies of binding for each scaffold across ALDH1 isozymes. Average free energies of binding for each scaffold and the sample sizes for each scaffold (*n*) are shown, with the name for scaffolds displayed along the x-axis and their corresponding mean free energy of bindings on the y-axis. A) Mean free energies of binding for scaffolds against ALDH1A1. Scaffold 2 had statistically lower mean free energies of binding than scaffolds 1 (p = 0.0038) and 3 (p = 0.0086). B) Mean free energies of binding for scaffolds against ALDH1A3. Scaffold 2 had a lower mean free energy of binding compared to scaffolds 1 (p = 0.067) and 3 (p = 0.052) with p-values close to the 0.05 cutoff. C) Mean free energies of binding for scaffolds against ALDH1A2. Scaffold 2 did not have a lower mean free energy of binding than scaffolds 1 (p=0.243) and 3 (0.241) that were statistically significant. D) Mean free energies of binding for scaffolds against ALDH1B1. Scaffold 2 did not have a lower mean free energy of binding than scaffolds 1 (p = 0.312) and 3 (p = 0.467) that was statistically significant in ALDH1B1 either. Pairwise two-sample, one-tailed *t*-tests were conducted to compare the scaffolds' free energies of binding. The error bars represent standard deviation.

0.675 (ALDH1A1) and 0.321, 0.652 (ALDH1A3) respectively, p <0.001) (**Figure 6D, F, K, M**). No correlation was found between topological polar surface area (TPSA) and DEAB analogs' free energies of binding to ALDH1A1 (Pearson's correlation, R^2 = 0.067 and p = 0.056 (**Figure 6E**). Yet, a weak correlation was found between TPSA and DEAB analogs' free energies to ALDH1A3 (Pearson's correlation, R^2 = 0.086 and p = 0.030) (**Figure 6L**).

Analyzing ligand-receptor interactions

Molecular docking was conducted using UCSF Chimera software to visualize docking of ALDH1 isozymes with DEAB analogs. We analyzed the four docking poses of DEAB on ALDH1A1 and ALDH1A3, ligand 11 (2-[2-Amino-2-(3-methylnaphthalen-1-yl)ethoxy]-4-(diethylamino) benzaldehyde) – which has the lowest binding energy for ALDH1A1 among 55 tested ligands – on ALDH1A1, and ligand 23 (2-[3-[Bis(pyridin-2-ylmethyl)amino]propyl]-4-(diethylamino)benzaldehyde) – which has lowest binding energy for ALDH1A3 among 55 tested ligands – on ALDH1A3. Hydrogen bonds were uncommon between the ALDH1A1 and ALDH1A3 isozymes and the ligands but van der Waals interactions were more common and especially pronounced among molecules with bulky R-groups (**Figure 7**). However, further pharmacokinetic tests for ligands 11 and 23 conducted through the pkCSM web server suggested that these ligands display high toxicity and low viability as drug-like compounds, hence they were only used for docking pose analysis and

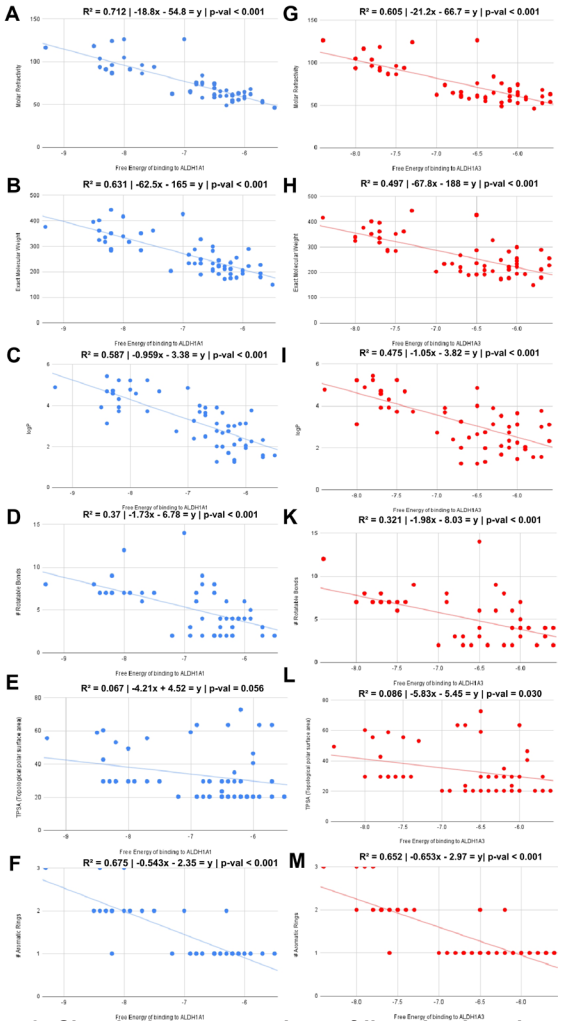

Figure 6: Chemical property values of ligands plotted against free energies of binding to ALDH1A1 and ALDH1A3. ALDH1A1 is marked in blue, ALDH1A3 is marked in red, number is represented by #, A) Molar refractivity of the DEAB analogs is inversely correlated with their binding energies against ALDH1A1 (R^2 = 0.712, p < 0.001). B) Exact molecular weight of DEAB analogs is inversely correlated with their binding energies against ALDH1A1 (R^2 = 0.631, p < 0.001). C) logP of DEAB analogs is inversely correlated with their binding energies against ALDH1A1 (R^2 = 0.587, p < 0.001). D) Number of rotatable bonds of DEAB analogs is inversely correlated with their binding energies against ALDH1A1 (R^2 = 0.37, p < 0.001) E) Topological polar surface area (TPSA) of DEAB analogs was not found to correlate with their binding energies against ALDH1A1 (R^2 = 0.067, p = 0.056). F) Number of aromatic rings of DEAB analogs is inversely correlated with their binding energies against ALDH1A1 (R^2 = 0.675, p < 0.001). G) Molar refractivity of DEAB analogs is inversely correlated with their binding energy against ALDH1A3 (R^2 = 0.605, p < 0.001). H) Exact molecular weight of DEAB analogs is inversely correlated with their binding energies against ALDH1A3 (R^2 = 0.497, p < 0.001). I) logP of DEAB analogs is inversely correlated with their binding energy against ALDH1A3 (R^2 = 0.475, p < 0.001). K) Number of rotatable bonds of DEAB analogs is inversely correlated with their binding energies against ALDH1A3 (R^2 = 0.321, p < 0.001). L) Topological polar surface area (TPSA) of DEAB analogs shows a weak inverse correlation with their binding energies against ALDH1A3 (R^2 = 0.086, p = 0.030). M) Number of aromatic rings of DEAB analogs is inversely correlated with their binding energies against ALDH1A3 (R^2 = 0.652, p < 0.001). The R^2 and p-values were calculated using a significance test for Pearson's correlation coefficient.

https://doi.org/10.59720/24-343

Ligand	Water solubility (log Mol/L)	Caco-2 permeability (log Papp in 10^{-6} cm/s)	Intestinal absorption (human) (% absorbed)	P-glycoprotein substrate (Yes/No)	AMES toxicity (Yes/No)	hERG I inhibitor (Yes/No)	hERG II inhibitor (Yes/No)	Oral Rat Chronic Toxicity (LOAEL) (log mg/kg_bw/day)	Hepatotoxicity (Yes/No)
16	-5.107	2.065	94.209	No	No	No	No	1.151	No
17	-5.006	1.927	94.968	No	No	No	No	1.738	No
18	-5.146	1.921	95.069	No	No	No	No	2.044	No
24	-5.006	1.927	94.968	No	No	No	No	1.738	No
25	-3.829	1.294	93.441	Yes	No	No	No	1.12	No
11	-4.238	0.587	91.923	Yes	Yes	Yes	Yes	0.94	No
23	-3.636	1.087	94.997	Yes	No	No	Yes	1.05	Yes

Table 1: Pharmacokinetic properties of DEAB analogs predicted using pkCSM web server, corresponding to the structures shown in Figure 8. Bioavailability (water solubility, Caco-2 permeability, human intestinal absorption) and toxicity metrics (P-glycoprotein substrate status, AMES toxicity, hERG I and II inhibition, lowest observed adverse effect level (LOAEL) for oral rat chronic toxicity, and hepatotoxicity) for ligands 16, 17, 18, 24, 25, 11, and 23. Ligands colored in green are deemed non-toxic toward human cell culture. Ligands colored in red are deemed toxic toward human cell culture.

are not recommended for lead compound development (24) (**Table 1**).

Pharmacokinetic screening analysis

For further testing, a pharmacokinetic property screen was conducted through the pkCSM webserver; 35 out of 55 ligands tested negative in the *Salmonella typhimurium* reverse mutation assay (AMES), human ether-a-go-go-related gene (hERG) I and II inhibition, hepatotoxicity, and passed Lipinski's Rule of Five (24). Ligands 11 and 23 both tested positive for AMES toxicity and hERG II inhibition and hence were not included in the list, as they were deemed toxic (24) (**Table 1**). Among the 35 ligands deemed as non-toxic, ligands 16 and 17 were singled out for having the lowest free energies of binding for both ALDH1A1 and ALDH1A3, with both scoring -8.3 kcal/mol and -7.7 kcal/mol for the respective isozymes (**Figure 8A**). Both these ligands were scaffold 2 ligands with ethers for the R-group that branched to a halogenated benzene, unlike the experimentally tested analogs, which had an unbranched halogen bonded to DEAB's aromatic ring (**Figure 2**). For their low free energy of binding, low toxicity, and drug-like properties, these two ligands are ideal for future development and potential *in vitro* testing.

DISCUSSION

Analogs bearing nonpolar and bulky functional groups exhibited an enhanced binding affinity toward ALDH1 active sites. We initially split our 55 ligands into three different substitution patterns: scaffold 1 with 10, scaffold 2 with 36, and scaffold 3 with 9 ligands. Docking demonstrated that ligands with scaffold 2 had a lower average free energy of binding to ALDH1A1 and ALDH1A3 when compared to ligands with the other scaffolds, indicating that having substituents on the benzene ring may enhance protein interactions through resonance stabilization and pi-pi interactions (**Figure 5**). Further pharmacokinetic screening tests, conducted through the pkCSM web server, deemed ligands 16 and 17 as non-toxic drugs, suggesting their potential as drug candidates for further investigation in both *in vitro* and *in vivo* assays (24) (**Table 1**).

Our observation of enhanced binding affinity led by the incorporation of nonpolar, bulky side chains is consistent with the literature, where a previous study showed that the addition of bulky R-groups led to more potent inhibition (7). For one, high molecular weights are associated with a larger ligand with more atoms, that hence could have more contact with the

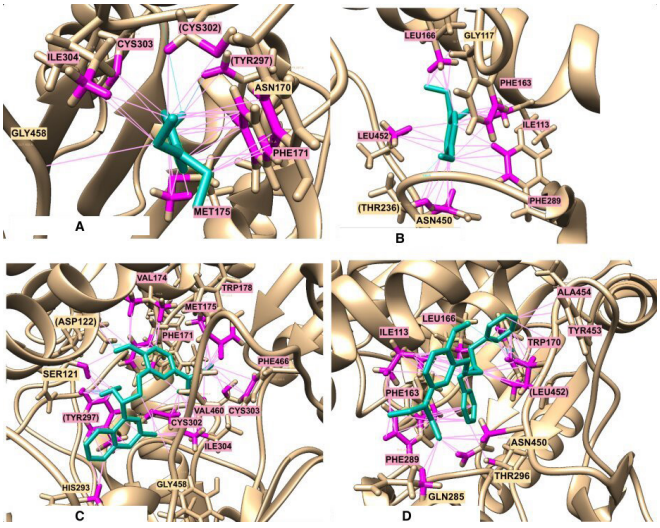

Figure 7: Binding pose of DEAB and best-performing analog to ALDH1A1 and ALDH1A3 isozymes. Ligands and hydrogen bonds are colored cyan; receptors are colored beige; van der Waals interactions between ligands and residues are colored in pink; in text, hydrophobic residues are highlighted red; neutral, hydrophilic residues are yellow; hydrogen-bonding residues are bracketed. A) DEAB on ALDH1A1. B) DEAB on ALDH1A3. C) Ligand 11, 2-[2-Amino-2-(3-methylnaphthalen-1-yl)ethoxy]-4-(diethylamino)benzaldehyde, on ALDH1A1. D) Ligand 23, 2-[3-[Bis(pyridin-2-ylmethyl)amino] propyl]-4-(diethylamino)benzaldehyde, on ALDH1A3.

https://doi.org/10.59720/24-343

enzyme, which could contribute to increased binding affinity. Higher molar refractivity contributes in a similar manner, where a molecule's higher volume and polarizability would strengthen London dispersion forces and strengthen drug-receptor interactions. The favorability of higher logP values, which signify a more non-polar, hydrophobic compound, is supported by the abundant distribution of hydrophobic residues in the binding pockets of ALDH isoenzymes (**Figure 7**). In addition, the fact that the ALDH1 substrate entry channel was found to work better for larger aldehydes, favoring higher volume substrates, supports this observation as well (15).

The presence of nonpolar residues on the binding pocket sites of ALDH1 isoenzymes suggests that these enzymes tend to create hydrophobic interactions with their ligands, especially when they exist in the aqueous environment of the cytosol. In addition, this assumption is supported by the observed trend that ligands with low free binding energy toward ALDH1A1 and ALDH1A3 have logP values higher than 2.5, meaning these ligands are hydrophobic. Also, the fact that van der Waals interactions predominantly appear in the interactions between our ligands and ALDH1 isoenzymes provides further evidence for the existence of hydrophobic interactions between ALDH1A1 and ALDH1A3 with DEAB and its analogs (**Figure 7**).

Moreover, the binding pockets of ALDH1A1 and ALDH1A3 isoenzymes contain nonpolar, aromatic residues, such as Phe171 in ALDH1A1 as well as Phe163 and Phe289 in ALDH1A3. These suggest the existence of pi-pi interactions between ligands and binding pockets of ALDH1 isoenzymes. Ligands with low free binding energy also tended to have more aromatic rings than those with higher free binding energy (**Figure 5**). One possible explanation for this observed trend is the non-polarity of aromatic groups. In a hydrophilic aqueous environment like the cytosol, hydrophobic, pi-pi interactions are favored because they help minimize the exposure of hydrophobic regions to the aqueous surroundings. This enhances the stability and binding affinity of the ligand to the receptor within the cell culture environment (8).

The top five best-performing non-toxic ligands, including ligands 16, 17 and 11, contain the motif of an ether bond connecting the benzene ring at the scaffold and the meta-R-group (**Figure 8).** This result suggests an enhanced binding affinity of ligands due to the additional hydrogen bond caused by the oxygen at the ether bond. Indeed, that theory is further supported by the pervasive distribution of residues promoting hydrogen bonding – including Asp122, Tyr297, Leu452, Thr236, and Cys302 – at ALDH1 active sites (**Figure 7**). Interestingly, both ligands 16 and 17 share a common feature: a chlorine atom positioned at the ortho position to the ether bond (**Figure 8A**). Since enhanced binding affinity correlates with a higher molar refractivity, the existence of a chlorine atom at this position likely increases the molar refractivity of the ligands, thereby elevating their binding affinity. However, the addition of another halogen – fluorine – at the other ortho position to the ether bond does not demonstrate any enhanced binding affinity benefits (**Figure 8A**). Similarly, changing the position of the chlorine atom from ortho to para to the ether bond resulted in a lowered binding affinity, as ligand 24 has a generally lower docking score than ligand 17 (**Figure 8A**). These inspections, derived from non-toxic ligands, provide essential insights for designing non-toxic inhibitors targeting ALDH1 isoenzymes.

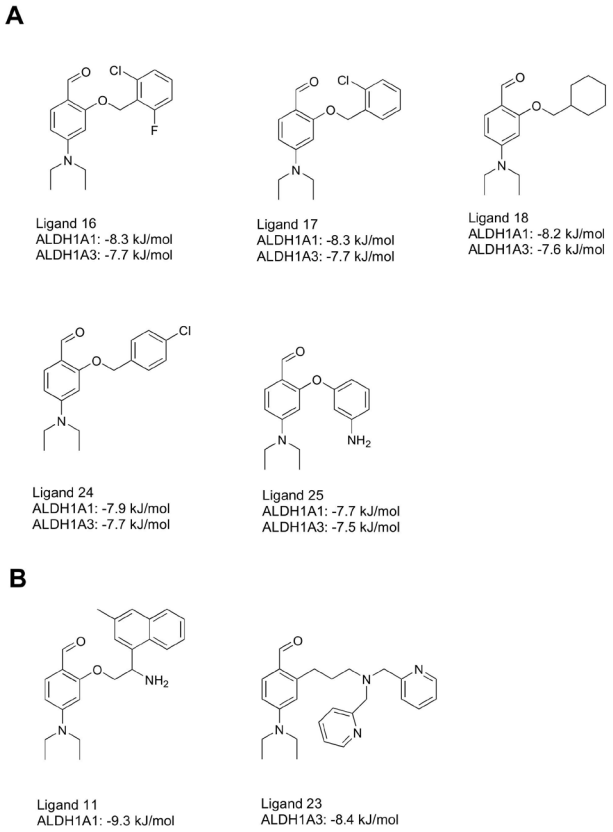

Figure 8: Structures of DEAB analogs. A) Top 5 best performing ligands that satisfies Lipinski's rule of five and passed all toxicity tests. Each ligand's docking scores against ALDH1A1 and ALDH1A3 are listed below their designation. B) Structures of ligands 11 and 23 and their docking scores against ALDH1A1 and ALDH1A3, respectively.

However, depending solely on molecular docking to predict the inhibitory effect of analogs is insufficient. Since molecular docking only provides information about the binding affinity of ligands to the protein receptor, questions about whether ligands 16 and 17 can bind stably to the ALDH1 active sites or act as irreversible or reversible inhibitors remain unresolved. Molecular docking, thus, may be limited, and molecular dynamics simulations, which utilize the computational force fields to predict the behavior of a molecule over time, may provide insights about the binding stability of ligands 16 and 17 to ALDH1A1 and ALDH1A3. Besides, *in vitro* experiments analyzing the enzyme kinetics of ligands 16 and 17 interacting with ALDHs would be required to further investigate the inhibitory mechanism and efficacy of these two analogs.

Research into the inhibition of ALDH1 isoenzymes is crucial for combating tumor proliferation and enhancing the efficacy of chemotherapy by improving drug sensitivity.

Based on our findings, future work should focus on designing and screening new DEAB analogs, leveraging the binding trends identified in this study. Further studies could use the structures of ligands 16 and 17 – which we found to be high-affinity, non-toxic inhibitors – as a foundation for ALDH1 isozyme inhibitors. Modifying them to include bulkier, nonpolar and aromatic ether groups in the R-group substituent may improve their potency as drugs. Additionally, pharmacokinetic testing is essential to ensure that the emerging drugs are

https://doi.org/10.59720/24-343

bioavailable and non-toxic. While predictive models offer a preliminary assessment of a drug's non-toxicity, *in vitro* testing of these novel DEAB analogs would be crucial for evaluating their *in vivo* viability. This approach would build on our computational findings and further elucidate binding affinities towards ALDH1 isoenzymes, possibly leading to more robust drug development strategies.

Our work highlights the impact of SAR studies, as we have successfully identified analogs derived from scaffold 2 that exhibited high *in silico* binding affinity to ALDH1A1 and ALDH1A3. As successes in drug modification require a large meta-analysis of SAR studies, our effort in modifying DEAB analogues fits into a bigger theme of research in finding new cancer treatments. The discovery of DEAB analogs derived from scaffold 2 – particularly ligands 16 and 17 – lay the groundwork for scientists to develop better ALDHs inhibitors with higher inhibitory efficacy as a new source of cancer therapeutics.

MATERIALS AND METHODS

Compiling DEAB analogs for testing

A library of 55 ligands was compiled from PubChem and Ibrahim *et al*. through a search for structures similar to DEAB (7, 16). These ligands were categorized into scaffolds 1, 2, and 3 (**Figure 4**). The R-group type was manually extracted from each ligand for trend analysis and categorized into functional groups like ethers, alkyl chains, rings, halogens, nitrogen groups, or a combination of groups. The simplified molecular input line entry system (SMILES) string for each ligand was extracted using a web scraper running Selenium on the PubChem website (16).

Bash docking

The 3D structures for our ALDH enzymes were downloaded in Protein Data Bank (PDB) format from the Research Collaboratory for Structural Bioinformatics (RCSB) PDB (codes for ALDH1A1: 4WJ9, ALDH1A2: 6ALJ, ALDH1A3: 7QK9, ALDH1B1: 7MJC) (17). The ALDH proteins were processed with Dock Prep on UCSF Chimera by removing waters, and external ligands, adding hydrogens and Gasteiger charges and were converted to mol2 format (18). For each ligand, the 3D conformer (SDF file format) was sourced from PubChem (16). Then, we prepared each ligand for docking by converting the mol2 format into the Protein Data Bank, partial charge (Q), and atom types (T) (PDBQT) format using OpenBabel (19). For our docking procedure, we utilized Dockstring, a Python library that serves as a wrapper around Autodock Vina, to automatically dock multiple ligands on the target proteins (20, 21). Our grid box parameters were based on the ALDH enzyme's catalytic site, which involves the Cys303 active residue. The lowest free binding energies for each ligand's conformer was recorded. A two-sample, one-tailed *t*-test was performed to measure statistical differences between each scaffold's average free energies of binding after docking (22).

Chemical property analysis

After docking, ligands were assessed for multiple characteristics, including molecular weight, Wildman-Crippen log partition coefficient (logP), an indication of drug hydrophobicity; number of hydrogen bond donors, number of hydrogen bond acceptors, number of rotatable bonds, number of aromatic rings, Wildman-Crippen molar refractivity value, and topological polar surface area (TPSA), using the RDKit Python package (23). We then validated which ligands fulfilled Lipinski's rule of five for drug-like compounds, and we graphed each molecular property against the free energy of binding (ΔG) for chemical property trends analysis. Through Google Sheets correlation (CORREL) and *t*-distribution (TDIST) functions, significance tests for Pearson's correlation coefficient were conducted to calculate the *R2* and *p*-value between each ligand's docking score and each of the listed chemical properties (22).

Pharmacokinetic property analysis

Pharmacokinetic properties of each ligand, including water solubility (log Mol/L), Caco-2 permeability (log Papp in 10-6 cm/s), intestinal absorption (Human) (% Absorbed), P-glycoprotein substrate (Yes/No), AMES toxicity (Yes/No), hERG I inhibitor (Yes/No), hERG II inhibitor (Yes/No), oral rat chronic toxicity (LOAEL) (log mg/kg_bw/day), and hepatotoxicity (Yes/No), were determined using the pkCSM web server (24). These metrics indicated the prediction for bioavailability and toxicity of each ligand. Ligands, which were deemed positive for AMES toxicity, hERG I-II inhibitors, and hepatotoxicity, were considered toxic toward human cell culture.

Visualization of docking poses and Ligand-Receptor interactions

UCSF Chimera (1.17.3) was used to visualize the docking poses of DEAB and other analogs on the ALDH1A1 and ALDH1A3 isozymes (18). A region within 4.0 Å between the protein and ligand was used to detect hydrogen pseudobonds. For finding van der Waals forces, contacts were calculated with a cutoff value of –0.4 Å and an allowance of 0.0 Å.

Received: August 3, 2024
Accepted: January 10, 2025
Published: July 9, 2025

REFERENCES

1. Koppaka, Vindhya, *et al*. "Aldehyde Dehydrogenase Inhibitors: A Comprehensive Review of the Pharmacology, Mechanism of Action, Substrate Specificity, and Clinical Application." *Pharmacological Reviews*, vol. 64, no. 3, Jul. 2012, https://doi.org/10.1124/pr.111.005538
2. Shortall, Kim, *et al*. "Insights into Aldehyde Dehydrogenase Enzymes: A Structural Perspective." *Frontiers in Molecular Biosciences*, vol. 13, 14 May. 2021, https://doi.org/10.3389/fmolb.2021.659550
3. Wei, Yaolu, *et al*. "Aldh1: A Potential Therapeutic Target for Cancer Stem Cells in Solid Tumors." *Frontiers in Oncology*, vol. 12, 12 Oct. 2022, https://doi.org/10.3389/fonc.2022.1026278
4. Vassalli, Giuseppe. "Aldehyde Dehydrogenases: Not Just Markers, but Functional Regulators of Stem Cells." *Stem Cells International*, 13 Jan. 2019, https://doi.org/10.1155/2019/3904645.
5. Moreb, Jan S, *et al*. "ALDH Isozymes Downregulation Affects Cell Growth, Cell Motility and Gene Expression in Lung Cancer Cells." *Molecular Cancer*, vol. 7, no. 87, 24 Nov. 2008, https://doi.org/10.1186/1476-4598-7-87
6. Javed, Shifa, *et al*. "Aldh1 & CD133 in Invasive Cervical

https://doi.org/10.59720/24-343

Carcinoma & Their Association with the Outcome of Chemoradiation Therapy." *The Indian Journal of Medical Research*, vol. 2, no. 154, Aug. 2021, pp. 367-374. https://doi.org/10.4103/ijmr.IJMR_709_20
7. Ibrahim, Ali I M, *et al*. "Expansion of the 4-(Diethylamino) Benzaldehyde Scaffold to Explore the Impact on Aldehyde Dehydrogenase Activity and Antiproliferative Activity in Prostate Cancer." *Journal of Medicinal Chemistry*, vol. 65, no. 5, 10 Mar. 2022, pp. 3833-3848, https://doi.org/10.1021/acs.jmedchem.1c01367
8. Qiu, Yan, *et al*. "The Expression of Aldehyde Dehydrogenase Family in Breast Cancer." *Journal of Breast Cancer*, vol. 17, no. 1, Mar. 2014, pp. 54-60. https://doi.org/10.4048/jbc.2014.17.1.54
9. Singh, S, *et al*. "Acetaldehyde and Retinaldehyde-Metabolizing Enzymes in Colon and Pancreatic Cancers." *Advances in Experimental Medicine and Biology*, vol. 815, Nov. 2014, pp. 281-294, https://doi.org/10.1007/978-3-319-09614-8_16
10. Wang, Ling, *et al*. "Aldehyde Dehydrogenase 1 in Gastric Cancer." *Journal of Oncology*, Mar. 2022, https://doi.org/10.1155/2022/5734549
11. Chen, Miao-Fen, *et al*. "Role of Aldh1 in the Prognosis of Esophageal Cancer and Its Relationship with Tumor Microenvironment." *Molecular Carcinogenesis*, Sep. 2017, pp. 78-88, https://doi.org/10.1002/mc.22733
12. Dong, Yue, *et al*. "Aldehyde Dehydrogenase 1 Isoenzyme Expression as a Marker of Cancer Stem Cells Correlates to Histopathological Features in Head and Neck Cancer: A Meta-Analysis." *PloS One*, vol. 12, no. 11, Nov. 2017, https://doi.org/10.1371/journal.pone.0187615
13. Morgan, Cynthia A, *et al*. "N,N-Diethylaminobenzaldehyde (DEAB) as a Substrate and Mechanism-Based Inhibitor for Human ALDH Isoenzymes." *Chemico-Biological Interactions*, vol. 234, 5 June 2015, pp. 18-28, https://doi.org/10.1016/j.cbi.2014.12.008.
14. Zhou, Lei, *et al*. "Identification of Cancer-Type Specific Expression Patterns for Active Aldehyde Dehydrogenase (ALDH) Isoforms in ALDEFLUOR Assay." *Cell Biology and Toxicology*, vol. 35, Sep. 2019, pp. 161-177. https://doi.org/10.1007/s10565-018-9444-y
15. Shortall, Kim, *et al*. "Insights into Aldehyde Dehydrogenase Enzymes: A Structural Perspective." *Frontiers*, vol. 8, 28 Apr. 2021, https://doi.org/10.3389/fmolb.2021.659550
16. Kim, Sunghwan, *et al*. "PubChem 2025 update." *Nucleic Acids Research*, vol. 53, 6 Jan. 2025, https://doi.org/10.1093/nar/gkae1059
17. Berman, Helen M, *et al*. "The Protein Data Bank." *OUP Academic*, Oxford University Press, vol. 28, no. 1, 1 Jan. 2000, pp. 235-242. https://doi.org/10.1093/nar/28.1.235
18. Petterson, Eric F, *et al*. "UCSF Chimera--a Visualization System for Exploratory Research and Analysis." *Journal of Computational Chemistry*, vol. 25, Jul. 2004, pp. 1605-1612. https://doi.org/10.1002/jcc.20084
19. O'Boyle, Noel M, *et al*. "Open Babel: An Open Chemical Toolbox." *Journal of Cheminformatics*, vol. 3, no. 33, 7 Oct. 2011, https://doi.org/10.1186/1758-2946-3-33
20. Eberhardt, Jerome, *et al*. "AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings." *Journal of Chemical Information and Modeling*, vol. 61, no. 8, Jul. 2021, https://doi.org/10.1021/acs.jcim.1c00203
21. Trott, Oleg, *et al*. "Autodock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading." *Journal of Computational Chemistry*, vol. 31, no. 2, 30 Jan. 2010, pp. 455-461. https://doi.org/10.1002/jcc.21334
22. Student. "The Probable Error of a Mean." *Biometrika*, vol. 6, no. 1, 1908, pp. 1-25., https://doi.org/10.2307/2331554
23. Landrum, Greg, *et al*. "rdkit/rdkit: 2020_03_1 (Q1 2020) Release." Zenodo, 29 Mar. 2020. https://doi.org/10.5281/zenodo.3732262
24. Pires, Douglas E V, *et al*. "PKCSM: Predicting Small-Molecule Pharmacokinetic and Toxicity Properties Using Graph-Based Signatures." *Journal of Medicinal Chemistry*, vol. 58, no. 9, 14 May 2015, pp. 4066-4072. https://doi.org/10.1021/acs.jmedchem.5b00104

Appendix

Github repository for the code

Link to Github page containing the code used in this project: https://github.com/kennardliong/jei-scripts