JOURNAL OF
EMERGING INVESTIGATORS

Article

High-throughput virtual screening of novel dihydropyrimidine monastrol analogs reveals robust structure-activity relationship to kinesin Eg5 binding thermodynamics

Tyler Shern[1,4], Ansh Rai[1,4], Krithikaa Premnath[2,4], Audrey Kwan[3,4], Ria Kolala[1,4], Ishani Ashok[1,4], Edward Njoo[4]

[1]Mission San Jose High School, Fremont, CA
[2]Dougherty Valley High School, San Ramon, CA
[3]Dublin High School, Dublin, CA
[4]Department of Chemistry, Biochemistry & Physics, Aspiring Scholars Directed Research Program, Fremont, CA

SUMMARY

As cancer continues to take millions of lives worldwide, the need to create effective therapeutics for the disease persists. The kinesin Eg5 assembly motor protein is a promising target for cancer therapeutics as inhibition of this protein leads to cell cycle arrest. Monastrol, a small dihydropyrimidine-based molecule capable of inhibiting the kinesin Eg5 function, has attracted the attention of medicinal chemists with its potency, affinity, and specificity to the highly targeted loop5/α2/α3 allosteric binding pocket. In this work, we employed high-throughput virtual screening (HTVS) to identify potential small molecule Eg5 inhibitors from a designed set of novel dihydropyrimidine analogs structurally similar to monastrol. Density functional theory (DFT) calculations and protein-ligand docking experiments revealed that the analogs with geranyl ester substitutions exhibited the greatest binding affinities to the allosteric binding pocket of kinesin Eg5. In-depth analysis of the binding pocket amino acid residues and calculations of the cLogP value for each compound demonstrated qualitatively and quantitatively that strong hydrophobic interactions of the ester functionality with kinesin Eg5 are of great significance in the improved binding of dihydropyrimidine analogs. This establishment of a quantitative structure-activity relationship to kinesin Eg5 binding thermodynamics using HTVS revealed the discovery of improved dihydropyrimidine-based inhibitors capable of advancing society's progress in the fight against cancer.

INTRODUCTION

In 2019, cancer was responsible for over 600,000 deaths throughout the United States (1). Though cancer is currently the second highest leading cause of death, the mortality rates for this disease have declined by 27% over the past 25 years, and these improvements in survival rates have been attributed mainly to advances in early detection and treatment (2). Many of the successful small molecule therapeutics for cancer have been the product of years of intensive drug discovery research, involving the identification of potential candidates through screening of naturally derived or synthetic compounds.

Kinesin Eg5 is a motor protein involved in the assembly and separation of mitotic spindle fibers in the cell cycle and plays a critical role in the establishment of spindle bipolarity (3, 4). Eg5 is generally not expressed in non-proliferating adult tissues and thus results in diminished toxicity when treated with Eg5-targeted therapies, especially when compared to other modern anti-mitotic therapeutics available on the market (5). Researchers have considered the inhibition of kinesin Eg5 as a high potential avenue for cancer therapy, especially with Eg5 overexpressed in breast carcinogenesis, laryngeal squamous cell carcinoma, astrocytic neoplasm, prostate cancer, bladder cancer, and renal cell carcinoma (6).

In 2000, the Mitchison group reported the discovery of monastrol, a small molecule dihydropyrimidine (DHPM) that allosterically inhibits kinesin Eg5 (Figure 1) (7). Following this revelation, DHPMs have gained significant interest in medicinal chemistry. The synthesis of other DHPMs structurally similar to monastrol has demonstrated great potential in the development of increasingly potent anticancer agents capable of treating aggressive glioma, renal, and breast cancers in past studies (8).

Monastrol binds to the loop5/α2/α3 allosteric site of kinesin Eg5 through hydrophobic interactions and hydrogen bonding, inducing conformational changes in Eg5 and preventing continued mitotic division (9). With kinesin Eg5 expression considered to be generally overexpressed in neoplastic tissue, the high selectivity of monastrol to mitotically active cancer cells suggests strengthened responses and fewer cytotoxic side effects when compared to other anticancer agents (6). As previous in vitro and in vivo studies have concluded that the S-monastrol is a more potent inhibitor of kinesin-Eg5 than the R-enantiomer, this work focuses solely on the former (10). Given the significance of biological activity that monastrol produces, it has been of deep interest to use the DHPM scaffold to create more potent inhibitors of kinesin Eg5.

High-throughput virtual screening (HTVS) provides an efficient computational method to discover the most effective compounds from massive libraries of analogs to a set of finite user-defined conditions. Molecular docking is a technique in which the thermodynamic efficiency of molecular interactions is calculated, shedding time-efficient, detailed insight into the specificity of binding for improved lead optimization. These values can be used in the prediction of the binding thermodynamics and preferred binding poses of small molecule ligands to protein targets (11). In this work, we hypothesize that the binding affinity of such analogs is primarily related to hydrophobic interactions that the ester functionality is engaged in with hydrophobic residues of the reported allosteric binding pocket.

This study focused on the computational screening of systematic modifications of the aromatic and ester constituents of monastrol, as displayed in Figure 2. The analogs presented in this study encompass a variety of aromatic and ester substitutions (Figure 3). We implemented HTVS to establish a robust, quantitative structure-activity relationship (QSAR) between the binding affinities of our novel dihydropyrimidine analogs and kinesin Eg5. We assessed 100 analogs for increased antiproliferative abilities in hopes of discovering improved small molecule cancer therapeutics. The results not only indicated that analogs with geranyl substitutions represent lead compounds that should be synthesized, but also revealed that enhanced kinesin Eg5 binding thermodynamics are attributed greatly to hydrophobic interactions between the ester functionality and Eg5.

RESULTS

The first computational docking software used was Swissdock. Each of the 100 analogs was docked onto the kinesin Eg5 receptor (PDB: 3HQD) with the grid box center having coordinates of (21.738, 26.509, 51.081) and the grid box dimensions being (40, 44, 40). Predicted free energies of binding (ΔG) and the differences in the binding affinities between each analog and the original monastrol molecule are reported in Table 1. As AR12 has the same chemical structure as monastrol, we compared each of the Swissdock binding energies for the other 99 analogs against the control monastrol ΔG value of -7.31 kcal/mol.

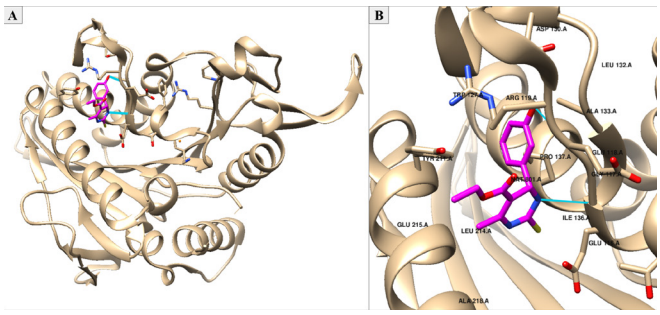

Figure 1: Human kinesin Eg5 motor protein structure. (A) The crystal structure of the human kinesin Eg5 motor protein bound to monastrol in the loop5/α2/α3 allosteric binding pocket. The magenta structure is monastrol, while the gold structure is the kinesin Eg5 protein. We extracted the crystallized protein structure from the Protein Data Bank (PDB: 1X88) and visualized it using UCSF Chimera. (B) A closer visualization of the binding site, with monastrol engaging in several hydrophobic interactions with nearby residues including TRP211, ILE136, PRO137, LEU214, TYR211, ARG119, TRP127, and ALA133. Hydrogen bonding (cyan sticks) occurs between the phenolic oxygen of monastrol and ARG119 as well as between a nitrogen of monastrol and GLY117.

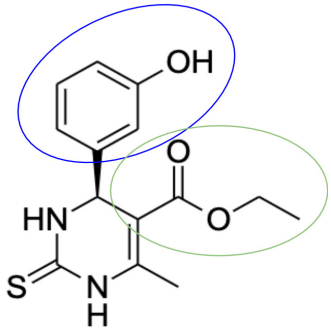

Figure 2: Structure of the S-enantiomer of monastrol, highlighting the 3-hydroxyphenyl aryl (circled in blue) and ethyl ester (circled in green) groups. We modified these circled aryl and ester functionalities during the high throughput virtual screening process to create 100 novel dihydropyrimidine analogs using the molecular editor software Avogadro. We created and edited this chemical structure using ChemDraw.

The best analogs depicted by the Swissdock computations were those with the decyl and geranyl ester substitutions (T1-T20), as indicated by the common trend of ΔG values being less than -8 kcal/mol. As these analogs reported the most negative ΔG values, there were increased amounts of energy released in the exothermic processes, forming comparatively more stable protein-ligand complexes than the other analogs.

Our results also revealed that the analogs R1-R20 consisting of phenoxyethyl and butoxyethyl ester substitutions exhibited relatively high binding affinities too, with the majority of analogs having ΔG values less than -8 kcal/mol. The analogs that had the least binding affinities and resulted in the most positive ΔG values were analogs AR1-AR20, which consisted of methyl and ethyl ester substitutions. With the decyl, geranyl, phenoxyethyl, and butoxyethyl analogs possessing greater amounts of carbon in the ester group than the methyl and ester analogs, the data suggests that the magnitude of ΔG values increases with additional carbons attached to the growing ester chain.

With a ΔG of -8.71 kcal/mol, T18 (which consisted of geranyl ester and 2-nitro aryl substitutions) exhibited the greatest binding affinity to the kinesin Eg5 receptor (Figure 4). When docked to Eg5, T18 presented a binding energy that was 1.40 kcal/mol greater than the control binding affinity of monastrol as computationally determined by Swissdock, a significant difference especially compared to the other analogs with various ester substitutions. The structure-activity relationship of this analog as computed by Swissdock suggests a higher likelihood of exhibiting a greater biological

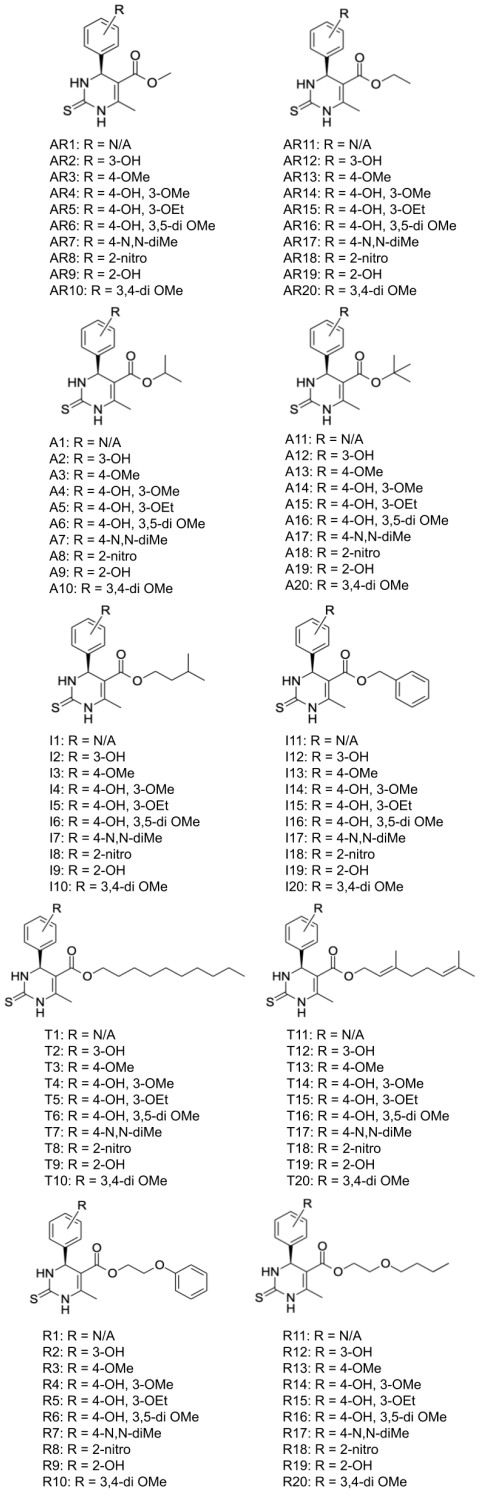

Figure 3: The 100 DHPM analogs with selected ester and aryl substitutions for the S-enantiomer of monastrol. The ester substitutions we made include methyl, ethyl, isopropyl, tert-butyl, isoamyl, benzyl, decyl, geranyl, phenoxyethyl, and butoxyethyl functional groups. The aryl substitutions that we made include phenyl, 3-hydroxy, 4-methoxy, 4-hydroxy-3-methoxy, 4-hydroxy-3-methoxy, 4-hydroxy-3,5-dimethoxy, 4-N,N-dimethylamino, 2-nitro, 2-hydroxy, and 3,4-dimethoxy functional groups. We created the chemical structures and the detailed aryl substitution text using ChemDraw.

potency than the other molecules. Conversely, with methyl ester and phenyl substitutions as well as a ΔG of -6.98 kcal/mol, we discovered that AR1 was the analog with the lowest binding affinity to kinesin Eg5 (Figure 5); when computed by Swissdock, the binding affinity of AR1 was 0.33 kcal/mol less than that of monastrol, indicating an even weaker interaction with kinesin Eg5 than the control.

We then used the Docking Incrementally (DINC) protein ligand docking software to conduct the next series of computational docking experiments. We docked each analog onto the same receptor of kinesin Eg5 using the appropriate grid box coordinates, with the center of the grid box having coordinates of (21.738, 26.509, 51.081) and the grid box dimensions being (40, 44, 40). The thermodynamic results for each analog are documented in Table 1. Again, as AR12 has the same chemical structure as monastrol, we compared each of the DINC binding energies for the other 99 analogs against the control ΔG value of -6.0 kcal/mol.

We color coded each of these values using the same set of ΔG critical values as in the Swissdock experiments. However, the DINC results depicted an overall diminished magnitude of ΔG values in comparison to the Swissdock thermodynamic outputs. Similar to the Swissdock results, analogs T11-20 were the most thermodynamically efficient binders, indicating once again that the analogs with the highest binding affinities were those with the geranyl ester substitutions. In contrast to Swissdock, however, the analogs with decyl ester substitutions (T1-T10) docked with DINC did not reveal substantial binding affinities. The R1-10 analogs containing phenoxyethyl ester substitutions displayed a similar trend of greater ΔG values with DINC in comparison to the other analogs, suggesting that these are also potential lead compounds that should be synthesized and biologically tested. The most unsuccessful binders were overall seen again to be AR1-20, the analogs with methyl and ethyl ester substitutions.

With geranyl ester and 4-hydroxy-3,5-dimethoxy aryl substitutions, we determined T16 as the most thermodynamically effective binder as it exhibited a ΔG value of -8.3 kcal/mol (Figure 6); the binding affinity for T16 was 2.3 kcal/mol greater than the control binding energy of monastrol determined using DINC. This again suggests how the geranyl ester functionality serves a key role in the strengthened binding to kinesin Eg5 and indicates a possible greater biological potency than the other analogs. With a ΔG value of -5.5 kcal/mol, we discovered that AR16 (which consisted of ethyl ester and 4-hydroxy-3,5-dimethoxy aryl substitutions) was the analog with the lowest binding affinity to kinesin Eg5 (Figure 7); the binding affinity of AR16 was 0.4 kcal/mol less than that of monastrol when calculated with DINC, indicating even weaker interactions with kinesin Eg5 than the control.

Following the molecular docking of the analogs using Swissdock and DINC, we employed the OSIRIS Property Explorer to determine the calculated logarithm of the partition coefficient (cLogP) for each analog and evaluate the relative

Analog Name	DINC	DINC: Difference with Monastrol	Swissdock	Swissdock: Difference with Monastrol	cLogP Value	Analog Name	DINC	DINC: Difference with Monastrol	Swissdock	Swissdock: Difference with Monastrol	cLogP Value
AR1	-5.8	0.2	-6.98	0.33	1.66	AR11	-6.1	-0.1	-7.14	0.17	2.07
AR2	-5.8	0.2	-7.37	-0.06	1.32	AR12	-6	0	-7.31	0	1.72
AR3	-5.8	0.2	-7.36	-0.05	1.59	AR13	-5.7	0.3	-7.57	-0.26	2
AR4	-5.8	0.2	-7.29	0.02	1.25	AR14	-6.2	-0.2	-7.78	-0.47	1.65
AR5	-6	0	-7.84	-0.53	1.65	AR15	-6.2	-0.2	-7.56	-0.25	2.34
AR6	-5.6	0.4	-7.42	-0.11	1.18	AR16	-5.5	0.5	-7.66	-0.35	1.58
AR7	-5.7	0.3	-7.4	-0.09	1.56	AR17	-6	0	-7.33	-0.02	1.97
AR8	-6.1	-0.1	-7.56	-0.25	0.74	AR18	-6.1	-0.1	-7.88	-0.57	1.15
AR9	-5.8	0.2	-7.08	0.23	1.32	AR19	-5.9	0.1	-7.05	0.26	2.06
AR10	-5.6	0.4	-7.7	-0.39	0.97	AR20	-5.6	0.4	-7.73	-0.42	1.93
A1	-6	0	-7.14	0.17	2.43	A11	-6.4	-0.4	-7.32	-0.01	2.81
A2	-6.1	-0.1	-7.53	-0.22	2.08	A12	-6.4	-0.4	-7.63	-0.32	2.46
A3	-6.2	-0.2	-7.45	-0.14	2.36	A13	-6.3	-0.3	-7.42	-0.11	2.74
A4	-5.7	0.3	-7.76	-0.45	2.01	A14	-6.2	-0.2	-7.63	-0.32	2.39
A5	-6.1	-0.1	-7.57	-0.26	2.42	A15	-6.5	-0.5	-7.78	-0.47	2.8
A6	-5.8	0.2	-7.57	-0.26	1.94	A16	-5.7	0.3	-8.03	-0.72	2.32
A7	-6.3	-0.3	-7.3	0.01	2.32	A17	-6.4	-0.4	-7.54	-0.23	2.71
A8	-6.3	-0.3	-7.55	-0.24	1.51	A18	-6.7	-0.7	-7.58	-0.27	1.89
A9	-6.3	-0.3	-7.6	-0.29	2.06	A19	-6.5	-0.5	-7.38	-0.07	2.46
A10	-5.7	0.3	-7.76	-0.45	2.29	A20	-5.9	0.1	-7.82	-0.51	2.67
I1	-6.8	-0.8	-7.36	-0.05	3.2	I11	-6.5	-0.5	-7.53	-0.22	3.08
I2	-6.8	-0.8	-7.6	-0.29	2.85	I12	-6.8	-0.8	-7.58	-0.27	2.73
I3	-7	-1	-7.67	-0.36	3.13	I13	-7.3	-1.3	-7.74	-0.43	3.01
I4	-7	-1	-7.78	-0.47	2.78	I14	-6.2	-0.2	-7.59	-0.28	2.67
I5	-7.2	-1.2	-7.86	-0.55	3.19	I15	-6.7	-0.7	-7.75	-0.44	3.07
I6	-6.3	-0.3	-7.91	-0.6	2.71	I16	-6.3	-0.3	-8.03	-0.72	2.59
I7	-6.3	-0.3	-7.84	-0.53	3.09	I17	-6.2	-0.2	-7.88	-0.57	2.98
I8	-7	-1	-8.04	-0.73	2.27	I18	-7.7	-1.7	-7.97	-0.66	2.16
I9	-7.1	-1.1	-7.65	-0.34	2.85	I19	-7.7	-1.7	-7.91	-0.6	2.73
I10	-6.3	-0.3	-8	-0.69	3.06	I20	-6.2	-0.2	-7.84	-0.53	2.94
T1	-7.1	-1.1	-8.3	-0.99	5.7	T11	-7.4	-1.4	-8.29	-0.98	5.5
T2	-6.9	-0.9	-8.25	-0.94	5.36	T12	-7.7	-1.7	-8.23	-0.92	5.16
T3	-6.9	-0.9	-8.35	-1.04	5.63	T13	-8.1	-2.1	-8.29	-0.98	5.43
T4	-6.9	-0.9	-8.11	-0.8	5.29	T14	-7.3	-1.3	-8.41	-1.1	5.09
T5	-6.4	-0.4	-8.12	-0.81	5.69	T15	-8.1	-2.1	-8.62	-1.31	5.49
T6	-6.8	-0.8	-8.63	-1.32	5.22	T16	-8.3	-2.3	-8.46	-1.15	5.02
T7	-6.9	-0.9	-8.17	-0.86	5.6	T17	-8.1	-2.1	-8.31	-1	5.4
T8	-6.8	-0.8	-8.25	-0.94	4.78	T18	-8.1	-2.1	-8.71	-1.4	4.58
T9	-7.1	-1.1	-8.16	-0.85	5.36	T19	-7.5	-1.5	-8.1	-0.79	5.16
T10	-6.7	-0.7	-8.56	-1.25	5.56	T20	-7.6	-1.6	-8.33	-1.02	5.36
R1	-7.4	-1.4	-8.13	-0.82	3.09	R11	-7	-1	-7.8	-0.49	2.89
R2	-7.5	-1.5	-8.09	-0.78	2.74	R12	-7.1	-1.1	-8.08	-0.77	2.54
R3	-7.4	-1.4	-7.87	-0.56	3.02	R13	-6.6	-0.6	-8.1	-0.79	2.82
R4	-7.6	-1.6	-8.07	-0.76	2.67	R14	-7.6	-1.6	-8.22	-0.91	2.47
R5	-7.8	-1.8	-8.36	-1.05	3.08	R15	-7.6	-1.6	-8.4	-1.09	2.88
R6	-7.4	-1.4	-8.43	-1.12	2.6	R16	-6.9	-0.9	-8.23	-0.92	2.4
R7	-7.3	-1.3	-8.69	-1.38	2.96	R17	-6.9	-0.9	-8.1	-0.79	2.78
R8	-8	-2	-8.14	-0.83	2.17	R18	-7.3	-1.3	-7.73	-0.42	1.96
R9	-7.3	-1.3	-7.92	-0.61	2.74	R19	-7.1	-1.1	-8.08	-0.77	2.54
R10	-7.6	-1.6	-8.12	-0.81	2.95	R20	-6.7	-0.7	-8.13	-0.82	2.75

	ΔG less than -8.00 kcal/mol
	ΔG between -7.00 and -7.99 kcal/mol
	ΔG between -6.00 and -6.99 kcal/mol
	ΔG less than -5.99 kcal/mol

Table 1: A heatmap of DINC and Swissdock results color-coded based on the calculated ΔG value. We docked each analog (entitled AR1, AR2, etc.) onto the human kinesin Eg5 protein, with the pose of the most negative binding affinity selected. For each analog, there exists the DINC binding affinity value, the difference in binding affinity (using DINC) between the modeled analog and monastrol, the Swissdock binding affinity value, the difference in binding affinity (using Swissdock) between the modeled analog and monastrol, and the cLogP value for each analog calculated through the OSIRIS Property Explorer. The ΔG values from the DINC results are overall more positive than those attained using Swissdock.

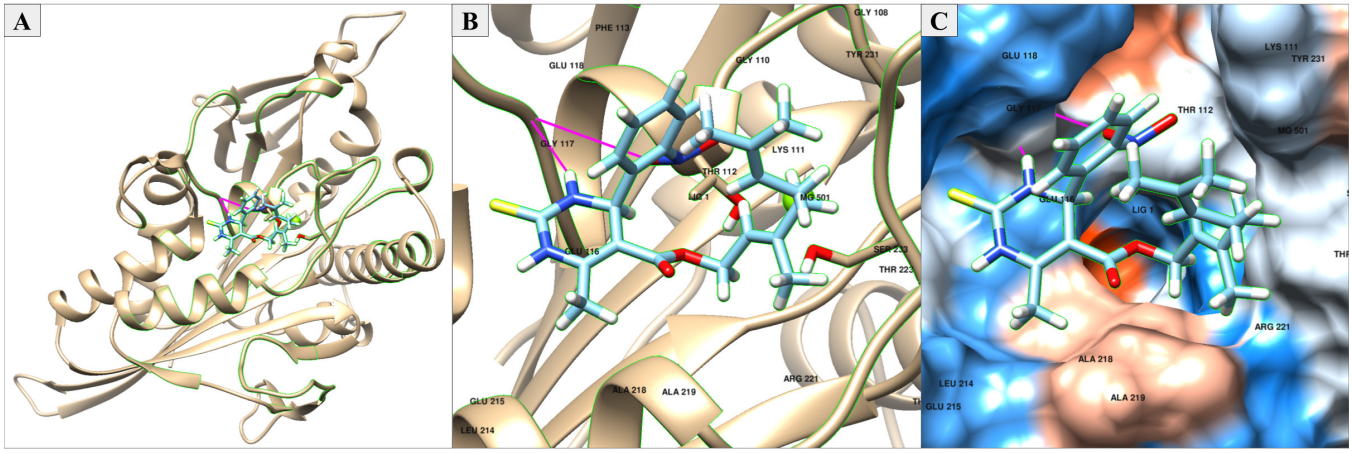

Figure 4: Swissdock results for the T18 analog. (A) The docked structure of the ligand with the highest binding affinity (T18) according to the Swissdock ΔG values obtained using UCSF Chimera, with a ΔG value of -8.71 kcal/mol. The cyan structure is the T18 analog, while the gold structure is the human kinesin Eg5 protein. (B) A detailed visualization of the protein-ligand interaction in the binding pocket, with significant hydrophobic interactions and two hydrogen bonds (magenta) occurring within 4.0 angstroms of the ligand. (C) A hydrophobicity surface model is depicted, with hydrophobic interactions (red), hydrophilic interactions (white), and neutral interactions (white) shown between T18 and the amino acid residues within the loop5/α2/α3 allosteric binding site.

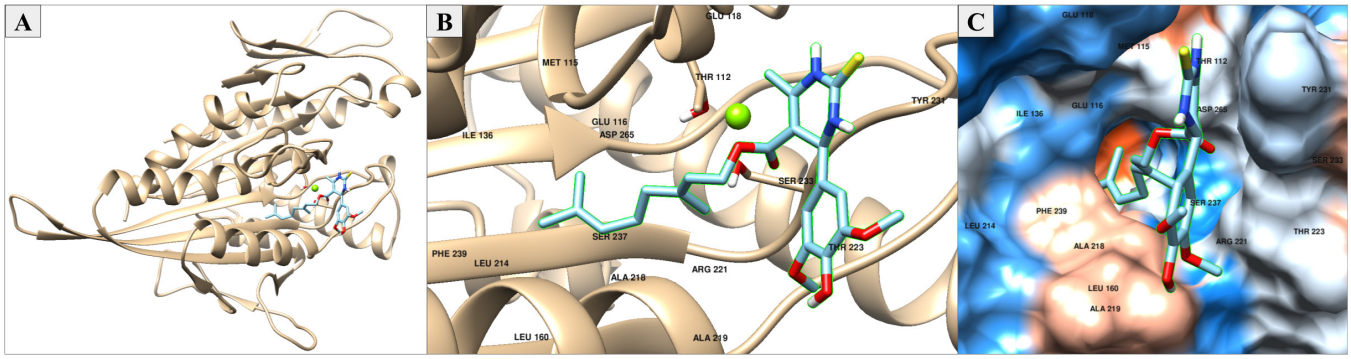

Figure 5: Swissdock results for the AR1 analog. (A) The docked structure of the ligand with the lowest binding affinity (AR1) according to the Swissdock ΔG values obtained using UCSF Chimera, with a ΔG value of -6.98 kcal/mol. The cyan structure is the AR1 analog, while the gold structure is the human kinesin Eg5 protein. (B) A detailed visualization of the protein-ligand interaction in the binding pocket, with hydrophobic interactions, hydrophilic interactions, and hydrogen bonding (magenta) occurring within 4.0 angstroms of the ligand. (C) A hydrophobicity surface model is depicted, with hydrophobic interactions (red), hydrophilic interactions (white), and neutral interactions (white) shown between AR1 and the amino acid residues within the loop5/α2/α3 allosteric binding site.

amount of hydrophobic interactions that each could engage in. We documented the cLogP value for each analog in Table 1. All of the constructed dihydropyrimidine analogs exhibited positive cLogP values, indicating a tendency for these molecules to have hydrophobic cores. We compared these molecules to the control cLogP value (AR12) of 1.72 to indicate the relative extent of hydrophobic interactions occurring in the binding pocket. With cLogP values all greater than 4.58, T1-20 (the decyl and geranyl analogs) exhibited the highest values compared to the other analogs, indicating that these molecules had the most substantial hydrophobicities. T18, the analog with the greatest binding affinity discovered through Swissdock, demonstrated a cLogP value of 4.58, which was greater than that of monastrol by 2.86. Though this factor of 2.86 may not initially resemble a statistically significant difference, it is important to note that the partition coefficients have been logarithmically transformed; thus, an increase in cLogP of 2.86 is actually an increase in hydrophobicity by a magnitude of 102.86. Similarly, we calculated T16, the analog with the greatest binding affinity discovered through DINC, as having a cLogP value of 5.02, which was greater than that of monastrol by 3.30. In the analysis of the molecules with the lowest binding affinities determined by Swissdock and DINC, we discovered that AR1 and AR16 displayed cLogP values of 1.66 and 1.58 respectively, representing diminished potentials to bind hydrophobically with kinesin Eg5 than even monastrol itself. In combination with the corroborating data indicating that the geranyl analogs had the greatest binding affinities, we presumed that improved binding to the receptor of the kinesin Eg5 protein is based on greater inherent abilities for strong hydrophobic interactions to occur between the dihydropyrimidine molecule and the protein.

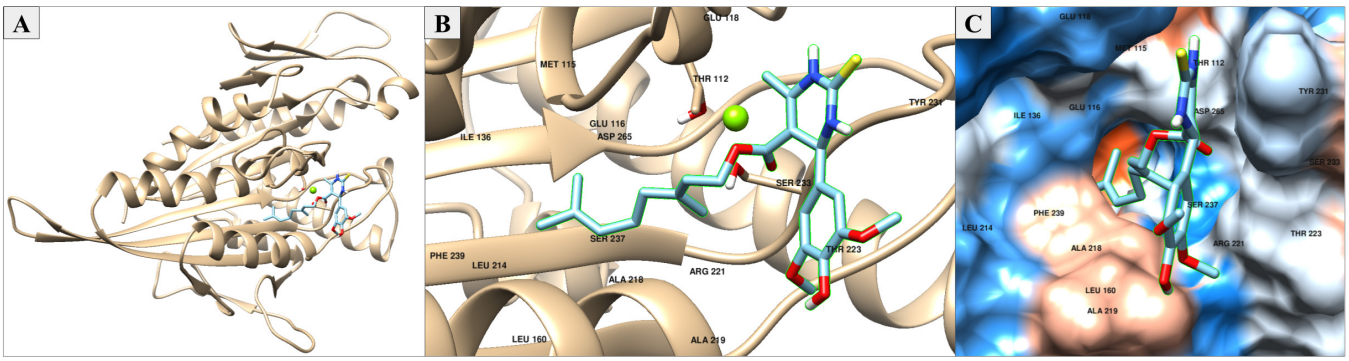

Figure 6: DINC results for the T16 analog. (A) The docked structure of the ligand with the highest binding affinity (T16) according to the DINC ΔG values obtained using UCSF Chimera, with a ΔG value of -8.3 kcal/mol. The cyan structure is the T16 analog, while the gold structure is the human kinesin Eg5 protein. (B) A detailed visualization of the protein-ligand interaction in the binding pocket, with significant hydrophobic interactions occurring within 4.0 angstroms of the ligand. (C) A hydrophobicity surface model is depicted, with hydrophobic interactions (red), hydrophilic interactions (white), and neutral interactions (white) shown between T16 and the amino acid residues within the loop5/α2/α3 allosteric binding site.

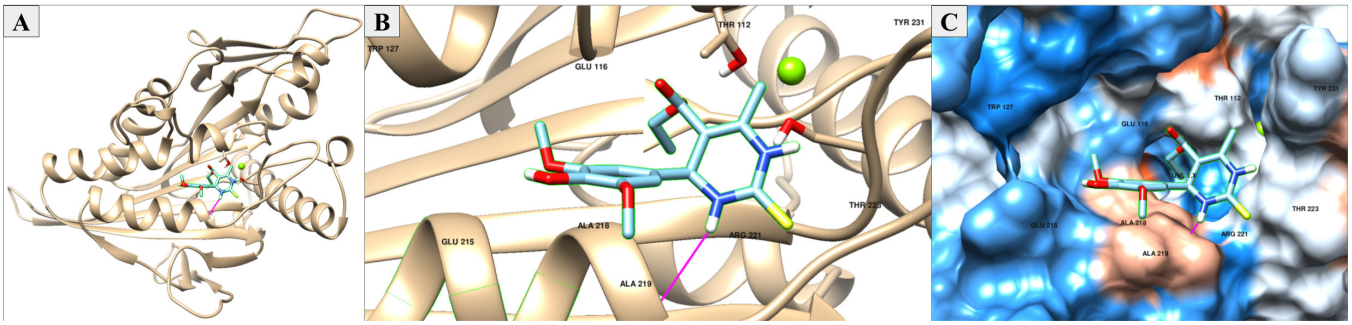

Figure 7: DINC results for the AR16 analog. (A) The docked structure of the ligand with the lowest binding affinity (AR16) according to the DINC ΔG values obtained using UCSF Chimera, with a ΔG value of -5.5 kcal/mol. The cyan structure is the AR16 analog, while the gold structure is the human kinesin Eg5 protein. (B) A detailed visualization of the protein-ligand interaction in the binding pocket, with hydrophilic interactions, hydrophobic interactions and hydrogen bonding (magenta) occurring within 4.0 angstroms of the ligand. (C) A hydrophobicity surface model is depicted, with hydrophobic interactions (red), hydrophilic interactions (white), and neutral interactions (white) shown between AR16 and the amino acid residues within the loop5/α2/α3 allosteric binding site.

DISCUSSION

While the ΔG values given by each docking algorithm were slightly different, a common trend emerged in the role of specific functionalities in binding affinity. As seen from the strongly negative ΔG values calculated by Swissdock and DINC, the geranyl ester substitution analogs were the best binders to kinesin Eg5, likely due to the characteristic ability of nonpolar geranyl groups for increased hydrophobic interactions with the binding pocket. All of the geranyl analogs displayed relatively high cLogP values (all greater than 4.58) when compared to the cLogP values of other analogs with varying ester substitutions; this confirmed the great amount of hydrophobic interactions that the geranyl analogs have the potential to engage in within the binding pocket.

The Swissdock results also indicated that the decyl analogs were particularly effective binders to kinesin Eg5 as they also exhibited strong negative ΔG values. Because of the heightened nonpolar abilities of the decyl group, we presumed that this ester functionality leads to stronger hydrophobic interactions to occur during protein-ligand binding. The specific interactions between the amino acid residues of the receptor and the decyl analogs can again be attributed to strong hydrophobic interactions as suggested by the high cLogP values (all greater than 4.78) calculated for each molecule.

Both Swissdock and DINC results corroborated the methyl and ethyl analogs as the worst binders to kinesin Eg5 as indicated by the least negative ΔG values. With the cLogP values calculated for these analogs relatively low (all less than 2.34) compared to analogs with other ester substitutions, the results suggested that the diminished hydrophobicities of the methyl and ethyl groups correlate with weaker intermolecular interactions during binding.

In the binding pocket analysis of T18, the analog with the greatest binding affinity according to Swissdock, we discovered the molecule engaging in hydrophobic interactions and hydrogen bonding with kinesin Eg5. Given that the aliphatic tail was in close proximity (within 4.0 angstroms) to the hydrophobic LEU214, ALA218, ALA219, GLY110, GLY117, PHE113, and TYR231 residues, it suggested there exists a high level of hydrophobic interactions between the ligand and receptor (Figure 4). Additionally, two hydrogen bonds occurred

between T18 and GLY117, as indicated by the magenta bonds in Figure 4B. The strong hydrophobic interactions (in red) with the aforementioned leucines, alanines, glycines, phenylalanines, and tyrosines in the binding pocket can be visualized in Figure 4C, confirming the high hydrophobicity of T18 denoted by the cLogP value of 4.58.

Though we noticed that AR1 (the analog with the lowest binding affinity according to Swissdock) also engaged in hydrophobic interactions and hydrogen bonding, we discovered substantial hydrophilic interactions in the binding pocket too. While T18 hydrogen bonded with GLY117, a hydrogen bond instead occurred between AR1 and GLU118, as indicated in Figure 5. As the aliphatic tail was verified to be in close proximity to the hydrophobic LEU214, LEU171, GLY117, GLY217, ALA218, ALA219, and TYR231 residues, it established that hydrophobic interactions did occur to a small extent between the ligand and receptor (Figure 5B). However, upon closer examination of the hydrophobicity surface in Figure 5C, we distinguished that a significant number of hydrophilic residues (in blue) of kinesin Eg5 including ARG119, ARG221, TRP127, GLU116, and GLU118 also interacts with the ligand, likely resulting in its decreased binding affinity. The increased amount of visualized hydrophilic interactions confirmed the weakly hydrophobic nature of AR1 implied by the cLogP value of 1.66.

Similar to the analysis of T18, we discovered that T16 (the analog with the greatest binding affinity according to DINC) also engaged in substantial hydrophobic interactions with kinesin Eg5 as the aliphatic tail was in close proximity to the hydrophobic ALA218, ALA219, TYR231, PHE239, LEU160, LEU214, MET115, and ILE136 residues (Figure 6). These strong hydrophobic interactions with the aforementioned amino acids in the binding pocket can be visualized in Figure 6C, confirming the high hydrophobicity indicated by the T16 cLogP value of 5.02.

In the binding pocket analysis of AR16, the analog with the lowest binding affinity according to DINC, we again discovered that although the molecule did engage in hydrophobic interactions, a considerable amount of hydrogen bonding and hydrophilic interactions occurred. As seen in Figure 7, a hydrogen bond was visualized between AR16 and ALA219. Given that the aliphatic tail was in close proximity to the hydrophobic ALA218, ALA219, and TYR231 residues, it suggested that minimal hydrophobic interactions occurred between the ligand and receptor (Figure 7B). However, a detailed inspection of the hydrophobicity surface in Figure 7C revealed a significant number of hydrophilic regions including TRP127, GLU116, GLU215, ARG221, THR112, and THR223 interacting with the ligand, likely causing its decreased binding affinity. These increased amounts of visualized hydrophilic interactions confirmed the diminished hydrophobicity of AR16 as denoted by the cLogP value of 1.58.

A possibility for the inconsistent values between the DINC and Swissdock thermodynamic outputs is the unique computational algorithms implemented by each software in the calculation of binding poses and energies. DINC is generally used for larger ligands, which could explain possible inaccuracies in binding affinities when compared to those outputted by Swissdock. With the algorithms executed by Swissdock generally used for smaller ligands like monastrol, these values are more likely to be accurate and closer to the actual experimentally derived binding affinities. However, the overall trends of binding affinities for the ester substitutions were relatively consistent, particularly with the indication that the geranyl analogs were lead compounds that should be synthesized and biologically screened while the methyl and ethyl analogs should not. Future studies using high-throughput virtual screening could employ the use of alternative molecular docking software packages to examine other binding affinity differences with varying algorithms and computations.

With molecular docking, most of the interactions are entropy-driven as two separate entities of the ligand and the receptor bind together to form a single protein-ligand complex. This previously rotatable and flexible entity is now constrained with the attractions in the binding pocket, resulting in an entropic penalty to occur with fewer microstates. Due to the docking algorithms used per software, we were unable to ascertain the relative entropic and enthalpic contributions; the final configurations are a combination of the two different factors, though the overall binding pose analysis should rely primarily on entropic interactions. Additionally, these simulations can only provide insight into the binding thermodynamics and not the binding kinetics, representing another limitation with this technique.

Further computational experiments that could be performed in the future to verify the importance of the side group modifications on the monastrol analogs include molecular dynamics simulations. These simulations could provide insight into the effects of binding on the time-resolved behavior of the binding site. Particularly, conducting molecular dynamics simulations for the top geranyl compounds discovered could provide further analysis into how binding affects the movement of the active site.

The increased binding affinities and high cLogP values of the geranyl analogs screened in this study suggests that hydrophobic interactions of the ester functionality are key in the binding of dihydropyrimidine monastrol analogs to the allosteric binding pocket of Eg5. These are potential leads in the development of potent Eg5 inhibitors, and the creation of new analogs with these nonpolar functionalities could potentially yield anticancer molecules with greater antiproliferative properties than monastrol itself. Future high-throughput virtual screening studies conducted with monastrol and dihydropyrimidine analogs should focus on the substitution of the ester functionality in favor of other side chains with increased hydrophobic properties. The QSAR uncovered by our high-throughput virtual screening efforts informs the future chemical synthesis of geranyl analogs with greater potency and biological activity. Using high-throughput

virtual screening to design, discover, and analyze effective small molecule Eg5 inhibitors has the potential to advance humanity's efforts in combating cancer and revolutionize the available treatments for this family of deadly diseases.

MATERIALS AND METHODS

Avogadro

Each analog to be screened was constructed virtually on Avogadro (version 1.2.0), an open access three-dimensional molecular editing and visualization software for use in the modeling of molecular compounds (12). Following the construction of each analog, an initial molecular mechanics geometry optimization was performed. This preprocessing step for the ligands allowed for a basic adjustment of human error when constructing the molecules, especially in terms of correcting the bond lengths and bond angles associated with each molecular geometry and bond type.

ORCA

To identify the ideal binding pose, we first determined the most quantum mechanically favorable geometry for each analog using DFT molecular geometry optimizations, which were conducted using ORCA (version 4.2.0), an ab initio quantum mechanical molecular modeling software (13). In order to run these geometric optimization calculations, input files for each structure suited for ORCA were generated by Avogadro. To simulate the aqueous environment in which monastrol binds to kinesin Eg5 in the human body, we used a conductor-like polarizable continuum (CPCM) solvation model. This was an implicit solvation model, which treated the solvent environment as a continuous medium with a particular dielectric and polarizability rather than individual, explicitly defined solvent molecules. A B3LYP hybrid functional and def2-SVP basis set were chosen. The DFT calculations were all conducted on a MacBook Pro with a 2.3 GHz Intel Core i5 processor, 4 cores, 8 GB of RAM, and 512 GB of flash storage, with each calculation running for an estimated 15 minutes. After computing the atomic positions in the most quantum mechanically stable states through submission of the input files into ORCA, the analogs were prepared for protein-ligand docking.

Molecular Docking

Molecular docking was used to predict the binding affinities of ligands bound to particular targets on the basis of specific thermodynamic factors that contribute to and detract from the free energy of interaction between the target and ligand. These factors include noncovalent interactions such as electrostatic attraction and repulsion, pi-stack interactions, hydrophobic interactions, and hydrogen bonding networks. The docking outputs were visualized on UCSF Chimera (version 1.13.1) (14). By observing the binding pocket of monastrol itself to the human motor protein kinesin Eg5 (PDB: 1X88), we were able to determine the appropriate grid box center and dimensions at which to bind the rest of our analogs by calculating the atomic distances between the kinesin protein and monastrol using UCSF Chimera. The center of the grid box was calculated as having coordinates of (21.738, 26.509, 51.081) and the grid box dimensions were determined to be (40, 44, 40), which provided a detailed specification of the proper allosteric binding site for ligand docking. This grid box covered the entirety of the ligand and only the observed allosteric binding pocket of the kinesin Eg5 protein. As we used the crystal structure of the kinesin Eg5 protein (PDB: 3HQD) as determined by X-ray diffraction for the docking of the analogs, no further preprocessing steps beyond the specification of the grid box were required for the protein prior to docking with both Swissdock and DINC (15). The quantum mechanically optimized ligand structures from the DFT calculations were then docked into the processed kinesin Eg5 protein. The docking software used generated a list of binding poses with different binding affinities; for the purpose of this study, we selected the highest calculated binding affinity to be the representative value for each compound, and alternative poses were not factored into the QSAR.

Swissdock

Swissdock, developed by the Swiss Institute of Bioinformatics, is a web-based server for molecular docking that utilizes an algorithm generating binding modes in the vicinity of a localized target cavity (16). Using their potential energies, binding modes with the most favorable energy values are analyzed and clustered, which can be viewed on the server or on other molecular visualization softwares. In this study, we docked each of the novel analogs into the allosteric binding pocket of human kinesin Eg5 (PDB: 3HQD) using the previously determined grid box coordinates to obtain the predicted binding affinities. Swissdock outputted varying conformations of the ligand on the receptor based on different binding modes with the most favorable energies clustered together. The predicted ΔG value of each dihydropyrimidine analog to Eg5 was reported in kcal/mol and represented the binding affinity of the ligand.

Docking Incrementally

Docking Incrementally (DINC), originally developed by the Kavraki Lab, was the second docking software used to computationally dock the analogs (17). DINC (version 2.0) uses a meta-docking algorithm called AutoDock Vina, another docking software which predicts thermodynamically favorable interactions between small molecule ligands and protein targets (17). DINC treated the ligand as a superposition of a rigid body component and a single rotatable component and was chosen to more accurately model the kinesin-ligand system, since many of our analogs had a large number of rotatable bonds. Each single partial solution was overlapped as a fragment until the entire ligand had been reconstructed. In this study, the analogs were docked into the allosteric binding pocket of kinesin Eg5 (PDB: 3HQD) using the aforementioned grid box coordinates to determine the predicted binding

affinities.

OSIRIS Property Explorer - cLogP Calculations

With monastrol known to interact with the allosteric binding pocket through hydrophobic interactions, we also sought to investigate the degree of hydrophobicity of our dihydropyrimidine analogs to provide a QSAR. Using the OSIRIS Property Explorer on the Organic Chemistry Portal, cLogP values, the logarithms of the partition coefficient between water and n-octanol for a given compound, were calculated to provide insight on the hydrophilicity of each analog (18). Negative cLogP values indicated greater hydrophilic properties of the molecule, while positive cLogP values indicated greater hydrophobic properties. Molecules with large, positive cLogP values were considered to have high hydrophobicities and this result indicated a greater affinity of the analog for stronger hydrophobic interactions in the binding pocket. Binding affinities and cLogP values were subsequently analyzed for emerging trends linking hydrophobicity and improved binding to kinesin Eg5. Additionally, we assumed differences in the cLogP values per analog to be solely due to the aliphatic side chain or aromatic substitutions.

Binding Pocket Analysis

Using UCSF Chimera to visualize the molecular docking results, the binding pocket in which the dihydropyrimidine analogs were docked was closely examined and analyzed. With typical hydrogen bonds having a bond distance between 2.4-3.5 angstroms, hydrophobic interactions having a bond distance between 3.3-4.0 angstroms, and pi-stack interactions having a bond distance between 3.3-3.8 angstroms, only amino acid residues specified within 4.0 angstroms of the docked ligand were considered to be taking part in potential non-covalent interactions (19, 20, 21). These amino acid residues were then analyzed for various properties that could have explained the calculated binding affinities on both Swissdock and DINC, with a particular focus on hydrophobic surfaces due to the significant amount of hydrophobic interactions monastrol engages in with kinesin Eg5.

REFERENCES

1. Siegel, Rebecca L., et al. “Cancer Statistics, 2019.” CA: A Cancer Journal for Clinicians, vol. 69, no. 1, 2019, pp. 7–34., doi:10.3322/caac.21551.
2. Simon, Stacy. “Facts & Figures 2019: US Cancer Death Rate Has Dropped 27% in 25 Years.” American Cancer Society, American Cancer Society, 2019, www.cancer.org/latest-news/facts-and-figures-2019.html.
3. Mayer, T. U. “Small Molecule Inhibitor of Mitotic Spindle Bipolarity Identified in a Phenotype-Based Screen.” Science, vol. 286, no. 5441, 1999, pp. 971–974., doi:10.1126/science.286.5441.971.
4. Liu, Min, et al. “Non-Canonical Functions of the Mitotic Kinesin Eg5.” Thoracic Cancer, vol. 9, no. 8, 2018, pp. 904–910., doi:10.1111/1759-7714.12792.
5. Wang, Yufei, et al. “Eg5 Inhibitor YL001 Induces Mitotic Arrest and Inhibits Tumor Proliferation.” Oncotarget, vol. 8, no. 26, 2017, pp. 42510–42524., doi:10.18632/oncotarget.17207.
6. Gonçalves, Itamar Luís, et al. “Effect of N-1 Arylation of Monastrol on Kinesin Eg5 Inhibition in Glioma Cell Lines.” MedChemComm, vol. 9, no. 6, 2018, pp. 995–1010., doi:10.1039/c8md00095f.
7. Kapoor, Tarun M., et al. “Probing Spindle Assembly Mechanisms with Monastrol, a Small Molecule Inhibitor of the Mitotic Kinesin, Eg5.” Journal of Cell Biology, vol. 150, no. 5, 2000, pp. 975–988., doi:10.1083/jcb.150.5.975.
8. Guido, Bruna C, et al. “Impact of Kinesin Eg5 Inhibition by 3,4-Dihydropyrimidin-2(1H)-One Derivatives on Various Breast Cancer Cell Features.” BMC Cancer, vol. 15, no. 1, 2015, doi:10.1186/s12885-015-1274-1.
9. Myers, Stephanie M, and Ian Collins. “Recent Findings and Future Directions for Interpolar Mitotic Kinesin Inhibitors in Cancer Therapy.” Future Medicinal Chemistry, vol. 8, no. 4, 2016, pp. 463–489., doi:10.4155/fmc.16.5.
10. Maliga, Zoltan, et al. “Evidence That Monastrol Is an Allosteric Inhibitor of the Mitotic Kinesin Eg5.” Chemistry & Biology, vol. 9, no. 9, 2002, pp. 989–996., doi:10.1016/s1074-5521(02)00212-0.
11. Pagadala, Nataraj S., et al. “Software for Molecular Docking: a Review.” Biophysical Reviews, vol. 9, no. 2, 2017, pp. 91–102., doi:10.1007/s12551-016-0247-1.
12. Hanwell, Marcus D, et al. “Avogadro: an Advanced Semantic Chemical Editor, Visualization, and Analysis Platform.” Journal of Cheminformatics, vol. 4, no. 1, 2012, doi:10.1186/1758-2946-4-17.
13. Neese, Frank. “The ORCA Program System.” WIREs Computational Molecular Science, vol. 2, no. 1, 2011, pp. 73–78., doi:10.1002/wcms.81.
14. Pettersen, Eric F., et al. “UCSF Chimera? A Visualization System for Exploratory Research and Analysis.” Journal of Computational Chemistry, vol. 25, no. 13, 2004, pp. 1605–1612., doi:10.1002/jcc.20084.
15. Parke, Courtney L., et al. “ATP Hydrolysis in Eg5 Kinesin Involves a Catalytic Two-Water Mechanism.” Journal of Biological Chemistry, vol. 285, no. 8, 2009, pp. 5859–5867., doi:10.1074/jbc.m109.071233.
16. Grosdidier, A., et al. “SwissDock, a Protein-Small Molecule Docking Web Service Based on EADock DSS.” Nucleic Acids Research, vol. 39, 2011, doi:10.1093/nar/gkr366.
17. Antunes, Dinler A., et al. “DINC 2.0: A New Protein–Peptide Docking Webserver Using an Incremental Approach.” Cancer Research, vol. 77, no. 21, 2017, doi:10.1158/0008-5472.can-17-0511.
18. Sander, Thomas. “CLogP Calculation.” Organic Chemistry, www.organic-chemistry.org/prog/peo/cLogP.

html.
19. Berg, Jeremy M., et al. Biochemistry. W.H. Freeman and Co., 2002.
20. Martz, Eric. "Noncovalent Bond Finder." University of Massachusetts, Amherst, 2002, www.umass.edu/microbio/chime/find-ncb/demo_txt.htm.
21. Janiak, Christoph. "A Critical Account on π–π Stacking in Metal Complexes with Aromatic Nitrogen-Containing Ligands." Journal of the Chemical Society, Dalton Transactions, no. 21, 2000, pp. 3885–3896., doi:10.1039/b003010o.

Article submitted: April 22, 2020
Article accepted: September 22, 2020
Article published: December 20, 2020