Virtual Screening of Hamigomycin B Natural Product Derivatives in 26 Kinase Proteins

Meagan Edmonds, Jesse D. Carrick, Derek J. Cashman

Tennessee Technological University Department of Chemistry, 55 University Drive, Cookeville, TN 38501-0001

Introduction

Hamigeromycin B and its analogs are synthetic natural product derivatives that may be useful at mediating signal transduction activity in human kinases. To study this potential activity, 11 Hamigomycin analogs were constructed using MOE 2019 and subjected to energy minimization using the AMBER14:EHT force field. Each analog was also compared against the parameters of Lipinski's Rules of Five to determine druggability.0 The compounds were docked into twenty-six human kinase structures obtained from the Protein Data Bank (www.rcsb.org) to obtain relative binding free energy scores using the Docking module of MOE 2019. The binding sites for these kinase proteins were analyzed using the Evolutionary Trace server (http://lichtargelab.org/software/ETserver; Baylor University; Houston, TX USA) and the Protein Frustratometer (http://frustratometer.qb.fcen.uba.ar/; EMBNet Argentina; Buenos Aires, Argentina) to characterize the energetics and evolutionary information of the amino acid residues for likely contributions to binding. Evolutionary Trace performs a multi-sequence alignment for each kinase in multiple species. It weighs the alignment against different species based on their relationship in the phylogenetic tree (e.g., more closely related species are weighted as more important than more divergent species). Frustration is based on energy landscape theory; it compares energy of the native state to a mutated state. Sites of high frustration can guide binding sites of proteins and small molecules, as well as sites of evolutionary importance.

Methods

The 3D X-ray crystal structures of 26 human kinases were obtained from the protein data bank (http://www.rcsb.org/) and proteins were modeled using MOE 2020 (Chemical Computing, Ltd., Montreal, Quebec, Canada). Each protein and small molecules were modeled and subjected to energy minimization using the AMBER14:EHT force field to convergence (0.01 kcal/mol). The Protein Frustratomer and Evolutionary Trace were used to identify residues surrounding the active site for docking. Frustrated residues were determined as those that displayed a configurational frustration density of 20% or greater, and mapped onto the 3D ribbon structure of each protein in green. The top 25% of evolutionarily important residues were identified as those with the lowest r_i importance scores. These residues were mapped onto the 3D ribbon structure in red. If amino acid residues showed both energetic frustration and evolutionary importance, they were mapped onto the 3D ribbon structure in blue.
These maps defined the active site for docking. The Docking module of MOE 2020 was used for docking studies. Ten compounds were originally built in MOE for the docking using the forcefield Amber:14 HBT. These structures were then docked into the active site of all three proteins using the Docking module of MOE 2020.

Using the refined protein structures originally from the PDB, each protein was opened, and any native ligands were removed from the structure. This also aided in identifying the active site residues, combined with evolutionary trace and frustration data. The force field used for docking was Amber14:EHT implemented in MOE 2020, and the active site residues were selected as the docking area. Initial docking obtained 30 docked conformations by placing triangles in the spaces defined by the van der Waal's surface area of the active site cavity London dispersion forces used to determine initial binding energy. These 30 conformations were further refined using induced fit and scored with the GBVI/WSA scoring method to calculate the binding free energy (ΔG_{bind}) in kcal/mol, and five docked conformations were obtained.

Compounds

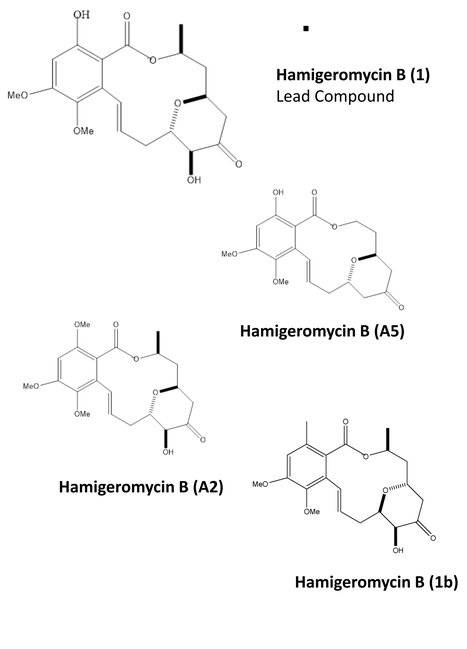

Figure 1. The figure above shows the 2D structure of the lead, or prototypical, compound (Hamigeromycin B (1) docked. Subsequent compounds were modified at various positions on these rings. The chemical structures of three of the top binding compounds in EGFR are shown illustrating their modifications.

Binding Free Energy

Compound Name	JNK	AKT1	Aurora-A	Hsp90	Aurora-B	VEGF-R2	CDK2	FAK	CDK4	B-RAF VE	CK2a	EphB4	ErbB2	VEGFR3	FLT3	INSR	MET	PDGF-RB	PLK1	SAK	TIE	COT	EGFR	SRC
B1	-6.9925	-5.8224	-7.2479	-7.7678	-8.0329	-7.2129	-7.4102	-7.1565	-5.8045	-7.3026	-7.0541	-7.3359	-6.6728	-5.3727	-6.5045	-6.9088	-7.8676	-5.2119	-7.9917	-5.7351	-6.6981	-6.543	-7.916	-6.843
A1	-9.0911	-5.0363	-6.2614	-6.5444	-7.2169	-6.8597	-6.5537	-6.3263	-5.7545	-6.0148	-6.538	-6.6742	-5.709	-5.3157	-6.7201	-6.3649	-6.4754	-4.2981	-6.6797	-4.72	-5.6377	-6.8071	-6.4416	-6.1478
A2	-7.5339	-5.9617	-8.1148	-8.3042	-8.3585	-8.5305	-7.8517	-7.8035	-6.1349	-7.8098	-6.6747	-8.1125	-6.9889	-5.3511	-7.7131	-7.1007	-7.845	-4.8816	-7.8515	-5.2724	-6.7101	-6.7443	-7.9268	-7.2751
A3	-7.6535	-6.3402	-7.7453	-7.9357	-7.836	-8.1353	-8.0356	-6.997	-5.6108	-7.1495	-7.4815	-7.6497	-6.6497	-5.78	-7.2738	-7.2868	-7.2874	-4.9948	-7.4482	-5.5068	-5.8434	-6.2676	-7.6214	-7.4408
A4	-7.0413	-6.1746	-7.7719	-8.0855	-8.4694	-7.7232	-8.0299	-7.2631	-6.1355	-7.9241	-7.5411	-7.2963	-6.8233	-5.7655	-7.097	-7.1246	-7.5866	-4.9173	-7.5321	-5.5119	-6.0626	-6.7942	-6.1859	-6.9387
A5	-6.7413	-6.6516	-7.5504	-7.6725	-7.8606	-7.3811	-7.6549	-7.1486	-6.1544	-7.8238	-7.0395	-7.2891	-7.06	-5.7679	-7.6999	-7.3127	-7.8555	-5.0883	-7.3516	-5.3296	-6.5106	-8.6352	-8.5605	-6.9674
A6	-7.3557	-6.239	-7.3981	-8.2281	-8.344	-7.8193	-7.8504	-6.4868	-6.1733	-7.1295	-7.4676	-7.1976	-6.801	-5.6342	-7.7557	-7.4986	-7.4473	-5.1161	-7.7984	-5.8371	-6.1173	-6.6462	-7.8812	-7.4965
A7	-6.591	-6.1105	-7.5517	-7.7094	-8.2928	-7.9605	-7.6169	-6.7095	-5.9411	-7.6361	-7.1163	-7.2596	-6.5654	-5.326	-7.4575	-7.3811	-7.5091	-5.3355	-7.5395	-5.5823	-5.8551	-8.5393	-7.1654	-7.094
A8	-6.9427	-6.273	-7.0964	-7.7658	-7.9415	-7.7252	-7.6602	-6.8561	-5.8656	-6.6247	-7.0105	-7.1456	-6.4184	-5.3024	-7.3565	-6.8647	-7.429	-4.689	-6.9755	-5.3761	-5.8168	-6.091	-7.1054	-7.0898
A9	-6.621	-6.1469	-7.256	-7.6003	-7.8216	-7.1523	-7.3756	-6.8453	-5.9272	-6.5886	-7.2008	-7.4937	-6.8272	-5.5899	-7.6721	-6.0724	-7.6042	-4.6425	-7.0224	-5.4606	-6.1391	-6.0912	-6.9313	-6.8231

Table 1. This table shows the binding free energy values for the original ten hamigeromycin B analogs versus twenty-six kinases docked. The lowest (most favorable) binding scores are highlighted in green; the highest (least favorable) binding scores are highlighted in red/orange.

Five kinases were found to have good binding affinity and docked in the second with the modified compounds – JNK, Aurora A, Aurora B, Hsp90, and EGFR. These five kinases were used for further binding studies based on modifications of the original lead compounds (see table 2).

Table 2. This table shows the binding free energy values of five kinases identified from the first round of docking and docked with twenty-three modified analogs of hamigeromycin B. The lowest (most favorable) binding scores are highlighted in green; the highest (least favorable) binding scores are highlighted in red/orange.

		ΔGbind (kcal/mol)			
Compound Name	JNK	Aurora-A	HSP90	Aurora-B	EGFR
B1a	-5.7347	-6.5247	-6.6464	-6.7113	-6.6701
B1b	-7.4548	-8.3305	-8.8746	-8.305	-9.6122
B1c	-7.4506	-7.8399	-8.1839	-7.8431	-9.2961
B1d	-7.0002	-7.9601	-8.0897	-7.8496	-8.9867
A1a	-7.0258	-7.977	-8.0808	-8.2553	-8.9502
A2a	-7.0331	-7.9133	-7.9589	-8.0399	-8.7971
A2b	-7.0365	-7.9064	-8.1655	-8.1835	-8.7462
A3a	-6.9795	-7.9175	-8.265	-8.1581	-8.7922
A3b	-7.0806	-7.6783	-7.9063	-8.0083	-8.7391
A4a	-7.3621	-8.0809	-8.6304	-8.3203	-9.0371
A4b	-6.5514	-7.9266	-8.3455	-8.2368	-8.6474
A4c	-6.6843	-7.9019	-8.4238	-8.2533	-8.5426
A5b	-6.9362	-7.9092	-8.4031	-8.2322	-8.577
A6b	-6.8592	-7.598	-7.9906	-8.0017	-8.1253
A7a	-6.4547	-7.5607	-8.023	-7.9667	-8.113
A7b	-6.4309	-7.591	-7.8354	-7.9964	-8.0981
A7c	-6.7347	-7.3727	-7.6164	-7.745	-8.0503
A8a	-6.6873	-7.352	-7.6809	-7.7322	-8.0897
A8b	-6.8361	-7.1207	-7.6515	-7.8417	-8.4814
A8c	-7.2647	-7.3115	-7.9418	-8.2441	-8.4553
A9a	-7.0956	-7.2901	-7.8337	-7.6027	-8.3389
A9b	-6.9133	-7.0536	-7.6406	-7.6326	-7.8445

Molecular Docking

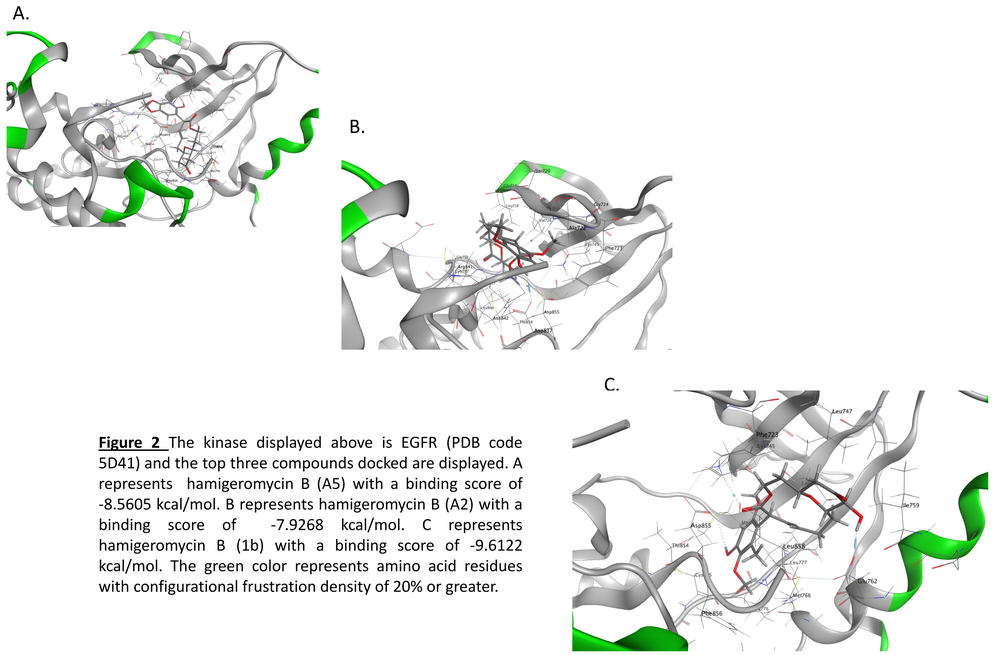

Figure 2 The kinase displayed above is EGFR (PDB code 5D41) and the top three compounds docked are displayed. A represents hamigeromycin B (A5) with a binding score of -8.5605 kcal/mol. B represents hamigeromycin B (A2) with a binding score of -7.9268 kcal/mol. C represents hamigeromycin B (1b) with a binding score of -9.6122 kcal/mol. The green color represents amino acid residues with configurational frustration density of 20% or greater.

Druggability

Lipinski's Rules	MW<500 Da	logp	H bond donors	H bond acceptors
B (1)	392.4	1.44	0	8
B (A1)	286.33	2.22	0	5
B (A2)	420.46	1.45	1	7
B (A3)	390.43	1.77	0	6
B (A4)	390.4	1.01	0	7
B (A5)	376.4	1.55	1	6
B (A6)	390.43	1.95	0	6
B (A7)	376.4	1.5	1	6
B (A8)	362.38	1.09	1	7
B (A9)	346.38	1.63	0	6
B1a	392.4	1.44	0	8
B (1)b	392.4	1.44	0	8
B (1)c	392.4	1.44	0	8
B (1)d	392.4	1.44	0	8
A1a	286.33	2.22	0	5
A2a	420.46	1.45	1	7
A2b	420.46	1.45	1	7
A3a	390.43	1.77	0	6
A3b	390.43	1.77	0	6
A4a	390.4	1.01	0	7
A4b	390.4	1.01	0	7
A4c	390.4	1.01	0	7
A5b	376.4	1.55	1	6
A6b	390.43	1.95	0	6
A7a	376.4	1.5	1	6
A7b	376.4	1.5	1	6
A7c	376.4	1.5	1	6
A8a	362.38	1.09	1	7
A8b	362.38	1.09	1	7
A8c	362.38	1.09	1	7
A9a	346.38	1.63	0	6
A9b	346.38	1.63	0	6

Table 3. This table shows the druggability of each designed analog using Lipinski's Rules of Five for all thirty-three compounds. This includes a molecular weight less than 500 Da, a Log P of less than 5, five or less H-bond donors, and ten or less H-bond acceptors.

Conclusion and Future Work

- The lowest docking scores were used to determine the best binding and orientation of each analog to each protein. Docking data suggests good binding of these analogs with five out of twenty-six human kinase proteins. Modifications improved targeting for EGFR over other proteins, which was one of the five kinases which were determined to have good binding affinity.
- .Of the thirty-three total compounds, hamigeromycin A5, hamigeromycin A2, and hamigeromycin B1b were the top binding compounds to the protein EGFR.
- Evolutionary Trace and Frustration seem to be good guide for the active sites of the kinases, as most kinases in the study has a good amount of frustrated and evolutionarily important residues in the active site.
- Future *in vitro* studies are being planned to confirm the computational docking results. Preliminary computational results will be disseminated.

References

1. Lipinski, C.A., Lombardo, F., Dominy, B.W., Feeney, P.J. (March 2001). "Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings". *Adv. Drug Deliv. Rev.* **46** (1–3): 3–26. doi:10.1016/S0169-409X(00)00129-

Acknowledgements

This research is partially supported by the Tennessee Tech University Department of Chemistry. MOE 2020 software was provided by a grant from the 2017 Governor's School for Emerging Technologies, Chemical Computing, Ltd. (Montreal, Quebec, Canada). Dr. Scott Northrup (Professor Emeritus at TTU). The authors gratefully acknowledge the late Dr. Scott Northrup (TN Tech Professor Emeritus) and the Tennessee Tech Capital Fund for computer equipment used in this project.