Synthesis and in silico screening of a library of β-carboline-containing compounds

Kay M. Brummond[*1], John R. Goodell[1], Matthew G. LaPorte[*1], Lirong Wang[2] and Xiang-Qun Xie[*2]

Full Research Paper

Address:
[1]Department of Chemistry, Chemical Methodologies and Library Development, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA and [2]Department of Pharmaceutical Sciences and Computational Chemical Genomics Center, School of Pharmacy; Drug Discovery Institute; Department of Computational Biology; University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA

Email:
Kay M. Brummond[*] - kbrummon@pitt.edu; Matthew G. LaPorte[*] - mgl12@pitt.edu; Xiang-Qun Xie[*] - xix15@pitt.edu

* Corresponding author

Abstract

The synthesis of a library of tetrahydro-β-carboline-containing compounds in milligram quantities is described. Among the unique heterocyclic frameworks are twelve tetrahydroindolizinoindoles, six tetrahydrocyclobutanindoloquinolizinones and three tetrahydrocyclopentenoneindolizinoindolones. These compounds were selected from a virtual combinatorial library of 11,478 compounds. Physical chemical properties were calculated and most of them are in accordance with Lipinski's rules. Virtual docking and ligand-based target evaluations were performed for the β-carboline library compounds and selected synthetic intermediates to assess the therapeutic potential of these small organic molecules. These compounds have been deposited into the NIH Molecular Repository (MLSMR) and may target proteins such as histone deacetylase 4, endothelial nitric oxide synthase, 5-hydroxytryptamine receptor 6 and mitogen-activated protein kinase 1. These in silico screening results aim to add value to the β-carboline library of compounds for those interested in probes of these targets.

Introduction

Identification of a comprehensive set of small organic molecules capable of selectively modifying the function of biological targets tremendously impacts modern medical research and drug discovery efforts [1]. Currently, this set of small molecules is largely occupied by in-house libraries and commercially available compounds. The NIH Roadmap initiative was established to address a recognized limitation of current compound diversity resulting in the Molecular Libraries Probe

Centers Network (MLPCN) which has, since its inception, garnered a library of over 370,000 chemically diverse small molecules in a central molecule repository [2]. This supply of compounds has been made possible by researchers across the disciplines, but largely by synthetic chemists who are preparing compounds with an eye towards biologically relevant targets. Another goal of the NIH Roadmap is the development of enabling methods for the synthesis of these structurally diverse compound libraries; amongst these methods, skeletal diversification strategies have emerged as particularly efficient for maximizing structural diversity [3].

Previous work in the Brummond laboratory has demonstrated that an allene-containing β-carboline provided a good starting point for synthesizing six novel types of hetero-frameworks, all skeletally unique [4]. Moreover, scope and limitation studies contributed to an understanding of chemistries that would possess the robustness necessary for library preparation. Information gained from these experiments was then utilized in the construction of a virtual library of 11,748 compounds. A diversity analysis was performed using B (Burden) C (CAS) UT (Pearlman at the University of Texas) metrics and Tanimoto coefficients (Tc) and this virtual compound library was mapped onto the existing chemical space of the NIH Molecular Libraries Small Molecule Repository (MLSMR) [4]. When considering the physical properties most important to bimolecular interactions, atomic Gasteiger–Hückel charges, polarizabilities, and hydrogen-bond acceptors, these virtual compounds were found to occupy new chemical space when compared to the 327,000 compounds in the MLSMR. A small subset of these compounds was subsequently identified as ones representing a maximally diverse chemical space. The synthesis of a modified subset of this virtual compound library is described within, where modifications were mainly driven by studies of compound stability. Furthermore, a high throughput, in silico screening analysis of this library identified a number of potential biological targets for the compounds.

Results and Discussion

Scaffolds **1**, **2** and **3** (Figure 1) were chosen for library preparation based upon favorable Tanimoto coefficient (Tc) scores when compared to the MLSMR, conformational constraints imposed by the β-carboline moiety, and the number of building blocks available for the diversifying elements R^1 and R^2.

The syntheses of tetrahydro-β-carbolines **6{1–16}** were accomplished in a manner entirely analogous to that reported previously (Table 1, entries 1–16) [4]. For example, the allenic methyl ester of tryptophan **4** was reacted with a number of aldehydes **5{1–15}** under acidic conditions to produce the corresponding products in yields ranging from 54–89%. A range of aldehydes were accommodated in the Pictet–Spengler reaction, including formaldehyde (Table 1, entry 1), alkyl aldehydes (Table 1, entries 2 and 3), aryl aldehydes with electron-withdrawing and electron-donating groups (Table 1, entries 4–7, 14 and 15), heteroaromatic aldehydes (Table 1, entries 8–13) and glyoxalates (Table 1, entry 16). Moreover, useful quantities of β-carboline-containing products were obtained (43–100 mg). For entries 2–15, mixtures of two diastereomers were obtained. Since the mixtures could not be readily separated by column chromatography, diastereomeric ratios were determined by ^{1}H NMR and were advanced without further purification. Reaction of allene **6** under the silver-nitrate-mediated cyclization conditions afforded the desired fused pyrrolines **1**. However, in the initial phases of this cyclization process, a color change was noted during the purification process. Indeed, when NMR stability studies were performed on the syn- and anti-pyrrolines **1{5}**, decomposition of both diastereomers was evident. Although it was generally difficult to isolate the individual diastereomers, they could be separated by column chromatography. It was found that anti-**1{5}** decomposed more rapidly than syn-**1{5}** during the ^{1}H NMR stability studies, when compared to an internal standard. These results combined with previously reported skeletal reorganization processes of functionalized β-carbolines, led to concerns about the long-term storage of these compounds and their inclusion in the MLSMR [5].

To increase the stability of this class of compounds, a toluenesulfonyl group was added to the indole nitrogen of **1{1–7}** to give *N*-tosyl-tetrahydro-β-carbolinepyrroline derivatives **7{1–7}**. These tosylated derivatives exhibited improved stability as evidenced by ^{1}H NMR (Table 1, entries 1–7, and Supporting Information File 1, S76–S81). Incorporation of the tosyl group
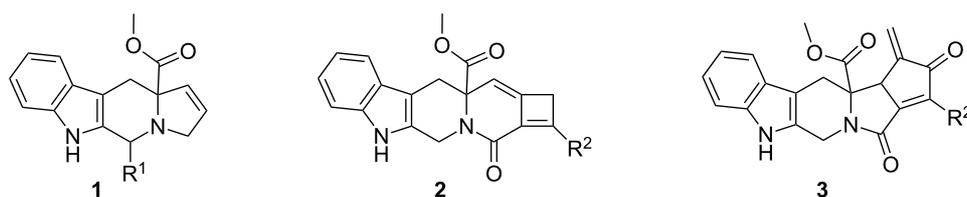

Figure 1: Tetrahydro-β-carboline containing scaffolds **1–3**.

Table 1: Library of *N*-tosyltetrahydro-β-carbolinepyrrolines **7{1–7}**.
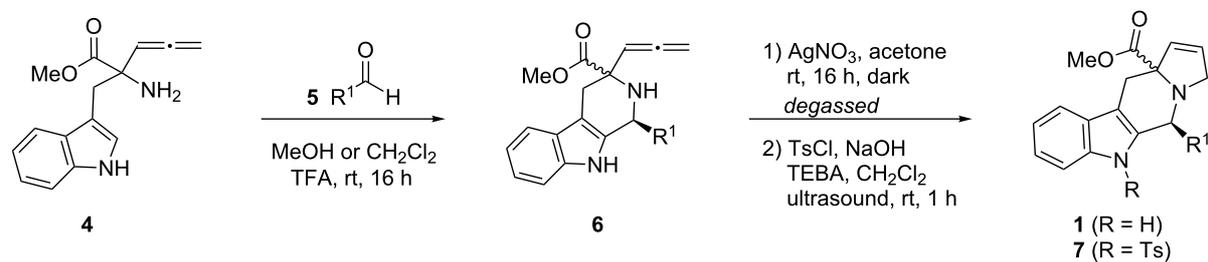
entry	R^1		yield of **6**[a] (%)	anti:syn[b]	yield of **7**[a] (%)	purity[c]
1	—H	**5{1}**	68 **6{1}**	NA	28 **7{1}**	98%
2	(cyclopropyl)	**5{2}**	74 **6{2}**	2.5:1	17 syn-**7{2}**	98%
3	(isobutyl)	**5{3}**	81 **6{3}**	2.1:1	23 syn-**7{3}**	98%
4	(phenyl)	**5{4}**	73 **6{4}**	2.4:1	11 anti-**7{4}** 28 syn-**7{4}**	98% 98%
5	(4-Cl-phenyl)	**5{5}**	75 **6{5}**	2.4:1	34 anti-**7{5}** 17 syn-**7{5}**	98% 98%
6	(4-CF$_3$-phenyl)	**5{6}**	74 **6{6}**	1.3:1	35 anti-**7{6}** 10 syn-**7{6}**	98% 98%
7	(3,4-diOMe-phenyl)	**5{7}**	54 **6{7}**	2.7:1	12 anti-**7{7}** 17 syn-**7{7}**	98% 98%
8	(pyridin-3-yl)	**5{8}**	89 **6{8}**	2.3:1	ND[d]	
9	(benzimidazol-2-yl)	**5{9}**	78 **6{9}**	1.2:1	ND[e]	
10	(1-Ph-3-Me-pyrazol-4-yl)	**5{10}**	80 **6{10}**	1:1.1	ND[d]	
11	(quinolin-4-yl)	**5{11}**	70 **6{11}**	1.3:1	ND[e]	
12	(quinolin-3-yl)	**5{12}**	80 **6{12}**	3.5:1	ND[e]	
13	(quinolin-2-yl)	**5{13}**	63 **6{13}**	1:1.4	ND[e]	
14	(naphthalen-1-yl)	**5{14}**	58 **6{14}**	2.3:1	ND[d]	

Table 1: Library of *N*-tosyltetrahydro-β-carbolinepyrrolines **7{1–7}**. (continued)
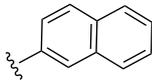
15		**5{15}**	80 **6{15}**	2.3:1	ND[d]	
16	CO$_2$Et	**5{16}**	88 **6{16}**	1.4:1	8 syn-**1{16}**	99%

[a]Isolated yields; [b]dr determined by ^{1}H NMR; [c]purity established by LCMS/ELS; [d]compound **6** detected; [e]compound **6** not detected.

also eased the chromatographic separation of the syn- and anti-isomers for entries 2–7, thus compounds **7{2–7}** were obtained as single diastereomers. Low to moderate yields for this two-step reaction sequence were attributed to a problematic tosylation due to the sterically hindered nature of the indole nitrogen atom. Moreover, unforeseen limitations were encountered for the heteroaromatic and naphthyl-containing β-carboline intermediates (Table 1, entries 8–15). While in some cases the intermediate pyrrolines **1** were observed (Table 1, entries 8, 10, 14, 15), the corresponding tosylated products were not obtained. The heteroaromatic examples (Table 1, entries 9, 11–13), did not undergo cyclization upon treatment with silver nitrate. For these cases, it was assumed that competing coordination of the heteroatom to the silver ion was an issue; however, attempts were not made to alter the reaction conditions for these substrates. Furthermore, conversion of the naphthalene-containing analogues **1{14}** and **1{15}** to their corresponding tosylates was not successful.

Next, compounds possessing the cyclobutene-fused β-carboline skeleton were assembled from the versatile allenyl intermediate **6**. For this subset of compounds, acylation of the sterically hindered amine of **6{1}** with the ynoic acids **8{1–8}**, by using bromo-tris-pyrrolidino-phosponium hexafluorophosphate (PyBroP), provided the requisite allene-yne substrates **9{1,2}–9{1,8}**. For the coupling reaction of the ynoic acids with an aryl group on the terminus of the alkyne **8{5–8}**, the desired allene-ynes **9{1,5}–9{1,8}** were afforded along with the [2 + 2] cycloadducts **2{1,5}–2{1,8}** (Table 2, entries 5–8). Previous studies with related tryptophan-substituted allene-ynes, required much higher temperatures (225 °C versus rt) to give the [2 + 2] cycloadducts, albeit none of these examples possessed two radical stabilizing groups on the alkyne. Because the calculations performed by Tantillo [6] regarding the thermal [2 + 2] cycloaddition reaction suggest that the energy barrier for this reaction should not be effected by the presence of two radical stabilizing groups over one, it was postulated that this cycloaddition process was facilitated by exposure to incident light [7]. Thus, the mixtures of compounds **2** and **9** were reconstituted in CH$_2$Cl$_2$ and placed in front of two 6 W UV lamps for 16 h at rt to afford the desired cyclobutenes (Method B, Table 2, entries

5–8). During optimization of the conditions for the [2 + 2] cycloaddition reaction (Method A), it was found that reducing the reaction temperature from 225 °C to 160 °C afforded cyclobutene **2{1,3}** in 73% yield (Table 2, entry 3). Similarly, allene-yne **9{1,5}** was subjected to the lower reaction temperature (Method A) to produce cyclobutene **2{1,5}** in 57% yield (Table 2, entry 5).

For the final library scaffold, a small subset of α-methylenecyclopentenone-containing tetrahydro-β-carbolines was synthesized. These compounds contain a general substructure that has recently been shown to inhibit DNA damage checkpoints [8]. Allene-ynes **9{1,3}–9{1,4}** undergo Pauson–Khand cyclocarbonylation reactions when treated with molybdenum hexacarbonyl in DMSO/toluene solutions (Table 3). Allene-yne **9{1,2}** afforded a mixture of four compounds comprising two diastereomers of **3{1,2}**, the 4-alkylidene cyclopentenone, resulting from the cyclocarbonylation reaction with the distal double of the allene, and a fourth compound, which could not be identified (see spectral data in Supporting Information File 1). Aryl-substituted alkynones **9{1,5}–9{1,8}** were not available for the molybdenum-mediated cyclocarbonylation process due to competing [2 + 2] cycloaddition reactions (Table 2).

The majority of these β-carboline-containing products exhibit acceptable calculated physical–chemical properties in accordance with Lipinski's rule of five (Figure 2) [9,10]. These favorable properties and structural novelty make these valuable candidates for deposition in the MLSCN for biological activity evaluation.

Diversity-oriented synthesis (DOS) has been employed to generate thousands of the organic compounds that have been deposited in the NIH molecular repository for medicinal chemistry research. Deciphering the therapeutic potential of this many compounds is a continuing challenge. By combining chemogenomics databases, such as Protein Data Base (PDB) and ChEMBL, it is possible to map new compounds into existing chemical space and to predict protein targets for new compounds, for which there are two complementary strategies that can be implemented. One is a structure-based docking

Table 2: Library of tetrahydro-β-carbolinecyclobutenes **2{1,3}–2{1,8}**.
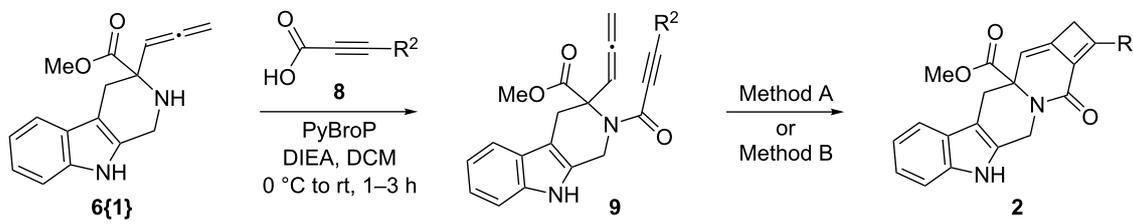
entry	R[2]		yield of **9**[a] (%)	method[b]	yield of **2**[a] (%)	purity[c]
1	⌇—H	**8{1}**	0 **9{1,1}**	—	—	—
2	⌇—TMS	**8{2}**	64 **9{1,2}**	A	ND[d]	—
3	⌇—Me	**8{3}**	75[e] **9{1,3}**	A	73 **2{1,3}**	99%
4	⌇—△	**8{4}**	71 **9{1,4}**	A	57 **2{1,4}**	99%
5	⌇—C6H5	**8{5}**	**9{1,5}** **2{1,5}**	A B	57[f] **2{1,5}** 40 **2{1,5}**	99%
6	⌇—C6H4—CF3	**8{6}**	**9{1,6}** **2{1,6}**	B	41 **2{1,6}**	99%
7	⌇—C6H4—OMe	**8{7}**	**9{1,7}** **2{1,7}**	B	30 **2{1,7}**	99%
8	⌇—thienyl-Me	**8{8}**	**9{1,8}**	B	33 **2{1,8}**	99%

[a]Isolated yield; [b]method A: μW, 160 °C, DMF, 10 min; method B: Placed in front of two 6 W UV lamps (245 nm), CH$_2$Cl$_2$, rt, 16 h, no stirring; [c]purity established by LCMS/ELS; [d]ND = not detected; [e]μW, 225 °C, DMF, 7 min, (39%); [f]separated **9{1,5}** (57% yield) from **2{1,5}** (18% yield) after the coupling reaction then submitted **9{1,5}** to method A to give **2{1,5}** in 68% yield. This compound was recombined with the previously isolated **2{1,5}** to afford a combined 57% yield.

strategy, in which a query compound is fit into a series of protein binding pockets to identify favorable compound–protein interactions. A second approach is a ligand-based strategy in which the structural similarities between a query compound and

a collection of bioactive compounds are identified. In the present study, both of these strategies were used to predict potential targets of the newly synthesized library of β-carboline-containing compounds.

Table 3: Synthesis of α-methylenecyclopentenone library **3{1,3}–3{1,4}**.
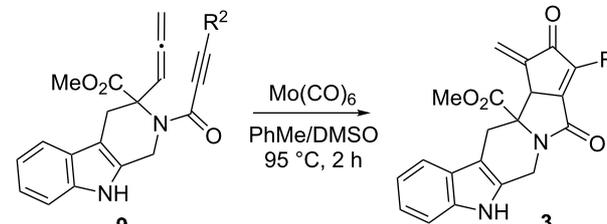
entry	R[2]		yield of **3**[a] (%)	purity[b]
1	⌇—Me	**9{1,3}**	48 **3{1,3}**	99%
2	⌇—△	**9{1,4}**	46 **3{1,4}**	99%

[a]Isolated yields; [b]purity determined by LCMS/ELS.

1{5}
ᵃ378.8, ᵇ4, ᶜ5.10, ᵈ48

7{1}
ᵃ422.5, ᵇ7, ᶜ3.02, ᵈ75

7{2}
ᵃ462.6, ᵇ7, ᶜ3.97, ᵈ64

7{3}
ᵃ478.6, ᵇ7, ᶜ4.32, ᵈ65

anti-**7{4}**
ᵃ498.6, ᵇ7, ᶜ4.13, ᵈ66
syn-**7{4}**
ᶜ4.49, ᵈ65

anti-**7{5}**
ᵃ533.0, ᵇ7, ᶜ4.51, ᵈ66
syn-**7{5}**
ᶜ4.98, ᵈ65

anti-**7{6}**
ᵃ566.6, ᵇ7, ᶜ4.93, ᵈ66
syn-**7{6}**
ᶜ5.71, ᵈ65

anti-**7{7}**
ᵃ558.6, ᵇ9, ᶜ4.13, ᵈ81
syn-**7{6}**
ᶜ4.62, ᵈ81

syn-**1{16}**
ᵃ340.4, ᵇ4, ᶜ3.79, ᵈ80

2{1,4}
ᵃ360.4, ᵇ5, ᶜ4.06, ᵈ72

2{1,5}
ᵃ396.4, ᵇ5, ᶜ4.51, ᵈ75

2{1,6}
ᵃ464.4, ᵇ5, ᶜ5.45, ᵈ75

2{1,7}
ᵃ426.5, ᵇ6, ᶜ4.59, ᵈ83

2{1,8}
ᵃ416.5, ᵇ5, ᶜ5.02, ᵈ72

3{1,3}
ᵃ362.4, ᵇ8, ᶜ2.50, ᵈ104

3{1,4}
ᵃ388.4, ᵇ8, ᶜ3.24, ᵈ99

Figure 2: Library of tetrahydro-β-carboline containing compounds **1–7** and calculated properties (ᵃmolecular weight; ᵇhydrogen-bond donor/acceptors; ᶜcLogP; ᵈpolar surface area (PSA)).

High-throughput docking studies for protein-target prediction of newly synthesized compounds

Molecular docking studies were performed with the 34 newly synthesized compounds, represented by scaffolds **1**, **2**, **3**, and allenyl precursors **4**, **6{1–16}**, **9{1–4}** to identify potential protein targets [11]. Protein structures were downloaded from the PDB [12] and the analysis was limited to a selection of the 607 proteins defined as "druggable" targets, in order to reduce computational time [13]. (The complete listing of these proteins and their PDB IDs are provided in Supporting Information File 2). The Surflex-dock module of the Sybyl software was employed for protein preparation and docking of the β-carboline library [14,15]. Water molecules and ligands were removed from the protein structures and the active site of each protein was defined by the corresponding residues around the cocrystal-

lized ligands. In-house algorithms were used to evaluate ligand-docking efficiency, and docking scores were used to assess and rank the protein targets.

A portion of the protein-scoring matrix is illustrated in Figure 3. Several interesting results emerged from this in silico analysis: (1) Twenty of the new compounds have docking scores greater

PDB ID	A	4	7{1}	7{7}	2{1,3}	2{1,4}	2{1,5}	2{1,6}	2{1,7}	2{1,8}
1nuh	7.23	6.45	-4.46	-25.14	3.98	3.49	5.4	3.91	6.17	1.79
1i7b	3.07	7.32	1.74	0.9	4.6	5.28	4.24	3.49	3.84	2.66
2xfo	-0.07	7.29	-4.96	-12.59	-4.95	2.88	1.81	-0.33	0.45	0.99
1d4l	5.78	7.17	3.06	3.08	3.68	3.55	3.31	4.98	3.19	2.29
2vgb	3.13	7.14	-7.1	-17.8	2.7	1.92	0.96	-0.13	-2.08	0.68
2qro	4.7	7.06	-21.42	-51.03	-3.69	-2.23	-2.36	-9.21	-13.18	-7.52
3gqy	4.98	4.96	7.75	0.67	3.66	4.38	3.3	0.84	-1.85	1.54
2vqq	5.24	5.07	1.31	7.14	6.74	7.59	7.69	7.85	7.99	7.03
1s9i	3.99	5.26	3.84	0.35	7.44	5.09	3.03	4.32	3.52	2.15
2q6b	4.76	5.49	5.15	-5.63	7	6.03	7.92	4.47	3.92	2.84
2v3e	4.93	4	2.54	-4.43	5.43	5.3	7.23	7.54	7.46	5.96
2vgq	5.04	6.27	3.42	1.9	3.76	3.87	4.02	7.95	4.35	3.04
3hhm	4.5	4.01	2.4	1.95	3.97	5.42	6.65	7.63	2.71	3.73
3hdn	4.39	4.22	3.63	1.21	2.68	2.86	2.29	7.56	6.14	1.92
3g7w	3.89	5.77	3.31	1.84	4.92	3.97	5.2	7.27	4.8	4.43
3eqc	6.69	5.33	1.59	-0.93	4.39	6.36	5.13	7.08	7.15	6.95
3fwl	5.9	4.39	1.99	1.97	4.03	4.07	6.34	5.18	7.04	3.96
1so2	5.05	6.12	2.3	1.88	6.34	4.21	5	6.2	5.29	7.22

PDB ID	2{1,8}	3{1,4}	6{3}	6{4}	6{6}	6{7}	6{9}	6{10}	6{12}	6{13}	6{15}
1d4l	2.29	3.02	3.26	3.55	4.55	3.51	5.15	4.83	7.3	5.22	4.1
1so2	7.22	3.84	3.84	6.32	7.25	6.08	8.17	5.78	6.15	5.77	6.43
3ljr	5.45	7.11	3.35	4.66	0.33	3.81	3.33	1.63	4.32	6.29	4.81
2zv2	2.53	3.45	7.37	3.36	2.02	2.83	1.37	4.2	2.05	4.38	1.75
2on3	2.58	2.8	2.13	7	3.02	6.2	4.17	-1.13	3.28	3.84	1.65
3fxw	2.59	3.47	3.09	5.02	5.39	7.65	5.13	4.53	5.69	5.94	3.08
3mhj	-5.26	3.16	3.5	2.03	6.9	7.17	4.23	3.44	6.6	5.53	6.17
1pq6	-1.13	-1.86	6.35	5.8	4.46	6.91	7.57	8.28	6.32	7.02	6.15
1t40	2.37	2.74	2.84	3.52	4.09	5.66	4.64	8.75	6.17	1.69	5.43
3ipq	-9.79	-3.95	5.21	5.47	4.69	6.55	6.9	8.65	5.44	4.37	6
2gvj	4.43	3.1	3.32	4.13	4.67	3.63	3.15	8.47	-7.88	4.49	4.3
1dgh	-1.94	4.06	5.08	4.02	4.75	4.15	6.06	7.74	3	7.17	3.27
3nos	4.35	2.27	3.31	3.84	6.79	6.67	3.47	7.35	7.27	5.18	7.6
1og5	5.8	4.21	3.96	5.08	5.07	5.91	5.92	7.24	5.31	4.78	5.7
1nf7	2.59	3.62	4.63	4.75	-1.19	4.97	0.46	7.21	-5.03	4.8	-1.1
1lpg	3.07	2.38	2	4.86	4.09	2.75	4.45	4.57	5.52	7.01	2.96

HR2 from PDB:2q6b   2{1,5}

TFG from PDB:2vqq   2{1,7}
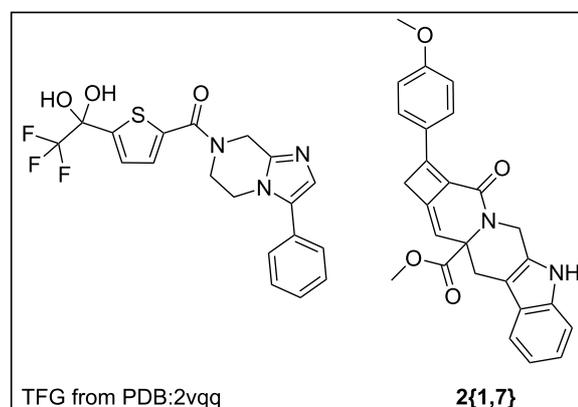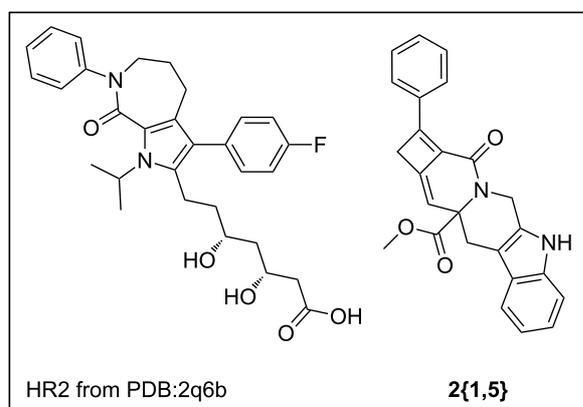
Figure 3: Results of high-throughput docking analysis. Top: A docking-score matrix arranged by compound IDs and PDB IDs; bottom: Structures of known ligands HR2 and TGF and the newly synthesized compounds **2{1,5}** and **2{1,7}**. Docking scores larger than 7.0 are red colored and can be mapped to K_d values less than 100 nM. The corresponding protein names of PDB IDs and the full docking-score matrix are listed in Supporting Information File 1.

than 7.0, a number that can be mapped to K_d values less than 100 nM, for several protein targets; (2) six compounds, **7{7}** and **2{1,4–1,8},** are predicted to be ligands for a single protein, human HDAC4 (PDBID:2vqq); (3) compounds **2{1,5}** and **2{1,7}** are predicted to be high-affinity ligands for 3-hydroxy-3-methylglutaryl-coenzyme A, reductase (PDBID:2q6b) and HDAC4, respectively, even though they are structurally different from the corresponding cocrystallized ligands HR2 and TGF; and (4) compounds **2{1,6}** and **6{10}** are predicted to be ligands for a total of 15 protein targets. This high number of potential protein targets may be due to the electronegative trifluoromethyl group on **2{1,6}** and the effect it would have on the α,β-unsaturated amide and the purported bioactivity of the pyrazole group of **6{10}**.

Ligand-based strategy for target prediction

Ligand-based target prediction algorithms have been developed based upon an established medicinal chemistry principle that structurally similar compounds, with comparable physical properties, should convey related biological properties [16,17]. In this study, structural similarities were calculated between the compounds of the β-carboline library and the bioactive compounds in the well-annotated database, ChEMBL version 13, the largest publicly available compound-target database, containing 1,143,682 distinct compounds, 8,845 targets and 6,933,068 bioactivity entries from 44,682 publications and PubChem bioassays [18,19]. The Openbabel FP2 fingerprint was used as a descriptor to assess similarities between molecules [20]. Tanimoto coefficients were calculated between the compounds of the β-carboline library and the ChEMBL database, and only β-carboline compounds with a Tc greater than 0.60 were considered for bioactivity analysis. A lower Tc threshold was used to identify a larger number of bioactivity targets. Table 4 lists the most promising bioactive targets for the newly synthesized β-carbolines together with the structurally similar lead compounds in ChEMBL along with their reported potency and literature citations. Several interesting results emerge from the comparison study performed, including a number of targets that the compounds should be screened against, such as C–C chemokine receptor type 3, gamma-

Table 4: Potential targets of β-carbolines based upon bioactivity data in ChEMBL.

target	CHEMBL compound	bioactivity type and reference	our compound	similarity score
C–C chemokine receptor type 3	CHEMBL33838	IC$_{50}$ = 325 nM [21]	**A**	0.79
Leishmania donovani	CHEMBL55830	IC$_{50}$ = 1.42 μM [22]	**4**	0.80
Acanthocheilonema viteae	CHEMBL44573	Activity = 94% [23]	**6{5}**	0.86
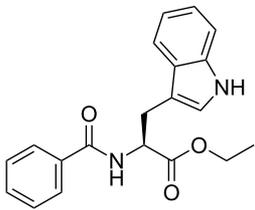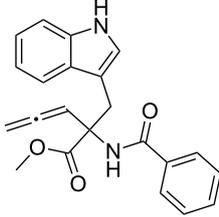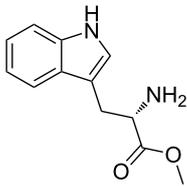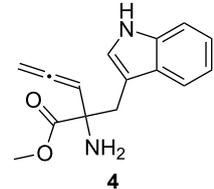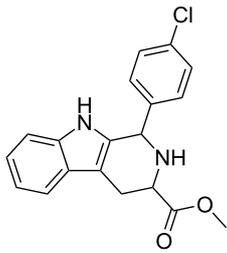
Table 4: Potential targets of β-carbolines based upon bioactivity data in ChEMBL. (continued)

Gamma-aminobutyric acid receptor subunit gamma-2	CHEMBL358326	IC_{50} = 250 nM [24]	**6{1}** 0.84
5-Hydroxytryptamine receptor 6	CHEMBL370935	K_i = 271.3 nM [25]	**7{1}** 0.69
3-Hydroxyacyl-CoA dehydrogenase type-2	CHEMBL1382101	Potency = 31.6 µM PubChem AID:893	**7{7}** 0.64
Benzodiazepine receptors	CHEMBL11901	K_i = 510 nM [26]	**1{16}** 0.72
Angiotensin-converting enzyme	CHEMBL148616	IC_{50} = 500 nM [27]	**9{1,3}** 0.71
Antithrombotic potency	CHEMBL1089460	IC_{50} = 8.56 nM [28]	**9{1,4}** 0.69
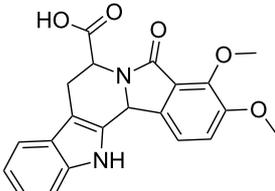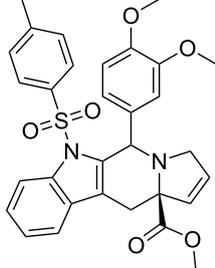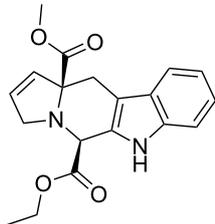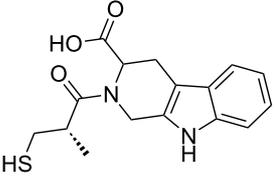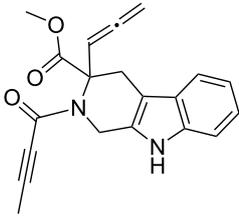
Table 4: Potential targets of β-carbolines based upon bioactivity data in ChEMBL. (continued)
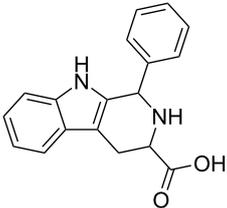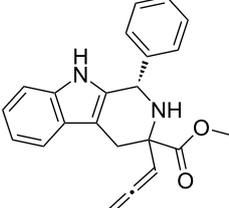
Mitogen-activated protein kinase 1	CHEMBL44295	Potency = 0.794 µM PubChem AID:995	**6{4}** 0.82
Breast adenocarcinoma cells	CHEMBL1650665	Inhibition = 78% [29]	**6{6}** 0.79
DNA polymerase iota	CHEMBL1360719	Potency = 1.78 µM PubChem AID:588590	**6{7}** 0.85
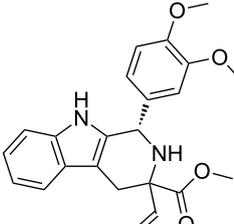
aminobutyric acid receptor subunit gamma-2, breast adenocarcinoma cells, 5-hydroxytryptamine receptor 6, angiotensin-converting enzyme, and DNA polymerase iota. Moreover, nine of the twelve compounds are represented by allene precursors, ones that were not originally considered in the diversity analysis.

Conclusion

A library of 34 β-carboline-containing compounds was synthesized utilizing a skeletal diversification strategy. High-throughput docking and ligand-based protocols were implemented to predict potential biological targets of the newly synthesized β-carbolines. The docking approach uses a structure-based technology to predict preferred interactions between compounds and protein targets, whereas the ligand-based method uses ligand similarity coefficients to identify potential biological targets. The complementary nature of these two protocols is evidenced by the fact that there was no overlap in the predicted biological targets. Furthermore, the in silico screening of these compounds is intended to add value to the library, by directing them to appropriate biological assays. Such strategies can also be used to explore the mechanisms of a biologically active compound in bioassays whose molecular target is as of yet unidentified.

Supporting Information

Supporting Information File 1

Experimental procedures and spectral data for compounds **1{5}**, **1{16}**, **2{1,3–8}**, **3{1,2–4}**, **4**, **6{1–10}**, **6{12–13}**, **6{15}**, **7{1–7}**, **9{1,2–4}**.
[http://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-8-117-S1.pdf]

Supporting Information File 2

The complete listing of the proteins and their PDB IDs (Targets Docking Score Matrix).
[http://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-8-117-S2.xls]

Acknowledgements

References

1. Stockwell, B. R. *Nature* **2004**, *432,* 846–854. doi:10.1038/nature03196
2. NIH Common Fund Office of Strategic Coordination. Molecular Libraries and Imaging. http://commonfund.nih.gov/molecularlibraries/overview.aspx (accessed June 14, 2012).
3. Dandapani, S.; Marcaurelle, L. A. *Curr. Opin. Chem. Biol.* **2010**, *14,* 362–370. doi:10.1016/j.cbpa.2010.03.018
4. Painter, T. O.; Wang, L.; Majumder, S.; Xie, X.-Q.; Brummond, K. M. *ACS Comb. Sci.* **2011**, *13,* 166–174. doi:10.1021/co100052s
5. Diness, F.; Beyer, J.; Meldal, M. *Chem.–Eur. J.* **2006**, *12,* 8056–8066. doi:10.1002/chem.200600138
6. Siebert, M. R.; Osbourn, J. M.; Brummond, K. M.; Tantillo, D. J. *J. Am. Chem. Soc.* **2010**, *132,* 11952–11966. doi:10.1021/ja102848z
7. Murakami, M.; Matsuda, T. Cycloaddition of Allenes. In *Modern Allene Chemistry;* Krause, N.; Hashmi, S. K., Eds.; Wiley-VCH: Weinheim, Germany, 2005; pp 727–733.
8. Huryn, D. M.; Brodsky, J. L.; Brummond, K. M.; Chambers, P. G.; Eyer, B.; Ireland, A. W.; Kawasumi, M.; LaPorte, M. G.; Lloyd, K.; Manteau, B.; Nghiem, P.; Quade, B.; Seguin, S. P.; Wipf, P. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108,* 6757–6762. doi:10.1073/pnas.1015251108
9. *Maestro 8.5.207-QikProp,* 2.1; Schrödinger, LLC: New York. For calculations of physicochemical properties.
10. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Delivery Rev.* **1997**, *23,* 3–25. doi:10.1016/S0169-409X(96)00423-1
11. Kitchen, D. B.; Decornez, H.; Furr, J. R.; Bajorath, J. *Nat. Rev. Drug Discovery* **2004**, *3,* 935–949. doi:10.1038/nrd1549
12. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28,* 235–242. doi:10.1093/nar/28.1.235
13. Gao, Z.; Li, H.; Zhang, H.; Liu, X.; Kang, L.; Luo, X.; Zhu, W.; Chen, K.; Wang, X.; Jiang, H. *BMC Bioinformatics* **2008**, *9,* 104. doi:10.1186/1471-2105-9-104
14. Jain, A. N. *J. Med. Chem.* **2003**, *46,* 499–511. doi:10.1021/jm020406h
15. Jain, A. N. *J. Comput.-Aided Mol. Des.* **2007**, *21,* 281–306. doi:10.1007/s10822-007-9114-2
16. Martin, Y. C.; Kofron, J. L.; Traphagen, L. M. *J. Med. Chem.* **2002**, *45,* 4350–4358. doi:10.1021/jm020155c
17. Xie, X.-Q.; Chen, J.-Z. *J. Chem. Inf. Model.* **2008**, *48,* 465–475. doi:10.1021/ci700193u
18. Bender, A. *Nat. Chem. Biol.* **2010**, *6,* 309. doi:10.1038/nchembio.354
19. Heikamp, K.; Bajorath, J. *J. Chem. Inf. Model.* **2011**, *51,* 1831–1839. doi:10.1021/ci200199u
   Heikamp et al. carried out a large-scale similarity search investigation on 266 well-defined compound activity classes extracted from the ChEMBL database using molecular fingerprints. In 203 of 266 test cases (~76%), a compound recovery rate of at least 50% was observed.
20. Guha, R.; Howard, M. T.; Hutchison, G. R.; Murray-Rust, P.; Rzepa, H.; Steinbeck, C.; Wegner, J.; Willighagen, E. L. *J. Chem. Inf. Model.* **2006**, *46,* 991–998. doi:10.1021/ci050400b
21. Dhanak, D.; Christmann, L. T.; Darcy, M. G.; Jurewicz, A. J.; Keenan, R. M.; Lee, J.; Sarau, H. M.; Widdowson, K. L.; White, J. R. *Bioorg. Med. Chem. Lett.* **2001**, *11,* 1441–1444. doi:10.1016/S0960-894X(01)00248-7
22. Singh, I. P.; Jain, S. K.; Kaur, A.; Singh, S.; Kumar, R.; Garg, P.; Sharma, S. S.; Arora, S. K. *Eur. J. Med. Chem.* **2010**, *45,* 3439–3445. doi:10.1016/j.ejmech.2010.04.033
23. Srivastava, S. K.; Agarwal, A.; Chauhan, P. M. S.; Agarwal, S. K.; Bhaduri, A. P.; Singh, S. N.; Fatima, N.; Chatterjee, R. K. *J. Med. Chem.* **1999**, *42,* 1667–1672. doi:10.1021/jm9800705
24. Lippke, K. P.; Schunack, W. G.; Wenning, W.; Mueller, W. E. *J. Med. Chem.* **1983**, *26,* 499–503. doi:10.1021/jm00358a008
25. Cole, D. C.; Lennox, W. J.; Stock, J. R.; Ellingboe, J. W.; Mazandarani, H.; Smith, D. L.; Zhang, G.; Tawa, G. J.; Schechter, L. E. *Bioorg. Med. Chem. Lett.* **2005**, *15,* 4780–4785. doi:10.1016/j.bmcl.2005.07.028
26. Cain, M.; Weber, R. W.; Guzman, F.; Cook, J. M.; Barker, S. A.; Rice, K. C.; Crawley, J. N.; Paul, S. M.; Skolnick, P. *J. Med. Chem.* **1982**, *25,* 1081–1091. doi:10.1021/jm00351a015
27. Ksander, G. M.; de Jesus, R.; Yuan, A.; Ghai, R. D.; McMartin, C.; Bohacek, R. *J. Med. Chem.* **1997**, *40,* 506–514. doi:10.1021/jm960583g
28. Liu, J.; Jiang, X.; Zhao, M.; Zhang, X.; Zheng, M.; Peng, L.; Peng, S. *J. Med. Chem.* **2010**, *53,* 3106–3116. doi:10.1021/jm901816j
29. Mohamed, H. A.; Girgis, N. M. R.; Wilcken, R.; Bauer, M. R.; Tinsley, H. N.; Gary, B. D.; Piazza, G. A.; Boeckler, F. M.; Abadi, A. H. *J. Med. Chem.* **2011**, *54,* 495–509. doi:10.1021/jm100842v