A Marine Natural Product Database

Jing Lei and Jiaju Zhou*

Laboratory of Computer Chemistry, Institute of Chemical Metallurgy, Chinese Academy of Sciences, P.O. Box 353, Beijing 100080, China

Received October 23, 2001

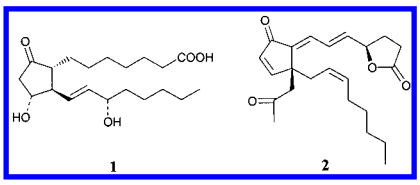
A database of marine natural products has been developed. The database contains approximately 6000 chemical compounds derived from over 10 000 marine-derived materials. For each compound, the structure, physical and chemical properties, marine source, and biological activities are given. A computer program for searching this database has also been developed and is described.

INTRODUCTION

The oceans of the planet represent a huge unexplored resource, and, as the exploitation of terrestrial resources proceeds, the marine environment offers a new frontier for research. From a biological perspective, the ocean is indeed a treasure: it contains perhaps 200 000 organisms, all of which survive in the terrestrially unusual conditions of higher salt content, low or zero light, high pressure, and unusually high or low temperatures. These properties of the marine environment make it very likely that marine organisms, and the chemicals they produce, will be quite different from those associated with terrestrial biology. Some examples of the interesting new types of chemicals found in the marine environment are shown in Table 1.

Serious exploration of the marine environment began only recently. Since 1960, some 10 000 marine-derived materials have been studied and about 6000 specific chemicals have been identified.[1] During the last 20 years, marine science has seen increasing levels of research. In 1998, some 840 new compounds were isolated and characterized. Of these 460 showed some bioactivity and indications of possible utility as, for example, anticancer, antibacterial, or cardiovascular agents,[2] and several marine chemicals are currently in clinical trials, especially those associated with cancer. Among the anticancer compounds, bryostatin 1 serves as a good example of past and current trends in marine biomedical research. Bryostatin 1 is a macrocyclic metabolite that was first isolated from bryozoan *Bugula neritina* from the Gulf of California and its structure was determined by X-ray crystallography in 1982. Bryostatin 1 was found to affect protein kinase C activity, which may be the mechanistic basis for both anticancer and immunostimulating activity. It has been developed for the treatment of melanoma, non-Hodgkins lymphoma, and renal cancer and is currently in phase 2 clinical trials. Recently, evidence favoring a symbiotic origin for bryostatin 1 has been present, opening the way for biotechnological manipulation of the biosynthetic genes. Furthermore, it has been shown that semisynthetic bryostatins retain the activity of the natural product.[3]

Marine chemicals often possess quite novel structures and this in turn leads to pronounced biological activity and novel pharmacology. The study of such chemicals therefore is a very promising endeavor. For example, prostaglandins such as prostaglandin E_1 (**1**), first isolated in 1934 from sheep spermatophore,[3] shows exciting uterine and antihypertensive activities. The cost of prostaglandin from this source is high because spermatophores are scarce and synthetic approaches to such compounds are difficult. Clavulactone II (**2**), isolated from the Okinawan soft coral *Clavularia viridis*,[4] is a natural prostanoid possessing a γ-lactone moiety in



the α-side-chain, and its discovery may provide a new, inexpensive source of prostanoids.

Marine chemicals often possess quite novel structures which in turn leads to pronounced biological activity and novel pharmacology. The study of such chemicals therefore is a very promising endeavor. There are three parallel tracks in marine natural products chemistry: marine toxins, marine biomedicinals, and marine chemical ecology. Integration of these three fields of study gives marine natural products chemistry its unique character and vigor.

The search among marine chemicals for medically useful agents involves two steps, discovering the type of biological activity and studying the pharmacological mechanism of the activity. Both of these tasks can be aided considerably by access to a database of marine-derived chemicals and their properties. Such a database has been developed for use in research into Traditional Chinese Medicines.[5,6]

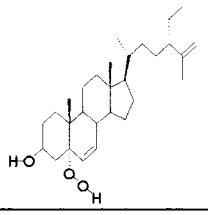
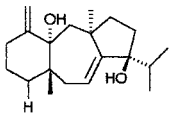
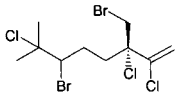
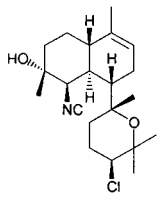
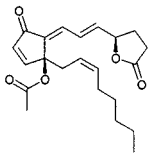
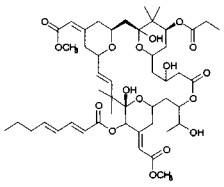
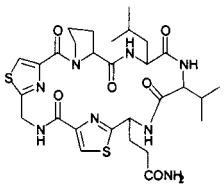
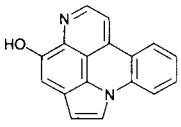
DATABASE CONTENT

To support our systematic study of marine natural products, we have developed and here describe a Marine Natural Products database. This database contains detailed information on the source of marine natural products, their structures, chemical and physical properties, and pharmacological

* Corresponding author phone: 0086-10-62626703; fax: 0086-10-62561822; e-mail: jjzhou@lcc.icm.ac.cn.

10.1021/ci010111x CCC: $22.00
Published on Web 03/30/2002

Table 1. Selected Marine Natural Products

(24S)-Ethyl-5α-hydroperoxycholesta-6,25-dien-3β-ol		Cytotoxic against P388, KB, A-549, HT-299 cell lines.	Green algae *Codium arabicum.*
Dolatriol		Cytotoxic	Brown algae *Dictyota divaricata.*
Halomon		Antineoplastic.with cell-specific cytotoxicity.	Red alga *Portieria homehanni*
O^9-Kalihinol Y		*Plasmodium falciparum* inhibitor	Philippine sponge *Phakellia pulcherrima*
Clavulactone II		Antihypertensive; uterine stimulant	Soft coral *Clavularia viridis*
Bryostatin I		Antitumor agent in Phase 2 trials	Japanese bryozen *Bugula neritina*
Dolastatine 10		Antitumor agent in Phase 1 & 2 trials. Antifungal	Sea hare *Dolabella uricularia*
Arnoamine A		Antitumor agent. Topoisomerase inhibitor.	Ascidian *Cystodytes* spp.

function. Most of the data has been derived from the secondary literature, i.e., books and reviews.

The database consists of four main segments, which are described below.

Table 2. Pharmacological Classes

abortifacients	3	antineoplastic (radiosource)	93
ACE inhibitors	4	antineoplastics	90
α-adrenergic agonists	9	antineutropenics	94
α-adrenergic blockers	10	antiosteoporotics	95
β-adrenergic agonists	122	antioxidant	275
β-adrenergic blockers	123	antipagetics	96
adrenocortical suppressants	5	antiparkinsonians	97
adrenocorticotropic hormone	6	antiperiodics	253
alcohol deterrents	257	antiperspirants	278
aldose reductase inhibitor	7	antipheochromocytoma	98
aldosterone antagonist	8	antipneumocystic	99
allergenic extracts	189	cholinesterase inhibitors	137
α-glucosidase inhibitor	11	cholinesterase reactivators	138
alzheimer-type dementia	190	CNS stimulants	139
amino acids	254	CNS, miscellaneous	287
aminoglycoside antibiotics	41	coccidiostats	263
amphenicol antibiotics	42	contact lens products	288
anabolic steroids	12	contraceptives	140
analgesic (dental)	13	cytoprotectants	141
analgesic (narcotic)	14	debridant	142
analgesic (non-narcotic)	15	decongestants	143
analgesic (topical)	219	deficiency anemias	289
analgesics, general	227	dental preparations	290
anaphylaxis treatment kit	264	dentistry	258
androgens	16	dentrifice/denture products	291
anesthesia, adjuncts to/analeptics	265	depigmentor	144
anesthetic (inhaled)	17	dermatitis suppressant	145
anesthetic (intravenous)	18	diagnostic acid	147
anesthetic (local)	19	diagnostic acid (radioisotope)	146
anesthetic (rectal)	266	digestive aids	148
anesthetic (topical)	267	disorders, acid/peptic	292
anesthetics, adjuncts	268	diuretics	149
angiotensin II antagonist	20	dopamine receptor agonists	150
anorectal products	269	dopamine receptor antagonists	151
anorexics	21	growth hormone secretion disorders	293
antacids	22	ear wax removal	249
anterior pituitary/hypothalamic function	270	emetics	153
anthelmintics (cestodes)	23	enkephalinase inihitors	154
antidiabetics	62	enzyme cofactors	156
antidiarrheals	63	enzyme inducers	157
antidiuretics	64	enzyme inhibitors	261
antidotes	65	enzymes	155
antidotes, general	272	estrogens	158
antidotes, specific	273	expectorants	159
antidyskinetics	66	extrapyrainidat movement disorders	294
antieczematics	67	fibrinogen receptor antagonists	160
antiemetics	68	gastric proton pump inhibitors	161
antiestrogens	69	gastric secretion inhibitors	162
antifibrotics	70	gastric secretion stimulant	163
antiflatulents	71	gastrointestinal, miscellaneous	295
antifungals	72	gastrointestinals	296
antiglaucoma	73	gastroprokinetics	164
antigonadotropin	74	gonad stimulation prinicipal	165
antigout	75	growth hormone inhibitors	166
antihistaminics	76	growth stimulants	168
antihyperlipoproteinemics	77	hematinics	169
antihyperphosphatemics	78	hematologics	297
antihypertensives	79	hematopoietics	170
antihyperthyroids	80	hemolytics	171
antihypotensives	81	hemorrhoid treatment	256
antihypothyroids	82	hemostatics	172
antiinflammatories	84	heparin antagonists	173
antimalarials	86	hepatoprotectants	174
antimanics	87	hepatotoxics	274
antimetabolites	275	histamine H_2 receptor antagonists	175
antimethemoglobinemics	88	HMG CoA reductase inhibitors	176
antimicrobials	276	homeopathic products	299
antimigraines	89	hormonal/biological response modifiers	300
antimycobacterials (inc1 anti leprosy)	277	immunomodulators	177
antineoplastic (hormonal)	91	immunosuppressants	178
antineoplastic (other)	92	insulin sensitizers	179
ion exchangers	180	relaxants/stimulants, urinary tract	320
keratolytics	181	relaxants/stimulants, uterine	321
lactation stimulating hormones	182	relief of pain	322
laxatives/cathartics	183	regulators of electrolytes/water balance	323
leprostatic antibiotics	45	replenishers	208
leukotriene antagonists	184	respiratory inhibitor	277
LH−RH agonists	185	respiratory stimulants	209
lincosamide antibiotics	46	respiratory tract	324
lipotropics	186	retroviral protease inhibitors	210
5-lipoxygenase inhibitors	2	RT inhibitors	211
lozenge products	306	rubifacients	238
lubricants	259	sclerosing agents	212
lupus erythematosus inhibitors	187	sedative, hypnotics	213
MAO inhibitors	193	serenics	214

Table 2. (Continued)

matrix metalloproteinase inhibitors	188	serotonin receptor agonists	215
medical research	244	serotonin receptor antagonists	216
menstrual products	307	serotonin reuptake inhibitors	217
metabolics/nutrients	308	serotonin uptake inhibitors	218
mineralocorticoids	191	skeletal muscle hyperactivity	325
miotics	192	skin/mucous membranes	326
mouth, canker sore products	309	sleep products (OTC)	327
mucolytics	194	spermaticides	240
muscle relaxants	195	steroidal antiinflammatories	85
myasthenia gravis	310	surgical aids	246
mydriatics	196	therapeutics − radiopharmaceuticals	328
narcotic antagonists	197	thrombolytics	220
neurologics	311	thromboxane inhibitors	221
neuromuscular blocker	278	thyroid hormones	222
nootropics	198	thyrotropic hormones	223
nutrition, enteral/parentenal	312	tocolytics	224
ocular anti-infective/antiinflammatory	313	topical analgesics	329
ophthalmics	314	topical anti-infectives	330
ophthalmics, miscellaneous	315	topical protectants	225
ophthalmics-antiallergy agents	316	topical steroids	331
otics	317	topisomerase inhibitors	226
otics, topical	318	ultaviolet screens	228
oxytocics	199	unclassified/miscellaneous	332
pediculosides	262	uricosurics	229
personal care products (vaginal)	319	vaccines	239
pharmaceutical aids	251	vasodilators, cerebral	230
pigmentation agents	200	vasodilators, coronary	231
plasma volume expanders	201	vasodilators, peripheral	232
potassium channel activators	202	vasoprotectants	233
progestogens	203	vertigo protectants	334
prolactin inhibitors	204	vitamins, vitamin sources	234
prostaglandins	205	vulneraries	235
protease inhibitors	206	Wilson's disease treatment	236
pulmonary surfactants	207	wounds and burns	243
5α-reductase inhibitor	1	xanthine oxidase inhibitors	237

Table 3. Biological Activity of Some Marine Toxins

toxin	biological activity
tetrodotoxin	blocks neurotransmission; has sodium channel-specific antiarrhythmic activity, possible antitumor activity and is implicated in food poisoning; highly toxic, $LD_{50} \sim 5-10\ \mu g/kg$ (mouse, oral)
saxitoxin	blocks voltage-sensitive sodium channels; high affinity ($K_D = 1.9$ nM)
aplysiatoxin	antileukocytothemic; $LD_{50} = 0.3$ mg/kg (mouse, oral)
holotoxin	inhibits oocyte maturation; antitumor agent
palytoxin	tumor promoter; binds to (Na^+/K^+) ATPase, $LD_{50} = 0.15\ \mu g/kg$.
nereistoxin	neurotoxins; $LD_{50} = 1.8$ mg/kg (mouse, oral)
ciguatoxins	anticholinesterase; highly toxic to humans

1. Identification. Every chemical compound is identified by its name, structure, and a structural type identifier (steroid, alkaloid, etc.). Chemical and physical properties, as available, are also described in detail in the file.

2. Biological Activity. The reported biological activity of the compound is carried in the database. The classification in *Drugs: Synonyms and Properties*,[7] with 253 separate classes is used here, as shown in Table 2.

Marine toxins are an important subgroup of marine chemicals and have received considerable attention because their toxicity may be a sign of a medically valuable biological activity. A number of marine toxins have been used[3,8] to probe biological and pharmacological systems, and some of these are listed in Table 3.

Among marine toxins, maitotoxin (**3**) (see Chart 1) is one of the most complex and lethal nonprotenaceous toxins known. This compound is a calcium channel activator which inhibits cell cycle progression through the G1/S and G2/M transitions and prevents CDC2 kinase activation in GH4Cl

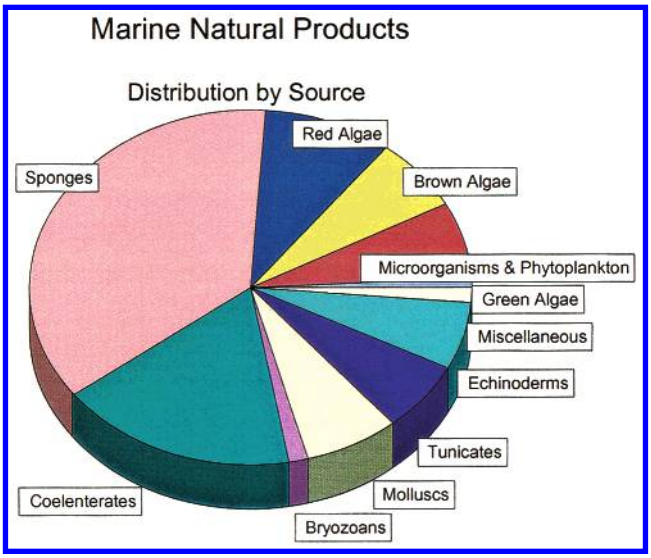


Figure 1. Marine organisms.

Cells.[9] It also stimulates phosphoinositide breakdown in smooth muscle cells, NCB-20 cells, and PC12 cells through a nifedipine-insensitive mechanism.[9] It was isolated from the epiphytic dinoflagellate *Gambierdiscus toxicus*, and the elucidation of its structure is a *tour de force* of modern structural chemistry.

3. Sources. Precise identification of the organism which is the source of the material in question requires a classification system for marine organisms. We use the system proposed by Faulkner[10] in which 11 distinct organisms are defined, as shown in Figure 1. The number of occurrences in each organism of compounds in the database is given in the figure. There is a fairly even distribution of compounds among these species although sponges and coelenterates are

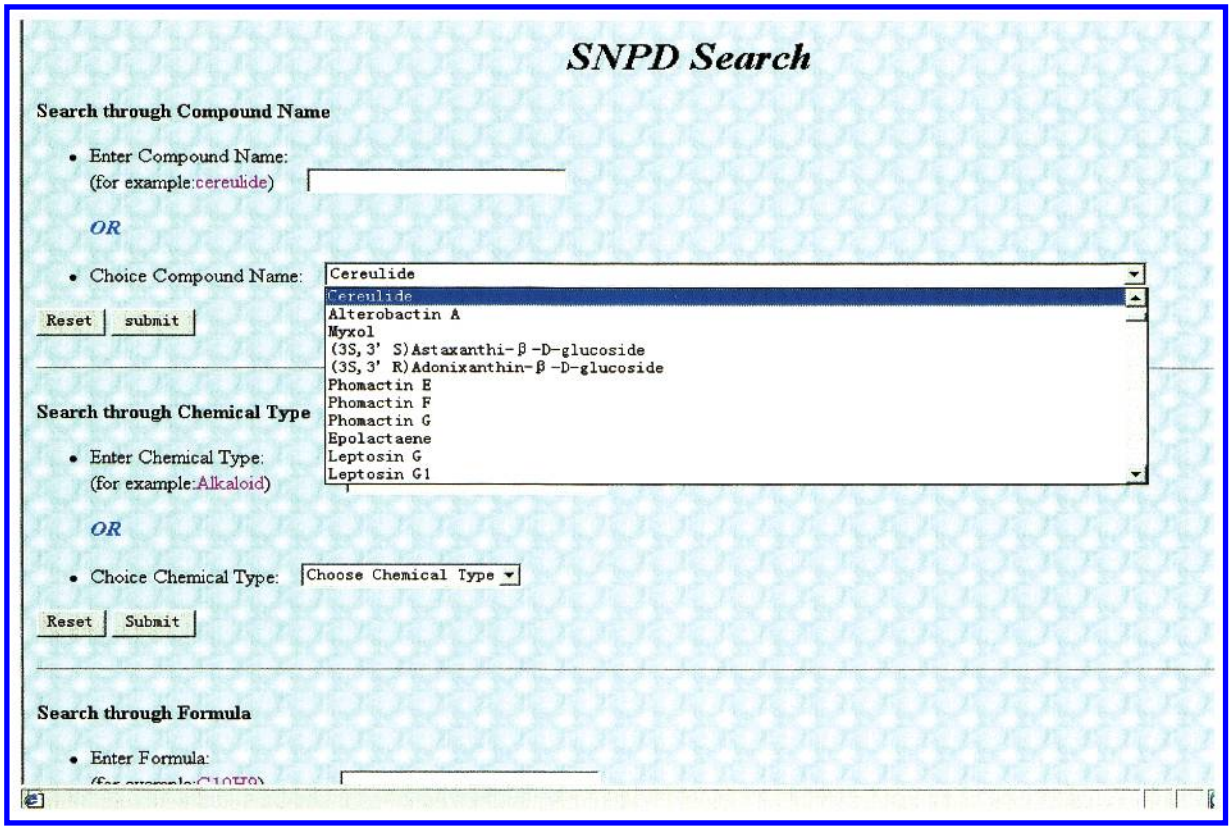
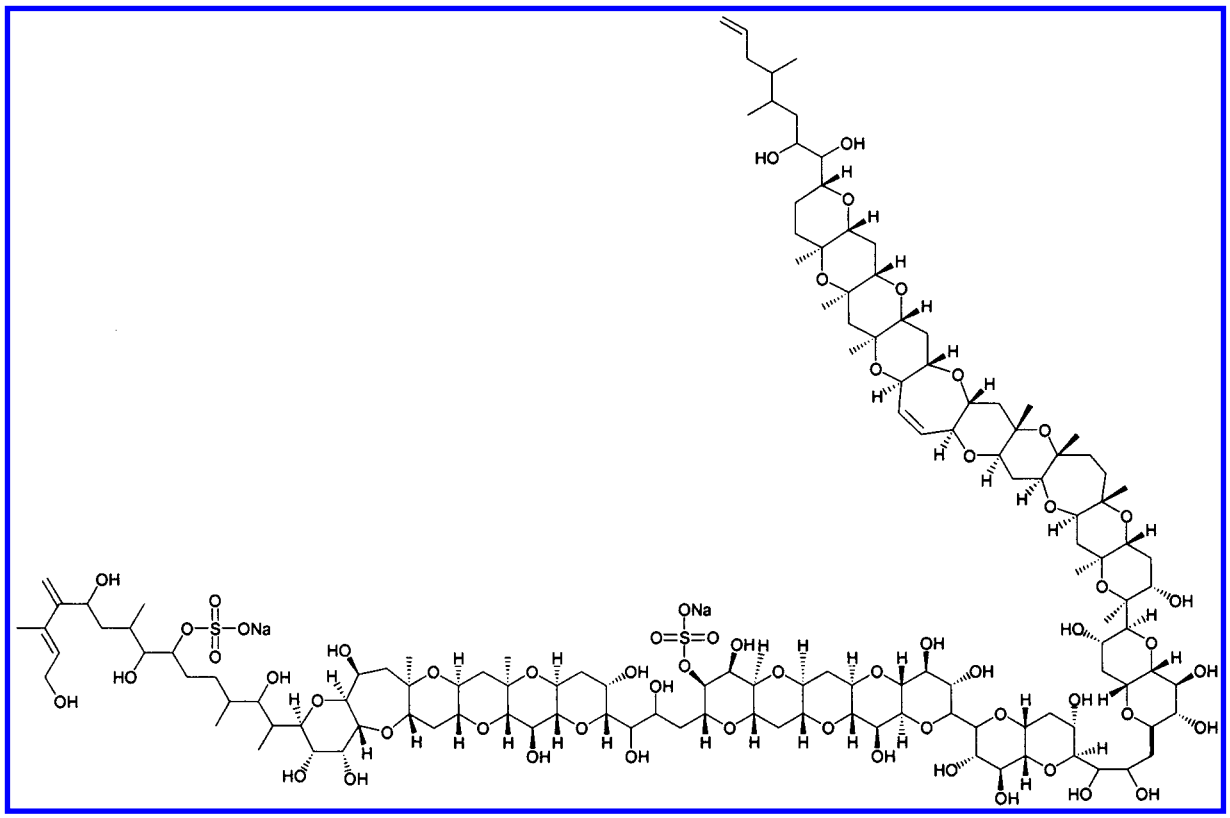


Figure 2. Search for "cereulide".

Chart 1



dominant, possibly as a reflection of the accessibility relative to bryozoans for example.

4. References. This segment contains information concerning the publication(s) from which the data in the record were taken.

DATABASE DESIGN

The Marine Natural Products Database is a cross-subject database with three subjects: biology, chemistry, and pharmacology. The biology sector is summarized by the classification of marine organism into the 11 categories

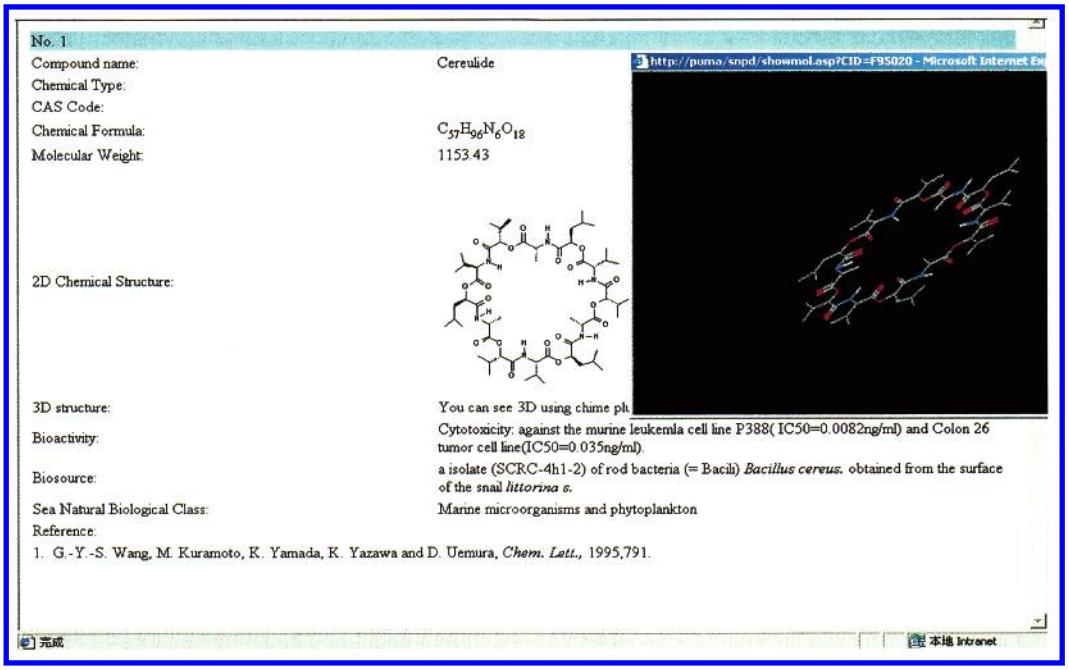


Figure 3. "Cereulide" query result.

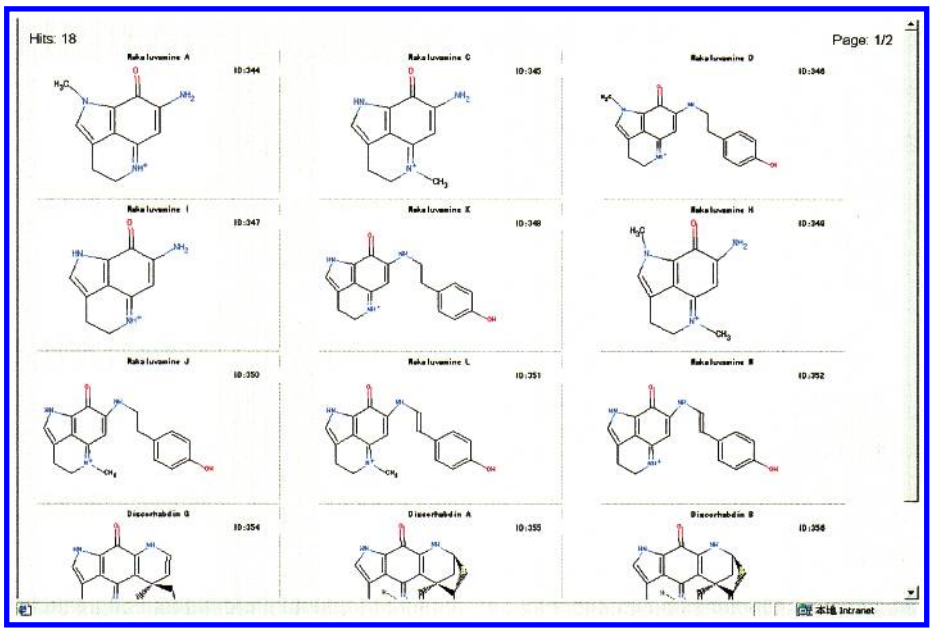


Figure 4. Eighteen 1,3,4,5-tetrahydropyrrolo[4,3,2-*de*]quinolines.

shown in Figure 2, and the database can be searched for members of any of these classes. The chemistry segment contains chemical data on the several thousand compounds which have been characterized as components of marine organisms. Searches for full or partial 2D or 3D structures are supported, as are searches by chemical name, molecular formula, and so on. All structures are developed and stored in the MolFile format.[11] Biological activity and pharmacology data for each compound are stored in the pharmacology section.

Structural data stored in the MolFile format can be visualized using Chime MDL software available from MDL.[12] This can be used in a stand alone PC or on the Web by means of Netscape or Internet Explorer browsers. Structures can be represented in a variety of styles such as ball-and-stick, wireframe, and so on, zoomed, and moved or rotated in any direction. Rasmol scripts or chemical structure markup language (CSML) can be used to change or argument the display mode. Geometries such as bond angles, torsion angles, interatomic separations, or hydrogen bond dimensions can be calculated as needed. Structures can be extracted from the database and transferred into ISIS/Draw[9−12] for copying or editing.

Searches for specific chemicals may be carried out with CAS Registry Number, chemical name, and full or partial structures. Chemical names may be entered for a search or may be selected from the database name dictionary, as shown in Figure 3. Data input is not case-sensitive.

Successively entered search-terms are combined in a Boolean AND operation unless otherwise specified by the user, who can select OR or NOT operators. Display of results from a search is controlled by the user; by default, the full record of a retrieved compound is display, as shown in Figure 3.

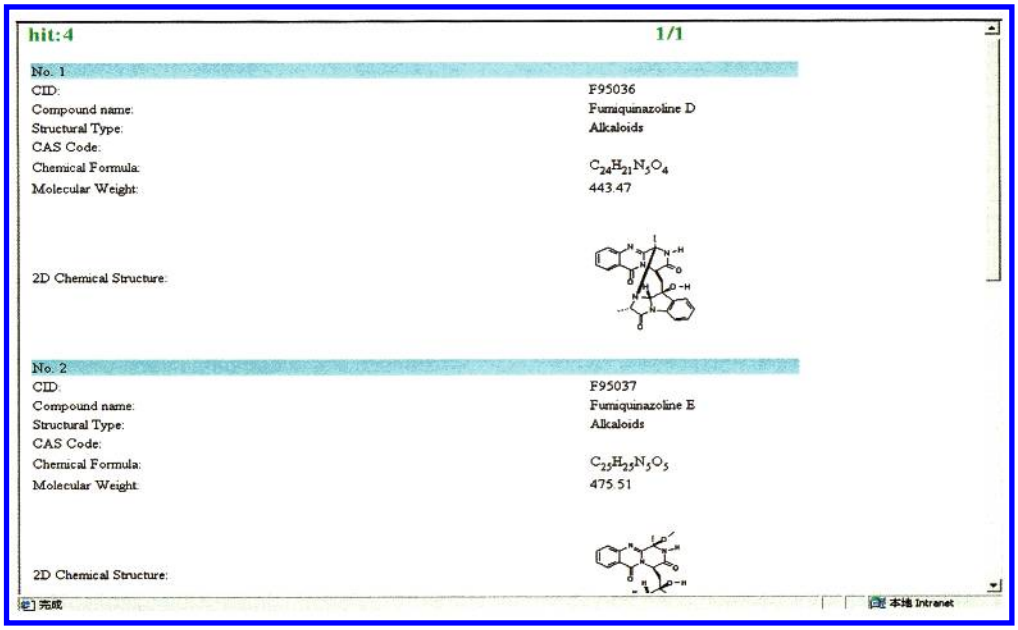
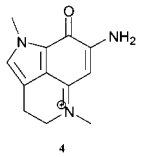


Figure 5. Fumiquinazoline alkaloids.

To carry out a structure-based search, the user must use a program such as ChemDraw[13] to enter the appropriate chemical structure, which is then used as the basis for a full or substructure search. An example of this is given in Figure 4 in which a search for the 1,3,4,5-tetrahydropyrrolo[4,3,2-*de*]quinoline nucleus[14](**4**) retrieves 18 different structures.



4

Structure and properties can be combined in searches. Thus the four "alkaloids" whose names contain the word "fumiquinazoline" can be retrieved, as shown in Figure 5.

CONCLUSIONS

Marine sources have the highest probability of yielding natural products with unprecedented carbon skeletons and interesting biological activity. The tools used to identify marine metabolites are constantly improving in scope and sensitivity, allowing the marine organisms to be explored.

The establishment of MNPD has great effect on studying marine organism and mining new molecules. Three application examples of the database are provided. Future developments will be planned in two directions: to establish 3D MNPD and to implement 3D conformation search and molecular filter.

ACKNOWLEDGMENT

The authors would like to acknowledge the assistance of the Editor, G. W. A. Miline, in preparing this paper.

REFERENCES AND NOTES

(1) Guan, H.; Geng, M.; Wang, C. 21th century, China Marine Drugs (in Chinese). *Chinese J. Marine Drugs*. **2000**, *19*, 44−47.

(2) Faulkner, D. J. Marine natural products. *Nat. Prod. Rep*. **2000**, *17*, 7−55.

(3) Faulkner, D. J. Highlights of marine natural products chemistry (1972−1999). *Nat. Prod. Rep*. **2000**, *17*, 1−6.

(4) Iguchi, K.; Iwashima, M.; Watanabe, K. Clavuloactone, New Marine Prostanoids with a γ-lactonic moiety in the α-Side-Chain from the Okinawan Soft Coral *Clavularia viridis*. *J. Nat. Prod*. **1995**, *58*, 790.

(5) Yan, X.; Zhou, J.; Zhi, H. Concept Design of Computer-Aided Study on Traditional Chinese Drug. *J. Chem. Inf. Comput. Sci*. **1999**, *39*, 86−89.

(6) Yan, X.; Zhou, J, Xie, G. *Traditional Chinese Medicines*; Ashgate Publishing: ISBN 0 566 08210 1.

(7) *Drugs: Synonyms and Properties*; Milne, G. W. A.,Ed.; Ashgate Pulishing: 2000.

(8) Takeshi, Y.; Michio, M. Marine Toxins. *Chem. Rev*. **1993**, *93*, 1897−1910.

(9) Van Dolah, F. M.; Ramsdell, J. S.; Maitotoxin, a calcium channel activator, inhibits cell cycle progression through the G1/S and G2/M transitions and prevents CDC2 kinase activation in GH4Cl cells. *J. Chem. Physiol*. **1996**, *166*, 49−56. Nakahata, N.; Yaginuma, T.; Ohizumi, Y. Maitotoxin-induced phosphoinositide hydrolysis is dependent on extracellular but not intracellular Ca2+ in human astrocytoma cells. *Jpn. J. Pharmacol*. **1999**, *81*, 240−3.

(10) Faulkner, D. J. Marine natural products. *Nat. Prod. Rep*. **1996**, *13*, 75.

(11) Dalby, A.; Nourse, J. G.; Hounshell, W. D.; Gushurst, A. K. I.; Grier, D. L.; Leland, B. A.; Laufer, J. Description of Several Chemical Structure File Formats Used by Computer Programs Developed at Molecular Design Limited. *J. Chem. Inf. Comput. Sci*. **1992**, *32*, 244−255.

(12) Molecular Design Ltd., 14600 Catalina Street, San Leandro, CA 94577.

(13) CambridgeSoft Corpn., 875 Massachusetts Avenue, Cambridge, MA 02139.

(14) Ding, Q.; Chichak, K.; Lown, J. W. Pyrroloquinoline and Pyridoacridine Alkaloids from marine sources. *Curr. Med. Chem*. **1999**, *6*, 1−27.

CI010111X