Send Orders for Reprints to reprints@benthamscience.net

Medicinal Chemistry, **XXXX**, *XX*, **1-10**

RESEARCH ARTICLE

GreenMolBD: Nature Derived Bioactive Molecules' Database

S. M. Zahid Hosen[1,*,#], Md. Junaid[1,#], Muhammad Shaiful Alam[1,#], Maruf Rubayed[1,#], Raju Dash[2], Rasheda Akter[1], Tania Sharmin[1], Nusrat Jahan Mouri[1], Mohammad Ali Moni[3], Mahmuda Khatun[1] and Mohammad Mostafa[1]

[1]Molecular Modeling Drug Design and Discovery Laboratory, Pharmacology Research Division, BCSIR Laboratories Chattogram, Bangladesh Council of Scientific and Industrial Research, Chattogram, Bangladesh; [2]Department of Anatomy, Dongguk University College of Medicine, Gyeongju 38066, Republic of Korea; [3]WHO Collaborating Centre on eHealth, UNSW Digital Health, School of Public Health and Community Medicine, Faculty of Medicine, UNSW Sydney, Australia

ARTICLE HISTORY

Received: September 01, 2021
Revised: September 13, 2021
Accepted: October 04, 2021

DOI:
10.2174/1573406418666211129103458

Abstract: ***Background*:** One of the essential resources for developing new drugs are naturally derived biologically active lead compounds. Biomedical researchers and pharmaceutical companies are highly interested in these plant-derived molecules to develop the new drug. In this process, collective information of the plants and their phytoconstituents with different properties and descriptors would greatly benefit the researchers to identify the hit, lead or drug-like compound.

***Aim and Objective*:** Therefore, the work intended to develop a unique and dynamic database GreenMolBD to provide collective information regarding medicinal plants, such as their profile, chemical constituents, and pharmacological evidence. We also aimed to present information of phytoconstituents, such as in silico description, quantum, drugability and biological target information.

***Methods*:** For data mining, we covered all accessible literature and books, and for in silico analysis, we employed a variety of well-known software and servers. The database is integrated by MySQL, HTML, PHP and JavaScript.

***Results*:** GreenMolBD is a freely accessible database and searchable by keywords, plant name, synonym, common name, family name, family synonym, compound name, IUPAC name, InChI Key, target name, and disease name. We have provided a complete profile of individual plants and each compound's physical, quantum, drug likeliness, and toxicity properties (48 type's descriptor) using in silico tools. A total of 1846 associated targets related to 6,864 compounds already explored in different studies are also incorporated and synchronized.

***Conclusion*:** This is the first evidence-based database of bioactive molecules from medicinal plants specially grown in Bangladesh, which may help explore and foster nature-inspired rational drug discovery.

Keywords: Plant database, compound database, natural products, Bangladeshi plant database, pharmacological evidence, *in-silico* properties.

1. INTRODUCTION

Numerous functionally significant plants have been widely explored and studied for human benefit. It has long been recognized that natural products (NPs) are a flourishing source of vast chemical diversity endowing the base for characterization of unique scaffold structures that function as initial points for influential drug design [1]. As biologically active NPs have a distinct selectivity towards cellular targets [2], they might offer selective ligands for disease-associated targets [3] and regulate the disease-associated pathways.

Advancement in large-scale network analysis enables researchers to begin the exploration of the mode of action of bioactive compounds in the context of biological networks, *e.g.,* protein-protein interaction network [4], drug-target network (DTN) [3, 5], disease pathways [6] and metabolic networks [7, 8]. Nevertheless, maximum investigations until now have focused on a small number of molecules. Researchers have investigated and statistically analyzed natural products under many categories, such as property distribution [9], chemical diversity [10-13], chemical space [14, 15], molecular scaffold [16-18], and contrast between NPs and other synthetic compound assemblies [18-20]. Yet, inclusive statistical analysis of natural products and contrast between NPs and other compounds have rarely been attempted as it was challenging to attain the large amount of data required (both annotations and structures).

*Address correspondence to this author at the Molecular Modeling Drug-design and Discovery Laboratory, Pharmacology Research Division, BCSIR Laboratories Chattogram, Bangladesh Council of Scientific and Industrial Research, Chattogram, Bangladesh; Tel: 01829675616;
E-mail: smzahidhosen@bcsir.gov.bd;
[#]*These authors are contributed equally to this work.*

1573-4064/XX $65.00+.00

Such developmental efforts and investigations can be assisted by the accessibility of data encompassing every plant's combined activity, various targets of its constituents, the regulated biological processes, biological pathways, and diseases for more varied classes of medicinal, and other functionally important plants. Recognized plant databases (*e.g.* The Plant List (http:/ /www.theplantlist.org), Plants of the World Online (http:/ /www.plantsoftheworldonline.org), the Plant Trait Database (https://www.try-db.org/ TryWeb/ Home.php), the PLANTS database (https:/ /plants.sc.egov.usda.gov), *etc*. provide wide-ranging information regarding the taxonomy, physical and chemical features, and other properties of a large number of plants; however, information regarding the plant constituents and action mechanisms is not included. Particular databases deliver other information about medicinal plants (*e.g.,* IMPPAT [21], KANpSAcK [22], NuBBE [23], BATMA-TCM [24], TCM @Taiwan [25], TM-MC [26], TCMSP [27], TCMID [28], TIPdb-3D [29], TCM-ID [30] and TCM Analyzer [31]). Among these databases, some of them provided information regarding the activities and targets of plant constituents. In specific, TCM Analyzer [31] and TCMSP [27] allow easy access to the information on various targets of every medicinal plant utilized in TCM and the associated diseases regulated by these plants.

Bangladesh is blessed with a multitude of medicinal plants belonging to more than 500 species. These ethnic medicinal plants have been extensively used in Ayurvedic, Unani, and Homeopathic medicine formulation in Bangladesh and have been playing a significant role by serving raw materials for contemporary medicinal preparations [32]. However, there is no detailed and classified data portal that organizes all the information of Bangladeshi plants as well as the information on a broad range of bioactive compounds, compound targets, ADME/T properties, drug-likeliness tests, and vendor names, *etc*. Hence, we developed a new database named GreenMolBD, which organized all the reported data regarding the Bangladeshi medicinal plants and *in silico* data regarding their phytochemical constituents. Although our data-portal is continuously updating, till now we have included 145 plants with 6, 837 unique compounds and their detailed taxonomical classification, physical properties, quantum information (heat of formation, dipole moment, electronegativity, HOMO-LUMO, and hardness), compound reactivity (DLP, nRFG, nLMR), TOPKAT (toxicity) properties, drug-likeliness test values, synonyms, compound sources, targets, and chemical vendors. Facilities have been set up for convenient access to the GreenMolBD database by multiple search modes, including 1) plant by names, synonym, common name or family, 2) compound by name, InChIkey, synonym, or IUPAC name, 3) targets of the compound, and 4) disease (Fig. **1**).

2. MATERIALS AND METHODS

2.1. Database Architecture and Web Interface

The GreenMolBD database is integrated by MySQL, an open-source relational database management system and Nginx, a web server that can also be used as a reverse proxy, load balancer, mail proxy, and HTTP cache, which works at the back end, and its web interfaces are built using HTML, PHP, and JavaScript as the front end. The database is de-

Fig. (1). The content architecture of GreenMolBD database *(A higher resolution / colour version of this figure is available in the electronic copy of the article).*

signed to enable access to its entries by different keywords, including the name of the plant, compound, target, and disease by using multiple searches and browse facilities.

2.2. Data Collection and Processing

2.2.1. Plant Selection

According to the National Encyclopedia of Bangladesh, more than 6,000 plant species are available in Bangladesh. From these, we have identified about 1000 Bangladeshi medicinal plants reported by Professor Dr. Abdul Ghani in his book titled "Medicinal plants of Bangladesh: chemical constituents and uses" [32] and Dr. Shaikh Bokhtear Uddin's database named "Medicinal plants of Bangladesh" (https://mpbd.cu.ac.bd). Among these plants, some are only available in Bangladesh. For the first phase of this database, we have reported 145 medicinal plants. A complete profile of each plant is created with useful information, such as common name, family, botanical description, availability, ethnobotanical use, chemical constituents, and pharmacological activity.

2.2.2. Plant Profile

To find the plants' common name, family name, botanical descriptor, availability, and ethnobotanical use, Google Scholar, books and literature databases were searched using the keywords "common name," "family name," "availability," and "ethnobotanical use" in combination with "plant name". Only reliable sources, such as scientific journals, credible databases, books, or official reports of government agencies, botanic gardens, international scientist networks, and charitable organizations, were selected. Information on pharmacological activities or diseases was acquired using the keywords "activity," "anticancer activity," "anti-diabetic activity," "antibacterial activity," "antifungal activity," "cytotoxic activity," "diuretic," "anti-malarial activity," "hepatoprotective activity," "liver disease," "kidney disease," "CNS activity," and "screening" combined with "plant name." The searched articles were evaluated to find extract type, plant parts, and activity or diseases. A total of 1,156 research articles related to our listed plants were found.

2.2.3. Compound Collection

The research articles on selected plants were used to find all the chemical constituents. To begin with, we looked for articles on chemical constituents from Google scholar, ACS, Science Direct, and PubMed databases using keywords, such as "plant name," "chemical constituents," "isolated compounds," "phytochemical constituents," "GC-MS," and "HPLC-NMR" combined with "plant name." From selected articles, we collected and enlisted all the chemical constituents and other information, like compound type and compound synonyms. A total of 11, 537 compounds were reported from 145 plants. These data were obtained from 1, 464 articles published in several reputed journals.

2.2.4. Compound Cross-Match

The listed compounds were found in PubChem (https://pubchem.ncbi.nlm.nih.gov/), FooDB (https://foodb.ca/), Chemspider (http://www.chemspider.com/), Spectrabase (https://spectrabase.com/) and Chembl (https://www. ebi.ac.uk/chembl/) databases using the compound name or compound synonyms, and their InChI and InChI key were collected. Certain compounds were not found in these databases, but their structures were provided in the respective literature. The structures of such compounds were drawn by the Chemdraw software, and the 2D or 3D structures were converted into the InChI key and InChI, respectively. The generated InChI keys were then used to find these compounds from previous databases. Finally, duplicate compounds were removed based on InChI key strings and all the unique compounds were listed by a unique identifier (GreenMolBD_CID). From a batch of 11,537 compounds, after removing duplicates, we enlisted 6,837 unique compounds for further analysis.

2.2.5. Compound Descriptors

All compounds have been classified using the ClassyFire web server [33]. Some of these compounds were not found in ClassyFire, and we classified those compounds according to the articles. The 2D structure of each compound was used to generate molecular weight, molecular mass, molecular formula, and a table of physical properties, like hydrogen bond acceptor count, hydrogen bond donor count, number of aromatic atoms, number of heavy atoms, number of aromatic bonds, number of rotatable bonds, log octanol/water partition coefficient, molar refractivity, topological polar surface area, mass density, van der Waals surface area, and van der Waals volume by using MOE software [34]. Isomeric SMILES were generated from the ChemDB chemoinformatics portal, [35] and InChI Key, InChI, Canonical SMILES, and IUPAC names were collected either from databases that were linked up or from Chemdraw [36]. The quantum information of the compound, like HOMO, LUMO, dipole moment, electronegativity, hardness, and heat of formation, was analyzed by Schrodinger-2018, using PM3 basis set and semiempirical methods. Compound reactivity table consisting of delocalized lone pair of electrons (DLP), number of reactive functional groups (nRFG) and number of lightly metabolic reaction (nLMR) along with two drug likeliness properties namely, Lipinski's rules of five and Jorgensen's rule of three, were calculated by the QikProp module of Schrodinger-2018. Other drug likeliness properties, like the Goshe filter and Veber filter, were calculated by Drug Likeness Tool (DruLiTo), whereas PAINS1, PAINS2, and PAINS3 were calculated by the Canvas module of Schrodinger-2018. Toxicity properties, like WOE prediction, Ames prediction, DTP prediction, rat oral LD50, rat inhalational LC50, chronic LOAEL, skin irritancy, skin sensitization, ocular irritancy, aerobic biodegradability prediction, Fathead Minnow LC50, and Daphnia EC50 were calculated by the Discovery Studio Client 2016.

2.2.6. Target Collection

To identify the protein targets of our selected natural compounds with available experimentally-determined data, we searched the PubMed database (https://www.ncbi.nlm.nih.gov/pubmed/), ACS (https://www.acs.org/content/acs/en.html), Science Direct (https://www.sciencedirect.com/), and Google Scholar (https://scholar.google.com/). We searched all articles studying the protein targets of collected compounds from these sites by using the combination of the keywords "plant name," "compound name," "activity," "target," and "activity of compound name." The literature was

evaluated for finding the target name, organism, gene symbol, half-maximal inhibition concentration IC50, equilibrium inhibition constant Ki, half-maximal effective concentration, and equilibrium binding constant Kd. Furthermore, we matched these data with BindingDB (https:/ /www. bindingdb.org/ bind/ index.jsp) [37] and Pubchem Bioassay (https:/ /www.ncbi.nlm.nih.gov/ pcassay/) [38] to cross-check. Finally, GreenMolBD target table was divided into the following columns, namely GreenMolBD target ID, Target Name, Gene Symbol, Organism, Protein Crystal Complex, Potency (nM), Ki (nM), IC50 (nM), Kd (50), EC50 (nM), Kon (m-1, s-1), koff (s-1), Biological Assay, DOI Article, PMID, PubChem AID.

2.2.7. Vendor Collection

The chemical vendor table was obtained from PubChem by using the InChI key.

3. RESULTS AND DISCUSSION

3.1 Data Content

GreenMolBD, freely accessible at https://greenmolbd.gov.bd, currently contains 145 plants' profiles with their 6, 837 unique compounds, 1, 846 target proteins, and 103 pharmacological activities (**Table S1**).

3.1.1. Plant Profile

A complete profile of each plant was constructed with diverse information, such as plant name, plant synonyms, common names, family name, family synonyms, local availability, worldwide availability, botanical description, common ethnobotanical use, chemical constituents, pharmacological evidence, and references. The pharmacological evidence table was organized with specific plant parts and extract types that were reportedly effective against several diseases. This provides more helpful information for ligand isolation and identification of single compound activity. A bar diagram detailing compound distribution along the various plant parts was designed and named "Compound in plant parts" (Fig. **2**).

3.1.2 Compound Profile

There were 11,537 compounds reported from 145 plants, duplicates were removed based on InChI key strings, resulting in 6837 unique compounds, which were enlisted in the database. These data were found from 1,464 articles published in reputed journals. Among them, 5,958 compounds were linked up with different databases, 5,868 with Pubchem, 63 with Chemspider, 22 with Spectrabase, 4 with FooDB, and 1 with ChEMBL whereas 879 compounds were not found in any database (up to December 31st, 2019). Their structures were generated using ChemDraw according to the literature. Target information of 1,316 compounds and vendors' information of 4,244 compounds were given.

3.1.3. Compound Descriptor

A total of 48 types of *in silico* descriptors are calculated using different computational platforms. There are 8 types of identifying properties in GreenMolBD CID and a reference in a compound webpage (Fig. **3a**). All compounds are classified into different chemical superclasses and classes using the ClassyFire web server [33], including alkaloids and derivatives, benzenoids, flavonoids, flavonoid glycosides, terpenoids, steroids and steroidal glycosides, benzenediols, coumarins and derivatives, indoles and derivatives, coumarinolignans, naphthalenes, lipids and lipid-like molecules, lignans, and related compounds, *etc*. A total of 5,801 compounds are classified by the Classyfire web server (Fig. **3b**). All *in silico* properties, like physical properties, quantum information, compound reactivity, toxicity properties, and drug-likeliness test tables are arranged with 12, 6, 3, 12, and 6 types of properties, respectively (Fig. **3c**). There are 5,715 compounds with m/w <500, 5,360 compounds satisfied Lipinski's rule of five, 5,515 satisfied Jorgensen's rules of three, 3,978 compounds are non-carcinogenic in Weight of Evidence (WOE) carcinogenic prediction, 6,546 compounds are non-mutagenic in Ames mutagenicity prediction, and 1,980 compounds are non-toxic in DTP prediction. Compound synonyms are given for every compound. A total of 2,16,280 compound synonyms are collected for 6,837 compounds (Fig. **3d**). For each compound, a table, entitled "compound in different plants," consisting of a list of different plants having the same compound with plant parts and references (Fig. **3e**) is included. Users will be redirected to individual plants by clicking on a plant name from the table.

3.1.4. Target Information

About 1,846 targets, 1,134 from humans and 712 from other organisms, are correlated with 1,316 compounds. The target table provides target names with the gene symbol, organism name, Protein Crystal Complex ID, and PubChem AID, which helps users to access the bioactive compound. By clicking on a target name, users will find a list of compounds associated with this target (Fig. **3f**).

3.1.5. Chemical Vendors

A table of chemical vendors is given with the maximum available information; 4,244 compounds are available from different purchasable sources.

3.2. Data Search and Browse

GreenMolBD entries can be searched on the home page (Fig. **4a**) by multiple search options, such as plant, compound, target, and disease. In the "Plant" section, terms, such as plant name, synonym, common name, family name, and family synonym can be used to search for a plant. A list of relevant plants will appear upon searching. Clicking on a specific plant name leads to the plant profile webpage. From the plant profile webpage, we can access the compounds page by clicking on a specific compound name or downloading the SDF of the list (Fig. **4b**).

Users can query a compound by name, synonym, IUPAC name, and InChI Key. This will result in a list of compounds or, in the case of an exact match, one specific compound. Users can access the individual compound profile page by clicking on a compound name or can download the SDF of the list (Fig. **4c**). A list of compounds will also appear upon entering specific target names on the target search panel. All these compounds are linked to their respective profiles with the option to download (Fig. **4d**). The search results of a disease name will lead to a list of plants, for which different

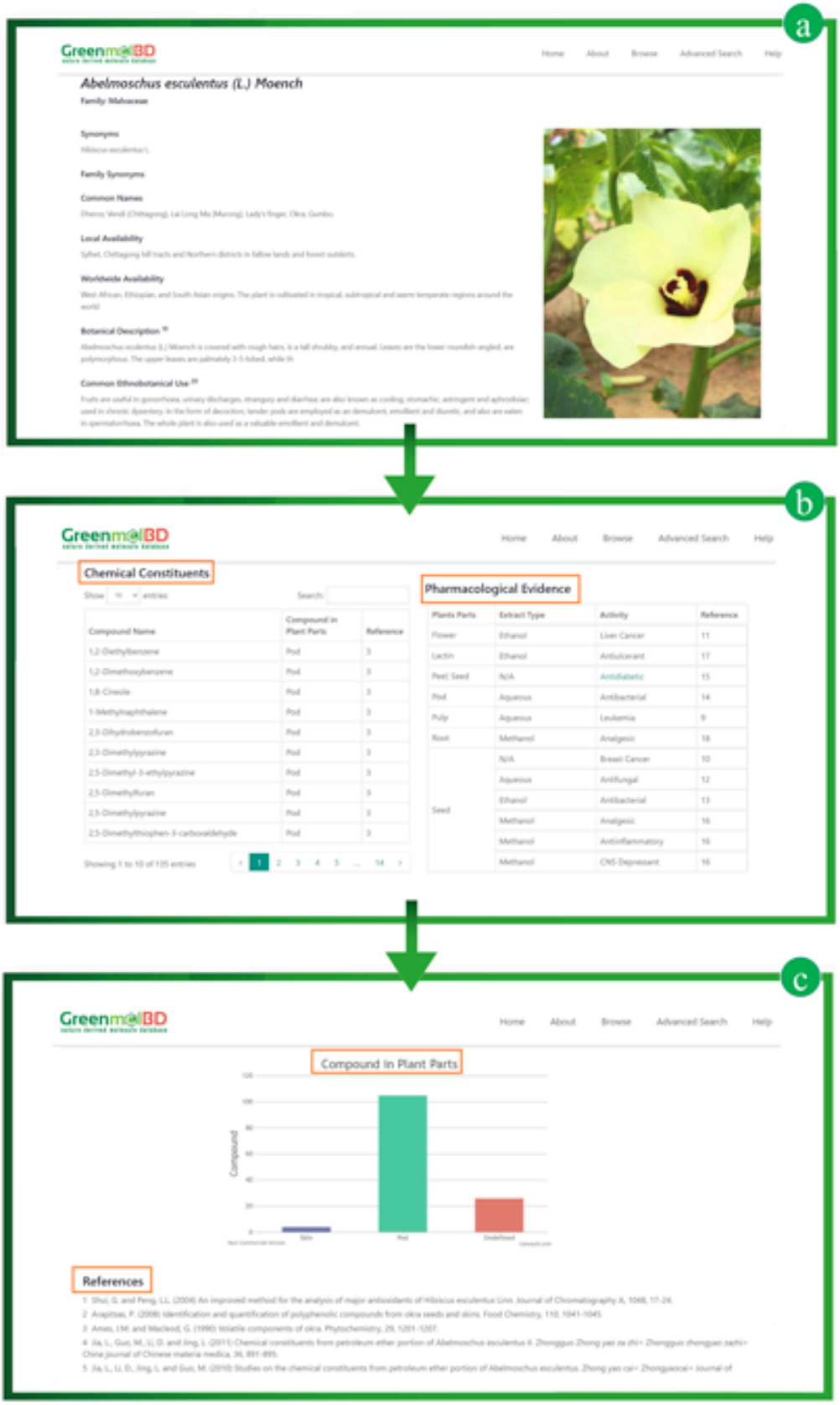

Fig. (2). Ingredients of individual plants **a**) Plant profile with common ethnobotanical use, **b**) List of chemical constituents with plant parts and references; pharmacological evidence table describing activities of plant parts on the basis of extract type *(A higher resolution / colour version of this figure is available in the electronic copy of the article).*

Fig. (3). Specific compound profile and its ingredients, **a**) 2D description (*e. g.* molecular formula, weight, IUPAC name, InChI, InChI Key, and SMILES) of a specific compound with cross-reference, **b**) Classification of individual compounds by Classyfire web server. **c**) 3D information (*e. g.* Physical properties, Quantum information, compound reactivity, toxicity properties, and drug-likeliness test) of specific compounds, **d**) Synonyms of a specific compound, **e**) Table of different plants having the same compound, **f**) Target information regarding specific compound Target list with a gene symbol, organism name, and reference *(A higher resolution / colour version of this figure is available in the electronic copy of the article)*

parts' extracts are explored and reported for query disease by pharmacological screening (Fig. **4e**).

In the browse panel, compounds can be browsed using several filters, such as classification, molecular weight, exact molecular mass, drug-likeness test properties, quantum information (dipole moment), and TOPKAT properties. Compounds, plants, or targets may also be browsed alphabetically (Fig. **5**).

The structure of chemical constituents can be downloaded from the compound profile web page as a single file and batch file from the browsing list. GreenMolBD provides 4 types of structures, such as Image, 2D, 3D (Molecular mechanics), and 3D (Semi-empirical).

3.3. Advanced Search

The incorporated advanced search option enables the user to select compounds based on their physicochemical properties, like molecular weight, XLogP, topological polar surface area, number of H-bond acceptor or donor, *etc*., quantum information, such as HOMO, LUMO, dipole moment, *etc*, drug-likeness test, including the rule of five, rule of three, Goshe filter and Veber filter. GreenMolBD also includes the "search by drawing" option in an advanced search (Fig. **6**).

CONCLUSION

GreenMolBD is the first database in Bangladesh attempting to collect and correlate all the phytochemical and

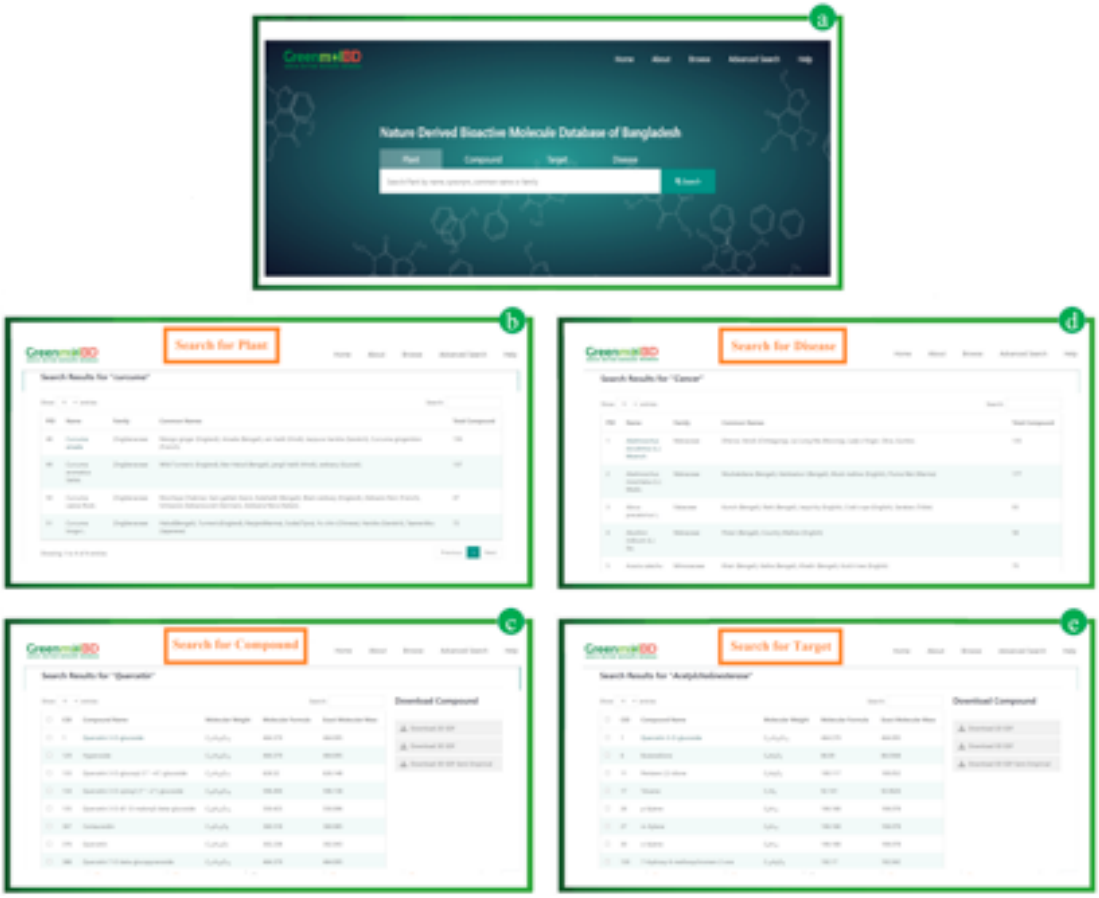

Fig. (4). Search by plant and schematic workflow of GreenMolBD, **a**) Home page selecting plant option in the search bar, **b**) Search results for plant search, showing a list of plants which matched with a searched hint, **c**) Showing the results of the searched compounds and their download bar, **d**) Showing list of compounds related to the searched target name, and **e**) Showing list of plants which have pharmacological evidence against the searched disease name *(A higher resolution / colour version of this figure is available in the electronic copy of the article).*

Fig (5). Browse panel, **a**) Browse by plants panel arranged by alphabetical order showing bars, **b**) Browse by compounds panel where all compounds are arranged by compound name, molecular weight, exact molecular mass, drug-likeliness test, quantum information, and toxicity properties, **c**) Browse by targets, where all targets ordered alphabetically, and **d**) Browse by activity, where the user will find the list of plant against the specific disease *(A higher resolution / colour version of this figure is available in the electronic copy of the article).*

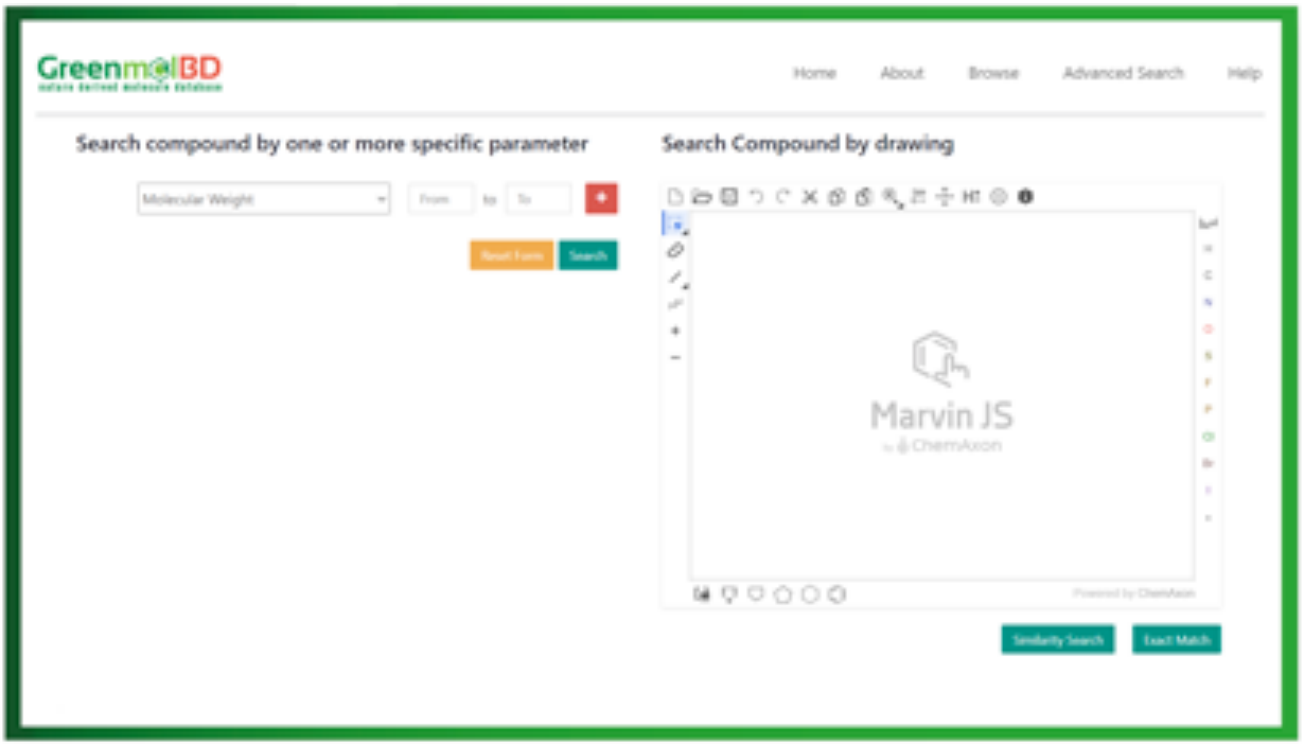

Fig. (6). Advanced search panel, where user can search compounds by 1 to 3 parameters (2D and 3D information of compounds) or by drawing a structure of the compound *(A higher resolution / colour version of this figure is available in the electronic copy of the article).*

pharmacological information regarding Bangladeshi plants and other organisms. In this phase, we provided 6, 837 unique, hit, and lead-like medicinal plants derived compounds with their reported 1,846 target proteins of several diseases. Along with individual plant profiles, we reviewed and reported *in vivo* and *in vitro* study data of plant parts with their extract type and pharmacological activity. Then, we enlisted all the reported, isolated compounds from individual plants along with the information regarding Plant parts from which they were isolated. We successfully designed an individual compound profile with 48 types of descriptors, based on which compound could be categorized and selected for further hit and lead optimization and development research. Furthermore, we also correlated all the associated targets of these compounds which have been reported in other studies as drug targets. Our data portal is consciously updating, and it will include all the medicinal plants and other organisms found in Bangladesh. We will also include an individual profile of targets with biological pathways for all the reported compounds. This rich and integrated collection of the phytochemical, pharmacological, and molecular descriptors information regarding bioactive ingredients in the GreenMolBD database will facilitate the discovery of natural products with therapeutic benefits.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

HUMAN AND ANIMAL RIGHTS

No Animals/Humans were used for studies that are base of this research.

CONSENT FOR PUBLICATION

Not applicable.

AVAILABILITY OF DATA AND MATERIALS

Not applicable.

FUNDING

This work was supported by the grant from the Bangladesh Council of Scientific and Industrial Research (https://mmddl.ctgbcsir.gov.bd/research-grants) program “Database development of nature-derived molecules for rational drug discovery in Bangladesh,” the financial year 2017-18 (Project.No.23, Reference No: 39.02.0000.11.014.007.2017/1358) by the ministry of science and technology, People’s Republic of Bangladesh.

CONFLICT OF INTEREST

The authors declare no conflict of interest, financial or otherwise.

ACKNOWLEDGEMENTS

The authors would like to thank the professional proofreader and grammatical support team of BCSIR, all their colleagues, friends, data survey and mining team, and researchers who provided them with published articles. They are also thankful to the Bangladesh Council of Scientific and Industrial Research for providing the funds to accomplish the first phase and continue with the study.

SUPPLEMENTARY MATERIALS

Supplementary material is available on the publisher's website along with the published article.

REFERENCES

[1] Koehn, F.E.; Carter, G.T. The evolving role of natural products in drug discovery. *Nat. Rev. Drug Discov.,* **2005**, *4*(3), 206-220. http://dx.doi.org/10.1038/nrd1657 PMID: 15729362

[2] Lagunin, A.; Filimonov, D.; Poroikov, V. Multi-targeted natural products evaluation based on biological activity prediction with PASS. *Curr. Pharm. Des.,* **2010**, *16*(15), 1703-1717. http://dx.doi.org/10.2174/138161210791164063 PMID: 20222853

[3] Clardy, J.; Walsh, C. Lessons from natural molecules. *Nature,* **2004**, *432*(7019), 829-837. http://dx.doi.org/10.1038/nature03194 PMID: 15602548

[4] Wang, X.; Wei, X.; Thijssen, B.; Das, J.; Lipkin, S.M.; Yu, H. Three-dimensional reconstruction of protein networks provides insight into human genetic disease. *Nat. Biotechnol.,* **2012**, *30*(2), 159-164. http://dx.doi.org/10.1038/nbt.2106 PMID: 22252508

[5] Vogt, I.; Mestres, J. Drug-target networks. *Mol. Inform.,* **2010**, *29*(1-2), 10-14. http://dx.doi.org/10.1002/minf.200900069 PMID: 27463845

[6] Li, Q.; Li, X.; Li, C.; Chen, L.; Song, J.; Tang, Y.; Xu, X. A network-based multi-target computational estimation scheme for anticoagulant activities of compounds. *Plos One,* **2011**, *6*(3), e14774(1-9). http://dx.doi.org/10.1371/journal.pone.0014774

[7] Yang, K.; Ma, W.; Liang, H.; Ouyang, Q.; Tang, C.; Lai, L. Dynamic simulations on the arachidonic acid metabolic network. *Plos Comput. Biol.,* **2007**, *3*(3), e55(523-530). http://dx.doi.org/10.1371/journal.pcbi.0030055

[8] Zhang, Y.; Thiele, I.; Weekes, D.; Li, Z.; Jaroszewski, L.; Ginalski, K.; Deacon, A.M.; Wooley, J.; Lesley, S.A.; Wilson, I.A.; Palsson, B.; Osterman, A.; Godzik, A. Three-dimensional structural view of the central metabolic network of *Thermotoga maritima. Science,* **2009**, *325*(5947), 1544-1549. http://dx.doi.org/10.1126/science.1174671 PMID: 19762644

[9] Feher, M.; Schmidt, J.M. Property distributions: Differences between drugs, natural products, and molecules from combinatorial chemistry. *J. Chem. Inf. Comput. Sci.,* **2003**, *43*(1), 218-227. http://dx.doi.org/10.1021/ci0200467 PMID: 12546556

[10] Hong, J. Role of natural product diversity in chemical biology. *Curr. Opin. Chem. Biol.,* **2011**, *15*(3), 350-354. http://dx.doi.org/10.1016/j.cbpa.2011.03.004 PMID: 21489856

[11] Firn, R.D.; Jones, C.G. Natural products-a simple model to explain chemical diversity. *Nat. Prod. Rep.,* **2003**, *20*(4), 382-391. http://dx.doi.org/10.1039/b208815k PMID: 12964834

[12] Basso, L.A.; da Silva, L.H.; Fett-Neto, A.G.; de Azevedo, W.F., Jr; Moreira, Ide.S.; Palma, M.S.; Calixto, J.B.; Astolfi Filho, S.; dos Santos, R.R.; Soares, M.B.P.; Santos, D.S. The use of biodiversity as source of new chemical entities against defined molecular targets for treatment of malaria, tuberculosis, and T-cell mediated diseases--a review. *Mem. Inst. Oswaldo Cruz,* **2005**, *100*(6), 475-506. http://dx.doi.org/10.1590/S0074-02762005000600001 PMID: 16302058

[13] Quinn, R.J.; Carroll, A.R.; Pham, N.B.; Baron, P.; Palframan, M.E.; Suraweera, L.; Pierens, G.K.; Muresan, S. Developing a drug-like natural product library. *J. Nat. Prod.,* **2008**, *71*(3), 464-468. http://dx.doi.org/10.1021/np070526y PMID: 18257534

[14] Dobson, C.M. Chemical space and biology. *Nature,* **2004**, *432*(7019), 824-828. http://dx.doi.org/10.1038/nature03192 PMID: 15602547

[15] Rosén, J.; Gottfries, J.; Muresan, S.; Backlund, A.; Oprea, T.I. Novel chemical space exploration *via* natural products. *J. Med. Chem.,* **2009**, *52*(7), 1953-1962. http://dx.doi.org/10.1021/jm801514w PMID: 19265440

[16] Yongye, A.B.; Waddell, J.; Medina-Franco, J.L. Molecular scaffold analysis of natural products databases in the public domain. *Chem. Biol. Drug Des.,* **2012**, *80*(5), 717-724. http://dx.doi.org/10.1111/cbdd.12011 PMID: 22863071

[17] Grabowski, K.; Baringhaus, K-H.; Schneider, G. Scaffold diversity of natural products: Inspiration for combinatorial library design. *Nat. Prod. Rep.,* **2008**, *25*(5), 892-904. http://dx.doi.org/10.1039/b715668p PMID: 18820757

[18] Lee, M.L.; Schneider, G. Scaffold architecture and pharmacophoric properties of natural products and trade drugs: Application in the design of natural product-based combinatorial libraries. *J. Comb. Chem.,* **2001**, *3*(3), 284-289. http://dx.doi.org/10.1021/cc000097l PMID: 11350252

[19] Grabowski, K.; Schneider, G.J.C.C.B. Properties and architecture of drugs and natural products revisited. *Curr. Chem. Biol.,* **2007**, *1*(1), 115-127.

[20] Henkel, T.; Brunne, R.M.; Müller, H.; Reichel, F. Statistical investigation into the structural complementarity of natural products and synthetic compounds. *Angew. Chem. Int. Ed. Engl.,* **1999**, *38*(5), 643-647. http://dx.doi.org/10.1002/(SICI)1521-3773(19990301)38:5643::AID-ANIE6433.0.CO;2-G PMID: 29711552

[21] Mohanraj, K.; Karthikeyan, B.S.; Vivek-Ananth, R.P.; Chand, R.P.B.; Aparna, S.R.; Mangalapandi, P.; Samal, A. IMPPAT: A curated database of Indian medicinal plants, phytochemistry and therapeutics. *Sci. Rep.,* **2018**, *8*(1), 4329. http://dx.doi.org/10.1038/s41598-018-22631-z PMID: 29531263

[22] Nakamura, Y.; Mochamad Afendi, F.; Kawsar Parvin, A.; Ono, N.; Tanaka, K.; Hirai Morita, A.; Sato, T.; Sugiura, T.; Altaf-Ul-Amin, M.; Kanaya, S. KNApSAcK metabolite activity database for retrieving the relationships between metabolites and biological activities. *Plant Cell Physiol.,* **2014**, *55*(1), e7(1-9). http://dx.doi.org/10.1093/pcp/pct176

[23] Valli, M.; dos Santos, R.N.; Figueira, L.D.; Nakajima, C.H.; Castro-Gamboa, I.; Andricopulo, A.D.; Bolzani, V.S. Development of a natural products database from the biodiversity of Brazil. *J. Nat. Prod.,* **2013**, *76*(3), 439-444. http://dx.doi.org/10.1021/np3006875 PMID: 23330984

[24] Liu, Z.; Guo, F.; Wang, Y.; Li, C.; Zhang, X.; Li, H.; Diao, L.; Gu, J.; Wang, W.; Li, D.; He, F. BATMAN-TCM: A bioinformatics analysis tool for molecular mechANism of traditional Chinese medicine. *Sci. Rep.,* **2016**, *6*(21146), 21146. http://dx.doi.org/10.1038/srep21146 PMID: 26879404

[25] Chen, C.Y.C. TCM Database@Taiwan:the world's largest traditional Chinese medicine database for drug screening *in silico*. *Plos One.,* **2011**, *6*(1), e15939(1-5).

[26] Kim, S-K.; Nam, S.; Jang, H.; Kim, A.; Lee, J-J.T.M-M.C. TM-MC: a database of medicinal materials and chemical compounds in Northeast Asian traditional medicine. *BMC Complement. Altern. Med.,* **2015**, *15*(218), 218. http://dx.doi.org/10.1186/s12906-015-0758-5 PMID: 26156871

[27] Ru, J.; Li, P.; Wang, J.; Zhou, W.; Li, B.; Huang, C.; Li, P.; Guo, Z.; Tao, W.; Yang, Y.; Xu, X.; Li, Y.; Wang, Y.; Yang, L. TCMSP: A database of systems pharmacology for drug discovery from herbal medicines. *J. Cheminform.,* **2014**, *6*(13), 13. http://dx.doi.org/10.1186/1758-2946-6-13 PMID: 24735618

[28] Huang, L.; Xie, D.; Yu, Y.; Liu, H.; Shi, Y.; Shi, T.; Wen, C. TCMID 2.0: A comprehensive resource for TCM. *Nucleic Acids Res.,* **2018**, *46*(D1), D1117-D1120. http://dx.doi.org/10.1093/nar/gkx1028 PMID: 29106634

[29] Tung, C.W.; Lin, Y.C.; Chang, H.S.; Wang, C.C.; Chen, I.S.; Jheng, J.L.; Li, J.H. TIPdb-3D: The three-dimensional structure database of phytochemicals from Taiwan indigenous plants. *Database (Oxford),* **2014**, *2014*, 1-5. http://dx.doi.org/10.1093/database/bau055 PMID: 24930145

[30] Chen, X.; Zhou, H.; Liu, Y.B.; Wang, J.F.; Li, H.; Ung, C.Y.; Han, L.Y.; Cao, Z.W.; Chen, Y.Z. Database of traditional Chinese medicine and its application to studies of mechanism and to prescription validation. *Br. J. Pharmacol.,* **2006**, *149*(8), 1092-1103. http://dx.doi.org/10.1038/sj.bjp.0706945 PMID: 17088869

[31] Liu, Z.; Du, J.; Yan, X.; Zhong, J.; Cui, L.; Lin, J.; Zeng, L.; Ding, P.; Chen, P.; Zhou, X.; Zhou, H.; Gu, Q.; Xu, J. TCMAnalyzer: A Chemo-and bioinformatics web service for analyzing traditional Chinese medicine. *J. Chem. Inf. Model.,* **2018**, *58*(3), 550-555. http://dx.doi.org/10.1021/acs.jcim.7b00549 PMID: 29420025

[32] Ghani, A. *Medicinal plants of Bangladesh: Chemical constituents and uses*; Asiatic society of Bangladesh: Dhaka, **1998**.

[33] Feunang, Y.D.; Eisner, R.; Knox, C.; Chepelev, L.; Hastings, J.; Owen, G.; Fahy, E.; Steinbeck, C.; Subramanian, S.; Bolton, E. ClassyFire: automated chemical classification with a comprehensive, computable taxonomy. *J. Cheminformatics.,* **2016**, *8*(61), 1-20.

[34] Vilar, S.; Cozza, G.; Moro, S. Medicinal chemistry and the Molecular Operating Environment (MOE): Application of QSAR and

molecular docking to drug discovery. *Curr. Top. Med. Chem.,* **2008**, *8*(18), 1555-1572. http://dx.doi.org/10.2174/156802608786786624 PMID: 19075767

[35] Chen, J.; Swamidass, S.J.; Dou, Y.; Bruand, J.; Baldi, P. ChemDB: A public database of small molecules and related chemoinformatics resources. *Bioinformatics,* **2005**, *21*(22), 4133-4139. http://dx.doi.org/10.1093/bioinformatics/bti683 PMID: 16174682

[36] Mills, N. ChemDraw Ultra 10.0 CambridgeSoft, 100 Cambridge-Park Drive, Cambridge, MA 02140. www.cambridgesoft.com. Commercial Price: $1910 for download, $2150 for CD-ROM; Academic Price: $710 for download, $800 for CD-ROM. *J. Am. Chem. Soc.,* **2006**, *128*(41), 13649-13650.

[37] Liu, T.; Lin, Y.; Wen, X.; Jorissen, R.N.; Gilson, M.K. BindingDB: A web-accessible database of experimentally determined protein-ligand binding affinities. *Nucleic Acids Res.,* **2007**, *35*(Database issue)(Suppl. 1), D198-D201. http://dx.doi.org/10.1093/nar/gkl999 PMID: 17145705

[38] Wang, Y.; Xiao, J.; Suzek, T.O.; Zhang, J.; Wang, J.; Zhou, Z.; Han, L.; Karapetyan, K.; Dracheva, S.; Shoemaker, B.A.; Bolton, E.; Gindulyte, A.; Bryant, S.H. PubChem's BioAssay database. *Nucleic Acids Res.*, **2012**, *40*, D400-D412. http://dx.doi.org/10.1093/nar/gkr1132 PMID: 22140110