

## Preface

The English edition *Handbook of Active Marine Natural Products (HAMNP) with 8 Volumes* is a selective version of the Marine Natural Products Dataset. The whole dataset was collected and developed by the Molecular Design Group, Institute of Process Engineering, Chinese Academy of Sciences during 1998–2016. Totally, it covers 19,722 entries of secondary metabolites from marine living things, where 8,350 compound entries have pharmacological activity data. The 8,350 compound entries were arranged into eight volumes to form the set of handbooks as follows:

*Volume 1: Terpenoids, Part 1*

*Volume 2: Terpenoids, Part 2*

*Volume 3: Alkaloids, Part 1*

*Volume 4: Alkaloids, Part 2*

*Volume 5: Polyketides and Steroids*

*Volume 6: Aliphatic Metabolites*

*Volume 7: O-Heterocycles and Aromatics*

*Volume 8: Peptides and Others*

This set of eight HAMNP books gathers the structure, origin, and bioactivity, as well as other relevant information, of 8,350 active marine natural products from 3,025 marine organisms.

The HAMNP handbooks represent a largest collection of active secondary metabolites from marine organisms, and all kinds of scientific data have been reorganized as well-formatted data so that the books became helpful to researchers as a convenient reference. The materials covered in these books include those through systematic collection up to 2012, and further accompanied with the latest data published in several core journals until 2016.

The work covered in these HAMNP books was accomplished in two phases. The initial phase ranged from 1998 to 2001 and the main phase from 2011 to 2018. In the original version of the dataset, more than 22,000 compounds have been collected, including duplicated compounds from different authors. The comprehensive data compilation process include data specification definition, cross-validation, assessment confirmation, identification of duplicated structures, and merging of relevant information, leading to the final accomplishment of the current 19,722 datasets.

In brief, the main compilation process of the HAMNP books is given as follows. First, collect the name list, origin, and structure of chemical compounds from successive annual reviews (see Core References R01 and R02 in Introduction) and literature reviews. Second, double-check the documents to verify and complete other information. Third, confirm the structural information and other types of data using orthogonal information from other sources with cross-validation methods. Fourth, the structures of more than 22,000 compounds are rechecked, and the information is integrated by manual identification and computer programs. Finally, the comprehensive information

on the 19,722 compounds constitutes the dataset. Here, 8,350 active sets were picked up from the dataset to form the current HAMNP handbooks.

Three problems need to be solved to compile a multidisciplinary reference book. First, every definition and concept should be explicit when expanding knowledge, connotation, and extension included, without any research details. Second, the reliability assessment is essential for all kinds of data, because the devil is in the detail. Third, it is essential to search, identify, and integrate data of duplicated chemical compounds. Fortunately, well-developed software packages can help us automatically identify the majority of duplicated chemical compounds. The remaining issues can be resolved along with manual processing.

It is the guiding principle of the author to make the book to be pithy, thorough, precise, and intelligible. In fact, we always view ourselves as HAMNP's readers, with the exclusive objective to let readers gain the most useful knowledge in the shortest possible time.

The core contents and highlights of the HAMNP books are the “three diversities,” that is, the diversity of chemical structures, the diversity of biological resources, and the diversity of pharmacological activities. In terms of chemical structure diversity, we refer to the classification system from references, then further improve and expand it based on the latest research and development to define our classification framework of structures. Once readers browse the contents of the books, the classification system is straightforward. For the diversity of biological resources, it is recommended to refer to Index 3 in each volume – Compound Marine Organism Source Index; and Index 4 in each volume – Compound Marine Source Sampling Geographic Location Index. For the diversity of pharmacological activities, it is recommended to refer to Index 5 in each volume – Compound Pharmacological Activity Index.

These HAMNP handbooks are expected to help readers who are engaged in research, in teaching, and in the development of marine natural products. It should also benefit college students, postgraduates, marine resource managers, and those who are interested in the chemistry and pharmacology of marine natural products. We would feel fortunate if it works as expected.



Jiaju Zhou  
Institute of Process Engineering (IPE),  
Chinese Academy of Sciences (CAS)  
February 2019