Chapter 53

Chemical Samples Recycling: The MDPI Samples Preservation and Exchange Project

S.-K. Lin

Abstract

Over the years, chemists have been known to make a significant contribution to science, and to generate new chemical substances as a result of their professional activities. Over 20 million compounds have been recorded in the literature, however, because of their use in chemical processes and their disposal in waste products, only a smaller number of these compounds are available for furture use. Therefore, there is a tremendous loss of molecular diversity and chemical heritage. Reproduction of samples, if successful, is expensive and time-consuming. It may also contribute to a substantial damage to the environment. Chemical sample archives are priceless resources that can tremendously facilitate and speed-up discovery of new drugs, and new compounds, as well as the development of many other chemical products, with industrial, pharmaceutical, agricultural, educational, and research applications. The Molecular Diversity Preservation International (MDPI) project has been successful in carrying out worldwide collections/deposits, exchanges, and "recycling" of a significant number of chemical samples.

53.1 Introduction

Since 1995, the collection and high throughput production of chemicals of high molecular diversity have been significantly increased, to cope with the crucial problem associated with chemical synthesis of new drugs and other chemical products needed in research and development. The related topic of molecular diversity has been so important that a journal *Molecular Diversity* (www.mdpi.org/modi) was launched (Lin 2003). In 1995, a nonprofit international organization, *Molecular Diversity Preservation International* (MDPI, www.mdpi.org) was founded in Switzerland, and started a worldwide repository and exchange of chemical samples as well as experimental data related to samples preparation and characterization. A chemistry journal, *Molecules* (www.mdpi.org/molecules) was launched by MDPI in 1996, with a unique editorial policy regarding the promotion these sample collection activities (Lin 1996a).

So far, we have already accepted for storage and registered more than 20 000 available chemical compound samples. *Molecules* is publishing its 8th volume in 2003. Many samples are supplied by the authors of *Molecules* and deposited as authentic reference samples.

53.2 Strategies and Activities of MDPI

53.2.1 Preserving Chemical Samples

The disposal of old chemical samples can be very expensive, and can also cause environmental damage. Recent studies have also demonstrated that environmental contamination by these chemicals can lead to a significant number of adverse health effects including various types of cancer in humans (Tchounwou 2003). For this and several other reasons, a new online periodical, *International Journal of Environmental Research and Public Health* (ISSN 1660-4601), has been launched (www.mdpi.net/ijerph).

Preservation of certain rare samples and their recycling is not only possible but also practically necessary. Similar to biospecies (WIPO 1994), samples of chemical species should be properly preserved due to many unique reasons. Chemical samples may decay. For this concern some scientists may not support the preservation of chemical samples. The preservation of biospecies also faced the sample stability problems. However, if some necessary measures are taken, the preservation of microorganisms (WIPO 1994) and chemical samples can be achieved. Regarding chemical samples, of course we do not collect the known unstable samples and known hazardous samples. Only adequately purified samples and carefully bottled samples are collected (e.g. if a compound is sensitive to light but otherwise very stable, this sample can be collected if it has been in a dark bottle). All sample vials should be properly sealed to prevent possible reaction with oxygen and moisture in the air. Stable samples, if carefully stored, can be preserved for many years. In pharmaceutical companies, almost all those properly archived organic samples prepared more than fifty years ago and stored even at room temperature so far have been found to be applicable and all can be used for such purposes as high throughput screening and useful for many

Box 53.1. Some statistics (some of this data is the present author's estimation)

- The average cost per sample of ordinary structural complexity prepared was estimated at around 7 500 USD (Xu and Hagler 2002).
- About 1/4 to 1/3 of a synthetic organic chemist's time is spent for preparing compounds according to literature. Samples are expensive, even prepared according to the literature, given that only 50 compounds are synthesized per chemist per year.
- In principle, the samples of published molecules can be reproduced. However, it is certain that many samples will never be reproduced if doing so is too expensive (e.g. it takes more than 30 steps), or may never be reproduced because of the incomplete record of experimental data (see Sect. 53.3).
- So far, more than 200 000 000 chemicals have been recorded according to Chemical Abstracts (see: www.cas.org/substance.html). There are 8 000 000 organic compounds recorded in Beilstein database (http://www.beilstein.com/products/xfire/). However, the number of chemicals commercially/readily available is much less than 100 000 before 1995, according to ACD (Available Chemicals Directory) database release. More than 99% of compounds recorded in literature exist only on paper; chemists discarded them.

other purposes. It is certain that most chemical samples, particularly organic compound samples can be stored for many years.

Many chemists, even many synthetic organic chemists, may insist that the publication of research papers is all they wanted. Traditionally a chemical synthesis is carried out for a target compound. Consequently no precursors prepared as intermediates are regarded as worthy to be preserved. Sometimes a compound is prepared for very specific academic and industrial purposes such as spectroscopic measurements or enzyme-ligand binding tests. Afterwards the samples may be regarded as useless and have to be discarded. In many cases such as the termination of a research project, graduation of a student, and retirement of a chemist or a professor, all the pertinent compounds accumulated may have to be discarded (Box 53.1 and 53.2). This is the loss of precious molecular diversity.

Molecular diversity has been a topic of increasing interest since 1995 (Waller 2002; Xu and Hagler 2002). Techniques of combinatorial chemical libraries have been developed to provide millions of compounds for pharmaceutical studies. The similarity and the gradual variation in the structures of chemical species in combinatorial chemical libraries permit adjustments designed to optimize structure-activity relationships. However, the high quality of a chemical library relies on the distinct differences of both the structures and properties of the collected samples (Lin 1996b,c). These high diversity compounds are isolated from natural sources, and/or traditionally or routinely prepared in the laboratories (Lin 1996a,b, 1997a,b, 1998).

Now, a rare sample as little as 1 mg can be used for dozens or even hundreds of bioactivity tests. Particularly in recent years, with the development of high throughput screening technology, the acquisition of chemical samples, even though the amount might be small, has become a bottleneck in the discovery process of new drugs and other chemical products. Absorption, distribution, metabolism, excretion, and toxicity (ADMET) studies of drugs in animal models can be performed before the active compound can be identified (Xu and Hagler 2002). Such studies may need large amounts of chemical samples. Many rare samples of smaller amounts should also be preserved for the purpose of pharmaceutical research.

Box 53.2. Several cases of samples loss

- *Case 1*. Summer 1994. The final work of a retired chemist in a pharmaceutical company (Ciba-Geigy) is to dissolve all of his 2 000 samples into acetone, which took him and his assistant for one week.
- *Case 2*. October 1995. A recently retired professor called the present author from Denmark and told him that he just threw away more than 2 000 samples because his successor needs the laboratory space and his wife did not permit him to move the sample home and place in the basement.
- *Case 3*. April 1997. One third of the samples had structural lists destroyed and ready to be trashed when we arrived in the chemistry department storage room. We were able to register 2/3 of all the 3 000 old samples properly. MDPI still accepted the rest 1/3 samples and gave separate registration numbers (they will be distributed for random screenings. Active hits will be analyzed and the structures will be elucidated afterwards).

Fortunately most synthetic organic chemists and natural product chemists have a tradition to store their chemical samples. They would like to preserve them for other people to use in future. However, there is usually no room for them to store their samples, after retirement. MDPI has provided the service to store these samples, as well as to develop and maintain a database. Therefore the collection of isolated natural products and individual rare samples of synthetic source have been MDPI's main service. Box 53.3 lists the criteria for sample registration and deposition.

In order to encourage chemists to deposit their rare samples, the MDPI services are free of charge to its contributors. In addition, MDPI pays postage for delivery by express carrier (TNT, Federal Express, UPS, etc.) of the compounds to MDPI. Contributors are rewarded 50% of the net profit of sample distribution services. Samples are usually divided into at least 5 unit amounts. The last remaining part is further divided into several parts. A typical unit price per unit amount is 100 USD/100 mg, and may vary based on user's recommendation. MDPI distributes compound samples worldwide. Chemists may also store compounds themselves for the registered samples, also free of charge, provided that they send the unit amount of the ordered compound promptly. They may distribute directly their compounds by citing MDPI registration number.

In order to provide services fare to contributors and to customers, and insure that the samples are correctly characterized, we publish unit amount and the unit price and preferably contributor's name and address in our database release. If the contributors agree, other chemists may obtain samples free of charge for academic research. Now all the samples can be searchable on the www.molmall.org server. Most of the structures of MDPI samples not published in the literature are given CAS Reg. numbers and the MDPI collection is included in CHEMCATS – Online Chemical Catalogs (www.cas.org/CASFILES/chemcats.html).

Many samples have historical significance and related to chemical heritage. For example samples prepared by Nobel laureates can be preserved forever. A chemical museum will be built for this purpose.

Unlike handling of bulk chemical storage and transportation, the storage, package and transportation of chemical samples in small amount are easier. However, because of their large number and diversity, we will face other problems. The management of chemical sample archives itself has been an important topic of research. MDPI plans to carry out more studies on this topic.

Box 53.3. Samples criteria

- All compounds, whether they are new or old, synthetic or natural products, published or unpublished.
- Authentic supporting samples for papers published in chemistry journals.
- Samples should be reasonably stable at room temperature, as pure as specified, not hazardous and preferably in solid (crystal) form.
- Catalysts and enzymes, macromolecules and polymers as well as other samples are also acceptable.

53.2.2
Preserving Experimental Data

In a published paper, the "experimental design and methods" section is the part of great interest to synthetic chemists. Traditionally, our contribution of chemical information has relied on the full description of chemical preparation and compound characterization. However, more and more chemistry journals ask authors to shorten their papers, mainly the experimental part, or include the research data in the "so-called "supplementary material". This makes the record of the knowledge incomplete and the reproduction of chemical preparation and characterization inconvenient. For example, experimental details of R. B. Woodward's several most important synthetic works have never been properly published (Cornforth 1993). In such cases, if a synthesis is very difficult and if there is no one else as resourceful as the original author, it is almost impossible to prepare the reported samples again.

Generally an ordinary and average synthetic chemist working in industry will prepare and fully characterize about 1 000 samples in his whole carrier (Box 53.2). On average, only a small part of this precious data of chemical synthesis is ever recorded in chemistry journals and patents. Thus, the precious expertise of experienced chemists can seldom be fully recognized, appreciated and benefited by others (Lin 1996a, 1997a,b, 1998). Another urgent task of our organization MDPI is to record all the experimental data as a complementary measure of the normal publication of the existing journals. Any scattered unassembled experimental data for individual compounds, which are conventionally not publishable, are particularly welcomed; to be published in a special form of one molecule per short note and given special page numbers (M1, M2, etc.). A large volume of (to say 1 million) structures will be easily published in this way (see: www.molbank.org). If every synthetic chemist contributes 100 such posters, this number will be easily reached within several years. Afterwards, an online-searchable database constructed from this data will be open to the public free of charge and it will be a very useful treasure to all chemists and other related scientists (Lin and Patiny 2000). The detailed procedure of the chemical reactions can make the reproduction of chemicals easier.

The structures of MDPI samples not published in the literature are given CAS Reg. The publication of the samples database at www.molmall.org is also a pertinent way of recognizing sample contributor's synthetic works, if they were never published before. The information regarding the sample availability becomes an important part of Chemical Abstracts Service (www.cas.org).

53.3
Conclusion

Compared to scientists of other fields of study, chemists have created a vast volume of chemical samples and experimental data, which are definitely unique. We should take care of our complete intellectual properties and act properly to preserve worldwide high-quality molecular diversity of both chemical information and chemical samples generated by us. It is never too late to act. To support the MDPI activities, the

recommendation is proposed: it should be an obligatory part of his/her final duty that a chemist finishing his/her works, and leaving the university or the company, deposits or registers all of his/her samples. He/she should also publish all his/her synthesis and structural characterization data. MDPI has been modifying both procedures for samples and data base development, to make it more and more convenient for chemists to make a significant contribution to chemical preservation, and scientific advancement.

Acknowledgement

The author wishes to thank Dr. Tchounwou for review and English corrections.

References

Cornforth JW (1993) The trouble with synthesis. Aust J Chem 46:157–170
Lin SK (1996a) A good yield and a high standard. Molecules 1:1–2
Lin SK (1996b) Molecular diversity assessment: logarithmic relations of information and species diversity and logarithmic relations of entropy and indistinguishability after rejection of Gibbs paradox of entropy of mixing. Molecules 1:57–67
Lin SK (1996c) Correlation of entropy with similarity and symmetry. J Chem Inf Comp Sci 36:367–376
Lin SK (1997a) Chemical information. Chem Eng News 75(21):4
Lin SK (1997b) Preserving and exploiting molecular diversity: deposit and exchange of chemical information and chemical samples. Molecules 2:1–2
Lin SK (1998) Chemical samples: chemists' and public concern, and preservation. Distributed at the MIT Forum on Chemicals and Society, Cambridge, MA, USA, 11–12 June 1998
Lin SK (2003) Editorial: what is molecular diversity. Mol Diversity 6:1
Lin SK, Patiny L (2000) MolBank: first fully web-based publication of chemical reaction data of individual molecules with structure search and submission. Internet J Chem 3:1
Tchounwou PB (2003) Editorial. Int J Environ Res Public Health 1:1
Waller, CL (2002) Recent advances in molecular diversity. Mol Diversity 5:173–174
World Intellectual Property Organization (WIPO) (1994) Guide to the deposit of microorganisms under the Budapest treaty. WIPO Publication No. 661 (E), Geneva
Xu J, Hagler A (2002) Chemoinformatics and drug discovery. Molecules 7:566–600