別添５

Ⅲ．研究成果の刊行に関する一覧表

書籍

著者名	論文ﾀｲﾄﾙ名	書籍全体の編集者名	書籍名	出版社名	出版地	出版年	ﾍﾟｰｼﾞ
Hayashi, M., and Y. Sakuratani	Development of an Evaluation Support System for Estimating Repeated-Dose Toxicity of Chemicals Based on Chemical Structure.	Wilson, A.G. E.	New Horizons in Predictive Toxicology	RSC Publishing	Croatia, Rijeka	2011	26-37

雑誌

発表者名	論文ﾀｲﾄﾙ名	発表誌名	巻号	ﾍﾟｰｼﾞ	出版年
Mekenyan OG, Petkov P, Kotov S, Stoeva S, Kamenska VB, Dimitrov S, Honma M, Hayashi M, Benigni R, Donner M, Patlewicz GY.	Investigating the relationship between in vitro - in vivo genotoxicity: Derivation of mechanistic QSAR models for in vivo liver genotoxicity and in vivo bone marrow micronucleus formation which encompass metabolism.	*Chem Res Toxicol.*	25(2)	277-296	2012
Matsumoto M, Harada T, Shibuya T, Hamada S, Honma M, and Hirose A.	A chemical category approach of genotoxicity studies for branched alkylphenols.	*Bull, Natl. Inst. Health Sci.,*	129	68-75	2011
Sheila Galloway, Elisabeth Lorge, Marilyn J. Aardema, David Eastmond, Mick Fellows, Bob Heflich, David Kirkland, Dan D. Levy, Anthony Lynch, Daniel Marzin, Takeshi Morita, Maik Schuler, Günter Speit	Workshop summary: Top concentration for in vitro mammalian cell genotoxicity assays; and Report from working group on toxicity measures and top concentration for in vitro cytogenetics assays (chromosome aberrations and micronucleus)	*Mutation Research*	723	77-83	2011
R. Kikura-Hanajiri, M. Kawamura, A. Miyajima, M. Sunouchi, Y. Goda	Chiral analyses of dextromethorphan/levomethorphan and their metabolites in rat and human samples using LC-MS/MS.	*Anal Bioanal Chem*	400	165-174	2011
Takeshi Morita and Kaoru Morikawa	Expert Review for GHS Classification of Chemicals on Health Effects.	*Industrial Health*	49	559-565	2011
T. Morita, M. Honma, K. Morikawa	Effect of reducing the top concentration used in the in vitro chromosomal aberration test in CHL cells on the evaluation of industrial chemical genotoxicity.	*Mutation Research*	741	32-59	2012
T. Morita, J. T. MacGregor and M. Hayashi	Micronucleus assays in rodent tissues other than bone marrow.	*Mutagenesis*	26	223-230	2011

Yoshida M, Takahashi M, Inoue K, Nakae D, Nishikawa A.	Lack of chronic toxicity and carcinogenicity of dietary administrated catechin mixture in Wistar Hannover GALAS rats.	*J Toxicol Sci.*	36	297-311	2011
Ono, A., Takahashi, M., Hirose, A., Kamata, E., Kawamura, T., Yamazaki, T., Sato, K., Yamada, M., Fukumoto, T., Okamura, H., Mirokuji, Y. and Honma, M.	Validation of the (Q)SAR combination approach for mutagenicity prediction of flavor chemicals.	*Food Chem Toxicol*	50	1538 -1546	2012
Hirata-Koizumi, M., Fujii, S., Furukawa, M., Ono, A. and Hirose, A.	Two-generation reproductive toxicity study of aluminium sulfate in rats.	*Reprod Toxicol*	31	219-230	2011
Hirata-Koizumi, M., Fujii, S., Ono, A., Hirose, A., Imai, T., Ogawa, K., Ema, M. and Nishikawa, A.	Repeated dose and reproductive/developmental toxicity of perfluorooctadecanoic acid in rats.	*J Toxicol Sci*	37	63-79	2012
高橋 美加, 松本 真理子, 宮地 繁樹, 菅野 誠一郎, 菅谷 芳雄, 平田 睦子, 小野 敦, 鎌田 栄一,広瀬 明彦	OECD 化学物質対策の動向（第 17 報）－第 28 回 OECD 高生産量化学物質初期評価会議（2009 年パリ）	*化学生物総合管理学会誌*	7	47-54	2011
Kobayashi, K., Y. Sakuratani, T. Abe, K. Yamazaki, S. Nishikawa, J. Yamada, A. hirose, E. Kamata and M. Hayashi	Influence of coefficient of variation in determining significant of quantitative value obtained from 28-day repeated-dose toxicity studies in rats.	*J. Toxicol. Sci.*	36(1)	63-71	2011
林　真、櫻谷祐企	化学物質の安全性　－in silico　評価への挑戦	*構造活性フォーラム 2011*		29-42	2011
Thybaud, V, JT MacGregor, L Müller, R Crebelli, K Dearfield, G Douglas, PB Farmer, E Gocke, M Hayashi, DP Lovell, WK Lutz, D Marzin, M Moore, T Nohmi, DH Phillips, J Van Benthemo	Strategies in case of positive in vivo results in genotoxicity testing.	*Mutation Research*	723	121-128	2011
Hayashi, M, K Dearfield, P Kasper, D Lovell, H-J Martus and V Thybaud	Compilation and use of genetic toxicity historical control data.	*Mutation Research*	723	87-90	2011

Ⅳ．研究成果の刊行物・別刷

RSC Drug Discovery

Edited by Alan G. E. Wilson

New Horizons in Predictive Toxicology

Current Status and Application

RSC Publishing

CHAPTER 3

Development of an Evaluation Support System for Estimating Repeated-Dose Toxicity of Chemicals Based on Chemical Structure

MAKOTO HAYASHI*[,a,b] AND YUKI SAKURATANI[b]

[a] Biosafety Research Center, Foods, Drugs and Pesticides, 582-2, Shioshinden, Iwata, Shizuoka 437-1213, Japan; [b] National Institute of Technology and Evaluation, Chemical Management Center, 2-49-10, Nishihara, Shibuya-ku, Tokyo 151-0066, Japan

3.1 Introduction

In Japan, industrial chemicals are regulated by the Chemical Substances Control Law (CSCL) for the prevention of environmental pollution caused by chemical substances that exhibit a risk of impairing human health or interfering with the inhabitation and/or growth of flora and fauna. This law mandates prior examination of the hazardous properties of new chemical substances which are intended to be manufactured in or imported into Japan.

The safety evaluations of chemicals designated as "existing chemicals" under CSCL have been continuously conducted by the government since implementation of CSCL. As a result, various test data for existing chemicals were

RSC Drug Discovery Series No. 12
New Horizons in Predictive Toxicology: Current Status and Application
Edited by Alan G. E. Wilson

Published by the Royal Society of Chemistry, www.rsc.org

accumulated for certain endpoints such as biodegradation, bioconcentration, genotoxicity, repeated (mainly 28 days) dose toxicity and ecotoxicity. The test data for the existing chemicals under CSCL are disclosed on the websites of public institutions:[1,2]

Based on the test data for new and existing chemicals, the activities for developing evaluation methods by (quantitative) structure–activity relationship [(Q)SAR] and the category approach[3] have been conducted under governmental initiatives.[4–8] These methods are utilized in the prioritization of untested existing chemicals or the preparation of the reference information in the examination of new chemicals.

Repeated dose toxicity (RDT) testing is one of the important endpoints of CSCL. However, a reliable and satisfactory (Q)SAR or category approach method for evaluating the RDT of chemicals still remains undeveloped. Generally, the purpose of RDT is to determine values such as the no-observed-effect level (NOEL), no-observed-adverse-effect level (NOAEL), lowest observed effect level (LOEL), and lowest observed adverse effect level (LOAEL). These values are determined by experts in the field based on test data of various endpoints related to hematology, blood biochemistry, and histopathology. Therefore, it is difficult to directly correlate the NOEL and other parameters with chemical structure since all these parameters are influenced by many toxicological events. In order to make an *in silico* evaluation of RDT for target chemicals from their chemical structure, a comprehensive collaboration of experts possessing a variety of knowledge on the analogue chemicals is necessary.

The project, "Development of hazard assessment techniques using structure–activity relationship methods",[9] sponsored by New Energy and Industrial Technology Development Organization (NEDO) in Japan, aims to develop the "Hazard Evaluation Support System Integrated Platform (HESS)"[10] for providing decision support information to experts to evaluate the RDT of chemicals by the category approach (project period: 2007–2011). This chapter contains an overview of HESS and describes the methodology for evaluating RDT by the category approach developed in this project. The trial version of the system has been developed and will be available in 2012.

3.2 Overview of the Hazard Evaluation Support System Integrated Platform (HESS)

As shown in Figure 3.1, HESS has two main databases. One is the Toxicity Knowledge Information Database containing RDT test reports and toxicity mechanism information. The other is the Metabolism Knowledge Information Database containing metabolic maps and Adsorption, Distribution, Metabolism, and Excretion (ADME) information. HESS has two support tools for estimating the RDT of chemicals by using information from these databases—the Bayesian Net RDT Prediction Model and the Category Approach Support Function. HESS is designed to be compatible with the Organisation for

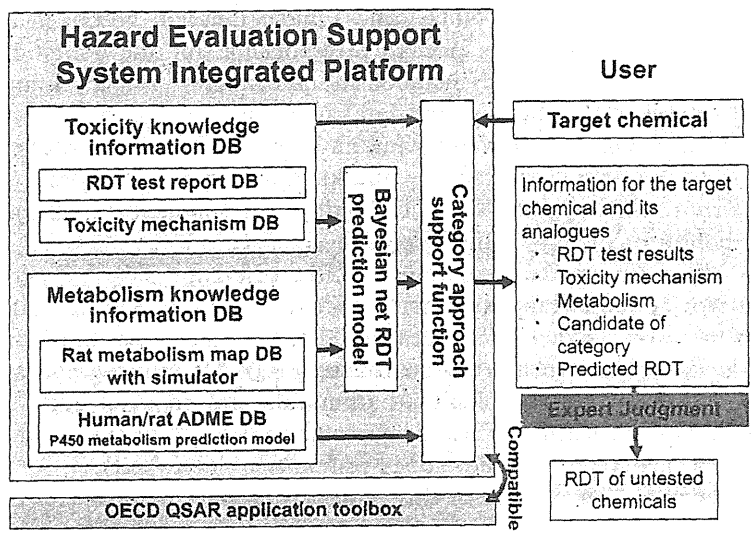


Figure 3.1 Structure of the Hazard Evaluation Support System Integrated Platform (HESS).

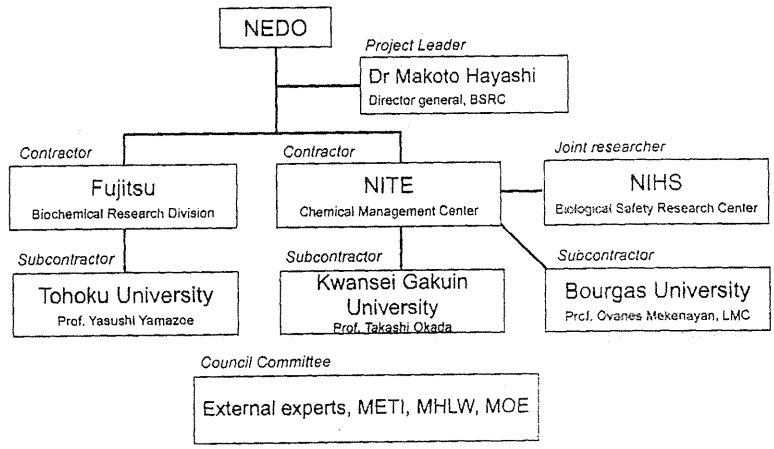


Figure 3.2 Organization of the "Development of hazard assessment techniques using structure–activity relationship methods" project.

Economic Co-operation and Development (OECD) (Q)SAR Application Toolbox.[11]

Figure 3.2 shows the organizational structure of the project. The project leader is Dr Makoto Hayashi who is the Director General of Biosafety Research Center, Foods, Drugs and Pesticides (BSRC) and former Head of the Division of Genetics and Mutagenesis, National Institute of Health Sciences (NIHS). The main contractors of the project are National Institute of

Technology and Evaluation (NITE) and Fujitsu Limited. NITE, in a joint research effort with NIHS, is primarily responsible for the collection and systemization of toxicity and metabolism information. Fujitsu Limited is primarily responsible for building the database systems. A group of experts led by Prof. Takashi Okada of Kwansei Gakuin University and a group of experts led by Prof. Yasushi Yamazoe of Tohoku University are responsible for development of the Bayesian Net RDT Prediction system and development of the metabolism prediction method, respectively.

HESS is not designed to be an "automatic prediction system". Rather, the aim is to provide support for experts such as toxicologists, pathologists and risk assessors in making toxicological decisions. Therefore, it is very important for the system to be easy to use by such experts. As such, the system is being developed primarily by toxicologists, pathologists and risk assessors. This is the most important feature of our project and is largely different from other projects in which systems are developed primarily by IT experts.

The other important feature of this project is its international focus. Since this system is designed to be compatible with the OECD (Q)SAR Application Toolbox and is intended to be used internationally, Professor Menkenyan's team from the Laboratory of Mathematical Chemistry (LMC) of Bourgas Prof. Assen Zlatarov University, responsible for building the OECD (Q)SAR Application Toolbox system, was invited to join the project. Accordingly, the project is coordinating closely with OECD activities.

In addition, recommendations from regulatory authorities are also being taken into consideration in the development of HESS. External experts in relevant fields and members of three ministries [The Ministry of Economy, Trade and Industry (METI); The Ministry of Health, Labour and Welfare (MHLW); and The Ministry of Environment (MOE)] are participating in the supporting committee of the project.

The following is an overview of each part of HESS:

3.2.1 Repeated-Dose Toxicity Test Report Database

This database contains RDT test reports (mainly 28-day RDT using rats with oral administration) for about 500 chemicals. The test data for the existing chemicals under CSCL comprise the main part of the database (about 300 chemicals). All these tests were conducted in compliance with GLP principles under the auspices of the Japanese government (MHLW, METI and NITE). In addition, other RTD test reports with detailed data and high reliability are included in the system (*e.g.*, US National Toxicology Program (NTP) data,[12] data from journal papers, *etc.*).

In this database, test report information is organized and stored in a uniform format in order to facilitate comparison of test data between chemicals. The data for hematological examinations, blood chemical examinations and histopathological examinations are grouped into individual data tables and incorporated into the RDT Test Report Database. In these group tables, the

data showing statistically significant differences from the control group are noted with a mark. However, it is difficult to compare the data among test reports using different statistical analyses methods because the statistically significant differences depend on the analysis method used.[13] Therefore, in this database, data determined to be a toxicological effect by experts are marked differently.

Histopathological terms used in the test reports generally differed between the various laboratories. In order to make searches of histopathological findings more complete and effective, we developed a thesaurus as part of the database search engine.[14]

3.2.2 Toxicity Mechanism Database

The purpose of the database is to provide the rationale for the categorization of toxicity at repeated dose levels. This database contains information from about 220 original papers that suggests the mechanism of the toxicological effects observed in the collected RDT test reports of about 70 chemicals. The toxicity mechanism information covers molecular level mechanisms, cellular level mechanisms and biological level mechanisms and contains *in vitro*, *in vivo* and *ex vivo* test results, signal transfer pathways, *etc.* We expect the database to provide appropriate information for use with the adverse outcome pathways (AOP).

3.2.3 Rat Metabolism Map Database with Metabolic Simulator

This database contains metabolic maps for about 680 chemicals with repeated dose toxicity test reports. All parental chemicals and metabolites in the database can be searched by chemical structure, and the database can be used for investigating metabolites that cause toxicological effects. In addition, a simulator for estimating metabolites in rat liver has been built based on the metabolic maps.

3.2.4 Human/Rat ADME Database (P450 Metabolism Prediction Model)

This database contains information from about 280 original papers related to ADME information on humans and rats for about 80 chemicals with RDT test reports. A model for predicting metabolites of human P450 (CYP2E1, *etc.*) based on the substrate structure was developed. The prediction results for about 100 chemicals with repeated dose toxicity tests are also included in the database. This database can be used for comparing species differences in toxicity between humans and rats based on their different metabolisms.

3.2.5 Bayesian Net RDT Prediction Model

In this model, causality of toxicity is expressed as a network of conditional probabilities (Bayesian network) to predict the probability for a chemical to induce a specified toxicity.[15] Detailed repeated dose toxicity test data (hematological examinations, blood chemical examinations and histopathological examinations) were analyzed by the cascade model,[16] a data mining method to extract structure alerts used in the Bayesian Net RDT Prediction Model.[17] The model was built using the expertise of toxicologists and pathologists.

3.2.6 Category Approach Support Function

Chemicals considered to have the same AOP in RDT tests were categorized based on information in the Toxicity Knowledge Information Database and the Metabolism Knowledge Information Database. The definitions of the categories (structure, parameter, mechanism of action and metabolism boundaries) were entered in the system as a category library of 36 categories.

The system operates in the following way: When a user inputs a target chemical for evaluation, the system searches the category library for possible categories to which the target chemical belongs. Information about the target chemical and its analogues, extracted from the Toxicity Knowledge Information Database and the Metabolism Knowledge Information Database, are displayed in a format that can be easily analyzed by the system user. The user can also investigate the category to which the target chemical belongs and fill in data gaps using data from the analogue chemicals.

3.3 Example of the Category Approach for Evaluating Repeated-Dose Toxicity Tests

It is important to establish a methodology for applying the category approach to complex endpoints in the field of chemical management. OECD is proposing using AOP principles in applying the category approach to complex endpoints.[18] The AOP indicates the mechanistic pathway to the outcome in the test from the molecular initiating event. We also included AOP principles in applying the category approach to repeated dose toxicity tests based on the OECD proposal.[19]

The outcomes observed in repeated dose toxicity tests largely depend on the dose levels tested. They are a complex combination of the findings recognized as the effects of the test chemical. Accordingly, we described an adverse effect as a combination of findings. We also described the mechanism by which these adverse effects are induced from molecular level to *in vivo* level as the AOP for repeated dose toxicity.

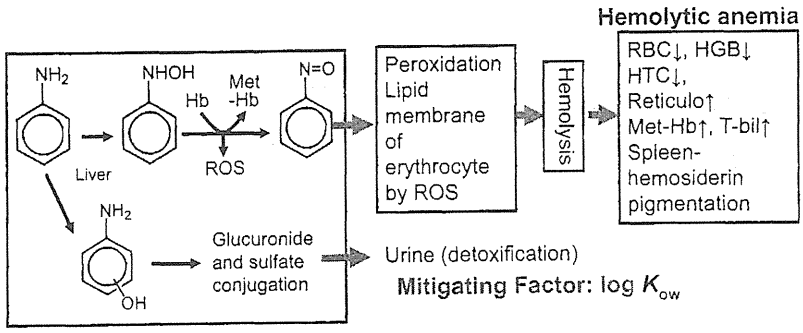


Figure 3.3 AOP for describing the hemolytic anemia induced by anilines.

Figure 3.3 shows the AOP for describing the hemolytic anemia induced by anilines in repeated dose toxicity tests in rats. *N*-hydroxyanilines, metabolites of anilines in the liver, react with hemoglobin to produce methemoglobin (Met-Hb) and reactive oxygen species (ROS). ROS then causes peroxidation of lipids in the membranes of erythrocytes, resulting in hemolytic anemia. When hemolytic anemia is induced, hematological examinations yield the following observations: a decrease in the number of erythrocytes (RBC↓), a decrease in hemoglobin (HGB↓), a decrease in hematocrit (HTC↓) and an increase in reticulocytes (Reticulo↑). In addition, an increase in total bilirubin (T-bil↑) can be observed in blood biochemical examinations, and pigmentation of hemosiderin in the spleen is typically observed in histopathological examinations.

In order to create a category of anilines that induce hemolytic anemia, it is necessary to specify anilines that induce hemolytic anemia by the same pathway as shown in Figure 3.3. Table 3.1 shows some anilines with information related to hemolytic anemia. This information is contained in the Toxicity Knowledge Information Database and the Metabolism Knowledge Information Database. The findings related to hemolytic anemia shown in the table is obtained from RDT tests in male rats. The test for chemicals No. 4 and No. 8 is the Combined Repeat-Dose and Reproductive/Developmental Toxicity Screening test. The tests for the other chemicals are 28-day repeated dose toxicity tests.

This table shows that, as a rule, anilines with the potential to induce hemolytic anemia have a log $K_{ow} > 1$ (Nos. 1–8). On the other hand, two chemicals (amino phenols Nos. 9 and 10) with low log K_{ow} values of 0.24 showed only a weak potential to induce hemolytic anemia at high-dose levels. Three chemicals, amino benzene sulfonic acids (Nos. 11–13) that had even lower log K_{ow} values (negative values) lack the potential to induce hemolytic anemia even at high-dose levels. The reason aminophenols and amino benzene sulfonic acids exhibit weak or no potential to induce hemolytic anemia can be explained by their high excretion rate in urine due to their high water solubility.[20] These results indicate that a low log K_{ow} value could be used as a mitigating factor in the formation of category boundaries.

Table 3.1 Evidence related to the adverse outcome pathway in Figure 3.3 for anilines.

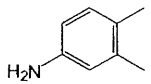
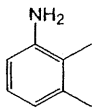
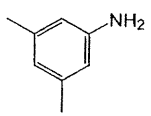
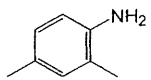
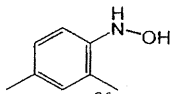
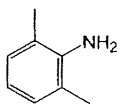
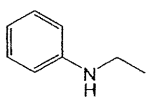
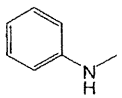
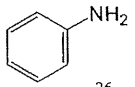
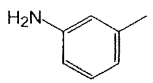
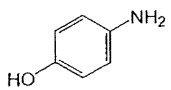
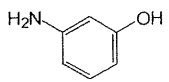
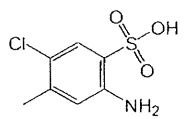
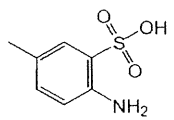
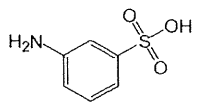
No.	*Structure*	*Observed findings related to hemolytic anemia in the repeated dose toxicity tests for male rats (dose mg^{-1} kg^{-1} day^{-1})*	*log K_{ow}**	*Mechanism information*		*Metabolism information*
				Met-Hb	*HBI*[23–24]	
1		RBC ↓ • Hgb ↓ · Hct ↓ · Ret ↑ Spleen pigmentation (250)	2.17	*in vitro*[21] *in vivo*[21]	0.7	—
2		Hgb ↓ (60) RBC ↓ · Hct ↓ · Ret ↑ · MetHb ↑ T-Bil ↑ (300)	2.17	*in vitro*[21] *in vivo*[21]	—	—
3		Hgb ↓ · Hct ↓ · Ret ↑ Spleen pigmentation (60) RBC ↓ (360)	2.17	*in vitro*[21] *in vivo*[21]	14	—
4		Bil ↑ (2) Hgb ↓ (10)	2.17	*in vitro*[21] *in vivo*[21]	2.3	 *in vitro*[25]
5		RBC ↓ · Hgb ↓ · Ret ↑ · Met-Hb ↑ Spleen pigmentation (250)	2.17	*in vitro*[21] *in vivo*[21]	1.1	—
6		RBC ↓ (5) Hgb ↓ · Hct ↓ · Ret ↑ · Met-Hgb ↑ Spleen pigmentation (25)	2.11	—	—	—

Table 3.1 (*Continued*)

No.	Structure	*Observed findings related to hemolytic anemia in the repeated dose toxicity tests for male rats (dose mg^{-1} kg^{-1} day^{-1})*	log K_{ow}*	*Mechanism information*		*Metabolism information*
				Met-Hb	*HBI*[23–24]	
7		RBC ↓ · Hgb ↓ · Hct ↓ · Ret ↑ Spleen pigmentation (25) Bil ↑ (125)	1.62	—	—	 *in vitro*[26]
8		Spleen pigmentation (30) RBC ↓ · Hgb ↓ · Hct ↓ · Bil ↑ (100)	1.62	*in vitro*[22] *in vivo*[22]	4.9	—
9		RBC ↓ (500)	0.24	—	—	—
10		Bil ↑ · Spleen pigmentation (720)	0.24	—	—	—
11		—	−0.89	—	—	—
12			−1.53	—	—	—
13		—	−2.08	—	—	—

*Calculated by KOWWIN 1.67 (US EPA).

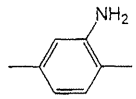
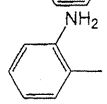
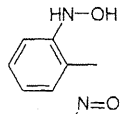
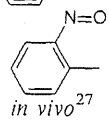
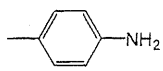
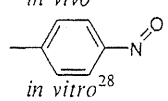
Mechanism and metabolism information can be used as supporting evidence for the pathway shown in Figure 3.3. For example, Met-Hb is not measured in many repeated dose toxicity tests, but the production of Met-Hb can be strong evidence for the pathway shown in Figure 3.3. However, such an information gap can be complemented by other studies.[21,22] For example, hemoglobin binding index (HBI)[23,24] values measured in other studies and the presence of metabolites such as *N*-hydroxy-, or nitroso-amines observed in other studies can be evidence for the pathway.

In comparison to the other anilines in the Table 3.1, 2,4-dimethylaniline (No. 4 in Table 3.1) did not show the same apparent potential to induce hemolytic anemia. This can be explained by the fact that the chemical was tested at rather low dose levels. However, the possible potential for this chemical to induce hemolytic anemia can be predicted by the pathway shown in Figure 3.3 by the supporting evidence that the production of Met-Hb was observed[21] as a result of the production of *N*-hydroxy-2,4-dimethylaniline as a metabolite[25] of the target chemical.

In the case of *N*-methylaniline (No. 7), the existence of aniline as a metabolite[26] in the *in vitro* test supported the fact that *N*-methylaniline induces hemolytic anemia. Therefore, it is suggested that the boundary of the category can be extended from primary aniline to *N*-alkyl aniline.

Based on the above discussions, chemical Nos. 1–8 in Table 3.1 can be specified as the category members inducing hemolytic anemia by the pathway shown in Figure 3.3. The boundary of the category can be defined not only by

Table 3.2 Evidence related to the adverse outcome pathway in Figure 3.3 for untested anilines.

No.	Structure	*Observed findings related to hemolytic anemia in the repeated dose toxicity tests for male rats (dose mg^{-1} kg^{-1} day^{-1})*	log K_{ow}*	*Mechanism information*		*Metabolism information*
				Met-Hb	*HBI*[23,24]	
14		Untested	2.17	*in vitro*[21] *in vivo*[21]	7.3	—
15		Untested	1.62	—	4.0	  *in vivo*[27]
16		Untested	1.62	—	4.3	 *in vitro*[28]

*Calculated by KOWWIN 1.67 (US EPA).

the chemical structure but also by the activity of the chemicals. In the case of the members mentioned above, the boundary of the category can be defined as monocyclic anilines and alkylated anilines with log $K_{ow} > 1$. In HESS, the boundaries of RDT categories defined by such an approach are registered in the category library.

It is also necessary to consider AOPs using evidence other than actual RDT data in order to find an appropriate category for chemicals lacking experimental data. For example, from the structure of an untested chemical, we can predict the reactivity in the molecular initiating event in an AOP based on, for example, the nature and position of a substitution group. Table 3.2 shows examples of chemicals in a category defined by the method described above and the supporting evidence. Based on the RDT test data of the category members (Nos. 1–8 in Table 3.1), these chemicals (Nos. 14–16 in Table 3.2) can be estimated to induce hemolytic anemia at dose levels less than 250 mg kg^{-1} day^{-1}.

3.4 Conclusions

In this chapter, we have given an overview of HESS and the category approach for evaluating RDT tests, as well as an example of categorization by this system. Because this system is designed to provide detailed RDT test data in conjunction with mechanism and metabolism information, the system can provide reliable information based on scientific evidence to the experts who evaluate the safety of chemicals. Completion of the final version of the system is planned for 2012.

Acknowledgements

This work was supported by a NEDO grant for the "Development of hazard assessment techniques using structure–activity relationship methods" project.

References

1. http://dra4.nihs.go.jp/mhlw_data/jsp/SearchPageENG.jsp.
2. http://www.safe.nite.go.jp/jcheck/Top.do.
3. OECD, Manuals for Investigation of HPV Chemicals, OECD, Paris, 2007, Ch 3.2.
4. http://kate.nies.go.jp/.
5. http://qsar.cerij.or.jp/cgi-bin/QSAR/index.cgi?e.
6. A. Hirose, E. Kamata, H. Akiyama, M. Takahashi, T. Morita, M. Ema and M. Hayasi, 6th World Congress on Alternatives & Animal Use in the Life Sciences Abstracts, Japanese Society of Alternatives to Animal Experiments, Tokyo, 2007, p. 80.
7. Y. Sakuratani, J. Yamada, K. Kasai, Y. Noguchi and T. Nishihara, *SAR QSAR Environ. Res.*, 2005, **16**, 403.
8. http://www.safe.nite.go.jp/english/kasin.html.

9. http://www.nedo.go.jp/activities/portal/p07033.html.
10. M. Hayashi, *ALTEX*, 2009, **26**, 85.
11. Guidance Document for Using the OECD (Q)SAR Application Toolbox to Develop Chemical Categories According to the OECD Guidance on Grouping of Chemicals, Organisation for Economic Cooperation and Development, Paris, 2009.
12. http://ntp.niehs.nih.gov/.
13. K. Kobayashi, Y. Sakuratani, T. Abe, S. Nishikawa, J. Yamada, A. Hirose, E. Kamata and M. Hayashi, *J. Toxicol. Sci.*, 2010, **35**, 79.
14. S. Nishikawa, T. Yamashita, T. Imai, M. Yoshida, Y. Sakuratanil, J. Yamada, A. Maekawa and M. Hayashi, *J. Toxicol. Sci.*, 2010, **35**, 295.
15. T. Okada, N. Ohmori, S. Mori, H. Horikawa, M. Yamakawa, Y. Sakuratani, T. Yamada and M. Hayashi, *ALTEX*, 2009, **26**, 86.
16. T. Okada, *J. Comput.-Aided Chem.*, 2002, **2**, 79.
17. T. Okada, M. Yamakawa, N. Ohmori, S. Mori, H. Horikawa, T. Hayashi and S. Fujishima, *Chem. Cent. J.*, 2010, **4**, 1.
18. http://www.oecd.org/document/23/0,3343,en_2649_34365_33957015_1_1_1_1,00.html.
19. Y. Sakuratani, S. Sato, S Nishikawa, J. Yamada, A. Maekawa and M. Hayashi, *ALTEX*, 2009, **26**, 80.
20. Y. Sakuratani, S. Sato, S. Nishikawa, J. Yamada, A. Maekawa and M. Hayashi, *SAR QSAR Environ. Res.*, 2008, **19**, 681.
21. D. Cauchon and K. Krishnan, *J. Appl. Toxicol.*, 1997, **17**, 397.
22. M. Kiese, *Pharmacol. Rev.*, 1966, **18**, 1091.
23. G. Sabbioni, *Environ. Health Perspect.*, 1993, **99**, 213.
24. G. Sabbioni, *Environ. Health Perspect.*, 1994, **102**, S6.
25. T. Nohmi, R. Miyata, K. Yoshikawa, M. Nakadate and M. Ishidate, *Biochem. Pharmacol.*, 1983, **32**, 735.
26. R. Kato, M. Takayanagi and T. Ohshima, *Jpn. J. Phamcol.*, 1969, **19**, 53.
27. O. S. Son, D. W. Everetf and E. S. Fiala, *Xenobiotica*, 1980, **10**, 457.
28. B. Tyrakowska, S. Boeren, B. Geurtsen and I. M. C. M. Rietjens, *Drug Metabol. Disposition*, 1993, **21**, 508.

Chemical Research in Toxicology

Article

pubs.acs.org/crt

Investigating the Relationship between in Vitro–in Vivo Genotoxicity: Derivation of Mechanistic QSAR Models for in Vivo Liver Genotoxicity and in Vivo Bone Marrow Micronucleus Formation Which Encompass Metabolism

Ovanes G. Mekenyan,[†] Petko I. Petkov,[†] Stefan V. Kotov,[†] Stoyanka Stoeva,[†] Verginia B. Kamenska,[†] Sabcho D. Dimitrov,[†] Masamitsu Honma,[‡] Makoto Hayashi,[‡] Romualdo Benigni,[§] E. Maria Donner,[∥] and Grace Patlewicz*[,∥]

[†]Laboratory of Mathematical Chemistry (LMC), As. Zlatarov University, Bourgas, Bulgaria

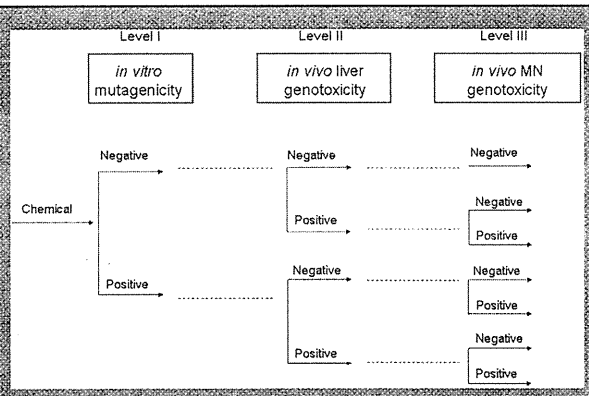
[‡]Division of Genetics and Mutagenesis, National Institute of Health Sciences, Tokyo, Japan

[§]Environment and Health Department, Istituto Superiore di Sanita', Rome, Italy

[∥]DuPont Haskell Global Centers for Health and Environmental Sciences, Newark, Delaware 19714-0050, United States

S *Supporting Information*

ABSTRACT: Strategic testing as part of an integrated testing strategy (ITS) to maximize information and avoid the use of animals where possible is fast becoming the norm with the advent of new legislation such as REACH. Genotoxicity is an area where regulatory testing is clearly defined as part of ITS schemes. Under REACH, the specific information requirements depend on the tonnage manufactured or imported. Two types of test systems exist to meet these information requirements, in vivo genotoxicity assays, which take into account the whole animal, and in vitro assays, which are conducted outside the living mammalian organism using microbial or mammalian cells under appropriate culturing conditions. Clearly, with these different broad experimental categories, results for a given chemical can often differ, which present challenges in the interpretation as well as in attempting to model the results in silico. This study attempted to compare the differences between in vitro and in vivo genotoxicity results, to rationalize these differences with plausible hypothesis in concert with available data. Two proof of concept (Q)SAR models were developed, one for in vivo genotoxicity effects in liver and a second for in vivo micronucleus formation in bone marrow. These "mechanistic models" will be of practical value in testing strategies, and both have been implemented into the TIMES software platform (http://oasis-lmc.org) to help predict the genotoxicity outcome of newly untested chemicals.



INTRODUCTION

Terms of Reference: Genotoxicity versus Mutagenicity.

Carcinogenicity and mutagenicity are among the toxicological end points that pose the highest concern for human health and are subject to regulatory testing for hazard and risk assessment. Much of the data that are currently available in the public domain have thus been derived from tests conducted to investigate potentially harmful effects on genetic material, that is, genotoxicity or mutagenicity. Since both terms, mutagenicity and genotoxicity, will be referenced in this paper, working definitions are given. According to academic definitions, genetic alterations that are fixed and can be inherited are termed mutations. These include different types of events such as base substitutions and deletions, structural chromosomal aberrations (CAs) (break and rearrangements), and numerical CAs (loss or gain of chromosomes, i.e., aneuploidy). The assays established to evaluate these events are described in brief. Genotoxicity is considered as a broader term—aside from mutations, it also encompasses other alterations of genetic material that are not fixed and are not inherited, such as DNA damage. Genotoxicity may or may not be transformed into mutations by the cell's machinery during cell replication, and it may be an indication of potential carcinogenesis associated with the exposure to a chemical agent. Appropriate in vivo experimental test systems used to evaluate genotoxicity include the bone marrow in vivo micronucleus test (MNT) assay, the unscheduled DNA synthesis (UDS) assay, and the alkaline single-cell gel electrophoresis assay (Comet assay). These tests are relevant to assess DNA-damaging and DNA-repair processes in specific organs of investigation in the whole animal such as liver. Therefore, the term *liver genotoxicity* was regarded as appropriate for the purposes of this study, although, overall, a wide array of other

Received: June 3, 2011

events aside from mutations are encompassed in these test systems.

Current Quantitative Structure–Activity Relationship (QSAR) Approaches. The importance of assessing genotoxicity coupled with the availability of experimental data has prompted many in silico studies. James and Elisabeth Millers's "electrophilic theory" introduced a chemical concept to help rationalize the mode of action of genotoxic carcinogens.[1] This prompted many evaluations to derive so-called structural alerts (SA), simple yet effective means of encoding qualitative mechanistic understanding for predicting potential mutagenicity/carcinogenicity. Seminal efforts include SA for carcinogenicity by John Ashby,[2] who subsequently extended his list with additional SA.[3] Bailey et al. compiled a set of 33 SAs for regulatory use within the U.S. Food and Drug Administration (FDA), which was predominantly based on the Ashby alerts.[4] Kazius et al. evaluated a mutagenicity database comprising 4337 mutagens and nonmutagens taken from the Toxnet database (http:/toxnet.nlm.nih.gov/) and derived 29 SAs for mutagenicity with associated detoxification fragments.[5] Some of these alerts exist in software platforms to enable routine use; for example, 17 SAs for mutagenicity are implemented into the OASIS tissue metabolism simulator (TIMES) software.[6] Benigni et al. combined the published information from Ashby, Bailey et al., and Kazius et al. with additional information from the OncoLogic (U.S. EPA) software (http://www.epa.gov/oppt/sf/pubs/oncologic.htm)[7] to arrive at a list of 33 SA for carcinogens and mutagens.[8]

Current quantitative strategies include (Q)SARs and expert systems. Two types of (Q)SAR models, local and global, exist to estimate the mutagenic potential of chemicals. Local (Q)SARs provide estimated results for closely related (congeneric) chemical structures. Such models are most predictive, but only if the essential features of the model domains are clearly represented. Models based on physicochemical descriptors with clear mechanistic meaning are particularly helpful in rationalizing genotoxic outcome as exemplified by Chung et al.[9] Other local models are based on mathematical representations of chemical structure, for example, topological indices, and thus are more difficult to interpret.[10]

Global (Q)SARs aim to provide mutagenicity estimations for a diverse (noncongeneric) set of chemicals. Such (Q)SARs may be additionally encoded into expert systems. For example, TOPKAT empirically makes predictions for a range of different end points including Ames mutagenicity and rodent carcinogenicity.[11] Other expert systems such as TIMES attempt to provide clear mechanistic meaning through the use of SAs, which address the reactivity toward DNA and/or proteins.[12,13] TIMES also includes 3D QSARs to underpin some of the available SAs. All of the aforementioned (Q)SARs have typically been derived on Ames (*Salmonella* mutagenicity data). TIMES includes a platform for in vitro CA data in addition to that for Ames.[13] There is a paucity of models for in vivo genotoxicity, but as highlighted in the survey by Benigni et al., there is only one publically available model for in vivo micronucleus.[14] The scarcity of such models may be due in part to experimental data being less readily available but also due to the complexicity of how to rationalize and interpret the outputs from the different test systems.

Our own investigation aims to fill in the above in vitro–in vivo genotoxicity gap by considering both the available test systems and how they are currently applied to formulate an approach for modeling in vivo genotoxicity. For convenience, we considered the REACH ITS[15] for mutagenicity since this described the typical assays used and how their outcomes should be interpreted for subsequent decision making. The actual experimental test systems are assumed to be reasonably familiar and are only briefly described in the next section.

Experimental Assays and Data for Rodent Mutagenicity and Genotoxicity. Integrated testing strategies, notably those described in the REACH Technical guidance,[15] outline the in vitro and in vivo systems that are most frequently used to evaluate the mutagenic potential of chemical substances. The in vitro systems include the bacterial reverse mutation test (Ames), an in vitro mammalian cell gene mutation test [such as the mouse lymphoma or hypoxanthine–guanine phosphoribosyltransferase (*hprt*) assay], the in vitro mammalian chromosome aberration (CA) test, and the in vitro MNT.[15] The Ames test uses amino acid-requiring strains of bacteria to detect (reverse) gene mutations (point and frameshift mutations). The in vitro mouse lymphoma assay (MLA), when correctly performed, detects structural chromosome aberrations, aneuploidy, and recombination events (e.g., such as gene conversion) that result in loss of heterozygosity. The *hprt* test identifies chemicals that induce gene mutations in the *hprt* gene of established cell lines. The in vitro mammalian CA test detects structural chromosome aberrations and increases in polyploidy. The in vitro MNT has the potential to detect both clastogenic (chromosome aberrations) and aneugenic (chromosome lagging due to dysfunction of mitotic apparatus) chemicals.

The scheme under REACH can be summarized as follows. As a first tier, three in vitro tests are recommended, which includes an Ames test, a mouse micronucleus/CA, and a mouse lymphoma/HRPT assay. If the results from all three tests are negative, then no more testing is merited, and a conclusion of nongenotoxicity can be made for the substance under study. If one or more tests are positive, then in vivo testing may be instigated. Obviously metabolism, pharmacokinetics, and toxicokinetics factors [absorption, distribution, metabolism, excretion (ADME)] are all inherent features in the in vivo genotoxicity tests, although the genetic end points for the tests address different genetic mechanisms. The UDS in vivo assay is used to evaluate the role of DNA repair. The in vivo Comet assay is a sensitive technique for the detection of DNA strand breaks; thus, it can be used for measuring DNA strand breaks in any tissue of an animal. Site-specific effects at contact tissues or the target tissue where the test compound accumulates or induces toxicity can be readily assessed. The specificity of the contact tissue under investigation is also feasible for the transgenic rodent gene mutation test (TGR), which measures gene mutations in vivo. However, the in vivo MNT is probably the most widely used test.[16] When performed appropriately, it detects both clastogenicity and aneugenicity.[17] The frequency of micronucleated polychromatic erythrocytes is traditionally determined from bone marrow samples, but with the emerging automated scoring methods, the emphasis is moving to assessing the induction of micronuclei in immature erythrocytes in peripheral blood samples.[18]

Most of the established in vitro mutagenicity tests, which are used for regulatory purposes, exhibit relatively high sensitivity for detection of genotoxic carcinogens.[19] However, particularly those based on cultured mammalian cells are thought to produce a remarkably high occurrence of irrelevant positive results (i.e., exhibit low specificity), when compared with rodent carcinogenicity.[19,20] To increase the specificity of predictions, regulators tend to interpret in vitro positive results in an in vivo perspective, that is, in vivo confirmation of in vitro mutagens.

In addition, in vivo tests can also be utilized to identify chemicals producing in vivo only positive results (i.e., chemicals for which mutagenicity is not or poorly detected in vitro). Only a very limited number of chemicals have been found to be genotoxic in vivo and not in the standard in vitro tests. Most of these are pharmaceuticals such as atovaquone (95233-18-4), which is designed to affect pathways of cellular regulation, including cell cycle regulation. One of the most preferred in vivo assays, complementing genotoxicity test batteries, is the in vivo bone marrow MNT. The preference of this assay is attributed to both its wide mutagenicity range assessment (clastogenicity and aneugenicity) and its remarkably high specificity in concordance with the genotoxic carcinogenicity model, although it shows low sensitivity.[14,21] Therefore, it may be appropriate to include a second in vivo test if a positive in vitro result has not been adequately confirmed by the in vivo bone marrow MNT test. The UDS test is one complement to the bone marrow MNT since it is a surrogate in vivo gene mutation assay[21] measuring DNA excision repair of induced DNA damage. The utility of the Comet and the TGR assays to detect genotoxic damage in specific tissues, specifically DNA strand breaks and gene mutations has also been recognized.[15] Thus, an evaluation of in vivo genotoxicity potential could involve integrating outcomes from MNT and either UDS, Comet, and TGR tests depending on the outcomes that have been observed in vitro. UDS, Comet, and TGR can also be undertaken to address in vivo liver genotoxicity. Such tissue-specific assays are useful in in vivo follow-up tests especially since the liver is an organ of high metabolic capacity and therefore is frequently subjected to significant toxic overload.

Aims of the Study. Bearing in mind the way in which these different assays are integrated together, our goal was to investigate the in vitro and in vivo relationship, the so-termed in vitro–in vivo "gap" to inform the development of mechanistic (Q)SAR model(s). A large body of data covering in vitro mutagenicity, in vivo (liver) genotoxicity, and in vivo bone marrow MNT test results was collected for the same set of substances. The scope of the investigation can be summarized in the following three questions: (a) To what extent are in vitro mutagenic chemicals in vivo (liver) genotoxic, that is, what in vivo detoxification pathways exist? (b) To what extent are in vivo (liver) genotoxic chemicals in vivo bone marrow MNT positive? (c) Are there in vitro nonmutagenic chemicals that are in vivo liver or bone marrow genotoxic; that is, what in vivo bioactivation pathways exist? These questions were structured into a workflow (Figure 1) and enabled a stepwise evaluation of the in vitro–in vivo gap.

MATERIALS AND METHODS

Compilation of Data Set. Our training set comprised 557 chemicals ("557 list") with in vivo MNT data (Appendix I of the Supporting Information lists the substances and their overall calls). In vitro mutagenicity and in vivo (liver) data were collected for the same set of substances to the extent possible. This helped maximize the overlap between chemicals with various genotoxicity effects and the in vivo MNT data set. Documented in vitro mutagenicity data from multiple literature sources were identified for 397 noncongeneric chemicals within the training set (Appendix II of the Supporting Information). Positive calls were categorized by the digit 1, negative calls by 0, and N/A signified "no data available", based on the literature searches that were performed. Our in vitro data comprised that from the Ames assay, the CA assay, and the MLA, since these are the typical assays considered under REACH. Out of necessity and as typically the case for modeling efforts, reported study results were accepted as

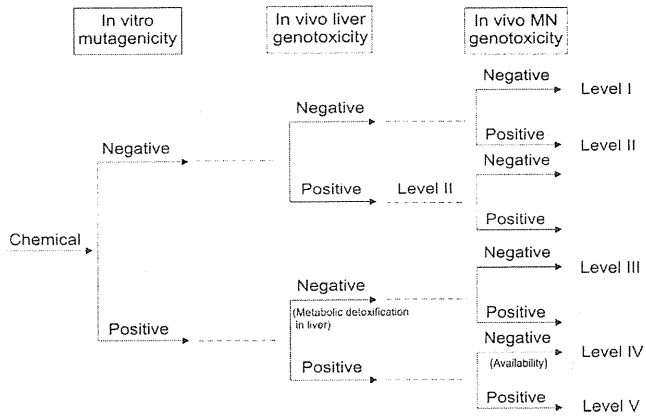


Figure 1. Workflow outlining the in vivo–in vitro gap.

reported, although an extensive effort was made in expert judgment and evaluation of the data quality and correctness of the calls.

Ames results with the rat liver S9 metabolic activation system were available for 283 noncongeneric chemicals. Of these chemicals, 109 (38%) were associated with positive calls and 174 (62%) with negative calls. Documented in vitro CA test data were identified for 296 chemicals, of which 186 (63%) were positive and 110 (37%) were considered negative. Data from 194 chemicals had been assessed in the in vitro MLA. The majority of the chemicals tested positive (148 chemicals, i.e., 76%) and 46 chemicals (24%) tested negative. For the 397 in vitro mutagenicity data, these comprised 267 positive calls (68%) and 124 negative calls (32%), and six calls were inconclusive. These substances were ethylene dichloride (107-06-2), sulfan blue (129-17-9), thiabendazole (148-79-8), methyl parathion (298-00-0), dibutylnitrosamine (924-16-3), C.I. direct black 38 (1937-37-7). In these six cases, only Ames and in vitro CA test outcomes were available with positive calls in Ames and negative calls in in vitro CA tests.

Results from in vivo Comet, UDS, and TGR assays were also collected to help evaluate in vivo liver genotoxic potential. Data were available for 185 diverse chemicals, which are listed in Appendix III of the Supporting Information. The Comet assay provided liver genotoxicity assignments for 127 (69%) of the 185 chemicals. Of the 127 chemicals, 78 (61%) were positive, and 49 (39%) were negative. The TGR comprised rodent liver genotoxicity data for 34 (18%) of the 185 chemicals; 27 (80%) of these were reported as positive, and 7 (20%) were negative. The in vivo UDS assay was associated with the least amount of liver genotoxicity data, only 24 (13%) of the 185 chemicals had overall calls, and five of them were observed to be positive in this assay (21%), and 19 were (79%) negative in this assay. Overall, of the 185 substances with liver assignments, 109 were associated with positive calls (59%) and 76 with negative calls (41%). The "557 list" included almost equal numbers of positive (267 chemicals, i.e., 48%) and negative (290 chemicals, i.e., 52%) MNT assignations performed in either bone marrow or peripheral blood. Figure 2 summarizes the distribution of assignments in each of the test systems.

The evaluation of this investigation was often hampered by conflicting in vivo MNT data available in the public domain. The compromised quality of these MNT data was attributed to the fact that many chemicals had been evaluated in the early 1980s; when species (rat vs mouse) and gender (male vs female) differences may not always have been considered, etc. To date, the validity of the in vivo MNT data has only been verified for chemicals where the in vitro mutagenicity outcome appeared to be negative, relative to the in vivo case (in either liver or bone marrow), where the genotoxicity result was positive. Expert judgment was relied upon to consider whether there were factors resulting in inconsistent in vitro results as compared with the in vivo situation, for example, rodent species differences, nonphysiological culture conditions, etc.

To illustrate the structural diversity of the training set, the 557 list was profiled against the set of DNA and protein binding alerts available within the OECD Toolbox v2.1. The distribution chart is