VALIDOVANÉ POČÍTAČOVÉ MODELY – QUANTITATIVE STRUCTURE – ACTIVITY RELATIONSHIP TOOL BOX

MARIÁN RUCKI a MILOŇ TICHÝ

Státní zdravotní ústav, Šrobárova 28, 100 42 Praha 10
mtichy@szu.cz

Došlo 19.1.09, přijato 20.4.09.

Klíčová slova: QSAR, Tool Box, validace, legislativa, OECD

Obsah

1. Úvod
2. (Q)SAR Application Tool Box a jeho vývoj
3. Kategorie chemických sloučenin
4. Využití Tool Boxu
5. Budoucnost systému Tool Box
6. Závěr

1. Úvod

Tento příspěvek se zabývá legalizací počítačových výsledků o nebezpečnosti chemických sloučenin pro legislativní použití. Co je zapotřebí, aby takové výsledky byly stejně použitelné jako výsledky tradičních testů *in vivo* a *in vitro?* Pro chemické bezpečnostní listy, legislativní posuzování nezávadnosti chemických přípravků a potravin máme nový druh údajů, a to odhady výpočtem *in silico*. Část dětství v historii QSAR (z angl. Quantitative Structure – Activity Relationbships, kvantitativní vztahy mezi chemickou strukturou látek a jejich biologickou účinností), začínající koncem 50. let a v 60. letech, tak končí.

V listopadu 2002 na jednání 34. společné schůze organizací Chemicals Committee a Working Party on Chemicals, Pesticides and Biotechnology v Paříži bylo dohodnuto ustavení skupiny expertů při Organizaci pro hospodářskou spolupráci a rozvoj OECD (Organization for Economic Co-operation and Development), která se nyní zabývá problematikou legalizace výsledků modelů QSAR. Následovala další setkání v Ispře (Itálie), která organizovalo Evropské centrum pro validaci alternativních metod při Spojených výzkumných ústavech Evropské komise ECVAM/JRC EC (European Centre for Validation of Alternative Methods / Joint Research Centre at European Committee) v roce 2003 a další opět v Paříži v září 2004. Toto poslední setkání organizoval sekretariát OECD spolu s JRC EC/ECVAM. Projednávala se legalizace výsledků analýzy QSAR jako alternativní metody pro testování toxicity *in silico* v návaznosti na jednání dalších skupin o bezpečnosti expozice chemickým látkám a hledání principů validace metod *in vitro*, které probíhalo ve městě Setubal v Portugalsku v roce 2002 (odtud název Setubalské principy). Závěrem byl formulován návrh principů validace metod (Q)SAR (tzn. metod QSAR i SAR), které musí být splněny, aby výsledky odhadem pomocí modelů QSAR byly přijaty stejně jako výsledky alternativních metod, spočívajících na využití nižších organismů, tkání, buněk apod.

Tato schůzka byla pro další vývoj legalizace výsledků modely QSAR velmi významná. Sešli se zde právníci, úředníci státních a vládních správ z nadnárodních i soukromých organizací s odborníky na výpočetní modely pro odhad toxicity výpočtem (jak zní celý termín), chemiky a toxikology. Státy byly zastoupené v celém spektru OECD, ale i mimo něj: od Kanady přes Evropskou unii, po Japonsko a Koreu. Byly hledány a vybírány alternativní metody rychlejší než metody klasické, zároveň nejméně stejně informativní a levné, již dobře propracované, poznané a vyvinuté.

Po bouřlivé diskusi se podařilo zformulovat pět kriterií „Setubalských principů pro validaci QSAR“, nověji již „QECD principy“. Je vhodné upozornit, že i v takovéto diskusi se projevily obchodní a politické motivy. Principy[1] v tom pořadí, jak jsou uváděny, zní:

1. Musí být přesně definovaný účinek, jehož velikost je odhadována.
   Potíž je v tom, že odhadovaný účinek může být měřen podle různých experimentálních protokolů a za různých experimentálních podmínek.
2. Musí být jednoznačně uveden algoritmus, který byl k tvorbě modelu použit.
   Je známo, že u mnohých komerčních modelů tato informace není, firmy ji nepodávají a veřejně přístupná není. Bez této informace však nemůže být model nezávisle ověřen. To může být omezující při přijetí výsledků pro právní účely.
3. Musí být definována oblast použitelnosti.
   QSAR modely jsou bezprostředně spojeny s typy chemické struktury, fyzikálně-chemickými vlastnostmi nebo mechanismem účinku, pro které byly tyto modely vytvořeny. Nejde jen o skupiny chemických látek, o rozsah velikostí fyzikálně-chemických vlastností, experimentální podmínky, ale při měření velikosti biologického účinku i o přesně vymezený biologický objekt, včetně pohlaví, věku, podmínek měření, jako je teplota a tlak.
4. Musí existovat vhodné míry statistického hodnocení dobré shody, robustnosti a schopností odhadovat (predictivity).
   Jde o vnitřní i vnější validaci. O tomto bodu existuje již řada studií, množství návodů a průvodců, obsahující i doporučované statistické indexy a způsoby validace[2,3] (křížová validace – cross validation, křížový

korelační koeficient Q a jeho kvadrát, korelační koeficient r a jeho kvadrát, indexy schopnosti odhadu, jako jsou PRESS = Σ[(exp – cal)(exp – cal)] Predictive REsidual Sum of Squares a PRESD = (PRESS/n)^{½} Predictive REsidual Standard Deviation, což jsou standardní chyba a standardní rozptyl odhadu, apod.).

5. Uvést mechanistickou interpretaci měřeného účinku, pokud ji lze najít.
   Toto kriterium nemá za účel vyčlenit modely bez interpretace účinku, ale přinutit k zamyšlení, jaký mohou mít vztah použité molekulové deskriptory k měřenému účinku, ať jde o fyzikálně-chemické indexy nebo molekulové fragmenty a že tato skutečnost (zamyšlení) je dokumentována.

2. (Q)SAR Application Tool Box a jeho vývoj

Na základě těchto principů se začalo s vývojem „nástroje“, který se nazval (Q)SAR Application Tool Box[4] (dále bude v textu pro zkrácení uváděn jen jako Tool Box). Jde o softwarovou aplikaci, která by měla být využívána vládními úředníky, úředníky chemického průmyslu a dalšími pracovníky, kteří potřebují ke své činnosti znát údaje o toxických a ekotoxických vlastnostech chemických sloučenin a to i pro sloučeniny, které testovány nebyly. Tool Box je soubor nástrojů pro tvorbu validovaných modelů QSAR, který bude tvořen ve více etapách. První verze byla zveřejněna v březnu 2008, aby byl prověřen koncept tohoto nástroje a první upravená verse 1.1 v prosinci 2008. Tvorbě komplexního systému pro vyhledání potřebných toxikologických údajů bylo věnováno již hodně úsilí[5].

Důvodů, proč hledat alternativní metody pro testování toxicity chemických sloučenin bylo hned několik. Byl to požadavek ukotvený v direktivě 86/609/EEC o ochraně zvířat proti týrání. Impulsem byla práce, kterou publikovali Burch a Russell[6]. Ta obsahovala požadavek 3R (tří er), z angl. Reduction, Refinement, Replacement, tedy snížit, zlepšit, nahradit pokusná zvířata při získávání informací o účincích sloučenin na zdraví. Dále to bylo právní nařízení o omezení testování kosmetických přípravků na zvířatech (76/768/EEC/EEC). V neposlední řadě to je program chemické bezpečnosti REACH (Registration, Evaluation, Authorisation and Restriction of Chemicals) č. 1907/2006/EC. Ten ukládá obchodníkům a průmyslovým pracovníkům nejen pravidla o zacházení a distribuci chemických sloučenin, ale požaduje snížení testování na obratlovcích. Tomu mají odpomoci nejen alternativní metody, ale i sdílení výsledků toxikologického testování všemi, kteří je požadují, aby nebylo testování téže látky nebo přípravku prováděno vícekrát, než je nutné.

Modely (Q)SAR a jiné metody pro odhad toxických účinků výpočtem byly vysvětlovány v několika kurzech. Ty byly zaměřeny zejména na pracovníky ze státního aparátu a z průmyslu členských zemí EU, kteří se zabývají právními předpoklady pro přijetí výsledků testování toxicity. První kurz byl organizován v Sofii v říjnu 2005 a byl zaměřený na teorii i využití modelů pro odhad toxicity výpočtem – *in silico* technika. Druhý kurz, kterého se zúčastnilo již 39 posluchačů z 13 zemí, následoval v červenci 2006 v Ispře (Itálie). Ve třetím kurzu, který proběhl opět v Ispře v červnu 2007, byly týmem z European Chemical Bureau představeny softwarové systémy, expertní systém a datový soubor AMBIT a organizacemi CEFIC (European Chemical Industry Council, European Committee, Brusel) a OECD navržený a vyvinutý Tool Box. Tool Box je samostatný počítačový program pracující pod Windows, vyžaduje procesor pracující při 2 GHz nebo rychleji, 8 Gb místa na pevném disku a 1 Gb pro operační RAM. Je k dostání bez poplatků, protože jeho vývoj byl hrazený Evropskou unií. Verzi 1.1 (prosinec 2008) lze nalézt na adrese www.oecd.org/env/existingchemicals/qsar spolu s návodem na instalaci nebo přes vyhledávač Internetu pod heslem QSAR Tool Box, kde najdete vysvětlení jeho funkce a jsou uvedeny nejnovější informace a verze.

Smyslem projektu Tool Box je, aby validace a využívání modelů QSAR byly snadno dostupné, transparentní a pro administrativní pracovníky pochopitelné, a zároveň méně nákladné. Tool Box využívá informace a nástroje z různých zdrojů tím, že jsou do něho zabudovány a spojeny do logického celku. Jsou to údaje toxikologických datových souborů, údaje, které byly někým získány dřívějšími pokusy na zvířatech nebo pokusy *in vitro*, datový soubor a expertní vyhledávající program AMBIT (CEFIC) a údaje, které dosud v souborech zahrnuty nejsou. Řada organizací se zabývá tím, že je vyhledává v literatuře a tvoří nové toxikologické datové soubory (rozšířený AMBIT) (granty podporované OECD a CEFIC). Této práce se zúčastnilo více pracovních skupin, nejenom z OECD. Tak spolupracovaly např. European Chemical Bureau, dánská EPA, US EPA, Environment Canada, NITE Japan, European Chemical Industry Council i velké organizace chemického průmyslu, které poskytovaly experimentální údaje jimi dříve získané, jako jsou ExxonMobil, Unilever a další.

3. Kategorie chemických sloučenin

Klíčová část činnosti Tool Boxu je schopnost seskupovat chemické látky do chemických kategorií. Chemická kategorie je skupina chemických sloučenin, jejichž fyzikálně-chemické, toxikologické, ekotoxikologické vlastnosti nebo jejich osud v životním i pracovním prostředí je podobný nebo se chovají podle společného vzoru jako výsledek strukturní podobnosti.

4. Využití Tool Boxu

Uživatel může Tool Box využívat nejen pro popis chemické struktury sloučeniny, ale i pro nalezení vhodných experimentálních výsledků, pro vyhledávání, zda sloučenina již není v národních seznamech nebo v existujících chemických kategoriích nebo zda nebyla hodnocena nějakou jinou organizací. Tool Box vyhledá v seznamech možné analogy, seskupí sloučeniny podle mechanismu účinku nebo na základě strukturní podobnosti

a podle společného metabolitu. Z těchto skupin vybere chemické sloučeniny s anomálním metabolismem nebo anomálním mechanismem toxického působení.

Chybějící údaje v kategoriích doplní pomocí modelů QSAR, analýzy trendů nebo pomocí techniky „read-across“ nebo pomocí knihovny modelů QSAR. V kategoriích nemusí být údaje pro všechny měřené účinky sloučenin, které jsou v kategorii obsažené, ale ostatních musí být dostatek, aby chybějící mohly být odhadnuty. Tool Box uspořádá odpovědi i pro tisk nebo export svých výsledků.

Tool Box tak dovoluje uživateli systematicky kategorizovat chemické sloučeniny do skupin podle jednotlivých účinků pro všechny členy kategorie za předpokladu obecného chemického nebo toxického účinku nebo typu účinku nebo strukturní podobnosti nebo společného metabolitu. Stejným způsobem je Tool Box schopen rychle vyhodnotit obecné toxikologické chování pro všechny členy kategorie nebo odhadnout shodné trendy mezi důležitými účinky, a to i pro legislativní použití.

Technika kategorizace v Tool Boxu přináší možnost soustředit se na vlastnosti chemických látek, jako je toxický efekt nebo typ mechanismu účinku, dovoluje provést odhad pro všechny členy kategorie, aniž by všechny požadované údaje existovaly. Tím šetří náklady finanční, časové i pokusná zvířata. Spolehlivý odhad nebezpečnosti chemických sloučenin umožňuje mechanistickým porovnáváním analogů bez nutnosti dalších pokusů.

To je umožněno tím, že Tool Box obsahuje velké množství datových souborů a experimentálních i odhadnutých výsledků fyzikálně-chemických vlastností sloučenin, údajů o osudu látek v životním i pracovním prostředí a o toxikologických i ekotoxikologických účincích sloučenin. Tool Box je otevřený systém, do kterého lze importovat další datové soubory v případě, že jsou pro účely Tool Boxu poskytnuty. Čerpá výhody z příspěvků expertů ve vládách, státních i soukromých organizacích, ale i nevládních neziskových organizací, jako jsou různá občanská sdružení, ekologické organizace nebo nadace. Tím je splněna i podmínka, aby systém byl rychle veřejně přístupný a použitelný.

5. Budoucnost systému Tool Box

Druhá fáze činností spojených s QSAR Tool Boxem bude probíhat po čtyři roky v letech 2008 až 2012 a nové verze 2.0 a 3.0 se poskytnou veřejnosti v roce 2010 a 2012. Budou nalezeny na již výše uvedené adrese www.oecd.org/env/existingchemicals/qsar. Činnost bude zaměřena na vývoj potřebných informačních technologií, na rozšíření a vývoj nových schopností systému, na doplňování datových souborů, na vytvoření (Q)SAR knihovny a expertního systému na vyhledávání dat a zejména na osvětu, školení, kurzy a výuku o jeho využívání pro legislativní účely pro potřebné pracovníky. Dlouhodobým cílem je zajištění techniky kategorizace pro doplňování údajů o všech organických sloučeninách a pro všechny toxické účinky, které legislativa požaduje. Při OECD se proto tvoří samostatná skupina, zabývající se rozvojem a využitím Tool Boxu.

6. Závěr

V příspěvku je popsán způsob, jak dospět k validaci výsledků modelů Quantitative Structure –Activity Relationships (QSAR) pro legislativní účely. Základní impuls k vytvoření systému QSAR Application Tool Box přišel od OECD. Jde o otevřený systém, který umožňuje vytvářet modely QSAR validované pro legislativní použití a který bude dále rozvíjen.

Tato práce byla podpořena granty GA ČR 203/06/1265, IGA MZ ČR NS/9647-4 a MŠMT 2B008075 a Státním zdravotním ústavem v Praze.

QSAR Tool Box byl vyvinut v Laboratory of Mathematical Chemistry, University „Prof. Assen Zlatarov“, Burgas, Bulharsko za podpory OECD, CEFIC, European Chemical Bureau, dánské EPA, US EPA, Environment Canada, NITE Japan jako zdrojů informací a krytí nákladů.

LITERATURA

1. Report from the Expert group on (quatitative structure – activity relationships(Q)SAR) on the principles for the validation of (Q)SARs. (ENV/JM/MONO), Series on Testing and Assessment, No. 49, OECD, Environment Directorate, 2004.
2. Eriksson L., Jaworska J., Worth A. P., Cronin M. T. D., McDowel R. M., Gramatica P.: Environ. Health Persp. *111*, 1361 (2003).
3. Wold S.: Quant. Str. Act. Relat. *10*, 191 (1991).
4. US EPA/EC Joint project on the evaluation of (Quantitative Structure Activity Relationships. (ENV/JM/MONO), Series on Testing and Assessment, No. 88, OECD, Environment Directorate, 2008.
5. LRI LandmarkProjects: Alternatives to animal testing (AMBIT and OECD Toolbox) (Watze de Wolf, DuPont Coordination Centre), CEFIC Annual Workshop, Brussels, November 20-21, 2008.
6. Russell W. M. S., Burch R. L.: *The Principles of Human Experimental Techniques*. Methuen, London 1959.

M. Rucki and M. Tichý *(National Institute of Public Health, Prague, Czech Republic):* **Tools for Development of Validated Computer Models – QSAR Application Tool Box**

A way to legalization results of Quantitative Structure – Activity Relationships (QSAR) models is described. A basic impulse has come from OECD followed by the other countries of the world. Final tool, QSAR Application Tool Box, has been developed and as an open system will be developed in the future, too.