

UNIVERZA V LJUBLJANI
FAKULTETA ZA FARMACIJO

SIMONA JURČIČ (DOLENEC)

**NAPOVEDOVANJE KANCEROGENOSTI IN MUTAGENOSTI
SPOJIN Z RAČUNALNIŠKO APLIKACIJO TOXTREE**

**PREDICTION OF CARCINOGENICITY AND MUTAGENICITY OF
COMPOUNDS WITH TOXTREE EXPERT SYSTEM**

Ljubljana, 2010

Diplomsko delo sem izdelala na Fakulteti za farmacijo Univerze v Ljubljani pod mentorstvom prof. dr. Marije Sollner Dolenc, mag. farm..

Zahvala

Iskreno se zahvaljujem svoji mentorici prof. dr. Mariji Sollner Dolenc, mag. farm., za vso strokovno pomoč, vzpodbudo in nasvete, ki so me pri pripravi diplomske naloge usmerjali na pravo pot. Zahvala velja tudi prof. dr. Juretu Zupanu in doc. dr. Marjanu Vračku Grobelšku iz Kemijskega inštituta v Ljubljani, za pomoč pri navajanju ustreznega izrazoslovja na področju izgradnje QSAR modelov.

Rada bi se zahvalila možu Borutu, ki me je tekom študija bodril s svojo potrpežljivostjo in optimizmom ter vseskozi verjel, da mi bo kljub materinstvu uspelo dokončati študij.

Posebej velika zahvala gre mojim staršem, predvsem mami Janji, ki mi je z vsakodnevnim varstvom štiriletne hčerke Nike in dve leti starega sina Blaža, veliko pripomogla k želji po čimprejšnjem dokončanju študija.

Izjavljam, da sem diplomsko delo izdelala sama, pod mentorstvom prof. dr. Marije Sollner Dolenc, mag. farm..

Simona Jurčič

Ljubljana, april 2010

Predsednik komisije: prof. dr. Stane Srčič, mag. farm.

Članica komisije: izr. prof. dr. Irena Mlinarič Raščan, mag. farm.

VSEBINA

POVZETEK	VI
SEZNAM OKRAJŠAV	VII
1 UVOD	1
1.1 REACH sistem in alternativne metode določanja toksičnosti kemikalij	1
1.2 QSAR modeli	4
1.2.1 Zgodovinsko ozadje	4
1.2.2 Definicije in opis metod za napovedovanje toksičnih lastnosti spojin	5
1.2.3 Delitev QSAR modelov	6
1.3 Izgradnja QSAR modelov	7
1.3.1 Tradicionalen način izgradnje	7
1.3.2 Modeliranje z odločitvijo (Consensus modeling)	11
1.4 Uporaba QSAR modelov	12
1.4.1 Splošna uporaba	12
1.4.2 Uporaba v regulativi	12
1.4.3 Uporaba v toksikologiji	13
1.5 Ekspertni sistemi	13
1.5.1 Opis ekspertnih sistemov	13
1.5.2 Delitev ekspertnih sistemov	14
1.5.3 Razlike med ekspertnimi sistemi	14
1.5.4 Napovednost ekspertnih sistemov	16
1.6 Razpoložljiva programska oprema	19
1.6.1 Programska oprema za razvoj QSAR modelov	19
1.6.2 Programska oprema za napovedovanje specifičnih toksičnih lastnosti	20
1.6.3 Veljavna, prosto odstopna programska oprema v Evropski uniji	20
1.7 Toxtree	21
1.7.1 Opis in struktura programa / aplikacije	21
1.7.2 Analizna metoda za napovedovanje toksičnih lastnosti	21
1.7.3 Mutagenost in kancerogenost spojin	22
1.7.4 Osnova za delovanje modula za napovedovanje mutagenosti in kancerogenosti	23
1.7.5 Izidi modula za napovedovanje mutagenosti in kancerogenosti	25

2	NAMEN DELA	26
3	MATERIALI IN METODE	28
3.1	Izbor spojin za testiranje	28
3.2	Testiranje spojin	30
3.3	Analiza rezultatov testiranja	34
4	REZULTATI IN RAZPRAVA	36
4.1	Občutljivost za napovednost	37
4.2	Specifičnost za napovednost	44
4.3	Točnost za napovednost	50
4.4	Ugotavljanje prisotnosti komponente metabolične aktivacije v programskem modulu	52
5	SKLEP	55
6	LITERATURA	57
7	PRILOGA I	63
8	PRILOGA II	69

KAZALO PREGLEDNIC

Preglednica 1: Nekateri od toksičnih lastnosti spojin, ki jih lahko napovedujemo z različnimi ekspertnimi sistemi.....	15
Preglednica 2: Rezultati testiranja programov TOPKAT, CASETOX in DEREK, za namen ovrednotenja le teh z vidika občutljivosti, specifičnosti in točnosti, za napovedovanje mutagenosti spojin.....	18
Preglednica 3: Rezultati testiranja spojin prvega niza.....	38
Preglednica 4: Razvrstitev spojin po strukturnih elementih, ki jih prepozna Toxtree.....	40
Preglednica 5: Rezultati testiranja spojin drugega niza.....	45
Preglednica 6: Razvrstitev drugega niza spojin po strukturnih elementih, ki jih prepozna Toxtree.....	47
Preglednica 7: Spojine prvega niza, ki se morajo predhodno metabolično aktivirati, da postanejo mutagene in kancerogene.	54

POVZETEK

Svet je dandanes preplavljen s celo paleto najrazličnejših kemikalij, ki jih dnevno srečujemo praktično na vseh področjih našega delovanja. Nekatere so bolj, druge pa manj nevarne za naše zdravje in druge žive organizme. Zaradi njihove velike porabe in vse večjih potreb po ugotavljanju tveganja ter ocenjevanju njihove varnosti, se je tako v Evropski uniji 1. junija 2008 začela v celoti izvajati nova uredba (EC) No 1907/2006 o registraciji, evalvaciji, avtorizaciji in omejevanju kemikalij (REACH), ki je namenjena zagotavljanju dodatnih informacij o kemikalijah, varni uporabi kemikalij in konkurenčnosti evropske industrije. V interesu časa, nižjih stroškov in blaginje živali REACH predvideva obširno uporabo alternativnih metod za napovedovanje toksičnosti spojin. Metode za napovedovanje toksičnih lastnosti spojin vključujejo med drugim tudi različne SAR in QSAR modele ter računalniško osnovane ekspertne sisteme. Ekspertni sistemi, prav tako kot SAR in QSAR modeli, omogočajo napovedovanje toksičnih lastnosti spojin na osnovi kemijske strukture. Ker se med seboj razlikujejo v načinu določanja potencialne toksičnosti spojin in ker so zgrajeni na popolnoma različnih učnih nizih spojin ter bazah pravil, imajo posledično tudi različno stopnjo napovednosti.

V okviru diplomske naloge smo vrednotili napovednost preko spleta prosto dostopnega računalniškega programa Toxtree, ki je odobren s strani Evropske komisije, natančneje njegovega programskega modula, ki se nanaša na vrednotenje kancerogenosti in mutagenosti spojin (Benigni / Bossa rulebase for mutagenicity and carcinogenicity). Napovednost smo vrednotili iz vidika občutljivosti, specifičnosti in točnosti programskega modula, pri čemer smo testirali 68 spojin, ki smo jih razdelili v dva testna niza. Ob enem smo preverili tudi, kako ustrezno programski modul ovrednoti spojine, ki šele po predhodni metabolični aktivaciji postanejo mutagene in kancerogene. Po končani analizi rezultatov testiranja spojin smo izračunali, da znaša občutljivost proučevanega programskega modula, ki jo izražamo kot odstotek biološko aktivnih-pozitivnih spojin, ki jih programski modul pravilno opredeli kot mutagene in kancerogene, 90,3 %, specifičnost, ki predstavlja odstotek biološko neaktivnih-negativnih spojin, ki jih programski modul pravilno opredeli kot nekancerogene in nemutagene, 88,9 %, točnost, ki je izražena kot odstotek vseh spojin, ki jih modul pravilno opredeli, 89,7 %, točnost napovedi za spojine, ki se metabolično aktivirajo pa 84,2 %.

SEZNAM OKRAJŠAV

2-AAF	2-acetilaminofluoren
AD	Applicability Domain (uporabna domena)
ANN	Artificial Neural Network (umetne nevronske mreže)
ASCII	American Standard Code for Information Interchange (Ameriški standardni nabor za izmenjavo informacij)
CPDB	The Carcinogenic Potency Database (podatkovna baza o rakotvornosti spojin)
DAB	N,N-dimetil-4-aminoazobenzen
DEN	N,N-dietilnitrozamin
DT	Decision Tree (odločitveno drevo)
EC	European Commission (Evropska komisija)
ECHA	European Chemicals Agency (Evropska agencija za kemikalije)
ECVAM	European Centre for the Validation of Alternative Methods (Evropski center za validacijo alternativnih metod)
FFBPNN	Feedforward Backpropagation Neural Network (nevronske mreže z vzvratnim širjenjem napake)
GRNN	General Regression Neural Network (regresijske nevronske mreže)
HF	Hierarchical framework (hierarhični delovni okvir)
IARC	International Agency for Research on Cancer (Mednarodna agencija za raziskovanje raka)
IE	Inference Engine (mehanizem sklepanja)
IHCP	Institute for Health and Consumer Protection (Inštitut za zdravje in varovanje potrošnikov)

JRC	Joint Research Centre (Skupno raziskovalno središče Evropske unije)
KB	Knowledge-Base (baza znanja)
kNN	K-Nearest Neighbor (metoda najbližjih sosedov)
LDA	Linear Discriminant Analysis (linearna diskriminacijska analiza)
LFER	Linear free energy relationship (povezava z linearno prosto energijo)
LR	Logistic Regression (logistična regresija)
MLR	Muliple Linear Regression (multipla linearna regresija)
NNK	4-(metilnitrozamino)-1-(3-piridil)-1-butanon
OECD	Organisation for Economic Co-operation and Developement (Organizacija za gospodarsko sodelovanje in razvoj)
PLS	Partial Least Squares (regresija delnih najmanjših kvadratov)
PNN	Probabilistic Neural Network (verjetnostne nevronske mreže)
QSAR	Qantitative Structure-Activity Relationship (kvantitativen odnos med strukturo in delovanjem)
QSTR	Quantitative Structure Toxicity Relationship (kvantitativen odnos med strukturo in toksičnim delovanjem)
REACH	Registration, Evaluation, Authorisation and Restriction of Chemicals (registracija, evalvacija, avtorizacija in omejevanje kemikalij)
SA	Structural Alert (strukturni element)
SAR	Structure-Activity Relationship (odnos med strukturo in delovanjem)
SLR	Simple Linear Regression (enostavna linearna regresija)

TCDD	2,3,7,8-tetraklorodibenzo-p-dioksin
UNCED	United Nations Conference on Environment and Development (Konferenca Združenih narodov o okolju in razvoju)

1 UVOD

1.1 REACH sistem in alternativne metode določanja toksičnosti kemikalij

Današnja moderna industrializirana družba je za svoj obstoj vedno bolj odvisna od kemikalij. Medtem ko jih je dnevno na različnih področjih našega delovanja v uporabi veliko število, mnoge nove na številnih tehnoloških področjih (tudi v farmaciji), raziskovalci še odkrivajo in razvijajo (1). Prav zaradi omenjenih dejstev je družba v zadnjih nekaj desetletjih postala upravičeno zaskrbljena glede njihovih možnih vplivov na okolje in zdravje človeka (2). Tako je bil že leta 1992 na Konferenci Združenih narodov o okolju in razvoju (UNCED) v Rio de Janeiro-u dosežen dogovor glede akcijskega načrta za nadaljnji razvoj na področju kemikalij in nastala je Agenda 21 (3,4). V njej so natančno predstavili upravljanje s kemikalijami in povečali mednarodno potrebo po ugotavljanju tveganja pri izpostavljanju nevarnim snovem. Najpomembnejši korak, ki je vodil od besed k dejanjem na področju Evropske Unije, pa je bil narejen februarja leta 2001, ko je Evropska komisija izdala Belo knjigo (White Paper) z naslovom, Strategija za prihodnost kemijske politike (Strategy for a Future Chemical Policy) (5). Ta je bila pozneje tudi odobrena s strani članic Evropske Unije (6,7). Glavni namen Bele knjige je bil zagotoviti visoko stopnjo zaščite okolja in zdravja ljudi, ob sočasno nemotenem stopnjevanju inovativnosti in tekmovalnosti znotraj evropske kemijske industrije. V njej je naveden tudi predlog o ustanovitvi novega sistema **REACH** (**R**egistration, **E**valuation, **A**uthorisation and **R**estriction of **C**hemicals), ki bi obravnaval tako že obstoječe spojine (to so tiste, ki so se na trgu pojavile že pred 18. 9. 1981 in imajo pomanjkljive podatke o oceni varnosti) kot tudi nove spojine, ki so na tržišče prišle po prej omenjenem datumu (6). Tako je 29. oktobra 2003 Evropska komisija že sprejela zakonodajni predlog za ustanovitev REACH-a, 18. 12. 2006 pa je izšla uredba (EC) No 1907/2006 o registraciji, evalvaciji, avtorizaciji in omejevanju kemikalij (REACH) (8). 1. junija 2007 je omenjena uredba stopila v pravo veljavo, leto kasneje, 1. junija 2008, pa se je začela v celoti izvajati. Uredba velja za vse nevarne snovi kot tudi za tiste, ki niso označene kot nevarne (9). Sistem REACH je namenjen zagotavljanju dodatnih informacij o kemikalijah, varni uporabi kemikalij in konkurenčnosti evropske industrije (10). V interesu časa, nižjih stroškov in blaginje živali, REACH predvideva obširno uporabo ocenjevalnih metod, kot so tudi QSAR metode in metode grupiranja (2,8,11).

V okviru nove zakonodaje sta danes proizvodnja in uvoz snovi (snov kot taka ali v pripravkih) v količinah ene tone ali več letno na proizvajalca oziroma uvoznika mogoča le, če je snov registrirana pri Evropski agenciji za kemikalije (ECHA), ali pa je bila snov pred-registrirana med 1. 6. 2008 in 1. 12. 2008 (9). Evropska agencija za kemikalije, s sedežem v Helsinkih, je namenjena upravljanju s tehničnimi, znanstvenimi in administrativnimi vidiki sistema REACH. V ta namen izvaja postopke registracije, evalvacije, avtorizacije in omejevanja kemikalij ter tako zagotavlja skladnost v vsej Evropski uniji (12).

Nov sistem kontrole kemikalij (REACH) torej pospešuje razvoj ter uzakonitev alternativnih metod in omejuje uporabo živali v testiranjih. V današnjem času vse pomembnejši poudarek na etično upravičljivi uporabi živali v poskusih vključuje tudi težnje po izboljšanju dobrobiti laboratorijskih živali in doslednem upoštevanju 3R načel (angl. reduction – zmanjšanje, refinement – izboljšanje, replacement – zamenjava) pri razvoju testnih metod. Koncept **3R** predstavljajo tri načela, in sicer: **REDUCTION** (izbrati poskus, ki zahteva najmanjše število živali), **REFINEMENT** (izboljšati pogoje bivanja živali in postopke, ki se izvajajo na živalih tako, da bodo utrpele čim manj nelagodja, stresa, bolečin ali trpljenja) in **REPLACEMENT** (zamenjati predvideno metodo z alternativno, ki da enak rezultat oziroma enako količino podatkov brez uporabe živali) (13,14).

Alternativna metoda je katerakoli metoda, ki lahko nadomesti, zmanjša ali spremeni uporabo testov na živalih (*in vivo*) v biomedicinskih raziskavah, testiranju ali izobraževanju. Alternativne metode temeljijo na integriranih in inteligentnih testnih strategijah in vključujejo različne *in silico* napovedne metode (SAR in QSAR modeli, ekspertni sistemi) in sofisticirane *in vitro* postopke/teste (testi na izoliranih tkivih in organih, perfuzijski sistemi, ki bazirajo na celičnih kulturah, testi na različnih celičnih linijah, testi na bakterijskih sevih, itd.) (14,15). Danes lahko nekatera testiranja na živalih v celoti nadomestimo z uporabo ene alternativne metode (npr. test na kožno korozivnost in test fototoksičnosti), medtem ko je določanje systemske (med katere sodita tudi karcinogenost in reproduktivna toksičnost) in kronične toksičnosti veliko zahtevnejše in je zato testiranje na živalih zaenkrat edina ustrezna metoda (tovrstne teste na živalih lahko zamenjamo samo z integriranimi testnimi strategijami, ki vključujejo QSAR modele, metode grupiranja in *in vitro* metode, in sicer z namenom pridobiti nekatere podatke za

napovedovanje in ocenjevanje tveganja) (14). Pogosto zahteva določanje nekaterih toksičnih učinkov uporabo integriranih testnih strategij, kot je na primer kombinacija *in vitro* ter *in silico* metod.

Alternativne teste danes že uporabljamo za napovedovanje kožne iritacije in korozije (EpiskinTM, EpidermTM, TER test–test transkutane električne resistance podganje kože, itd.), kožne senzibilnosti (LLNA test– Local Lymph Node Assay), mutagenosti (*in vitro* testi za genske (točkovne) mutacije na bakterijah npr. Ames-ov test; *in vitro* testi kromosomskih aberacij na sesalčjih celicah; itd.) in še nekaterih drugih toksičnih lastnosti (<http://ecb.jrc.ec.europa.eu/test-methods/annex5/>). Ti testi vseeno vedno niso uporabni kot alternativa *in vivo* testiranjem, ampak lahko z njihovo pomočjo največkrat le zmanjšamo obseg *in vivo* testov.

Z validacijo alternativnih metod se v Evropski uniji ukvarja Evropski center za validacijo alternativnih metod (ECVAM), ki je bil ustanovljen leta 1992 in deluje pod okriljem Inštituta za zdravje in varovanje potrošnikov (IHCP) znotraj Skupnega raziskovalnega središča Evropske unije (JRC) in širše, znotraj Evropske komisije (EC) (14).

1.2 QSAR modeli

1.2.1 Zgodovinsko ozadje

Začetki razvoja QSAR modelov segajo v sredino 19. stoletja in v začetek 20. stoletja, ko so odkrili fizikalne osnove kemijskih lastnosti spojin in fizikalne osnove interakcij molekul z biološkimi sistemi (16). Že pred letom 1858 je ruski kemik Borodin dejal, da so toksikološke lastnosti in kemijska struktura spojin tesno povezane. Kmalu za tem sta Julius Lothar Meyer in Dmitriy Mendeleev na osnovi podobnosti razvrstila takrat znane elemente v periodni sistem elementov. Razvrščanje elementov po skupinah je predstavljalo nekakšno osnovo za napovedno kemijo. Tako so preko poznavanja enostavnih lastnosti spojin tvorili enostavna pravila (1).

Sledilo je odkritje empiričnih korelacij med toksičnostjo in enostavnimi fizikalno-kemijskimi lastnostmi na majhnem nizu spojin. Leta 1863 je Cros opazil obratno sorazmerno zvezo med toksičnostjo alkohola pri sesalcih in topnostjo v vodi; leta 1869 sta Crum-Brown in Frazer ugotovila, da so paralizirajoče lastnosti niza kvaterniranih strihninov odvisne od narave kvarterne skupine; Richardson je poročal o odnosu med narkotičnim delovanjem alkoholov in molekulsko maso; 1893 je Richet opazil odnos med toksičnostjo enostavnih organskih spojin in vodotopnostjo; Meyer in Overton pa sta neodvisno dokazala, da je narkotična aktivnost spojin odvisna od porazdelitvenega koeficienta med oljem in vodo. Tako so znanstveniki na osnovi empiričnih fizikalno-kemijskih parametrov spojinam določali toksične lastnosti (1).

Glavni konceptualni okvir za fizikalno osnovo interakcij med ksenobiotiki in biološkimi makromolekulami, na osnovi katerega se je razvil SAR, pa predstavljata, na prelomu 20. stoletja postavljena, valenčna teorija formacije kemijske vezi Gilbert Newton Lewis-a in Langley-eva teorija receptorjev. Nato je leta 1930 Hammet na podlagi povezave z linearno prosto energijo (LFER) zagotovil termodinamsko osnovo parametričnih postopkov, ki določajo SAR. Deset let kasneje pa sta Taft in Corwin Hansch z uporabo steričnih in hidrofilnih deskriptorjev potrdila veljavnost LFER in tako zaznamovala dejanski začetek razvoja kvantitativnih odnosov med strukturo in delovanjem (QSAR). Z različnimi QSAR modeli je bilo tako mogoče napovedovati fizikalno-kemijske, farmakološke in toksikološke lastnosti spojin. Kmalu za tem je leta 1962 Corwin Hansch objavil prvi formalni dokument o QSAR. V njem je poudaril vlogo porazdelitvenega koeficienta med oktanolom in vodo. Kasneje pa je objavil še metodo za izračun hidrofobnih fragmentnih konstant. Hanscha tako štejemo za očeta modernega QSAR (16).

Ko so s toksičnimi lastnostmi spojin in z eksperimentalno določenimi substituentnimi konstantami potrdili Hanschevo hipotezo, so se raziskovalci začeli ukvarjati s problemom identifikacije dodatnih kemijskih, strukturno osnovanih deskriptorjev in z odkrivanjem novih pristopov, kako s pomočjo izračunanih topoloških in geometričnih lastnosti, učinkovito opisati strukturo molekul. Toksične lastnosti spojin so torej začeli določati neposredno s pomočjo teoretičnih molekulskih deskriptorjev (1).

Razvijati in uporabljati so začeli bolj prefinjene statistične metode, kot so regresija delnih najmanjših kvadratov, umetne nevronske mreže (ANN), itd.. Na novo pa so odkrili tudi številne nove molekulske deskriptorje in računske pristope, ki jih še danes uporabljamo za napovedovanje različnih toksičnih lastnosti spojin (16).

1.2.2 Definicije in opis metod za napovedovanje toksičnih lastnosti spojin

Metode za napovedovanje toksičnih lastnosti spojin vključujejo formalizirane metode, med katere spadajo različni SAR in QSAR modeli ter računalniško osnovani ekspertni sistemi (3,17). Z njimi lahko na osnovi povezave med strukturo in delovanjem neke skupine spojin, napovemo biološko aktivnost njim podobnih spojin.

SAR in QSAR sta teoretična modela, ki ju lahko uporabljamo za napovedovanje fizikalno-kemijskih lastnosti in bioloških lastnosti molekul, na osnovi poznavanja njihove kemijske strukture (3,14). Razvili so ju na osnovi eksperimentalnih podatkov modelnih substanc (17).

SAR je kvalitativna povezava med kemijsko strukturo in delovanjem spojine, ki vsebuje strukturo z izkazanim določenim biološkim učinkom (16). Kvalitativen odnos je izpeljan iz podatkov, ki omogočajo binarno napoved (»da« in »ne« podatki oziroma podatki, ki nam povedo ali je spojina na primer kancerogena ali ni) (18). SAR nam omogoča določitev farmakofornega dela spojine.

QSAR je matematični model, ki z določeno enačbo kvantificira odnos med kvantitativno opredeljenimi (izmerjenimi ali izračunanimi) fizikalno-kemijskimi lastnostmi (molekulski deskriptorji) spojine in njenim številčno opredeljenim (izmerjenim) biološkim učinkom (npr. toksična lastnost spojine). Biološki učinek je torej funkcija strukturne spremembe (16). Kvantitativen odnos je v tem primeru osnovan na podatkih o jakosti/moči toksičnosti (podatki so tu izraženi s številčnimi vrednostmi), s katerimi lahko znanje o toksičnosti spojin še nadgradimo (18).

QSAR model običajno vsebuje tri glavne elemente:

- zunanjo obliko spojine (molekulski deskriptorji in strukturni fragmenti),
- lastnost ali učinek spojine, ki ga napovedujemo (toksična lastnost spojine ali odgovor)
- algoritem, ki pretvori molekulski deskriptor v napoved, z modelom določenih toksičnih lastnosti spojine (3).

V znanstveni literaturi je bilo do danes objavljenih že več kot 20.000 QSAR modelov. Vzporedno z njimi pa so kot tržne produkte razvili ekspertne sisteme (3).

Ekspertni sistemi predstavljajo raznovrsten nabor računalniško osnovanih ocenjevalnih metod, ki so izdelane na osnovi enotne uporabe baz podatkov (eksperimentalni podatki o toksičnosti) in baz pravil izpeljanih iz teh podatkov. Slednja pravila so lahko osnovana na matematični indukciji oziroma na statističnem pristopu (npr. princip QSAR modelov) in/ali na obstoječem strokovnem znanju (opis aktivnosti s SAR modeli ter druga pravila sklepanja) (3).

1.2.3 Delitev QSAR modelov

❖ Glede načina izgradnje QSAR modela

QSAR modele lahko glede načina izgradnje in namena uporabe kategoriziramo na tri tipe:

- modeli grupiranja,
- klasifikacijski modeli,
- regresijski modeli (16).

QSAR modeli grupiranja spojin pri preiskovanju podatkovnih vzorcev uporabljajo nenadzorovane učne tehnike na osnovi deskriptorjev, brez prištevanja teh vzorcev specifični kategoriji toksičnih lastnosti spojin.

V nasprotju z modeli grupiranja so klasifikacijske metode nadzorovane učne tehnike, ki grupirajo spojine v znane kategorije toksičnih lastnosti spojin (npr. zelo aktivne, zmerno aktivne in neaktivne).

Regresijski modeli pa poskušajo vpeljati kvantitativen odnos med strukturo in delovanjem določene spojine, in sicer preko številnih linearnih regresijskih metod ter zadnje čase tudi preko umetnih nevronske mrež (ANN).

Največkrat uporabljeni so klasifikacijski in regresijski modeli.

❖ Glede odnosa v QSAR modelu

Odnos v QSAR modelu lahko predstavlja eno od treh oblik:

- teoretični QSAR model, pri katerem enačba bazira na osnovnih fizikalnih načelih; model je omejen na fizikalno-kemijske in kinetične lastnosti,
- statistični (empirični) SAR ali QSAR model, ki je bil razvit z uporabo statističnih metod na učnem nizu spojin; številni statistični modeli imajo teoretično osnovo in vključujejo QSAR modele razvite po Hanschevem pristopu,
- QSAR model osnovan na pravilih sklepanja, ki so določena z izkustvi/dokazi ali pa temeljijo na strokovnem znanju/presoji (3).

1.3 Izgradnja QSAR modelov

Razvoj QSAR modelov zahteva strokovno znanje molekulskega modeliranja in statistike, v povezavi s strokovnim znanjem, ki bazira na razumevanju kemije, biologije in toksikologije.

V grobem ločimo tradicionalen način izgradnje QSAR modelov in modeliranje z odločitvijo (consensus modeling). Modele, ki nastanejo pri posameznem tipu izgradnje, lahko potem kombiniramo in združujemo med seboj v hierarhične delovne okvirje (HF) (16).

1.3.1 Tradicionalen način izgradnje

Tradicionalen način izgradnje QSAR modelov se začne z zbiranjem pomembnih bioloških podatkov iz zanesljivih podatkovnih baz. Sledi izločanje podatkov nizke kvalitete, ki bi lahko negativno vplivali na kakovost modela. Naslednji korak je izbira reprezentativnih spojin v učni in validacijski niz, za postopno kalibriranje in validacijo QSAR modela (16). Nato za proučevane spojine izračunamo molekulske deskriptorje (odvisne spremenljivke), ki ponazarjajo fizikalno-kemijske in strukturne lastnosti le teh (18). Tiste deskriptorje, ki so nepotrebni in vsebujejo malo informacij, izločimo še pred procesom modeliranja (16). Sledi izbor določene lastnosti (v našem primeru ene izmed toksičnih lastnosti), za določanje katere bo model namenjen (neodvisna spremenljivka) (18). Naslednjo stopnjo pri razvoju modela, ki povezuje biološke lastnosti s fizikalno-kemijskimi in strukturnimi lastnostmi spojin, predstavljajo metode za analizo podatkov ali statistične regresijske metode (le te korelirajo spremembe v strukturi s spremembami tarčne lastnosti). Med procesom modeliranja hkrati izboljšujemo nujne parametre analizne metode in izbiramo

pomembne skupine deskriptorjev. Za izgradnjo končnega QSAR modela potrebujemo optimalen niz parametrov analize metode in molekulskih deskriptorjev, zato posamezne cikle, v katerih določamo lastnosti predhodno s testi že okarakteriziranih spojin, večkrat ponovimo (16). Izdelan QSAR model mora nato biti podvržen različnim validacijskim metodam. Ko je QSAR model validiran, ga lahko uporabimo za napovedovanje z modelom točno določene lastnosti oziroma več določenih lastnosti spojin, ki se nahajajo v domeni učnega in validacijskega niza (16).

❖ Podatkovne baze

Pri izgradnji QSAR modela je zelo pomemben izbor kvalitetnih bioloških podatkov iz zanesljivih in verodostojnih podatkovnih baz. Številne podatkovne baze so dostopne preko interneta (»on-line«; med njimi je veliko prosto dostopnih). To zelo olajša iskanje in ponoven priklic informacij o spojinah in njihovih strukturnih analogih (3).

V nadaljevanju so navedene nekatere podatkovne baze in spletne strani, na katerih jih lahko najdemo. To so:

- Danish (Q)SAR database (<http://ecb.jrc.ec.europa.eu/>),
- ChemFinder (<http://www.chemfinder.com>),
- ChemIDPlus (<http://toxnet.nlm.nih.gov>),
- DSSTox (Distributed Structure-Searchable Toxicity) (<http://www.epa.gov/ncct/dsstox/index.html>),
- Ambit (<http://ambit.acad.bg>),
- ISSCAN (http://www.epa.gov/NCCT/dsstox/sdf_isscan_external.html),
- Carcinogenic Potency Database (CPDB) (<http://potency.berkeley.edu/cpdb.html>),
- Leadscope (<http://www.leadscope.com>).

❖ Molekulski deskriptorji

Izgradnja QSAR modelov vedno zahteva generiranje molekulskih deskriptorjev. Za posamezno spojino lahko danes neposredno iz kemijske strukture izračunamo več sto ali celo več tisoč molekulskih deskriptorjev. To nam omogočajo enostavni in uporabnikom prijazni računalniški programi. Splošno uporabne programe vključujejo: EPI Suite, DRAGON, Adriana.Code, Molconn-Z, TSAR, MDL QSAR, MOPAC, QSAR Builder, itd. (3).

Molekulske deskriptorje delimo na pet skupin:

- konstitucijski (kemijska sestava spojine: molekulska masa, število atomov/vezi in število akceptorskih/donorskih vodikovih vezi, itd.),
- topološki (2D strukturna formula, razporeditev atomov v molekuli: Kier-Hallovi indeksi/kazalci, obseg razvejanosti),
- geometrijski (3D struktura molekule, odražajo dejansko velikost molekule: molekulska prostornina, površina molekule, topilom dostopno področje, področji polarnosti in nepolarnosti),
- elektrostatski (porazdelitev naboja: delni naboji atomov, elektronegativnost),
- kvantnokemijski (elektronska struktura: HOMO-LUMO energije, ionizacijski potencial).

Pri izgradnji QSAR modelov nas zanimajo tisti molekulske deskriptorji, ki korelirajo s fizikalno-kemijskimi lastnostmi, biološko aktivnostjo in toksičnostjo tarčnih spojin. Ko imamo izbrane ustrezne molekulske deskriptorje, lahko izdelamo QSAR model, ki bo napovedoval želeno lastnost spojin (16).

❖ Metode za analizo/obdelavo podatkov

Metode za analizo podatkov ločimo na tiste za reševanje regresijskih problemov in na tiste za reševanje klasifikacijskih problemov (16).

a) analize metode za reševanje regresijskih problemov:

- enostavna linearna regresija (SLR),
- multipla linearna regresija (MLR),
- regresija delnih najmanjših kvadratov (PLS),
- nevronske mreže z vzratnim širjenjem napake (FFBPNN),
- regresijske nevronske mreže (GRNN).

b) analize metode za reševanje klasifikacijskih problemov:

- linearna diskriminacijska analiza (LDA),
- logistična regresija (LR),
- metoda odločitvenega drevesa (DT),
- metoda najbližjih sosedov (kNN),
- verjetnostne nevronske mreže (PNN).

Multipla linearna regresija in logistična regresija sta tradicionalni in tudi najpogostejše uporabljeni metodi za analizo podatkov; medtem ko na primer regresijske nevronske mreže spadajo med novejšje metode (16).

❖ Validacija

Validacija QSAR modela je postopek, pri katerem vrednotimo napovednost in mehanistično interpretacijo modela za določen namen. Pri tej definiciji napovednost obsega primernost, robustnost in sposobnost napovedovanja; namen pa predstavlja toksično lastnost spojine in uporabno domeno (3). Če model ni validiran, ima zelo majhno uporabno vrednost za napovedovanje lastnosti neokarakteriziranih spojin.

Model lahko validiramo s pomočjo interne/prečne/navskrižne validacije in eksterne validacije. Čeprav imata oba postopka v splošnem številne skupne značilnosti, se razlikujeta v zmožnosti in učinkovitosti pri ocenjevanju celokupne napovedne točnosti modela, uporabne domene modela in možnosti korelacije med izvrševanjem (uporabo modela) (16).

Kadar imamo zagotovljeno zadostno število podatkov, lahko model validiramo z napovedovanjem lastnosti spojinam, ki imajo nam poznan učinek in jih nismo uporabili pri izgradnji modela. Takšnemu postopku pravimo eksterna validacija in je povsem podoben uporabi modela v realnosti. Glavni predpogoj za sprejemljivo eksterno validacijo je ustrezno načrtovanje testnega niza spojin, v smislu velikosti in raznolikosti ter ustreznosti za predvideno uporabo. Nesmiselno je namreč načrtovati testni niz, na da bi vedeli v kakšne namene bomo QSAR model uporabljali. Najpogostejše se za določitev testnega niza pri eksterni validaciji uporabi naključen izbor določenega števila spojin iz baze podatkov. Eksterna validacija v splošnem dobro zagotavlja vrednotenje celokupne napovednosti modela, majhnega pomena pa je pri vrednotenju uporabne domene.

Pri interni validaciji pa del spojin iz učnega niza izločimo in jim nato z modelom, zgrajenim na osnovi ostalih spojin (tistih, ki niso bile izločene iz učnega niza), napovedujemo lastnosti. Če je vsaka spojina med postopkom izločena enkrat in proces ponovimo za vsako spojino, tak postopek imenujemo validacija z izpuščanjem objektov (leave-one-out). Če učni niz razdelimo na N skupin s približno enakim številom spojin in proces ponovimo za vsako skupino, takemu postopku pravimo N-kratna navskrižna

validacija. Največkrat je za vrednotenje napovednih sposobnosti klasifikacijskega modela uporabljena 10-kratna navskrižna validacija. Rezultati navskrižne validacije so odvisni od naključne razdelitve baze podatkov, zato jo je priporočljivo ponoviti večkrat. Povprečni rezultat multiple navskrižne validacije zagotavlja nepristransko oceno modelne napovednosti. V primerjavi z eksterno validacijo, navskrižna validacija zagotavlja sistematsko merjenje modelne zmogljivosti brez izgube niza spojin za testiranje. (16)

Glede validacije je Organizacija za gospodarsko sodelovanje in razvoj (OECD) marca 2002 v Setubal-u na Portugalskem predlagala šest načel. Nato pa jih je novembra leta 2004 na politični ravni sprejela pet (dva od prejšnjih šestih načel je skupina strokovnjakov združila v enega). OECD je ta načela postavila zato, da bi pospešila uporabo QSAR modelov v regulatorne namene. QSAR model naj bi vseboval naslednje informacije:

- definirano toksično lastnost spojine (endpoint),
- nedvoumen oz. jasen algoritem,
- definirano uporabno domeno,
- ustrezno stopnjo prileganja, robustnosti in napovednosti,
- mehanistično informacijo, če je možno (3,19).

1.3.2 Modeliranje z odločitvijo (Consensus modeling)

Modeliranje z odločitvijo se zadnja leta razvija predvsem na področju statistike z namenom izboljšanja posameznih napovedi, in sicer s kombiniranjem multiplih individualnih modelov. Odobrena domneva modeliranja z odločitvijo je, da je SAR mnogo bolje identificiran in kodiran z multiplimi modeli kot s posameznimi modeli. Te prednosti pa lahko izkoristimo samo, če so modeli med seboj neodvisni. To pomeni, da kodirajo kompleksen odnos med kemijsko strukturo in biološko aktivnostjo na različne načine. Tako naj bi bil idealen kombiniran sistem sestavljen iz majhnega števila točnih in neodvisnih modelov (16).

Izgradnja posameznih modelov lahko bazira na:

- različnem kreiranju učnega niza (metoda ponovnega vzorčenja - resampling metoda, metoda bagging and boosting),
- izboru različnih skupin deskriptorjev (metodi odločitvenih dreves imenovani metoda Random Forest in metoda Decision Forest) (16).

1.4 Uporaba QSAR modelov

1.4.1 Splošna uporaba

QSAR modeli nam omogočajo napovedovanje bioloških lastnosti kemijsko že okarakteriziranih spojin, ki še niso bile biološko testirane, v veliko pomoč pa so nam tudi pri pridobivanju informacij o molekulskih karakteristikah spojin, ki pomembno vplivajo na biološke lastnosti (16). QSAR modeli so tudi tako imenovani strukturni elementi ekspertnih sistemov, ki za napovedovanje toksičnih lastnosti spojin uporabljajo statističen pristop (npr. TOPKAT, MCASE, Leadscape) (1).

QSAR modele uporabljamo:

- v farmacevtski industriji (pri odkrivanju novih učinkovin in za izboljšanje lastnosti obstoječih spojin),
- v toksikologiji (za napovedovanje toksičnih lastnosti spojin),
- v regulativi (za razvrščanje netestiranih spojin pri bolj obširnih in dražjih evalvacijah) (16).

1.4.2 Uporaba v regulativi

Načeloma lahko QSAR modele pri implementaciji kemijske zakonodaje, torej v regulativi, uporabimo z namenom (3,16):

- zagotavljanja informacij za uporabo pri postopkih razvrščanja (priority setting),
- vodenja eksperimentalnega načrta testiranja ali testne strategije,
- izboljšanja evalvacije/vrednotenja obstoječih testnih podatkov (npr. QSAR modeli nam pomagajo izbrati veljavne in značilne podatke ter nam priskrbeti mehanistične informacije),
- da podpiramo grupiranje kemikalij v skupine ali kemijske družine (glede na strukturo, delovanje, mehanizem delovanja, toksičnost, itd.), tako da ni potrebno, da testiramo vse kemikalije iz iste skupine za vsako posamezno toksično lastnost spojine,
- zapolnjevanja manjkajočih podatkov, ki jih potrebujemo pri klasifikaciji in označevanju,
- zapolnjevanja manjkajočih podatkov, ki jih potrebujemo pri ocenjevanju tveganja,
- zagotavljanja informacij o fizikalno-kemijskih lastnostih spojin, stopnji toksičnosti, biokinetičnih procesih ter porazdelitvi in usodi v okolju.

1.4.3 Uporaba v toksikologiji

Na področju toksikologije lahko z različnimi QSAR modeli (ti so na omenjenem področju pogosto imenovani tudi QSTR - Quantitative Structure Toxicity Relationship modeli oz. modeli, ki opisujejo odnos med strukturo spojin in njihovo toksičnostjo) in ekspertnimi sistemi napovedujemo različne toksične lastnosti spojin (endpoint) (16).

Nekatere od njih so:

- mutagenost,
- kancerogenost,
- teratogenost,
- toksičnost za vodne organizme,
- biorazgradljivost,
- biokoncentracija,
- bioakumulacija,
- maksimalni tolerantni odmerki,
- akutna toksičnost,
- kardiotoksičnost (16).

Ob tem se moramo zavedati, da je posamezen QSAR model ali ekspertni sistem, ki je namenjen napovedovanju toksičnosti spojin, uporaben izključno samo za napovedovanje določenih toksičnih lastnosti spojin (le te posamezen model natančno določa) in to za skupine tistih spojin, ki so strukturno podobne spojinam, na osnovi katerih je bil ta model zgrajen (uporabna domena modela) (2). V nasprotnem primeru model ne daje ustreznih oziroma želenih rezultatov.

1.5 Ekspertni sistemi

1.5.1 Opis ekspertnih sistemov

Poleg QSAR modelov, se za ocenjevanje različnih toksičnih lastnosti spojin vse pogosteje uporabljajo tudi ekspertni sistemi (1,2). To so računalniško osnovani programi, ki omogočajo napovedovanje toksičnosti spojin na osnovi kemijske strukture. Številni so prosto dostopni tudi preko interneta. Kot je bilo že omenjeno, so izdelani na osnovi enotne uporabe baz podatkov (eksperimentalni podatki o toksičnosti) in baz pravil izpeljanih iz omenjenih podatkov, in sicer z metodo modeliranja s sklepanjem ali pa z uporabo njemu sorodnih pristopov. Na splošno ekspertni sistemi vsebujejo tri različne dele: skupek/izbor

podatkov in pravil, sistem za izvrševanje sklepanja in generator pravil (1). Prvi ekspertni sistem, o katerem so poročali, je bil TOX-MATCH (16).

Ekspertni sistemi vršijo logično sklepanje na ravni človeškega. Nanašajo se na matematične in logične tehnike ter obsegajo dve glavni komponenti:

- knowledge-base (KB; baza znanja)
- inference engine (IE; ustrezno logično sklepanje oz. mehanizem sklepanja) (1,16).

Ekspertni sistemi so torej pomembni umetni inteligentni sistemi. Dostopni so tudi ljudem, ki na tem področju niso strokovnjaki. Njihova uporaba prevladuje na številnih področjih: pri načrtovanju spojin, v analitski kemiji, pri pojasnjevanju kemijske strukture in še marsikje. Shranjeni so kot računalniško vidna zbirka vedno širše prepletenih IF - THEN (če – potem) konstruktov. Kadar moramo napovedati toksično lastnost neke spojine, ekspertni sistem razčleni njeno strukturo in jo primerja z razpoložljivimi pravili iz baze znanja. Pravila, ki jih dana struktura izzove so zbrana in predstavljena mehanizmu sklepanja, ki opravi še kvalitativno ocenjevanje potencialne toksičnosti (16).

1.5.2 Delitev ekspertnih sistemov

Glede na to, na kakšen način ekspertni sistem določa potencialno toksičnost in na čem temelji, ločimo tri tipe (1,3,16):

- ekspertni sistemi, ki temeljijo na QSAR pravilih in uporabljajo statističen pristop oz. statistične metode (**TOPKAT**, **LeadScope**),
- ekspertni sistemi, ki temeljijo na obstojećem strokovnem znanju in pravilih, ki so medsebojno neodvisna; le-ti na osnovi strukturnih elementov (SA) iščejo podobnosti med molekulami (**TOX-MATCH**, **DEREK**, **Derek for Windows**, **Oncologic**, **HazardExpert**) - knowledge-based oziroma rule-based sistemi,
- hibridni sistemi (**MCASE**, **TIMES**, **ECOSAR**).

Poleg navedenih treh vrst ekspertnih sistemov poznamo še tiste, ki kodirajo določene mehanistične procese (1,2). V to skupino spada ekspertni sistem **COMPACT**. Ta zahteva in vključuje veliko več znanja o ključnih biokemijskih procesih delovanja spojin (2).

1.5.3 Razlike med ekspertnimi sistemi

Ekspertni sistem je iznajdljiv/heurističen (sposoben je sklepanja tako s pomočjo teorije kot tudi s pomočjo strokovnega znanja), jasen (razloži svoj način sklepanja) in prilagodljiv

(lahko v obstoječo bazo vključi novo znanje) (1). Med seboj se razlikujejo v naslednjih sposobnostih:

- nekateri imajo sposobnost, da sami tvorijo pravila (CASE), drugi pa se opirajo na strokovno znanje (slednje velja za bolj zanesljivo),
- sposobnost napovedovanja toksičnosti metaboličnih produktov (CASE, TOPKAT),
- nekateri ekspertni sistemi zaznajo fragmente (biofobe), ki zmanjšajo toksičnost preiskovane spojine (CASE, TOPKAT); na ta način zmanjšajo možnost lažno pozitivnih rezultatov,
- kako se sistem odzove v primeru večjega števila aktivnih fragmentov; lahko uporabi največjo aktivnost, lahko izbiro prepusti uporabniku (DEREK), lahko pa molekulo v celoti prouči (COMPACT),
- sposobnost kvantifikacije toksičnega delovanja (HazardExpert)
- ali upoštevajo velikost strukturnega elementa (smatra se namreč, da so večji fragmenti manj pomembni, saj jih najdemo v premajhnem številu spojin),
- razlika med njimi je tudi v bazah podatkov in v bazah pravil (1).

Tudi s posameznimi ekspertnimi sistemi lahko napovedujemo samo določene toksične lastnosti spojin, zato bo v nadaljevanju predstavljeno, katere toksične lastnosti lahko napovemo z nekaterimi ekspertnimi sistemi (Preglednica 1) (3,16).

Preglednica 1: Nekateri od toksičnih lastnosti spojin, ki jih lahko napovedujemo z različnimi ekspertnimi sistemi.

	Mutag.	Kancerog.	Teratog.	Biorazgrad.	Iritirajoče za oči	Kožna senzibilnost
TOPKAT	x	x	x	x	x	X
MCASE	x	x	x	x	x	X
DEREK	x	x				X
OncoLogic		x				
HazardExpert	x	x	x			
TIMES	x					X
ToxTree	x	x		x	x	X

1.5.4 Napovednost ekspertnih sistemov

Na splošno velja, da so ekspertni sistemi v literaturi pomanjkljivo opisani, predvsem njihova struktura oziroma mehanizem logičnega sklepanja (IE), ki sicer predstavlja največjo razliko med ekspertnimi sistemi. Zato je primerjavo med njimi najlažje narediti na osnovi različnih napovednosti, in sicer za določeno toksično lastnost/endpoint, ki jo je posamezen ekspertni sistem sposoben napovedati. Največ avtorjev poroča o dobri napovednosti (točnosti) ekspertnih sistemov, ki je pogosto večja kot 90 % (1).

Napovednost lahko ovrednotimo z vidika točnosti, občutljivosti in specifičnosti (1). Za posamezen ekspertni sistem pa lahko izračunamo tudi njegovo uporabnost (20).

Točnost predstavlja odstotek vseh spojin, ki jih program pravilno opredeli.

$$\text{Točnost} = ((N_{A+S+} + N_{A-S-}) / N_{\text{eval}}) \times 100$$

Občutljivost predstavlja odstotek biološko aktivnih-pozitivnih spojin, ki jih program pravilno opredeli.

$$\text{Občutljivost} = (N_{A+S+} / N_{A+}) \times 100$$

Specifičnost pa je odstotek biološko neaktivnih-negativnih spojin, ki jih program pravilno opredeli

$$\text{Specifičnost} = (N_{A-S-} / N_{A-}) \times 100$$

$$\text{Uporabnost} = (N_{\text{eval}} / N_{\text{all}}) \times 100$$

Z omenjenimi formulami lahko določamo napovednost posameznega ekspertnega sistema za ocenjevanje genotoksičnosti. Oznake v navedenih formulah pa predstavljajo:

$N_{A+} \Rightarrow$ število spojin, ki pri Ames-ovem testu dajo pozitiven rezultat,

$N_{A-} \Rightarrow$ število spojin, ki pri Ames-ovem testu dajo negativen rezultat,

$N_{A+S+} \Rightarrow$ število spojin, ki dajo pozitiven rezultat tako pri Ames-ovem testu kot tudi pri testiranju z ekspertnim sistemom,

$N_{A-S-} \Rightarrow$ število spojin, ki dajo negativen rezultat tako pri Ames-ovem testu kot tudi pri testiranju z ekspertnim sistemom,

$N_{\text{eval}} \Rightarrow$ število vseh testiranih spojin z ekspertnim sistemom,

$N_{\text{all}} \Rightarrow$ celotno število spojin testnega seta, z objavljenimi podatki Ames-ovega testa (20).

Največjo oviro pri vrednotenju napovednosti ekspertnih sistemov predstavlja izbor ustreznih spojin za testiranje. Izbrati moramo namreč tiste, ki niso bile uporabljene pri izgradnji ekspertnega sistema. V nasprotnem primeru lahko dobimo previsoke ocene za napovednost (16). Prav to predstavlja velik problem pri določanju napovednosti ekspertnega sistema za napovedovanje mutagenosti zdravilnih učinkovin, saj je v praksi samo 10 % teh (zdravilnih učinkovin) testiranih in ovrednotenih kot bakterijski mutageni. Tako bi lahko vsak ekspertni sistem, ki bi ocenil vse testirane spojine kot negativne, imel 90 % točnost, kar pa ni ustrezno (16).

V eni od študij so raziskovalci ovrednotili in med seboj primerjali tri najpogosteje uporabljene komercialne ekspertne sisteme za napovedovanje mutagenosti spojin; TOPKAT, DEREK in CASETOX (nespremenljivi del programa MCASE) (21). Omenjene programe so ovrednotili na podlagi njihove sposobnosti za napovedovanje Ames-ovega testa mutagenosti. Pri tem so uporabili dva testna niza. Prvi testni niz je sestavljalo 520 lastnih spojin (kandidatov za zdravilne učinkovine), drugi testni niz pa je vseboval 94 komercialnih spojin. Iz obeh testnih nizov so izločili tiste spojine, ki so bile del učnega niza pri izgradnji TOPKAT-a in CASETOX-a; pri vrednotenju DEREK-a pa niso izločili nobene spojine. Iz testnih nizov so izločili tudi spojine, ki so vsebovale anorganske atome, ki jih program ni prepoznal, spojine, ki so bile glede na Ames-ov test nezanesljive, izločili pa so tudi kemijske mešanice več spojin. Vsakega od omenjenih treh ekspertnih sistemov so ovrednotili s stališča občutljivosti, specifičnosti in točnosti. Rezultati testiranja so navedeni v preglednici na naslednji strani (Preglednica 2) (21).

Preglednica 2: Rezultati testiranja programov TOPKAT, CASETOX in DEREK, za namen ovrednotenja le teh z vidika občutljivosti, specifičnosti in točnosti, za napovedovanje mutagenosti spojin.

Program	Testni niz	Št. sp.	Št. testiranih sp. (%)	Občutljivost (%)	Specifičnost (%)	Točnost (%)
TOPKAT	1	520	319 (61)	63 (32/51)	76 (204/268)	74
	2	94	68 (72)	59 (13/22)	80 (37/46)	74
CASETOX	1	520	213 (41)	50 (16/32)	86 (156/181)	81
	2	94	70 (75)	40 (8/20)	90 (45/50)	76
DEREK	1	520	520 (100)	28 (21/76)	80 (353/444)	72
	2	94	94 (100)	63 (19/30)	81 (52/64)	76

Kot je razvidno iz zgornje preglednice so se vsi trije programi izkazali kot slabo občutljivi za testni niz kandidatov za zdravilne učinkovine (nizka frekvenca mutagenov). Občutljivost je znašala 50 % za CASETOX in 63 % za TOPKAT. Izjema je bil ekspertni sistem DEREK, katerega občutljivost je bila le 28 %. Razlog za nizko občutljivost tega programa je bilo veliko število sorodnih spojin znotraj testnega niza, za katere program ni našel ustreznega pravila, ki bi omenjene spojine lahko povežalo z mutagenostjo in jih je zato označil kot negativne (16). Specifičnost programov je bila med 76 % in 90 % (odvisno od programa in testnega niza), točnost pa je znašala med 72 % in 81 % (odvisno od programa in testnega niza) (21).

Pomembna ugotovitev te študije je bila tudi, da so si napovednosti komercialno dostopnih programov za napovedovanje mutagenosti po Ames-u podobne, ne glede na to, ali testirano spojino uporabljamo v farmaciji, kmetijstvu ali industriji. Obenem so ugotovili, da kombiniranje programov izboljša točnost in specifičnost napovedi, medtem ko občutljivost ne preseže praga 50 % (21).

1.6 Razpoložljiva programska oprema

Danes je preko spleta dostopnih že veliko število računalniških programov. Ti so uporabni tako za razvoj QSAR modelov, kot tudi za napovedovanje specifičnih toksičnih lastnosti.

1.6.1 Programska oprema za razvoj QSAR modelov

Na tem področju obstaja veliko število tako komercialnih kot tudi prosto dostopnih računalniških programov. Nekateri so osredotočeni/specializirani za določeno analizno metodo (npr. metoda odločitvenega drevesa pri programu C4.5), medtem ko ostali pri svojem delovanju uporabljajo več analiznih metod (npr. ELECTRAS vsebuje številne analizne metode kot so MLR, PLS in FFBPNN; vsebuje pa tudi modul za računanje molekulskih deskriptorjev) (16). Programi nam torej omogočajo razvoj različnih algoritmov, vizualizacijo podatkov in struktur molekul, molekulsko modeliranje; validacijske metode; uporabljajo različne analizne metode, itd.. Zavedati se moramo, da posamezen program izvaja samo določene od zgoraj navedenih aktivnosti, ki so z opisom/definicijo računalniškega programa točno določene (16).

Eden izmed boljših internetnih virov omenjene programske opreme za razvoj QSAR modelov, je spletna stran The Cheminformatics and QSAR Society (<http://www.qsar.org>) (16).

V nadaljevanju je navedena programska oprema za razvoj QSAR modelov (16):

- C4.5,
- SVM^{light},
- ELECTRAS (Electronic Data Analysis Service),
- Weka,
- PHAKISO,
- NeuroSolutions,
- METLAB,
- Sybyl.

Slaba stran večine teh programov je, da nimajo zmožnosti izračunavanja molekulskih deskriptorjev, zato potrebujemo dodatne programe, kot so DRAGON, Molconn-Z, MODEL (Molecular Descriptor Lab) (16).

1.6.2 Programska oprema za napovedovanje specifičnih toksičnih lastnosti

Programska oprema je naslednja:

- DEREK (Deductive Estimation of Risk from Existing Knowledge),
- MCASE (Multiple Computer Automated Structure Evaluation),
- TOPKAT (Toxicity Prediction by Komputer Assisted Technology),
- ECOSAR (Ecological Structure Activity Relationship),
- HazardExpert,
- CSGenoTox,
- MetaDrug,
- Lazar,
- Toxtree (16).

1.6.3 Veljavna, prosto odstopna programska oprema v Evropski uniji

V zadnjem času so torej razvili številne računalniške programe, s katerimi je možno napovedovati toksične lastnosti spojin. Pri Skupnem raziskovalnem središču Evropske unije (JRC), ki spada pod okrilje Evropske komisije (EC), so odobrili naslednjo (uporabnikom prijazno in prosto dostopno) programsko opremo in podatkovne baze, ki jih lahko uporabljamo pri izvrševanju REACH-a (22):

- Toxtree,
- Toxmatch,
- DART (Decision Analysis by Ranking Techniques),
- CRAFT (Chemical Reactivity and Fate Tool),
- Danish (Q)SAR Database,
- JRC QSAR Model Database.

Omenjeno programsko opremo in podatkovne baze najdemo na spletni strani <http://ecb.jrc.ec.europa.eu/qsar/qsar-tools/>.

1.7 Toxtree

1.7.1 Opis in struktura programa / aplikacije

Toxtree je prilagodljiv in uporabnikom prijazen računalniški program / računalniška aplikacija, ki je ob enem tudi veljaven in odobren s strani Evropske komisije (EC). Razvilo ga je podjetje Ideacon Ltd iz Sofije v Bolgariji v skladu s pogodbo s Skupnim raziskovalnim središčem (JRC). Namenjen je strokovnjakom in vsakomur, ki ga zanima uporaba računalniških ocenjevalnih metod pri vrednotenju kemijske toksičnosti (23). Na spletu je prosto dostopen in ga najdemo na internetni strani Evropskega urada za kemikalije (Ex-ECB) (<http://ecb.jrc.ec.europa.eu/>). Julija 2009 je začela veljati nova verzija tega programa, in sicer Toxtree (Version 1.60), ki je zamenjala starejšo različico programa Toxtree (Version 1.5).

Program vsebuje naslednjih 9 ločenih modulov:

- Cramer-jeva pravila,
- Verhaar-jeva shema,
- napovedovanje dražilnega učinka na koži,
- napovedovanje dražilnega učinka na očeh,
- Benigni / Bossa-ejeva pravila za mutagenost in kancerogenost,
- strukturni elementi biološke razgradljivosti,
- strukturni elementi toksičnosti pri testu na podganjih mikronukleusih,
- razširjena Cramer-jeva pravila,
- strukturni elementi toksičnosti za identifikacijo Michael-ovih akceptorjev (24).

1.7.2 Analizna metoda za napovedovanje toksičnih lastnosti

Toxtree je preko posameznih modulov (le-te so razvile različne organizacije v različnih krajih po svetu) in z uporabo strukturnih pravil, sposoben razvrščanja spojin v skupine ter napovedovanja različnih toksičnih lastnosti spojin. Pri tem program uporablja metodo odločitvenega drevesa (DT) (22,23). Ta pristop spada med analizne metode za reševanje klasifikacijskih problemov (16).

Metoda odločitvenega drevesa ima drevesu podobno strukturo. Veja v odločitvenem drevesu predstavlja skupino določenih razredov podatkov, list pa predstavlja specifičen podatkovni razred. Posamezen vozle odločanja določa test, ki ga je potrebno izvesti pri posamezni vrednosti deskriptorja. Razvrščanje posameznih spojin z vektorjem x (komponente tega vektorja so molekularni deskriptorji), se s to metodo prične pri drevesnih

koreninah in se nato postopoma pomika navzgor po drevesu, dokler ne doseže lista. Na vsakem vozlu, ki ne vodi do lista, program izvede test in proces razvrščanja se nadaljuje na veji, ki jo test izbere. Ko je list dosežen, program napove podatkovni razred dane spojine, ki ga predstavlja omenjeni list (16).

1.7.3 Mutagenost in kancerogenost spojin

Mutagenost in kancerogenost sta eni izmed toksičnih lastnosti spojin, ki za zdravje človeka predstavljata veliko skrb in sta zato tudi predmet validiranih testnih metod petega aneksa (Annex V) Direktive 67/548/EEC (25). Kot kancerogeno lahko označimo tisto snov, ki povzroča maligno transformacijo celice in s tem rakavo obolenje. Mutagena snov pa je tista snov, ki povzroči nastanek genskih sprememb (največkrat mutacij). Ob tem se moramo zavedati, da mutagena snov ni nujno tudi kancerogena. Mutagena snov je kancerogena samo v primeru mutacije protoonkogenov in tumor zavirajočih genov (26).

Medtem, ko lahko mutagenost spojin napovemo z relativno enostavnimi metodami (najpogosteje uporabljena *in vitro* testa sta Ames-ov test, ki poteka na bakterijskem sevu *Salmonella typhimurium* in test na bakteriji *Escherichia coli*), so testi za napovedovanje kancerogenosti mnogo bolj zahtevni (potekajo na živalih) (3,27); zanje namreč potrebujemo daljše časovno obdobje (ponavadi 2 leti za testiranje na podganah in 18 mesecev za testiranje na miših) in veliko število testnih živali (miši in podgane obeh spolov, tri skupine različnih odmerkov in kontrolna skupina ter približno 50 živali na spol na skupino) (3). Vse to skupaj zahteva velike vsote denarja.

Kot je bilo že omenjeno, se eden od modulov računalniškega programa Toxtree nanaša tudi na napovedovanje kancerogenosti in mutagenosti spojin. Ta kodira Benigni / Bossa-ejeva pravila za kancerogenost in mutagenost, ki sta jih razvila Romualdo Benigni in Cecilia Bossa (Istituto Superiore di Sanita') v Rimu, v Italiji (27). Modul omogoča uporabnikom enostaven način napovedovanja kancerogenosti in mutagenosti spojin na osnovi poznavanja njihove strukture.

QSAR modeli za napovedovanje mutagenosti in kancerogenosti vključujejo tako kvalitativne (SAR, strukturni elementi) kot tudi kvantitativne (QSAR) odnose med strukturo in delovanjem (3). Oba pristopa sta v programu Toxtree tudi uporabljena.

1.7.4 Osnova za delovanje modula za napovedovanje mutagenosti in kancerogenosti

Glavno orodje modula za napovedovanje mutagenosti in kancerogenosti spojin je spisek **strukturnih elementov** (SA) za kancerogenost (27). Strukturni elementi so funkcionalne skupine molekul ali podstrukture, ki so povezane s kancerogenostjo spojin. Vključujejo jih številni tako komercialni (DEREK) kot tudi nekomercialni (Oncologic) računalniški programi. Niz spojin, ki jih označuje isti strukturni element, določa družino oziroma razred spojin, ki izkazujejo enak mehanizem delovanja. Toda ker je aktivnost posameznega strukturnega elementa lahko spremenjena ob prisotnosti neke druge podstrukture v bližini le-tega na molekuli (npr. prisotnost velike skupine na orto položaju ob aromatski amino skupini), potrebujemo še dodatno orodje, QSAR model, ki nam pomaga rešiti ta problem (27). Na aktivnost strukturnih elementov (aktivnost lahko zmanjšajo ali pa ovirajo njihovo delovanje) lahko vplivajo tudi naslednji faktorji:

- velika molekulska masa spojine (posledica le-te je manjša absorpcija),
- fizikalno stanje spojine (vpliva na sposobnost spojine pri doseganju ključnih mest delovanja),
- topnost spojine (velika hidrofilnost – majhna absorpcija, hitra eliminacija),
- kemijska reaktivnost spojine (spojine, ki so zelo reaktivne, ne morejo biti kancerogene, ker spontano hidrolizirajo in polimerizirajo ter zato, ker še pred dosegom mesta delovanja v celici reagirajo z drugimi nekritičnimi celičnimi komponentami),
- geometrija kemijske spojine (veliko število znanih kancerogenov in mutagenov ima planarno strukturo, elektrofilno funkcionalno skupino in ustrezno velikost za pravilno umeščanje/interkeliranje v molekulo DNA) (27).

V literaturi lahko najdemo veliko število različnih seznamov s strukturnimi elementi. Kot osnovo za izgradnjo računalniškega programa Toxtree pa so uporabili naslednje sezname:

- John Ashby-jev seznam (vsebuje 19 strukturnih elementov),
- Bailey-ev seznam (vsebuje 33 strukturnih elementov),
- Kazius-ov seznam (vsebuje 29 strukturnih elementov),
- novejši Kazius-ov seznam (vsebuje 6 kompleksnih strukturnih elementov).

Iz omenjenih štirih seznamov so za izgradnjo programa izbrali samo tiste strukturne elemente, ki se nahajajo v velikih ter zanesljivih podatkovnih bazah, in sicer ISSCAN, CPDB in Toxnet-Kazius. Pri tem je bila velika pozornost namenjena trem parametrom za

vrednotenje ustreznosti strukturnih elementov, in sicer občutljivosti, specifičnosti in točnosti.

Pri izgradnji programa je kot vir informacij o strukturnih elementih služil tudi ekspertni sistem Oncologic (27).

V kolikor program prepozna enega ali več strukturnih elementov, sporoči potencialno kancerogenost spojine. Spisek strukturnih elementov danes v glavnem vključuje takšne, ki jih glede na mehanizem delovanja označimo kot genotoksične kancerogene (neposredno povzročijo poškodbo na DNA; so elektrofilni ali pa se metabolično aktivirajo v elektrofilne aktivne intermediate), vanj pa so vključeni tudi tisti strukturni elementi, ki predstavljajo negenotoksične ali epigenetske kancerogene (z DNA ne interagirajo direktno in zato ne povzročijo direktne poškodbe DNA; pri standardnih testih mutagenosti dajejo negativen rezultat) (27). Vseh skupaj je v programu Toxtree identificiranih 33 strukturnih elementov; od teh jih pet pripada skupini negenotoksičnih kancerogenov.

Kot dodatek k strukturnim elementom program vključuje tudi naslednje **QSAR modele**:

- QSAR model za mutagenost aromatskih aminov na *Salmonella typhimurium* sevu TA100 (Ames test),
- QSAR model za kancerogenost aromatskih aminov na podganah
- QSAR model za mutagenost α,β -nenasičenih aldehydov na *Salmonella typhimurium* sevu TA100 (Ames test) (27).

Ti QSAR modeli so bili izdelani s kanonično diskriminantno analizo za številne individualne kemijske razrede mutagenov in kancerogenov, ki vključujejo aromatske amine, nitroarene, kinoline, triazene, policiklične aromatske ogljikovodike, laktone in aldehide. Nekateri QSAR modeli opišejo stopnjo jakosti aktivne spojine, medtem ko imajo drugi cilj, opredeliti spojino kot aktivno ali neaktivno. Z ozirom na raziskavo, ki jo je opravil Evropski urad za kemikalije v sodelovanju z Inštitutom Superiore di Sanita, imajo QSAR modeli za razlikovanje med aktivnimi in neaktivnimi spojinami večjo napovedno vrednost kot tisti za opis stopnje jakosti delovanja. QSAR analiza zagotavlja bolj izpopolnjeno ocenjevanje kot je tisto, ki uporablja samo strukturne elemente, zato naj bi imela tudi bolj pomembno vlogo (27).

Ko želimo napovedati biološko aktivnost nove spojine, se moramo zavedati, da mora le ta vsebovati karakteristike učnega niza, oziroma, da mora biti znotraj uporabne domene posameznega modela.

1.7.5 Izidi modula za napovedovanje mutagenosti in kancerogenosti

Končni izidi modula za napovedovanje mutagenosti in kancerogenosti so lahko naslednji:

- ni zaznanega strukturnega elementa toksičnosti za kancerogenost,
- zaznan eden ali več strukturnih elementov za genotoksično ali negenotoksično kancerogenost,
- zaznani strukturni elementi, ki so sorodni aromatskim aminom ali α,β -nenasičenim aldehidom; v tem primeru je spojina podvržena še QSAR analizi, ki pa ima lahko za rezultat pozitiven ali negativen izid (binaren odgovor) (27).

Tako program po opravljenem testiranju spojin, končne rezultate poda z eno (lahko tudi več) od naslednjih opredelitev:

- 1 - strukturni element za genotoksično kancerogenost,
- 2 - strukturni element za negenotoksično kancerogenost,
- 3 - ni strukturnih elementov za kancerogeno aktivnost,
- 4 - spojina je potencialni *S. typhimurium* TA100 mutagen, ki bazira na QSAR,
- 5 - majhna verjetnost, da je spojina *S. typhimurium* TA100 mutagen, ki bazira na QSAR,
- 6 - spojina je potencialni kancerogen, ki bazira na QSAR,
- 7 - majhna verjetnost, da je spojina kancerogen, ki bazira na QSAR,
- 8 - za boljšo oceno bi morali uporabiti QSAR izračune,
- 9 - napaka pri uporabi metode odločitvenega drevesa (DT).

2 NAMEN DELA

Na področju ocenjevanja varnosti kemikalij so se z ustanovitvijo in s pričetkom veljave novega sistema REACH, v vse večji meri začele razvijati in uporabljati alternativne metode. Te vključujejo med drugim tudi številne SAR in QSAR modele ter ekspertne sisteme. Tako lahko na enostaven način, z uporabo določenega (izbranega) računalniškega programa, napovemo lastnosti (tako fizikalno-kemijske kot tudi biološke lastnosti) preiskovanih spojin.

Na spletu je danes prosto dostopnih kar nekaj računalniških programov oziroma ekspertnih sistemov za napovedovanje toksičnih lastnosti spojin (Toxtree, ECOSAR, Toxmatch, Lazar, itd.) (3,22,28). Med seboj se razlikujejo v tem, katere toksične lastnosti lahko s posameznim ekspertnim sistemom določamo, ločijo pa se tudi po načinu določanja lastnosti (v našem primeru toksičnih lastnosti) preiskovanim spojinam, kar posledično pripelje do različnih stopenj napovednosti posameznega programa oziroma posameznega programskega modula.

Glavni cilj diplomske naloge bo ovrednotiti napovednost prosto dostopnega spletnega računalniškega programa Toxtree; natančneje njegovega modula, ki se nanaša na napovedovanje kancerogenosti in mutagenosti spojin (Benigni / Bossa-ejeva pravila za mutagenost in kancerogenost). Napovednost bomo vrednotili iz vidika občutljivosti, specifičnosti in točnosti omenjenega programskega modula.

Opravili bomo testiranje skupno 68-ih naključno izbranih spojin. Najprej bomo na skupini 34-ih spojin (prvi niz spojin, naključno izbranih iz različnih literaturnih virov in podatkovnih baz, ki so navedene v poglavju »Materiali in metode«, na strani 28), ki so v literaturi opisane kot dejansko/dokazano mutagene in kancerogene, želeli ovrednotiti **občutljivost** za napovednost modula (odstotek biološko aktivnih-pozitivnih spojin, ki jih programski modul pravilno opredeli kot mutagene in kancerogene). Nato pa bomo na skupini 34-ih spojin (drugi niz naključno izbranih spojin), za katere menimo, da niso kancerogene in mutagene oziroma v literaturi nismo zasledili, da bi bile kancerogene in mutagene ter so dnevno v uporabi (prehranski aditivi, kemikalije, zdravilne učinkovine), želeli ugotoviti **specifičnost** za napovednost modula (odstotek biološko neaktivnih-negativnih spojin, ki jih programski modul pravilno opredeli kot nekancerogene in

nemutagene). Zanimala nas bo tudi **točnost** za napovednost modula (odstotek vseh spojin, ki jih modul pravilno opredeli).

Med drugim bomo preverili ali programski modul ustrezno ovrednoti spojine, ki se morajo predhodno metabolično aktivirati, da postanejo mutagene in kancerogene. Na ta način bomo ocenili, kako ustrezno oziroma, če sploh je, v omenjeni modul vgrajena komponenta metabolične aktivacije in izračunali odstotek pravilno ovrednotenih spojin tega tipa.

3 MATERIALI IN METODE

3.1 Izbor spojin za testiranje

V skladu z namenom diplomske naloge smo se odločili sestaviti dva niza naključno izbranih spojin. Prvi niz sestavlja 34 spojin, ki so v literaturi opredeljene kot mutagene in kancerogene (Priloga I). Kot vir za prvi niz spojin smo uporabili naslednjo literaturo:

- Auterhoff H, Knabe J, Höltje HD. Lehrbuch der Pharmazeutischen Chemie. 14. Auflage, Wissenschaftliche Verlagsgesellschaft mbH Stuttgart, Stuttgart, 1999.
- Mulder GJ, Dencker L. Pharmaceutical Toxicology (Safety sciences of drugs). Pharmaceutical Press, London, 2006.
- Timbrell J. Introduction to Toxicology. 3rd Edition, Informa Healthcare USA, Inc., 2009.
- Klaasen CD. Casarett & Doull's Toxicology, the basic science of poisons, 6th Edition, McGraw-Hill Companies, Inc., USA, 2001.

Uporabili smo tudi podatkovno bazo IARC (International Agency for Research on Cancer) Monographs on the Evaluation of Carcinogenic Risks to Humans; <http://monographs.iarc.fr/ENG/Classification/crthgr01.php>.

Vse spojine iz prvega niza, razen klornafazina, smo našli v podatkovni bazi CPDB (The Carcinogenic Potency Database; Lois Swirsky Gold); <http://potency.berkeley.edu/>.

Drugi niz sestavlja 34 spojin, ki jih dnevno uporablja večja populacija ljudi (prehranski aditivi, kemikalije in zdravilne učinkovine) (Priloga II). Zanje v literaturi nismo zasledili, da bile mutagene in kancerogene. Aditive smo izbrali iz E-number indeksa, ki se nahaja na spletni strani <http://www.ukfoodguide.net/enumeric.htm>. Na tej internetni strani lahko za določene aditive izvemo, v katerih izdelkih se nahajajo, kakšni so njihovi neželeni učinki in morebitno prepoved uporabe le teh v določenih državah po svetu. Vse izbrane spojine drugega niza, razen etil in metilparabena smo našli tudi v CPDB.

Na spletu torej lahko najdemo kar nekaj prosto dostopnih podatkovnih baz, ki nam omogočajo evalvacijo rezultatov testiranja želenega ekspertnega sistema. Te so naslednje:

- CPDB (The Carcinogenic Potency Database; <http://potency.berkeley.edu/>),
- EPA-HEN (US EPA, Air Risk Information Support Center, Health Effects Notebook for Hazardous Air Pollutants; <http://www.epa.gov/ttnatw01/hlthef/hapindex.html>),

- HAZMAP (A Relational Database of Hazardous Chemicals and Occupational Diseases; <http://hazmap.nlm.nih.gov/hazmapadv.html>),
- IARC (International Agency for Research on Cancer; <http://monographs.iarc.fr/ENG/Classification/index.php>),
- IRIS (Integrated Risk Information System; <http://www.epa.gov/NCEA/iris/>),
- OEHHA (Office of Environmental Health Hazard Assessment, Proposition 65; http://www.oehha.ca.gov/prop65/prop65_list/files/P65single020510.pdf),
- TOXNET (Toxicology Data Network; <http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?CCRIS>).

Pri našem praktičnem delu smo za evalvacijo rezultatov testiranja uporabili podatkovni bazi TOXNET in CPDB.

V obeh nizih smo spojine razvrstili v skupine glede na strukturne elemente, ki jih zasledimo v Benigni / Bossa zbirki pravil (ta vsebuje 33 strukturnih elementov, od katerih jih pet spada med negenotoksične kancerogene) (27). Toxtree-jev programski modul za napovedovanje mutagenosti in kancerogenosti v glavnem deluje na prepoznavi enega ali več strukturnih elementov v strukturi spojine, poleg tega pa vključuje tudi tri QSAR modele (27).

3.2 Testiranje spojin

Najprej smo v podatkovni bazi CPDB spojine obeh nizov preverili glede mutagenosti in kancerogenosti (spojine, ki jih v tej bazi nismo našli, smo poiskali v podatkovni bazi TOXNET). Podatkovna baza CPDB vsebuje rezultate številnih testiranj kancerogenosti na živalih, rezultate testiranj mutagenosti na bakteriji *Salmonella typhimurium* TA100 in ob enem tudi SMILES-e, ki smo jih potrebovali za testiranje spojin z računalniškim programom Toxtree (namesto strukturne formule spojin smo v program vnašali njihove SMILES kode).

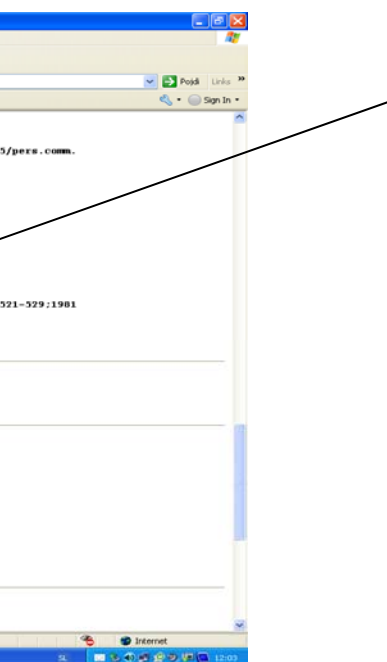
SMILES (Simplified Molecular Input Line Entry Specification) je koda, ki podrobno in nedvoumno opisuje kemijske molekule in reakcije ob uporabi kratkih ASCII (American Standard Code for Information Interchange) nizov/serij (te nize sestavljajo simboli, ki predstavljajo atome in vezi v molekuli). SMILES-e sta v poznih 80-ih letih razvila Arthur Weininger in David Weininger. Od takrat dalje so kode SMILES modificirali in razširili tudi drugi strokovnjaki; najbolj pomembno pa je na razvoj vplivala družba Daylight Chemical Information Systems Inc. (29,30). SMILES kodo lahko prenesemo v vsak urejevalec molekul, ki nam kodo pretvori v dvodimenzionalen ali tridimenzionalen model molekule, na katero se nanaša omenjena koda. SMILES-e lahko uporabljamo v naslednjih primerih:

- služijo nam lahko kot ključ za vstop v podatkovne baze,
- za izmenjavo kemijskih informacij med raziskovalci,
- kot vstopni sistem za kemijski podatek,
- so eden od jezikov ekspertnih sistemov v kemiji (31).

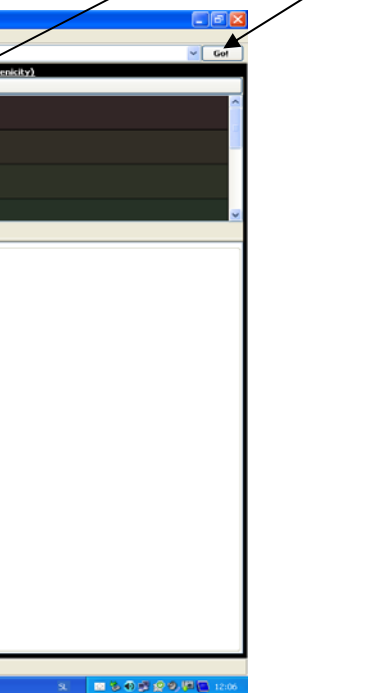
Po preverjanju podatkov o mutagenosti in kancerogenosti spojin v CPDB, smo iz spletne strani <http://ecb.jrc.ec.europa.eu/qsar/qsar-tools/index.php?c=TOXTREE>, na osebni računalnik naložili program Toxtree (Version 1.6). Po zagonu omenjenega programa smo izbrali polje »Method«, nato pa smo med vsemi možnostmi, ki jih ta ponuja, izbrali »Select a decision tree«. Tako nam je program ponudil vse svoje programske module. Med njimi smo izbrali modul »Benigni / Bossa rulebase for mutagenicity and carcinogenicity«, ki nam je tekom diplomske naloge služil kot glavno orodje za testiranje spojin.

ih:

- wni bazi CPDB.



- ramu Toxtree; nato izberemo
e strukturna formula spojine



- Izberemo polje »Estimate« in program nam ovrednoti spojino z enim od devetih možnih izidov.

The screenshot shows the ToxTree v1.60 interface. The main window displays a decision tree for estimating toxic hazard. The left sidebar lists available structure attributes like SMILES, CAS, and IUPAC. The main area displays a decision tree with nodes for structural alerts and QSAR models. A chemical structure of a phenol derivative is shown on the left. The bottom status bar indicates 'Completed'.

- Na koncu sledi še zbiranje rezultatov testiranja spojin in vpisovanje le teh v dve ločeni tabeli (ena za prvi niz spojin in ena za drugi niz spojin).

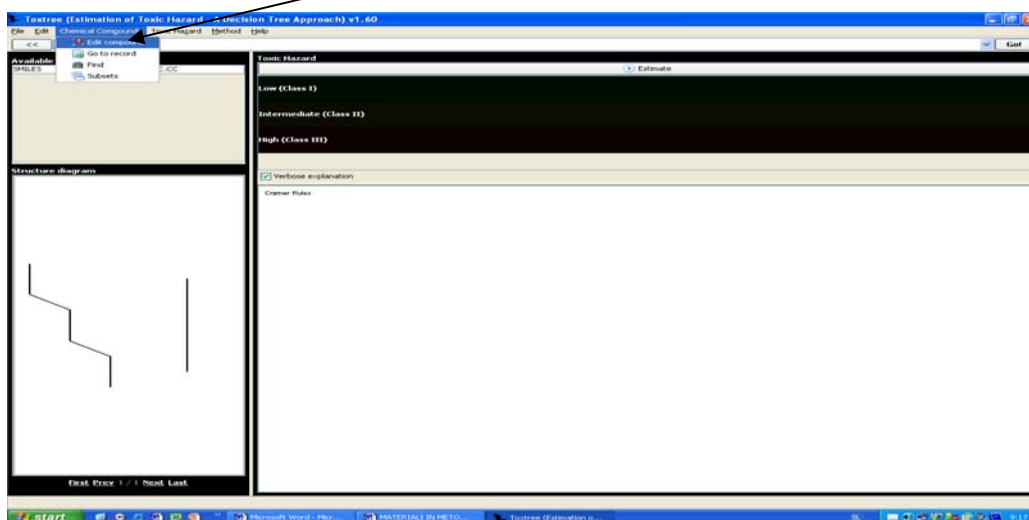
V primeru, da spojine, ki jo želimo testirati, v CPDB ne najdemo, le tej lahko SMILES kodo določimo tako, da spojino narišemo z enim od številnih računalniških programov za predstavitev ali editiranje kemijskih molekul (npr. Molinspiration WebME Editor 2.6).

The screenshot shows the Molinspiration WebME Editor 2.6 web application. The interface includes a toolbar with various editing tools and a central canvas displaying a chemical structure. The bottom status bar indicates 'Completed'.

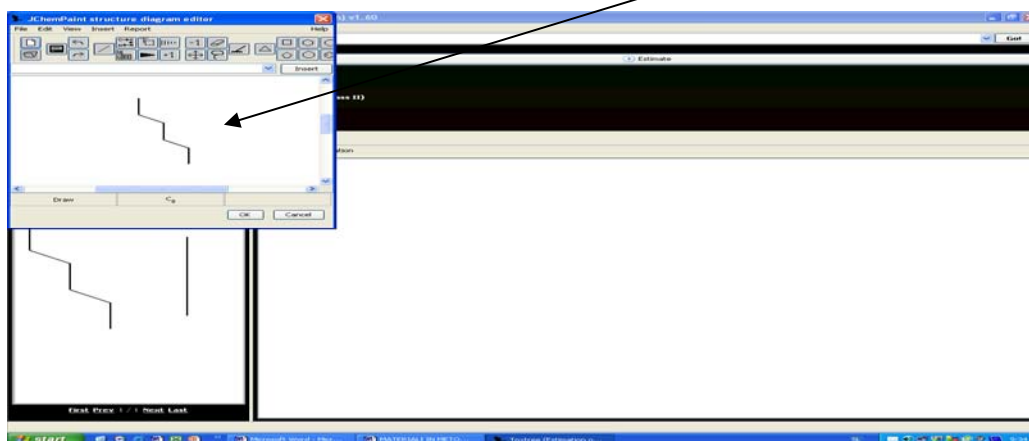
Na podlagi strukturne formule spojine nam program po izboru ikone »SMILES«, poda SMILES kodo, ki ustreza narisani spojini.

Kode nekaterih spojin lahko najdemo tudi na spletni strani Wikipedia-e (http://en.wikipedia.org/wiki/Main_Page), pri čemer v iskalnik omenjene spletne strani odtipkamo ime želene spojine. Tako se nam na ekranu izpišejo številne karakteristike spojine, med katerimi je pri določenih spojinah zapisana tudi SMILES koda.

Na tem mestu moramo omeniti, da spojino, ki jo želimo testirati, lahko narišemo tudi v programu Toxtree in sicer s pomočjo programskega orodja »Jchem Paint structure diagram editor«. To storimo tako, da izberemo polje »Chemical Compounds« in nato »Edit Compound«.



Na levi strani zaslona se pojavi okno, v katerem lahko narišemo željeno spojino.



3.3 Analiza rezultatov testiranja

V skladu z namenom diplomske naloge smo na podlagi pridobljenih rezultatov testiranja izbranih spojin, želeli ovrednotiti napovednost Toxtree-jevega programskega modula za napovedovanje mutagenosti in kancerogenosti spojin. Čeprav smo v literaturi zasledili, da lahko to lastnost vrednotimo tudi s pomočjo validacije, tako interne/prečne/navskrižne validacije (velja za najmanj zanesljivo, a največkrat uporabljeno metodo za vrednotenje napovednosti), kot tudi prospektivne eksterne validacije (najbolj eksaktna znanstvena metoda za vrednotenje napovednosti) in retrospektivne eksterne validacije (28,32,33), smo se raje odločili za vrednotenje napovednosti z vidika občutljivosti, specifičnosti in točnosti (1,20,21). Slednja metoda se nam je zdela bolj primerna, predvsem pa veliko bolj enostavna za uporabo.

Toxtree-jev programski modul za napovedovanje mutagenosti in kancerogenosti ovrednoti testirane spojine z enim od naslednjih devetih možnih izidov:

- 1 - strukturni element za genotoksično kancerogenost,
- 2 - strukturni element za negenotoksično kancerogenost,
- 3 - ni strukturnih elementov za kancerogeno aktivnost,
- 4 - spojina je potencialni *S. typhimurium* TA100 mutagen, ki bazira na QSAR,
- 5 - majhna verjetnost, da je spojina *S. typhimurium* TA100 mutagen, ki bazira na QSAR,
- 6 - spojina je potencialni kancerogen, ki bazira na QSAR,
- 7 - majhna verjetnost, da je spojina kancerogen, ki bazira na QSAR,
- 8 - za boljšo oceno bi morali uporabiti QSAR izračune,
- 9 - napaka pri uporabi metode odločitvenega drevesa (DT).

Po vnosu rezultatov testiranja spojin v dve ločeni tabeli v program Microsoft Excel (ločeni tabeli za prvi in drugi niz spojin), smo napovednost Toxtree-jevega programskega modula za napovedovanje mutagenosti in kancerogenosti (Benigni / Bossa rulebase for mutagenicity and carcinogenicity), kot je bilo že prej omenjeno, ovrednotili z vidika občutljivosti, specifičnosti in točnosti. Pri tem smo za izračun omenjenih treh parametrov, uporabili naslednje formule:

$$\text{Točnost} = ((N_{M+K+} + N_{M-K-}) / N_{\text{eval}}) \times 100$$

$$\text{Občutljivost} = (N_{M+K+} / N_{LM+LK+}) \times 100$$

$$\text{Specifičnost} = (N_{M-K-} / N_{LM-LK-}) \times 100$$

Oznake v navedenih formulah predstavljajo:

$N_{LM+LK+} \Rightarrow$ število spojin, ki so v literaturi opredeljene kot mutagene in kancerogene,

$N_{LM-LK-} \Rightarrow$ število spojin, za katere v literaturi nismo zasledili, da bi bile mutagene in kancerogene,

$N_{M+K+} \Rightarrow$ število spojin, ki so v literaturi opredeljene kot mutagene in kancerogene in so tudi pri testiranju z ekspertnim sistemom dale enak (pozitiven) izid,

$N_{M-K-} \Rightarrow$ število spojin, ki v literaturi niso opredeljene kot mutagene in kancerogene in so tudi pri testiranju z ekspertnim sistemom dale enak (negativen) izid,

$N_{eval} \Rightarrow$ število vseh testiranih spojin z ekspertnim sistemom.

Če je program ovrednotil spojino z izidoma 1 ali 2, smo spojino šteli kot mutageno in kancerogeno, če jo je označil z izidom 3, pa smo jo šteli kot nemutageno in nekancerogeno. V nekaj primerih smo dobili tudi izide 4, 5, 6 in 7, ki pa nikoli niso bili izraženi sami zase, ampak so v vseh primerih stali poleg izida 1; tudi vse take spojine smo šteli kot mutagene in kancerogene.

Pri izračunu točnosti, specifičnosti in občutljivosti, smo upoštevali literaturne podatke, o mutagenosti in kancerogenosti spojin iz CPDB. V primeru, da spojine v tej bazi nismo našli, pa smo upoštevali podatke iz literaturnih virov izbora spojin, navedenih na začetku tega poglavja (na strani 28).

Kot mutageno in kancerogeno smo smatrali tisto spojino prvega niza, ki je bila v CPDB za obe lastnosti opredeljena kot pozitivna, kot nemutageno in nekancerogeno pa smo smatrali tisto spojino drugega niza, ki je bila v CPDB za obe lastnosti opredeljena kot negativna ali pa za ti dve lastnosti v tej bazi nismo našli podatka.

Na podlagi literaturnih podatkov o metabolični aktivaciji smo naredili dodatno tabelo, v katero smo vstavili rezultate testiranj vseh spojin iz prvega niza, ki se morajo predhodno metabolično aktivirati, da postanejo mutagene in kancerogene. Tako smo preverili, kako ustrezno oziroma, če sploh je, v omenjeni programski modul za napovedovanje mutagenosti in kancerogenosti vgrajena komponenta metabolične aktivacije. S pomočjo teh rezultatov smo izračunali, kolikšen je odstotek pravilno ovrednotenih spojin z omenjeno karakteristiko.

4 REZULTATI IN RAZPRAVA

Po ustanovitvi novega sistema REACH (Registration, Evaluation, Authorisation and Restriction of Chemicals), ki se je v praksi začel v celoti izvajati 1. junija 2008, so se za ocenjevanje oziroma napovedovanje toksičnih lastnosti spojin v vse večji meri začele uporabljati in razvijati alternativne metode. Te vključujejo tudi različne SAR in QSAR modele ter računalniško osnovane ekspertne sisteme, s katerimi lahko na osnovi povezave med strukturo in delovanjem neke skupine spojin, napovemo fizikalno-kemijske in biološke lastnosti njim podobnih spojin. Rezultate omenjenih metod lahko uporabljamo v regulatorne namene samo v primeru, če je model ustrezno validiran in primerno dokumentiran. Le na ta način je možno izvesti zanesljivo napoved in pridobiti ustrezne rezultate (34). Kritično ocenjevanje napovednosti QSAR modelov je torej prvi korak na poti k verodostojni in zanesljivi uporabi takšnih računskih tehnik. Vrednotimo jo lahko s pomočjo interne validacije, eksterne validacije (28,32,33) ali pa s stališča občutljivosti, specifičnosti in točnosti (1,20,21).

S pomočjo QSAR modelov in ekspertnih sistemov lahko danes napovemo celo kopico toksičnih lastnosti, med katerimi po pomembnosti močno izstopata mutagenost in kancerogenost spojin. V začetku leta 2008 so tako poročali že o približno 78 nekomercialnih QSAR modelih za napovedovanje mutagenosti in kancerogenosti (pri tem se moramo zavedati, da obstaja tudi veliko število komercialnih QSAR modelov) (28). Mi smo se v skladu z namenom diplomske naloge odločili za vrednotenje napovednosti nekomercialnega (prosto dostopnega) računalniškega programa Toxtree in sicer njegovega programskega modula, ki omogoča napovedovanje mutagenosti in kancerogenosti spojin (Benigni / Bossa rulebase for mutagenicity and carcinogenicity).

Toxtree je računalniška aplikacija, s katero lahko napovedujemo različne tipe toksičnih lastnosti spojin, z uporabo pristopa odločitvenega drevesa (DT). Program sestavljajo različni moduli, od katerih se eden nanaša tudi na napovedovanje mutagenosti in kancerogenosti spojin. Glavno orodje tega modula je seznam 33-ih strukturnih elementov, katerim pa so pridruženi še trije QSAR modeli (27). V kolikor program v strukturi spojine, ki jo testiramo, prepozna enega ali več strukturnih elementov iz prej omenjenega seznama, takoj sporoči potencialno mutagenost oziroma kancerogenost spojine.

Napovednost programskega modula smo se odločili vrednotiti iz vidika občutljivosti, specifičnosti in točnosti, saj se nam je ta način s stališča števila predvidenih spojin za

testiranje (68 naključno izbranih spojin razdeljenih v dva testna niza), zahtevnosti izvedbe in interpretacije rezultatov, zdel najprimernejši.

4.1 Občutljivost za napovednost

Rezultati testiranja prvega niza, ki zajema 34 spojin, so podani v preglednici na naslednji strani (Preglednica 3). V njej so zbrana vsa imena in CAS številke posameznih spojin, morebitna potreba po predhodni metabolični aktivaciji spojin za njihovo mutageno in kancerogeno delovanje (to smo črpali iz knjige Casarett & Doull's Toxicology, the basic science of poisons, avtorja Klaasen-a (35) in iz baze člankov, ki jih nudi ScienceDirect (<http://www.sciencedirect.com/>)), v preglednici so navedeni tudi podatki o mutagenosti in kancerogenosti spojin, pridobljeni iz CPDB, in nenazadnje še rezultati testiranja spojin s Toxtree-jevim programskim modulom za napovedovanje mutagenosti in kancerogenosti.

Na enem mestu smo zbrali tudi vse strukturne formule spojin, SMILES kode, ki smo jih uporabili pri njihovem testiranju in namene uporabe posameznih spojin (Priloga I).

Pred obdelavo rezultatov, smo spojine prvega niza, po strukturnih elementih, razvrstili v skupine (Preglednica 4). V prvem stolpcu preglednice 4 so po vrsti zbrani Benigni / Bossa-jevi strukturni elementi; med njimi so z modro barvo označeni strukturni elementi za negenotoksično kancerogenost, vsi ostali pa predstavljajo strukturne elemente za genotoksično kancerogenost. V tretjem stolpcu so zajete tiste testirane spojine, ki so na Benigni / Bossa seznamu prisotne kot predstavniki posameznega strukturnega elementa, v zadnjem stolpcu pa so zajete še vse ostale spojine, ki so bile predmet testiranja. Na tem mestu je potrebno poudariti, da nam štirih spojin, in sicer akrilonitrila, benzena, 1,3-butadiena in formaldehida, ni uspelo razvrstiti v nobeno od skupin, vendar smo jih zaradi izrazito mutagenih in kancerogenih lastnosti, ki smo jih zasledili v literaturi, vseeno testirali.

Preglednica 3: Rezultati testiranja spojin prvega niza.

spojina	CAS	metabolična aktivacija ^(a)		CPDB ^(b)		Toxtree ^(c)								
		Douls	ScienceDirect	kancer.	mutag.	1	2	3	4	5	6	7	8	9
benzidin	92-87-5	+	+	+	+	+			+			+		
p-aminodifenil	92-67-1	+		+	+	+			+		+			
2-naftilamin	91-59-8	+		+	+	+			+		+			
stiren oksid	96-09-3			+	+	+								
N,N-dietilnitrozamin	55-18-5		+	+	+	+								
dibenz(a,h)antracen	53-70-3			+	+	+								
DAB	60-11-7	+	+	+	+	+			+		+			
furfural	98-01-1			+	+	+								
melfalan	148-82-3			+	+	+								
ciklofosfamid	50-18-0		+	+	+	+								
1,2-propilen oksid	75-56-9			+	+	+								
bromodiklorometan	75-27-4		+	+	+	+								
azatioprin	446-86-6			+	+	+								
8-metoksi psoralen	298-81-7			+	+	+								
mitomicin C	50-07-7			+	+	+								
NNK (nitrozamin)	64091-91-4		+	+	/	+								
4,4-metilen-bis(2-kloroanilin)	101-14-4			+	+	+			+		+			
klornafazin	494-03-1			/	/	+								
vinil klorid	75-01-4		+	+	+	+								
aflatoksin B1	1162-65-8	+	+	+	+	+								
uretan	51-79-6		+	+	+	+								
benz[a]piren	50-32-8	+		+	+	+								
fenacetin	62-44-2			+	+	+				+	+			
2-acetilaminofluoren	53-96-3		+	+	+	+			+		+			
1,2-dibromoetan	106-93-4	+	+	+	+	+								
klorambucil	305-03-3			+	+	+								

spojina	CAS	metabolična aktivacija ^(a)		CPDB ^(b)		Toxtree ^(c)								
		Douls	ScienceDirect	kancer.	mutag.	1	2	3	4	5	6	7	8	9
tiotepa	52-24-4			+	+	+								
kumarin	91-64-5		+	+	+	+								
TCDD	1746-01-6		+	+	-		+							
mireks	2385-85-5			+	-		+							
akrilonitril	107-13-1		+	+	+			+						
benzen	71-43-2		+	+	-			+						
1,3-butadien	106-99-0		+	+	+			+						
formaldehid	50-00-0			+	+			+						

^(a)Podatki o metabolični aktivaciji spojin, ki smo jih črpali iz dveh virov; Casarett & Doull's Toxicology, the basic science of poisons, avtorja Klaasen-a (35) in iz baze člankov, ki jih nudi ScienceDirect (<http://www.sciencedirect.com/>). Z znakom + so označene spojine, ki se morajo predhodno metabolično aktivirati, da postanejo mutagene in kancerogene.

^(b)Podatki o mutagenosti in kancerogenosti spojin, ki smo jih črpali iz podatkovne baze CPDB (The Carcinogenic Potency Database). Z znakom + so označene mutagene oziroma kancerogene spojine, z – so označene nemutagene oziroma nekancerogene spojine, oznaka / pa pomeni, da podatka o mutagenosti oziroma kancerogenosti v CPDB nismo našli.

^(c)Rezultati testiranja prvega niza spojin s Toxtree-jevim programskim modulom za napovedovanje mutagenosti in kancerogenosti. Številke 1 – 9 označujejo devet možnih izidov modula, ki so opisani v poglavju »Materiali in metode«, natančneje v podpoglavju »Analiza rezultatov testiranja« na strani 34.

Preglednica 4: Razvrstitev prvega niza spojin po strukturnih elementih, ki jih prepozna Toxtree.

Skupine strukturnih elementov ^(a)	Št. strukt.elem. ^(b)	Sp. iz B/B zbirke strukt. elem. ^(c)	Ostale spojine ^(d)
acil halidi	SA 1		
alkilni ali benzilni estri sulfon. ali fosfon. kislin	SA 2		
N-metilol derivati	SA 3		
monohaloalkeni	SA 4	vinilklorid	
S ali N iperiti	SA 5	klorambucil	melfalan;ciklofosfamid; klornafazin
propiolaktoni ali propiosultoni	SA 6		
epoksidi in aziridini	SA 7	tiotepa	stiren oksid; 1,2-propilen oksid; mitomicinC
alifatski halogeni	SA 8	1,2-dibromoetan	bromodiklorometan
alkil nitriti	SA 9		
α,β -nenasičeni karbonili	SA 10		
enostavni aldehidi	SA 11		furfural
kinoni	SA 12		mitomicin C
hidrazini	SA 13		
alifatske azo in azoksi spojine	SA 14		
izocianati in izotiocianati	SA 15		
alkil karbamati in tiokarbamati	SA 16	uretan	mitomicin C
tiokarbonili	SA 17		
policiklični aromatski ogljikovodiki	SA 18	benz[a]piren	dibenz(a,h)antracen
heterociklični policiklični aromatski CH	SA 19		
halogenirani cikloalkani	SA 20	mireks	
alkil in aril nitrozo skupine	SA 21		N,N-dietilnitrozamin; 4-(metilnitrozamino)-1-(3-piridil)-1-butanon
azidi in triazeni	SA 22		
alifatska N-nitro skupina	SA 23		
α,β -nenasičena alifatske alkoksi skupina	SA 24		
aromska nitrozo skupina	SA 25		
aromatski obroč N-oksidi	SA 26		
nitro aromati	SA 27		azatioprin

Skupine strukturnih elementov ^(a)	Št. strukt.elem. ^(b)	Sp. iz B/B zbirke strukt. elem. ^(c)	Ostale spojine ^(d)
primarni arom. amini, hidroksil amini in derivati	SA 28		benzidin; p-aminodifenil; 2-naftilamin; 4,4-metilen-bis(2-kloroanilin)
aromatski mono in dialkil amini	SA 28bis		N,N-dimetil-4-aminoazobenzen
aromatski N-acilamini	SA 28ter	fenacetin; 2-acetilaminofluoren	
aromatske diazo spojine	SA 29		N,N-dimetil-4-aminoazobenzen
kumarini in furokumarini	SA 30	aflatoksin B1; kumarin	8-metoksipsoralen
halogenirani benzeni	SA 31a		
halogenirani PAH	SA 31b		
halogenirani dibenzoksini	SA 31c	2,3,7,8-tetraklorodibenzo-p-dioksin	

^(a) Skupine strukturnih elementov, ki so navedene na Benigni / Bossa-ejevem seznamu (27) in jih prepozna program Toxtree. Modro so označeni strukturni elementi za negenotoksično kancerogenost, vsi ostali pa so strukturni elementi za genotoksično kancerogenost.

^(b) Številka skupine posameznega strukturnega elementa iz Benigni / Bossa-ejevega seznama.

^(c) Testirane spojine, ki so na Benigni / Bossa-ejevem seznamu prisotne kot predstavniki posameznega strukturnega elementa.

^(d) Razvrstitev vseh preostalih testiranih spojin po strukturnih elementih.

Na podlagi rezultatov testiranja prvega niza spojin (Preglednica 3), smo s pomočjo spodnje formule izračunali občutljivost Toxtree-jevega programskega modula za napovedovanje mutagenosti in kancerogenosti spojin. Občutljivost, ki v našem primeru predstavlja odstotek biološko aktivnih-pozitivnih spojin, ki jih je programski modul pravilno opredelil kot mutagene in kancerogene, je znašala **90,3 %**.

$$\text{Občutljivost} = (N_{M+K+} / N_{LM+LK+}) \times 100$$

$$\text{Občutljivost} = (28 / 31) \times 100$$

$$\text{Občutljivost} = 90,3 \%$$

Pri izračunu smo upoštevali rezultate testiranja vseh spojin razen TCDD (2,3,7,8-tetraklorodibenzo-p-dioksin), mireksa in benzena. Omenjene tri spojine smo izločili, ker smo pri pregledu literaturnih podatkov o njihovi mutagenosti in kancerogenosti ugotovili, da spadajo v skupino nemutagenih kemijskih kancerogenov, torej nimajo obeh preiskovanih lastnosti pozitivno izraženih (to je bila naša zahteva za pozitivno opredelitev spojin glede mutagenosti in kancerogenosti). Ob tem moramo poudariti, da je programski modul spojin TCDD in mireks, označil z izidom 2, kar pomeni, da vsebujeta strukturne elemente za negenotoksično kancerogenost (pravilna opredelitev programskega modula). Benzen pa je programski modul označil z izidom 3, kar pomeni, da spojina ne vsebuje strukturnih elementov za kancerogeno aktivnost. Na primeru te spojine je programski modul storil napako. Res je, da benzen v svoji strukturi ne vsebuje nobenega strukturnega elementa iz Benigni / Bossa-ejevega seznama, je pa dobro znano, da se ta spojina metabolično aktivira, pri čemer nastaneta p-benzokinon in o-benzokinon, ki pa povzročata levkemijo (35,36).

V zgornjo formulo za izračun občutljivosti smo v imenovalnik vstavili število spojin iz preglednice 3, ki so v literaturi opredeljene kot mutagene in kancerogene (N_{LM+LK+}). Teh je bilo 31. V števec pa smo vstavili število spojin, ki so v literaturi opredeljene kot mutagene in kancerogene in so tudi pri testiranju z ekspertnim sistemom dale enak (pozitiven) izid (N_{M+K+}). Teh je bilo skupaj 28. Programski modul je torej za tri spojine (od 31-ih), in sicer za akrilonitril, 1,3-butadien in formaldehid, podal napačno napoved o njihovi mutagenosti

in kancerogenosti. Po preverjanju morebitne prisotnosti strukturnih elementov v omenjenih treh spojinah (Preglednica 4), smo ugotovili, da le-te v svoji strukturi ne vsebujejo nobenega od strukturnih elementov iz Benigni / Bossa-ejevega seznama. To je bil glavni razlog, da programski modul spojin ni mogel prepoznati kot mutagenih in kancerogenih. Modul namreč spojine najprej opredeli s stališča strukturnih elementov in šele potem, v primeru prisotnosti aromatskih aminov ali α,β -nenasičenih aldehydov v strukturi, spojino dodatno testira s tremi QSAR modeli, ki so mu na voljo. V teh treh primerih in tudi na primeru benzena (tudi ta ne vsebuje strukturnih elementov iz Benigni / Bossa-ejevega seznama), bi bilo veliko bolje, če bi modul uporabniku sporočil, da takšne spojine ne more testirati, ker je ni v njegovi uporabni domeni, kot pa, da zanjo poda napačno napoved. Po drugi strani je jasno, da bi v primeru prisotnosti na primer benzena ali formaldehida kot posameznega strukturnega elementa na Benigni / Bossa-ejevem seznamu, programski modul vse spojine, ki imajo v svoji strukturi fragment benzenovega obroča ali aldehydno skupino, opredelil kot mutagene in kancerogene. V tem primeru bi pri napovedovanju mutagenosti in kancerogenosti, programski modul naredili še večjo napako kot jo sicer, saj je spojin, ki v svoji strukturi vsebujejo enega od teh dveh fragmentov, zelo veliko. Zavedati se moramo tudi, da so spojine takšnega tipa že tako »enostavne«, da jih bolj ne moremo poenostaviti (če seveda hočemo obdržati enak tip spojine), ampak jim nasprotno, posamezne fragmente lahko samo še dodajamo.

Pri obdelavi rezultatov testiranja prvega niza spojin in izračunu občutljivosti je potrebno omeniti, da smo spojine iz preglednice 3, in sicer tiste od vinil klorida do mireksa (12 spojin), srečali že med predstavniki posameznih strukturnih elementov na Benigni / Bossa-ejevem seznamu (27). V primeru morebitne uporabe le teh pri izgradnji, bi lahko dobili previsoke vrednosti za napovednost; to predstavlja tudi največji problem pri vrednotenju napovednosti posameznega ekspertnega sistema (16). Ker nismo nikjer zasledili, da naj bi omenjene spojine bile uporabljene pri izgradnji programskega modula, smo jih pri ocenjevanju napovednosti vseeno uporabili. Iz tega lahko sklepamo, da lahko strokovnjaki, ki so sodelovali pri izgradnji programskega modula, najbolj ustrezno ovrednotijo njegovo napovednost, saj le oni najboljše vedo, katere spojine so bile uporabljene pri njegovi izgradnji.

4.2 Specifičnost za napovednost

Pri določanju specifičnosti smo uporabili rezultate testiranja drugega testnega niza (Preglednica 5). Ta prav tako kot prvi zajema 34 spojin. V preglednici 5 so zbrana vsa imena in CAS številke posameznih spojin, podatki o morebitni mutagenosti in kancerogenosti spojin, pridobljeni iz CPDB, in seveda rezultati testiranja proučevanega programskega modula.

Na enem mestu smo zbrali tudi vse strukturne formule testiranih spojin, njihove SMILES kode in namene uporabe posameznih spojin (Priloga II).

Pred obdelavo rezultatov smo spojine drugega niza, po strukturnih elementih razvrstili v skupine (Preglednica 6). V prvem stolpcu preglednice 6 so po vrsti zbrani Benigni / Bossa-ejevi strukturni elementi; med njimi so z modro barvo označeni strukturni elementi za negenotoksično kancerogenost, vsi ostali pa predstavljajo strukturne elemente za genotoksično kancerogenost. V tretjem stolpcu so zapisane vse tiste spojine drugega niza, ki vsebujejo katerega od strukturnih elementov. V drugem testnem nizu smo uspeli razvrstiti le štiri spojine, in sicer sukralozo, FD&C Red no.3 (E127, eritrozin), FD&C Blue no. 2 (E132, indigo karmin) ter acetaminofen.

Preglednica 5: Rezultati testiranja spojin drugega niza.

spojina	CAS	CPDB ^(a)		Toxtree ^(b)								
		kancer.	mutag.	1	2	3	4	5	6	7	8	9
acetilsalicilna kislina	50-78-2	-	-			+						
acetaminofen	103-90-2	+	-	+				+	+			
piroksikam	36322-90-4	-	/			+						
atropin	51-55-8	-	/			+						
fenotiazin	92-84-2	-	-			+						
anetol	104-46-1	-	-			+						
acetonitril	75-05-8	-	-			+						
fenol	108-95-2	-	-			+						
FD&C Yellow no. 5 (E102, tartrazin)	1934-21-0	-	-			+						
FD&C Yellow no. 6 (E110, Sunset Yellow)	2783-94-0	-	-			+						
FD&C Red no. 3 (E127, eritrozin)	16423-68-0	-	-	+								
FD&C Blue no. 2 (E132, indigo karmin)	860-22-0	-	-	+								
sorbinska kislina (E200)	110-44-1	-	-			+						
benzojska kislina (E210)	65-85-0	-	-			+						
etilparaben (E214)	120-47-8	/	/			+						
metilparaben (E218)	99-76-3	/	/			+						
bifenil (E230)	92-52-4	-	-			+						
heksametilentetramin (E239)	100-97-0	-	+			+						
L-askorbinska kislina (E300)	50-81-7	-	-			+						
natrijev eritorbat (E316)	6381-77-7	-	-			+						
butiliran hidroksianizol (E320)	25013-16-5	+	-			+						
butiliran hidroksitoluen (E321)	128-37-0	+	-			+						
citronska kislina (330)	77-92-9	-	/			+						
glicerol (E422)	56-81-5	-	-			+						
propil galat (E310)	121-79-9	-	-			+						
4-heksilresorcinol (E586)	136-77-6	-	-			+						

spojina	CAS	CPDB ^(a)		Toxtree ^(b)								
		kancer.	mutag.	1	2	3	4	5	6	7	8	9
glutaminska kislina (E620)	56-86-0	-	-			+						
glicin (E640)	56-40-6	+	-			+						
aspartam (E951)	22839-47-0	-	/			+						
natrijev ciklamat (E952)	139-05-9	+	/			+						
izomalt (E953)	64519-82-0	-	/			+						
saharin (E954)	81-07-2	-	-			+						
sukraloza (E955)	56038-13-2	-	/	+								
propilen glikol (E1520)	57-55-6	-	-			+						

^(a)Podatki o mutagenosti in kancerogenosti spojin, ki smo jih črpali iz podatkovne baze CPDB (The Carcinogenic Potency Database). Z znakom + so označene mutagene oziroma kancerogene spojine, z – so označene nemutagene oziroma nekancerogene spojine, oznaka / pa pomeni, da podatka o mutagenosti oziroma kancerogenosti v CPDB nismo našli.

^(b)Rezultati testiranja drugega niza spojin s Toxtree-jevim programskim modulom za napovedovanje mutagenosti in kancerogenosti. Številke 1 – 9 označujejo devet možnih izidov modula, ki so opisani v poglavju »Materiali in metode«, natančneje v podpoglavju »Analiza rezultatov testiranja« na strani 34.

Preglednica 6: Razvrstitev drugega niza spojin po strukturnih elementih, ki jih prepozna Toxtree.

Skupine strukturnih elementov ^(a)	Št. strukt.elem. ^(b)	Spojina ^(c)
acil halidi	SA 1	
alkilni ali benzilni estri sulfon. ali fosfon. kislin	SA 2	
N-metilol derivati	SA 3	
monohaloalkeni	SA 4	
S ali N iperiti	SA 5	
propiolaktoni ali propiosultoni	SA 6	
epoksidi in aziridini	SA 7	
alifatski halogeni	SA 8	sukraloza
alkil nitriti	SA 9	
α,β -nenasičeni karbonili	SA 10	E132, E127
enostavni aldehidi	SA 11	
kinoni	SA 12	
hidrazini	SA 13	
alifatske azo in azoksi spojine	SA 14	
izocianati in izotiocianati	SA 15	
alkil karbamati in tiokarbamati	SA 16	
tiokarbonili	SA 17	
policiklični aromatski ogljikovodiki	SA 18	
heterociklični policiklični aromatski CH	SA 19	
halogenirani cikloalkani	SA 20	
alkil in aril nitrozo skupine	SA 21	
azidi in triazeni	SA 22	
alifatska N-nitro skupina	SA 23	
α,β -nenasičena alifatske alkoksi skupina	SA 24	
aromska nitrozo skupina	SA 25	
aromatski obroč N-oksid	SA 26	
nitro aromati	SA 27	
primarni arom. amini, hidroksil amini in derivati	SA 28	
aromatski mono in dialkil amini	SA 28bis	
aromatski N-acilamini	SA 28ter	acetaminofen
aromske diazo spojine	SA 29	
kumarini in furokumarini	SA 30	
halogenirani benzeni	SA 31a	
halogenirani PAH	SA 31b	
halogenirani dibenzoksini	SA 31c	

^(a)Skupine strukturnih elementov, ki so navedene na Benigni / Bossa-ejevem seznamu (27) in jih prepozna program Toxtree. Modro so označeni strukturni elementi za negenotoksično kancerogenost, vsi ostali pa so strukturni elementi za genotoksično kancerogenost.

^(b)Številka skupine posameznega strukturnega elementa iz Benigni / Bossa-ejevega seznama.

^(c)Spojine drugega niza, ki nam jih je uspelo razvrstiti v eno od skupin strukturnih elementov.

Na podlagi rezultatov testiranja drugega niza spojin (Preglednica 5), smo s pomočjo spodnje formule izračunali specifičnost Toxtree-jevega programskega modula za napovedovanje mutagenosti in kancerogenosti spojin. Specifičnost, ki v našem primeru predstavlja odstotek biološko neaktivnih-negativnih spojin, ki jih je programski modul pravilno opredelil kot nekancerogene in nemutagene, je znašala **88,9 %**.

$$\text{Specifičnost} = (N_{M-K-} / N_{LM-LK-}) \times 100$$

$$\text{Specifičnost} = (24 / 27) \times 100$$

$$\text{Specifičnost} = 88,9 \%$$

Pri izračunu specifičnosti od 34-ih spojin nismo upoštevali rezultatov testiranja šestih spojin, saj smo zanje v podatkovni bazi CPDB, vsaj za eno od obeh proučevanih lastnosti ugotovili, da je pozitivna (Preglednica 5). Te spojine so bile naslednje: acetaminofen, heksametilentetramin, butiliran hidroksianizol, butiliran hidroksitoluen, glicin in natrijev ciklamat. Že predhodno smo namreč postavili zahtevo, da mora biti spojina v podatkovni bazi CPDB, za obe lastnosti opredeljena negativno – takšne spojine smo šteli kot nemutagene in nekancerogene. Pri izračunu nismo upoštevali še rezultata testiranja saharina. Tu smo naredili izjemo; saharin je namreč v CPDB za obe lastnosti opredeljen negativno, v knjigi avtorja Klaasen-a, Casarett & Doull's Toxicology, the basic science of poisons (35), pa je označen kot nemutagen kemijski kancerogen in kot promotor (snov, ki povzroči klonalno razširitev mutiranih celic) ter ga zato pri izračunu specifičnosti nismo upoštevali. Vseh sedem zgoraj omenjenih spojin, razen acetaminofena, je programski modul označil z izidom 3 (ni strukturnih elementov za kancerogeno aktivnost), čeprav so v podatkovni bazi CPDB vsaj za eno od proučevanih lastnosti bile opredeljene pozitivno. Izjema je bil torej acetaminofen, v katerem je programski modul prepoznal enega od strukturnih elementov in ga je zato označil z izidom 1 (strukturni element za genotoksično kancerogenost). Acetaminofen, ki ga vsi bolje poznamo kot paracetamol, je torej kljub obširni uporabi kot šibki analgetik, tako v podatkovni bazi CPDB kot tudi glede na rezultate našega testiranja, kancerogen. Na to lastnost bi bilo potrebno ljudi, ki jemljejo paracetamol, v lekarni vedno znova opozoriti in jim dati jasna navodila glede odmerjanja. Na splošno lahko rečemo, da je za delovanje spojin v človeškem organizmu bistvenega

pomena velikost odmerka (tu ima pomembno vlogo prag toksičnosti), pomembni pa so tudi pogostost odmerjanja, način aplikacije, stanje (zdravstveno; morebitna nosečnost) ter starost človeka. Če se na tem mestu osredotočimo na prag toksičnosti (to je tista koncentracija snovi, pod katero ni izraženih toksičnih učinkov), lahko rečemo, da ta odpove pri snoveh, ki povzročajo raka. Ne moremo namreč vedeti, kako visok odmerek je potreben za razvoj rakavega obolenja. Odvisnosti toksičnega učinka od odmerka torej na primeru mutagenih in kancerogenih snovi ne moremo določiti. Pomembno pa je poudariti, da se ob doslednem upoštevanju navodil jemanja (ustrezna complianca), v našem primeru paracetamola, ta neželen učinek naj ne bi smel pojaviti, saj bi v nasprotnem primeru morali to zdravilno učinkovino že zdavnaj umakniti iz prometa.

V imenovalcu zgornje formule za izračun specifičnosti smo vstavili število nemutagenih in nekancerogenih spojin (N_{LM-LK-}) drugega testnega niza. Teh je bilo 27. V števec pa smo vstavili število spojin, ki so v literaturi opredeljene kot nemutagene in nekancerogene in so tudi pri testiranju z ekspertnim sistemom dale enak (negativen) izid (N_{M-K-}). Teh je bilo skupaj 24. Programski modul je torej za tri spojine, in sicer za FD&C Red no.3 (E127, eritrozin), FD&C Blue no.2 (E132, indigo karmin) in sukralozo, glede na obravnavane literaturne podatke in podatke iz CPDB, podal napačno napoved o njihovi mutagenosti in kancerogenosti. Označil jih je namreč z izidom 1 (strukturni element za genotoksično kancerogenost), saj je v njih prepoznal enega od strukturnih elementov iz Benigni / Bossa-ejevega seznama le teh (Preglednica 6). Glede na to, da gre pri omenjenih treh spojinah za prehranske aditive, bi bilo napoved programskega modula smiselno upoštevati in tem spojinam, pri proučevanju njihovih stranskih učinkov, nameniti več pozornosti.

Na tem mestu bi se bilo potrebno zamisliti nad dejstvom, kaj vse človek tekom svojega življenja vnese v svoj organizem, pa naj bo to s hrano, z zdravili ali pa zaradi stika z najrazličnejšimi kemikalijami, ki jih zasledimo povsod okrog nas.

4.3 Točnost za napovednost

Točnost programskega modula smo določili s pomočjo spodnje formule.

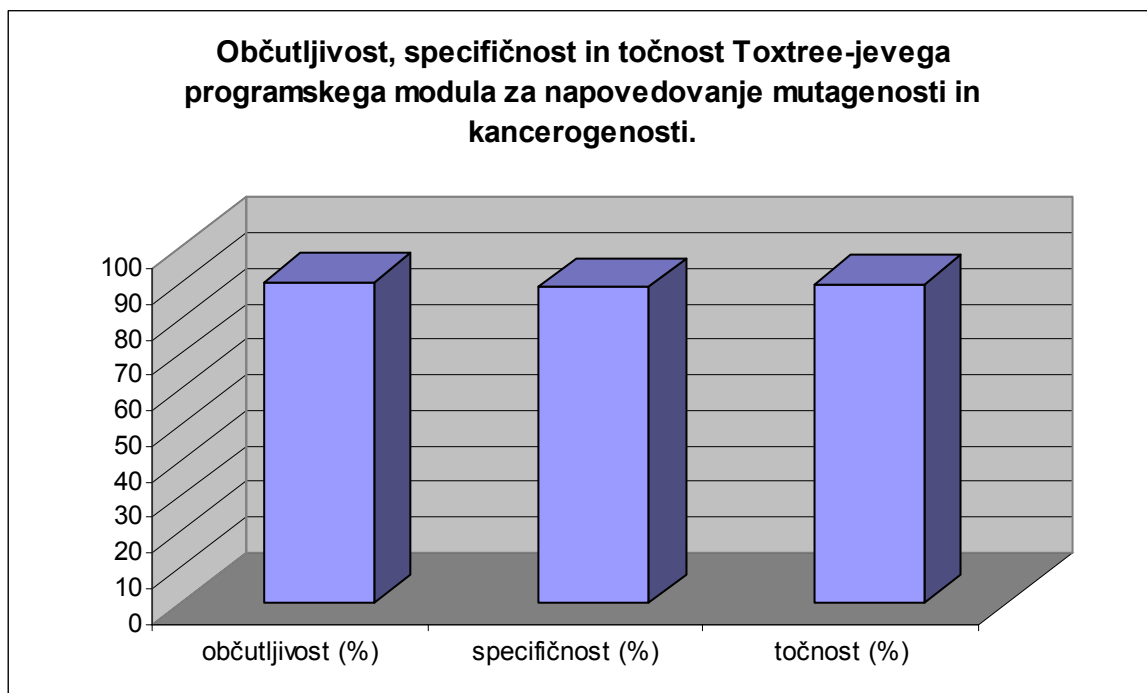
$$\text{Točnost} = ((N_{M+K+} + N_{M-K-}) / N_{\text{eval}}) \times 100$$

$$\text{Točnost} = ((28 + 24) / (31 + 27)) \times 100$$

$$\text{Točnost} = 89,7 \%$$

Od vseh spojin je torej programski modul pravilno opredelil **89,7 %** testiranih spojin, ki smo jih vključili v izračune občutljivosti in specifičnosti.

Če na enem mestu povzamemo izračunane vrednosti testiranja za občutljivost (90,3 %), specifičnost (88,9 %) in točnost (89,7 %), lahko ugotovimo, da so te vrednosti precej visoke. To v prvi vrsti prav gotovo zavisi od izbora spojin obeh testnih nizov in od števila testiranih spojin. Celokupna napovednost Toxtree-jevega programskega modula za napovedovanje mutagenosti in kancerogenosti spojin, je torej glede na splošno podano literaturno vrednost za napovednost, ta naj bi bila po poročanju številnih avtorjev pogosto večja kot 90 %, povsem ustrezna.



V literaturi sicer glede vrednotenja napovednosti Toxtree-jevega programskega modula za napovedovanje mutagenosti in kancerogenosti spojin, nismo zasledili veliko podatkov, zato našega rezultata ne moremo primerjati z literaturnimi vrednostmi. Našli smo le en podatek o točnosti, natančneje o točnosti napovedi strukturnih elementov, ki so vključeni na Benigni / Bossa-ejev seznam. Navedeni sta bili ločeni vrednosti napovednosti (celokupne točnosti) za kancerogenost (70 %) in mutagenost (78 %) omenjenega programskega modula (27,28).

V primerjavi z napovednostjo drugih ekspertnih sistemov, s katerimi prav tako lahko napovedujemo mutagenost in kancerogenost spojin (npr. TOPKAT, MCASE, DEREK), je izračunana napovednost (točnost napovedi) za Toxtree-jev programski modul nekoliko višja (21). Ob tem se moramo zavedati, da ima tu glavni pomen izbor in število testiranih spojin. Zato lahko rečemo, da je takšna primerjava med ekspertnimi sistemi, ki napovedujejo isto lastnost spojin, napovednost pa vrednotijo s pomočjo popolnoma različnih testnih nizov, lahko le informativnega značaja. Najbolj ustrezna primerjava (direktna) med ekspertnimi sistemi bi bila prav gotovo tista, ki bi med seboj primerjala rezultate testiranj različnih ekspertnih sistemov pridobljenih na istem testnem nizu spojin (pri čemer bi seveda morali upoštevati uporabno domeno posameznega ekspertnega sistema) (37). Čeprav bi to zahtevalo veliko dela, predvsem pri izboru dovolj velikega testnega niza spojin, bi se to splačalo izvesti, saj bi na ta način ugotovili, s katerim programom lahko najbolj natančno napovemo mutagenost in kancerogenost spojin.

Pri vrednotenju napovednosti se moramo zavedati, da je povsem nerealno pričakovati večjo točnost napovedi računalniškega programa, kot pa je točnost eksperimentalnih podatkov, na podlagi katerih je bil program zgrajen. Reproducibilnost/obnovljivost med različnimi laboratoriji ali pa znotraj istega laboratorija ob različnih časih, in sicer na istem nizu spojin, za katere ni bil opravljen še noben test, znaša 85%. V primerjavi z 85% zanesljivostjo bioloških analiz, ki so uporabljene pri izgradnji modela, komercialni računalniški programi dosegajo 60 % do 90 % celokupno točnost napovedi (odvisno od niza podatkov, ki ga vrednotijo) (16).

Napovednost računalniških programov lahko izboljšamo na dva načina, in sicer z neprestanim izpopolnjevanjem in dopolnjevanjem baz podatkov ter baz pravil izpeljanih iz

omenjenih podatkov (znotraj posameznega ekspertnega sistema) in s kombiniranjem različnih individualnih modelov. Slednji način močno izboljša celokupno točnost in specifičnost in na ta način zmanjša omejitve posameznih modelov (21). Z izboljševanjem in izpopolnjevanjem ekspertnih sistemov se danes strokovnjaki stalno ukvarjajo, zato imajo tovrstne alternativne metode pri napovedovanju toksičnih lastnosti spojin tudi vedno večjo vrednost in pomen.

4.4 Ugotavljanje prisotnosti komponente metabolične aktivacije v programskem modulu

Kot smo že pri določanju občutljivosti izpostavili, smo za vse spojine prvega niza v literaturi preverili, katera od njih se mora predhodno metabolično aktivirati, da postane mutagena in kancerogena. Spojine, ki izkazujejo takšno lastnost, smo zbrali v posebni preglednici (Preglednica 7).

Pri izračunu točnosti napovedi za omenjene spojine, smo v števec vstavili število spojin, ki jih je programski modul pravilno opredelil – 16 spojin, v imenovalce pa smo vstavili število vseh spojin prvega niza, ki za kancerogeno in mutageno delovanje potrebujejo predhodno metabolično aktivacijo – 19 spojin. Tako je točnost napovedi iz vidika metabolične aktivacije znašala **84,2 %**. Programski modul je samo na primeru treh spojin, in sicer akrilonitrila, benzena in 1,3-butadiena, storil napako. Za te tri spojine smo že pri določanju občutljivosti ugotovili, da ne vsebujejo nobenega strukturnega elementa iz Benigni / Bossa-ejevega seznama, zato jih je programski modul enostavno opredelil negativno. V primeru, da bi v program vnesli metabolični produkt posamezne spojine, ki v vseh treh primerih vsebuje enega od strukturnih elementov (pri benzenu je to kinonska struktura – SA 12, pri akrilonitrilu in 1,3-butadienu pa je to epoksidna struktura – SA 7), pa bi jo le ta pravilno opredelil. Na podlagi teh rezultatov smo ugotovili, da je pred napovedovanjem mutagenosti in kancerogenosti spojin potrebno v literaturi preveriti, kako se spojina v organizmu metabolizira in kakšni so njeni metabolični produkti. V primeru, da programski modul opredeli spojino kot nemutageno in nekancerogeno, je še vedno smiselno testirati metabolične produkte obravnavane spojine, saj so le ti v večini primerov bistvenega pomena za njeno toksično delovanje.

Kljub temu da nikjer nismo zasledili, da bi bila v Toxtree-jev programski modul za napovedovanje mutagenosti in kancerogenosti, vgrajena komponenta metabolične aktivacije, nas je visok odstotek, 84,2 %, kar presenetil. Takšen rezultat lahko pripisujemo dejstvu, da je večina obravnavanih spojin vsebovala enega od strukturnih elementov in jih je zato programski modul, ki v glavnem deluje na podlagi le teh, pravilno opredelil.

Preglednica 7: Spojine prvega niza, ki se morajo predhodno metabolično aktivirati, da postanejo mutagene in kancerogene.

spojina	CAS	metabolična aktivacija ^(a)		CPDB		Toxtree ^(b)								
		Douls	ScienceDirect/(ref.)	kancer.	mutag.	1	2	3	4	5	6	7	8	9
benzidin	92-87-5	+	+/ (38)	+	+	+			+			+		
p-aminodifenil	92-67-1	+		+	+	+			+		+			
2-naftilamin	91-59-8	+		+	+	+			+		+			
N,N-dietilnitrozamin	55-18-5		+/ (39)	+	+	+								
DAB	60-11-7	+	+/ (40)	+	+	+			+		+			
ciklofosfamid	50-18-0		+/ (41)	+	+	+								
bromodiklorometan	75-27-4		+/ (42)	+	+	+								
NNK (nitrozamin)	64091-91-4		+/ (43)	+	/	+								
vinil klorid	75-01-4		+/ (44)	+	+	+								
aflatoksin B1	1162-65-8	+	+/ (45)	+	+	+								
uretan	51-79-6		+/ (46)	+	+	+								
benz[a]piren	50-32-8	+		+	+	+								
2-acetilaminofluoren	53-96-3		+/ (47)	+	+	+			+		+			
1,2-dibromoetan	106-93-4	+	+/ (48)	+	+	+								
kumarin	91-64-5		+/ (49)	+	+	+								
TCDD	1746-01-6		+/ (50)	+	-		+							
akrilonitril	107-13-1		+/ (51)	+	+			+						
benzen	71-43-2		+/ (36)	+	-			+						
1,3-butadien	106-99-0		+/ (52)	+	+			+						

^(a) Podatki o metabolični aktivaciji spojin, ki smo jih črpali iz dveh virov; Casarett & Doull's Toxicology, the basic science of poisons, avtorja Klaasen-a (35) in iz baze člankov, ki jih nudi ScienceDirect (<http://www.sciencedirect.com/>); + označuje, kje smo našli podatek o metabolični aktivaciji; s številko v oklepaju pa je označena referenca (članek naveden v literaturi), ki se nanaša na metabolično aktivacijo posamezne spojine.

^(b) Številke 1 – 9 označujejo možne izide modula, ki so opisani v poglavju »Materiali in metode«, natančneje v podpoglavju »Analiza rezultatov testiranja« na strani 34.

5 SKLEP

Cilj diplomskega dela je bil ovrednotiti napovednost Toxtree-jevega programskega modula za napovedovanje mutagenosti in kancerogenosti spojin (Benigni / Bossa rulebase for mutagenicity and carcinogenicity). Na podlagi podatkov v literaturnih člankih o različnih načinih določanja napovednosti, smo se odločili za vrednotenje omenjene karakteristike iz vidika občutljivosti, specifičnosti in točnosti. Pri tem smo opravili testiranje 68-ih naključno izbranih spojin, ki smo jih razdelili na dva testna niza. Prvi testni niz je sestavljalo 34 spojin, za katere smo v literaturi zasledili, da so dokazano mutagene in kancerogene; drugi testni niz pa je vseboval 34 spojin, ki jih vsakodnevno uporablja veliko število ljudi (prehranski aditivi, kemikalije in zdravilne učinkovine) in zanje v literaturi nismo zasledili, da bi bile mutagene in kancerogene. Vse spojine smo glede mutagenosti in kancerogenosti preverili v podatkovni bazi CPDB.

Za vse spojine prvega niza smo, bodisi v bazi člankov, ki jih nudi podatkovna baza ScienceDirect, bodisi v knjigi Casarett & Doull's Toxicology, the basic science of poisons, avtorja Klaasen-a, preverili, katera od njih se mora predhodno metabolično aktivirati, da postane mutagena in kancerogena. Vse spojine z omenjeno karakteristiko smo še dodatno obravnavali in tako na podlagi rezultatov testiranja s programskim modulom preverili, v koliko primerih modul spojine ustrezno opredeli (v koliko odstotkih poda točno napoved).

Rezultati testiranja spojin obeh testnih nizov so pokazali, da so vrednosti za občutljivost (90,3 %), specifičnost (88,9 %) in točnost (89,7 %), z ozirom na literaturne podatke drugih ekspertnih sistemov pri napovedovanju mutagenosti in kancerogenosti, zadovoljivi (21). Čeprav je program prosto dostopen vsakomur, lahko z njegovim modulom, po naših izračunih, pri 89,7 % testiranih spojin, mutagenost in kancerogenost ustrezno napovemo. Programski modul je pravilno ocenil tudi 84,2 % spojin, ki šele po predhodni metabolični aktivaciji postanejo mutagene in kancerogene. Kljub dokaj visokemu odstotku bi bilo pred vsakim testiranjem spojin z omenjenim programskim modulom v literaturi smiselno preveriti njihove metabolične produkte in tudi te testirati. Na ta način, bi lahko dobili še višji odstotek pravilnih napovedi.

Glede na to, da danes obstaja že veliko število tako komercialnih (TOPKAT, MCASE, DEREK, HazardExpert) kot tudi nekomercialnih (OncoLogic, Lazar) računalniških

programov za napovedovanje mutagenosti in kancerogenosti, bi bilo smiselno enak niz spojin testirati še z drugimi. Na ta način bi lahko na osnovi rezultatov testiranja spojin z različnimi programi, s pomočjo primerjave ugotovili, kateri od njih izkazuje najboljšo napovednost. Poleg tega bi lahko ocenili, kateri pristop napovedovanja mutagenosti in kancerogenosti spojin (različni računalniški programi oziroma ekspertni sistemi namreč uporabljajo pri napovedovanju različne pristope in različna pravila) danes velja za najustreznejšega. Nenazadnje pa bi lahko ugotovili tudi, ali imajo računalniški programi, ki napovedujejo samo mutagenost ali kancerogenost spojin (tisti, ki so specialni za napovedovanje te vrste toksičnosti) v primerjavi s tistimi, ki jih sestavlja več različnih modulov za napovedovanje (tisti, ki so sposobni napovedati več različnih toksičnih lastnosti), značilno višjo stopnjo napovednosti ali ne.

6 LITERATURA

1. Benfenati E, Gini G. Computational predictive programs (expert systems) in toxicology. *Toxicology* 1997; 119: 213-25.
2. Gini G, Testaguzza V, Benfenati E, Todeschini R. Hybrid toxicology expert system: architecture and implementation of a multi-domain hybrid expert system for toxicology. *Chemom Intell Lab Sys* 1998; 43: 135-45.
3. Van Leeuwen CJ, Vermeire TG. Risk Assessment of Chemicals: An Introduction. Second Edition, Springer, Dordrecht, 2007.
4. <http://habitat.igc.org/agenda21/> (september 2009)
5. Garner I, Spielman H, Hoefer T, Liebsch M. Regulatory use of QSARs in assessment strategies. *SAR QSAR Environ Res* 2004; 15: 359-66.
6. Combes R, Barratt M, Balls M. An Overall Strategy for the Testing of Chemicals for Human Hazard and Risk Assessment under the EU REACH System. *ATLA Altern Lab Anim* 2003; 31: 7-19.
7. Vogelgesang J. The EC White Paper on Strategy for a Future Chemicals Policy. *ATLA Altern Lab Anim* 2002; 30: 211-12.
8. <http://ecb.jrc.ec.europa.eu/qsar/> (avgust 2009)
9. http://www.uk.gov.si/si/delovna_podrocja/reach/ (september 2009)
10. Schaafsma G, Kroese ED, Tielemans EL, Van de Sandt JJ, Van Leeuwen CJ. REACH, non testing approaches and the urgent need for a change in mind set. *Regul Toxicol Pharmacol* 2009; 53: 70-80.

11. http://www.oecd.org/document/6/0,3343,en_2649_34379_43087494_1_1_1_1,00.html (september 2009)
12. http://echa.europa.eu/home_sl.asp (september 2009)
13. Sollner Dolenc, Marija. Use of QSAR in risk assessment: predavanje na izobraževanju z naslovom: Basics of Toxicology with Introduction to Risk Assessment STOX 2008: program. Bled, September 23, 2008.
14. Hartung T, Bremer S, Casati S, Coecke S, Corvi R, Fortaner S, Gribaldo L, Halder M, Roi AJ, Prieto P, Sabbioni E, Worth A, Zuang V. ECVAM's Response to the Changing Political Environment for Alternatives: Consequences of the European Union Chemicals and Cosmetic Policies. ATLA Altern Lab Anim 2003; 31: 473-81.
15. Bhogal N, Gridon C, Combes R, Balls M. Toxicity testing: creating a revolution based on new technologies. Trends Biotechnol 2005; 23: 299-307.
16. Ekins S. Computational Toxicology: Risk Assessment for Pharmaceutical and Environmental Chemicals. John Wiley & Sons, Inc., New Jersey, 2007.
17. <http://ecb.jrc.ec.europa.eu/tgd/> (avgust 2009)
18. http://www.oecd.org/document/29/0,3343,en_2649_34379_42675741_1_1_1_1,00.html (Introduction to (Quantitative) Structure Activity Relationships, september 2009)
19. http://www.oecd.org/document/4/0,3343,en_2649_34379_42926724_1_1_1_1,00.html (september 2009)
20. Hayashi M, Kamata E, Hirose A, Takahashi M, Morita T, Ema M. In silico assessment of chemicals mutagenesis in comparison with results of Salmonella microsome assay on 909 chemicals. Mutat Res 2005; 588: 129-35.

21. White AC, Mueller RA, Gallavan RH, Aaron S, Wilson AGE. A multiple in silico program approach for the prediction of mutagenicity from chemical structure. *Mutat Res* 2003; 539: 77-89.
22. Pavan M, Worth AP: Publicly-accessible QSAR software tools developed by the Joint Research Centre. *SAR QSAR Environ Res* 2008; 19:785-99.
23. <http://ecb.jrc.ec.europa.eu/qsar/qsar-tools/index.php?c=TOXTREE> (september 2009)
24. <http://toxtree.sourceforge.net/> (september 2009)
25. <http://ecb.jrc.ec.europa.eu/testing-methods/annex5/> (oktober 2009)
26. Rang HP, Dale MM, Ritter JM, Flower RJ. Rang and Dale's Pharmacology 6th Edition. Churchill Livingstone Elsevier 2007.
27. Benigni R, Bossa C, Jeliaskova N, Netzeva T, Worth A. The Benigni / Bossa rulebase for mutagenicity and carcinogenicity – a module of Toxtree. JRC Scientific and Technical Reports 2008.
28. Valerio LG Jr. In silico toxicology for the pharmaceutical sciences. *Toxicol Appl Pharmacol* 2009; 241: 356-70.
29. <http://www.daylight.com/smiles/index.html> (november 2009)
30. http://en.wikipedia.org/wiki/Simplified_molecular_input_line_entry_specification (november 2009)
31. <http://www.daylight.com/dayhtml/doc/theory/theory.smiles.html> (november 2009)
32. Greene N. Computer systems for the prediction of toxicity: an update. *Adv Drug Deliv Rev* 2002; 54: 417-31.

33. Benigni R. Computational prediction of drug toxicity: the case of mutagenicity and carcinogenicity. *Drug Discov Today* 2004; 1: 457-62.
34. Tsakovska I, Worth A. The Use of Computational Methods for the Assessment of Chemicals in REACH. *BIO Automation* 2009; 13: 151-62.
35. Klaasen CD. Casarett & Doull's Toxicology, the basic science of poisons, 6th Edition, Mc Graw-Hill Companies, Inc., USA, 2001.
36. Xie Z, Zhang Y, Guliaev AB, Shen H, Hang B, Singer B, Wang Z. The p-benzoquinone DNA adducts derived from benzene are highly mutagenic. *DNA Repair (Amst)* 2005; 4: 1399-409.
37. Snyder RD, Smith MD. Computational prediction of genotoxicity: room for improvement. *Drug Discov Today* 2005; 10: 1119-24.
38. Lynn RK, Garvie-Gould CT, Milam DF, Scott KF, Eastman CL, Ilias AM, Rodgers RM. Disposition of the aromatic amine, benzidine, in the rat: Characterization of mutagenic urinary and biliary metabolites. *Toxicol Appl Pharmacol* 1984; 72: 1-14.
39. Aiub CA, Mazzei JL, Pinto LF, Felzenszwalb I. Evaluation of nitroreductase and acetyltransferase participation in N-nitrosodiethylamine genotoxicity. *Chem Biol Interact* 2006; 161: 146-54.
40. Katlubar FF, Ketterer B, Flammang TJ, Christodoulides L. Formation of 3-(glutathion-S-yl)-N-methyl-4-aminoazobenzene and inhibition of aminoazo dye-nucleic acid binding in vitro by reaction of glutathione with metabolically-generated N-methyl-4-aminoazobenzene-N-sulfate. *Chem Biol Interact* 1980; 31: 265-78.
41. Nau H, Spielman H, Lo Turco Morther CM, Winckler K, Riedel L, Obe G. Mutagenic, teratogenic and pharmacokinetic properties of cyclophosphamide and some of its deuterated derivatives. *Mutat Res* 1982; 95: 105-18.

42. Allis JW, Anderson BP, Zhao G, Ross TM, Pegram RA. Evidence for the involvement of CYP1A2 in the metabolism of bromodichloromethane in rat liver. *Toxicology* 2002; 176: 25-37.
43. Maser E, Stinner B, Atalla A. Carbonyl reduction of 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) by cytosolic enzymes in human liver and lung. *Cancer Lett* 2000;148: 135-44.
44. Vainio H. Vinyl chloride and vinyl benzene (styrene) – Metabolism, mutagenicity and carcinogenicity. *Chem Biol Interact* 1978; 22: 117-24.
45. Daniels JM, Massey TE. Modulation of aflatoxin B1 biotransformation in rabbit pulmonary and hepatic microsomes. *Toxicology* 1992; 74: 19-32.
46. Sakano K, Oikawa S, Hiraku Y, Kawanishi S. Metabolism of carcinogenic urethane to nitric oxide is involved in oxidative DNA damage. *Free Radic Biol Med* 2002;33: 703-14.
47. Heflich RH, Neft RE. Genetic toxicity of 2-acetylaminofluorene, 2-aminofluorene and some of their metabolites and model metabolites. *Mutat Res* 1994; 318:73-174.
48. van Bladeren PJ, Breimer DD, Rotteveel-Smijs GM, de Jong RA, Buijs W, van der Gen A, Mohn GR. The role of glutathione conjugation in the mutagenicity of 1,2-dibromoethane. *Biochem Pharmacol* 1980; 29: 2975-82.
49. Lake BG. Coumarin Metabolism, Toxicity and Carcinogenicity: Relevance for Human Risk Assessment. *Food Chem Toxicol* 1999; 37: 423-53.
50. Atalla A, Maser E. Carbonyl reduction of the tobacco-specific nitrosamine 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) in cytosol of mouse liver and lung. *Toxicology* 1999; 139: 155-66.

51. Léonard A, Gerber GB, Stecca C, Rueff J, Borba H, Farmer PB, Sram RJ, Czeizel AE, Kalina I. Mutagenicity, carcinogenicity and teratogenicity of acrylonitrile. *Mutat Res* 1999; 436: 263-83.
52. Filser JG, Hutzler C, Meischner V, Veereshwarayya V, Csanády GA. Metabolism of 1,3-butadiene to toxicologically relevant metabolites in single-exposed mice and rats. *Chem Biol Interact* 2007; 166: 93-103.

Literatura za uporabo spojin: http://en.wikipedia.org/wiki/Main_Page (november 2009).

7 PRILOGA I

SMILES kode, CAS številke, nameni uporabe in strukturne formule spojin **prvega** testnega niza.

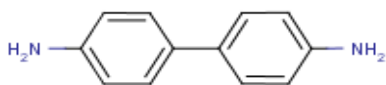
- BENZIDIN

SMILES:

NC1=CC=C(C2=CC=C(N)C=C2)C=C1

CAS: 92-87-5

Uporaba: intermediat, barvilo

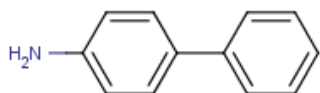


- P-AMINODIFENIL

SMILES: NC1=CC=C(C=C1)C2=CC=CC=C2

CAS: 92-67-1

Uporaba: plemenitenje pnevmatik

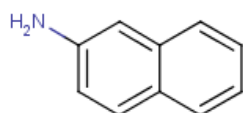


- 2-NAFTILAMIN

SMILES: C12=C(C=CC(=C1)N)C=CC=C2

CAS: 91-59-8

Uporaba: intermediat, barvilo

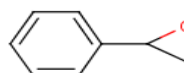


- STIREN OKSID

SMILES: C1C(C2=CC=CC=C2)O1

CAS: 96-09-3

Uporaba: intermediat

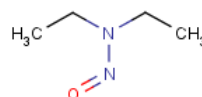


- N,N-DIETILNITROZAMIN
(DEN)

SMILES: CCN(CC)N=O

CAS: 55-18-5

Najdemo ga v hrani (pivo, ribe, meso), v
tobačnem dimu.



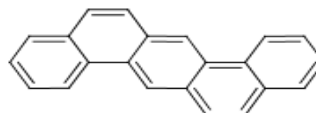
- DIBENZ(a,h)ANTRACEN

SMILES: C1=C2C=CC3=CC=CC=C3C2=CC4=C

C=C5C(=C14)C=CC=C5

CAS: 53-70-3

Uporaba: v literaturi ni znane uporabe



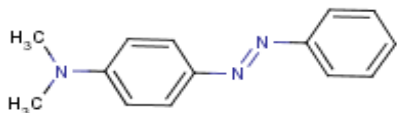
- N,N-DIMETIL-4-

AMINOAZOBENZEN (DAB)

SMILES: CN(C1=CC=C(C=C1)/N=N/C2=CC=C(C=C2)C

CAS: 60-11-7

Uporaba: pH indikator

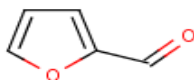


- FURFURAL

SMILES: O=CC1=CC=CO1

CAS: 98-01-1

Uporaba: topilo, intermediat v organski kemiji

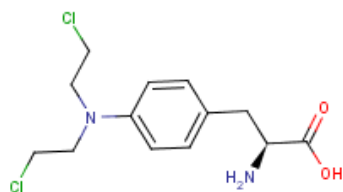


- MELFALAN

SMILES: OC(=O)[C@@H](N)Cc1ccc(cc1)N(CCCl)CCCl

CAS: 148-82-3

Uporaba: alkilirajoča zdravilna učinkovina

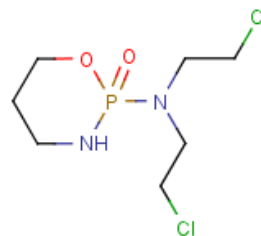


- CIKLOFOSFAMID

SMILES: C1CCN(CCCl)[P]1(=O)NCCCO1

CAS: 50-18-0

Uporaba: alkilirajoča zdravilna učinkovina

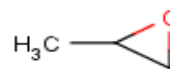


- 1,2-PROPILEN OKSID

SMILES: CC1CO1

CAS: 75-56-9

Uporaba: proizvodnja poliuretanske plastike; fumigant

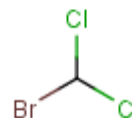


- BROMODIKLOROMETAN

SMILES: ClC(Cl)Br

CAS: 75-27-4

Uporaba: reagent, intermediat v organski kemiji

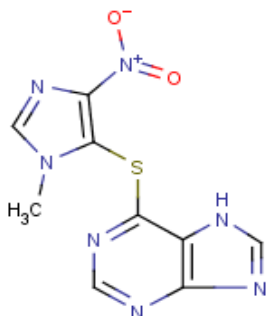


- AZATIOPRIN

SMILES: CN1C(=C(N=C1)[N+](=O)[O-])SC2=NC=NC3=C2N=CN3

CAS: 446-86-6

Uporaba: imunosupresiv

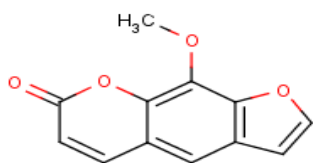


- 8-METOKSIPSORALEN

SMILES: COC1=C2C(=CC3=C1OC(=C3)C=CC(=O)O2

CAS: 298-81-7

Uporaba: zdravilna učinkovina za zdravljenje psoriaze, ekcemov, vitiliga

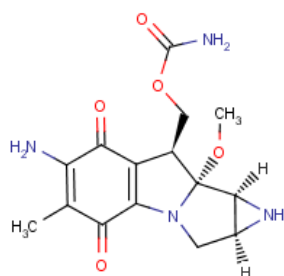


- MITOMICIN C

SMILES: NC(=O)OC[C@@H]3C=1C(=O)C(/N)=C(/C)C(=O)C=1N2C[C@@H]4N[C@@H]4[C@@]23OC

CAS: 50-07-7

Uporaba: citostatični antibiotik

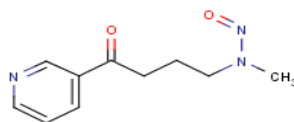


- 4-(METILNITROZAMINO)-1-(3-PIRIDIL)-1-BUTANON ali NNK

SMILES: O=C(C1=CC=CN=C1)CCCN(N=O)C

CAS: 64091-91-4

Najdemo ga v cigaretah, tobačnem dimu.

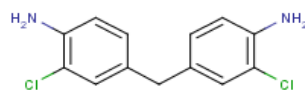


- 4,4-METILEN-BIS(2-KLOROANILIN)

SMILES: ClC1=CC(=CC(=C1)CC2=CC(=CC(=C2)N)Cl)N

CAS: 101-14-4

Uporaba: pri sintezi poliuretanov

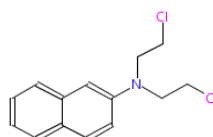


- KLORNAFAZIN

SMILES: C1=CC=C2C=C(C(=CC2=C1)N(CCCl)CCCl

CAS: 494-03-1

Uporaba: včasih so ga uporabljali za zdravljenje Hodgkin-ove bolezni in policitemije

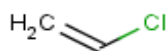


- VINIL KLORID

SMILES: C=CCl

CAS: 75-01-4

Uporaba: intermediat, proizvodnja polimerov

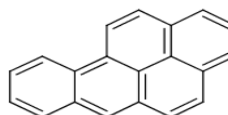


- BENZ[A]PIREN

SMILES: C1=CC2=CC=CC3=CC=C4C(=C23)C1=C5C(=C4)C=CC=C5

CAS: 50-32-8

Najdemo ga v premogovem katranu, cigaretne dimu, izpušnih plinih dizelskega goriva.

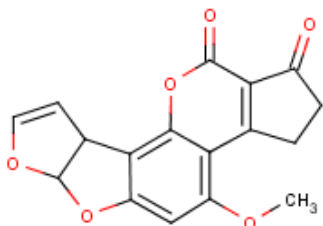


- AFLATOKSIN B1

SMILES: C12=C3C(C4=C(C(O3)=O)C(=O)CC4)=C(C=C1OC5C2C=CO5)OC

CAS: 1162-65-8

Opis: mikotoksin

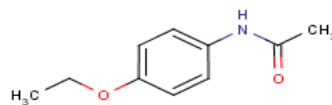


- FENACETIN

SMILES: CC(=O)NC1=CC=C(C=C1)OCC

CAS: 62-44-2

Uporaba: včasih so ga uporabljali kot nesteroidno protivnetno zdravilno učinkovino (NSAID)

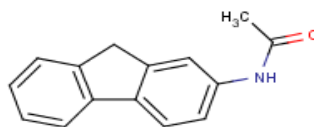


- 2-ACETILAMINOFLUOREN (2-AAF)

SMILES: C12C3=C(C=CC=C3)CC1=CC(=CC=2)NC(C)=O

CAS: 53-96-3

Uporaba: biokemično orodje pri študijah kancerogenov

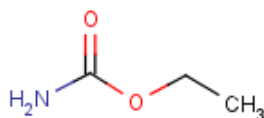


- URETAN

SMILES: NC(=O)OCC

CAS: 51-79-6

Uporaba: sinteza poliuretanov



- 1,2-DIBROMOETAN

SMILES: BrCCBr

CAS: 106-93-4

Uporaba: dodatek k bencinu, pesticid, fumigant;
reagent pri organskih sintezah

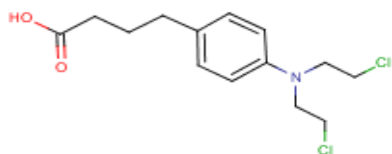


- KLORAMBUCIL

SMILES: ClCCN(C1=CC=C(C=C1)CCCC(=O)O)CCCl

CAS: 305-03-3

Uporaba: alkilirajoči citostatik



- TIOTEPA

SMILES: S=P(N1CC1)(N1CC1)N1CC1

CAS: 52-24-4

Uporaba: alkilirajoči citostatik

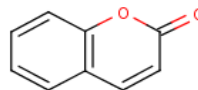


- KUMARIN

SMILES: O=C1OC2=C(C=CC=C2)C=C1

CAS: 91-64-5

Uporaba: v kozmetiki; rodenticid



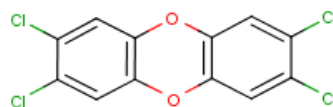
- 2,3,7,8-

TETRAKLORODIBENZO-p-
DIOKSIN (TCDD)

SMILES: ClC1=CC2=C(C=C1Cl)OC3=C(C=C(C(=C3)Cl)Cl)O2

CAS: 1746-01-6

Uporaba: herbicid, defoliant

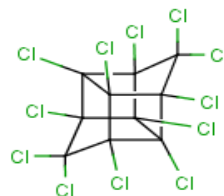


- MIREKS

SMILES: ClC5C1(Cl)C4(Cl)C2(Cl)C1(Cl)C(Cl)(Cl)C5(Cl)C2(Cl)C3(Cl)C4(Cl)Cl

CAS: 2385-85-5

Uporaba: pesticid, insekticid

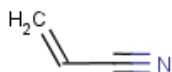


- AKRILONITRIL

SMILES: C=CC#N

CAS: 107-13-1

Uporaba: monomer v proizvodnji sintetičnih polimerov

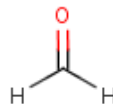


- FORMALDEHID

SMILES: C=O

CAS: 50-00-0

Uporaba: reagent pri sintezi številnih kompleksnih spojin in materialov; v tekstilni industriji; konzervans v cepivih; biocid

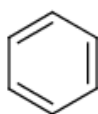


- BENZEN

SMILES: C1=CC=CC=C1

CAS: 71-43-2

Uporaba: intermedijat, topilo

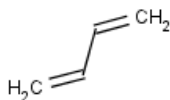


- 1,3-BUTADIEN

SMILES: C=CC=C

CAS: 106-99-0

Uporaba: monomer pri sintezi umetnega kavčuka (gume)



8 PRILOGA II

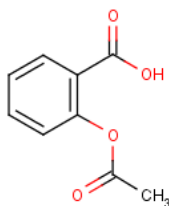
SMILES kode, CAS številke, nameni uporabe in strukturne formule spojin **drugega** testnega niza.

- ACETILSALICILNA KISLINA

SMILES: OC(=O)C1=CC(=CC=C1)OC(=O)C

CAS: 50-78-2

Uporaba: nesteroidni antirevmatik (NSAID)

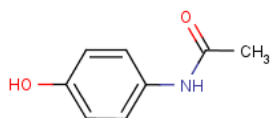


- ACETAMINOFEN

SMILES: C1(=CC=C(C=C1)O)NC(C)=O

CAS: 103-90-2

Uporaba: NSAID

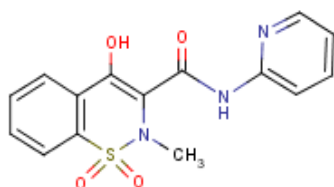


- PIROKSIKAM

SMILES: OC1=C(C(=O)NC2=NC=CC=C2)N(S(=O)(=O)C3=C1C=CC=C3)C

CAS: 36322-90-4

Uporaba: NSAID

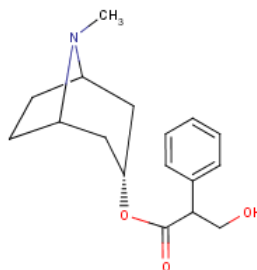


- ATROPIN

SMILES: CN3[C@@H]1CC[C@H]3C[C@H](C1)OC(=O)C(CO)c2ccccc2

CAS: 51-55-8

Uporaba: midriatik

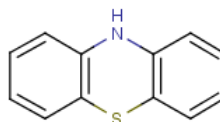


- FENOTIAZIN

SMILES: N1C2=C(C=CC=C2)SC3=CC=CC=C13

CAS: 92-84-2

Uporaba: antipsihotik

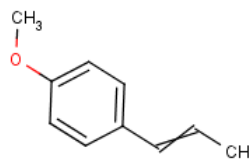


- ANETOL

SMILES: CC=CC1=CC=C(C=C1)OC

CAS: 104-46-1

Uporaba: v alkoholnih pijačah (aroma, sladilo)

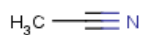


- ACETONITRIL

SMILES: CC#N

CAS: 75-05-8

Uporaba: topilo v sintezni kemiji, topilo pri HPLC

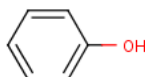


- FENOL

SMILES: OC1=CC=CC=C1

CAS: 108-95-2

Uporaba: topilo

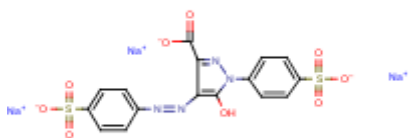


- FD&C Yellow no. 5 (E102, tartrazin)

SMILES: N1(C2=CC=C(C=C2)S(=O)(=O)[O-])(C(C(=N=N/C3=CC=C(C=C3)S(=O)(=O)[O-])C(=N1)C(=O)[O-])=O).[Na+].[Na+].[Na+]

CAS: 1934-21-0

Uporaba: barvilo

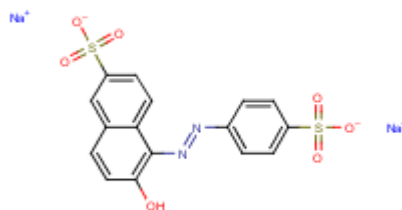


- FD&C Yellow no. 6 (E110, Sunset Yellow)

SMILES: O=S(C1=CC=C2C(C=CC(O)=C2\N=N/C3=CC=C(S(=O)([O-])=O)C=C3)=C1)([O-])=O.[Na+].[Na+]

CAS: 2783-94-0

Uporaba: barvilo

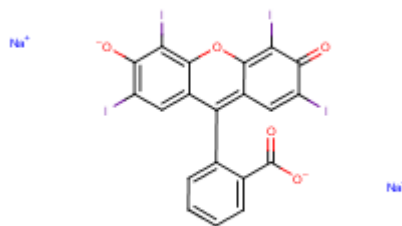


- FD&C Red no. 3 (E127, eritrozina)

SMILES: [O-]C1=C(I)C=C(C(C2=C(C([O-])=O)C=CC=C2)=C3C=C(C(C(I)=C3O4)=O)I)C4=C1I.[Na+].[Na+]

CAS: 16432-68-0

Uporaba: barvilo

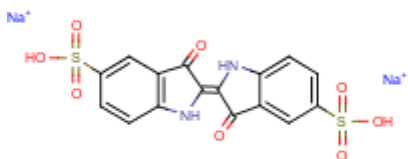


- FD&C Blue no. 2 (E132, indigo karmin)

SMILES: C1(=C2/C(C3=CC(S(=O)(=O)[O-])=CC=C3N2)=O)/C(C4=CC(S(=O)(=O)[O-])=CC=C4N1)=O.[Na+].[Na+]

CAS: 860-22-0

Uporaba: barvilo

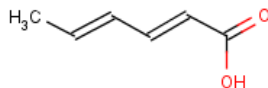


- SORBINSKA KISLINA (E200)

SMILES: OC(=O)C=CC=CC

CAS: 110-44-1

Uporaba: konzervans

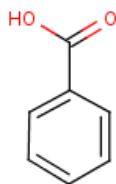


- BENZOJSKA KISLINA (E210)

SMILES: OC(=O)C1=CC=CC=C1

CAS: 65-85-0

Uporaba: konzervans

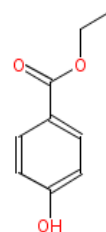


- ETILPARABEN (E214)

SMILES: CCOC(=O)C1=CC=C(C=C1)O

CAS: 120-47-8

Uporaba: konzervans

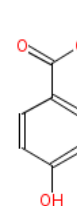


- METILPARABEN (E218)

SMILES: COC(=O)C1=CC=C(C=C1)O

CAS: 99-76-3

Uporaba: konzervans

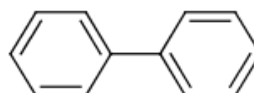


- BIFENIL (E230)

SMILES: C1=C(C=CC=C1)C2=CC=CC=C2

CAS: 92-52-4

Uporaba: konzervans



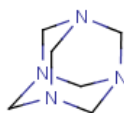
- HEKSAMETILENTETRAMIN

ali HEKSAMIN (E239)

SMILES: C1N2CN3CN(C2)CN1C3

CAS: 100-97-0

Uporaba: konzervans

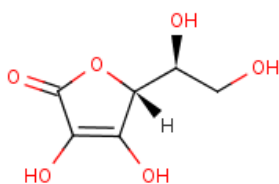


- L-ASKORBINSKA KISLINA
(E300)

SMILES: OC([C@H]([C@H](O)CO)O1)=C(O)C
I=O

CAS: 50-81-7

Uporaba: antioksidant

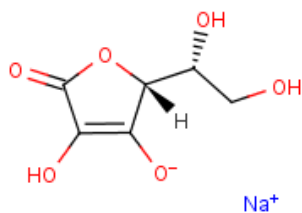


- NATRIJEV ERITORBAT (E316)

SMILES: OC1(=C(O)C(=O)O[C@H]1[C@@H](
C[O-])O).[Na+]

CAS: 6381-77-7

Uporaba: antioksidant



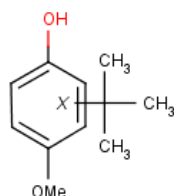
- BUTILIRAN

HIDROKSIANIZOL ali BHA
(E320)

SMILES: OC1=CC=C(C=C1C(C)(C)C)OC

CAS: 25013-16-5

Uporaba: antioksidant



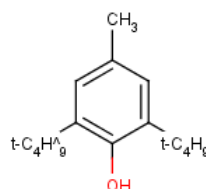
- BUTILIRAN

HIDROKSITOLUEN ali BHT
(E321)

SMILES: OC1=C(C=C(C=C1C(C)(C)C)C(C)(
C)C

CAS: 128-37-0

Uporaba: antioksidant

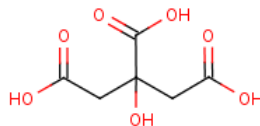


- CITRONSKA KISLINA (E330)

SMILES: OC(CC(=O)O)(CC(=O)O)C(=O)O

CAS: 77-92-9

Uporaba: regulator kislosti

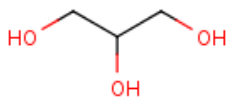


- GLICEROL (E422)

SMILES: OCC(O)CO

CAS: 56-81-5

Uporaba: mehčalo

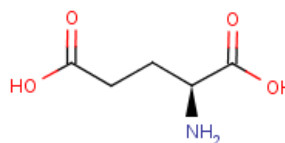


- GLUTAMINSKA KISLINA (E620)

SMILES: O=C(O)CC[C@H](N)C(O)=O

CAS: 56-86-0

Uporaba: ojačevalec okusa



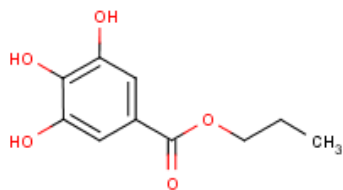
- PROPIL GALAT (E310)

SMILES:

O=C(C1=CC(=C(C(=C1)O)O)O)OCCCC

CAS: 121-79-9

Uporaba: antioksidant

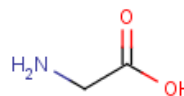


- GLICIN (E640)

SMILES: NCC(O)=O

CAS: 56-40-6

Uporaba: ojačevalec okusa

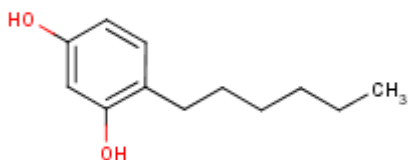


- 4-HEKSILRESORCINOL (E586)

SMILES: OC1=CC(C=CC(=C1)O)CCCCC

CAS: 136-77-6

Uporaba: antioksidant

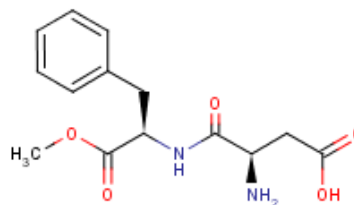


- ASPARTAM (E951)

SMILES: O=C([C@H](CC1=CC=CC=C1)NC(=O)[C@H](CC(=O)O)N)OC

CAS: 22839-47-0

Uporaba: sladilo

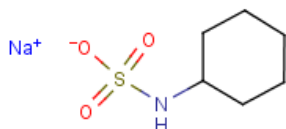


- NATRIJEV CIKLAMAT (E952)

SMILES: C1(NS(=O)(=O)[O-])CCCCC1.[Na+]

CAS: 139-05-9

Uporaba: sladilo

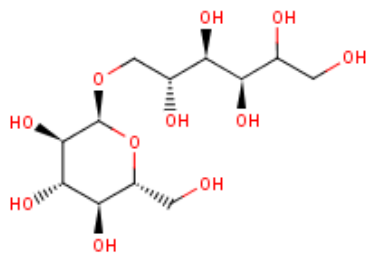


- IZOMALT (E953)

SMILES: OC[C@H](O)[C@@H](O)[C@H](O)[C@H](O)CO[C@H]1O[C@H](CO)[C@@H](O)[C@H](O)[C@H]1O

CAS: 64519-82-0

Uporaba: sladilo

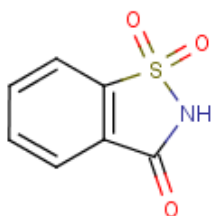


- SAHARIN (E954)

SMILES: C12=C(C(=O)NS1(=O)=O)C=CC=C2

CAS: 81-07-2

Uporaba: sladilo

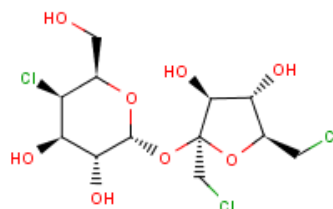


- SUKRALOZA (E955)

SMILES: O[C@@H]1[C@@](O[C@@H](O[C@H](CO)[C@@H](O)[C@H]2O)(CCl)O[C@H](CCl)[C@H]1O

CAS: 56038-13-2

Uporaba: sladilo



- PROPILEN GLIKOL (E1520)

SMILES: CC(CO)O

CAS: 57-55-6

Uporaba: vlažilo

