연구보고서

급성독성 화학물질의 유해성 예측프로그램 적용연구

김현옥 · 임경택 · 이인섭

산업재해예방
안전보건공단
산업안전보건연구원

요 약 문

연구기간	2019년 1월 ~ 2019년 11월
핵심단어	급성독성, 화학물질, 유해성, 예측, 프로그램
연구과제명	급성독성 화학물질의 유해성 예측프로그램 적용연구

1. 연구배경

본 연구는 유해성 예측프로그램(전산예측 및 분자 모델링)을 도입·적용하여 급성독성, 변이원성 및 발암성 등의 건강유해성 정보가 부족한 반도체 및 LCD 제조공정 등에서 사용하는 신규화학물질에 대한 건강유해성 정보를 생산·제공하고, 예측프로그램을 활용하여 화학물질을 효율적으로 관리하는 방안을 마련하고자 하였다.

2. 주요 연구내용

1) 1992년 이후 사업주가 고용노동부에 제출하여 우리 연구원에서 검토하고 있는 "신규화학물질 유해성·위험성 조사 및 확인" 사업결과 중 반도체 및 LCD 등 전자산업 관련 신규화학물질 1,754종 중 중복물질(139종)과 정보보호 요청 및 고분자화합물 등으로 CAS No.가 없는 물질(789종)을 제외한 826종을 1차 대상물질로 선정하고, SMILES string을 검색하여 Derek Nexus, Danish QSAR, Vega QSAR, T.E.S.T (EPA) 등 4개의 예측프로그램으로 사전 예측한 결과 모든 프로그램에서 예측 가능한 132종을 최종

연구대상으로 선정하였다.

2) 문헌조사 및 기존 연구결과를 바탕으로 국제적으로 검증된 신뢰성 있는 독성예측프로그램을 선정하여 각 QSARs 모델의 구조, 화학물질 구조분류 및 특성, 예측 범위(domain) 등 모델의 특징을 분석하였다.
3) 본 연구에서는 각 예측프로그램간 결과 비교를 위해 유료로 구입한 Derek Nexus와 무료 제공프로그램인 Danish QSAR, Vega QSAR, T.E.S.T (EPA), QSAR Toolbox를 사용하였고, 화학물질의 독성정도를 예측하기 위하여 Toxtree를 사용하였다.
4) 변이원성의 경우 5개의 예측프로그램을 통해 예측하였고, 모든 예측프로그램에서 양성인 물질은 9종이었으며, 3개의 예측프로그램을 통해 예측된 발암성의 경우 모든 프로그램에서 양성인 물질이 11종이었다. 급성 경구독성의 경우에도 3개의 예측프로그램을 활용하였으며, 모든 프로그램에서 GHS 분류기준에 따른 급성 경구독성($0 < LD_{50} \leq 2{,}000$ mg/kg)으로 예측된 물질은 33종이었다.
5) 변이원성 예측결과는 Derek Nexus 및 QSAR Toolbox 등을 활용한 기존 연구에서도 예측 정확도가 85% 이상으로 나타났고, 5개의 예측프로그램을 활용한 결과이므로 모든 프로그램에서 양성으로 예측된 물질(9종)의 경우 MSDS 등을 통하여 사업장 및 근로자에게 제공할 수 있을 것으로 사료된다.
6) 발암성 및 급성 경구독성 예측결과는 유전독성 및 발암성 시험 후보물질 선정을 위한 사전 스크리닝용으로 활용하는 것은 가능하나, 사업장 및 근로자에게 제공하기 위해서는 예측프로그램별 기존 시험자료와 예측결과를 비교하여 예측 정확도를 평가하는 등 신뢰도 확인을 위한 추가적인 연구가 필요할 것으로 사료된다.

3. 연구 활용방안

- 제 언

신규화학물질 관련 동물실험을 대체하고, 유전독성 및 발암성 시험 후보물질 선정을 위한 사전 스크리닝용으로 활용할 수 있도록 발암성 및 혼합물 관련 예측프로그램에 대한 추가적인 연구가 필요하다.

- 개선방안 또는 정책방안

유전독성(복귀돌연변이 및 소핵) 시험성적서 제출을 생략하고 있는 1톤 미만 신규화학물질의 경우 근로자 건강장해 예방을 위하여 유해성 예측프로그램을 활용하는 방안 마련이 필요하다.

- 활 용

양성으로 예측된 변이원성 예측결과는 안전보건공단 MSDS 또는 기술자료 등을 통하여 사업장 및 근로자에게 제공하고, 본 연구결과는 한국산업보건학회 등 학회발표, 국내외 우수 학술지 논문게재 및 관련 후속연구의 기초연구로 활용하고자 한다.

4. 연락처

- 산업안전보건연구원 유해성연구부 김현옥
- ☎ 042) 869. 0321
- E-mail : okhokim@kosha.or.kr

차 례

요약문 i

Ⅰ. 서 론 1

1. 연구 배경 및 필요성 1
2. 연구 목적 및 범위 4

Ⅱ. 연구 방법 5

1. 유해성 예측프로그램 및 대상물질 모집단 구축 5
2. 유해성 예측프로그램 도입 및 적용 7

Ⅲ. 연구 결과 9

1. 적용 QSARs 모델에 대한 사전 고찰 9
2. 대상물질 선정 결과 17
3. 대상물질에 대한 QSARs 모델 적용 결과 19
4. 각 예측프로그램을 이용한 예측결과 비교·검토 29

Ⅳ. 고 찰 ····· 33

Ⅴ. 결 론 ····· 37

참고문헌 ····· 39

Abstract ····· 47

부록 ····· 49

부록 1. 변이원성 예측결과 ····· 50
부록 2. 발암성 예측결과 ····· 59
부록 3. 급성 경구독성 예측결과 ····· 64
부록 4. 대상물질의 용도, 작용기 및 Toxtree 예측결과 ····· 69

표 차례

<표 1> 연구대상 물질 선정을 위한 사전 예측 결과 ····· 6
<표 2> 1992~2018년 전자산업 관련 신규화학물질 DB 구축 현황 ····· 17
<표 3> 최종대상물질의 작용기(functional group) 분포 현황 ····· 18
<표 4> Derek Nexus 예측결과 ····· 20
<표 5> Danish QSAR 급성 경구독성 예측결과 ····· 22
<표 6> Vega QSAR 예측결과 ····· 23
<표 7> T.E.S.T (EPA) 급성 경구독성 예측결과 ····· 24
<표 8> T.E.S.T (EPA) 및 QSAR Toolbox의 변이원성 예측결과 ····· 27
<표 9> QSAR Toolbox 급성 경구독성 예측결과 ····· 27
<표 10> GHS 분류기준에 따른 급성 경구독성 예측물질 ····· 32

그림 차례

[그림 1] Derek Nexus 프로그램 화면 ····· 10
[그림 2] Danish QSAR 프로그램 화면 ····· 11
[그림 3] Vega QSAR 프로그램 화면 ····· 13
[그림 4] T.E.S.T (EPA) 프로그램 화면 ····· 14
[그림 5] QSAR Toolbox 프로그램 화면 ····· 15
[그림 6] Toxtree 프로그램 화면 ····· 16

약어의 정리(Abbreviations) 및 용어의 정의(Definitions)

ECHA	유럽화학물질청(European Chemicals Agency)
EPA	미국 환경부(United States Environmental Protection Agency)
GHS	화학물질의 분류 및 표지에 관한 세계조화시스템(Globally Harmonized System of classification and labelling of chemicals)
MSDS	물질안전보건자료(Material Safety Data Sheets)
NIH(NLM)	미국 국립의학도서관(U.S. National Library of Medicine)
OECD	경제협력기구(Organization for Economic Co-operation and Development)
QSARs	정략적 구조-활성관계(Quantitative structure-activity relationships)
QSAR Toolbox	OECD Toolbox라고도 하며, OECD 및 ECHA에서 개발한 독성예측프로그램
QSNR	양적구조 나노독성관계(Quantitative structure nanotoxicity relationships)
REACH	신화학물질관리제도(Registration, Evaluation, Authorization and Restriction of CHemicals)
SMILES	단순화된 분자 입력라인 기입사항(Simplified Molecular Input Line Entry Specification)의 약어로 화학기호를 일렬로 늘어세우는 방식으로 분자구조를 나타내는 화학물질 표기 시스템
T.E.S.T.	EPA에서 개발한 독성예측프로그램(Toxicity Estimation Software Tool)

Ⅰ. 서 론

1. 연구 배경 및 필요성

기존 생명공학의 연구방법 중 살아있는 생명체(척추동물 이상의 고등동물)를 대상으로 한 실험의 경우를 "*In Vivo*"라 하며, 세포, 미생물 등 시험관에서 생체와 유사한 환경을 인위적으로 조성하여 실험하는 방법을 "*In Vitro*"라 한다. 최근 컴퓨터 기술의 발달로 컴퓨터 시뮬레이션을 이용하는 방법이 있는데, 이러한 가상환경에서의 실험방법을 인실리코(*In Silico*)라 한다.

인실리코*(In silico)* 독성예측의 목표는 전산 모델링(computational modeling)인 QSARs로 예측하고, 알고리즘을 독성 데이터와 정보로 예측하기 위한 것이다. 이는 동물실험의 대안 중 하나로 여겨지며, 광범위한 용어로 화학물질의 독성 혹은 이의 효능과 그 화학물질의 구조를 연관 짓는 다양한 전산기법(computational techniques)을 활용한다. 기존의 실험과 비교했을 때 인실리코(*in silico*) 기법에는 비용 효율성, 속도 및 동물실험의 감소 등 확실히 많은 장점이 있다. 전산기법을 사용하면 화학물질, 혼합물 및 나노물질과 같은 여러 물질들의 독성을 양적 및 질적으로 예측하게 된다. 이런 방법들은 각각의 강점, 한계, 적용 범위, 해석의 특수성이 있으며, 각각의 목표는 당면한 문제에 대해 가장 효과적인 방법을 찾는 것이 필요하다.

In silico 모델링 방법의 사례로 우선 나노독성 분야의 "양적구조 나노독성관계(QSNR; 나노 QSAR이라고도 함)"가 있으며, 이 모델은 나노물질의 크기, 모양, 표면적, 용해도 등과 같은 나노물질의 특이적 기술자(descriptor[1])를 사용하

1) a word or phrase used to describe or refer to something

며, 나노물질의 구조 및 물리화학적 특성과 관련된다. 크기와 표면전하[2]를 사용하여 은나노 입자에 대한 EC_{50}(50% 효소활성 억제를 위한 유효농도)을 예측하기 위해 선형 QSNR 모델의 예가 제공되었으며, 양자화학적 기술자(descriptor)를 사용하는 QSNR 모델은 풀러렌 C_{60} 유도체에 대한 결합친화성을 예측하고 금속산화물 나노입자의 세포독성을 예측하기 위해 개발되었다.

화학물질의 독성은 다른 화학물질과의 상호 작용에 의해 영향을 받는다. 예를 들어, 혼합물은 각 화학물질의 NOAEL[3]에 큰 영향을 주지 않을 것이며, 독성평가는 혼합물의 부작용을 과소평가하거나 간과할 수 있다. 예를 들어, 납(Pb)의 독성은 다른 금속의 동시투여와 함께 증가한다는 것이 밝혀졌으며, 혼합물의 독성을 연구하기 위한 '누적 위험 평가(Cumulative Risk Assessment)'가 개발되었다. 그러나 화학물질의 노출유형 및 복잡한 상호작용으로 인해 혼합물의 독성에 대한 실험 데이터가 부족하며, 이러한 요소의 모든 조합을 시험하는 것은 불가능하다. 단일 화학물질에 대한 방법은 혼합 효과를 결정하기가 어렵기 때문에 혼합물에 적용할 수 없다.

In silico 기계학습(machine learning) 방법은 질적 분류 및 양적 회귀분석에 널리 적용되어 왔으며 경고 구조(structural alert)[4]는 최적화를 위한 보완도구로 간주될 수 있다. 여기서는 다양한 독성예측을 위해 만들어진 예측모델들과 더불어 사용 가능한 데이터베이스와 웹 서버도 제시하였다. 이 방법과 모델들은 향후 화학물질 안전성 평가를 위해 앞으로 많은 발전이 기대되는 분야이다. 일반적으로 시험관 내(in vitro) 및 생체 내(in vivo) 검사는 다양한 독성 및 약물 부작용을 포함하여 화학물질 안전성을 조사하기 위해 수행되며, 최근에는

2) 전하(電荷, electric charge)는 전기현상을 일으키는 주체적인 원인으로, 어떤 물질이 갖고 있는 전기의 양

3) 최대무독성용량(No Observed Adverse Effect Level, NOAEL) : 독성시험 시 대조군에 비해 바람직하지 않은 영향을 나타내지 않는 통계학적으로 유의한 차이를 보이지 않은 최대 투여용량(mg/kg bw/day)

4) 경고 구조는 독성학 및 규제 결정에서 잠재적인 화학적 유해성 또는 화합물을 적합한 범주로 분류할 수 있는 간단하고 투명한 수단으로 광범위하게 수용됨

비용을 줄이기 위해(Huh et al., 2011) “체내 장기“와 같은 시험관 내 모델을 개발하려는 노력도 있지만 여전히 많은 비용과 시간이 소요된다. 실험방법과 비교하면 전산기법은 자연친화적이고, 빠르고, 싸고, 정확하며, 무엇보다도 화합물이 합성되기 전에 수행될 수 있기 때문에 큰 이점을 보이고 있다(Segall and Barber, 2014). 최근 몇 년 동안 화학물질 안전성 평가 분야에서 기계학습 방법 및 경고 구조와 함께 화학물질 독성예측 모델들이 개발되었으며, 공개적으로 무료로 사용할 수 있도록 admetSAR이라는 웹 서버도 개발되었다(Cheng et al., 2012b).

독성예측 분야의 진전으로, 발암성 물질은 인간의 건강에 심각한 영향을 미치므로 화학물질 개발에서 점점 더 중요해지고 있다. 대부분의 예측모델은 data sources로 발암 가능성 데이터베이스(Carcinogenic Potency Database; CPDB)를 사용하며 TD_{50}[5] 값에 따라 발암물질 또는 비발암물질이 1,500개가 넘는다. 화학물질 발암성을 예측하는 모델을 만들기 위해 여러 논문들에서 프로토콜을 공유했으며, 최근에는 발암성을 온라인으로 예측할 수 있는 웹 서버인 CarcinoPred-EL을 개발했다(Zhang et al., 2017). 화학물질의 발암성 기전은 (1) DNA를 손상시키는 화학물질의 돌연변이 유발성에 의해 주로 야기되는 유전독성(Fan et al., in press)과 (2) 다른 특정 메커니즘을 통해 행동하는 비유전독성 발암물질(Golbamaki and Benfenati, 2016)로 분류할 수 있다. 현재 복귀돌연변이원성(Ames mutagens)이 있는 8,000종이 넘는 화합물이 알려져 있으며, 최근 몇 년 동안 이러한 독성 데이터들을 활용한 예측 모델과 경고 구조가 확대되었다(Yang et al., 2017a).

화학물질의 노출 경로에 따라 급성 독성은 경구, 경피 및 흡입으로 구분할 수 있으며 그 중 급성 경구독성이 계산 예측에서 가장 널리 연구되고 있다. Zhu 등은 LD_{50} 값을 갖는 7,385종의 화합물을 수집하고 급성 경구독성 예측을

5) 주어진 표적 부위에 대해, 대조군 동물에 종양이 없다면, TD_{50}은 종양을 유도하는 만성 용량(mg/kg 체중/일)임. TD_{50}은 LD_{50}과 유사하며 TD_{50} 값이 낮으면 강력한 발암물질을 나타내는 반면 높은 값은 약한 물질을 나타냄

위한 몇 가지 모델을 구축했다. 데이터 세트를 바탕으로 LD_{50} 또는 그 독성 카테고리를 예측하기 위해 분류 및 회귀 모델을 구성하는 여러 가지 기계학습 방법을 개발하고 적용했으며(Xu et al., 2017), 이 모델은 웹 서버 (http://www.pkumdl.cn/DLAOT/DLAOThome.php)에서 무료로 사용할 수 있다.

2. 연구 목적 및 범위

유해성 예측프로그램(유해성 전산예측 및 분자 모델링)을 도입·적용하여 반도체 및 LCD 제조공정 등에서 사용하는 신규화학물질에 대한 건강유해성 정보를 생산·제공하고, 향후 예측프로그램을 활용하여 건강유해성이 미확인된 화학물질을 효율적으로 관리하는 방안에 대하여 고찰하고자 하였다.

또한, 본 연구는 유해성이 명확히 밝혀지지 않은 신규화학물질로 인한 급성독성 및 사망사고의 효율적 예방 필요(관련 : 중기사업계획-화학물질연구)에 의해 기획되었으며, 연구원 산업화학연구실 사업혁신 방안인 “빅데이터를 활용한 흡입독성 시험물질 선정의 타당성 강화(사회적 이슈, 국민적 관심사를 고려한 흡입독성시험 추진)”의 선행연구로 수행하고자 하였으며, 안전보건공단 산재예방 신사업 및 업무개선 공모 입상내용(‘18.12.03)의 실현을 위한 결과물로 제시하고자 하였다.

Ⅱ. 연구 방법

1. 유해성 예측프로그램 및 대상물질 모집단 구축

1) 대상물질 모집단 선정

(1) 1992년 이후 사업주가 고용노동부에 제출하여 우리 연구원에서 검토하고 있는 "신규화학물질 유해성·위험성 조사 및 확인" 사업결과 중 반도체 및 LCD 등 전자산업 관련 화학물질 1,754종 중 중복물질(139종)과 정보보호 요청 및 고분자화합물 등으로 CAS No.가 없는 물질(789종)을 제외하고 CAS No.가 있는 826종을 1차적으로 선별하였다.

(2) 1차 선별물질을 대상으로 예측프로그램에 공통 적용되는 입력 쿼리인 SMILES string 또는 MOL file을 검색하기 위해 NIH의 ChemIDPlus 등을 검색한 결과, SMILES string은 395종, MOL file은 126종이 검색되었다. MOL file이 있는 경우 SMILES string이 모두 있었기 때문에 395종을 2차로 선별하였다.

(3) Derek Nexus, Danish QSAR, Vega QSAR, T.E.S.T (EPA) 등 4개의 예측프로그램으로 2차 선별물질 395종을 사전 예측한 결과 <표 1>과 같이 예측되었다. 4개의 프로그램에서 모두 예측 가능한 물질 132종을 최종 연구대상으로 선정하였다.

<표 1> 연구대상 물질 선정을 위한 사전 예측 결과

프로그램명	사전 예측 대상	예측 가능 물질	예측율	비고
Derek Nexus	395종	358종	90.6%	Certain level에서 예측불가(13종)
Danish QSAR	395종	187종	47.3%	
Vega QSAR	395종	261종	66.1%	
T.E.S.T (EPA)	395종	154종	39.0%	

2) 독성 예측 QSARs 프로그램 및 모델 선정

(1) 문헌 조사 및 기존 연구결과를 바탕으로 국제적으로 검증된 신뢰성 있는 독성예측프로그램을 선정하여 각 QSARs 모델의 구조, 화학물질 구조 분류 및 특성, 예측 범위(domain) 등 모델의 특징을 분석하였다.

(2) 각 예측프로그램간 결과 비교를 위해, 본 연구에서는 유료로 구입한 Derek Nexus와 무료 제공프로그램인 Danish QSAR, Vega QSAR, T.E.S.T (EPA), QSAR Toolbox를 사용하였다.

(3) 또한, 화학물질의 독성정도를 예측하기 위하여 Toxtree를 사용하였다.

2. 유해성 예측프로그램 도입 및 적용

1) 예측결과 도출 및 비교분석

Derek Nexus 등 선정된 QSARs 모델을 활용하여 연구대상 물질에 대한 예측결과를 확보하고, 독성별 각 예측모델 간의 예측 차이를 비교·분석하였다. 추가로, 사업장 제출자료 및 안전보건공단 MSDS 상의 시험자료와 일치 또는 불일치되는 정도를 비교·분석하였다.

2) 화학물질관리를 위한 예측프로그램 활용방안 모색

화학물질 유해성 예측프로그램을 도입·적용하여 반도체 및 LCD 제조공정 등에서 사용하는 신규화학물질에 대한 유해성 정보를 사업장에 제공하는 방안과 우리 연구원에서 수행하고 있는 유전독성 및 발암성 시험 후보물질 선정에 활용할 수 있는 방안 및 산업안전보건법에 따른 신규화학물질 유해성·위험성 조사보고서 제출 제도에 활용하는 방안 등을 모색하였다.

Ⅲ. 연구 결과

1. 적용 QSARs 모델에 대한 사전 고찰

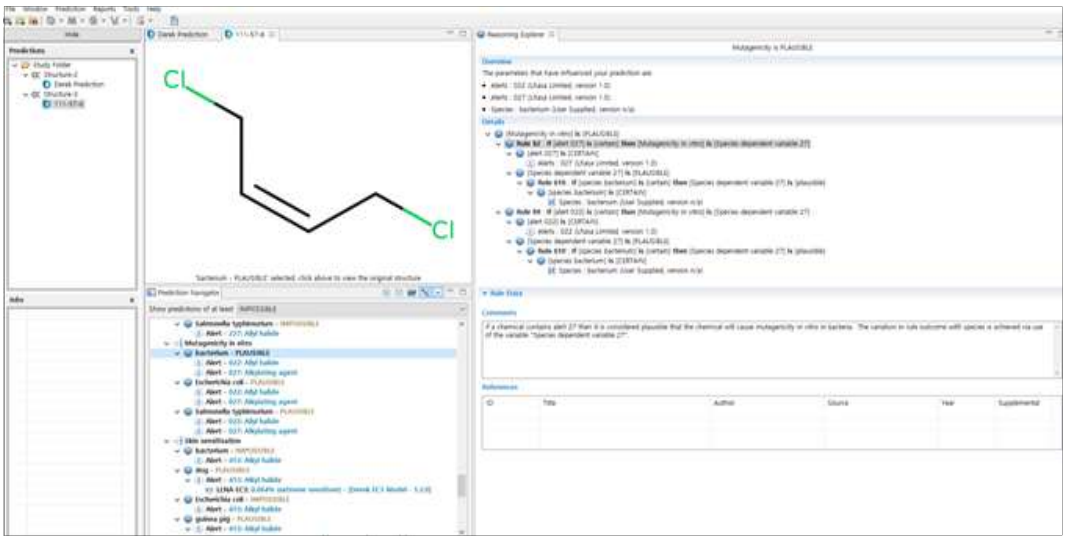
QSARs란 화학물질의 구조 및 물리화학적 성질과 생체 내 독성 영향 등과의 상관관계 또는 관련성을 바탕으로 시험 결과를 추정하는 것으로, 다양한 알고리즘을 통해 예측결과 값을 도출한다. 일반적으로 예측방법에는 전문적 경험규칙 기반(Expert rule-based)과 통계 기반(Statistical-based) 방법으로 나누어진다. 전문적 경험규칙을 기반으로 하는 모델의 경우, 독성학적 정보, 전문가의 의견, 퍼지이론을 바탕으로 하는 'If-then-else'규칙으로 예측을 수행한다. 반면 통계 기반 모델은 통계 기법, 귀납 규칙, 인공지능, 패턴 기술을 이용한 방정식 기법으로 예측을 수행한다(Serafimova et al., 2010; 식품의약품안전처, 2016; 국립환경과학원, 2016).

본 연구에서는 화학물질의 유해성 예측을 위하여 대표적인 경험규칙 기반 모델인 Derek Nexus, Danish QSAR, Vega QSAR, T.E.S.T (EPA), QSAR Toolbox를 사용하였고, 화학물질의 독성정도를 예측하기 위하여 Toxtree를 추가로 사용하였다.

1) Derek Nexus

Lhasa Ltd에서 개발한 Derek은 전문적 경험 규칙에 기반을 둔 QSAR 모델의 한 종류이다. Derek은 화학물질 구조를 일련의 분류군으로 나누고 이를 바탕으로 구조와 생물학적 활성의 상관관계를 예측하는 Knowledge base의 독성

예측 프로그램으로 유전독성을 비롯한 다양한 독성 지표에 대해 예측이 가능하다. 대상 화학물질 내 독성을 유발하는 Toxicophore를 확인하여 독성 예측하는데, Alert는 분자 특성을 결정짓는 구조를 확인하는 것으로 이를 통해 독성을 추론하고, Reasoning을 통한 Rule set에 입각하여 Alert을 보완한다(Serafimova et al., 2010). 이때, 위험성 평가는 Data 검토를 포함한 전문가 해설과 OECD Principle에 입각한 구조 활성 관계 분석을 기반으로 하며 물리·화학적 특성과 전문적 원칙을 적용하여 의미 있는 예측이 가능하다. 또한 Alert 성능에 대한 의견 요약과 검증 설명을 통해 예측에 대한 강한 신뢰성이 확보된다. 특히 동 개발사에서는 Sarah 및 Toxtree 등 다른 예측 프로그램과 복수 적용을 통해 예측력을 향상시키는 연구가 진행되고 있다(Capina et al., 2017; Hemingway et al., 2017).

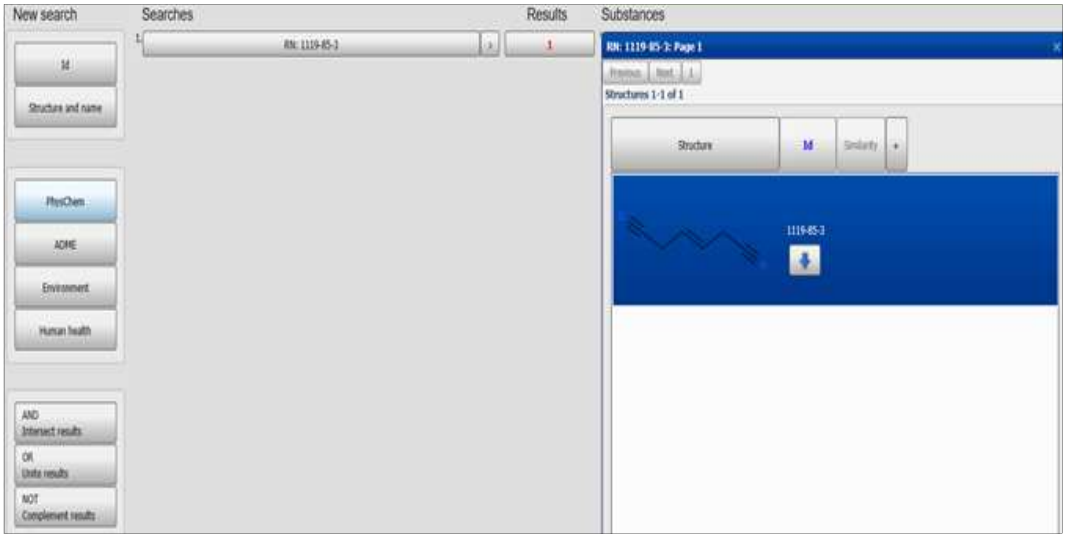


[그림 1] Derek Nexus 프로그램 화면

2) Danish QSAR

Repository 기반의 모델로써 600,000종 이상의 화학물질 데이터를 보유하고 있으며, REACH에 등록된 모든 단일 유기 화학 구조(약 80,000종)가 Structure set에 포함되어 있다. 또한 200여 개의 QSAR 모델을 프로그램 내에 탑재하고 있다. Endpoints는 Leadscope, CASE Ultra, SCi QSAR 총 3가지 소프트웨어 시스템으로 모델링 되었으며, 세 가지 모델의 통합 결과로 배터리 예측이 이뤄진다. 배터리 예측의 장점은 각 모델의 예측값 불일치를 감소시켜 결과의 정확성을 향상시키고 적용 가능한 도메인을 확장하는 역할을 하는 것이다.

입력 쿼리는 화학물질명, Structure, CAS No., SMILES string, Mol file 등의 형태이며, 예측범위는 물리·화학적 특성, 급성독성, 변이원성, 피부부식성·자극성, 환경독성 등이다.



[그림 2] Danish QSAR 프로그램 화면

3) Vega QSAR Ver 1.1.5 BETA22

REACH non-animal alternative methods 요구 조건에 최적화된 예측 결과를 제공하며, CAESAR, SARpy, EPI Suite 등 다양한 QSAR 모델의 결과를 통합하여 Read-across 결과를 도출하는 독립적인 모델이다. 프로그램은 Applicability Domain Index를 통해 모델 예측의 신뢰성을 평가하며, 이를 위해 완전히 독립된 알고리즘을 사용한다. 이 알고리즘은 각각의 QSAR 모델에 개별적으로 작동하며 유사한 화합물을 보여주고 유사한 화합물에 대한 QSAR 결과를 평가하여 예측 대상 화학물질과 관련된 화합물의 화학적 특징을 분석한다.

대부분의 기존 QSAR 모델은 예측을 위해 3가지 구성성분을 기초로 하고 있다. 첫 번째로는 유전독성이나 발암성과 같이 예측결과로 얻고자하는 물질의 특성이며, 두 번째로는 화학물질의 정보, 세 번째는 앞의 두 가지를 이어줄 수 있는 기능이다. Vega QSAR는 이 세 가지의 변수를 통합하여 단일 값으로 나타내고자 Applicable domain index(ADI)를 개발하였다. 타 프로그램이 물질간의 구조 유사성을 비교 예측하여 적용가능 영역을 in과 out의 결과로 나타낼 때, Vega QSAR는 정량적 수치인 ADI를 통해 예측의 신뢰성을 향상시킨다.

Vega QSAR 내 4백만 개 화합물 정보를 활용함으로써 구조 유사성 알고리즘을 최적화하며, Batch process 적용으로 많은 수의 물질에 대한 동시예측이 가능하다. 입력 쿼리는 SMILES string을 사용하고, 예측범위는 변이원성, 발암성, 피부 과민성, BCF, log P 등이 해당한다.

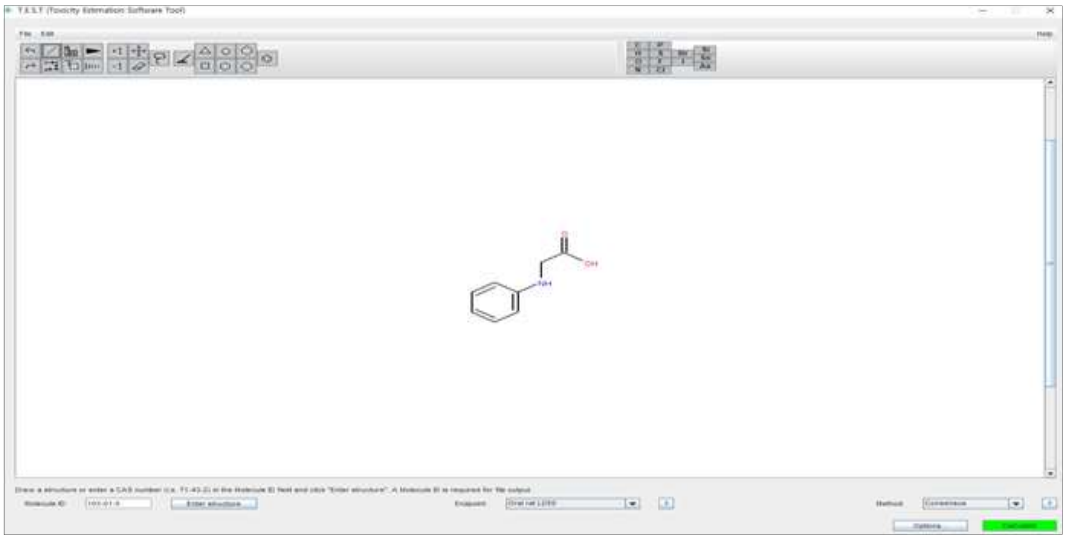
[그림 3] Vega QSAR 프로그램 화면

4) T.E.S.T (EPA) Ver 4.2.1

The Chemistry Development Kit (CDK) 자바 오픈 소스 사이트와 화학물질 데이터가 연동되어 있으며, 각기 다른 8개의 특징적 QSAR 예측 방식[6]을 적용할 수 있다. 다양한 예측 방식의 적용을 통해 결과에 대한 신뢰도를 강화할 수 있다는 장점을 가지며, Statistical external validation을 사용하여 결과의 유효성을 검증한다. 프로그램 예측의 정확도는 세 가지의 통계값으로 결정된다. 첫 번째는 일치성(concordance)으로 정확하게 예측된 화합물의 비율이며, 두 번째는 민감도(Sensitivity)로 활성(active)으로 예측된 화합물이 실험값으로 활성을 나타낼 비율이다. 마지막으로 특이도(Specificity)는 비활성(Inactive)으로 예측된 화합물이 실험값으로 비활성을 나타낼 비율을 의미한다.

입력 쿼리는 Structure, CAS No., SMILES string, Mol file 등이 있으며, 예측 범위는 급성 경구독성, 변이원성, 환경 독성 등이 해당한다.

6) 8개의 QSAR 예측 방식 : Hierarchical, FDA, Single model, Group contribution, Nearest neighbor, Consensus, Random forest, Mode of actioin method



[그림 4] T.E.S.T (EPA) 프로그램 화면

5) QSAR Toolbox Version 4.3.1

OECD 및 ECHA에서 법적규제를 목적으로 개발한 독성예측 프로그램으로써 정부, 산업체 등 이해관계자가 화학물질의 유해성을 평가하는 데 활용되고 있다.

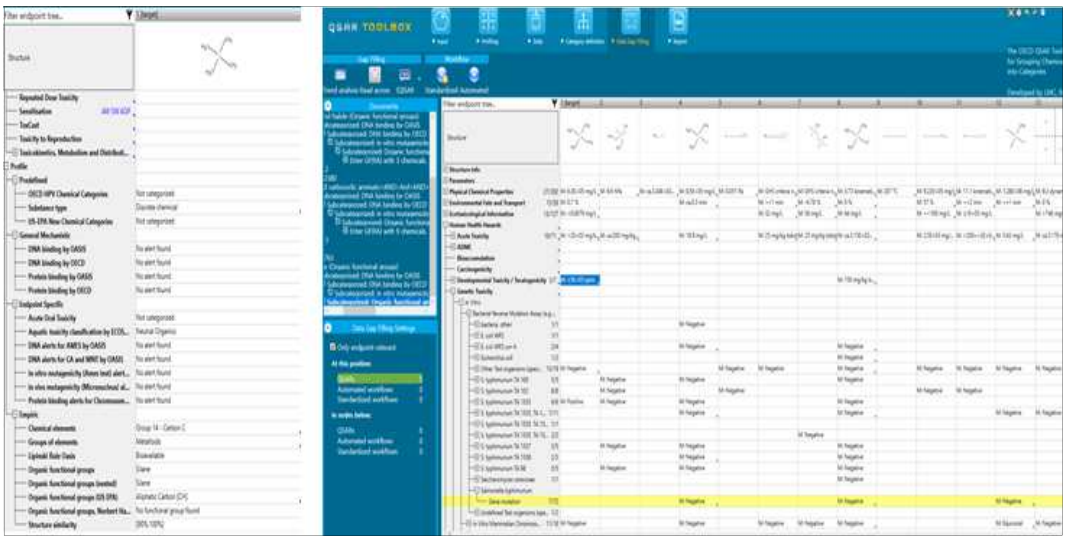
QSAR toolbox는 크게 3가지의 주요 특징을 갖고 있다. 첫째, 예측 대상물질과 관련된 구조적 특성 및 잠재적 메커니즘 또는 작용방식을 식별한다. 둘째, 구조적 특성 및 잠재적 메커니즘이 동일한 다른 화학물질을 검색한다. 셋째, 기존 실험 데이터를 사용하여 데이터 갭을 채운다.

예측에 사용되는 Training sets(categories)은 유동적으로 결정된다는 점에서 여타 예측 프로그램과는 차별성을 띠고 있다. 또한 개별 예측의 결과는 일관성을 유지하며, 그룹핑과 데이터 갭 채우기 단계에서 Computational method가 적용되어 데이터의 신뢰성이 보장된다.

가장 최근 개발된 Version 4.3.1은 2개의 DB (ADME, pKa OASIS), 159개의 QSAR 모델, 5개의 프로파일러가 추가되었으며, 업데이트된 16개의 프로파일러

와 7개 반응 대사 시뮬레이터 등의 향상된 기능을 활용할 수 있다.

입력 쿼리는 화학물질명, Structure, CAS No., SMILES string 등을 사용하고, 예측범위는 급성독성, 변이원성, 발암성, 피부 부식성·자극성, 환경독성, 물리·화학적 특성 등이 해당된다.

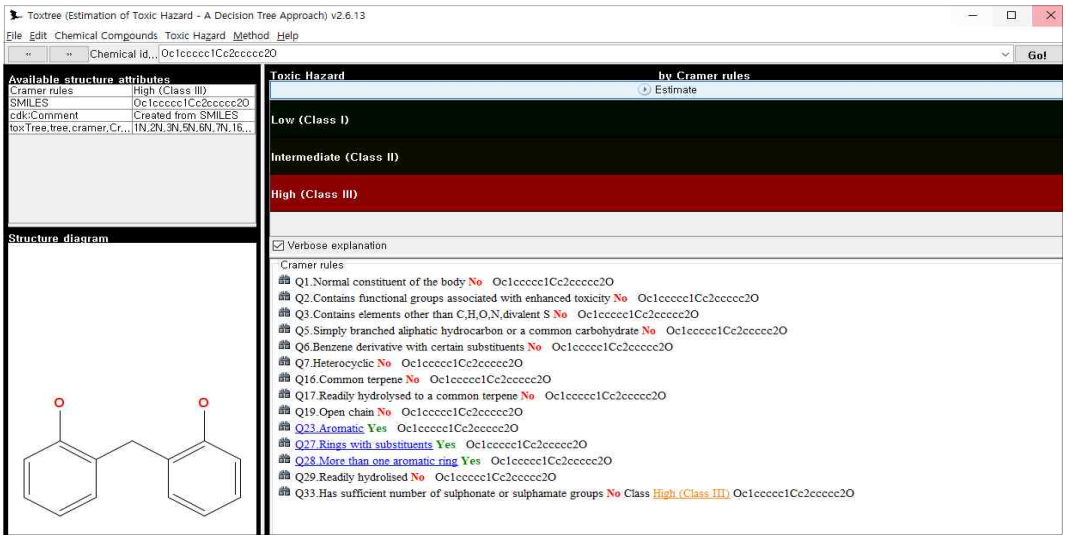


[그림 5] QSAR Toolbox 프로그램 화면

6) Toxtree 2.6.13

Toxtree는 독립적 QSAR 모델로써 화학물질의 독성을 Decision tree 형식으로 진행하는 오픈소스 프로그램이다. 유전독성과 발암성을 예측하기 위한 Decision tree는 European Comission report, “The Benigni/Bossa Rulebase for mutagenicity and Carcinogenicity - a module of Toxtree”를 기반으로 적용된다. 프로그램 내 활용 가능한 Plug-in으로 Cramer-rule이 존재한다. 분석대상 화학물질의 독성정도를 표현하기 위하여 3개의 Class로 나누게 되며, 예측을 위해서는 33개의 구조규칙이 적용된다.

입력 쿼리는 화학물질명, CAS No., SMILES string 등이 가능하고, 화합물의 시스템적 독성 예측이 가능하다.



[그림 6] Toxtree 프로그램 화면

2. 대상물질 선정 결과

1) 1차 및 2차 대상물질 선별

1992년 이후 사업주가 고용노동부에 제출하여 우리 연구원에서 검토하고 있는 "신규화학물질 유해성·위험성 조사 및 확인" 사업결과 중 반도체 및 LCD 등 전자산업 관련 화학물질 1,754종을 DB로 구축하였다. 중복물질 139종을 제외한 DB 구축 현황은 <표 2>와 같다.

중복물질을 제외한 DB 구축물질 1,615종 중 정보보호(영업비밀) 등으로 CAS No.를 제출하지 않았거나 고분자화합물로 인해 CAS No.가 미부여된 789종을 제외하고, CAS No.가 있는 물질 826종을 1차로 선별하였다.

<표 2> 1992~2018년 전자산업 관련 신규화학물질 DB 구축 현황

[단위 : 종, %]

구분	합계			정보보호 대상			정보보호 비대상		
	계	일반 화학물질	고분자 화합물	계	일반 화학물질	고분자 화합물	계	일반 화학물질	고분자 화합물
계	1,615 (100)	984 (60.9)	631 (39.1)	682 (42.2)	355 (22.0)	327 (20.2)	933 (57.8)	629 (39.0)	304 (18.8)
CAS No. 有	826 (51.1)	594 (36.7)	232 (14.4)	192 (11.9)	133 (8.2)	59 (3.7)	634 (39.3)	461 (28.6)	173 (10.7)
CAS No. 無	789 (48.9)	390 (24.2)	399 (24.7)	490 (30.3)	222 (13.8)	268 (16.5)	299 (18.5)	168 (10.4)	131 (8.1)

1차 선별물질을 대상으로 예측프로그램의 입력 쿼리인 SMILES string이나 MOL file을 검색한 결과 395종(47.8%)이 검색되었다. SMILES string 등의 검색 비율을 높이고자 전문가 심의에서 제안해 주신 NIH의 Chemical Identifier Resolver(https://cactus.nci.nih.gov/chemical/structure) 온라인 사이트를 활용한 결과, 17종이 검색되었으나 기존의 395종에 포함되어 있었다. 이에 2차 대상물질로는 SMILES string이 있는 395종을 선별하였다.

2) 최종대상물질 선정

SMILES string이 있는 395종을 대상으로 Derek Nexus, Danish QSAR, Vega QSAR, T.E.S.T (EPA) 등 4개의 예측프로그램으로 사전 예측한 결과, 모든 프로그램에서 예측 가능한 132종을 최종 연구대상으로 선정하였다.

<표 3> 최종대상물질의 작용기(functional group) 분포 현황

[단위 : 종]

연번	작용기	물질수	연번	작용기	물질수
1	alcohol	7	21	aryl ketone	2
2	alkane	4	22	aryl naphtalene	1
3	alkane ether	1	23	aryl naphtalene phenol	1
4	alkene	14	24	aryl phenol	1
5	alkene alkylhalide	1	25	carbonate dioxolane	1
6	alkene aryl alcohol	1	26	carboxylic acid	5
7	alkene carboxylicacid	2	27	carnonic acid derivative, urea	1
8	alkene ether	3	28	cycloalkane oxolane	1
9	alkene ketone	1	29	dioxane	1
10	alkenyl halide	1	30	ester	2
11	alkyl	3	31	ether	4
12	alkyl arenes	1	32	imide pyrrolidine	1
13	alkyl halide	10	33	ketone	2
14	alkyl ether	1	34	naphthalene aryl	1
15	amine	7	35	phenyl	2
16	amine aryl	1	36	phenyl cyanide	1
17	anthracene	1	37	phenyl ester	1
18	aryl	36	38	silane	4
19	aryl biphenyl	1	39	siloxane	2
20	aryl ether	1	40	sulfonyl halide	1

3. 대상물질에 대한 QSARs 모델 적용 결과

1) Derek Nexus

Derek Nexus 프로그램을 이용하여 변이원성(복귀돌연변이) 및 발암성을 예측하였다. 예측을 위해 선택된 endpoints는 Genotoxicity, Carcinogenicity이다. Genotoxicity 적용을 위한 Species로 Escherichia coli, Salmonella typhimurium을 선택했으며, Carcinogenicity는 rat에 대하여 실시하였다.

Derek Nexus 프로그램의 예측 결과는 9개[7]의 세부 단계로 나뉜다. 이 중 예측에 대한 확실한 증거가 있는 "Certain", 적어도 한 개 이상의 강력한 증거가 있으며 이에 대해 논란의 여지가 없는 "Probable", 예측을 지지하는 증거가 있는 "Plausible", 예측결과에 동일한 가중치의 증거가 있는 "Equivocal"까지 양성으로 분류하였다.

---

7) ① Certain : There is proof that the proposition is true.
② Probable : There is at least one strong argument that the proposition is true and there are no arguments against it.
③ Plausible : The weight of evidence supports the proposition.
④ Equivocal : There is an equal weight of evidence for and against the proposition.
⑤ Doubted : The weight of evidence opposes the proposition.
⑥ Improbable : There is at least one strong argument that the proposition is false and there are no arguments that it is true.
⑦ Impossible : There is proof that the proposition is false.
⑧ Inactive, no misclassified or unclassified features : The query structure does not match any structural alerts or examples in Derek which show activity in a bacterial reverse mutation assay(Ames test). Additionally, the query structure does not contain any unclassified or misclassified features.
⑨ Inactive, contains unclassified features : some features in the molecule have not been found in the Lhasa reference set. The prediction remains negative and the unclassified features are highlighted to enable the negative prediction to be verified by expert assessment.

또한, Inactive 예측결과 중 “No misclassified or unclassified features”로 예측된 물질은 음성으로, “Contains unclassified features”로 예측된 물질은 N/A로 분류하였다. 예측결과는 <표 4>와 같다.

<표 4> Derek Nexus 예측결과

구 분	변이원성(복귀돌연변이)		발암성
	Escherichia Coli	Salmonella typhimurium	Rat
계	132종	132종	132종
양성	25종 (18.9%)	25종 (18.9%)	34종 (25.76%)
음성	90종 (68.2%)	90종 (68.2%)	1종 (0.8%)
N/A	17종 (12.8%)	17종 (12.8%)	97종 (73.5%)

변이원성의 경우 Escherichia coli와 Salmonella typhimurium의 예측결과는 동일하였다. 다만, 1종(N-hydroxybenzamide; 495-18-1)이 Escherichia coli에서는 “Probable”이었고, Salmonella typhimurium에서는 “Certain”이었으며, 양성으로 예측된 25종의 Toxtree 예측결과는 “Class 3”가 24종, “Class 1”이 1종으로 예측되었다. “Class 1”으로 예측된 물질(Methylglyoxal; 78-98-8)의 경우 MSDS 상의 변이원성 시험결과(Salmonella typhimurium)는 양성인 것으로 나타났다.

발암성의 경우 예측 가능한 물질이 26.5%(35종)에 불과하였고, 35종의 Toxtree 예측결과 “Class 3”이 34종, “Class 1”이 1종 이었다. 특히, 음성으로 예측된 물질(1-Bromo-3-chlorobenzene; 108-37-2)은 “Class 3”으로 예측되었고, 양성으로 예측된 물질(2,4-diethyl-1,5-pentanediol; 57987-55-0)은 “Class 1”로 예측되었다.

2) Danish QSAR

Danish QSAR를 이용하여 변이원성, 급성 경구독성 및 발암성을 예측하였다. 유전독성(복귀돌연변이, In vitro 염색체이상 시험, In vivo 소핵 시험) 예측을 위해 Genotoxicty를 endpoint로 선택했으며, CASE ULTRA, Leadscope, SciQSAR, Experimental(from training set)을 통합한 Battery 시스템 결과를 활용하여 적용범위(Inside applicable domain)에 해당하는 결과를 정리하였다. 발암성은 FDA RCA(US Food and Drug Administration as part of Research cooperation Agreement)의 Data set을 바탕으로 CASE Ultra와 Leadscope 프로그램에 적용한 결과를 정리하였다. 급성 경구독성 결과는 ACD Labs[8] 데이터를 활용하여 Rat(oral)을 대상으로 LD_{50}(mg/kg/d)값을 예측하였다.

변이원성 및 발암성 예측결과는 양성과 음성으로 단순하게 판단하기 어렵기 때문에 다른 예측프로그램의 결과와 비교하는 용도로 사용하였고, 변이원성 예측결과는 부록 1, 발암성 예측결과는 부록 2에 나타내었다.

급성 경구독성 예측결과는 <표 5>와 같이 GHS 분류기준[9]에 따라 분류하였다. 급성 경구독성물질로 예측된 80종에 대한 Toxtree 예측결과, “Class 3”이 74종, “Class 1”이 6종이었고, “Class 1”로 예측된 물질은 모두 급성 경구독성 구분 4에 해당하였다.

---

8) ACD Labs : Advanced Chemistry Development, Inc. 분자 특성 모델링 및 설계 소프트웨어를 통해 데이터 처리 및 지식 관리를 위한 솔루션을 제공하는 연구 개발 회사

9) 급성 경구독성 GHS 분류기준
① 구분 1 : $0 < LD_{50} \leq 5$ mg/kg, ② 구분 2 : $5 < LD_{50} \leq 50$ mg/kg,
③ 구분 3 : $50 < LD_{50} \leq 300$ mg/kg, ④ 구분 4 : $300 < LD_{50} \leq 2{,}000$ mg/kg,
⑤ 구분 외 : $LD_{50} > 2{,}000$ mg/kg

<표 5> Danish QSAR 급성 경구독성 예측결과

Danish QSAR / Toxtree	계	구분 1	구분 2	구분 3	구분 4	구분 외
계	132종 (100%)	2종 (1.5%)	-	9종 (6.8%)	69종 (52.3%)	52종 (39.4%)
Class 1	17종 (12.9%)	-	-	-	6종 (4.6%)	11종 (8.3%)
Class 2	1종 (0.8%)	-	-	-	-	1종 (0.8%)
Class 3	114종 (86.3%)	2종 (1.5%)	-	9종 (6.8%)	63종 (47.7%)	40종 (30.3%)

3) Vega QSAR Ver 1.1.5 BETA22

Vega QSAR를 활용하여 변이원성과 발암성을 예측하였다. Mutagenicity를 end point로 설정하였으며, 예측을 위해서는 프로그램 내 탑재된 CAESAR, SArPy, EPI Suite, ISS, KNN/Read-Across 등의 모델이 적용된다. Vega QSAR는 독자적인 예측을 수행하기 위해서 QSAR 모델의 개념적 속성, 화학정보를 통합한 알고리즘을 개발하였다. 이러한 매개변수들은 ADI(Applicability Domain Index)라는 단일값으로 병합된다. ADI 지수는 0에서 1부터 표기될 수 있으며 예측결과값이 0.85이상일 때 적용 가능한 영역 내에 있다는 것을 의미한다. 따라서 ADI 값의 범위에 따라 예측 결과값의 신뢰도를 정량적인 수치로 확인할 수 있었다. 또한 Similarity value로 표현되는 화합물 구조 유사성 지수 역시 0~1 사이의 값을 가지며 0.75이하 값은 예측 목표물질과 구조 유사성이 현저히 낮아지기 때문에 예측결과를 정리할 때 유의하였다. 발암성 예측은 CAESAR, ISS, IRFMN/Antares, IRFMN/ISSCAN-CGX 예측 모델이 적용되어 결과를 나타냈다.

최종결과를 정리하기 위해서 Measured Applicablility Domain Scores를 참고했으며, 예측 모델별로 예측대상 유사구조의 실험값을 비교한 Similarity index, 유사구조와의 예측 정확성을 나타낸 Accuracy index, 유사구조와 예측대상의 구조 일치성을 지수로 표현한 Concordance index를 확인함으로써 신뢰할 수

있는 예측결과인지 판단하는 과정을 거쳤다. 유사화합물질을 예측하는 데는 QSAR와 Read-across를 통합한 결과가 사용되며, 이는 예측 평가도를 강화하는 역할을 한다. 예측결과 및 신뢰도는 <표 6>과 같다.

<표 6> Vega QSAR 예측결과

[단위 : 종, %]

구 분	변이원성(복귀돌연변이)					발암성				
신뢰도	계	1	2	3	실험값	계	1	2	3	실험값
계	132 (100)	28 (21.2)	45 (34.1)	51 (38.6)	8 (6.1)	132 (100)	47 (35.6)	51 (38.6)	29 (22.0)	5 (3.8)
양성(○)	29 (22.0)	4 (3.0)	10 (7.6)	11 (8.3)	4 (3.0)	61 (46.2)	17 (12.9)	25 (18.9)	16 (12.1)	3 (2.3)
음성(X)	103 (78.0)	24 (18.2)	35 (26.5)	40 (30.3)	4 (3.0)	65 (49.2)	26 (19.7)	24 (18.2)	13 (9.9)	2 (1.5)
양성(△)	-	-	-	-	-	6 (4.6)	4 (3.0)	2 (1.5)	-	-

변이원성의 경우 양성으로 예측된 29종의 Toxtree 예측결과는 “Class 3”이 28종, “Class 1”이 1종으로 예측되었으며, “Class 1”로 예측된 물질은 Derek 예측결과와 동일한 Methylglyoxal(78-98-8)이었다.

발암성의 경우 양성이 46.2%(61종), 음성이 49.2%(65종), 양성으로 의심되는 물질이 4.6%(6종)로 예측되었고, 양성 의심물질을 포함한 67종의 Toxtree 예측결과 “Class 3”이 64종, “Class 1”이 3종 이었다. 양성으로 예측된 물질(61종) 중 공단 MSDS 상의 발암성 물질이 4종 포함되어 있었으며, 모두 “Class 3”으로 예측된 물질이었다.

4) T.E.S.T (EPA) Ver 4.2.1

T.E.S.T (EPA) 프로그램으로 변이원성과 급성 경구독성을 예측할 수 있었다. 본 프로그램에서 선택할 수 있는 8개의 QSAR 예측 방법에서 Consensus

method를 선택하여 예측했으며, 이는 T.E.S.T가 제공하고 있는 Hierarchical, FDA, Single-model, Group contribution, Nearest neighbor, Mode of action 등의 method 중에서 가장 통합적이고 정확한 결과를 얻을 수 있는 방법이기 때문이다. 이 방법은 각각의 method에서 예측 적용이 가능한 영역에 있는 독성치를 고려하여 평균값을 이용하는 방식이다. 신뢰도가 높은 결과만 예측에 적용되므로 예측 정확도 및 범위가 가장 넓은 방법으로 알려져 있다. Endpoint는 Ames mutagenicity와 acute toxicity(rat, oral) 두 가지를 선택하여 예측하였다.

변이원성 예측결과는 <표 8>과 같이 양성이 23.5%(31종), 음성이 73.5%(97종), 예측이 불가능한 물질이 3.0%(4종)로 예측되었고, 양성으로 예측된 31종의 Toxtree 예측결과는 “Class 3”이 29종, “Class 1”이 2종 이었다. “Class 1”로 예측된 물질은 Derek 및 Vega QSAR 예측결과와 동일한 Methylglyoxal(78-98-8)과 2-acryloyloxyethyl 2-hydroxyethyl phthalate(38056-88-1)이었다.

급성 경구독성의 경우 예측이 불가능한 물질 2종을 제외한 예측결과는 <표 7>과 같다. 급성 경구독성물질로 예측된 82종에 대한 Toxtree 예측결과, “Class 3”이 78종, “Class 2”가 1종, “Class 1”이 3종이었고, “Class 1” 및 “Class 2”로 예측된 물질은 모두 급성 경구독성 구분 4에 해당하였다.

<표 7> T.E.S.T (EPA) 급성 경구독성 예측결과

T.E.S.T (EPA) / Toxtree	계	구분 1	구분 2	구분 3	구분 4	구분 외
계	130종 (100%)	-	-	8종 (6.2%)	74종 (56.9%)	48종 (36.9%)
Class 1	17종 (13.1%)	-	-	-	3종 (2.3%)	14종 (10.8%)
Class 2	1종 (0.8%)	-	-	-	1종 (0.8%)	-
Class 3	112종 (86.1%)	-	-	8종 (6.2%)	70종 (53.8%)	34종 (26.1%)

5) QSAR Toolbox 4.3.1

QSAR Toolbox 4.3.1 프로그램을 활용하여 유전독성과 급성 경구독성을 endpoint로 설정하여 예측결과를 정리했다. QSAR Toolbox 예측을 위해서는 크게 5가지 단계가 수행되어야 한다. 유전독성 예측을 위한 단계는 다음과 같다.

첫 번째 Input module에서 화학물질 정보를 입력하고, 검색된 화학물질의 target-endpoint를 Genetic toxicity로 설정하여 In Vitro-Gene mutatioin을 선택한 후 Test type-Bacterial reverse mutation Assay(e.g. Ames Test) species-Salmonella typhimurium로 설정하였다.

두 번째 단계인 Profiling module은 선택된 화학물질에 대한 특정 구조와 반응 메커니즘에 적합한 profiling method를 제시한다.

유전독성 예측을 위한 method는 DNA alerts for AMES by OASIS, DNA binding by OASIS, DNA binding by OECD, in vitro mutagenicity(Ames test) alerts by ISS 등이 존재하며, profiling의 결과로 예측하고자 하는 물질들의 우선 선별 및 비교순위가 결정되기 때문에 선택한 endpoint의 신뢰성을 확보하기 위해 수행해야 할 중요한 단계 중 하나이다.

세 번째 단계는 Data module로써 화학물질 정보를 수집하는 단계이다. Data module에는 Toolbox 내 이용 가능한 모든 실험데이터가 포함되어 있으며, 인체 유해성을 예측하기 위한 39,130개의 화학물질과 1,038,214개의 endpoint 값을 제공하고 있다.

유전독성 예측에서 활용 가능한 databases는 Genotoxicity OASIS, Bacterial mutagenicity ISSTY, Toxicity Japan MHLW 등이 있다.

네 번째 Category definition 단계는 목표 화학물질에 관련된 독성학적으로 의미 있는 화학물질을 grouping하는 단계이다. 유전독성 예측을 위해서 “Organic functional groups”를 적용했으며, 선택 결과로 예측하고자 하는 물질과 유사한 화학물질로 분류되었지만, 구조 또는 메커니즘적으로 다른 물질들은 Subcategorize 기능을 이용하여 제거하였다.

QSAR Toolbox는 예측 마지막 단계인 Data gap filling을 위해 Trend analysis, Read across, (Q)SAR 총 세 가지 방법을 제공하고 있다.

Trend analysis 방법은 급성 수생환경 유해성과 같이 endpoint와 생체 매개변수 사이의 선형관계가 예상되는 독성 예측에 적용 가능한 반면, Read across 방법은 생체 매개변수와 가장 근접한 독성 값 평균을 나타내기 때문에 유전독성과 급성 경구독성을 예측하는 경우 Read across 방법을 선택하여 결과를 정리하였다. Read across의 예측결과로는 log Kow[10] 값과 해당 예측 값의 범위가 나타나며, 결과 값이 유효성을 가지려면 log Kow 범위 안에 예측 값이 포함되어야 한다.

급성 경구독성을 예측하기 위한 단계는 다음과 같다.

Input module은 유전독성과 동일한 과정으로 수행하며, Profiling module 설정을 위해 endpoint를 Acute toxicity-Rat(oral) LD_{50}값으로 선택했다. 예측 적합한 Profiling method는 Acute oral Toxicity로 제시되며 Databases는 Acute Oral toxicity DB를 활용했다.

Category definition module 단계에서 “Organic functional groups”를 선택하였으며, Read-across 방식을 적용하였다. 예측결과는 부록에 나타내었다.

변이원성 예측결과는 <표 8>과 같이 양성이 12.1%(16종), 음성이 69.7%(92종), 예측이 불가능한 물질이 18.2%(24종)로 예측되었고, 양성으로 예측된 16종의 Toxtree 예측결과는 “Class 3”이 15종, “Class 1”이 1종 이었다. “Class 1”로 예측된 물질은 Derek, Vega QSAR 및 T.E.S.T (EPA) 예측결과와 동일한 Methylglyoxal(78-98-8)이었다.

10) log kow : 옥탄올-물 분배계수를 뜻하며, 두 혼합되지 않은 옥탄올과 물에서의 용질 분포를 나타내는 계수이다.

<표 8> T.E.S.T (EPA) 및 QSAR Toolbox의 변이원성 예측결과

구 분	T.E.S.T (EPA)	QSAR Toolbox
계	132종	132종
양성	31종 (23.5%)	16종 (12.1%)
음성	97종 (73.5%)	92종 (69.7%)
N/A	4종 (3.0%)	24종 (18.2%)

급성 경구독성의 경우 log Kow 범위 안에 예측 값이 없는 물질(9종)과 구조가 없어 예측이 불가능한 물질(10종)을 제외하고 113종에 대한 예측결과를 <표 9>에 나타내었다. 급성 경구독성물질로 예측된 62종에 대한 Toxtree 예측결과, “Class 3”이 54종, “Class 1”이 8종이었고, “Class 1”로 예측된 물질은 모두 급성 경구독성 구분 4에 해당하였다.

<표 9> QSAR Toolbox 급성 경구독성 예측결과

QSAR Toolbox / Toxtree	계	구분 1	구분 2	구분 3	구분 4	구분 외
계	113종 (100%)	-	-	4종 (3.6%)	58종 (51.3%)	51종 (45.1%)
Class 1	14종 (12.4%)	-	-	-	8종 (7.1%)	6종 (5.3%)
Class 2	1종 (0.9%)	-	-	-	-	1종 (0.9%)
Class 3	98종 (86.7%)	-	-	4종 (3.6%)	50종 (44.2%)	44종 (38.9%)

6) Toxtree 2.3.1

독성정도를 예측하는 기능인 Cramer rule을 적용하여 132종의 물질을 예측한 결과, Class 1으로 분류된 물질이 17종이었으며, Class 2는 1종, Class 3는

114종이 분류되었다.

Class 1에 해당되는 물질은 단순 화학구조를 가지며, 경미한 급성 경구독성을 띠는 특성이 있다. Class 2는 확실한 독성 구조를 갖고 있지 않은 중간체 화학물질을 의미하며, Class 3은 강한 초기 독성을 지니고 있진 않지만, 확실한 독성 작용기를 포함하고 있거나, 중요한 독성을 가질 가능성이 있는 경우를 의미한다.

4. 각 예측프로그램을 이용한 예측결과 비교·검토

1) 변이원성(복귀돌연변이)

변이원성 관련 복귀돌연변이의 경우 5개의 예측프로그램을 이용하여 예측 가능하였다.

Derek Nexus에서 양성으로 예측된 25종을 기준으로 비교한 결과 Danish QSAR에서는 17종이 양성 가능성이 있는 물질로 예측되었고, Vega QSAR에서는 19종, T.E.S.T (EPA)는 16종, QSAR Toolbox는 12종으로 나타났다.

모든 프로그램에서 양성으로 예측된 물질은 9종으로 ① 1-Piperidinecarboxylic acid, 4-hydroxy-, 1,1-dimethylethyl ester(110-57-6), ② pyrene(129-00-0), ③ 9-bromoanthracene(1564-64-3), ④ chloromethyl)-1,1′-biphenyl(1667-10-3), ⑤ 2,2′-[methylenebis(phenyleneoxymethylene)]bisoxirane(39817-09-9), ⑥ anthracene-9-carbaldehyde(642-31-9) ⑦ 2-[2,5-di- tert-butyl-4-(oxiran-2-ylmethoxy)phenoxy-methyl]oxirane(64777-22-6), ⑧ 1-Nitronaphthalene(86-57-7), ⑨2,3-Dibromo-1-propanol(96-13-9)) 이었으며, Toxtree 예측결과 모두 “Class 3”에 해당하였다.

양성으로 예측된 9종을 시험자료와 비교한 결과, 사업장에서 제출된 자료에서 양성인 물질이 3종(chloromethyl)-1,1′-biphenyl / 1-Nitronaphthalene / 2,3-Dibromo-1-propanol), 음성인 물질이 3종(pyrene / 9-bromoanthracene /anthracene-9-carbaldehyde), 안전보건공단 MSDS 상에서 양성인 물질이 1종(2,3-Dibromo-1-propanol)으로 나타났다.

2) 발암성

발암성의 경우 Derek Nexus, Danish QSAR, Vega QSAR 등 3개의 예측프로그램에서 예측이 가능하였다.

Derek Nexus에서 양성으로 예측된 34종을 기준으로 비교한 결과 Danish QSAR에서는 22종이 양성 가능성이 있는 물질로 예측되었고, Vega QSAR에서는 20종이 양성인 것으로 나타났다.

3개의 모든 프로그램에서 양성으로 예측된 물질은 11종으로 ① 1-Piperidinecarboxylic acid, 4-hydroxy-, 1,1-dimethylethyl ester;N-boc-4-hydroxypiperidine(110-57-6), ② chloromethyl)-1,1′-biphenyl(1667-10-3), ③ [3-(2,3-epoxypropoxy)propyl] diethoxymethylsilane(2897-60-1), ④ 2,2′-[methylenebis (phenyleneoxymethylene)] bisoxirane(39817-09-9), ⑤ 2-(p-methoxystyryl) -4,6-bis(trichloromethyl)-1,3,5-triazine(42573-57-9), ⑥ Cyclotetrasiloxane, 2,4,6,8-tetramethyl-2-[3-(oxiranylmethoxy) propyl]-(61614-55-9), ⑦ 2-[2,5-di-tert-butyl-4-(oxiran-2-ylmethoxy)phenoxymethyl] oxirane(64777-22-6), ⑧ 1,1,2,3,4,4-hexafluorobuta-1,3-diene(685-63-2), ⑨ propyl 3,4,5-tris[[(6-diazo-5,6-dihydro-5-oxo-1-naphthyl)sulphonyl]oxy]benzoate(73003-79-9), ⑩ 3,3′-(3,3′,5,5′-Tetramethyl-1,1′-biphenyl-4,4′-diylbisoxy)bis(1,2-epoxypropane) (85954-11-6), ⑪ 2,3-Dibromo-1-propanol(96-13-9) 이었으며, Toxtree 예측결과 모두 “Class 3”에 해당하였다.

양성으로 예측된 11종 중 사업장에서 제출된 자료에서 양성인 물질이 1종(2,3-Dibromo-1-propanol), 안전보건공단 MSDS 상에서 발암성인 물질이 2종(1-Piperidinecarboxylic acid, 4-hydroxy-, 1,1-dimethylethyl ester;N-boc-4-hydroxypiperidine / 3,3′-(3,3′,5,5′-Tetramethyl-1,1′-biphenyl-4,4′-diylbisoxy)bis(1,2-epoxypropane))으로 나타났다.

3) 급성 경구독성

급성독성의 경우 경구독성에 대하여 Danish QSAR, T.E.S.T (EPA), QSAR Toolbox 등 3개의 예측프로그램에서 예측이 가능하였다.

Danish QSAR는 132종 모두 예측되었고, T.E.S.T (EPA)는 130종, QSAR Toolbox는 113종이 예측 가능하였다.

Danish QSAR의 경우 GHS 분류기준의 구분 1에 해당하는 물질이 2종 (1-(1-methyl-2-propoxyethoxy)propan-2-ol / 2-(piperidin-1-yl)ethanol), 구분 3에 해당하는 물질이 9종, 구분 4에 해당하는 물질이 69종 예측되었다.

T.E.S.T (EPA)의 경우 GHS 분류기준의 구분 3에 해당하는 물질이 8종, 구분 4에 해당하는 물질이 74종 예측되었고, QSAR Toolbox의 경우 GHS 분류기준의 구분 3에 해당하는 물질이 4종, 구분 4에 해당하는 물질이 58종 예측되었다.

GHS 분류기준에 따라 Danish QSAR, T.E.S.T (EPA), QSAR Toolbox에서 모두 급성 경구독성($0 < LD_{50} \leq 2{,}000$ mg/kg)으로 예측된 물질은 33종이었으며, 해당물질에 대한 Toxtree 예측결과는 “Class 3”이 31종, “Class 1”이 2종이었다. 급성 경구독성물질로 예측된 33종 중 사업장에서 제출된 자료에서 급성 경구독성인 물질은 10종, 공단 MSDS 자료에서 급성 경구독성인 물질은 9종으로 <표 10>과 같다.

<표 10> GHS 분류기준에 따른 급성 경구독성 예측물질

연번	CAS No.	물질명	Danish QSAR	T.E.S.T (EPA)	QSAR Toolbox	Toxtree	사업장 제출자료	공단 MSDS
1	10147-36-1	propanesulphonyl chloride	920	871	726	Class 3	LD_{50} 500 mg/kg 실험종: Rat	
2	106-45-6	p-Toluenethiol	250	308	245	Class 3	LD_{50} 300~2,000 mg/kg 실험종: Rat	
3	1079-66-9	chlorodiphenylphosphine	480	500	1,880	Class 3	LD_{50} 316 mg/kg	LD_{50} 316 mg/kg 실험종: Rat
4	110-57-6	1-Piperidinecarboxylic acid, 4-hydroxy-, 1,1-dimethylethyl ester	160	475	1,350	Class 3		LD_{50} 300~2,000 mg/kg 실험종: Rat
5	110-87-2	3,4-dihydro-2H-pyran	1,700	1,123	1,360	Class 3		
6	1119-85-3	trans-3-Hexenedinitrile	440	289	1,500	Class 3		
7	120-73-0	purine	320	615	917	Class 3		
8	132-64-9	Dibenzofuran	620	587	1,520	Class 3	LD_{50} > 2,000 mg/kg 실험종: Rat	LD_{50} 500 mg/kg 실험종: Rat
9	137-07-5	2-aminothiophenol	180	290	628	Class 3		LD_{LO} 500 mg/kg 실험종: Rat
10	150-68-5	Urea, N'-(4-chlorophenyl)-N,N-dimethyl-	1,300	584	808	Class 3		LD_{50} 1,053 mg/kg 실험종: Rat
11	15112-89-7	N,N,N',N',N'',N''-hexamethylsilanetriamine	1,100	1,496	1,120	Class 3		
12	1680-21-3	1,2-ethanediylbis(oxy-2,1-ethanediyl) diacrylate	1,200	772	1,490	Class 1	LD_{50} 300~2,000 mg/kg 실험종: Rat	
13	17351-75-6	1,4-Cyclohexanedimethanol divinyl ether, mixture of isomers	2,000	1,905	1,720	Class 3		LD_{50} > 5,000 mg/kg 실험종: Rat
14	1739-84-0	1,2-dimethylimidazole	820	559	757	Class 3	LD_{50} 1,300 mg/kg	LD_{50} 1,300 mg/kg 실험종: Rat
15	1763-23-1	perfluorooctane-1-sulfonic acid heptadecafluorooctane-1-sulphonicacid	1,500	207	435	Class 3		LD_{50} 251 mg/kg 실험종: Rat
16	18039-42-4	5-Phenyl-1H-tetrazole	1,900	1,075	753	Class 3		
17	205-43-6	benzo[b]naphtho[1,2-d]thiophene	1,500	707	295	Class 3		LD_{50} > 2,000 mg/kg 실험종: Rat
18	2128-93-0	4-phenylbenzophenone	1,900	1,042	1,410	Class 3	LD_{50} > 4,700 mg/kg	
19	2382-96-9	benzoxazole-2-thiol	220	331	1,050	Class 3	LD_{50} > 2,000 mg/kg 실험종: Rat	
20	27607-77-8	trimethylsilyl trifluoromethanesulphonate	1,300	1,357	1,850	Class 3		
21	326-62-5	2-fluorophenylacetonitrile	160	600	1,050	Class 3	LD_{50} 500 mg/kg 실험종: Rat	
22	3586-14-9	phenyl m-tolyl ether	1,800	1,919	816	Class 3	LD_{50} > 2,000 mg/kg 실험종: Rat	LD_{50} 2,509 mg/kg 실험종: Mouse
23	42867-40-3	3,3-dimethyl-2-oxobutyronitrile	410	104	796	Class 3	LD_{50} 50 mg/kg 실험종: Rat	
24	491-45-2	Chemical Name (1,1'-Biphenyl)-2,3',4,5',6-pentol	1,500	1,636	935	Class 3		
25	50940-49-3	Butanedioic acid, 1-[2-[(1-oxo-2-propen-1-yl)oxy]ethyl] ester	1,300	542	552	Class 1	LD_{50} 300~2,000 mg/kg 실험종: Rat	
26	534-85-0	N-phenyl-o-phenylenediamine	540	719	1,060	Class 3		
27	542-91-6	N-phenyl-o-phenylenediamine	1,600	1,425	975	Class 3	LD_{50} > 2,000 mg/kg 실험종: Rat	
28	545-06-2	Trichloroacetonitrile	540	231	250	Class 3	LD_{50} 250 mg/kg	LD_{50} 250 mg/kg 실험종: Rat
29	559-40-0	octaphenylcyclotetrasiloxane	820	516	495	Class 3		
30	670-83-7	1h-imidazole, 2,4-diphenyl-	740	1,777	126	Class 3		
31	6789-99-7	4,5,6,7-tetrahydro-1H-benzotriazole	330	304	326	Class 3		
32	71255-78-2	1,3,5-Triazine, 2-(1,3-benzodioxol-5-yl)-4,6-bis(trichloromethyl)-	1,100	1,762	1,200	Class 3	LD_{50} > 2,000 mg/kg	
33	96-13-9	2,3-Dibromo-1-propanol	580	821	1,010	Class 3	LD_{50} 200 mg/kg 실험종: Rat	LD_{50} 681 mg/kg 실험종: Rat

Ⅳ. 고 찰

화학물질의 독성을 결정하는 것은 인간, 동물, 식물 또는 환경에 대한 유해한 영향을 식별하는 데 필요하다. 동물 모델은 오랫동안 독성시험에 사용되었지만, 시간, 생명윤리적 문제 및 재정적 부담으로 인해 제한점이 있어 화학물질의 독성을 평가하기 위한 수학적 계산방법이 유용하게 사용되고 있다. 인실리코 독성학(*In silico* toxicology)은 화학물질의 독성을 분석, 시뮬레이션, 시각화 또는 예측을 위한 전산학적 방법을 사용하는 독성평가의 한 유형으로, 기존의 독성시험을 보완하여 독성예측, 유해성평가 화학물질의 우선순위 지정, 독성시험지침 작성 및 신약설계 등을 목표로 하고 있으며, 규제 혹은 새로운 성분 스크리닝(screening)의 측면으로 볼 때 독성 및 효능을 예측하기 위한 인실리코(*in silico*) 접근은 점차 증가하게 되었다.

유럽연합은 이미 2013년부터 화장품 관련 동물실험 규제조치를 시행하고 있으며, 미국 환경보호청은 2008년부터 '독성예보관(ToxCast; Toxicity Forecater)'이라고 불리는 중장기 연구개발 프로젝트를 통해 동물을 이용하지 않고 컴퓨터, 유전체(Genome), 단백체(Proteome) 등을 대상으로 짧은 시간에 다수의 화학물질이 미치는 유해한 영향을 평가·예측할 수 있는 방법을 개발 중이며, 일본은 인공지능(AI) 및 빅데이터 기반의 화학물질의 안전성평가 방법을 개발해 2019년부터 단계적으로 도입할 예정이라고 발표한 바 있다.

또한, EU REACH 및 환경부 화평법에서도 QSARs 등 비시험방법에 의한 자료제출을 허용하고 있다. 다만, QSARs를 이용하여 실험결과를 대신할 경우에는 과학적 유효성이 확립된 QSAR 모델로부터 결과가 유도되어야 하며, 대상 화학물질이 QSAR 모델의 적용 가능한 영역 내에 포함되어야 하고, 결과가 분류·표지 및/또는 위해성평가를 위해 적절하여야 한다. 또한 적용된 방법에 대

해 신뢰할 수 있는 문서가 제공되어야 한다.

안전보건분야에서 동물실험을 대체하여 QSARs을 활용할 수 있는 분야는 산업안전보건연구원에서 수행하고 있는 유전독성 및 발암성 시험의 후보물질을 선정하기 위한 사전 스크리닝용이나 안전보건공단의 MSDS를 통해 건강유해성 정보를 제공하기 위해 사용할 수 있다. 제도적으로는 화평법과 같이 신규화학물질 유해성·위험성 조사보고서 제출서류로 허용할 수 있으나, 법 규제 목적으로 활용하기 위해서는 화학물질의 구조 특성에 따른 선별적 또는 배터리식 예측프로그램 활용 등에 대한 추가적인 연구와 QSARs 사용지침서 개발·보급 및 전문가 양성교육 등이 선행되어야 할 것으로 사료된다.

다만, 유전독성(복귀돌연변이 및 소핵) 시험성적서 제출이 생략되는 1톤 미만의 신규화학물질을 대상으로 유전독성 관련 정보를 제공하는데 활용할 수 있을 것으로 사료된다. 특히, 최근 3년('16년~'18년) 동안 제출된 신규화학물질(1,198종) 중 1톤 미만의 신규화학물질이 58.4%(674종)를 차지하고 있으므로 향후 예측프로그램을 적극 활용한다면 유전독성이 우려되는 신규화학물질로부터 근로자의 건강장해를 예방할 수 있을 것으로 기대된다.

신규화학물질 등에 대한 변이원성, 발암성 및 급성 경구독성 예측결과를 사업장에 제공하기 위해서는 Toxtree 등으로 화학물질의 독성정도를 우선 평가한 후 고독성이 예측된 물질을 대상으로 2개 이상의 예측프로그램 결과를 비교·분석하는 배터리식 예측프로그램 활용이 필요하다. 또한, 예측결과가 음성인 경우에는 근로자 건강장해 우려가 남아있기 때문에 독성시험을 통해 최종 확인 후 사업장에 제공하여야 할 것으로 사료된다.

다만, 변이원성의 경우 성창훈 등(국립환경과학원, 2016)이 유해성심사를 통

해 보유하고 있는 Ames 시험결과와 TOPKAT® 및 Derek의 예측결과를 비교·분석한 결과와 홍문기 등(산업안전보건연구원, 2018)이 QSAR Toolbox 예측결과를 안전보건공단 MSDS 상의 복귀돌연변이 시험결과와 비교·분석한 결과에서 Derek의 예측 정확도가 85.4%, QSAR Toolbox의 예측 정확도가 87.6%로 높게 나타났으므로 본 연구에서 Derek 및 QSAR Toolbox 등을 활용하여 예측된 변이원성 결과는 사업장에 제공하여도 무방할 것으로 판단된다.

본 연구에서 수행된 발암성 및 급성 경구독성 예측결과는 연구원에서 수행예정인 유전독성 및 발암성 시험의 후보물질 선정을 위한 사전 스크리닝용으로 활용할 수 있으나, 사업장에 제공하기 위해서는 예측프로그램별 예측 정확도 평가 등 신뢰도 확인을 위한 추가적인 연구가 필요하다.

Ⅴ. 결 론

반도체 및 LCD 등 전자산업에서 사용되는 신규화학물질 1,754종을 대상으로 유해성 예측프로그램을 이용하는데 있어서 중복물질 139종을 제외한 1,615종 중 정보보호(영업비밀) 요청 및 고분자화합물 등으로 CAS No.가 없는 물질이 48.9%(789종)에 달했고, CAS No.가 있는 물질(826종) 중에서도 QSAR 예측이 어려운 고분자화합물이 14.4%(232종)에 해당하였다.

예측프로그램의 입력 쿼리인 SMILES string을 파악하여 Derek Nexus, Danish QSAR, Vega QSAR, T.E.S.T (EPA)로 사전 예측된 132종을 대상으로 변이원성, 발암성 및 급성 경구독성을 예측하였다.

변이원성의 경우 Derek Nexus, Danish QSAR, Vega QSAR, T.E.S.T (EPA), QSAR Toolbox 등 5개 예측프로그램을 사용한 결과 모든 프로그램에서 양성인 물질이 9종으로 나타났고, 발암성은 Derek Nexus, Danish QSAR, Vega QSAR를 활용한 결과 11종이 모든 프로그램에서 양성이었고, 급성 경구독성은 Danish QSAR, T.E.S.T (EPA), QSAR Toolbox를 활용한 결과 GHS 분류기준으로 급성 경구독성($0 < LD_{50} \leq 2{,}000$ mg/kg)으로 예측된 물질은 33종이었다. 추가로 화학물질의 독성정도를 예측하기 위하여 Toxtree를 활용한 결과, 변이원성 및 발암성에서 양성으로 예측된 물질은 모두 “Class 3”에 해당하였으며, 급성 독성물질로 예측된 물질도 93.9%(31종)가 “Class 3”에 해당하였다.

본 연구에서 수행된 변이원성 예측결과는 5개의 예측프로그램을 활용한 결과이며, Derek Nexus 및 QSAR Toolbox를 활용한 기존 연구에서도 예측 정확도가 85% 이상으로 나타났으므로 모든 프로그램에서 양성으로 예측된 물질(9

종)의 경우 안전보건공단의 MSDS나 기술자료 등을 통하여 사업장 및 근로자에게 제공할 수 있을 것으로 사료된다.

발암성 및 급성 경구독성 예측결과는 산업안전보건연구원에서 수행하고 있는 유전독성 및 발암성 시험 후보물질 선정을 위한 사전 스크리닝용으로 활용 가능하나 사업장 등에 제공하기 위해서는 기존 시험자료와 예측결과를 비교하여 예측 정확도를 평가하는 등 신뢰도 확인을 위한 연구가 선행되어야 할 것으로 사료된다.

안전보건분야에서 화학물질 관리를 위해 예측프로그램을 활용하는 방안으로는 산업안전보건연구원에서 수행하고 있는 유전독성 및 발암성 시험의 후보물질 선정을 위한 사전 스크리닝용이나 안전보건공단의 MSDS를 통해 건강유해성 정보를 제공하기 위하여 사용할 수 있으며, 유전독성 시험성적서 제출을 생략하고 있는 1톤 미만의 신규화학물질을 대상으로 변이원성을 예측하여 사업장 및 근로자 등에게 건강장해 예방자료로 제공할 수 있을 것으로 사료된다.

참고문헌

Ames B.N, Durston W.E, Yamasaki E, Lee F.D. Carcinogens are mutagens: a simple test system combining liver homogenates for activation and bacteria for detection. Proc Natl Acad Sci. 1973;70(8):2281 – 2285.

Arockiasamy D.L, Radhika S, Parthasarathi R, Nair B.U. Synthesis and DNA-binding studies of two ruthenium(II) complexes of an intercalating ligand. Eur J Med Chem. 2009;44(5):2044 – 2051.

Assessment and control of DNA reactive(mutagenic) impurities in pharmaceuticals to limit potential carcinogenic risk. In: Guideline I.H.T, ed. International Conference on Harmonization of Technical Requirements for Registration of pharmaceuticals for Human Use(ICH), Geneva. 2014.

Barber C, Amberg A, Custer L, Dobo K.L, Glowienke S, Van Gompel J, et al. Establishing best practise in the application of expert review of mutagenicity under ICH M7. Regul Toxicol Pharmacol. 2015;73(1):367 – 377.

Benigni R, Battistelli C.L, Bossa C, Colafranceschi M, Tcheremenskaia O. Mutagenicity, carcinogenicity, and other end points. Comput Toxicol. 2013;II:67 – 98.

Benigni R, Bossa C. Mechanisms of chemical carcinogenicity and mutagenicity: a review with implications for predictive toxicology. Chem Rev. 2011;111(4):2507 – 2536.

Brooks B.R, Brooks C.L, MacKerell A.D, Nilsson L, Petrella R.J, Roux B, et al. CHARMM: the biomolecular simulation program. J Comput Chem. 2009;30(10):1545 – 1614.

Cherkasov A, Muratov E.N, Fourches D, Varnek A, Baskin I.I, Cronin M, et al. QSAR Modeling: where have you been? Where are you going to? J Med Chem. 2014;57(12):4977–5010.

Claxton L.D, de Umbuzeiro G.A, DeMarini D.M. The Salmonella mutagenicity assay: the stethoscope of genetic toxicology for the 21st century. Environ Health Perspect. 2010;118(11):1515.

Committee IS. Genotoxicity: a standard battery for genotoxicity testing of pharmaceuticals. ICH harmonised tripartite guideline. 1997:1-6.

Cronin M.T, Jaworska J.S, Walker J.D, Comber M.H, Watts C.D, Worth A.P. Use of QSARs in international decision-making frameworks to predict health effects of chemical substances. Environ Health Perspect. 2003;111(10):1391.

Cronin M.T.D. In silico tools for toxicity prediction. In: New horizons in predictive toxicology: current status and application. The Royal Society of Chemistry; 2012:9–25.

Cunny H, Hodgson E. Toxicity testing. A textbook of modern toxicology. 2004:353.

Dearfield K.L, Gollapudi B.B, Bemis J.C, Benz R.D, Douglas G.R, Elespuru R.K, et al. Next generation testing strategy for assessment of genomic damage: a conceptual framework and considerations. Environ Mol Mutagen. 2016;58:264–283.

Ekins S, Nikolsky Y, Nikolskaya T. Techniques: application of systems biology to absorption, distribution, metabolism, excretion and toxicity. Trends Pharmacol Sci. 2005;26(4):202–209.

Giarratano J.C, Riley G. Expert systems. PWS publishing Co.; 1998.

Greene N, Dobo K.L, Kenyon M.O, Cheung J, Munzner J, Sobol Z, et al. A

practical application of two in silico systems for identification of potentially mutagenic impurities. Regul Toxicol Pharmacol. 2015;72(2):335 – 349.

Greene N. Computer systems for the prediction of toxicity: an update. Adv Drug Deliv Rev. 2002;54(3):417-431.

Guideline IHT. Guidance on specific aspects of regulatory genotoxicity tests for pharmaceuticals S2A. Adv Drug Deliv Rev. 2002;54(3):417 – 431.

Hansch C, Hoekman D, Gao H, Comparative Q.S.A.R. Toward a deeper understanding of chemicobiological interactions. Chem Rev. 1996;96(3):1045 – 1076.

Hansch C, Hoekman D, Leo A, Weininger D, Selassie C.D. Chem-bioinformatics: comparative QSAR at the interface between chemistry and biology. Chem Rev. 2002;102(3):783 – 812.

Hansch C, Maloney P.P, Fujita T, Muir R.M. Correlation of biological activity of phenoxyacetic acids with Hammett substituent constants and partition coefficients. Nature. 1962;194(4824):178 – 180.

Hess B, Kutzner C, Van Der Spoel D, Lindahl E. GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. J Chem Theory Comput. 2008;4(3):435 – 447.

Holt P.A, Chaires J.B, Trent J.O. Molecular docking of intercalators and groove-binders to nucleic acids using Autodock and Surflex. J Chem Inf Model. 2008;48(8):1602 – 1615.

Hsu K.-H, Su B.-H, Tu Y.-S, Lin O.A, Tseng Y.J. Mutagenicity in a molecule: identification of core structural features of mutagenicity using a scaffold analysis. PLoS One. 2016;11(2):e0148900.

Kavlock R, Dix D. Computational toxicology as implemented by the US EPA: providing high throughput decision support tools for screening and assessing chemical exposure, hazard and risk. J Toxicol Environ Health B.

2010;13(2–4):197–217.

Knudsen T.B, Keller D.A, Sander M, Carney E.W, Doerrer N.G, Eaton D.L, et and molecular dynamics. Comput Sys Pharmacol Toxicol. 2017:174–190.

Kotandeniya D, Murphy D, Yan S, Park S, Seneviratne U, Koopmeiners J.S, et al. Kinetics of O 6-pyridyloxobutyl-2′-deoxyguanosine repair by human O 6-alkylguanine DNA alkyltransferase. Biochemistry. 2013;52(23):4075-4088.

Kuhn H, Demidov V.V, Coull J.M, Fiandaca M.J, Gildea B.D, Frank-Kamenetskii M.D. Hybridization of DNA and PNA molecular beacons to single-stranded and double-stranded DNA targets. J Am Chem Soc. 2002;124(6):1097–1103.

Lapenna S, Gatnik M.F, Worth A.P. Review of QSAR models and software tools for predicting acute and chronic systemic toxicity. Luxembourg: Publications Office of the European Union; 2010.

Lionta E, Spyrou G, Vassilatis D.K, Cournia Z. Structure-based virtual screening for drug discovery: principles, applications and recent advances. Curr Top Med Chem. 2014;14(16):1923–1938.

Miller E.C, Miller J.A. Mechanisms of chemical carcinogenesis. Cancer. 1981;47(S5):1055–1064.

Miller E.C, Miller J.A. Searches for ultimate chemical carcinogens and their reactions with cellular macromolecules. Cancer. 1981;47(10):2327–2345.

Modi S, Li J, Malcomber S, Moore C, Scott A, White A, et al. Integrated in silico approaches for the prediction of Ames test mutagenicity. J Comput Aided Mol Design. 2012;26(9):1017–1033.

Muster W, Breidenbach A, Fischer H, Kirchner S, Müller L, Pähler A. Computational toxicology in drug development. Drug Discov Today. 2008;13(7):303–310.

OECD, bacterial reverse mutation test, 1997, July, 21st.

OECD, Guidance notes for analysis and evaluation of chronic toxicity and carcinogenicity studies, 2002, December, 12th

Parthasarathi R, Elango M, Padmanabhan J, Subramanian V, Roy D, Sarkar U, et al. Application of quantum chemical descriptors in computational medicinal chemistry and chemoinformatics. 2006.

Patlewicz G. et al., An evaluation of the implementation of the Cramer classification scheme in the toxtree software, SAR and QSAR in environmental research, 2008, 495-524

Pearlman D.A, Case D.A, Caldwell J.W, Ross W.S, Cheatham T.E, DeBolt S, et al. AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. Comput Phys Commun. 1995;91(1–3):1–41.

Phillips J.C, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, et al. Scalable molecular dynamics with NAMD. J Comput Chem. 2005;26(16):1781–1802.

Raies A.B, Bajic V.B. In silico toxicology: computational methods for the prediction of chemical toxicity. Wiley Interdiscip Rev Computa Mol Sci. 2016;6(2):147-172.

Randić M. Generalized molecular descriptors. J Math Chem. 1991;7(1):155–168.

Ridings J, Barratt M, Cary R, Earnshaw C, Eggington C, Ellis M, et al. Computer prediction of possible toxic action from chemical structure: an update on the DEREK system. Toxicology. 1996;106(1):267–279.

Segall M.D, Barber C. Addressing toxicity risk when designing and selecting compounds in early drug discovery. Drug Discov Today. 2014;19(5):688–693.

Singh S, Handa T, Narayanam M, Sahu A, Junwal M, Shah R.P. A critical review on the use of modern sophisticated hyphenated tools in the characterization of impurities and degradation products. J Pharm Biomed Anal. 2012;69:148 – 173.

Snyder R.D, Smith M.D. Computational prediction of genotoxicity: room for improvement. Drug Discov Today. 2005;10(16):1119 – 1124.

Sutter A, Amberg A, Boyer S, Brigo A, Contrera J.F, Custer L.L, et al. Use of in silico systems and expert knowledge for structure-based assessment of potentially mutagenic impurities. Regul Toxicol Pharmacol. 2013;67(1):39 – 52.

Uma Maheswari P, Rajendiran V, Palaniandavar M, Parthasarathi R, Subramanian V. Enantiopreferential DNA Binding:[{(5, 6-dmp) 2Ru} 2(μ-bpm)] 4+ induces a B-to-Z conformational change on DNA. Bull Chem Soc Japan. 2005;78(5):835 – 844.

Vijayalakshmi R, Kanthimathi M, Parthasarathi R, Nair B.U. Interaction of chromium(III) complex of chiral binaphthyl tetradentate ligand with DNA. Bioorg Med Chem. 2006;14(10):3300 – 3306.

Warshel A, Levitt M. Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. J Mol Biol. 1976;103(2):227 – 249.

Wetzel S, Klein K, Renner S, Rauh D, Oprea T.I, Mutzel P, et al. Interactive exploration of chemical space with Scaffold Hunter. Nat Chem Biol. 2009;5(8):581 – 583.

Wijeyesakere S.J, Richardson R.J. In silico chemical-protein docking and molecular dynamics. Comput Sys Pharmacol Toxicol. 2017;174-190.

Yen low et al., predicting drug-induced hepatotoxicity using QSAR and toxicogenomics approaches, Chemical research in toxicology, 2011, 1251-1262

Zeiger E. History and rationale of genetic toxicity testing: an impersonal, and sometimes personal, view. Environ Mol Mutagen . 2004;44(5):363 – 371.

성창호, 박선영, 김경태 등, QSAR 프로그램을 통한 유전독성 예측 향상 연구, 국립환경과학원, 2016.

성창호, 박선영, 최윤석 등, QSAR 프로그램을 통한 유전독성 예측력 향상 연구(Ⅱ), 국립환경과학원, 2017.

성창호, 박선영, 이진욱 등, QSAR을 활용한 독성물질탐색 예측 향상 방안 연구(Ⅰ), 국립환경과학원, 2018.

홍문기, 조지훈, 최보경 등, 화학물질정보 분야에서의 OECD Toolbox의 활용방안 연구, 산업안전보건연구원, 2018.

환경부, 신규화학물질 유해성심사제도 개선방안 연구, 2001

환경부, 미국의 화학물질 관리제도 – EPA와 OSHA의 제도를 중심으로, 1994

기타 인터넷 사이트 참고

http://alttox.org/mapp/toxicity-endpoints-tests/

http://www.ich.org/products/guidelines/safety/article/safety-guidelines.html

https://www.qsartoolbox.org/home

http://toxtree.sourceforge.net/index.html

https://www.vegahub.eu/

https://www.epa.gov/chemical-research/toxicity-estimation-software-tool-test

https://cactus.nci.nih.gov/chemical/structure

https://cdk.github.io/

https://preadmet.bmdrc.kr/

Abstract

A Study on application of hazard prediction program for acute toxic chemicals

Hyun-Ock Kim, Kyung-Taek Rim, In-Seob Lee

Chemicals Research Bureau, Occupational Safety and Health Research Institute, KOSHA
#339-30, Exporo Yuseong-Gu, Daejeon 34122, Korea

Objectives : This study aims at providing an effective method for managing toxic chemicals through the introduction and application of chemical hazard prediction programs. The predicted toxic values are acute-toxicity, mutagenicity, and carcinogenicity.

Methods : Since 1992, The project called "Hazard·Risk investigation and verification of new chemical substances" has been conducted by the Occupational Safety and Health Research Institute.
This project established databases for chemicals which frequently used in the electronics industry. (eg. semiconductors, LCD)
In this paper, 132 predictable chemicals are selected and applied to the chemical hazard prediction programs(Derek Nexus, Danish QSAR, Vega QSAR, T.E.S.T (EPA), QSAR Toolbox). Lastly, the Toxtree program is used to estimate the toxicity level of chemicals.

Results : In the case of mutagenicity, 9 chemicals are commonly "positive" when applied to 5 prediction programs. and carcinogenicity "positive" prediction results are 11 chemicals using Derek Nexus, Danish QSAR, Vega QSAR. To predict acute toxicity, Danish QSAR, T.E.S.T (EPA), QSAR Toolbox programs are available. As a result, 33 chemicals are included ($0 < LD_{50} \leq 2,000$ mg/kg) criteria in GHS classification category. In addition, toxic level grouping was conducted through the Toxtree program. "positive" results for mutagenicity and carcinogenicity (9 and 11 chemicals respectively) are all classified as "Class 3" and 31 out of 33 acute toxicity chemicals are classified as "Class 3" which accounts for 93.9%.

Conclusions : The application methods of prediction programs for the management of chemicals are divided into three major ways.
first, predicted results are available for preliminary screening of candidate chemicals for genetic toxicity and carcinogenicity animal tests conducted by the Occupational Safety and Health Research Institute.
second, results can be provided through MSDS of the Korea Occupational Safety and Health Agency(KOSHA) to provide health risk information.
finally, predicted data can be offered to new chemicals (with less than 1ton of manufacture and import) which omit submission of genotoxicity test reports for health hazard prevention.

Keywords : Aucte Toxic, Chemicals, Hazard, Prediction, Program

부록

부록 1. 변이원성 예측결과

부록 2. 발암성 예측결과

부록 3. 급성 경구독성 예측결과

부록 4. 대상물질의 용도, 작용기 및 Toxtree 예측결과

부록 1. 변이원성 예측결과

연번	Derek Nexus		Danish QSAR			Vega QSAR		T.E.S.T (EPA)	QSAR Toolbox		
	복귀돌연변이		변이원성			변이원성		변이원성	변이원성	log kow	
	E.Coli	*Salmonella typhimurium*	복귀돌연변이 (X/5)	염색체이상 (Y/2)	소핵 (Z/1)	복귀돌연변이	신뢰도	복귀돌연변이	복귀돌연변이	범위	값 (calcultated)
1	음성	음성	POS (1) NEG (4)	NEG (1)	-	음성	3	음성	음성	3.06~4.34	3.06
2	음성	음성	POS (2) NEG (3)	POS (1)	-	양성	3	음성	X (구조없음)	-	-
3	양성 (Equivocal)	양성 (Equivocal)	NEG (1)	NEG (1)	NEG (1)	양성	2	양성	음성	1.27~3.09	2.26
4	음성	음성	NEG (3)	NEG (1)	-	음성	2	음성	음성	-0.11~1.88	0.88
5	음성	음성	NEG (1)	-	-	음성	3	음성	음성	3.1~3.29	3.23
6	음성	음성	-	-	-	음성	1	음성	음성	5.25~7.93	6.70
7	N/A (contains unclassified features)	N/A (contains unclassified features)	POS (1) NEG (3)	-	-	음성	1	음성	음성	4~4.04	4.02
8	음성	음성	POS (1) NEG (3)	NEG (1)	POS (1)	음성	3	음성	음성	3.5~3.55	3.53
9	음성	음성	NEG (1)	-	-	양성	2	음성	음성	0.86~2.69	1.55
10	**양성 (Plausible)**	**양성 (Plausible)**	**POS (1) NEG (1)**	**-**	**NEG (1)**	**양성**	**3 (exp)**	**양성**	**양성**	**2.6~2.61**	**2.60**
11	음성	음성	POS (1) NEG (4)	NEG (1)	NEG (1)	음성	1	음성	음성	6.78~7.2	6.99
12	음성	음성	NEG (1)	-	-	음성	2	음성	음성	1.19~1.23	1.21
13	음성	음성	POS (2) NEG (3)	-	-	음성	3	음성	음성	0.01~1.11	0.13
14	N/A (contains unclassified features)	N/A (contains unclassified features)	POS (1) NEG (2)	POS (1)	-	음성	2	음성	양성	-0.83~-0.77	-0.82

연번	Derek Nexus		Danish QSAR			Vega QSAR		T.E.S.T (EPA)	QSAR Toolbox		
	복귀돌연변이		변이원성			변이원성		변이원성	변이원성	log kow	
	E.Coli	*Salmonella typhimurium*	복귀돌연변이 (X/5)	염색체이상 (Y/2)	소핵 (Z/1)	복귀돌연변이	신뢰도	복귀돌연변이	복귀돌연변이	범위	값 (calcultated)
15	음성	음성	POS (2) NEG (2)	NEG (1)	-	음성	2	음성	X (out of domain)	6.81~9.25	11.90
16	**양성 (Plausible)**	**양성 (Plausible)**	**POS (2) NEG (3)**	**POS (1) NEG (1)**	**-**	**양성**	**3 (exp)**	**양성**	**양성**	**4.89~5.14**	**4.93**
17	음성	음성	-	NEG (1)	-	양성	1	양성	음성	3.43~4.36	3.71
18	음성	음성	POS (1) NEG (3)	NEG (2)	-	음성	2	음성	음성	1.55~4.31	3.69
19	양성 (Plausible)	양성 (Plausible)	NEG (2)	POS (2)	-	음성	2	음성	양성	-0.64~5.2	1.77
20	음성	음성	NEG (1)	NEG (1)	-	음성	2	양성	X (구조없음)	-	-
21	음성	음성	POS (1) NEG (2)	NEG (1)	POS (1)	음성	3	음성	음성	4.38~4.47	4.42
22	음성	음성	POS (1) NEG (3)	POS (1)	POS (1)	음성	exp	음성	음성	1.7~2.58	2.03
23	N/A (contains unclassified features)	N/A (contains unclassified features)	-	-	-	음성	1	음성	음성	-1.91~-1.55	-1.68
24	**양성 (Plausible)**	**양성 (Plausible)**	**POS (1) NEG (2)**	**POS (1) NEG (1)**	**POS (1)**	**양성**	**2**	**양성**	**양성**	**4.31~6.17**	**5.24**
25	N/A (contains unclassified features)	N/A (contains unclassified features)	POS (1) NEG (1)	NEG (1)	POS (1)	양성	3	양성	X (구조없음)	-	-
26	음성	음성	NEG (2)	NEG (2)	NEG (1)	음성	3	음성	음성	-1.52~-0.68	-1.23
27	N/A (contains unclassified features)	N/A (contains unclassified features)	NEG (3)	NEG (4)	-	음성	1	음성	음성	5.65~5.76	5.72

연번	Derek Nexus		Danish QSAR			Vega QSAR		T.E.S.T (EPA)	QSAR Toolbox		
	복귀돌연변이		변이원성			변이원성		변이원성	변이원성	log kow	
	E.Coli	*Salmonella typhimurium*	복귀돌연변이 (X/5)	염색체이상 (Y/2)	소핵 (Z/1)	복귀돌연변이	신뢰도	복귀돌연변이	복귀돌연변이	범위	값 (calcultated)
28	양성 (Plausible)	양성 (Plausible)	POS (1)	-	-	양성	3	양성	양성	5.03~5.73	5.36
29	음성	음성	NEG (3)	POS (2)	-	음성	3	음성	음성	0.29~2.65	0.57
30	음성	음성	NEG (1)	-	-	음성	2	음성	X (out of domain)	3.16~3.17	3.19
31	음성	음성	NEG (2)	-	NEG (1)	음성	3	음성	음성	0.93~1.45	1.15
32	N/A (contains unclassified features)	N/A (contains unclassified features)	-	-	-	음성	1	음성	X (구조없음)	-	-
33	음성	음성	NEG (1)	NEG (2)	-	음성	2	음성	음성	4.32~4.63	4.49
34	음성	음성	NEG (1)	-	-	음성	1	음성	음성	6.15~8.16	8.04
35	N/A (contains unclassified features)	N/A (contains unclassified features)	NEG (1)	-	-	음성	1	음성	양성	-0.34~1.14	0.18
36	N/A (contains unclassified features)	N/A (contains unclassified features)	POS (1) NEG (1)	-	-	음성	1	음성	음성	9.37~9.93	9.85
37	음성	음성	NEG (2)	NEG (2)	NEG (1)	음성	3	음성	X (out of domain)	4.24~6.82	10.10
38	N/A (contains unclassified features)	N/A (contains unclassified features)	NEG (3)	POS (2)	NEG (1)	음성	exp	음성	음성	3.13~4.69	3.57
39	음성	음성	POS (1)	POS (1)	-	양성	3	양성	양성	5.16~5.93	5.34
40	N/A (contains unclassified features)	N/A (contains unclassified features)	-	-	-	음성	1	음성	음성	1.14~2.01	1.43

연번	Derek Nexus		Danish QSAR			Vega QSAR		T.E.S.T (EPA)	QSAR Toolbox		
	복귀돌연변이		변이원성			변이원성		변이원성	변이원성	log kow	
	E.Coli	*Salmonella typhimurium*	복귀돌연변이 (X/5)	염색체이상 (Y/2)	소핵 (Z/1)	복귀돌연변이	신뢰도	복귀돌연변이	복귀돌연변이	범위	값 (calcultated)
41	음성	음성	NEG (2)	NEG (1)	POS (1)	음성	exp	양성	음성	3.96~4.89	4.65
42	음성	음성	POS (1) NEG (3)	NEG (1)	-	음성	2	음성	음성	4.04~5.42	4.91
43	음성	음성	POS (1) NEG (1)	NEG (1)	NEG (1)	음성	3	음성	음성	1.81~3.43	2.11
44	음성	음성	-	NEG (1)	-	양성	3	양성	X (구조없음)	-	-
45	음성	음성	POS (1) NEG (3)	NEG (2)	-	음성	3	음성	음성	3.9~4	3.94
46	음성	음성	-	-	-	음성	1	음성	음성	2.39~4.03	2.41
47	N/A (contains unclassified features)	N/A (contains unclassified features)	POS (1) NEG (2)	POS (1) NEG (1)	-	음성	1	음성	X (구조없음)	-	-
48	N/A (contains unclassified features)	N/A (contains unclassified features)	POS (1) NEG (1)	-	-	음성	1	음성	음성	8.39	8.39
49	음성	음성	POS (1) NEG (1)	-	-	음성	3	음성	음성	3.2~4.83	4.04
50	음성	음성	NEG (3)	NEG (2)	-	음성	3	양성	음성	1.55~1.98	1.84
51	N/A (contains unclassified features)	N/A (contains unclassified features)	-	NEG (1)	POS (1)	양성	1	음성	X (out of domain)	3.33	4.27
52	양성 (Plausible)	양성 (Plausible)	POS (2) NEG (1)	-	-	양성	1	음성	음성	1.23~1.5	1.32
53	음성	음성	NEG (3)	NEG (1)	-	음성	1	음성	음성	4.74~6.61	5.77
54	음성	음성	NEG (2)	NEG (1)	NEG (1)	음성	2	음성	음성	0.35~0.87	0.64
55	음성	음성	POS (1) NEG (2)	NEG (2)	-	음성	2	음성	음성	0.34~0.56	0.43

연번	Derek Nexus		Danish QSAR			Vega QSAR		T.E.S.T (EPA)	QSAR Toolbox		
	복귀돌연변이		변이원성			변이원성		변이원성	변이원성	log kow	
	E.Coli	*Salmonella typhimurium*	복귀돌연변이 (X/5)	염색체이상 (Y/2)	소핵 (Z/1)	복귀돌연변이	신뢰도	복귀돌연변이	복귀돌연변이	범위	값 (calcultated)
56	음성	음성	NEG (2)	POS (2)	NEG (1)	음성	3	음성	음성	2.67~3.31	3.00
57	음성	음성	NEG (4)	POS (1)	-	음성	3	음성	음성	1.93~2.13	2.01
58	음성	음성	POS (1) NEG (4)	-	-	음성	2	음성	음성	1.29~2.76	1.76
59	음성	음성	POS (1) NEG (3)	-	-	음성	2	음성	X (out of domain)	1.23~1.59	-1.18
60	음성	음성	NEG (3)	-	-	음성	2	음성	음성	4.2~4.84	4.60
61	음성	음성	NEG (2)	POS (1)	-	음성	2	음성	음성	3.11~5.54	4.61
62	음성	음성	POS (1) NEG (1)	POS (2)	-	음성	3	양성	음성	0.97~1.22	1.08
63	음성	음성	POS (1) NEG (2)	NEG (1)	NEG (1)	음성	2	음성	음성	-1.33~2.03	0.38
64	음성	음성	NEG (3)	POS (2)	-	음성	3	음성	음성	-4.56~4.94	0.94
65	**양성 (Plausible)**	**양성 (Plausible)**	**POS (2) NEG (1)**	**POS (2)**	**POS (1)**	**양성**	**3**	**양성**	**양성**	**3.09~3.44**	**3.26**
66	N/A (contains unclassified features)	N/A (contains unclassified features)	POS (1) NEG (1)	NEG (1)	-	음성	1	음성	음성	2.49~2.84	2.56
67	음성	음성	POS (1) NEG (4)	POS (1)	-	음성	2	음성	X (구조없음)	-	-
68	양성 (Plausible)	양성 (Plausible)	POS (1) NEG (3)	-	-	양성	2	양성	음성	5.67~6.33	6.22
69	음성	음성	POS (1) NEG (2)	-	NEG (1)	음성	3	음성	음성	-1.08~3.25	-0.34
70	음성	음성	POS (1) NEG (1)	NEG (1)	-	음성	2	음성	X (out of domain)	2.39	0.19

연번	Derek Nexus		Danish QSAR			Vega QSAR		T.E.S.T (EPA)	QSAR Toolbox		
	복귀돌연변이		변이원성			변이원성		변이원성	변이원성	log kow	
	E.Coli	*Salmonella typhimurium*	복귀돌연변이 (X/5)	염색체이상 (Y/2)	소핵 (Z/1)	복귀돌연변이	신뢰도	복귀돌연변이	복귀돌연변이	범위	값 (calcultated)
71	음성	음성	NEG (2)	NEG (1)	-	음성	2	음성	음성	-3.66~3.05	0.57
72	음성	음성	POS (1) NEG (1)	NEG (1)	-	음성	exp	음성	양성	3~3.91	3.55
73	음성	음성	NEG (3)	-	-	음성	2	양성	음성	0.59~2.8	1.36
74	양성 (Probable)	양성 (Certain)	POS (4)	-	-	양성	exp	음성	양성	-0.91~3.96	0.61
75	음성	음성	NEG (3)	NEG (2)	-	음성	3	음성	음성	2.39~3.31	2.85
76	음성	음성	POS (2) NEG (2)	NEG (1)	NEG (1)	음성	3	음성	음성	0.04~1.08	0.81
77	음성	음성	NEG (2)	POS (2)	NEG (1)	음성	3	음성	음성	0.29~0.72	0.50
78	양성 (Equivocal)	양성 (Equivocal)	-	-	-	음성	1	양성	음성	-1.19	-1.19
79	음성	음성	NEG (3)	POS (1)	NEG (1)	음성	3	음성	음성	4.5~4.94	4.58
80	음성	음성	-	-	-	음성	2	음성	음성	-0.11~0.35	0.14
81	음성	음성	NEG (1)	-	-	음성	2	음성	음성	-0.93~ -0.59	-0.76
82	음성	음성	NEG (2)	POS (1)	-	음성	3	음성	음성	2.37~2.38	2.37
83	음성	음성	-	-	-	음성	1	음성	음성	2.24~3.11	2.61
84	음성	음성	-	-	-	양성	3	양성	음성	1.12~1.27	1.21
85	음성	음성	POS (1)	-	-	음성	1	적용안됨	X (out of domain)	7.76~12	16.50
86	양성 (Plausible)	양성 (Plausible)	NEG (1)	-	-	음성	2	음성	음성	3.41~3.53	3.47
87	음성	음성	POS (1) NEG (2)	NEG (2)	NEG (1)	음성	3	음성	X (구조없음)	-	-
88	음성	음성	POS (1) NEG (2)	POS (1) NEG (1)	POS (1)	음성	2	음성	음성	3.72~4.14	4.06
89	음성	음성	NEG (1)	NEG (2)	NEG (1)	음성	3	음성	음성	0.12~4.09	0.56

연번	Derek Nexus		Danish QSAR			Vega QSAR		T.E.S.T (EPA)	QSAR Toolbox		
	복귀돌연변이		변이원성			변이원성		변이원성	변이원성	log kow	
	E.Coli	*Salmonella typhimurium*	복귀돌연변이 (X/5)	염색체이상 (Y/2)	소핵 (Z/1)	복귀돌연변이	신뢰도	복귀돌연변이	복귀돌연변이	범위	값 (calcultated)
90	음성	음성	NEG (3)	NEG (1)	-	음성	3	음성	음성	4~4.1	4.05
91	N/A (contains unclassified features)	N/A (contains unclassified features)	POS (1) NEG (1)	-	POS (1)	음성	1	음성	음성	9.5~11.3	10.40
92	N/A (contains unclassified features)	N/A (contains unclassified features)	-	-	-	음성	2	음성	X (out of domain)	8.2~10.2	11.00
93	양성 (Plausible)	양성 (Plausible)	POS (1) NEG (2)	POS (1)	-	양성	3	양성	음성	5.04~5.08	5.08
94	음성	음성	POS (1) NEG (4)	NEG (1)	NEG (1)	음성	3	음성	음성	3.82~3.97	3.93
95	음성	음성	NEG (1)	-	-	음성	2	음성	음성	2.4~2.41	2.41
96	음성	음성	POS (1) NEG (4)	NEG (1)	NEG (1)	음성	3	음성	음성	6.79~7.25	7.12
97	양성 (Plausible)	양성 (Plausible)	POS (3) NEG (1)	-	-	양성	1	적용안됨	X (구조없음)	-	-
98	음성	음성	NEG (1)	NEG (1)	NEG (1)	음성	2	음성	음성	-0.93~1.21	0.05
99	음성	음성	POS (1) NEG (4)	NEG (1)	-	음성	3	음성	음성	1.29	2.83
100	음성	음성	NEG (2)	-	-	음성	3	음성	음성	-0.34~1.87	0.72
101	음성	음성	POS (2) NEG (2)	-	-	음성	3	음성	음성	2.43~2.56	2.48
102	음성	음성	NEG (2)	-	-	음성	3	음성	음성	-0.21~0.19	-0.2
103	**양성 (Plausible)**	**양성 (Plausible)**	**POS (2) NEG (2)**	**POS (1)**	**-**	**양성**	**2**	**양성**	**양성**	**3.97~4.07**	**4.06**
104	**양성 (Plausible)**	**양성 (Plausible)**	**POS (2) NEG (1)**	**-**	**POS (1)**	**양성**	**2**	**양성**	**양성**	**4.99~5.1**	**5.05**

연번	Derek Nexus		Danish QSAR			Vega QSAR		T.E.S.T (EPA)	QSAR Toolbox		
	복귀돌연변이		변이원성			변이원성		변이원성	변이원성	log kow	
	E.Coli	*Salmonella typhimurium*	복귀돌연변이 (X/5)	염색체이상 (Y/2)	소핵 (Z/1)	복귀돌연변이	신뢰도	복귀돌연변이	복귀돌연변이	범위	값 (calcultated)
105	음성	음성	POS (1) NEG (1)	NEG (1)	-	음성	2	음성	음성	1.81~1.87	1.83
106	음성	음성	NEG (2)	POS (1)	-	양성	3	양성	X (out of domain)	2.71~3.4	3.59
107	N/A (contains unclassified features)	N/A (contains unclassified features)	-	-	-	음성	2	음성	음성	1.92~2.09	1.96
108	양성 (Plausible)	양성 (Plausible)	POS (1) NEG (3)	-	-	음성	2	음성	음성	-1.51~6.36	2.09
109	음성	음성	NEG (3)	-	-	음성	3	음성	음성	0.34~0.64	0.50
110	양성 (Plausible)	양성 (Plausible)	-	POS (1) NEG (1)	-	음성	2	음성	음성	5.31~5.43	5.43
111	음성	음성	POS (1) NEG (3)	-	NEG (1)	음성	3	음성	음성	4.71~8.23	7.12
112	양성 (Plausible)	양성 (Plausible)	-	NEG (1)	-	양성	2	음성	X (out of domain)	-1.14~1.54	3.51
113	음성	음성	POS (1) NEG (7)	POS (5) NEG (1)	POS (1)	음성	3	음성	음성	4.64~5.62	5.13
114	음성	음성	POS (3) NEG (11)	POS (1) NEG (6)	NEG (4)	음성	2	음성	음성	0.94~1.1	0.99
115	음성	음성	POS (4)	-	POS (1)	음성	1	음성	음성	2.48~3.15	2.72
116	음성	음성	POS (3) NEG (7)	POS (1)	POS (1) NEG (1)	음성	3	음성	음성	-0.07~0.09	0.01
117	음성	음성	NEG (8)	NEG (3)	-	음성	1	음성	음성	1.51~4.74	2.87
118	음성	음성	POS (1)	-	-	음성	1	적용안됨	음성	7.76~12	11.70
119	음성	음성	POS (1) NEG (2)	POS (1)	NEG (1)	음성	3	양성	음성	-0.03~0.87	0.21

연번	Derek Nexus		Danish QSAR			Vega QSAR		T.E.S.T (EPA)	QSAR Toolbox		
	복귀돌연변이		변이원성			변이원성		변이원성	변이원성	log kow	
	E.Coli	*Salmonella typhimurium*	복귀돌연변이 (X/5)	염색체이상 (Y/2)	소핵 (Z/1)	복귀돌연변이	신뢰도	복귀돌연변이	복귀돌연변이	범위	값 (calcultated)
120	음성	음성	POS (1)	POS (1)	-	음성	2	양성	X (out of domain)	3.28~3.28	0.83
121	양성 (Plausible)	양성 (Plausible)	NEG (1)	-	NEG (1)	양성	exp	양성	양성	-1.66~-1.16	-1.5
122	음성	음성	NEG (3)	NEG (1)	-	음성	2	음성	음성	9.37~10.4	9.99
123	음성	음성	POS (1) NEG (2)	NEG (1)	NEG (1)	음성	3	음성	음성	2.72~3.84	3.15
124	음성	음성	NEG (4)	NEG (1)	-	음성	3	양성	음성	1.91~2.5	2.21
125	양성 (Plausible)	양성 (Plausible)	POS (1) NEG (1)	POS (1)	POS (1)	양성	2	양성	X (out of domain)	5.12~5.17	5.19
126	**양성 (Plausible)**	**양성 (Plausible)**	**POS (3)**	**POS (1)**	**-**	**양성**	**exp**	**양성**	**양성**	**2.8~3.16**	**2.99**
127	음성	음성	-	-	POS (1)	양성	2	양성	X (out of domain)	4.36	4.60
128	양성 (Plausible)	양성 (Plausible)	-	-	POS (1)	양성	2	음성	X (out of domain)	-2.64~-2.24	-3.13
129	음성	음성	POS (1) NEG (1)	NEG (2)	-	음성	3	적용안됨	X (구조없음)	-	-
130	양성 (Plausible)	양성 (Plausible)	POS (1) NEG (1)	NEG (1)	-	음성	1	양성	음성	4.02~4.28	4.15
131	**양성 (Plausible)**	**양성 (Plausible)**	**POS (2) NEG (1)**	**POS (1) NEG (1)**	**NEG (1)**	**양성**	**exp**	**양성**	**양성**	**0.78~0.96**	**0.96**
132	음성	음성	-	-	-	음성	1	음성	음성	2.01~2.62	2.17

※ 음영 표시된 물질은 5개의 예측프로그램에서 복귀돌연변이 예측결과 모두 변이원성(양성)인 물질

부록 2. 발암성 예측결과

연번	Derek Nexus	Danish QSAR	Vega QSAR		공단 MSDS
	Rat	(C/14)	발암성	신뢰도	
1	N/A	POS (1) NEG (8)	X	2	
2	N/A	NEG (9)	O	3	
3	N/A	NEG (9)	△	1	
4	N/A	NEG (13)	△	2	
5	N/A	-	X	2	
6	N/A	-	△	1	
7	N/A	-	X	1	
8	음성 (Impossible)	NEG (7)	O	3	
9	N/A	POS (2) NEG (3)	O	3	
10	양성 (Plausible)	POS (10)	O	1	IARC 3 EU CLP 1B
11	N/A	NEG (10)	△	1	
12	N/A	-	△	1	
13	N/A	NEG (1)	O	2	
14	N/A	POS (2) NEG (4)	O	2	
15	N/A	NEG (7)	△	2	
16	N/A	POS (2) NEG (4)	O	3	IARC 3
17	N/A	POS (9)	O	2	
18	양성 (Plausible)	POS (2) NEG (4)	X	2	
19	N/A	NEG (1)	O	2	
20	N/A	POS (1) NEG (6)	O	3	
21	양성 (Plausible)	NEG (7)	X	2	
22	N/A	POS (1) NEG (6)	O	exp	IARC 3 EU CLP 2
23	N/A	-	O	1	
24	N/A	POS (4) NEG (4)	O	3	
25	N/A	POS (8)	X	3	
26	N/A	NEG (10)	X	2	
27	N/A	NEG (7)	X	1	

연번	Derek Nexus	Danish QSAR	Vega QSAR		공단 MSDS
	Rat	(C/14)	발암성	신뢰도	
28	양성 (Plausible)	POS (5)	O	3	
29	양성 (Equivocal)	POS (2) NEG (5)	X	2	
30	N/A	-	O	2	
31	N/A	NEG (7)	X	3	
32	N/A	NEG (6)	O	1	
33	N/A	NEG (3)	O	1	EU CLP 2
34	N/A	-	X	1	
35	N/A	NEG (7)	X	3	
36	N/A	-	X	1	
37	N/A	NEG (12)	X	2	
38	N/A	POS (1) NEG (10)	X	2	
39	N/A	POS (3)	O	2	
40	N/A	-	X	1	
41	N/A	NEG (3)	X	1	
42	N/A	POS (2) NEG (2)	X	1	
43	N/A	NEG (2)	X	2	
44	N/A	POS (4)	O	2	
45	N/A	NEG (14)	X	3	
46	N/A	POS (1)	X	1	
47	N/A	NEG (3)	O	1	
48	N/A	-	X	1	
49	N/A	POS (1) NEG (3)	X	2	
50	N/A	NEG (7)	X	2	
51	N/A	NEG (2)	X	1	IARC 3
52	양성 (Plausible)	POS (6)	O	1	
53	N/A	NEG (8)	O	1	
54	N/A	POS (5) NEG (7)	X	2	
55	N/A	NEG (8)	O	3	
56	양성 (Equivocal)	POS (2) NEG (3)	X	3	
57	N/A	NEG (13)	O	2	
58	N/A	POS (1)	X	3	

연번	Derek Nexus	Danish QSAR	Vega QSAR		공단 MSDS
	Rat	(C/14)	발암성	신뢰도	
59	N/A	NEG (2)	X	2	
60	N/A	NEG (6)	X	1	
61	양성 (Plausible)	POS (4) NEG (5)	X	3	
62	양성 (Equivocal)	POS (3) NEG (2)	X	2	
63	양성 (Plausible)	POS (3) NEG (2)	X	2	
64	N/A	NEG (12)	O	2	
65	양성 (Plausible)	POS (10)	O	3	
66	N/A	NEG (2)	O	1	
67	양성 (Plausible)	NEG (12)	O	2	
68	양성 (Plausible)	POS (4) NEG (2)	O	1	
69	N/A	NEG (4)	X	1	
70	N/A	NEG (8)	X	2	
71	N/A	NEG (7)	X	2	
72	N/A	POS (2) NEG (3)	X	1	
73	양성 (Plausible)	NEG (9)	O	3	
74	N/A	POS (1) NEG (1)	X	1	
75	N/A	NEG (7)	O	3	
76	N/A	NEG (7)	X	2	
77	양성 (Equivocal)	POS (2) NEG (10)	X	2	
78	N/A	-	O	1	
79	양성 (Equivocal)	POS (2) NEG (7)	X	2	
80	N/A	POS (1) NEG (4)	O	2	
81	N/A	NEG (14)	X	2	
82	N/A	POS (1) NEG (9)	O	2	
83	N/A	-	X	1	
84	양성 (Plausible)	NEG (2)	O	2	IARC 3
85	N/A	-	X	1	

연번	Derek Nexus	Danish QSAR	Vega QSAR		공단 MSDS
	Rat	(C/14)	발암성	신뢰도	
86	양성 (Plausible)	NEG (2)	O	1	
87	양성 (Plausible)	POS (1) NEG (8)	X	2	
88	N/A	POS (5)	O	2	
89	N/A	NEG (9)	X	3	
90	양성 (Plausible)	NEG (7)	O	1	
91	N/A	-	X	1	
92	N/A	-	X	1	
93	N/A	NEG (5)	O	3	
94	N/A	NEG (10)	X	1	
95	N/A	-	O	2	
96	양성 (Plausible)	NEG (10)	X	3	
97	양성 (Plausible)	POS (6)	O	1	
98	N/A	NEG (11)	X	exp	
99	N/A	NEG (7)	O	2	
100	N/A	NEG (8)	X	1	
101	양성 (Plausible)	NEG (6)	O	2	
102	N/A	NEG (11)	X	2	
103	N/A	POS (5) NEG (7)	O	2	
104	양성 (Plausible)	POS (9) NEG (2)	O	3	
105	N/A	NEG (12)	O	2	
106	N/A	NEG (7)	O	2	
107	N/A	-	X	1	
108	양성 (Plausible)	POS (4) NEG (7)	O	2	ACGIH A4
109	N/A	POS (1) NEG (12)	O	exp	
110	양성 (Plausible)	NEG (6)	O	1	
111	N/A	NEG (11)	X	3	
112	양성 (Plausible)	POS (1) NEG (5)	O	2	
113	양성 (Plausible)	POS (2) NEG (3)	X	3	

연번	Derek Nexus	Danish QSAR	Vega QSAR		공단 MSDS
	Rat	(C/14)	발암성	신뢰도	
114	N/A	POS (1) NEG (8)	X	3	
115	N/A	-	X	1	
116	N/A	NEG (14)	X	3	
117	N/A	NEG (7)	O	1	
118	N/A	-	X	1	
119	N/A	POS (5) NEG (7)	X	1	
120	N/A	NEG (6)	O	2	
121	N/A	NEG (7)	O	1	IARC 3
122	N/A	POS (1) NEG (9)	O	3	
123	양성 (Plausible)	NEG (11)	X	2	
124	N/A	NEG (2)	O	2	
125	**양성 (Plausible)**	**POS (12)**	**O**	**3**	**EU CLP 2**
126	양성 (Plausible)	POS (1) NEG (7)	X	exp	IARC 3
127	N/A	POS (9)	O	3	
128	양성 (Plausible)	NEG (6)	O	2	
129	N/A	NEG (10)	X	2	
130	양성 (Plausible)	NEG (6)	O	1	
131	**양성 (Plausible)**	**POS (12)**	**O**	**exp**	
132	N/A	-	X	1	

※ 음영 표시된 물질은 3개의 예측프로그램에서 모두 발암성(양성)인 물질

부록 3. 급성 경구독성 예측결과

연번	Danish QSAR	T.E.S.T (EPA)	QSAR Toolbox			공단 MSDS
	경구(Rat) LD50 mg/kg	경구(Rat) LD50 mg/kg	경구(Rat) LD50 mg/kg	log kow		경구
				범위	값 (calcultated)	
1	1,900	3,350	1,570	3.06~6.92	3.06	
2	180	1,080	X (구조없음)	-	-	LD50 300~2,000 ml/kg 실험종 : Rat
3	920	871	726	1.7~2.38	2.26	
4	3,000	912	4,110	0.23~2.87	0.88	
5	250	308	245	3.18~3.29	3.23	
6	1,600	3,241	2,610	6.62	6.70	
7	480	500	1,880	4.02~4.02	4.02	LD50 316 mg/kg 실험종 : Rat
8	2,200	912	1,370	3.52~3.54	3.53	
9	2,900	826	1,070	1.33~1.94	1.55	
10	160	475	1,350	2.6~2.6	2.60	LD50 300~2,000 mg/kg 실험종 : Rat
11	1,900	440	5,580	6.98~7	6.99	
12	1,700	1,123	1,360	0.78~1.37	1.21	
13	440	289	1,500	0.13~0.13	0.13	
14	320	615	917	-0.82~-0.28	-0.82	
15	1,100	842	X (out of domain)	-0.1~5.62	11.90	LD50 > 5,000 mg/kg 실험종 : Rat
16	2,300	1,013	616	4.02~6.62	4.93	LD50 800 mg/kg 실험종 : Mouse (Rat : LD50 2,700 mg/kg)
17	620	587	1,520	3.71~3.71	3.71	LD50 500 mg/kg 실험종 : Rat
18	3,300	2,652	5,050	3.69~3.73	3.69	
19	180	290	628	1.7~1.8	1.77	LDL0 500 mg/kg 실험종 : Rat
20	2,600	658	X (구조없음)	-	-	
21	3,800	1,080	1,120	4.41~4.43	4.42	
22	1,300	584	808	2~2.15	2.03	LD50 1,053 mg/kg 실험종 : Rat
23	1,100	1,496	1,120	-1.71~-1.66	-1.68	
24	970	3,305	360	5.19~5.27	5.24	
25	820	165	X (구조없음)	-	-	LD50 300~2,000 mg/kg
26	5,500	12,083	1,950	-2.01~-0.32	-1.23	LD50 > 2,000 mg/kg

연번	Danish QSAR	T.E.S.T (EPA)	QSAR Toolbox			공단 MSDS
	경구(Rat) LD50 mg/kg	경구(Rat) LD50 mg/kg	경구(Rat) LD50 mg/kg	log kow		경구
				범위	값 (calcultated)	
27	3,400	2,908	2,850	5.71~5.73	5.72	
28	1,800	1,665	2,120	5.18~5.49	5.36	
29	1,200	772	1,490	0.29~0.94	0.57	
30	2,000	1,905	1,720	3.16~3.2	3.19	LD50 > 5,000 mg/kg 실험종 : Rat
31	820	559	757	1.03~1.39	1.15	LD50 1,300 mg/kg 실험종 : Rat
32	720	850	X (구조없음)	-	-	
33	1,500	207	435	1.82~4.49	4.49	LD50 251 mg/kg 실험종 : Rat
34	1,900	3,512	1,560	8.02~8.06	8.04	
35	1,900	1,075	753	0.04~0.27	0.18	
36	1,100	3,201	1,940	9.69~10	9.85	
37	9,300	4,673	1,390	6.42~10.1	10.10	
38	7,700	4,561	1,040	2.21~4.5	3.57	
39	1,500	707	295	5.23~5.41	5.34	LD50 > 2,000 mg/kg 실험종 : Rat
40	1,100	1,980	3,260	1.37~1.49	1.43	
41	2,100	2,043	2,470	4.57~4.74	4.65	LD50 > 2,000 mg/kg
42	1,900	1,042	1,410	4.9~4.98	4.91	
43	390	704	2,300	2.09~2.29	2.11	
44	780	918	X (구조없음)	-	-	
45	1,800	1,665	4,440	3.14~4.28	3.94	
46	220	331	1,050	2.31~2.59	2.41	
47	240	1,029	X (구조없음)	-	-	
48	1,000	1,076	5,380	8.3~8.48	8.39	
49	2,500	1,185	5,880	3.76~4.32	4.04	
50	1,400	1,941	3,200	1.84~1.84	1.84	
51	1,300	1,357	1,850	2.65~5.6	4.27	
52	4,100	11,122	5,230	1.72~2.02	1.87	
53	1,200	69	15,400	5.69~5.85	5.77	LD50 > 2,000 mg/kg 실험종 : Rat
54	2.3	3,736	3,760	0.57~0.73	0.64	LD50 1,498 mg/kg 실험종 : Rat
55	2.3	2,446	337	0.39~1.15	0.43	

연번	Danish QSAR	T.E.S.T (EPA)	QSAR Toolbox			공단 MSDS
	경구(Rat) LD50 mg/kg	경구(Rat) LD50 mg/kg	경구(Rat) LD50 mg/kg	log kow		경구
				범위	값 (calcultated)	
56	2,000	5,474	5,250	2.86~3.11	3.00	
57	5,400	2,066	9,410	1.72~2.29	2.01	
58	160	600	1,050	1.29~2.21	1.76	
59	740	1,864	3,530	-1.43~-0.91	-1.18	
60	1,800	1,919	816	4.58~4.64	4.60	LD50 2,509 mg/kg 실험종 : Mouse
61	2,300	N/A	1,070	3.67~5.15	4.61	
62	4,800	3,191	1,490	0.91~1.22	1.08	
63	12,000	1,752	1,750	0.27~0.48	0.38	
64	6,400	8,202	3,190	0.31~2.43	0.94	LD50 > 2,000 mg/kg 실험종 : Rat
65	3,200	3,471	6,660	2.16~4.35	3.26	
66	2,100	511	1,020	2.39~2.73	2.56	
67	5,900	3,811	X (구조없음)	-	-	LD50 > 2,000 mg/kg 실험종 : Rat
68	2,100	4,085	421	6.16~6.3	6.22	
69	410	104	796	-0.75~0.28	-0.34	
70	200,000	2,498	445	-0.98~4.37	0.19	
71	6,500	5,508	1,130	0.08~2.1	0.57	
72	670	2,117	5,300	3.51~3.58	3.55	
73	1,500	1,636	935	0.97~1.68	1.36	
74	2,300	1,085	382	0.6~0.62	0.61	
75	2,800	3,852	2,320	2.72~2.99	2.85	LD50 11,300 mg/kg 실험종 : Rat
76	6,200	4,299	1,170	0.04~1.65	0.81	
77	1,300	542	552	0.27~0.73	0.50	
78	780	1,544	7,290	-1.12~-1.18	-1.19	
79	2,000	8,635	5,500	3.67~5.15	4.58	
80	1,300	2,368	1,800	0.11~0.17	0.14	
81	2,600	1,940	2,170	-0.91~-0.63	-0.76	
82	540	719	1,060	1.63~3.28	2.37	
83	1,600	1,425	975	2.61~2.61	2.61	
84	540	231	250	0.78~1.63	1.21	LD50 250 mg/kg 실험종 : Rat
85	6,900	N/A	X (out of domain)	11.3~15	16.50	LD50 4,640 mg/kg 실험종 : Rat
86	820	516	495	3.45~3.49	3.47	

연번	Danish QSAR	T.E.S.T (EPA)	QSAR Toolbox			공단 MSDS
	경구(Rat) LD50 mg/kg	경구(Rat) LD50 mg/kg	경구(Rat) LD50 mg/kg	log kow		경구
				범위	값 (calcultated)	
87	2,500	3,049	X (구조없음)	-	-	
88	1,100	1,122	2,190	3.81~4.06	4.06	
89	1,800	2,547	1,300	0.48~0.64	0.56	
90	1,600	638	7,960	3.98~4.07	4.05	
91	1,400	3,698	X (out of domain)	10.2~10.4	10.40	
92	1,700	1,966	X (out of domain)	9.36~10.7	11.00	LD50 > 5,000 mg/kg
93	800	3,140	5,670	2.19~7.82	5.08	LD50 50~300 mg/kg 실험종 : Rabbit
94	2,400	2,904	2,260	3.9~3.94	3.93	
95	130	937	X (out of domain)	3.65~3.65	2.41	
96	2,200	1,267	2,810	6.92~7.23	7.12	
97	4,300	1,739	X (구조없음)	-	-	
98	3,000	1,730	2,888	-1.14~0.85	0.05	
99	2,800	418	2,888	2.16~2.2	2.18	LD50 300~2,000 mg/kg 실험종 : Rat
100	6,700	3,902	1,290	0.33~1.12	0.72	
101	1,300	657	3,570	1.81~3.12	2.48	
102	1,100	1,104	2,180	-0.28~0.2	-0.2	
103	2,300	1,710	4,350	3.94~4.18	4.06	
104	2,200	14,618	3,400	4.71~5.45	5.05	
105	700	1,677	X (out of domain)	2.05~2.24	1.83	
106	740	1,777	126	3.56~3.63	3.59	
107	330	304	326	1.79~2.1	1.96	
108	4,300	336	X (out of domain)	3.38~4.72	2.09	
109	2,400	3,660	1,010	0.47~0.53	0.50	
110	1,100	1,762	1,200	5.27~5.62	5.43	
111	3,300	12,178	X (out of domain)	2.84~6.6	7.12	LD50 > 2,000 mg/kg
112	340	1,989	4,690	3.5~3.52	3.51	
113	2,500	321	1,130	4.09~6.13	5.13	
114	3,700	3,929	3,380	0.77~1.27	0.99	
115	1,300	3,930	3,330	2.65~2.8	2.72	

연번	Danish QSAR	T.E.S.T (EPA)	QSAR Toolbox			공단 MSDS
	경구(Rat) LD50 mg/kg	경구(Rat) LD50 mg/kg	경구(Rat) LD50 mg/kg	log kow		경구
				범위	값 (calcultated)	
116	2,200	2,466	1,960	-0.02~0.04	0.01	
117	400	2,840	9,390	1.51~3.93	2.87	
118	9,100	185	8,490	10.9~12.5	11.70	
119	2,800	3,957	3,960	0.15~0.26	0.21	
120	4,100	2,841	1,360	-0.5~1.9	0.83	
121	1,100	1,134	2,780	-1.86~-0.93	-1.5	LD50 531 mg/kg 실험종 : Rat
122	6,500	2,657	10,100	9.34~10.4	9.99	
123	2,500	581	3,410	3.09~3.32	3.15	
124	1,500	1,388	X (out of domain)	2.33~3.18	2.21	LD50 660 mg/kg 실험종 : Rat
125	4,500	2,939	1,120	5.16~5.22	5.19	
126	1,700	1,245	2,160	2.99~8.86	2.99	LD50 150 mg/kg 실험종 : Rat
127	580	676	3,110	4.55~4.61	4.60	
128	180	3,447	5,320	-3.56~-2.42	-3.13	
129	3,200	7,568	X (구조없음)	-	-	
130	1,800	3,657	9,250	3.88~4.36	4.15	
131	580	821	1,010	0.78~1.11	0.96	LD50 681 mg/kg 실험종 : Rat
132	1,400	1,463	2,030	2.11~2.24	2.17	

※ 음영 표시된 물질은 3개의 예측프로그램에서 모두 급성 경구독성(GHS 기준: 구분1~4)인 물질

부록 4. 대상물질의 용도, 작용기 및 Toxtree 예측결과

연번	사 용 용 도	고분자 여부	작용기 (functional group)	Toxtree (Toxic Hazard-Cramer rules)
1	반도체 : 바니쉬	X	aryl	Class 3
2	전자 : 소재 원료	X	amine	Class 3
3	전자 : 도공용액	X	sulfonyl halide	Class 3
4	전자 : 증감제	X	amine	Class 3
5	반도체 : 소재 제조용 중간체	X	alkyl arenes	Class 3
6	반도체 : 코팅제	X	siloxane	Class 3
7	반도체 : 소재 원료	X	aryl	Class 3
8	OLED : 소재 중간체	X	aryl	Class 3
9	반도체 : 포토레지스트 원료	X	alcohol	Class 3
10	전자 : 첨가제 원료	X	alkene alkylhalide	Class 3
11	반도체 : 감광제	X	aryl	Class 3
12	반도체 : 소재 모노머	X	alkene	Class 3
13	전자 : 첨가제	X	alkene	Class 3
14	반도체 : 포토레지스트 원료	X	amine	Class 3
15	반도체 : 감도조절제	X	aryl	Class 3
16	반도체 : 소재 원료	X	aryl	Class 3
17	전자 : 소재 원료	O	aryl	Class 3
18	반도체 : 코팅제	X	aryl ketone	Class 3
19	OLED : 소재 원료	X	amine	Class 3
20	전자 : 소재 원료	X	alkyl halide	Class 3
21	OLED : 소재 중간체	X	aryl	Class 3
22	전자 : 첨가제	X	aryl	Class 3
23	반도체 : 산화막 형성용	X	amine	Class 3
24	OLED : 소재 제조용 중간체	X	anthracene	Class 3
25	OLED : 소재 중간체	X	aryl	Class 3
26	LCD : 표면개질제		carboxylic acid	Class 3
27	전자 : 소재 원료	X	aryl	Class 3
28	반도체 : 봉지재	X	alkyl halide	Class 3
29	전자 : 소재 단량체	X	alkene ether	Class 1
30	반도체 : 포토레지스트	X	alkene	Class 3
31	전자 : 소재 원료	X	amine	Class 3
32	반도체 : 포토레지스트 원료	X	phenyl cyanide	Class 3
33	반도체 : 소재 원료	X	alkyl halide	Class 3
34	반도체 : 봉지재 원료	X	aryl	Class 3
35	전자 : 표면처리제	X	aryl	Class 3

연번	사 용 용 도	고분자 여부	작용기 (functional group)	Toxtree (Toxic Hazard-Cramer rules)
36	반도체 : 봉지재 원료	X	alkene	Class 3
37	전자 : 소재 원료	X	amine	Class 1
38	반도체 : 접착제 원료	X	alkane	Class 1
39	OLED : 소재 중간체	X	aryl	Class 3
40	반도체 : 세정제	X	silane	Class 3
41	OLED : 소재 원료	X	aryl	Class 3
42	반도체 : 광중합개시제	X	aryl	Class 3
43	반도체 : 소재 중간체	X	alkyl	Class 3
44	OLED : 소재 중간체	X	aryl	Class 3
45	반도체 : 광개시제	X	alkane	Class 3
46	반도체 : 첨가제	X	aryl	Class 3
47	전자 : 코팅제(절연제)	O	phenyl ester	Class 3
48	전자 : 밀봉재	X	alkene	Class 3
49	전자 : 첨가제	X	alkene	Class 2
50	LCD : 소재 원료	X	alcohol	Class 3
51	반도체 : 소재 합성용 원료	X	alkyl halide	Class 3
52	전자 : 첨가제	X	ether	Class 3
53	반도체 : 냉각 및 세정제	X	alkane	Class 3
54	반도체 : 세정제	X	alcohol	Class 3
55	반도체 : 첨가제	X	alcohol	Class 3
56	전자 : 점착제	X	alkene	Class 1
57	전자 : 소재 단량체	X	alkene	Class 1
58	전자 : 소재 합성용 중간체	X	aryl	Class 3
59	전자 : 첨가제	X	ether	Class 3
60	OLED : 소재 원료	X	alkyl	Class 3
61	전자 : 소재 단량체	X	alkene	Class 3
62	반도체 : 드라이필름	X	alkene aryl alcohol	Class 1
63	반도체 : 코팅 및 보호제	X	alcohol	Class 3
64	전자 : 코팅용	X	alkene ether	Class 1
65	반도체 : 접착제	X	aryl	Class 3
66	반도체 : 첨가제	X	alkyl halide	Class 3
67	전자 : 레지스트 원료	O	phenyl	Class 3
68	전자 : 레지스트 원료	X	alkene	Class 3
69	전자 : 소재 합성용 중간체	X	ketone	Class 3
70	전자 : 소재 원료	X	cycloalkane oxolane	Class 3
71	반도체 : 세정제	X	carbonate dioxolane	Class 3
72	반도체 : 소재 원료	X	aryl	Class 3

연번	사 용 용 도	고분자 여부	작용기 (functional group)	Toxtree (Toxic Hazard-Cramer rules)
73	반도체 : 촉진제	X	aryl	Class 3
74	반도체 : 연마액 원료	X	aryl	Class 3
75	전자 : 소재 합성용 원료	X	alkene	Class 3
76	전자 : 소재 원료	X	dioxane	Class 1
77	반도체 : 코팅제	X	alkene carboxylicacid	Class 1
78	OLED : 소재 원료	X	imide pyrrolidine	Class 3
79	전자 : 점도 증대용	X	alkene carboxylicacid	Class 1
80	LCD : 소재 원료	X	aryl	Class 3
81	LCD : 박리 용매	X	ether	Class 3
82	OLED : 소재 원료	X	amine aryl	Class 3
83	반도체 : 산화막 형성용	X	silane	Class 3
84	반도체 : 소재 원료	X	alkyl halide	Class 3
85	전자 : 칩 보호제	X	aryl	Class 3
86	반도체 : 드라이에칭용	X	alkyl halide	Class 3
87	반도체 : 납땜용	X	ester	Class 1
88	OLED : 소재 원료	X	aryl	Class 3
89	반도체 : 포토레지스트 원료	X	alcohol	Class 1
90	OLED : 소재 합성용 중간체	X	aryl	Class 3
91	반도체 : 봉지재 원료	X	alkene	Class 3
92	전자 : 희석제	X	alkene	Class 3
93	반도체 : 증감제	X	aryl	Class 3
94	반도체 : 포토레지스트	X	aryl phenol	Class 3
95	전자 : 소재 합성용 원료	X	aryl	Class 3
96	반도체 : 모노머 원료	X	aryl	Class 3
97	반도체 : 봉지재	X	siloxane	Class 3
98	LCD : 세정제	X	carboxylic acid	Class 3
99	반도체 : 소재 합성용 원료	X	aryl	Class 3
100	전자 : 전해액	X	ester	Class 1
101	전자 : 부식방지제	X	aryl	Class 3
102	반도체 : 세정제	X	carnonic acid derivative, urea	Class 3
103	반도체 : 고분자 원료	X	naphthalene aryl	Class 3
104	전자 : 소재 원료	X	aryl ether	Class 3
105	반도체 : 광개시제	X	aryl ketone	Class 3
106	반도체 : 방청제	X	aryl	Class 3
107	LCD : 첨가제	X	aryl	Class 3
108	반도체 : 에칭용	X	alkenyl halide	Class 3
109	반도체 : 분해제	X	carboxylic acid	Class 3

연번	사 용 용 도	고분자 여부	작용기 (functional group)	Toxtree (Toxic Hazard-Cramer rules)
110	LCD : 광개시제	X	alkyl halide	Class 3
111	OLED : 소재 원료	X	alkene	Class 1
112	반도체 : 포토레지스트 원료	X	alkene ketone	Class 3
113	전자 : 소재 원료	X	alkene ether	Class 3
114	전자 : 희석제	X	ether	Class 1
115	반도체 : 세정제	X	silane	Class 3
116	LCD : 세척제	X	carboxylic acid	Class 3
117	반도체 : 냉각 또는 세정용	X	alkyl halide	Class 3
118	전자 : 봉지재 원료	X	phenyl	Class 3
119	반도체 : 소재 원료	X	alkane ether	Class 3
120	반도체 : 포토레지스트 원료	X	aryl	Class 3
121	전자 : 경화제	X	ketone	Class 1
122	전자 : 소재 원료	X	aryl biphenyl	Class 3
123	반도체 : 소재 원료	X	carboxylic acid	Class 3
124	반도체 : 포토레지스트 원료	X	aryl naphtalene phenol	Class 3
125	반도체 : 소재 원료	X	alkyl ether	Class 3
126	OLED : 소재 원료	X	aryl naphtalene	Class 3
127	OLED : 소재 원료	X	aryl	Class 3
128	전자 : 접착제	X	alkane	Class 3
129	반도체 : 소재 원료	O	alkyl	Class 1
130	전자 : 반응개시제	X	alkyl halide	Class 3
131	반도체 : 납땜용	X	alcohol	Class 3
132	반도체 : 증착제	X	silane	Class 3

〈〈연 구 진〉〉

연 구 기 관 : 안전보건공단 산업안전보건연구원

연구책임자 : 김 현 옥(부 장, 산업화학연구실)
연 구 원 : 임 경 택(부 장, 산업화학연구실)
연 구 원 : 이 인 섭(실 장, 산업화학연구실)

〈〈연 구 기 간〉〉

2019. 1. 1 ~ 2019. 11. 30

본 연구보고서의 내용은 연구책임자의 개인적 견해이며, 우리 연구원의 공식견해와 다를 수도 있음을 알려드립니다.

산업안전보건연구원장

급성독성 화학물질의 유해성 예측프로그램 적용연구

2019-연구원-1615

발 행 일 : 2019년 11월 30일

발 행 인 : 산업안전보건연구원 원장 고재철

연구책임자 : 산업화학연구실 유해성연구부 부장 김현옥

발 행 처 : 안전보건공단 산업안전보건연구원

주 소 : (44429) 울산광역시 중구 종가로 400

전 화 : (042) 869-0321

F A X : (042) 863-9001

Homepage : http://oshri.kosha.or.kr