ФЕДЕРАЛЬНОЕ АГЕНТСТВО

ПО ТЕХНИЧЕСКОМУ РЕГУЛИРОВАНИЮ И МЕТРОЛОГИИ

РЕКОМЕНДАЦИИ ПО СТАНДАРТИЗАЦИИ

Р 1323565.1.027— 2019

Руководство по группировке схожих химических веществ в токсикологически значимые категории для устранения пробелов в информации о токсичности при помощи программного обеспечения ОЭСР QSAR Toolbox

Издание официальное

Москва
Стандартинформ
2019

gost.e-ecolog.ru

Предисловие

1 РАЗРАБОТАНЫ Ассоциацией «Некоммерческое партнерство Координационно-информационный центр государств — участников СНГ по сближению регуляторных практик» (Ассоциация «НП КИЦ СНГ»)

2 ВНЕСЕНЫ Техническим комитетом по стандартизации ТК 339 «Безопасность сырья, материалов и веществ»

3 УТВЕРЖДЕНЫ И ВВЕДЕНЫ В ДЕЙСТВИЕ Приказом Федерального агентства по техническому регулированию и метрологии от 28 ноября 2019 г. № 1278-ст

4 ВВЕДЕНЫ ВПЕРВЫЕ

*Правила применения настоящих рекомендаций установлены в статье 26 Федерального закона от 29 июня 2015 г. № 162-ФЗ «О стандартизации в Российской Федерации». Информация об изменениях к настоящим рекомендациям публикуется в ежегодном (по состоянию на 1 января текущего года) информационном указателе «Национальные стандарты», а официальный текст изменений и поправок — в ежемесячном информационном указателе «Национальные стандарты». В случае пересмотра (замены) или отмены настоящих рекомендаций соответствующее уведомление будет опубликовано в ближайшем выпуске ежемесячного информационного указателя «Национальные стандарты». Соответствующая информация, уведомление и тексты размещаются также в информационной системе общего пользования — на официальном сайте Федерального агентства по техническому регулированию и метрологии в сети Интернет (*gost.e-ecolog.ru

© Стандартинформ, оформление, 2019

Настоящие рекомендации не могут быть полностью или частично воспроизведены, тиражированы и распространены в качестве официального издания без разрешения Федерального агентства по техническому регулированию и метрологии

Содержание

1 Область применения 1
2 Термины и определения. 1
3 Общие положения 2
4 Модуль «Ввод данных» («Input»). 2
5 Модуль «Профилирование» («Profiling») 4
6 Модуль «Данные» («Data») 8
7 Модуль «Определение категории» («Category Definition») 16
8 Модуль «Сбор недостающих данных» («Data Gap Filling»). 23
9 Модуль «Отчет» («Report») 28
10 Доступ к программному обеспечению 29
Библиография 30

Введение

В настоящее время отмечается, что анализ и выявление опасных свойств новых химических веществ широко востребованы в различных отраслях промышленности. Лабораторные исследования (испытания) новых химических веществ достаточно дороги, и, принимая во внимание интенсификацию химической промышленности и прогнозы специалистов о синтезе свыше 300 миллионов новых химических веществ только в течение последующих 30 лет, проблема поиска доступных и быстрых методов по выявлению свойств неизученных химических веществ как никогда актуальна.

Для устранения пробелов в информации об опасных свойствах новых химических веществ, в частности о токсичности, целесообразно использование альтернативного лабораторным испытаниям аналогового подхода Quantitative Structure-Activity Relationship (QSAR). С помощью данного подхода возможно построение моделей, позволяющих по структурам химических соединений предсказывать их свойства.

Организацией экономического сотрудничества и развития (ОЭСР) разработана программа QSAR Toolbox, которая базируется на поиске количественных соотношений структура — свойство посредством методов математической статистики, позволяющих по описанию структур химических веществ предсказывать их свойства (в том числе по воздействию на организм). При этом программа QSAR Toolbox способна прогнозировать свойства как на качественном уровне (например, будет ли химическое вещество обладать тем или иным видом воздействия на организм), так и на количественном уровне (прогнозирование численных значений свойств).

Настоящее руководство знакомит пользователя с возможностями программного обеспечения ОЭСР QSAR Toolbox, которое может быть использовано для устранения пробелов в информации о токсичности химических веществ и решения вопроса о целесообразности проведения исследований in vitro и/или in vivo, что в свою очередь, например, может быть использовано при формировании отчета о химической безопасности для новых химических веществ в рамках процедуры нотификации новых химических веществ согласно техническому регламенту [1].

РЕКОМЕНДАЦИИ ПО СТАНДАРТИЗАЦИИ

Руководство по группировке схожих химических веществ в токсикологически значимые категории для устранения пробелов в информации о токсичности при помощи программного обеспечения ОЭСР QSAR Toolbox

Guideline on grouping of the similar chemicals into toxicologically relevant categories to fill in data gaps regarding to the toxicological information by using the OECD QSAR Toolbox software

Дата введения — 2020—05—01

1 Область применения

Настоящие рекомендации по стандартизации разработаны на основании англоязычного руководства ОЭСР по использованию функциональных возможностей программы QSAR Toolbox*.

Данное руководство направлено на работу с программой ОЭСР QSAR Toolbox, позволяющей группировать химические вещества в токсикологически значимые категории для устранения пробелов в информации о токсичности. Основная цель программы — предоставить возможность использовать методологию QSAR для группировки химических веществ по категориям и заполнения пробелов в информационных данных, используя анализ тенденций, метод аналогий и возможности программы QSAR Toolbox, базирующиеся на поиске количественных соотношений структура — свойство. Справочная информация о концепции химических категорий представлена в руководстве по группировке химических веществ.

Данное руководство помогает пошагово изучить основные функции панели инструментов программы QSAR Toolbox с помощью примеров. В дополнение ОЭСР рекомендует просматривать дополнительные материалы по изучению функций QSAR Toolbox**.

2 Термины и определения

В настоящих рекомендациях применены следующие термины с соответствующими определениями:

2.1 **аналоги:** Химические соединения одного структурного типа, отличающиеся по составу на определенное число структурных единиц.

2.2 **каноническая SMILES (Canonical SMILES):** Версия спецификации, включающая правила канонизации, позволяющие записать формулу молекулы любого вещества однозначным образом.

Примечание — Эти правила касаются выбора первого атома в записи, направления обхода циклов, выбора направления основной цепи при разветвлениях. Поскольку в разных пакетах молекулярного моделирования используются различные алгоритмы канонизации SMILES, вследствие чего могут получаться разные записи одной и той же молекулы, понятие «каноническая SMILES» не является абсолютным. Данная версия стандарта обычно применяется для индексирования и проверки уникальности молекул в базах данных.

---

* QSAR Toolbox Application manual gost.e-ecolog.ru manual_F1.compressed.pdf.

** Пользователю рекомендуется просматривать дополнительные материалы по изучению функций QSAR Toolbox, такие как: gost.e-ecolog.ru relationshipsprojectqsars.htm, gost.e-ecolog.ru

Издание официальное

2.3 **изомерическая SMILES (Isomeric SMILES):** Версия спецификации, позволяющая включать в запись данные об изотопном составе, конфигурации асимметрических атомов углерода и двойных связей.

Примечание — Особенность этой версии по сравнению с официальной номенклатурой ИЮПАК в том, что изомерическая SMILES позволяет хранить информацию о молекулах, для которых известны конфигурации лишь некоторых хиральных центров или двойных связей.

2.4

коэффициент распределения н-октанол/вода (K_{OW}): Соотношение равновесных концентраций вещества, растворенного в двухфазной системе, состоящей из двух практически не смешивающихся растворителей.

[ГОСТ 32291—2013, статья 2.1]

2.5 **dice:** Четыре фактора, определяющие оценку вероятности успеха проекта, основанного на объективных измерениях:

D — Продолжительность (Duration),

I — Ответственность (Integrity),

C — Приверженность (Commitment),

E — Усилие (Effort).

Примечание — Каждому фактору присваивается число от 1 (благоприятный) до 4 (маловероятный).

3 Общие положения

Программа QSAR Toolbox имеет несколько функций, позволяющих выполнять такие операции, как:

- идентифицировать аналоги для химического вещества, получать экспериментальные результаты, доступные для этих аналогов, и восполнять недостаток в данных, используя метод аналогий или анализ тенденций;
- классифицировать широкий спектр химических веществ, используя соответствующий инструмент или воспроизводя необходимую операцию;
- заполнять пробелы в данных для любого химического вещества, используя хранилище моделей QSAR;
- оценивать корректность потенциального аналога для метода аналогий;
- оценивать пригодность (точность) модели QSAR при заполнении пробелов в данных для конкретного химического вещества;
- строить модели QSAR.

Данное руководство распространяется на первую операцию, то есть выбор возможных аналогов для химического вещества: извлечение экспериментальных результатов, доступных для этих аналогов, и восполнение недостатка в данных, используя метод аналогий или анализ тенденций.

Рассмотренный в руководстве пример направлен на восполнение информационных пробелов по вопросу сенсибилизации кожи к воздействию 4-нитробензоилхлорида (CAS № 122-04-3).

4 Модуль «Ввод данных» («Input»)

При запуске программы QSAR Toolbox открывается панель инструментов. Для начала работы необходимо нажать кнопку «Ввод данных» («Input») (рисунок 1).

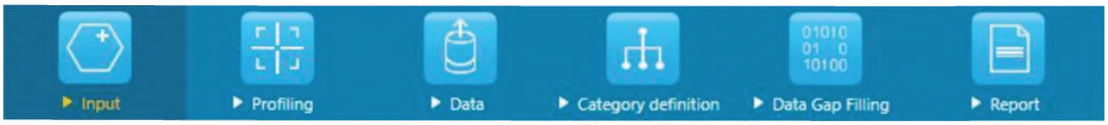
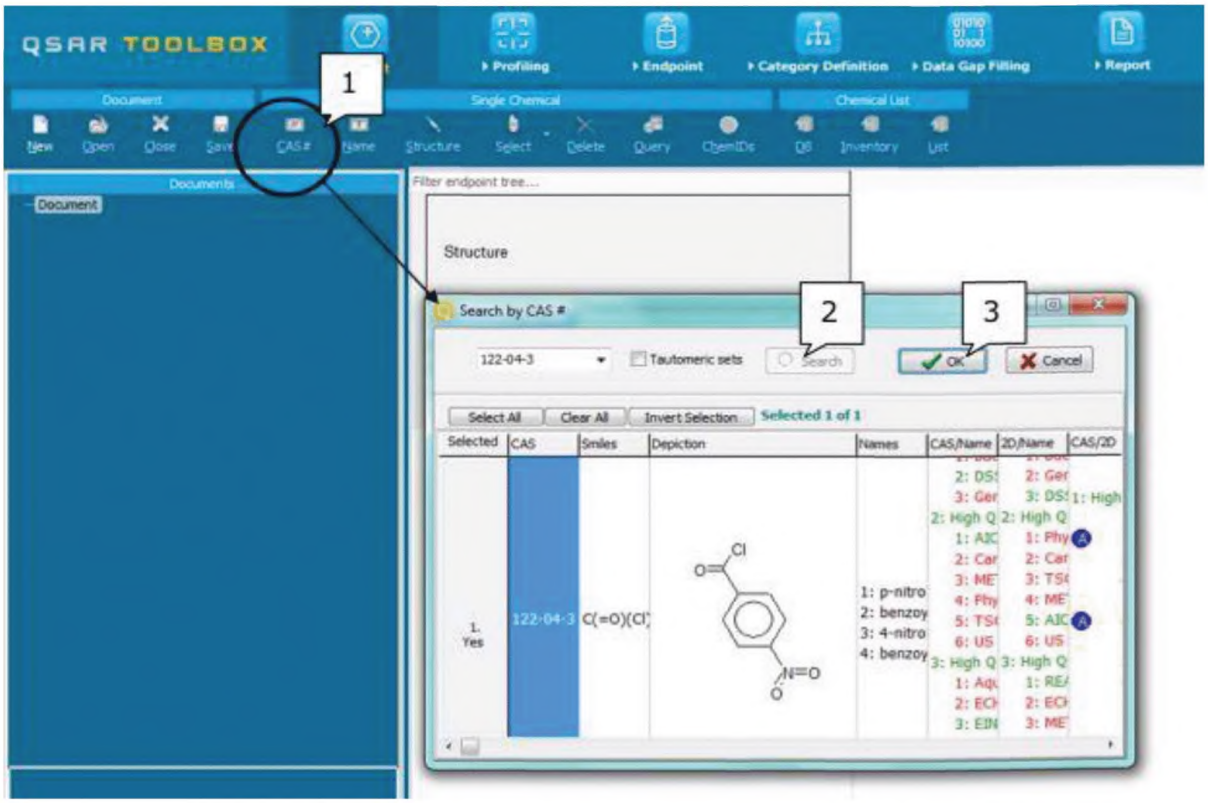


Рисунок 1 — Панель инструментов с выделенным модулем «Ввод данных»

Модуль «Ввод данных» предназначен для выбора необходимого химического вещества, с которым предстоит дальнейшая работа при помощи функциональных возможностей инструментария программы. Идентификация химического вещества может быть проведена по названию, номеру CAS, путем зарисовки структурной формулы, путем использования SMILES (спецификация упрощенного представления молекул в строке ввода) или путем выбора необходимого химического соединения из списка, базы данных, реестра или файла.

Пример — В начале работы необходимо нажать на кнопку «CAS #» 1, ввести номер CAS 4-нитробензоилхлорида (CAS № 122-04-3), нажать на кнопку «Search» («Поиск») 2 и после появления правильной химической структуры нажать на кнопку «OK» («Принять») 3 во всплывающем окне (рисунок 2).



***1* — кнопка «CAS #»; *2* — кнопка «Search» («Поиск»); *3* — кнопка «OK» («Принять»)**

Рисунок 2 — Пример работы в модуле «Ввод данных»

Если структура вещества имеет несколько номеров CAS или структура согласуется с несколькими веществами (например, в случае сложных химических соединений), происходит идентификация некоторого количества химических структур. При этом необходимо выбрать одно вещество для последующей обработки данных.

Затем выбранное вещество отображается в матрице данных, после чего необходимо перейти к следующему модулю (рисунок 3).

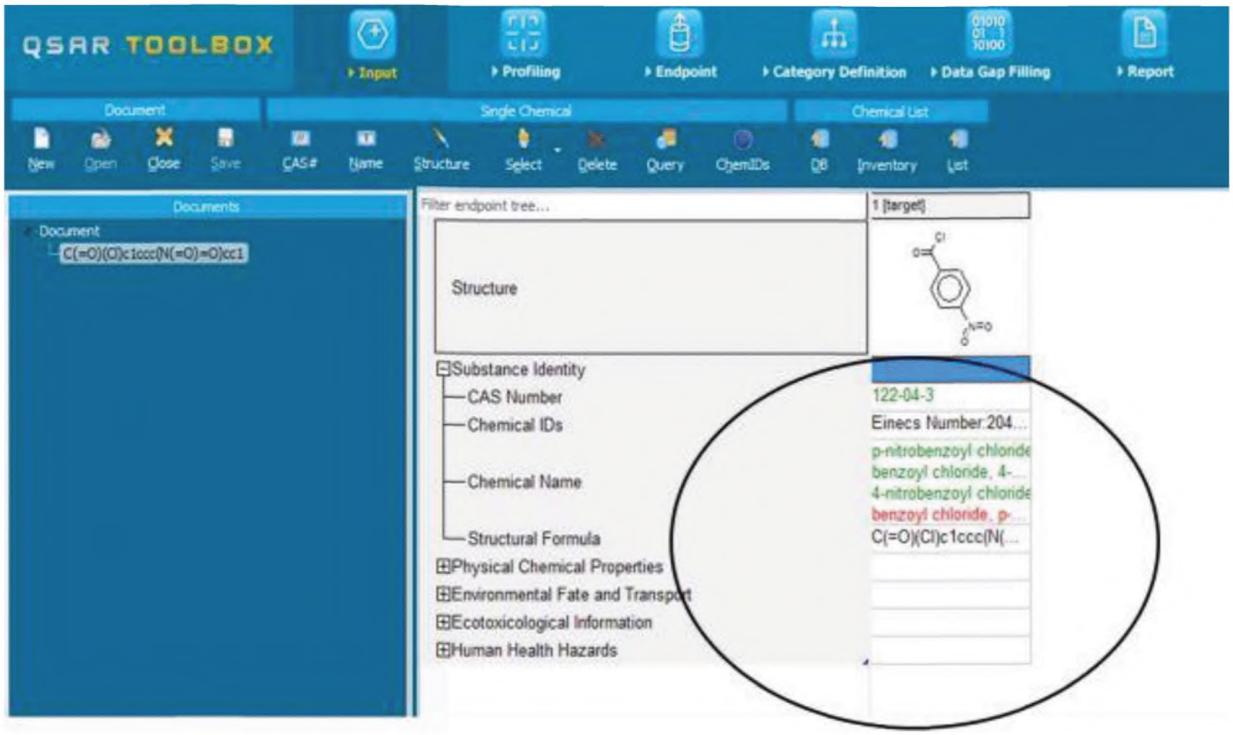


Рисунок 3 — Добавление идентифицированного вещества в матрицу данных

Цветовой код указывает на достоверность идентифицированного химического вещества:

- зеленый — между веществом и выбранной структурой существует высокая степень родства. Этот цвет присваивается, если структура вещества подтверждается в нескольких базах данных;

- желтый — между веществом и выбранной структурой существует средняя степень родства. Цвет присваивается, если структура вещества подтверждается в нескольких базах данных, однако достоверность баз данных не может быть установлена;

- красный — между веществом и выбранной структурой существует плохая степень родства. Цвет присваивается, если вещество соответствует разным структурам в отличных друг от друга базах данных.

5 Модуль «Профилирование» («Profiling»)

После отображения выбранного вещества в матрице данных необходимо перейти к следующему модулю «Профилирование», нажав на кнопку «Профилирование» («Profiling») на верхней панели инструментов (рисунок 4).

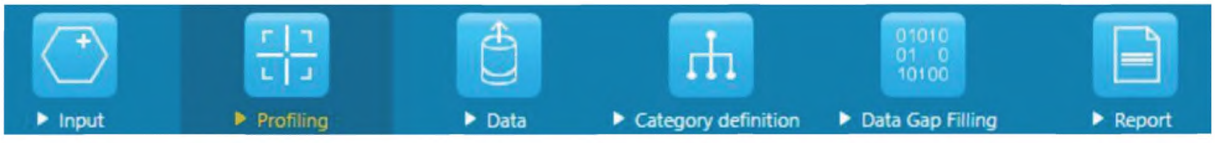
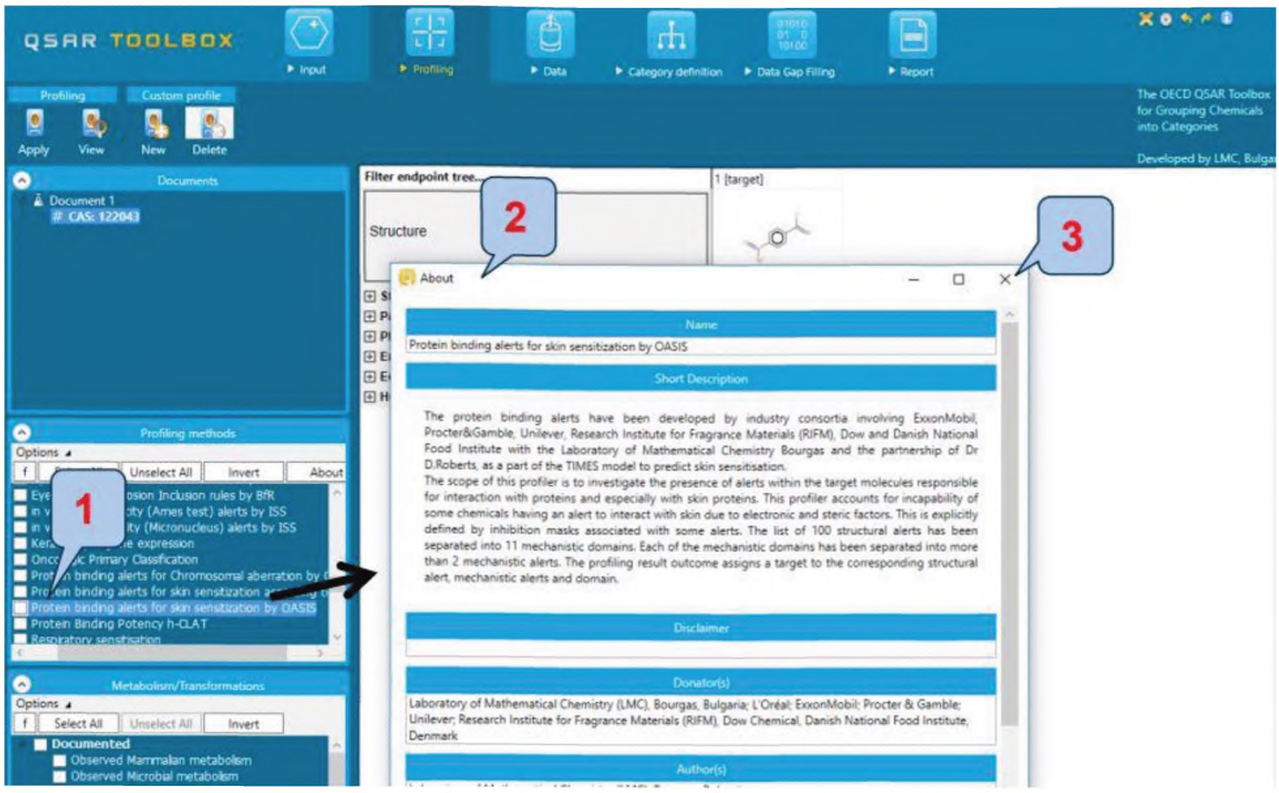


Рисунок 4 — Панель инструментов с выделенным модулем «Профилирование»

Модуль «Профилирование» позволяет получить информацию путем осуществления выборки из массива данных по определенному критерию на основе идентификации вещества или его структуры. Панель инструментов программы в настоящее время содержит большой список «Профилей», которые идентифицируют принадлежность к ранее определенным категориям (например, категории, опреде-

ленные US-EPA* для оценки новых химических веществ, или категории, оцененные в рамках программы ОЭСР по химическим веществам с высоким объемом производства), а также механизмам или принципам действий. Сводную информацию можно получить, выбрав название профиля *1*, затем щелкнуть по нему правой кнопкой мыши и нажать кнопку «About» («Узнать больше») 2 (рисунок 5).

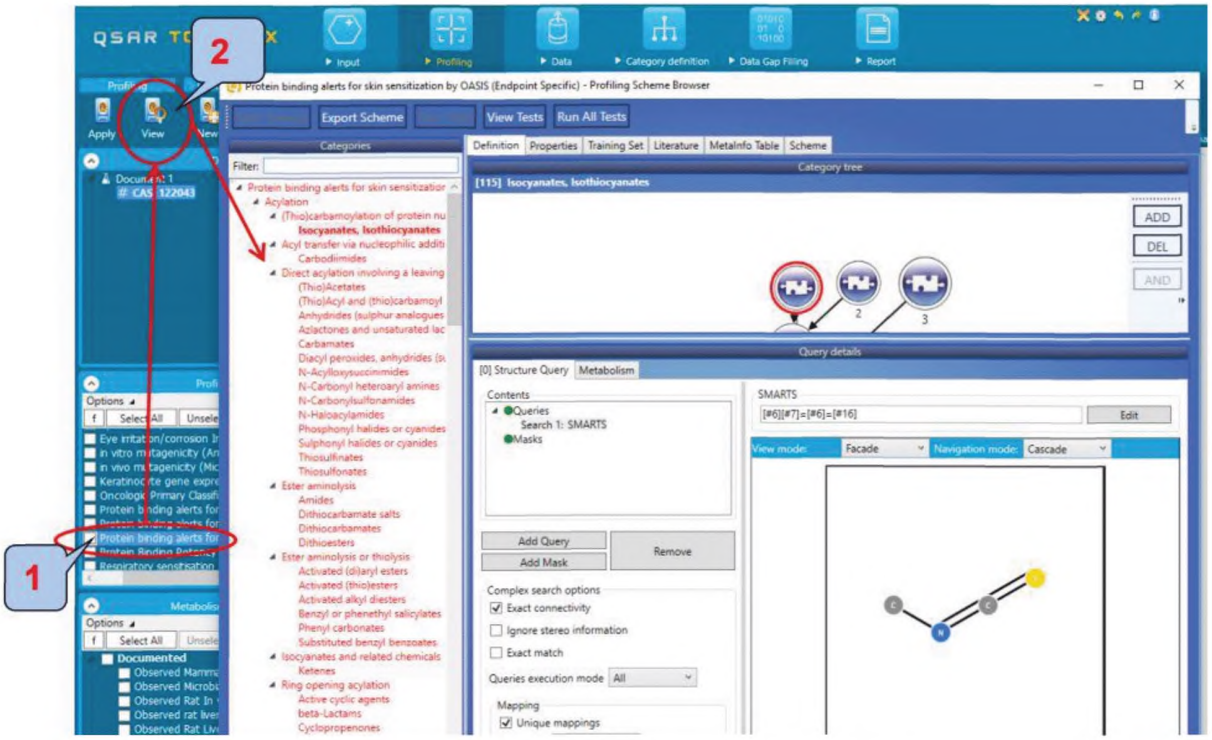


1 — название профиля; *2* — кнопка «About» («Узнать больше»); *3* — закрыть перед последующей работой

Рисунок 5 — Получение сводной информации о выбранном профиле

Для большинства профилей можно найти подробную справочную информацию, выбрав один из профилей, *1,* и нажав кнопку «View» («Просмотр») 2 (рисунок 6).

* Управление по охране окружающей среды США.



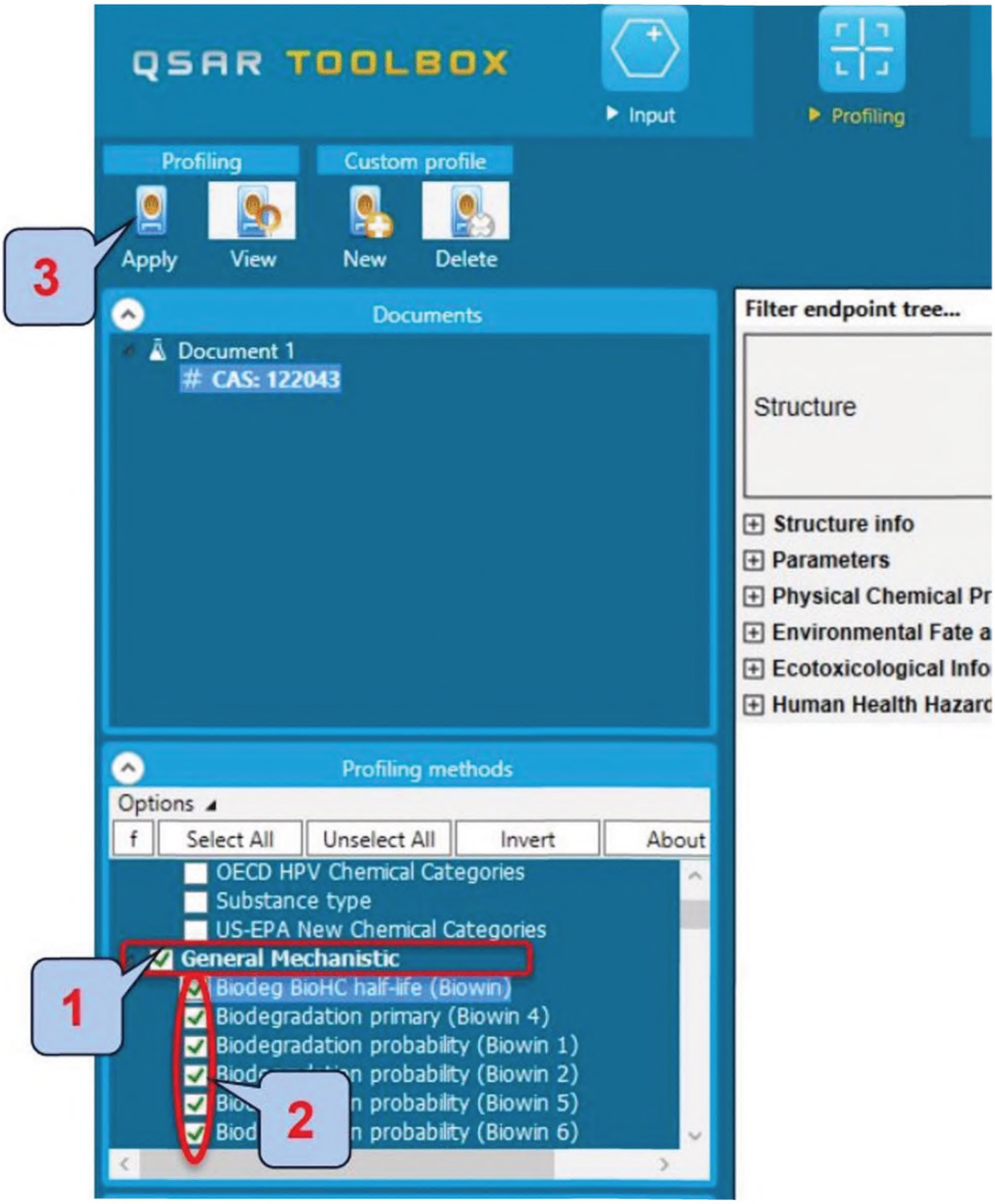
1 — название профиля; *2* — кнопка «View» («Просмотр»)

Рисунок 6 — Получение подробной справочной информации о выбранном профиле

Не все профили являются релевантными (подходящими) для решения всей совокупности конечных задач (для расчета всех параметров, например токсичности, химического вещества). Некоторые профили, например общие механистические, такие как профиль «Связывание белков» («Protein binding»), основанные на механизмах химических реакций, подходят для решения многих конечных задач (расчета кожной сенсибилизации, острой токсичности, токсичности в водной среде), некоторые профили подходят для решения лишь очень узконаправленных задач.

В разделе 7 представлены список профилей и их применимость для разных регуляторных целей.

Пример — После проверки всех профилей (1, 2) и нажатия кнопки «Apply» («Применить») 3 результат профилей отображается в матрице (рисунок 7).

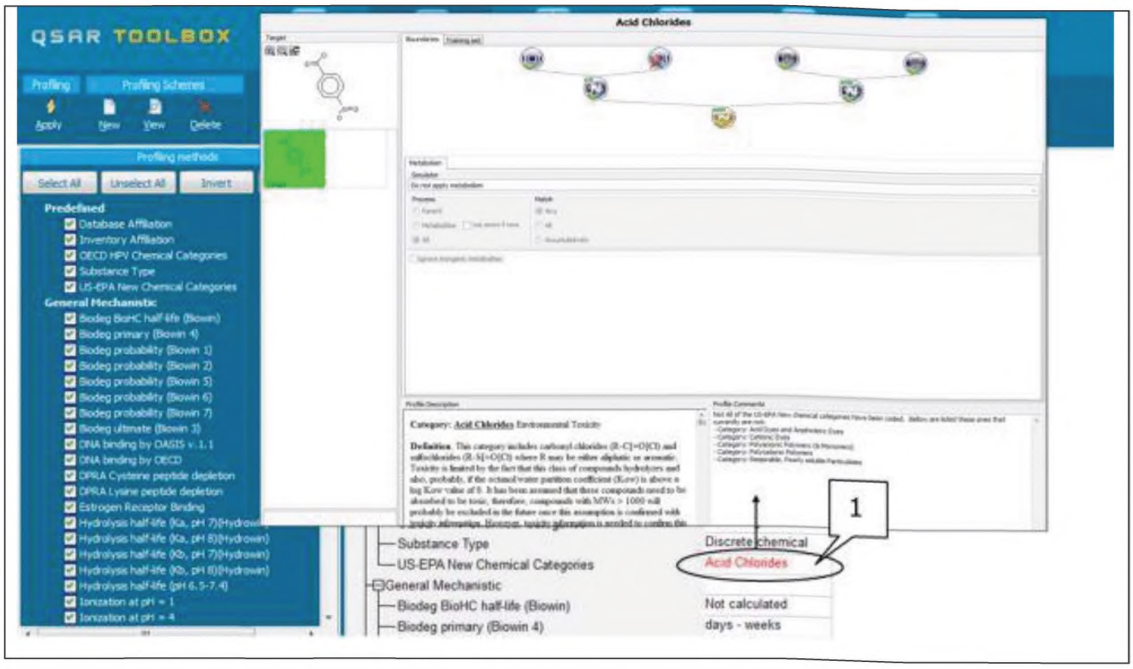


1 — название группы профилей; 2 — название профиля; 3 — кнопка «Apply» («Применить»)

Рисунок 7 — Пример работы в модуле «Профилирование»

Различные профили обнаруживают ряд специфических результатов: информация о химическом веществе представлена в нескольких базах данных, и данное вещество вписывается в ряд специальных категорий различных классификационных схем. Согласно профилю «US-EPA Новые химические категории» (US-EPA New Chemical Categories) рассматриваемое в примере химическое вещество, 4-ни-

тробензоилхлорид (CAS № 122-04-3), относится к категории Галогенангидриды. Двойной щелчок мыши по кнопке «Acid Chlorides» («Галогенангидриды»)1 на конкретном результате позволяет получить доступ к основе, детализирующей конкретные схемы для рассматриваемого вещества, а также справочную информацию о соответствующих предполагаемых реакциях (рисунок 8).



1 — **«Acid Chlorides» («Галогенангидриды»)**

Рисунок 8 — Пример работы в модуле «Профилирование»

6 Модуль «Данные» («Data»)

Чтобы перейти к следующему модулю «Данные», необходимо нажать на кнопку «Данные» («Data») на верхней панели инструментов (рисунок 9).



Рисунок 9 — Панель инструментов с выделенным модулем «Данные»

Модуль «Данные» предназначен для извлечения экспериментальных результатов из имеющихся баз данных. Версия 4.1 программы QSAR Toolbox содержит несколько постоянных баз данных (таблица 1).

Таблица 1 — Постоянные базы данных и их описание

База данных	Описание
Aquatic ECETOC	Токсичность в водной среде
Aquatic Japan MoE	Токсичность в водной среде
Aquatic OASIS	Токсичность в водной среде
Aquatic US-EPA ECOTOX	Токсичность в водной среде и биоконцентрация
Bacterial mutagenicity ISSSTY	Генная мутация, хромосомная аберрация
Bioaccumulation CANADA	Биоаккумуляция (для водных организмов)
Bioaccumulation fish CEFIC LRI	Биоаккумуляция (для рыбы)
Bioconcentration NITE	Коэффициент биоконцентрации
Biodegradation in soil OASIS	Биоразложение в почве
Biodegradation NITE	Биоразложение
Biota-Sediment Accumulation Factor	Факторы накопления биоты-осадка (BSAF)
Carcinogenic Potency Database (CPDB)	Долгосрочные исследования по канцерогенезу на животных
Carcinogenicity & mutagenicity ISSCAN	Мутагенность и канцерогенность (на грызунах)
Chemical Reactivity COLIPA	Данные для истощения цистеина и лизина на основе анализа реакционной способности пептидов
Dendritic cells COLIPA	Данные, связанные с двумя анализами: тест активации клеточной линии человека (h-CLAT) и тест U937 на кожную сенсибилизацию при миелоидном лейкозе (MUSST)
ECHA CHEM	Информация о химических веществах, произведенных или импортированных в Европу. Информация взята из регистрационных форм, представленных компаниями в Европейское химическое агентство (ECHA) в рамках Регламента REACH. Объем информации, доступной для каждого вещества, зависит от требований, указанных в REACH, — например, чем выше объем производства вещества, тем больше информации должны предоставить компании
Estrogen Receptor Binding Affinity OASIS	Аффинность связывания рецептора эстрогена, выраженная как относительная аффинность связывания в сравнении с аффинностью эстрадиола
Experimental pKa	Значение константы ионизации (pKa)
Eye irritation ECETOC	Раздражение глаз (у кролика)
GSH Experimental RC50	Абиотические тиоловые реакционноспособные электрофилы, реагирующие по механизму реакции присоединения Михаэля
Developmental toxicity ILSI	Дозозависимый эффект, связанный с лечением, для предельных значений жизнеспособности плода
Repeated Dose Toxicity HESS	Токсичность при многократном воздействии
Hydrolysis rate constant OASIS	Экспериментальные данные по нейтральному гидролизу
Cell Transformation Assay ISSCTA	Данные по четырем анализам клеточной трансформации in vitro: тест SHE (pH 6,7 и pH 7 или выше), тесты BALB / c 3T3, C3H / 10HT1 / 2 и Bhas 42
Keratinocyte gene expression Givaudan	EC1.5, EC2 и EC3 (концентрации, вызывающие 1,5-, 2- и 3-кратное увеличение индукции люциферазы)

Окончание таблицы 1

База данных	Описание
kM database Environment Canada	Константы скорости метаболической биотрансформации in vivo у водных организмов
Micronucleus ISSMIC	Исследование мутагенности микроядер (микронуклеус) in vivo (у грызунов)
Micronucleus OASIS	Исследование мутагенности микроядер (микронуклеус) in vivo (у грызунов)
MUNRO non-cancer EFSA	Данные NOEL*, LOEL**
Genotoxicity OASIS	Экспериментальные результаты для генотоксичности
Phys-chem EPISUITE	Физико-химические свойства
Rep Dose Tox Fraunhofer ITEM	Токсичность при многократном воздействии
Rodent Inhalation Toxicity Database	Данные исследований ингаляционного воздействия веществ (на крысах)
Skin irritation	Первичные показатели раздражения кожи при тестировании кожи на раздражение
Skin sensitization	Кожная и респираторная сенсибилизация
Skin sensitization ECETOC	Кожная сенсибилизация
Terrestrial US-EPA ECOTOX	Наземная экотоксикология, токсичность для млекопитающих и биоконцентрация
Toxicity Japan MHLW	Тест на токсичность однократной дозы и тесты на мутагенность
ToxRefDB US-EPA	Хроническая токсичность, данные тестов о репродуктивной токсичности и токсичности для развития
Yeast estrogen assay database University of Tennessee-Knoxville (USA)	Относительная экспрессия гена 17-бета-эстрадиола

Базы данных классифицируются по четырем основным направлениям:

- физико-химические свойства;
- состояние и перенос вещества в окружающей среде;
- экотоксикологическая информация;
- опасность для здоровья человека.

В настоящий момент по направлению исследования «физико-химические свойства» возможно рассчитать следующие параметры вещества:

- точка кипения;
- константа диссоциации (pKa);
- взрывоопасные свойства;
- точка плавления / замерзания;
- коэффициент разделения;
- давление пара;
- растворимость в воде.

По направлению «Состояние и перенос вещества в окружающей среде» возможно рассчитать следующие параметры вещества:

---

* No Observed Effect Level (УНВОЭ): уровень, не вызывающий видимых отрицательных эффектов; доза, не вызывающая отрицательной реакции; недействующая доза, т. е. без видимого эффекта; доза без видимого отрицательного эффекта.

** Lowest Observed Effect Level: наименьший наблюдаемый уровень воздействия; наименьшая доза, оказывающая действие.

- биоаккумуляция: водная;
- биодеградация;
- фоторазложение;
- стабильность в воде;
- транспортировка и распределение между компонентами природной среды.

По направлению «Экотоксикологическая информация» возможно рассчитать следующие параметры вещества:

- водная токсичность;
- наземная экотоксикология.

По направлению «Опасность для здоровья человека» возможно рассчитать следующие параметры вещества:

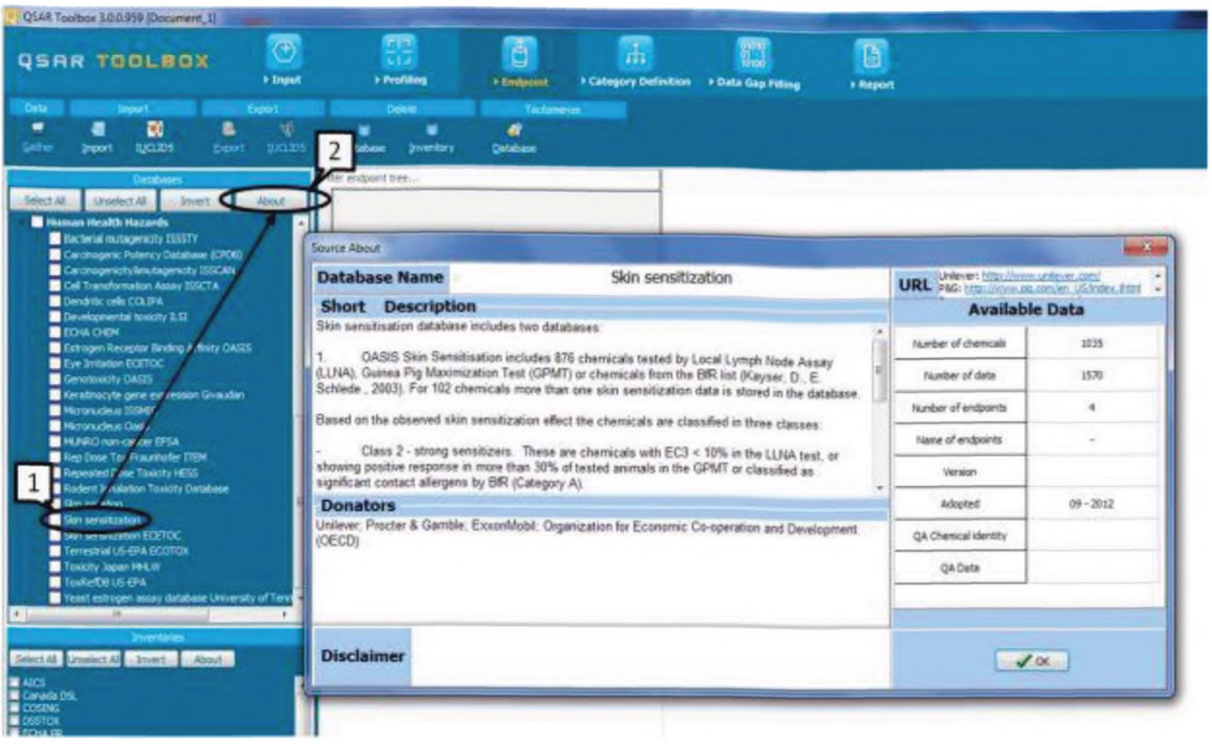
- острая токсичность;
- канцерогенность;
- токсичность для развития/тератогенность;
- генетическая токсичность;
- раздражение/коррозия;
- фотоиндуцированная токсичность;
- токсичность повторной дозы;
- аллергены;
- токсичность для размножения.

Наглядное представление размещения всех баз данных, реализованных в программе, представлено на рисунке 10.



Рисунок 10 — Наглядное представление размещения баз данных

Справочную информацию можно получить для каждой базы данных, выбрав название базы данных *1* и нажав кнопку «About» («Узнать больше») *2* (рисунок 11).



1 — название базы данных; *2* — кнопка «About» («Узнать больше»)

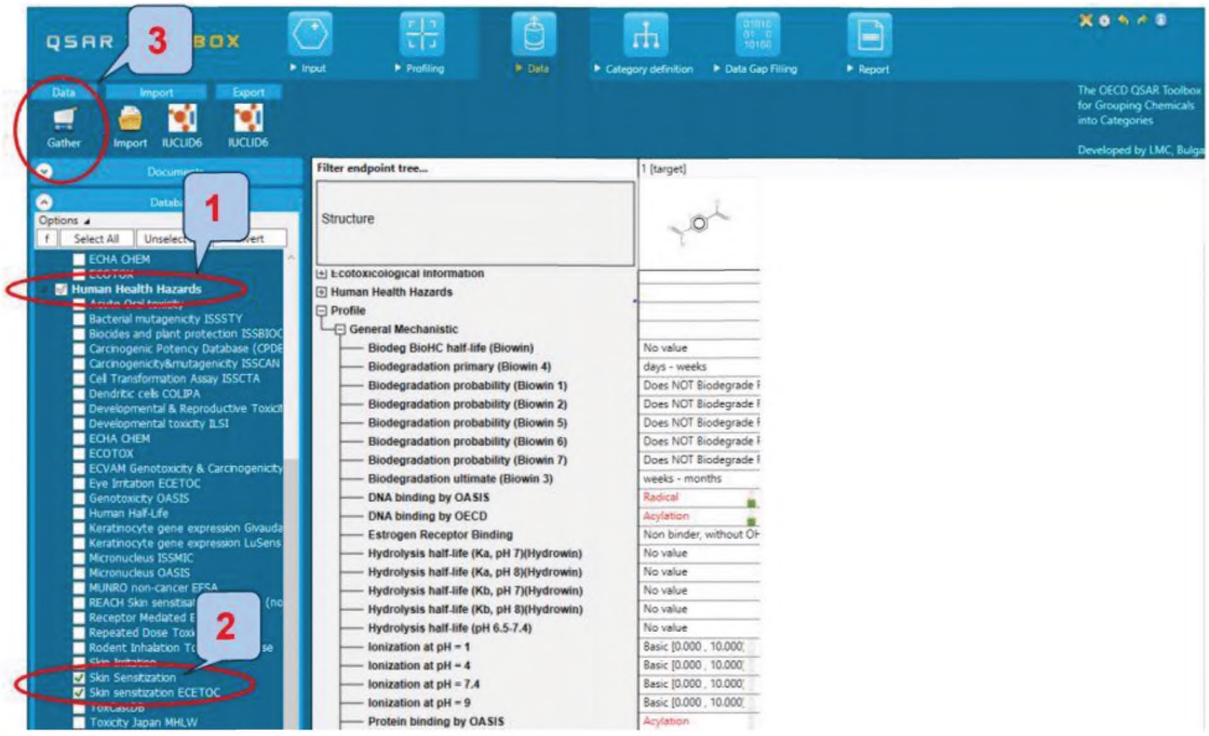
Рисунок 11 — Получение справочной информации о базе данных

Для того, чтобы выбрать, из каких баз данных следует извлекать данные, необходимо установить флажок рядом с названием базы данных.

Затем полученные данные отображаются в матрице данных для рассматриваемого химического вещества. При поиске аналогов для химической структуры, как описано в разделе 7, результаты в выбранных базах данных автоматически извлекаются и отображаются в матрице данных.

Пример — Представим, что в рассматриваемом примере для 4-нитробензоилхлорида (CAS № 122-04-3) необходимо ограничить поиск экспериментальными результатами по сенсибилизации. В настоящее время две базы данных — Skin sensitization ECETOC и Skin sensitization — содержат экспериментальные результаты по сенсибилизации.

По этой причине необходимо выбрать только существующие (в данном случае их две) базы данных 1, 2, а затем запустить запрос, нажав кнопку «Gather» («Собрать данные») 3 (рисунок 12).



***1* — направление исследования, *2* — база данных; *3* — кнопка «Gather» («Собрать данные»)**

Рисунок 12 — Пример работы в модуле «Данные»

Для запрашиваемого химического соединения не имеется экспериментальных данных по сенсибилизации, и потому данные отсутствуют.

В том случае, когда необходимо узнать, доступны ли экспериментальные данные по другим критериям выбора, необходимо выбрать другие базы данных, содержащие экспериментальные результаты, и повторно запустить запрос.

Перемещаться по списку данных необходимо, закрывая или открывая вкладки двойным кликом мыши. В рассматриваемом примере доступными являются результаты теста Эймса по изучению мутагенности.

Для быстрого поиска необходимой информации необходимо использовать опцию фильтра данных «Фильтр древа конечных точек» (Filter endpoint tree) в верхнем левом углу (на рисунке 13 область выделена красным). Например, можно ввести название видов животных, длительность обработки запроса, ожидаемый результат и так далее, для того чтобы отобразить варианты возможного выбора.

После двойного щелчка мыши по ячейке в матрице данных становится доступна дополнительная информация о результатах теста (рисунок 13).

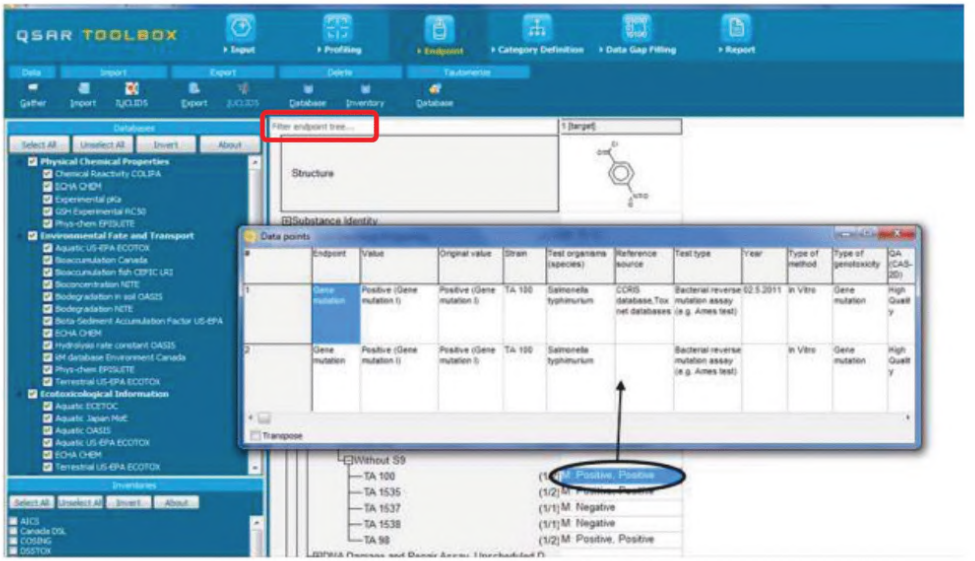


Рисунок 13 — Демонстрация дополнительной информации о результатах теста

Перед переходом к следующему модулю рекомендуется сбросить выбор отобранных ранее баз данных (рисунок 14), содержащих результаты по сенсибилизации, для того чтобы воспроизвести новые результаты, изложенные в следующих разделах.



Рисунок 14 — Сброс отобранных баз данных

7 Модуль «Определение категории» («Category Definition»)

Для перехода к следующему модулю необходимо нажать на кнопку «Определение категории» («Category Definition») на верхней панели инструментов (рисунок 15).

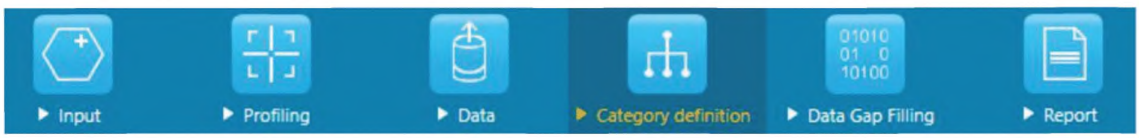


Рисунок 15 — Панель инструментов с выделенным модулем «Определение категории»

Модуль «Определение категории» позволяет группировать химические вещества по различным химическим категориям в соответствии с различными показателями сходства, например:

- определенные механизмы действия, по которым вещество вступает в химические взаимодействия;
- сходство по реакционной способности;
- сходство в строении или структуре химического соединения.

Именно так в ряду обширного перечня химические вещества могут быть сгруппированы в соответствии со степенью их показателя на токсичность в водных растворах. Или, рассматривая химическое соединение, для которого известен специфический механизм связывания белка, можно найти аналоги, которые могут связываться по такому же механизму и для которых имеются экспериментальные данные. Сводная информация об используемых методах группировки* приведена в таблице 2.

* Наиболее подходящий способ для поиска аналогов для рассматриваемого химического соединения определяется по результатам модуля «Профилирование» (раздел 5).

Таблица 2 — Сводная информация о методах группировки

Метод группировки	Сводная справочная информация
Общие методы группировки, связанные с механизмами химических реакций	
DNA Binding by OECD	Метод содержит категории или химические методы, относящиеся к связыванию ДНК. Особенно важен при изучении генотоксичности
DNA Binding by OASIS V.1.1	Метод содержит категории или химические методы связывания ДНК
DPRA Cysteine peptide depletion	Метод разработан в связи с внедрением AOP* для сенсибилизации кожи на основе данных, полученных по направленному анализу реакционной способности пептида DPRA**
DPRA Lysine peptide depletion	Метод разработан в связи с внедрением AOP для сенсибилизации кожи на основе данных, полученных по направленному анализу реакционной способности пептида DPRA
Estrogen Receptor Binding	Метод содержит простые категории для связывания рецептора эстрогена (ER). Используется при изучении репродуктивной токсичности у рыб и млекопитающих
Protein binding by OASIS v.1.1	Метод включает категории или химические механизмы связывания белка
Superfragment	Метод идентифицирует самую большую группу атомов, связанных между собой электронами. Это важно для идентификации химических веществ с различной степенью сольватации и поэтому может быть использовано для идентификации выбросов для биоаккумулирования
Biodeg probability from 1 to 7 (Biowin)	Модели, извлеченные из программы EPISUIT, связанной с изучением биоразлагаемости химических веществ
Hydrolysis half life (Ka) at pH7 and pH8	Данные профили на основе HYDROWIN, которые оценивают периоды полураспада в кислотно-катализируемом гидролизе, используя константы скорости полного гидролиза кислоты. Период полураспада для константы скорости кислотно-катализируемого гидролиза рассчитывают при pH 7 и pH 8
Hydrolysis half life (Kb) at pH7 and pH8	Профили на основе HYDROWIN, которые оценивают периоды полураспада в кислотно-катализируемом гидролизе, используя константы скорости полного кислотно-катализируемого гидролиза. Период полураспада для константы скорости кислотно-катализируемого гидролиза рассчитывают при pH 7 и pH 8
Ionization at pH 1, 4, 7.4, 9	Данные профили рассчитывают фракционную концентрацию ионизированных частиц молекулы (ионизация, %) при разных pH = 1, 4, 7.4 и 9
Методы группировки, связанные со специфическими направлениями исследований	
Acute Aquatic Toxicity MOA by OASIS	Метод классифицирует химические вещества по механизму острого токсического действия в водной среде, метод базируется на химических веществах, протестированных на рыбах семейства карповых (Pimephales promelas) до 1997 года. Данный метод особенно актуален при изучении острой водной токсичности. Поскольку этот метод только определяет механизмы действия токсина (по данным тестирования на рыбах), его лучше использовать в сочетании с другими методами группировки, связанными с острой токсичностью в водной среде
Aquatic Toxicity Classification by ECOSAR	ECOSAR — это метод для идентификации химических классов. Его использование особенно важно при изучении водной токсичности, поскольку данный метод часто используют, так как он наиболее надежный из механистических методов группировки

* Adverse Outcome Pathways: Путь негативного воздействия.

** Direct Peptide Reactivity Assay: Анализ реакционной способности, который оценивает способность химических веществ реагировать с белками.

Продолжение таблицы 2

Метод группировки	Сводная справочная информация
Acute Aquatic Toxicity Classification by Verhaar	Метод рассматривает реакционную способность, основываясь на массиве информации по данным тестов на острую токсичность у рыб вплоть до 1992 года. Этот метод особенно актуален при изучении острой токсичности. Поскольку этот метод определяет только широкие классы реакционной способности (инертные, менее инертные, реактивные и специально действующие химические вещества), его лучше использовать в сочетании с другими методами группировки, связанными с изучением острой токсичности в водной среде
Bioaccumulation — metabolism alerts	Методы основаны на структурных признаках, используемых моделью BCFBAF* для биоаккумуляции, разработанной для US EPA. Это особенно важно для определения выбросов при группировке химических веществ по аналогам для биоаккумуляции
Bioaccumulation — metabolism half life	
Biodegradation Fragments (BIOWIN MITI)	Метод основан на структурных признаках полного (быстрого) биологического разложения
Carcinogenicity (genotox and nongenotox) alerts by ISS	Метод для оценки канцерогенности, основанный на списке из 55 структурных признаков (SAs). Из них 35 извлекаются из модуля Toxtree, а 20 — являются новоизвлеченными. Большинство новых SAs относятся к негенотоксичной канцерогенности, тогда как SAs в исходном списке главным образом имеют отношение к генотоксичной. Структурные признаки для канцерогенности представляют собой молекулярные функциональные группы или субструктуры, которые, как известно, связаны с канцерогенной активностью химических веществ. При распознавании одного или более SAs, встроенных в молекулярную структуру, система маркирует потенциальную канцерогенность химического вещества
DNA alerts for AMES, MN and CA by OASIS v.1.1	Метод содержит категории или химические методы связывания ДНК. Этот метод особенно важен при изучении генотоксичности
Eye irritation/corrosion exclusion rules by BfR	Метод основан на физико-химических критериях исключения, применим для исследований по аспектам раздражения слизистой оболочки и разъедания глаз
Eye irritation/corrosion inclusion rules by BfR	Метод основан на структурных признаках для исследований по аспектам раздражения слизистой оболочки и разъедания глаз
in vitro mutagenicity (Ames test) alerts by ISS	Метод основан на модуле мутагенности/канцерогенности программы Toxtree. Он необходим для оценки мутагенности in vitro (тест Эймса) на основе списка из 30 структурных признаков (SAs). Структурные признаки для мутагенности представляют собой молекулярные функциональные группы или субструктуры, которые, как известно, связаны с мутагенной активностью химических веществ. По мере того как распознаются один или несколько SAs, встроенных в молекулярную структуру, система маркирует потенциальную мутагенность химического вещества
in vivo mutagenicity (Micronucleus) alerts by ISS	Метод основан на базе правил ToxMic программы Toxtree. Эта база правил содержит список из 35 структурных признаков (SAs) для предварительного отбора потенциальных мутагенов in vivo
Keratinocyte gene expression	Метод построен на основании анализа неблагоприятного исхода (AOP) при сенсибилизации кожи. Он разрабатывается на основе данных анализа KeratinoSens, который исследует возможность химических веществ индуцировать экспрессию репортерного гена люциферазы под контролем единственной копии ARE-элемента человеческого AKR1C2-гена, стойко встраиваемого в иммортализованные кератиноциты человека

* Существует несколько способов измерения и оценки биоаккумуляции и биоконцентрации. К ним относятся: коэффициенты распределения октанол-вода (KOW), коэффициенты биоконцентрации (BCF), факторы биоаккумуляции (BAF) и коэффициент накопления биоты-осадка (BSAF). Каждый из них может быть рассчитан с использованием эмпирических данных или измерений либо математических моделей.

Окончание таблицы 2

Метод группировки	Сводная справочная информация
Protein binding alerts for skin sensitization by OASIS v1.1	Метод группировки включает категории или химические механизмы связывания белка. Этот метод особенно актуален для кожной и респираторной сенсибилизации, острой токсичности в водной среде, а также для генетической токсичности
Oncologic Primary Classification	Метод основан на структурных признаках химических классов потенциальных канцерогенов
Skin irritation/corrosion exclusion rules by BfR	Метод основан на физико-химических критериях исключения, применим для исследований по аспектам раздражения и повреждений кожи
Skin irritation/corrosion inclusion rules by BfR	Метод основан на структурных признаках для исследований по аспектам раздражения и повреждений кожи
Эмпирические методы группировки	
Chemical elements	Профиль содержит все химические элементы из периодической таблицы, включающей восемнадцать групп
Organic functional groups	Система органических функциональных групп (OFG) разработана для того, чтобы внести классификацию и систематизировать различные функциональные группы и характерные структурные фрагменты в органических молекулах из большой базы данных, а также выявить схожие по структуре химические вещества
Organic functional groups (nested)	OFG* (составная) включает те же функциональные группы, что и OFG. Отличие состоит в том, что OFG (основная) отображает все функциональные группы, присутствующие в рассматриваемых соединениях, тогда как OFG (составная) не отображает функциональные группы, которые являются составными частями более крупных групп. Таким образом, любые «совпадающие» группы в OFG (составной) фактически образуют новые функциональные группы
Tautomers unstable	Метод нестабильного таутомера разработан на основе имеющихся данных и теоретических расчетов таутомерных форм в воде и газовой фазе. Нестабильные таутомерные формы представлены в виде отдельных категорий

Работая в модуле «Определение категории», необходимо знать:

- Если для рассматриваемого химического вещества известны конкретные механизмы возможных химических взаимодействий, которые имеют отношение к исследованию и изучению конкретных свойств, то рекомендуется проводить поиск химических веществ, которые имеют одинаковые механизмы химических взаимодействий. Результаты поиска могут быть затем уточнены путем устранения тех химических соединений, которые имеют наиболее разную структуру;

- Если для рассматриваемого химического вещества не изучены какие-либо конкретные механизмы химических взаимодействий, которые имеют отношение к изучению конкретных интересующих свойств, тогда рекомендуется проводить поиск химических веществ, структурно сходных с рассматриваемым химическим веществом. Результаты поиска могут быть затем уточнены путем исключения тех химических веществ, которые имеют отличительные механизмы химических взаимодействий.

Следует иметь в виду, что поиск аналогов осуществляется среди химических веществ, которые перечислены в выбранных базах данных «Databases» или перечне «Inventories» (поля отмечены красным на рисунке 16). Например, если выбраны только базы данных по сенсибилизации кожи «Skin sensitisation ECETOC» и «Skin sensitisation», аналоговый поиск будет проводиться только среди тех химических веществ, для которых имеются экспериментальные данные о сенсибилизации кожи в указанных базах данных.

* The Organic Functional Groups: Система органических функциональных групп.

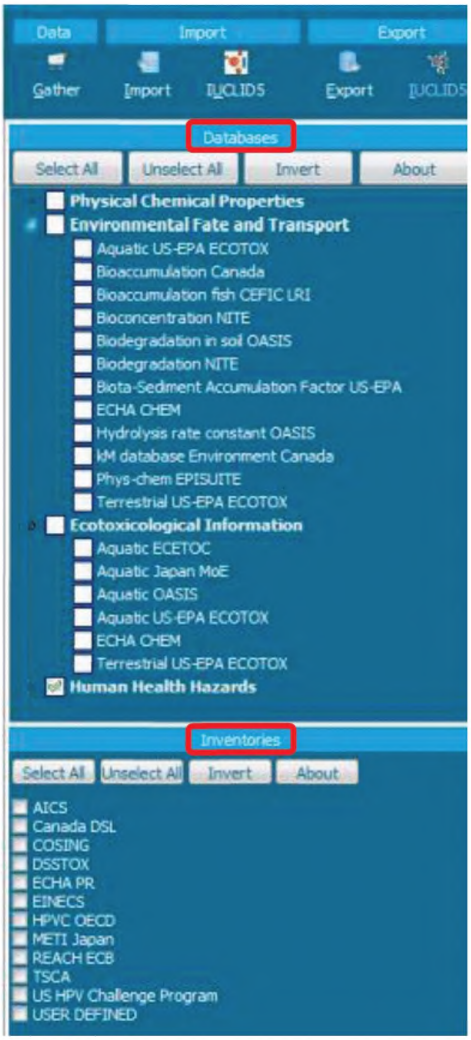


Рисунок 16 — Модуль «Определение категории»

Аналогичным образом может быть принято решение о расширении поиска, исходя из перечня химических веществ. Например, выбрав базы данных «Carcinogenicity & mutagenicity ISSCAN» и «Genotoxicity OASIS» для исследования аспектов канцерогенности и генотоксичности, а также перечень «DSL Canada», программа будет выполнять поиск аналогов как в этих базах данных, так и в указанном перечне.

При выборе базы данных и/или перечня для последующего поиска аналогов следует знать:

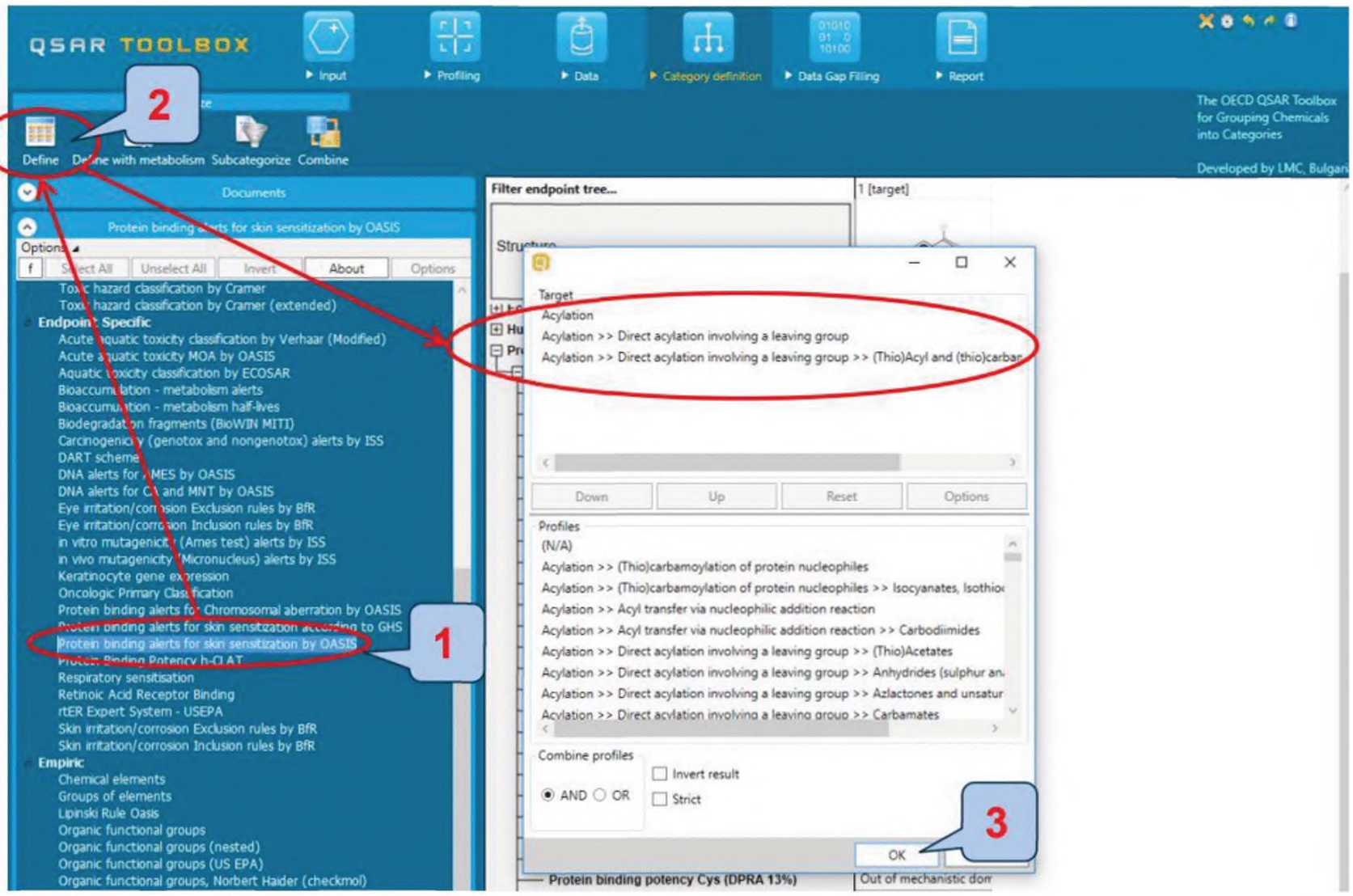
- Если целью является только поиск аналогов, для которых имеются экспериментальные данные по конкретным запросам и направлениям исследований, то следует выбрать только те базы данных, которые содержат результаты по этим запросам. Делать выбор в перечне не следует;
- Если целью является и поиск аналогов, для которых экспериментальных результатов нет, но которые предположительно полностью согласуются со свойствами рассматриваемого химического вещества, то рекомендуется выбрать соответствующий перечень.

Каждый перечень содержит от 5000 до 100000 веществ. Включение одного из них в запрос приведет к увеличению времени поиска.

Пример — Рассмотрим цель, состоящую из того, чтобы найти аналоги с имеющимися экспериментальными результатами по сенсибилизации, которые могут быть использованы в методе аналогий для нахождения необходимой информации.

Результат профилирования определяет наиболее подходящий способ поиска аналогов. Наиболее значимым профилем для рассмотрения сенсибилизации кожи является «Связывание с белками». Результаты профилирования (раздел 5) показали, что рассматриваемое вещество может связываться с белком по определенному механизму (нуклеофильное замещение в галоидангидриде карбоновой кислоты). Поэтому наиболее подходящим способом дальнейших действий является поиск химических веществ, которые могут связываться с белками по тому же механизму и для которых имеются экспериментальные результаты по сенсибилизации. Как видно в разделе 6, соответствующие базы данных выбирают по сенсибилизации «Skin sensitisation ECETOC» и «Skin sensitisation».

Затем необходимо выбрать подходящий метод группировки, включающий различные механизмы связывания белка «Protein Binding by OASIS v1.1» 1 и нажать кнопку «Define» («Определить») 2. Чтобы подтвердить выбор категорий для дальнейшего поиска, необходимо нажать кнопку «ОК» («Принять») 3 во всплывающем окне (рисунок 17).



1 — механизм связывания белка; *2* — кнопка «Define» («Определить»); *3* — кнопка «ОК» («Принять»)

Рисунок 17 — Пример работы в модуле «Определение категории»

После данных операций программа отображает те химические вещества, которые имеют такой же механизм связывания с белками, что и рассматриваемое соединение, и по которым имеются данные в ранее отобранных базах данных. В этом примере извлекается 9 химических структур (включая рассматриваемое вещество) (рисунок 18).



Рисунок 18 — Пример работы в модуле «Определение категории»

Название категории появится в окне «Отобранные категории» вместе с количеством веществ, относящихся к этой категории.

Программа автоматически предлагает выбрать направленный поиск, исходя из информации, которую следует извлечь. В данном случае необходимо выбрать конкретный запрос (кожная сенсибилизация в данном примере), выбрав вариант «Choose» («Выбрать»), либо по умолчанию выбрать извлечение данных по всем возможным запросам данного поиска «All endpoints» («Все результаты») (рисунок 19).



Рисунок 19 — Пример работы в модуле «Определение категорий»

В данном примере выбраны только те базы данных, которые содержали информацию по запросу сенсибилизации кожи, по этой причине обе опции дадут одинаковые результаты.

Экспериментальные результаты для аналогов отображаются в матрице. Чтобы получить дополнительную информацию по конкретному результату, необходимо дважды щелкнуть соответствующую ячейку в матрице (рисунок 20).

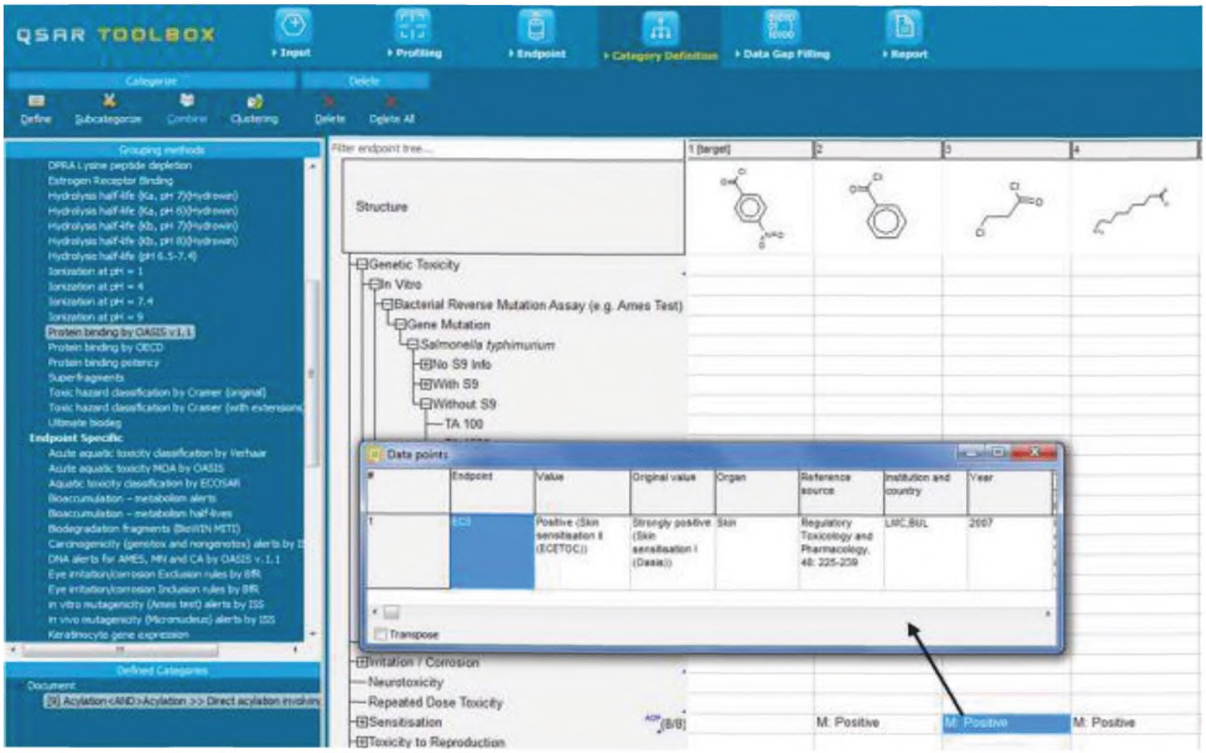


Рисунок 20 — Пример работы в модуле «Определение категорий»

Затем необходимо перейти к модулю «Сбор недостающих данных».

8 Модуль «Сбор недостающих данных» («Data Gap Filling»)

Для перехода к следующему модулю необходимо нажать на кнопку «Сбор недостающих данных» («Data Gap Filling») на верхней панели инструментов (рисунок 21).

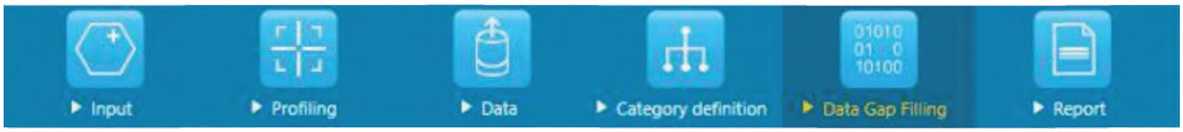


Рисунок 21 — Панель инструментов с выделенным модулем «Сбор недостающих данных»

Целью модуля «Сбор недостающих данных» является предоставление доступа к трем различным инструментам для сбора:

- метод аналогий;
- анализ тенденций;
- модели QSAR.

Метод аналогий и анализ тенденций используют имеющиеся результаты в матрице данных для сбора недостающей информации. Модели QSAR обеспечивают доступ к библиотеке внешних моделей QSAR, которые были интегрированы в программу. В зависимости от ситуации наиболее релевантный механизм для сбора недостающих данных выбирают, принимая во внимание следующее:

- метод аналогий представляет соответствующий метод сбора недостающих данных для «качественных» направлений исследований, таких как кожная сенсибилизация или мутагенность, для которых возможно ограниченное число вариаций результата (например, положительный результат, отрицательный или сомнительный). Кроме того, рекомендуется использовать метод аналогий для «количественных» направлений исследований (в частности, для рыб 96ч-ЛД50), если идентифицировано небольшое количество аналогов с имеющимися экспериментальными результатами;
- анализ тенденций является подходящим для сбора недостающих данных для «количественных» направлений исследований (в частности, для рыб 96ч-ЛД50), если идентифицировано большое количество аналогов с имеющимися экспериментальными результатами;
- модели QSAR могут использоваться для сбора недостающих данных, если для рассматриваемого химического вещества не найдено подходящих аналогов.

Пример — Если идентификация аналогов была выполнена в соответствии с конкретным механизмом химических превращений, рассматриваемое химическое вещество и его аналоги будут иметь одинаковые соответствующие механизмы. Ниже приводится иллюстрация действий на примере 4-нитробензоилхлорида (CAS № 122-04-3).

Программа определила семь дополнительных химических веществ, которые имеют такой же механизм связывания с белками, что и рассматриваемое соединение, и по которым имеются экспериментальные результаты. Сенсибилизация — это «качественный» запрос, поэтому недостающие данные можно заполнить методом аналогий.

После выделения ячейки в матрице, соответствующей недостатку данных для последующего восполнения, необходимо выбрать методы сбора недостающих данных 1, а затем нажать кнопку «Apply» («Загрузить») 2 (рисунок 22), в результате чего появится окно «Possible data inconsistency» («Возможная несогласованность данных») с указанием всех доступных таблиц. Подробная информация по каждой таблице доступна в разделе «Options» («Параметры»): подраздел «Units» («Единицы»). Все таблицы доступны к возможным преобразованиям. Таблица по сенсибилизации с наименьшим числом данных, а именно Skin sensitization ECETOC, выбрана по умолчанию. Программа рекомендует использовать метод аналогий с участием таблицы Skin sensitization ECETOC. Подтвердить использование таблицы Skin sensitization ECETOC необходимо, нажав кнопку «OK» («Принять») 3 (рисунок 22).

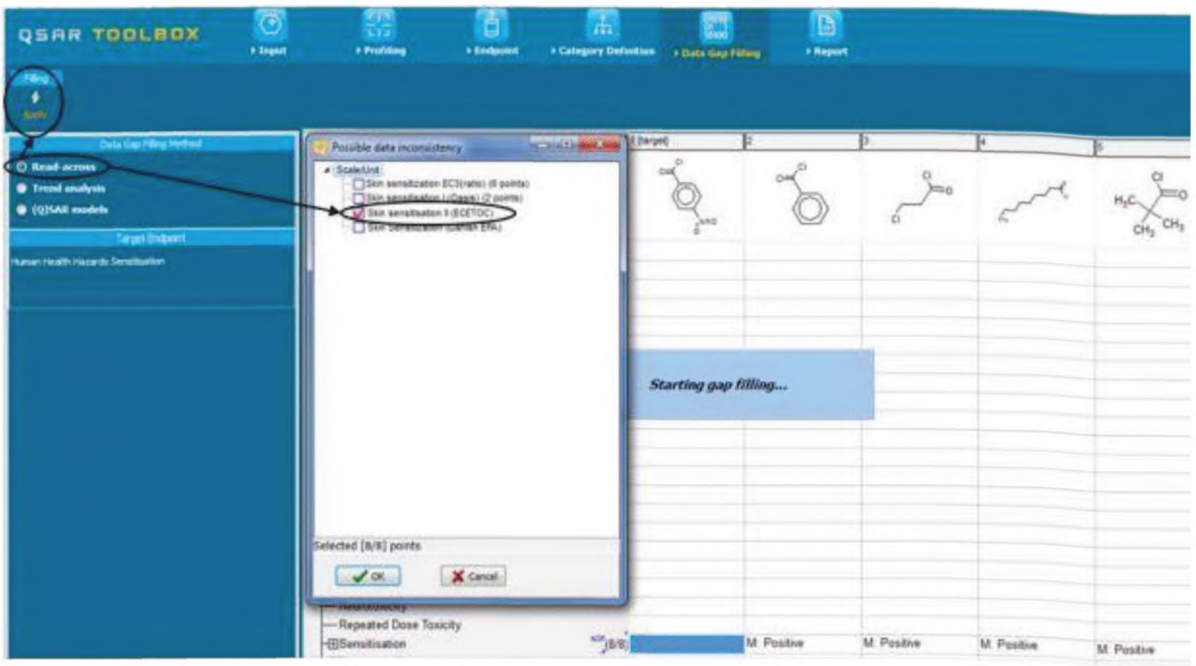


Рисунок 22 — Пример работы в модуле «Сбор недостающих данных»

После чего появится окно заполнения данных с графиком (рисунок 23).

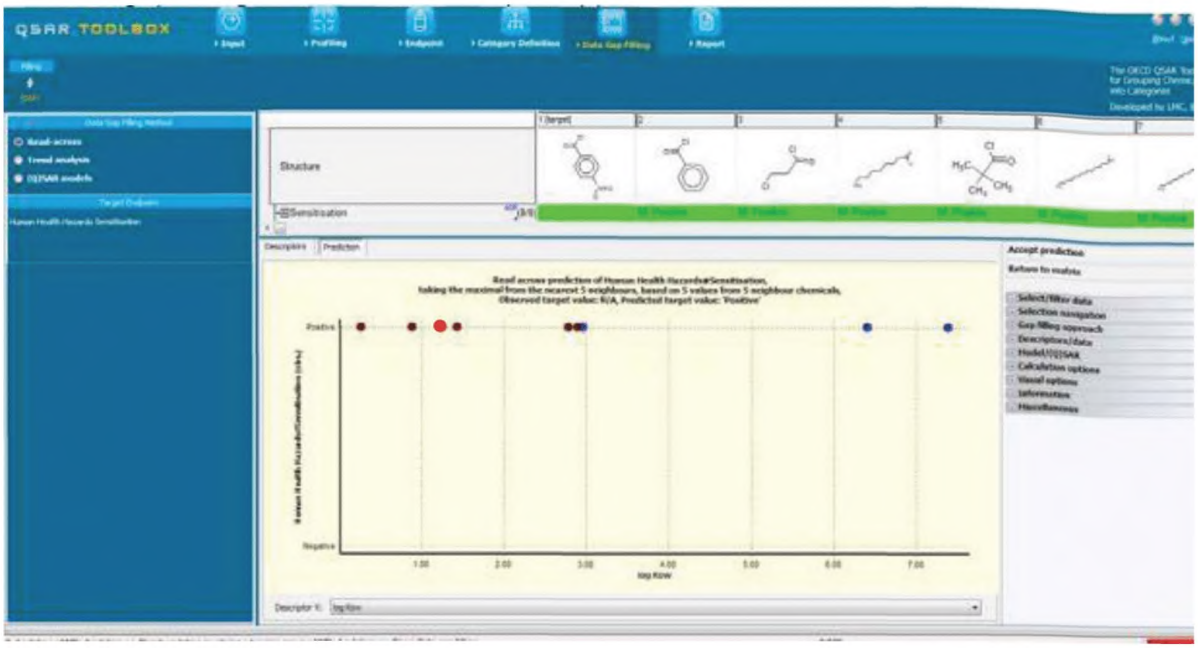


Рисунок 23 — Пример работы в модуле «Сбор недостающих данных»

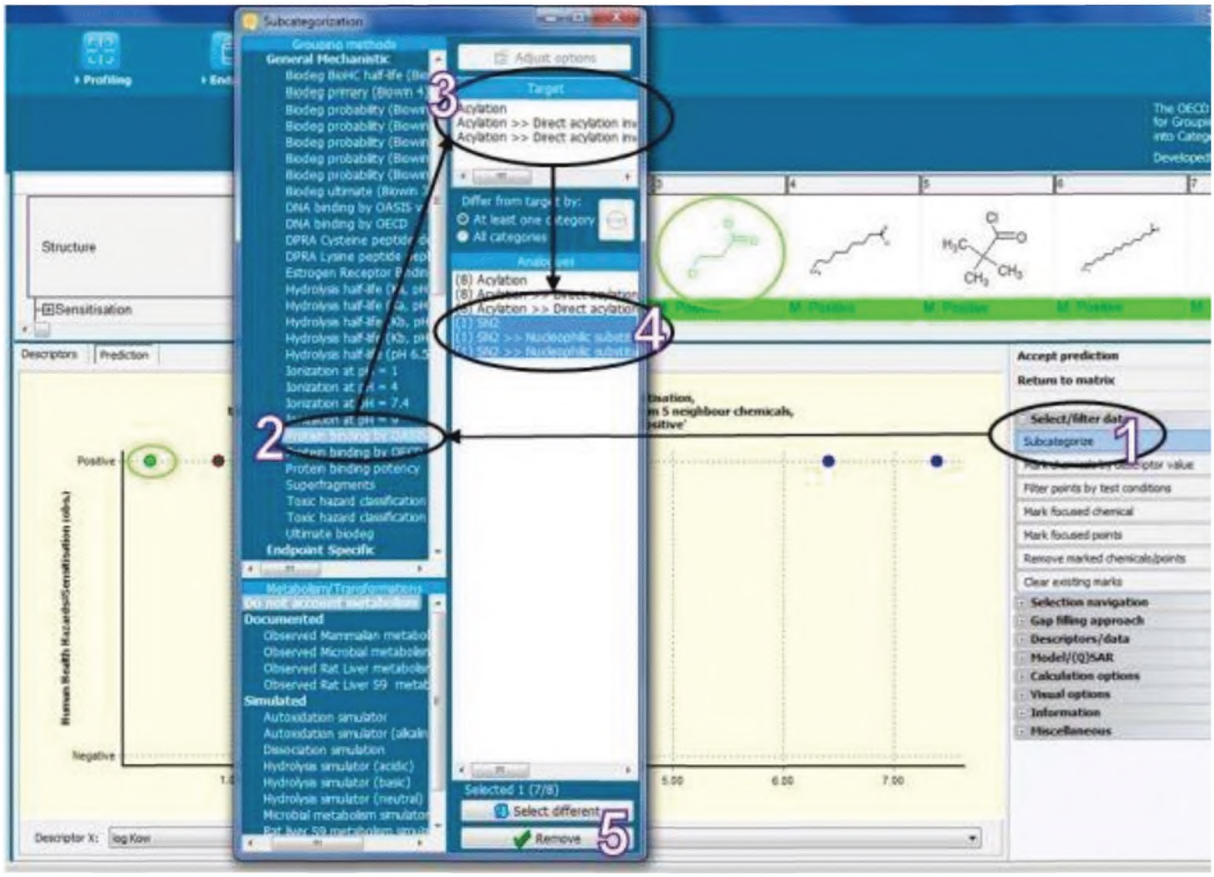
Полученный график отображает существующие экспериментальные результаты всех аналогов (ось Y) в соответствии с дескриптором (ось X). Дескриптором по умолчанию является log Kow. Темно-красные точки на графике представляют собой экспериментальные результаты, доступные для аналогов и которые используются для метода аналогий. Синие точки на графике представляют собой экспериментальные результаты, доступные для аналогов и которые не используются для метода

аналогий, потому что они наиболее удалены от рассматриваемого химического вещества, исходя из расчета дескриптора (ось X). Красная точка представляет собой оценочный результат для рассматриваемого химического вещества, рассчитанный на основании аналогов этого вещества (метод аналогий).

В этом конкретном примере все результаты аналогов положительны. Таким образом, для рассматриваемого химического вещества прогнозируется аналогичная активность в аспекте сенсибилизации. По умолчанию программа усредняет результат пяти «ближайших» аналогов (как определено дескриптором оси X), чтобы оценить результат для рассматриваемого химического вещества.

Затем необходимо проверить надежность аналогового подхода. Если идентификация аналогов была выполнена в соответствии с конкретным механизмом химического превращения, рассматриваемое вещество и аналоги будут уже иметь одинаковые соответствующие механизмы этого превращения. Тем не менее аналоги могут также обладать и другими механизмами возможных превращений из-за наличия дополнительных функциональных групп в составе их молекул.

Для того чтобы это проверить, необходимо открыть меню «Select/Filter Data» («Выбрать/Фильтровать данные») и перепрофилировать список идентифицированных аналогов, нажав кнопку «Subcategorize» («Подкатегория») 1, и выбрать метод группировки «Protein Binding by OASIS v.1.1» 2. Затем свойства аналогов сравниваются со свойствами рассматриваемого вещества 3. Как видно на рисунке 24, существует один аналог, отличающийся от рассматриваемого механизма связывания белка 4. Для удаления неподходящего вещества необходимо нажать кнопку «Remove» («Удалить») 5.

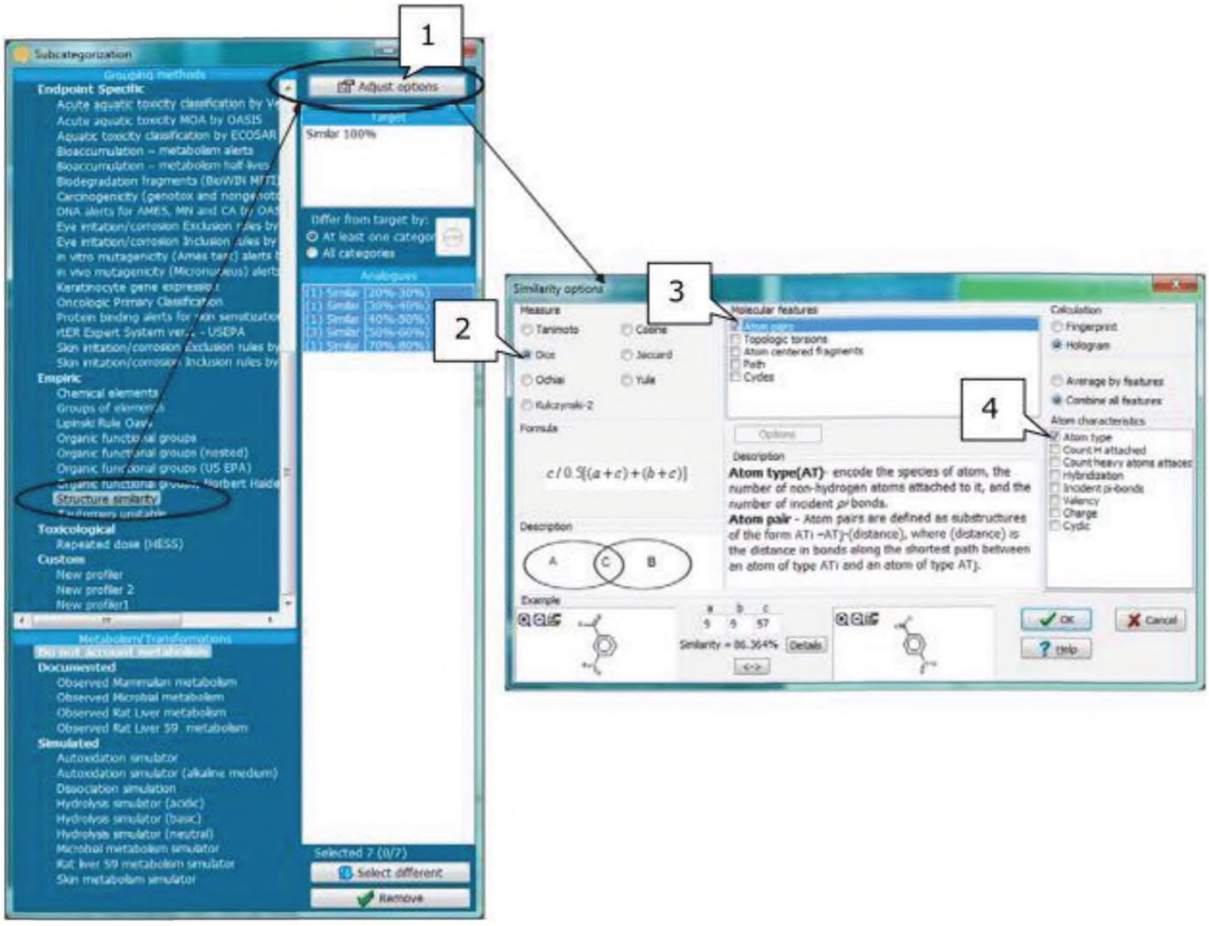


1 — кнопка «Subcategorize» («Подкатегория»); *2* — метод группировки «Protein Binding by OASIS v.1.1»; *3* — показаны свойства аналогов, сравниваемые со свойствами рассматриваемого вещества; *4* — пример существования одного аналога, отличающегося от рассматриваемого механизма связывания белка; *5* — кнопка «Remove» («Удалить»)

Рисунок 24 — Пример работы в модуле «Сбор недостающих данных»

Кроме того, необходимо идентифицировать найденные аналоги и удалить из списка те, которые имеют самое низкое структурное сходство с рассматриваемым соединением.

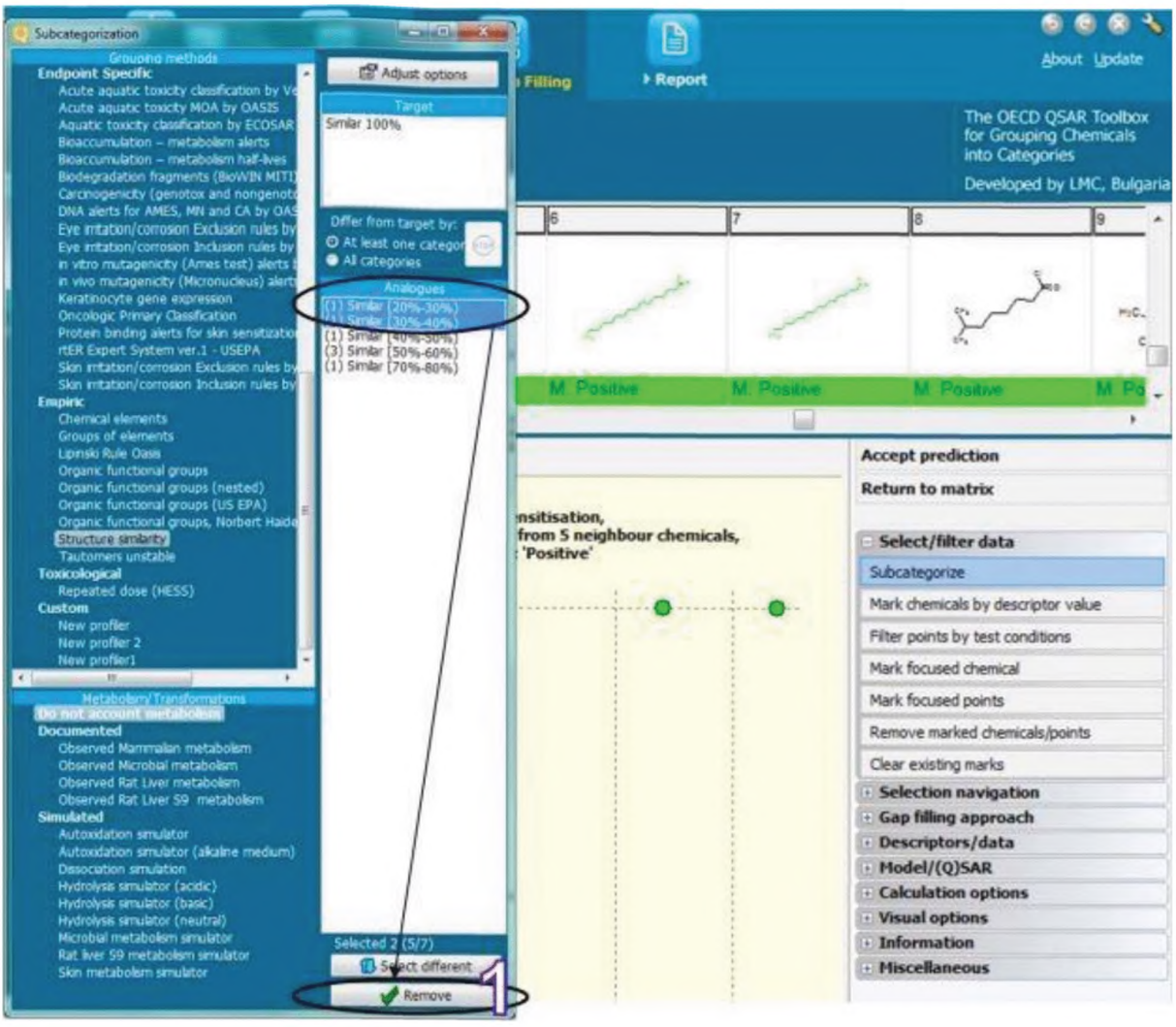
Настроить параметры подобия возможно при нажатии на кнопку «Adjust» («Установить») 1, после чего необходимо использовать метод Dice 2, выбрать атомарные пары «Atom pairs» («Атомные пары») 3, исходя из определенных молекулярных свойств, и указать тип атома «Atom type» («Тип атома») 4, исходя из его характеристичных свойств (рисунок 25).



1 — кнопка «Adjust» («Установить»); *2* — метод «Dice»; *3* — атомарные пары «Atom pairs» («Атомные пары»); *4* — тип атома «Atom type» («Тип атома»)

Рисунок 25 — Пример работы в модуле «Сбор недостающих данных»

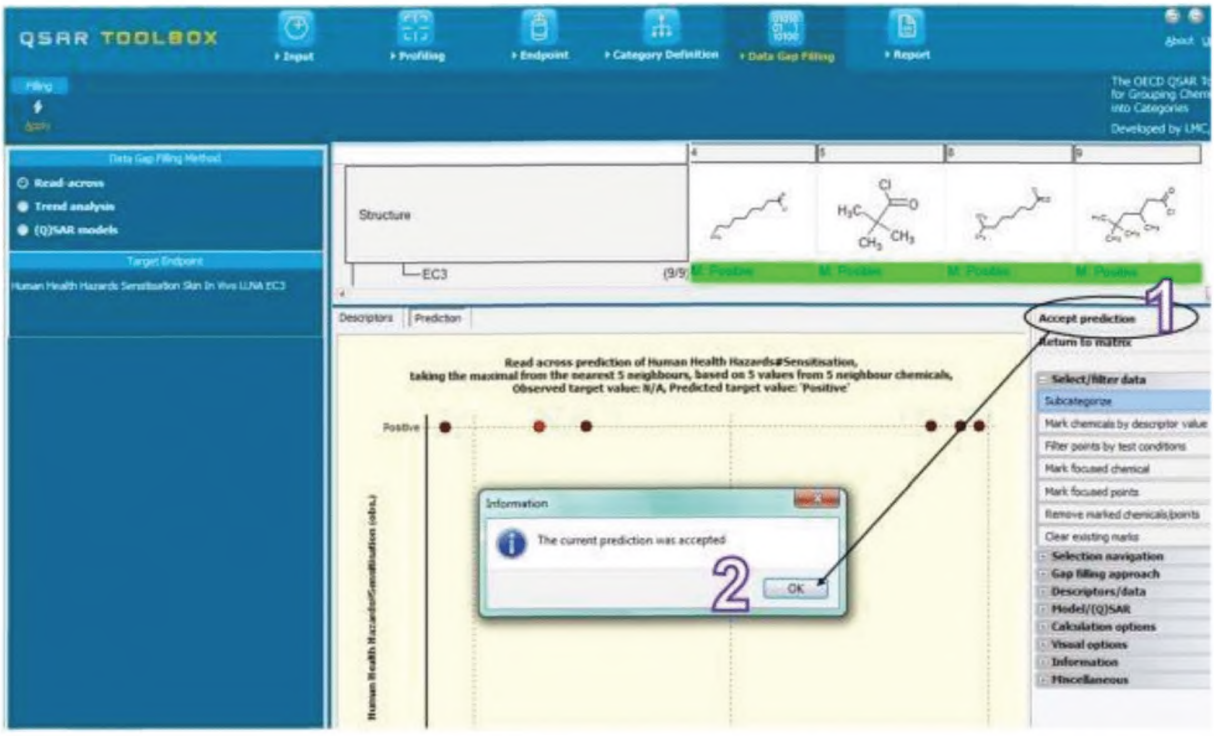
В данном примере имеется только одно химическое вещество, которое имеет очень высокое структурное сходство с рассматриваемым соединением. На данном этапе необходимо решить, сохранить ли только этот аналог только для метода аналогий или сохранить также и другие аналоги (рисунок 26). Нажатием кнопки «Remove» («Удалить») 1 можно убрать не подходящие для рассмотрения аналоги.



1 — кнопка «Remove» («Удалить»)

Рисунок 26 — Пример работы в модуле «Сбор недостающих данных»

Подтвердив, что аналоговый подход достаточно надежный как с точки зрения предполагаемых механизмов химических превращений, так и с точки зрения структурного сходства, расчетные данные можно принять, нажав на кнопку «Accept prediction» («Принять расчетные данные») 1 и нажав кнопку «ОК» («Принять») 2, в результате чего результат появится в матрице данных (рисунок 27).



1 — кнопка «Accept prediction» («Принять расчетные данные»); *2* — кнопка «OK» («Принять»)

Рисунок 27 — Пример работы в модуле «Сбор недостающих данных»

Происхождение результирующего значения, представленного в матрице, указывается в начале с прописной буквой:

- ***M — экспериментально найденный;***
- ***R — анализ аналогий;***
- ***T — анализ тенденций;***
- ***Q — для QSAR моделей.***

Нажимая кнопку «Return to Matrix» («Вернуться к матрице»), возможно закрыть просмотр и продолжить рабочий процесс.

Кроме того, можно экспортировать расчетные данные в IUCLID 5, перемещая курсор на столбец с рассматриваемым веществом и щелкая правой кнопкой мыши. Программа будет направлять с помощью разных шагов. Эти шаги включают выбор полученных значений для экспорта, возможность редактирования отчетной информации, такой как информация о веществе, категории, предварительных вычислениях и достоверности, выбор гармонизированного шаблона ОЭСР, окончательное оформление информации о записи рассматриваемого исследования, такой как достоверность, интерпретация результатов и тому подобного, а также выбор формата для экспорта, который может быть открыт через xml (*.i5z файл) или напрямую через веб-службы.

9 Модуль «Отчет» («Report»)

Для перехода к следующему модулю необходимо нажать на кнопку «Отчет» («Report») на верхней панели инструментов (рисунок 28).

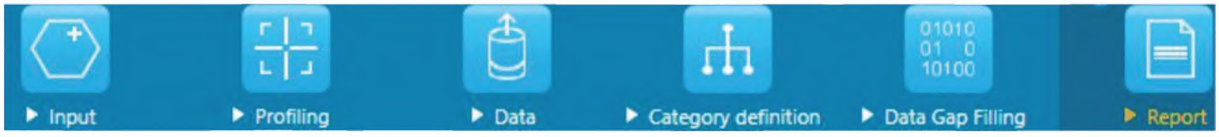
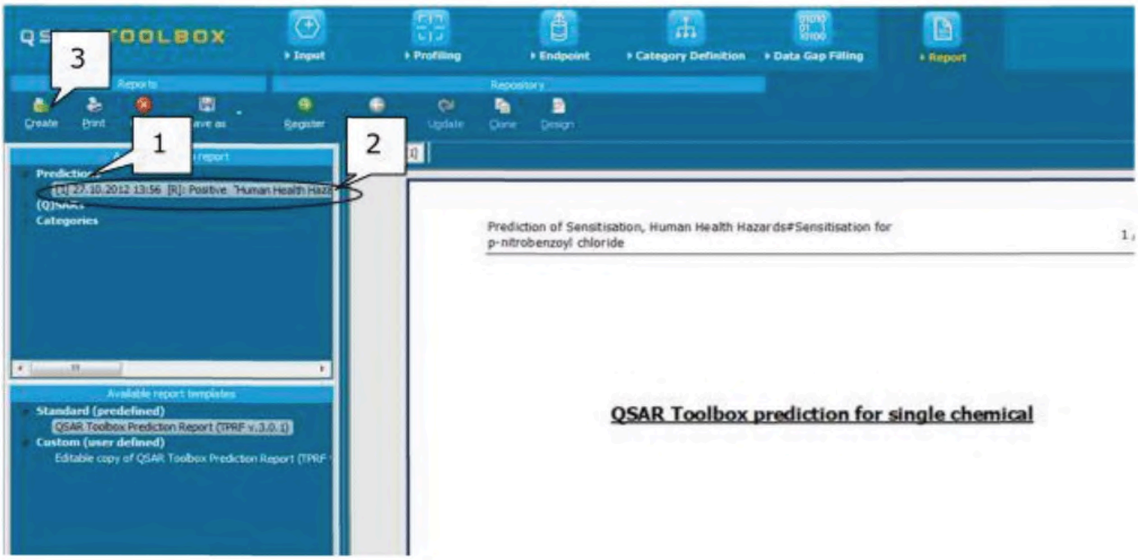


Рисунок 28 — Панель инструментов с выделенным модулем «Отчет»

С помощью модуля «Отчет» можно создавать отчеты по любому вычислению, выполняемому с помощью программы QSAR Toolbox. Панель инструментов содержит заранее установленные шаблоны отчетов, а также редактор шаблонов, с помощью которого возможно определять свои собственные задаваемые шаблоны.

Пример — Чтобы сформировать отчет по вычислениям для аспекта сенсибилизации кожи по отношению к 4-нитробензоилхлориду, необходимо в окне «Available data to report» («Доступные данные для отчета») двойным щелчком мыши открыть вычисления «Predictions» («Прогноз»), нажав кнопку 1, выбрать конкретное вычисление 2, а затем нажать на кнопку «Create» («Создать») 3 (рисунок 29).



***1* — кнопка вычисления «Predictions» («Прогноз»); *2* — конкретное вычисление; *3* — кнопка «Create» («Создать»)**

Рисунок 29 — Формирование отчета в модуле «Отчет»

10 Доступ к программному обеспечению

Программное обеспечение QSAR Toolbox является основным инструментом для нетестовых данных в ECHA и OECD, которые координируют работы по созданию и обновлению данной программы. Система находится в свободном доступе, и скачать ее для последующей работы возможно посредством электронной ссылки: gost.e-ecolog.ru

Список вспомогательной литературы, содержащей основные методики программного обеспечения QSAR и примеры расчетов различных параметров химического вещества на английском языке, доступен по электронной ссылке: gost.e-ecolog.ru

Библиография

[1] Технический регламент Евразийского экономического союза ТР ЕАЭС 041/2017 «О безопасности химической продукции»

УДК 504.05:004.18:006.354 ОКС 71.100.01

Ключевые слова: QSAR Toolbox, руководство по группировке схожих химических веществ, токсикологически значимые категории программного обеспечения ОЭСР

БЗ 12—2019/68

Редактор *Н.А. Аргунова*
Технический редактор *И.Е. Черепкова*
Корректор *Л.С. Лысенко*
Компьютерная верстка *Е.А. Кондрашовой*

Сдано в набор 03.12.2019. Подписано в печать 20.12.2019. Формат 60×84⅛. Гарнитура Ариал.
Усл. печ. л. 4,18. Уч.-изд. л. 3,55.

Подготовлено на основе электронной версии, предоставленной разработчиком стандарта

Создано в единичном исполнении во ФГУП «СТАНДАРТИНФОРМ»
для комплектования Федерального информационного фонда стандартов,

117418 Москва, Нахимовский пр-т, д. 31, к. 2.

gost.e-ecolog.ru gost.e-ecolog.ru

gost.e-ecolog.ru