

ChemSpider и кристаллографическая информация

ChemSpider (<http://www.chemspider.com/>)

В прошлом году *ChemSpider* перешел под крыло *Royal Society of Chemistry*, что придало сильный импульс его развитию, причем, похоже, взаимовыгодному. За это время на *ChemSpider* произошли такие значительные изменения, что этот сайт из разряда просто интересных мигрировал в категорию обязательных для использования. Хотя создатели позиционируют свой продукт всего лишь как "химическую поисковую систему", на самом деле уже теперь это — портал, предоставляющий огромное количество физико-химической, структурной, спектроскопической, библиографической, патентной информации в области органической химии.

Основное (но далеко не исчерпывающее) назначение *ChemSpider* — извлечение из многочисленных баз данных и концентрация в одном месте сведений об органических веществах. Существенно, что агрегация не сводится к чисто механическому суммированию; информация здесь просеивается и функционально обогащается.

В отличие от *ChemIDPlus*, *PubChem* и иных подобных справочников, ориентированных в основном на биохимическую аудиторию, *ChemSpider* накапливает широкий набор данных, как числовых, так и текстовых, графических и структурных.

В большинстве случаев ссылки на первоисточники здесь присутствуют в явной форме, что для пользователя значительно упрощает процедуру проверки достоверности информации.

В настоящее время (июнь 2010 г.) *ChemSpider* находится в стадии бета-тестирования, причем объем и конфигурация некоторых его разделов меняется чуть ли не еженедельно. Коррекция содержимого не успевает за ростом базы данных, и некоторые позиции требуют исправления или, по меньшей мере, комментариев профессионалов. Команда *ChemSpider*, в духе *Web 2.0*, привлекает онлайн-общество к рецензированию справочного материала; эффективность такого подхода явно недостаточна: на веб-страницах встречаются фрагменты некачественные, но имеющие пометку "*validated by experts*" или "*validated by users*". Понятно, что это — трудности роста, и пользователь обязательно должен их учитывать.

Объем кристаллографической информации, доступной через *ChemSpider*, еще относительно невелик.

Портал не формирует собственную базу данных, а предоставляет ссылки на те страницы иных ресурсов, где размещена кристаллографическая информация, касающаяся данного индивидуального вещества.

Целесообразность использования промежуточного инструмента — *ChemSpider* — для поиска кристаллографической информации состоит в следующем:

- *ChemSpider* имеет мощную поисковую систему, которая работает и с текстовыми, и со структурными запросами;
- *ChemSpider* предоставляет пользователю сведения из разных источников одновременно.

Особо следует подчеркнуть, что в исходных кристаллографических базах данных, которые в настоящее время подключены к *ChemSpider*, отсутствует (по состоянию на июнь 2010 г.) функция поиска информации по названию (брутто-, структурной формуле) вещества. По этой причине *ChemSpider* оказывается инструментом незаменимым.

В настоящее время (июнь 2010 г.) *ChemSpider* предоставляет доступ к базе данных [CrystalEye](#), а также к мизерной части ресурсов [Cambridge Structural Database \(CSD\)](#); ведутся работы по подключению базы данных [eCrystals](#).

[Краткое описание базы данных CSD](#) и [Краткое описание базы данных CrystalEye](#) имеются на нашем сайте.

Путь к кристаллографической информации на сайте *ChemSpider*:

- вывести на экран страницу заданного химического вещества, для чего провести поиск по любому из идентификаторов (по названию вещества, CASRN, SMILES, InChI или по структурной формуле),
- найти ссылку (если она присутствует на странице результатов поиска) и перейти к соответствующей странице во внешней базе данных.

На странице результатов поиска ссылки на внешние кристаллографические базы данных находятся (июнь 2010 г.) в разделе **Associated Data Sources and Commercial Suppliers** на следующих закладках:

а) **Xtal Structures** — ресурсы **Cambridge Structural Database (CSD)**, главным образом, демо-версии *Teaching Subset*.

Пример:

Открыть раздел

Перейти на закладку

Data Source	External ID(s)
Cambridge Crystallographic Data Centre	ETHANE01

Название внешнего ресурса

Идентификатор данного вещества в CSD. Щелкать по этой ссылке.

б) **Data Aggregators** — здесь следует искать строку базы данных **CrystalEye**.

Пример:

Открыть раздел

Перейти на закладку

Data Source	External ID(s)
PubChem	176
CrystalEye	N/A, N/A

Найти строку базы данных CrystalEye

Щелкнуть по ссылке. В данном примере база данных содержит две кристаллические модификации одного и того же вещества.