

Исследование механизмов реакций с использованием компьютерных программ

А.В.Зейгарник, Л.Г.Брук, О.Н.Темкин, В.А.Лихолобов, Л.И.Майер

Московская государственная академия тонкой химической технологии им М.В. Ломоносова
117571 Москва, просп. Вернадского, 86, факс (095)430–7983

Институт катализа им. Г.К.Борескова Сибирского отделения Российской академии наук
630090 Новосибирск, просп. акад. Лаврентьева, 5, факс (383–2)35–5776

Освещены основные этапы рациональной стратегии исследования механизмов реакций, ключевым этапом которой является выдвижение гипотез на начальной стадии исследования. Рассмотрены компьютерные программы, предназначенные для выдвижения гипотез о механизмах сложных реакций при участии пользователя. Обсуждены алгоритмы и программы эмпирического и формально-логического направлений, а также различные аспекты их применения. Проведено сравнение этих программ со сходными по структуре и алгоритмам программами компьютерного планирования органического синтеза.

Библиография — 119 ссылок.

Оглавление

I. Рациональная стратегия исследования механизмов каталитических реакций	125
II. Выдвижение гипотез о механизмах каталитических реакций	127
III. Программное обеспечение для выдвижения гипотез о механизмах сложных реакций	128
IV. Заключение	137

I. Рациональная стратегия исследования механизмов каталитических реакций

До недавнего времени доминирующим подходом к исследованию механизмов каталитических реакций был метод, который можно назвать «индуктивным». Его суть заключается в последовательном выполнении следующих этапов исследования. Сначала проводят эксперимент (главным образом кинетический). Находят математическое описание для результатов этого эксперимента. Исходя из математи-

ческого описания получают схему реакции. Затем в эту схему «вкладывают» химическое содержание, заменяя абстрактные символы (типа А, В, С, Х, Y) формулами конкретных веществ и получая таким образом механизм реакции. Подобный подход настолько прочно утвердился в сознании многих исследователей, что, несмотря на регулярно проводимые конференции, посвященные методическим вопросам кинетических исследований в области катализа, он практически не подвергался критике. В то же время познавательные возможности этого метода весьма ограничены. Очевидно, что для одних и тех же экспериментальных данных можно найти произвольное число способов их описания, из которого (особенно в сложных случаях) невозможно однозначно вывести схему реакции.

В реальной практике исследователи, изучая механизм реакции по приведенной выше схеме, изначально руководствуются какой-либо гипотезой о механизме (чаще всего имеется только одна гипотеза), обосновывая и усложняя ее по мере необходимости, если экспериментальные данные не укладываются в рамки имеющихся представлений. При этом подразумевается, что именно обработка кинетических зависимостей позволяет выявить направление модификации имеющейся гипотезы.

Анализ литературы по философии^{1–5} и применению вычислительных методов и математической статистики в химической кинетике^{6–10} ясно показывает, что индуктивный метод исследования механизмов реакций методологически ошибочен. Однако кризис индуктивного метода в кинетике наступил не только благодаря исследованиям философов и математиков, которые более других искушены в логике. Со временем объем знаний о механизмах реакций, в том числе каталитических, существенно расширился. По мере накопления этих знаний изменились и представления о сложности каталитических процессов. Сегодня исследователи

А.В.Зейгарник. Младший научный сотрудник лаборатории химической кинетики и катализа МГАТХТ. Телефон (095)437–1120, e-mail: andrew@azeigarnik.msk.ru. Область научных интересов: математическая химия, применение теории графов и комбинаторики в химии, разработка топологических индексов, применение компьютеров в химии.

Л.Г.Брук. Кандидат химических наук, старший научный сотрудник той же лаборатории. Область научных интересов: металлокомплексный катализ, химическая кинетика, реакции карбонилирования в растворах комплексов металлов, компьютерные методы в химии.

О.Н.Темкин. Доктор химических наук, профессор, руководитель той же лаборатории. Область научных интересов: металлокомплексный катализ, химия алкинов, применение теории графов в химии.

В.А.Лихолобов. Доктор химических наук, заместитель директора ИК СО РАН, заведующий лабораторией катализа комплексными соединениями металлов. Область научных интересов: катализ комплексами металлов, органический катализ кластерами.

Л.И.Майер. Кандидат химических наук, научный сотрудник лаборатории информационных систем и программного обеспечения ИК СО РАН. Область научных интересов: компьютерные методы в химии.

Дата поступления 10 мая 1995 г.

предлагают гораздо более сложные механизмы для описания каталитических реакций, чем раньше. В этих механизмах фигурируют интермедиаты более сложного строения и новые типы элементарных стадий. Они включают большее число интермедиатов и стадий, в связи с чем кинетические модели (математические уравнения), соответствующие таким механизмам, заметно усложнились. Стало ясно, что путь «от кинетики к механизму» перестает быть целесообразным для изучения механизма и построения его кинетической модели и вообще может быть продуктивным только в относительно простых случаях.

На смену индуктивному пришел дедуктивный метод, который можно изобразить в виде следующей логической последовательности:

выдвижение гипотез $\rightarrow$ планирование эксперимента

$\rightarrow$ эксперимент $\rightarrow$ дискриминация гипотез

$\rightarrow$ механизмы, не противоречащие эксперименту.

Этот метод кажется весьма естественным и находит широкое применение во многих областях науки. Интересно, что Американское химическое общество даже статьи рекомендует писать по плану, который соответствует описанным выше этапам исследования в рамках этого метода.¹⁰ Его несомненным преимуществом является то, что гипотезы формулируются в явном виде, что позволяет оценить широту «поискового пространства» и получить представление о ценности «выигравшей» гипотезы.

Исследователь фактически задает природе вопросы. По меткому выражению Д. Пойа, «природа может ответить «да» и «нет», но она шепчет один ответ и громогласно произносит другой; ее «да» условно, ее «нет» определено».¹¹ Таким образом, разница между двумя описанными здесь методами ясна: в случае «индуктивного» подхода исследователь находится в поисках весьма условного «да», а в случае «дедуктивного» метода — отбраковывает те гипотезы, для которых получено однозначное «нет». И тогда весьма важно знать не только гипотезы, оставшиеся после экспериментальной проверки, но и весь набор исходных гипотез, при этом метод работает тем лучше, чем полнее набор гипотез.

Популяризации дедуктивного метода было посвящено несколько статей.^{12–14} Сравнение двух стратегий позволило сделать важное заключение о том, что понятие «механизм реакции» объединяет два типа информации — топологическую и физико-химическую, которые являются взаимосвязанными и взаимообусловленными.¹⁴

Топологическая информация характеризует структуру механизма химической реакции (взаимосвязи интермедиатов, маршруты реакции, их сопряжение). Идентификация структуры механизма базируется на кинетических и физико-химических методах. Структура механизма может быть наглядно представлена с помощью различных графов, которые широко применяют в химии вообще^{15–20} и в теории механизмов химических реакций в частности.^{21–23}

Физико-химическая информация отражает конкретное химическое содержание механизма — состав и строение интермедиатов, их реакционную способность (значения кинетических констант, структуры переходных состояний). Топологическая и физико-химическая информации дополняют друг друга и не могут быть получены независимо, т.е. в общем случае нельзя сначала установить схему реакции, а затем вложить в нее химическое содержание.¹⁴

Важным доводом в пользу первоначального выдвижения гипотез являются постепенно складывающиеся представления о множественности механизмов в катализе. Под этим понятием мы понимаем возможность протекания реакции по нескольким различным механизмам (через разные интермедиаты) и возможность доминирования одного или нескольких таких механизмов над другими в зависимости от типа

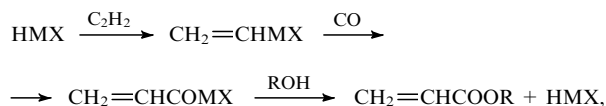
катализатора (в случае металлокомплексного катализа — от природы металла, степени его окисления и даже лигандного окружения).²⁴

Так, например, реакция

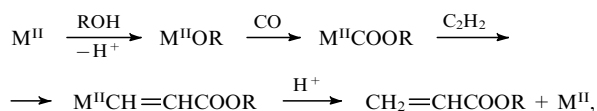


может протекать по трем механизмам:^{24, 25}

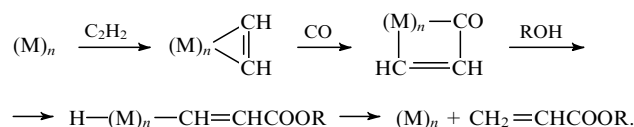
1) гидридному



2) алколюляльному

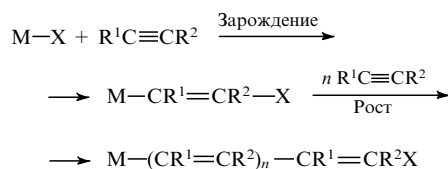


и 3) металлоциклическому

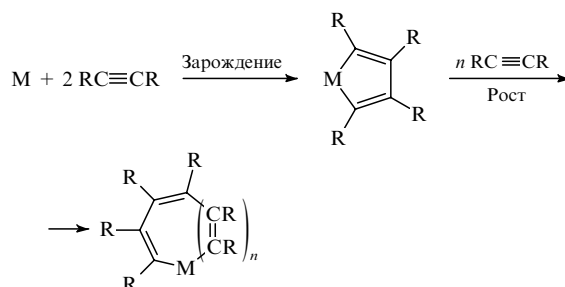


В случае комплексов Pd(0) более вероятен гидридный механизм, в случае Pd(II) — алколюляльный, а в растворах кластеров Pd(I) — металлоциклический.

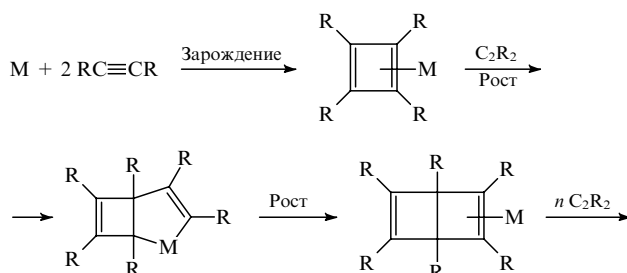
Особенно ярко множественность механизмов и их усложнение проявляются в реакциях полимеризации алкинов.²⁴ Наиболее простой механизм включает две стадии — зарождение (иницирование) и рост цепи.

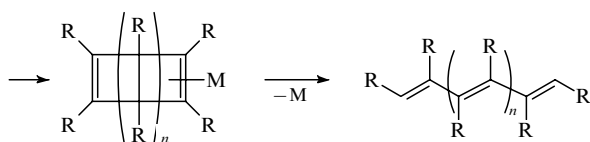


Другими механизмами полимеризации алкинов являются металлоциклический

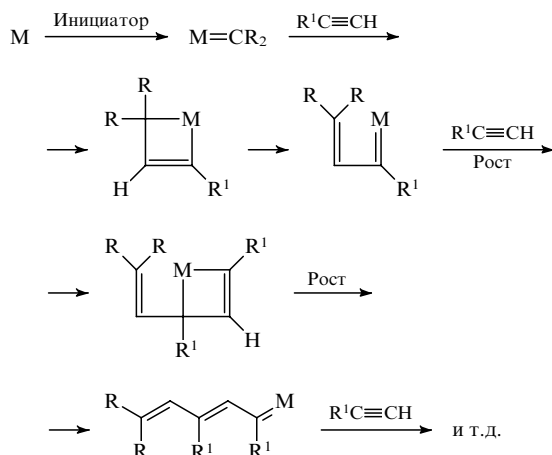


и 2+2-циклоприсоединение





Возможен также механизм, включающий несколько стадий роста и зарождения цепи.



Как видно из приведенных выше примеров, одна и та же реакция может протекать по различным механизмам. Исследователю обычно известны из литературы возможные пути реакции, и он явно или неявно использует их как гипотезы.

II. Выдвижение гипотез о механизмах каталитических реакций

Выдвижение гипотез — творческая задача, и ни сейчас, ни в будущем она не может быть полностью формализована. Однако некоторые элементы этой процедуры формализуемы. Более того, именно интуиция исследователя в сочетании с возможностью с помощью компьютера выполнять сложные алгоритмизированные процедуры и составляет основу успешного решения проблемы выдвижения гипотез о механизмах реакций.

В работе¹⁴ обсуждались три группы гипотез: 1) о природе катализатора, 2) об элементарных стадиях и 3) об узлах сопряжения. Кроме того, рассматривалась еще четвертая группа, как бы объединяющая в себе элементы этих трех групп: гипотезы о механизме реакции в целом. Каждая из групп гипотез может служить основой для планирования дискриминирующих экспериментов (здесь и далее термин «планирование эксперимента» мы применяем в более широком смысле, чем это принято у специалистов по статистике).

В случае гомогенного металлокомплексного катализа группа гипотез о природе катализатора включает следующие вопросы: состояние катализатора в растворе, наличие равновесий комплексообразования, специфическая роль растворителя, степень неидеальности раствора, влияние компонентов раствора на растворимость субстратов (если реакция гетерофазная), воздействие реакционной среды на комплексы металлов. В случае гетерогенного катализа рассматриваются состав и природа поверхности, фазовый состав катализатора, степень неоднородности поверхности и ее характер, адсорбция реагентов и продуктов, превращения поверхностных центров под действием реакционной среды.

Группа гипотез об элементарных стадиях механизма реакции позволяет проводить дискриминацию механизмов на уровне дискриминации отдельных стадий. При этом также необходим как можно более полный набор гипотез.

В сложных случаях стехиометрически неоднозначных реакций очень полезной оказывается третья группа гипо-

тез — гипотезы об узлах сопряжения в схеме механизма.¹³ Узел сопряжения — это фрагмент структуры механизма реакции, включающий такой интермедиат, который превращается в двух и более стадиях. Вид выражения для отношения скоростей образования продуктов по маршрутам, имеющим узел сопряжения, определяется структурой этого узла. Когда реакции находятся вдали от равновесия, а маршруты, по которым образуются продукты, включают необратимые стадии, отношения скоростей образования продуктов являются существенно более простыми функциями концентраций реагентов, чем сами скорости и даже выражения для селективности. Анализ узлов сопряжения целесообразно проводить, уже имея предположения о возможном механизме реакции (т.е. при планировании дискриминирующих экспериментов). Однако и при отсутствии общей схемы механизма такой анализ оказывается полезным.¹³ Примеры анализа узлов сопряжения приведены в работах^{13, 26–28}.

При наличии априорной информации об особенностях изучаемого процесса полезно перед выдвижением гипотез сформулировать требования, соблюдение которых является обязательным для всех гипотетических механизмов. Естественно, что чем больше объем априорной информации, тем меньше гипотез может быть предложено для проверки.

В настоящей работе мы ограничимся рассмотрением формальной стороны проблемы выдвижения гипотез и обсудим методы выдвижения гипотез о механизмах реакций с использованием компьютерных программ.

Большое количество публикаций посвящено проблеме компьютерного планирования органического синтеза (КПОС). Исследователи — пользователи программ КПОС — обычно решают следующие задачи: «Какие надо взять реагенты и какие провести реакции, чтобы простейшим способом осуществить синтез заданного вещества?» и «Что можно получить из заданного набора веществ?»^{29–33} По самым скромным подсчетам проблеме КПОС посвящено около 500 оригинальных статей и несколько десятков обзоров. Здесь мы ссылаемся лишь на те работы, которые содержат достаточно полную библиографию.

Методы КПОС во многом совпадают с методами компьютерного генерирования гипотез о механизмах реакций. Существует большое сходство и в алгоритмах программ. Кроме того, в обоих случаях неизбежно возникают проблемы, связанные с неэффективностью алгоритмов комбинаторного перебора. Чтобы избежать комбинаторного взрыва, особое внимание уделяют разработке эффективных критериев отбора, которые позволяют сократить время работы программы. Однако компьютерное моделирование механизмов реакций имеет и свою специфику.

Во-первых, в отличие от большинства алгоритмов КПОС, программы выдвижения гипотез о механизмах оперируют не реакциями синтетической химии, а элементарными реакциями. По крайней мере, разработчики программ для моделирования механизмов уделяют гораздо больше внимания проблеме элементарности. Во-вторых, центральным становится вопрос не о том, как проводить многостадийный синтез, а о том, какие альтернативные варианты поведения можно ожидать от химической системы, когда она «предоставлена сама себе» (после задания определенных условий). В-третьих, при компьютерном планировании синтеза используют такие критерии отбора, которые неприменимы к механизмам реакций. К ним относятся критерии, связанные с выбором «стратегии синтеза»^{31, 34–36} и с оценкой стоимости синтетических стадий;³⁷ критерии, характеризующие подобие промежуточных и конечных веществ^{29, 38–54} и использующие информационные индексы сложности и подобия для описания сложности синтеза^{55–57} и др. В некоторых программах часть этих критериев ошибочно распространяют на исследование механизмов. Такие критерии, как индексы сложности механизма, в принципе могут быть

использованы при исследовании механизмов реакций, но не в качестве критериев отбора, заложенных в программу, а для дискриминации гипотез, если необходимо выбрать простейший механизм для описания экспериментальных наблюдений. В противном случае правильная, но более сложная гипотеза может оказаться отброшенной, в то время как наборы более простых, но неверных гипотез будут подвергаться экспериментальной проверке.

При исследовании механизмов реакций на первый план выступают физико-химические критерии, связанные с оценкой реакционной способности, которые основаны, например, на расчетах электроотрицательности, поляризуемости,⁵⁸ энергии связей,^{59–64} термодинамических функций^{65,66} и других физико-химических величин. Кроме того, используют критерии, основанные на эвристических правилах, таких как правила Толмана,^{67–69} ограничения на молекулярность стадий и другие. Выбор критериев отбора является узловым моментом при разработке программного обеспечения для исследования механизмов реакций, так как они во многом определяют правдоподобие получаемых результатов и время работы программы.

Еще одним различием между программами КПОС и генерации механизмов является то, что полнота комбинаторно возможных гипотез о путях синтеза имеет меньшее значение, чем полнота гипотез о механизмах реакции.

Компьютерные программы, которые разрабатываются для генерации реакционных сетей и исследования возможных путей реакции, могут манипулировать химической информацией на более или менее обобщенном уровне. В одних случаях запись реакции осуществляется сжато (описывается только преобразование фрагмента молекулы или даже просто указывается, какие связи между какими атомами образуются или рвутся), а в других случаях — более подробно (указывается не только реагирующий фрагмент, но и все реагирующие молекулы или даже условия протекания реакции). В зависимости от того, насколько детально описывается реакция, говорят о большей или меньшей степени ее общности.

Разработчики программ сталкиваются с несколькими типичными проблемами. Генерация гипотез всегда связана с комбинаторным перебором. Зачастую алгоритм содержит несколько модулей, предназначенных для решения разных задач, связанных с перебором. Один из таких модулей обычно предназначен для решения задачи генерации интермедиатов, а другой — для проверки интермедиатов на изоморфизм (т.е. проверяется, не был ли данный интермедиат получен ранее). Для того чтобы уменьшить время работы программы, используют и специфические приемы, которые зависят от типа конкретной программы.

В принципе, существуют два типа программ: программы эмпирического направления и программы, тяготеющие к формально-логическому направлению.^{31,32} В первых используются библиотеки обобщенных реакций[†] (степень общности при этом может сильно различаться, но конкретные реакции обычно не рассматриваются, так как с их помощью могут быть получены только тривиальные гипотезы, которые уже и так описаны в литературе). Обобщенные реакции выступают в роли операторов, которые, действуя на фрагмент молекулы, дают новый фрагмент. Поскольку исходная молекула при этом известна, то становится известна и вся

реакция. Таким образом, алгоритм программы сводится к перебору фрагментов молекул, к которым могут быть применены заданные обобщенные реакции. Помимо всевозможных ограничений, задаваемых пользователем, роль фактора отбора может играть также степень общности.

В программах второго (формально-логического) направления для того, чтобы сформировать реакцию, используют логические правила (это могут быть стехиометрические соотношения, перебор всех возможных перераспределений связей и т.д.). Естественно, что программы второго направления характеризуются высокой степенью интерактивности, а также более высокими требованиями к заданию всевозможных исходных ограничений и использованию различных критериев отбора.

III. Программное обеспечение для выдвижения гипотез о механизмах сложных реакций

1. Программы формально-логического направления

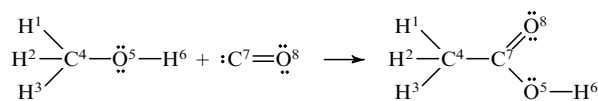
Одним из наиболее распространенных методов разработки программ исследования механизмов химических реакций является использование математической модели Дугунджи–Уги.³⁸ Этот метод был реализован в комплексе программ в рамках Мюнхенского проекта.²⁹

В модели Дугунджи–Уги, известной также как модель *be*-матрицы, используются матричные представления химических структур. Пусть задан фиксированный набор (не обязательно различных) атомов $\mathcal{A} = \{A_1, A_2, \dots, A_n\}$. Множество возможных наборов молекул, составленных из этих атомов, можно рассматривать как изомерные ансамбли молекул ЕМ. Каждый атом A_i ($A_i \in \mathcal{A}$) включается в любой изомерный ансамбль ЕМ один и только один раз. Полный набор из k изомерных ансамблей молекул образует семейство изомерных ансамблей $\text{FIEM}(\mathcal{A}) = \{\text{EM}_1, \text{EM}_2, \dots, \text{EM}_k\}$. Два изомерных ансамбля из одного семейства различны (неизоморфны), если они содержат разные наборы молекул или других свободных частиц. Исходные вещества и конечные продукты реакции могут быть представлены в виде структурных *be*-матриц двух изомерных ансамблей (для исходных веществ и для продуктов). Недиagonальные элементы *be*-матрицы отражают связность атомов из множества $\mathcal{A}$ в молекулах, входящих в соответствующий изомерный ансамбль. Диагональные элементы характеризуют число свободных валентных электронов соответствующего атома. Пусть *be*-матрица **B** описывает исходные вещества, а *be*-матрица **E** — конечные вещества в некоторой реакции $\text{EM}(\mathbf{B}) \rightarrow \text{EM}(\mathbf{E})$.[‡] Тогда перераспределение связей может быть описано фундаментальным уравнением модели Дугунджи–Уги

$$\mathbf{B} + \mathbf{R} = \mathbf{E},$$

где **R** — матрица, преобразующая **B** в **E** (*r*-матрица).³⁸

Рассмотрим в качестве примера реакцию карбонилирования метанола.



[†] Под обобщенной реакцией мы понимаем описание химической реакции, избавленное от тех или иных деталей, отличающих каждую конкретную реакцию от сходных с ней реакций. Такими деталями могут быть атомы или целые фрагменты, не принимающие непосредственного участия в перераспределении электронов или в формальном перераспределении связей. В обобщенных реакциях обозначения атомов могут быть заменены на символы. Форма записи обобщенной реакции может быть совсем непривычной для химиков: могут быть использованы матрицы, графы, файлы специального формата или отдельные их части.

[‡] Здесь и далее мы используем сокращения и обозначения, принятые в англоязычной литературе: ЕМ — ансамбль молекул (ensemble of molecules), FIEM — семейство изомерных ансамблей (family of isomeric ensembles of molecules). Латинские символы *be* и *r*, фигурирующие в названиях матриц, происходят от слов bond-electron matrix (матрица связей и электронов) и reaction matrix (реакционная матрица). Символы **B** и **E**, происходят от слов beginning (начало) и end (конец).

Схема 1

	H ¹	H ²	H ³	C ⁴	O ⁵	H ⁶	C ⁷	O ⁸
H ¹	0	0	0	1	0	0	0	0
H ²	0	0	0	1	0	0	0	0
H ³	0	0	0	1	0	0	0	0
C ⁴	1	1	1	0	1	0	0	0
O ⁵	0	0	0	1	4	1	0	0
H ⁶	0	0	0	0	1	0	0	0
C ⁷	0	0	0	0	0	0	2	2
O ⁸	0	0	0	0	0	0	2	4

B

+

	H ¹	H ²	H ³	C ⁴	O ⁵	H ⁶	C ⁷	O ⁸
H ¹	0	0	0	0	0	0	0	0
H ²	0	0	0	0	0	0	0	0
H ³	0	0	0	0	0	0	0	0
C ⁴	0	0	0	0	-1	0	1	0
O ⁵	0	0	0	-1	0	0	1	0
H ⁶	0	0	0	0	0	0	0	0
C ⁷	0	0	0	1	1	0	-2	0
O ⁸	0	0	0	0	0	0	0	0

R

=

	H ¹	H ²	H ³	C ⁴	O ⁵	H ⁶	C ⁷	O ⁸
H ¹	0	0	0	1	0	0	0	0
H ²	0	0	0	1	0	0	0	0
H ³	0	0	0	1	0	0	0	0
C ⁴	1	1	1	0	0	0	1	0
O ⁵	0	0	0	0	4	1	1	0
H ⁶	0	0	0	0	1	0	0	0
C ⁷	0	0	0	1	1	0	0	2
O ⁸	0	0	0	0	0	0	2	4

E

В этом случае набор атомов, составляющих изомерные ансамбли из семейства FIEM, может быть описан как $\langle H^1, H^2, H^3, C^4, O^5, H^6, C^7, O^8 \rangle$, а уравнение модели Дугунджи–Уги принимает вид, представленный на схеме 1.

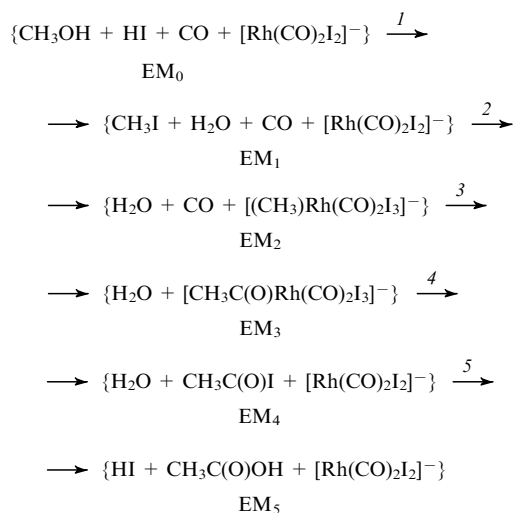
Матрица **R** может быть представлена в виде суммы *r*-матриц элементарных стадий механизма

$$\mathbf{R} = \mathbf{R}_1 + \mathbf{R}_2 + \dots + \mathbf{R}_S$$

(*S* — общее число элементарных стадий), а весь механизм — в виде последовательности трансформаций изомерных ансамблей.



Для каталитических реакций набор **A** должен быть дополнен атомами, входящими в состав катализатора. Например, описанная выше реакция протекает в присутствии каталитической системы Rh(I)—HI.^{70–72} Тогда изомерный ансамбль $EM_0 \equiv EM(\mathbf{B})$ должен включать частицы HI и $[Rh(CO)_2I_2]^-$. Упрощенный механизм реакции карбонилирования метанола можно записать следующим образом:⁷²



В данном примере каждая из пяти стадий механизма может быть представлена в матричной форме

$$\mathbf{B}_i + \mathbf{R}_i = \mathbf{E}_i, i = 1, 2, \dots, 5,$$

где $\mathbf{B}_i = \mathbf{E}_{i-1}$, а итоговое превращение — в виде

$$\mathbf{B}(EM_0) + \mathbf{R} = \mathbf{E}(EM_5).$$

$$\text{Здесь } \mathbf{R} = \sum_i \mathbf{R}_i.$$

В рамках Мюнхенского проекта разработана компьютерная программа RAIN для исследования механизмов реакций.^{29, 30, 73–78} Отличительной чертой этой программы является так называемый «билатеральный» поиск. Программа RAIN унаследовала от КПОС идею ретросинтеза,^{79–80} которая состоит в том, что можно вести поиск возможных исходных реагентов, исходя из продуктов (ретропоиск). В RAIN поиск осуществляется одновременно в прямом и обратном направлениях (билатеральный поиск). В программе RAIN используются генераторы реакций,⁷⁷ которые позволяют найти как прямые, так и ретрореакции. Генераторы реакций — это фактически свод логических правил, управляющих процессом формирования конкретной реакции (или ретрореакции). В основе таких правил лежит комбинаторное перечисление всех возможных способов перераспределения связей. Перечисление осуществляется с заданными пользователем ограничениями, основным из которых, по-видимому, следует считать таблицы переходов (transition tables). Эти таблицы, составляемые для атомов каждого типа, ограничивают число превращений (и ретропревращений) валентных состояний атомов в ходе протекания (генерирования) элементарных стадий. В процессе генерации реакций не используются никакие обобщенные реакции, поэтому программу RAIN можно отнести к программам формально-логического направления.

Идея билатерального поиска реализуется следующим образом. Задаются изомерные ансамбли молекул исходных ($EM(\mathbf{B})$) и конечных ($EM(\mathbf{E})$) веществ. Для ансамбля $EM(\mathbf{B})$ осуществляется поиск реакций, а для $EM(\mathbf{E})$ — ретрореакций. На каком-то этапе происходит совпадение двух (или более) промежуточных изомерных ансамблей, полученных путем поиска в различных направлениях. Другими словами, генерация путей реакции в направлении продуктов и в направлении реагентов должна привести к изоморфным изомерным ансамблям. Разработаны методы оптимизации поиска изомерных ансамблей молекул, сочленяющих дерево поиска с деревом ретропоиска³⁰ на основе принципа минимального химического расстояния.^{29, 30, 38–43} За величину химического расстояния принимают число электронов, которое должно быть перераспределено, чтобы превратить один ансамбль молекул в другой.

Действие этого принципа, как оптимизирующего фактора, основано на следующем предположении. Пусть некоторый ансамбль молекул подвергнут различным превращениям. Тогда процессы, которым соответствуют меньшие значения химических расстояний между реагентами и продуктами, будут протекать с более высокой вероятностью, чем процессы, в которых эти значения больше. Допустимые отклонения от этого правила составят, по мнению авторов работы⁷⁷, одну или две единицы химического расстояния (один или два электрона), т.е. реакция иногда может выбрать путь не минимальных структурных изменений,

однако эти отклонения от общего правила количественно не слишком велики.

Комбинаторный перебор в программе RAIN осуществляется при ограничениях, которые задает пользователь. В частности, кроме таблиц переходов вводят и другие формальные ограничения: максимальное число атомов, входящих в реагирующую часть молекулы; максимальное изменение порядка связи в ходе элементарной реакции; максимальное число связей, меняющих свой порядок; минимальный размер циклов углеродных скелетов молекул, которые содержат тройные или кумулированные двойные связи; максимальное число заряженных атомов; максимальное число ковалентных связей между гетероатомами; максимальное число циклов; максимальное число отдельных молекул в промежуточных ансамблях молекул и т. д. (всего 18 правил).

Билатеральный поиск имеет ряд недостатков. В отличие от КПОС, в случае исследования механизмов реакций необходимо осуществлять поиск в одном (лучше прямом) направлении, так как образующиеся в ходе реакции интермедиаты могут взаимодействовать между собой и участвовать в реакциях с исходными веществами. При этом возникает необходимость расширения множества $\mathcal{A}$ на каждом шаге генерации возможных ансамблей молекул (и, соответственно, увеличения размеров этих ансамблей), что невозможно при осуществлении билатерального поиска, который в рамках модели Дугунджи–Уги требует фиксации множества $\mathcal{A}$, не изменяемого в ходе работы программы. Как свидетельствуют авторы работы³², это впервые заметил Ионеда, но в статье, содержащей описание его программы GRACE,⁸¹ не рассматривался вопрос о возможных расширениях множества $\mathcal{A}$ в ходе ее работы.

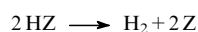
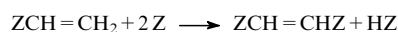
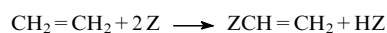
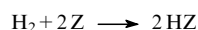
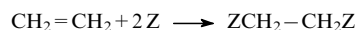
Программа GRACE, как и RAIN, принадлежит к программам формально-логического направления, однако в ней используется несколько иной принцип генерации реакций. Ионеда ввел представления о групповых реакционных матрицах (матрицах, описывающих превращения внутри функциональных групп молекул). После ввода ограничений программа GRACE идентифицирует функциональные группы внутри молекул и логико-комбинаторным путем генерирует групповые реакционные матрицы. Затем она обнаруживает все допустимые комбинации групповых реакционных матриц, из которых могут быть получены матрицы элементарных превращений (r -матрицы стадий). Таким образом, в программе GRACE, как и в программе RAIN, не используют ни библиотеки превращений функциональных групп, ни библиотеки элементарных реакций, которые являются их комбинациями.

В программе GRACE критерии отбора формулируются пользователем по заданной схеме. По умолчанию действуют следующие ограничения:

- 1) число атомов (реакционных центров), которые изменяют свое валентное состояние в ходе элементарных стадий, не должно превышать 4;
- 2) количество разорвавшихся (пониживших порядок) и образовавшихся (повысивших порядок) связей не должно быть больше 3;
- 3) молекулярность реакции не должна превышать 3;
- 4) число неспаренных электронов на одном атоме не должно быть больше 1;
- 5) число атомов с неспаренными электронами в одной молекуле не должно быть больше 4;
- 6) число валентных электронов у атомов углерода и водорода равно соответственно 4 и 1;
- 7) по крайней мере один активный центр катализатора должен участвовать в каждой стадии;
- 8) атомы водорода должны быть связаны с катализатором и не могут находиться в свободном состоянии в виде радикалов. В программе предусмотрена возможность изменения конкретных цифр, более того, можно вовсе не исполь-

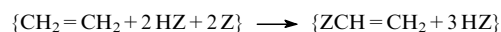
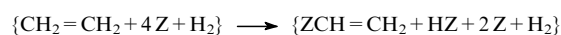
зовать то или иное ограничение. Кроме этих ограничений задается верхний предел сложности элементарного преобразования. В качестве меры сложности принята величина, вычисляемая по r -матрице и равная половине химического расстояния. Построение сети элементарных реакций (реакционной сети) ведется только в прямом направлении, что позволяет при необходимости расширять множество $\mathcal{A}$ в ходе ее наращивания.

Работу программы можно проиллюстрировать на примере получения реакционной сети гетерогенно-каталитического гидрирования этилена⁸¹



Z — активный центр катализатора.

Интересно, что набор тех же самых реакций, но записанный в виде преобразований ансамблей молекул, будет несколько больше в силу идентичности некоторых преобразований. Например, два следующих преобразования являются по сути одной и той же реакцией.

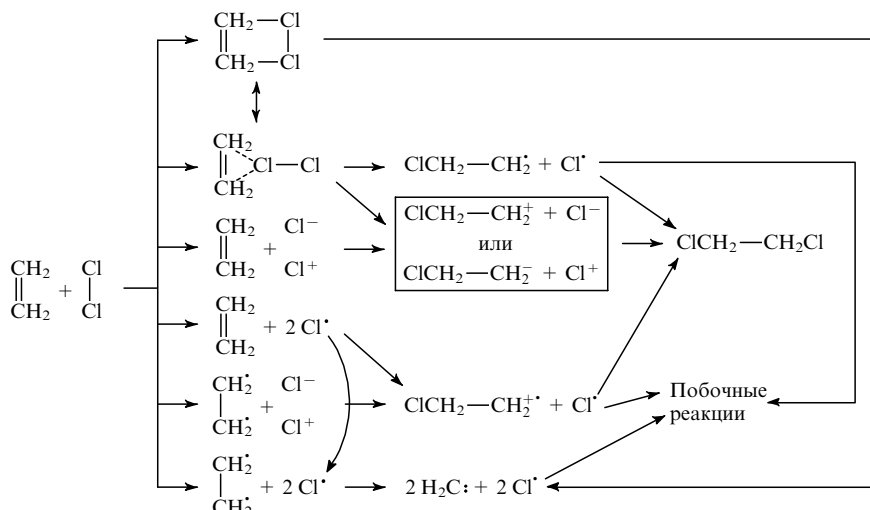


В этом смысле модель Дугунджи–Уги несколько избыточна. Действительно, если бы одна и та же реакция не дублировалась, потребовалось бы меньше времени для работы программы. При значительном числе элементарных стадий этот фактор играет существенную роль. В приведенном выше примере имеется семь реакций. Если представить участников реакций в виде ансамблей молекул, то требуется 11 преобразований таких ансамблей (т.е. в полтора раза больше). Пропорционально возрастет и время на осмысление результатов работы программы.

Модель Дугунджи–Уги была использована также в программе Дозморова.⁵⁵ Он ввел новую характеристику подобия, которая позволила уменьшить число генерируемых стадий и интермедиатов и которая является по своей структуре топологическим информационным индексом, учитывающим информационные вклады различной природы.⁸² Применение этой характеристики в качестве критерия отбора, на наш взгляд, не имеет достаточного физического обоснования. Детали алгоритма Дозморова не сообщает, но, по-видимому, в этой программе, как и в программе GRACE, используется поиск только в прямом направлении, а генерирование элементарных реакций ведется логико-комбинаторными методами (т.е. она тоже относится к формально-логическому направлению). Реакционная сеть «информационно» разрешенных стадий реакции хлорирования этилена, полученная с использованием данной программы, показана на схеме 2.

Таким образом, во всех программах формально-логического направления, основанных на модели Дугунджи–Уги, используются характеристики подобия для уменьшения количества «ложных» решений. В программе RAIN — это химическое расстояние (из всех преобразований ансамбля молекул выбирают только те, которые соответствуют минимальным значениям химического расстояния с погрешностью ± 2 единицы химического расстояния). В программе GRACE — это величина сложности элементарной реакции, равная половине химического расстояния (пользователь

Схема 2



задает верхний предел такой величины). В программе Дозморова — это информационный топологический индекс, с помощью которого оценивается «информационная разрешенность» элементарной реакции.

Сравнительно недавно появилась программа MECHEM, разработанная Вальдесом – Пересом.^{83–92} В основе ее алгоритма лежат принципы стехиометрии. Исходная версия программы оперировала не структурными формулами веществ, а их составами (брутто-формулами), в настоящее время программа позволяет распознавать и структуры веществ. Программа MECHEM принадлежит к программам формально-логического направления, поскольку в ней не используются никакие библиотеки обобщенных реакций. Генерация механизмов осуществляется комбинаторным перебором различных вариантов превращений веществ с использованием логических ограничений. В программе нет генераторов реакций как таковых. Происходит формирование «скелета» механизма реакции с последующей конкретизацией элементарных стадий и веществ. Модель формирования гипотез основана на следующих принципах.⁸⁵

1. Осуществляется комбинаторный поиск простейших механизмов (гипотез) с последовательным наращиванием числа стадий и веществ. Если механизмы с меньшим числом стадий и веществ не удовлетворяют всем введенным пользователем ограничениям или условию соблюдения элементного баланса во всех стадиях, то осуществляется переход к гипотезам с бóльшим числом стадий, а в рамках гипотез с одинаковым числом стадий — к гипотезам с бóльшим числом веществ. Таким образом, сложность гипотезы характеризуется числом веществ и стадий.

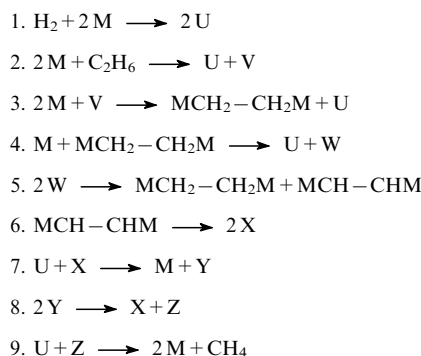
2. Сначала генерируется набор схем реакции, в которых вместо неизвестных интермедиатов, предлагаемых программой, фигурируют символы X, Y, Z и т.д. (wild cards). Затем эти символы заменяются формулами конкретных веществ. Их брутто-формулы вычисляют на основе соотношений элементного баланса (в последней версии программы МЕСНЕМ был использован алгоритм выявления структур веществ по их формулам,⁹⁰ при этом можно наложить ограничения на структуры веществ⁹³).

3. Молекулярность и число продуктов элементарной реакции не превышают 2; в одной и той же стадии одно и то же вещество не может одновременно выступать и в роли исходного вещества, и в роли продукта.

4. Механизмы генерируют так, что каждая последующая стадия не может содержать в качестве исходных веществ те, которые не были получены на предыдущих стадиях или не были заданы пользователем в качестве исходных. (Этот и некоторые другие принципы лежат в основе канонического представления механизмов.⁹¹)

Пользователь вводит в программу информацию об исходных реагентах и конечных продуктах, задает исходную форму катализатора (если он есть) и наблюдаемые интермедиаты, вводит ограничения, связанные с валентностью каждого атома, последовательностью появления интермедиатов в стадиях, возможностью (или невозможностью) наблюдать исходные вещества среди продуктов отдельных стадий, числом связей, изменяющих свой порядок и т.д. В программе используется большое число приемов (эвристик), позволяющих сделать алгоритм более эффективным.⁸⁹

С помощью программы МЕСЕМ были генерированы возможные механизмы реакции гидрогенолиза этана.^{88, 89} Были заданы следующие ограничения: 1) стехиометрия итоговой реакции соответствует уравнению $C_2H_6 + H_2 \rightarrow 2CH_4$; 2) один активный центр катализатора М может образовывать только одну связь (не обязательно одинарную); 3) в реакции есть интермедиаты с мостиковой структурой, MCH_2-CH_2M и $MCH-CHM$; 4) водород не образуется ни в одной из стадий; 5) все интермедиаты содержат хотя бы один центр катализатора М; 6) ни один из интермедиатов не содержит трех атомов углерода или металла; 7) в одной стадии образуется и рвется не более трех связей (независимо от их кратности); 8) интермедиат MCH_2-CH_2M должен появиться в последовательности стадий раньше, чем $MCH-CHM$. Кроме этих ограничений учитывалась также необходимость соблюдения элементного баланса в каждой стадии и условия, что валентности атомов водорода и углерода не превышают 1 и 4 соответственно. Среди прочих равных по сложности (т.е. с одинаковым числом стадий и веществ) была генерирована следующая схема гидрогенолиза этана:[†]



† Так как программа генерирует только бимолекулярные реакции (это один из принципов, заложенных в алгоритм формирования стадии), то, по-видимому, подразумевается, что $2M$ — это два связанных активных центра гетерогенного катализатора (M_2).

Здесь U, V, W, X, Y, Z — интермедиаты, которые были расшифрованы программой как HM, $\text{CH}_3\text{—CH}_2\text{M}$, $\text{CH}_2\text{—CHM}$, MCH, CH_2M и CH_3M соответственно. Стехиометрические числа стадий 1–9 равны соответственно 1, 1, 1, 2, 1, 1, 4, 2, 2, что приводит к итоговому уравнению, в котором все интермедиаты сокращаются. Интермедиаты MCH и CH_2M могут быть интерпретированы как карбиновые и карбеновые комплексы (на момент проведения этого исследования в программу еще не был заложен алгоритм идентификации порядков связей).

В целом программа MECHEM ориентирована на поиск гипотез, призванных объяснить результаты наблюдений (т.е. гипотезы следуют за наблюдениями). Как отмечалось ранее, методологически более корректна другая постановка задачи: какие варианты наблюдений следует ожидать? Отметим, что самым слабым местом этой программы является концепция выбора простейших механизмов. Разумно использовать критерий простоты на этапе дискриминации гипотез, а также в момент, когда экспериментальная проверка гипотез уже закончена, а оставшиеся гипотезы одинаково хорошо объясняют наблюдения и согласуются с общими представлениями о процессах данного типа. В этом случае, если не ясен дальнейший план экспериментов или если невозможно экспериментально дискриминировать гипотезы имеющимися методами, целесообразно применить критерии, характеризующие простоту механизма. Это могут быть критерии, основанные на подсчете числа стадий и веществ (как в MECHEM), или более сложные критерии.^{94–97}

Кроме рассмотренных здесь формально-логических методов получения реакционных сетей во многих работах генерацию механизмов реакции осуществляют путем совместного использования комбинаторных методов и стехиометрии,^{98–100} а также теории графов.¹⁰¹ На основе этих методов не было создано компьютерных программ или они не были достаточно хорошо описаны в литературе. Разработаны также методы «вырезания» из реакционных сетей механизмов, приводящих к заданным итоговым уравнениям.^{9,98,102–112} В этих методах предполагается, что известны все исходные, промежуточные и конечные вещества, а иногда и уравнения стадий. Поэтому, несмотря на полезность таких алгоритмов и программ при исследовании механизмов, их использование для генерации гипотез о механизмах химических реакций весьма ограничено.¹⁴

2. Программы эмпирического направления

Как указывалось выше, все программы, относящиеся к эмпирическому направлению, объединяет одно общее свойство: в них используются обобщенные реакции (transforms). Способ описания обобщенных реакций может быть при этом произвольным. Общий алгоритм состоит в узнавании фрагментов молекул, к которым могут быть применены те или иные обобщенные реакции (шаблоны реакций) с последующим генерированием конкретных реакций и структур веществ.

Одной из первых программ эмпирического направления является программа CAMEO.¹¹³ К моменту написания настоящего обзора опубликовано 25 статей. Здесь мы ссылаемся лишь на первую публикацию¹¹³ и на обзор¹¹⁴.

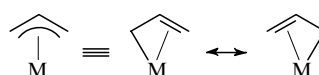
Программа CAMEO ориентирована на моделирование механизмов органических реакций. Характерно, что поиск каждого последующего фрагмента, к которому может быть применена та или иная обобщенная реакция, осуществляется в этой программе только в прямом направлении (при этом рассматриваются побочные маршруты и учитываются взаимодействия между интермедиатами).

Изначально задают исходные вещества и предполагаемые продукты реакции, условия протекания реакции, а также список модулей, которые необходимо задействовать. Программа содержит несколько модулей, каждый из которых

предназначен для описания определенного класса органических реакций. К настоящему моменту разработаны модули, предназначенные для исследования реакций, катализируемых основаниями и кислотами (реакций с участием нуклеофильных и электрофильных частиц); периклических реакций; окислительно-восстановительных реакций; реакций с участием свободных радикалов и цепных процессов; реакций карбеноидных частиц; реакций восстановления гидридами; «еновых» и «ретроеновых» синтезов. Молекулы представляют в виде специальных таблиц, в которых специфицированы атомы и связи.¹⁰⁸ В этих таблицах каждый атом имеет собственный порядковый номер, задан тип каждого атома (C, H, ...), его координаты на графическом дисплее и код заряда (цифры 1, 2 и 3 указывают на то, что при данном атоме должен стоять символ катиона, аниона или радикала соответственно; если атом нейтрален, то ему присваивается код 0), указано, между какими атомами есть связи и порядок этих связей. Каждой связи присвоен собственный номер и указано ее расположение относительно условной плоскости рисунка. Кроме того, учитываются правила отбора, включая правила, основанные на физико-химических принципах (оценках pK_a , ΔH_f , энергий HOMO–LUMO, теплот реакции и т.д.).

С помощью программы CAMEO можно определять наиболее вероятные места атаки активных частиц, проводить оценку сравнительной реакционной способности.

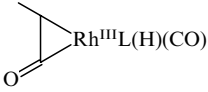
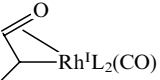
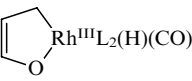
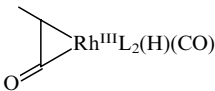
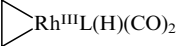
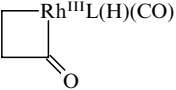
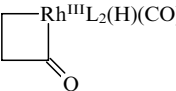
Программа TAMREAC^{67–69} ориентирована на исследователей в области металлоорганического катализа. К сожалению, авторы работ^{67–69} не опубликовали деталей алгоритма, поэтому трудно анализировать действительные возможности программы. Для компьютерного представления молекул в рамках этой программы используют таблицы связности атомов, в которых каждый атом характеризуют номером, символом (углерод, водород, металл, произвольный атом, произвольный электрофильный атом и т. п.). В таблицах также указано, с какими атомами и как связан данный элемент и каков его общий заряд (степень окисления). Кроме того, для описания η^2 -связей с металлами введены таблицы η^2 -связей, в которых указано, между какими атомами есть кратная связь и каким атомом она координирована. η^3 -Координация описывается с помощью двух изомерных структур.



В программе TAMREAC предусмотрена возможность генерации элементарных стадий, в каждой из которых фигурирует не более одного атома металла. Разработчики программы TAMREAC систематизировали большой массив элементарных согласованных реакций и оформили их в виде хорошо организованной библиотеки обобщенных реакций. В нее входят реакции окисления металла; восстановления металлического центра; комплексообразования, приводящие к увеличению координационного числа металла; диссоциации комплекса с образованием координационной вакансии; без изменения степени окисления и координационного числа металла; радикальные реакции и реакции переноса электрона. Библиотека состоит из файлов, каждый из которых содержит несколько обобщенных реакций, описанных в виде структурного преобразования реакционного фрагмента.

В программе TAMREAC заложен алгоритм распознавания лигандов, характеристик их донорных способностей и дентатности. На основе этих данных осуществляется подсчет числа валентных электронов атома металла и его степени окисления. Полученные сведения используются в критериях отбора. Например, автоматически отслеживается, чтобы число валентных электронов находилось в заданных пользователем пределах. В программу также заложен алгоритм для

Таблица 1. Список веществ, предлагаемых программой TAMREAC при генерации реакционной сети гидроформилирования этилена.⁶⁸

Номер соединения	Химическая формула	Номер соединения	Химическая формула
1	$\text{Rh}^{\text{I}}\text{L}_2(\text{H})(\text{CO})_2$	21	
2	$(\eta^2\text{-C}_2\text{H}_4)\text{Rh}^{\text{I}}\text{L}_2(\text{H})(\text{CO})_2$	22	$[\eta^2\text{-C}(\text{O})=\text{CHCH}_3]\text{Rh}^{\text{I}}\text{L}_2(\text{H})(\text{CO})$
3	$\text{Rh}^{\text{I}}\text{L}_2(\text{H})(\text{CO})_2$	23	$\text{O}=\text{CH}-\text{CH}(\text{CH}_3)-\text{Rh}^{\text{III}}\text{L}_2\text{H}_2\text{CO}$
4	$\text{C}_2\text{H}_5\text{Rh}^{\text{I}}\text{L}_2(\text{CO})_2$	24	
5	$\text{O}=\text{C}(\text{C}_2\text{H}_5)-\text{Rh}^{\text{I}}\text{L}_2(\text{CO})$	25	$\text{CH}_3\text{CH}=\text{CH}-\text{O}-\text{Rh}^{\text{I}}\text{L}_2(\text{CO})$
6	$\text{O}=\text{C}(\text{C}_2\text{H}_5)-\text{Rh}^{\text{III}}\text{L}_2\text{H}_2(\text{CO})$	26	
7	$\text{Rh}^{\text{I}}\text{L}_2(\text{H})(\text{CO})$	27	H_2
8	$\text{CH}_3\text{CH}=\text{Rh}^{\text{I}}\text{L}_2(\text{H})(\text{CO})_2$	28	$\text{HO}-\text{CH}=\text{CHCH}_2-\text{Rh}^{\text{I}}\text{L}_2(\text{CO})$
9		29	$\text{HO}-\text{CH}=\text{CHCH}_2-\text{Rh}^{\text{III}}\text{L}_2\text{H}_2(\text{CO})$
10	$[\eta^2\text{-C}(\text{O})=\text{CHCH}_3]-\text{Rh}^{\text{I}}\text{L}_2(\text{H})(\text{CO})$	30	$\text{CH}_3\text{CH}=\text{CH}-\text{O}-\text{Rh}^{\text{III}}\text{L}_2\text{H}_2(\text{CO})$
11	$\text{O}=\text{CH}-\text{CH}(\text{CH}_3)-\text{Rh}^{\text{I}}\text{L}_2\text{CO}$	31	
12		32	$[\eta^2\text{-CH}_2=\text{CH}-\text{C}(\text{O})]\text{Rh}^{\text{I}}\text{L}_2(\text{H})(\text{CO})$
13		33	$\text{CH}_3\text{CH}=\text{CH}-\text{OH}$
14		34	$\text{O}=\text{CHCH}_2\text{CH}_2-\text{Rh}^{\text{I}}\text{L}_2(\text{CO})$
15	L	35	$\text{O}=\text{CHCH}_2\text{CH}_2-\text{Rh}^{\text{III}}\text{L}_2\text{H}_2(\text{CO})$
16	$(\eta^2\text{-C}_2\text{H}_4)\text{Rh}^{\text{I}}\text{L}_2(\text{H})(\text{CO})_2$	36	
17	$\text{C}_2\text{H}_5\text{Rh}^{\text{I}}\text{L}_2(\text{CO})_2$	37	C_2H_4
18	$\text{C}_2\text{H}_5\text{Rh}^{\text{III}}\text{L}_2(\text{H})_2(\text{CO})_2$	38	C_2H_6
19	CO	39	$\text{CH}_2=\text{CH}-\text{CH}=\text{O}$
20	$\text{CH}_3\text{CH}=\text{Rh}^{\text{I}}\text{L}_2(\text{H})(\text{CO})_2$	40	$\text{C}_2\text{H}_5\text{CHO}$

Примечание. Исходные вещества $\text{Rh}^{\text{I}}\text{L}_2(\text{H})(\text{CO})_2$ ($\text{L} = \text{PPh}_3$), C_2H_2 , CO , H_2 .

распознавания геометрии каждого комплекса. Разработчики TAMREAC создали систему правил отбора, на основе которых с использованием данных о типе превращения (восстановительное элиминирование, внедрение, реакции образования металлоциклов и т.д.), координационном числе металла, числе валентных электронов атома металла и геометрии комплекса дается экспертная оценка возможности протекания реакции. Эта информация, формализованная в виде матриц отбора, обычно очень полезна для исследователя. Однако следует отметить, что данных, на основе которых делается заключение, явно недостаточно, поэтому такая оценка может быть лишь принята во внимание, но возможности ее использования в рамках конкретных исследований весьма ограничены. Необходимо также учесть, что указанные выше правила отбора обычно вырабатываются исключительно на основе экспериментальных наблюдений, которые часто не имеют достаточно надежного теоретического обоснования. Эти правила непрерывно обновляются и, вообще говоря, не имеют «силы закона». Исследуя конкретную систему, каждый исследователь, в конечном счете, имеет право самостоятельно решать вопрос о необходимости учета экспертных рекомендаций. Поэтому наиболее приемлем вариант, когда пользователь (исследователь) сам имеет

возможность формулировать правила отбора по различным предлагаемым программой схемам.

Программа TAMREAC была тестирована на большом количестве конкретных реакций, и в ряде случаев получены весьма интересные результаты. Например, эту программу использовали для моделирования реакционной сети гидроформилирования этилена в системе, содержащей $\text{HRh}(\text{CO})_2\text{L}_2$ ($\text{L} = \text{PPh}_3$).⁶⁸ В качестве исходных веществ были заданы этилен, H_2 , CO и $\text{HRh}(\text{CO})_2\text{L}_2$. Из результатов, приведенных в табл. 1 и 2, видно, что программа «синтезировала» довольно интересные побочные продукты, поиск которых в реальной каталитической системе поможет подтвердить или отвергнуть некоторые из предложенных маршрутов.

На основе модели Дугунджи–Уги и метода групповых матриц (аналогичных тем, которые были использованы в программе GRACE⁸¹) в Институте катализа СО РАН был разработан комплекс программ КОМСИКАТ для исследования механизмов каталитических реакций.^{59–64} В целом идеология программы КОМСИКАТ близка к реализованной в программе GRACE, но в ней используются не все комбинаторно возможные групповые матрицы, а только те, которые заданы. Поэтому ее следует отнести к эмпиричес-

Таблица 2. Реакционная сеть гидроформилирования этилена, генерированная программой TAMREAC.⁶⁸

№	Реакция ^a	№	Реакция ^a	№	Реакция ^a	№	Реакция ^a
1	1 → 3 + 15	23	7 + 39 → 32	45	13 → 12	67	23 → 7 + 40
2	1 → 7 + 19	24	7 + 40 → 6	46	13 + 15 → 14	65	23 → 11 + 27
3	1 + 37 → 2	25	7 + 40 → 23	47	14 → 7 + 36	69	24 → 11
4	2 → 1 + 37	26	8 → 4	48	14 → 13 + 15	70	24 → 25
5	2 → 4	27	8 → 9	49	15 + 16 → 2	71	25 + 27 → 30
6	2 → 16 + 15	28	8 → 20 + 15	50	16 → 3 + 37	72	25 → 24
7	3 + 15 → 1	29	9 → 10	51	16 → 9	73	25 → 26
8	3 + 37 → 16	30	9 → 11	52	16 → 12	74	26 → 7 + 31
9	4 → 2	31	9 → 16	53	16 → 20	75	26 → 25
10	4 → 5	32	9 → 21 + 15	54	17 + 15 → 4	76	26 → 28
11	4 → 8	33	10 → 5	55	17 → 20	77	28 → 26
12	4 → 17 + 15	34	10 → 9	56	17 + 27 → 18	78	28 + 27 → 29
13	5 → 4	35	10 → 11	57	18 → 3 + 38	79	29 → 7 + 33
14	5 → 10	36	10 → 22 + 15	58	18 → 17 + 27	80	29 → 28 + 27
15	5 + 27 → 6	37	11 → 9	59	20 + 15 → 8	81	30 → 7 + 33
16	6 → 5 + 27	38	11 → 10	60	20 → 16	82	30 → 25 + 27
17	6 → 7 + 40	39	11 → 24	61	20 → 17	83	32 → 7 + 39
18	7 + 19 → 1	40	11 → 32	62	20 → 21	84	32 → 11
19	7 + 31 → 26	41	11 + 27 → 23	63	21 + 15 → 9	85	32 → 34
20	7 + 33 → 30	42	12 → 9	64	21 → 22	86	34 + 27 → 35
21	7 + 33 → 29	43	12 → 13	65	22 + 15 → 10	87	34 → 32
22	7 + 36 → 14	44	12 → 16	66	22 → 21	88	35 → 7 + 40
						89	35 → 34 + 27

^a Номера соединений указаны в табл. 1.

кому направлению. Авторам проекта КОМСИКАТ на основе литературных данных удалось выделить характерные структурные типы элементарных преобразований с участием катализатора. В частности, авторы допустили, что процессы, протекающие в системе CO + H₂ с участием металла М, могут быть описаны пятью основными типами элементарных преобразований. В табл. 3 приведены использованные реакционные фрагменты и соответствующие блоки *r*-матриц (матрицы индексной группы). Диагональные элементы матриц индексной группы характеризуют изменения числа валентных электронов.

В программе КОМСИКАТ выбран энергетический критерий отбора маршрутов. Генерация реакционной сети происходила следующим образом. На первом этапе генерировались все возможные ансамбли молекул, которые могут быть получены на основе допустимых элементарных преобразований исходного ансамбля молекул. Из ансамблей молекул, полученных на первом этапе, отбирались только те, которые в пределах заданной точности соответствовали максимальной энергии атомизации (т.е. наиболее устойчивые ансамбли молекул),[†] а остальные отбрасывались. Для каждого промежуточного ансамбля молекул, оставшегося

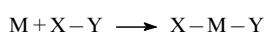
после такого отбора, находили новый набор ансамблей молекул, полученный путем допустимых элементарных преобразований, и уже среди них проводили отбор наиболее устойчивых. Генерация «ветви» реакционной сети прекращалась после выделения металла-катализатора в свободном виде (когда ансамбль молекул содержал свободный атом металла). В результате реакционная сеть состояла из промежуточных ансамблей молекул, которые соответствовали путям превращения их предшественников в наиболее устойчивые соединения. Энергию атомизации рассчитывали по методу взаимодействующих связей⁶⁰ с применением полуэмпирических параметров энергий связи атомов, представленных в ансамбле молекул. Для связей с участием атомов С, Н и О использовали параметры, полученные из экспериментальных термодинамических данных о теплотах образования соответствующих соединений. Так, параметр E_{C-O} находили из данных о теплотах образования альдегидов и кетонов, параметр E_{C-H} — из теплот образования углеводородов и спиртов, параметр E_{O-H} — из теплот образования спиртов и т.д.

На следующем этапе осуществляли подбор оптимального катализатора. С этой целью варьировали характеризующие металл параметры E_{M-x} (E_{M-C} , E_{M-H} , E_{M-O}) для каждого промежуточного ансамбля каждого маршрута в реакционной сети (в каждой линейной последовательности ансамблей молекул от исходного к конечному). Так как параметры E_{M-x} варьировали с некоторым шагом, то энергия атомизации каждого промежуточного ансамбля в маршруте расщеплялась на множество значений. Для каждого ансамбля выбирали то значение энергии атомизации, которое соответствовало условию равномерности распреде-

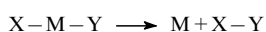
[†] Наиболее устойчивым считался ансамбль молекул, для которого сумма энергий атомизации входящих в него частиц была наибольшей. Множественность устойчивых ансамблей молекул возникает вследствие того, что атом металла заранее не известен, а известен только интервал возможных значений энергий атомизации, полученный путем варьирования полуэмпирических параметров энергий связей.⁶⁰

Таблица 3. Реакционные фрагменты и матрицы индексной группы, используемые в программе КОМСИКАТ.⁶⁴

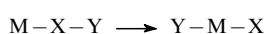
Схема перераспределения связей	Матрица индексной группы
--------------------------------	--------------------------



	M	X	Y
M	-2	1	1
X	1	0	-1
Y	1	-1	0



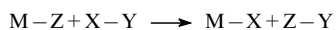
	M	X	Y
M	2	-1	-1
X	-1	0	1
Y	-1	1	0



	M	X	Y
M	-1	0	1
X	0	1	-1
Y	1	-1	0



	M	X	Y
M	-1	0	1
X	0	1	-1
Y	1	-1	0



	M	X	Y	Z
M	0	-1	1	0
X	-1	0	0	1
Y	1	0	0	-1
Z	0	1	-1	0

ления энергии по стадиям каталитического процесса. Другими словами, выбирали наиболее сглаженную ломаную линию, отвечающую энергиям атомизации промежуточных ансамблей в маршруте (рис. 1). Вычисленные для этой ломаной значения полуэмпирических параметров энергий связей сравнивали с табличными значениями для конкретных металлов. Металл считался оптимальным, если табличные значения параметров энергий связей для этого металла оказывались наиболее близкими к расчетным (т.е. полученным для наиболее сглаженной ломаной).

Так, из сопоставления расчетных значений параметров образования этиленгликоля в реакции между CO и H₂ в присутствии металла-катализатора и табличных значений, следует, что для этой реакции оптимальными катализаторами являются Rh, Co, Ir, Ru и V. Были найдены области значений параметров E_{M-X}, оптимальных для образования каждого из продуктов реакции гидрирования CO (метана, метанола, этана, этанола, уксусной кислоты, метилформата, этиленгликоля), и определены соответствующие им металлы. Из пересечения этих областей значений параметров были найдены величины E_{M-X} и металлы (Mn, Fe, Co, Ni, Ru, Rh), оптимальные для образования всех C₁ и C₂ продуктов.

Балтанас и Фромент¹¹⁵ разработали программу генерации реакционных сетей для решения частных задач, таких как исследование каталитической гидроизомеризации и гидро-

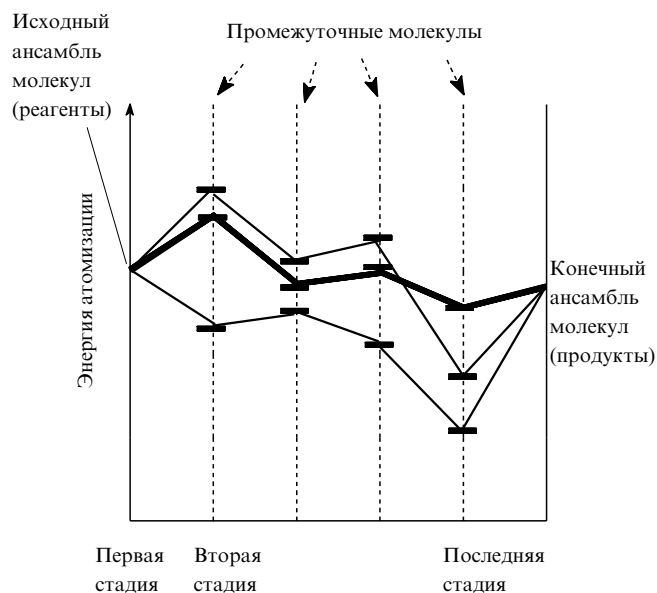


Рис. 1. Расщепление уровней при варьировании параметров E_{M-X}. Оптимальными являются параметры E_{M-X}, которые соответствуют наиболее «сглаженной» ломаной, показанной на рисунке более жирной линией. На пунктирных линиях показаны расщепленные уровни энергии атомизации ансамбля молекул, полученного на соответствующей стадии.

крекинга. В этой программе используется пять обобщенных реакций, которые последовательно применяются к различным веществам. Эти реакции предназначены только для исследования данных конкретных каталитических процессов. Описание веществ и реакций дано на языке булевых матриц, которые не учитывают кратности связей. С помощью данной программы в автоматическом режиме были выведены кинетические уравнения для полученной сети превращений. Основным недостатком программы является узость ее приложений.

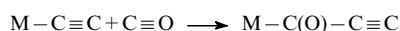
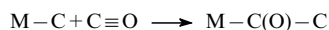
Группой Темкина (Московская государственная академия тонкой химической технологии им. М.В.Ломоносова) ведется разработка компьютерной программы ChemNet для выдвижения гипотез о механизмах каталитических реакций. Эта программа обсуждалась лишь в небольшом числе публикаций, часть из которых малоизвестна.^{14,116} Поэтому мы рассмотрим ее более подробно.

В середине 80-х годов была написана первая версия программы на языке ФОРТРАН IV для компьютеров, которые в настоящее время безнадежно устарели.¹¹⁷ Недавно разработана новая версия, рассчитанная на современные персональные компьютеры типа IBM PC и использующая основные идеи старой программы.

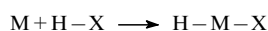
В качестве исходных данных вводится список веществ, участвующих в реакции в качестве реагентов и компонентов системы, и список элементарных преобразований (обобщенных реакций). Вещества задаются пользователем в виде структурных формул и преобразуются в записи, имеющие специальный формат. Часть записи содержит верхнюю треугольную матрицу смежности атомов, а другая часть — различные характеристики атомов (заряды, степени окисления и др.), кратности связей, а также многоцентровые связи. Элементарные преобразования являются аналогами химических реакций и описывают характерные черты происходящих изменений. При этом степень детализации реакции определяет пользователь. Чем более детально описано элементарное преобразование, тем меньше список веществ, к которым это элементарное преобразование применимо. Например, два элементарных преобразования

Таблица 4. Список веществ, предлагаемых программой ChemNet при генерации реакционной сети с использованием десяти элементарных преобразований и пяти исходных веществ (PdCl₂, CuCl, MeOH, HC≡CMe, CO).

Номер соединения	Химическая формула	Номер соединения	Химическая формула	Номер соединения	Химическая формула
1 ^a	PdCl ₂	17	MeC≡CPdCOOMe	33	MeOPdCH=C(Me)COOMe
2 ^a	CuCl	18	Pd(C≡CMe) ₂	34	MeOPdC(Me)=CHCOOMe
3 ^a	MeOH	19	ClPdCOC≡CMe	35	ClPdCH=C(Me)COC≡CMe
4 ^a	HC≡CMe	20	MeOPdC≡CMe	36	ClPdC(Me)=CHCOC≡CMe
5 ^a	CO	21	MeC≡CC(O)PdCOOMe	37	ClPdH
6	MeOPdCl	22	MeC≡CPdCOC≡CMe	38	MeOPdH
7	HCl	23	Pd(COC≡CMe) ₂	39	MeOC(O)OMe
8	MeOCu	24	MeOC(O)–C(O)OMe	40	HPdCOOMe
9	(MeO) ₂ Pd	25	Pd ⁰	41	HPdC≡CMe
10	ClPdCOOMe	26	MeC≡CCOOMe	42	HPdC(O)C≡CMe
11	CuCOOMe	27	MeC≡C–C≡CMe	43	PdH ₂
12	MeOPdCOOMe	28	MeOC(O)C(O)C≡CMe	44	HPdCH=C(Me)COOMe
13	Pd(COOMe) ₂	29	MeC≡CC(O)C≡CMe	45	HPdC(Me)=CHCOOMe
14	CuC≡CMe	30	MeC≡CC(O)C(O)C≡CMe	46	HPdCH=C(Me)COC≡CMe
15	ClPdC≡CMe	31	ClPdCH=C(Me)COOMe	47	HPdC(Me)=CHCOC≡CMe
16	MeOPdC≡CMe	32	ClPdC(Me)=C(H)COOMe		

^a Исходное вещество.

описывают одну и ту же реакцию с различной степенью детализации. Разница состоит лишь в том, что в первом случае программа «внедрит» СО во все соединения, содержащие связь М–С, а во втором — только в этильные производные металлов. При этом в первом случае программа будет «внедрять» СО также и во фрагменты типа М–С(О) до тех пор, пока не начнут действовать ограничения на число атомов в веществах. Предусмотрена запись элементарного преобразования как в символьном виде, так и с использованием обозначений конкретных атомов. Если элементарное преобразование записано в символьном виде, то необходимо определить также возможные значения каждого символа. Например, обобщенная реакция



записана в символьном виде. Символы М и Х необходимо расшифровать, указав, какие атомы металла М и какие атомы Х могут фигурировать в данной реакции.

Проведя генерацию реакционной сети «начерно», пользователь может изменить степень детализации в записи стадий, исключив тем самым заведомо ложные направления атаки тех или иных веществ.

Для описания веществ и элементарных преобразований разработан специальный единый формат.¹¹⁸ Вещества и реакции хранятся в памяти компьютера в виде специальных кодов, описание которых выходит за рамки настоящего обзора.

В программе ChemNet имеется множество правил отбора, которые рекомендуются пользователю.

1. В элементарной стадии изменение числа валентных электронов и степени окисления атома металла не должно по модулю превышать 2, что находится в согласии с правилами Толмана¹¹⁹.

2. Молекулярность элементарных реакций не должна превышать 2.

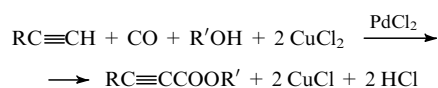
3. Пользователь задает максимальное число валентных электронов, максимальное координационное число и интервал возможных степеней окисления для каждого атома металла (по умолчанию используются величины, хранящиеся в специальном файле).

4. Пользователь задает максимальное число атомов и/или максимальные числа атомов определенного типа (С, Н, О, ...) в образующихся продуктах.

Генерация реакционной сети ограничивается также набором стадий, которые пользователь задает сам или выбирает из библиотеки. Кроме того, генерацию реакционной сети неявно ограничивает степень детализации в записи стадий.

После задания исходных веществ и элементарных преобразований программа осуществляет перебор исходных веществ, которые рассматриваются в качестве кандидатов на участие в каждом из элементарных преобразований. После того как для каждого элементарного преобразования найдены все сочетания веществ, к которым оно применимо, а новые вещества сохранены в памяти компьютера, программа добавляет все новые вещества к списку исходных веществ и повторяет всю процедуру заново. Комбинации веществ и стадий, проверенные на предыдущих этапах, заново не проверяются. Процесс останавливается, когда каждое вещество (исходное или полученное в ходе генерации сети) исследовано на возможность участия в каждом элементарном преобразовании, а новые вещества не образуются из-за действия ограничений. Общий алгоритм работы программы можно представить в виде блок-схемы, показанной на рис. 2.

Одним из тестов для проверки работы программы была реакция получения эфира алкилкарбоновой кислоты в системе CuCl₂–CuCl–PdCl₂–R'OH.



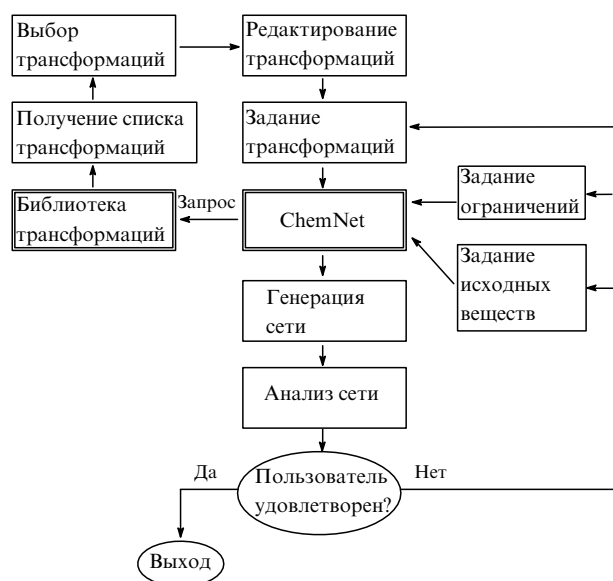


Рис. 2. Блок-схема интерактивной работы пользователя и программы ChemNet.

Были заданы следующие элементарные преобразования:

- $MCl + ROH \rightarrow MOR + HCl$ ($M = Pd, Cu; R = C$)
- $MOR + CO \rightarrow MCOOR$ ($M = Pd, Cu; R = C$)
- $MCl + RC \equiv CH \rightarrow MC \equiv CR + HCl$ ($M = Cu; R = C$)
- $CuC \equiv CR + Pd - Cl \rightarrow CuCl + PdC \equiv CR$ ($R = C$)
- $PdC \equiv CR + CO \rightarrow Pd - C(O) - C \equiv CR$ ($R = C$)
- $C - M - C \rightarrow M + C - C$ ($M = Pd$)

- $M - C(O) - + C \equiv C \rightarrow M - C = C - C(O) -$ ($M = Pd$)
- $M - C = CH \rightarrow C \equiv C + M - H$ ($M = Pd$)
- $Pd - C(O) - + ROH \rightarrow PdH + RO - C(O) -$ ($R = C$)
- $CuC + C - Pd - X \rightarrow CuX + C - Pd - C$ ($X = Cl$)

Были использованы следующие ограничения: максимальное число атомов в получаемых молекулах должно быть не больше 20, максимальная степень окисления металла в генерируемых продуктах для меди равна 1+, а для палладия — 2+, максимальные координационные числа Cu и Pd составляют соответственно 1 и 2.

Результаты генерации реакционной сети представлены в табл. 4 и 5. Эти результаты могут быть использованы экспериментатором при планировании дискриминирующих экспериментов.

IV. Заключение

Подводя итоги, отметим, что существует большое число компьютерных программ, которые дают возможность выдвигать гипотезы о механизмах сложных реакций и осуществлять генерацию реакционных сетей. Анализ программ, рассмотренных в настоящем обзоре, позволяет указать на несколько типичных присущих им недостатков.

1. Невозможность получения полных данных о превращениях интермедиатов, побочных реакциях и продуктах из-за необходимости жесткого задания числа атомов, входящих в изомерные ансамбли, и из-за билатеральности поиска (RAIN, программа Дозморова, GRACE).

2. Использование критерия простоты в качестве фактора отбора (MECNEM).

3. Узость спектра приложений (программа Фромента и Балтанаса).

Объединяет большинство рассмотренных программ то, что ключевым моментом при разработке методов выдвижения гипотез является применение эвристических правил,

Таблица 5. Реакционная сеть, генерированная программой ChemNet на основе заданных десяти элементарных преобразований (ЭП) и пяти исходных веществ ($PdCl_2$, $CuCl$, $MeOH$, $HC \equiv CMe$, CO).

ЭП	Реакция ^a	ЭП	Реакция ^a	ЭП	Реакция ^a	ЭП	Реакция ^a
1	3 + 1 → 6 + 7	5	22 + 5 → 23	9	10 + 3 → 37 + 39	10	10 + 3 → 7 + 12
1	3 + 2 → 7 + 8	5	13 → 24 + 25	9	12 + 3 → 38 + 39	1	15 + 3 → 7 + 16
1	6 + 3 → 7 + 9	6	17 → 25 + 26	9	13 + 3 → 39 + 40	1	19 + 3 → 7 + 20
1	6 + 5 → 10	6	18 → 25 + 27	9	17 + 3 → 39 + 41	1	31 + 3 → 7 + 33
2	8 + 5 → 11	6	21 → 25 + 28	9	19 + 3 → 26 + 37	1	32 + 3 → 7 + 34
2	9 + 5 → 12	6	22 → 25 + 29	9	20 + 3 → 26 + 38	1	37 + 3 → 7 + 38
2	12 + 5 → 13	6	23 → 25 + 30	9	21 + 3 → 26 + 40	1	16 + 5 → 17
2	4 + 2 → 7 + 14	6	10 + 4 → 31	9	21 + 3 → 39 + 42	2	20 + 5 → 21
3	14 + 1 → 2 + 15	7	10 + 4 → 32	9	22 + 3 → 26 + 41	2	38 + 5 → 40
4	14 + 6 → 2 + 16	7	12 + 4 → 33	9	23 + 3 → 26 + 42	2	37 + 14 → 2 + 41
4	14 + 10 → 2 + 17	7	12 + 4 → 34	9	40 + 3 → 39 + 43	4	41 + 5 → 42
4	15 + 14 → 2 + 18	7	19 + 4 → 35	9	42 + 3 → 26 + 43	5	40 + 4 → 44
4	15 + 5 → 19	7	19 + 4 → 36	9	11 + 10 → 2 + 13	7	40 + 4 → 45
5	16 + 5 → 20	7	32 → 26 + 37	10	15 + 11 → 2 + 17	7	42 + 4 → 46
5	17 + 5 → 21	8	34 → 26 + 38	10	19 + 11 → 2 + 21	7	42 + 4 → 47
5	18 + 5 → 22	8	36 → 29 + 37	10	19 + 14 → 2 + 22	7	45 → 26 + 43
						8	47 → 29 + 43

^a Номера соединений указаны в табл. 4.

которые в разной степени используются на различных этапах работы программ.

1. Использование библиотек трансформаций (генераторов реакций), полученных в результате обобщения имеющихся в литературе сведений о реакциях (TAMREAC, КОМСИКАТ, ChemNet и др.)

2. Использование в той или иной степени формализованных правил для прогнозирования вероятности протекания реакции (TAMREAC, SAMEO).

3. Использование эвристических правил для оптимизации, ускорения и ограничения поиска (RAIN, MECHEM).

4. Оценка термодинамических и других характеристик, полученных на основе полуэмпирических или полностью эмпирических методов, которые в свою очередь позволяют оценить возможность протекания отдельных стадий или целых маршрутов реакции (КОМСИКАТ, программа Фромента и Балтанаса, SAMEO).

Во всех случаях пользователю важно знать, какие эвристические правила заложены в программу, каковы возможности их модифицирования, можно ли отказаться от каких-либо из них. Желательно, чтобы такая же возможность существовала и в отношении действующих ограничений.

В программах эмпирического направления используются разные уровни общности элементарных реакций. В некоторых программах, например, в программе ChemNet, можно применять одновременно несколько уровней общности (т.е. реакции, в которых охарактеризовано только преобразование фрагментов, и реакции, записанные более детально).

Поскольку различные программы основаны на различных методах и эвристических правилах, иногда целесообразно комбинировать их в рамках единого пакета. В частности, нами ведется работа по созданию пакета программ, который объединил бы программы ChemNet и КОМСИКАТ. По нашему мнению, подобное объединение может обеспечить дополнительную гибкость при выборе средств для выдвижения гипотез.

Литература

1. K.R.Popper. *The Logic of Scientific Discovery*. Hutchinson, London, 1959
2. K.R.Popper. *Objective Knowledge*. Oxford University Press, Oxford, 1972
3. K.R.Popper. *Conjectures and Refutations: The Growth of Scientific Knowledge*. Harper and Row, New York, 1968
4. Л.Б.Баженов, Л.Х.Самородникий. *Вопр. философии*, **6**, 93 (1971)
5. D.Hodson. *Educ. Chem.*, **19**, 112 (1982)
6. В.В.Налимов. *Зав. лаб.*, **44**, 325 (1978)
7. Л.С.Полак, В.Я.Гольденберг, А.А. Левицкий. *Вычислительные методы в химической кинетике*. Наука, Москва, 1988
8. В.Н.Писаренко, А.Г.Погорелов. *Планирование кинетических исследований*. Наука, Москва, 1969
9. В.Г.Горский. *Планирование кинетических экспериментов*. Наука, Москва, 1984
10. *The ACS Style Guide: A Manual for Authors and Editors*. (Ed. V.S.Dodd). American Chemical Society, Washington, DC, 1986
11. Д.Пойа. *Математика и правдоподобные рассуждения*. Наука, Москва, 1975
12. О.Н.Темкин, С.М.Брайловский, Л.Г.Брук. В кн. *Химическая кинетика в катализе: Кинетические модели жидкофазных реакций*. (Под ред. С.Л.Кипермана). ИХФ, ИОХ АН СССР, Черноголовка, 1985. С.59
13. О.Н.Темкин, Л.Г.Брук, Д.Бончев. *Теорет. и эксперим. химия*, **24**, 282 (1988)
14. О.Н.Темкин, Л.Г.Брук, А.В.Зейгарник. *Кинетика и катализ*, **34**, 445 (1993)
15. *Chemical Applications of Graph Theory*. (Ed. A.T.Balaban). Academic Press, London, 1976
16. *Chemical Applications of Topology and Graph Theory*. (Ed. R.B.King). Elsevier, Amsterdam, 1983
17. *Применение теории графов в химии*. (Под ред. Н.С.Зефирова, С.И.Кучанова). Наука, Новосибирск, 1988
18. N.Trinajstić. *Chemical Graph Theory*. CRC Press, Boca Raton, FL, 1992
19. *Mathematical Chemistry: Graph Theory and Its Applications to Chemistry*. V.1, 2. (Eds D.Bonchev, D.H.Rouvray). Gordon and Breach, Chichester, 1991
20. *Computational Chemical Graph Theory* (Ed. D.H.Rouvray). Nova Publications, New York, 1990
21. А.В.Зейгарник, О.Н. Темкин. *Кинетика и катализ*, **35**, 691 (1994)
22. О.Н.Темкин, А.В.Зейгарник, Д.Г.Бончев, *J. Chem. Inf. Comput. Sci.*, **35**, 729 (1995)
23. А.Т.Балабан. *Graph Theoretical Approaches to Chemical Reactivity*. (Eds D.Bonchev, O.Mekenyan). Kluwer, Dordrecht, 1994. P.137
24. О.Н.Темкин, Г.К.Шестаков, Ю.А.Треглер. *Ацетилен: Химия. Механизмы реакций. Технология*. Химия, Москва, 1991
25. Т.Т.Зунг, Л.Г.Брук, О.Н.Темкин. *Изв. АН, Сер. хим.*, 1806 (1993)
26. T.J.Katz, J.McGinnis. *J. Am. Chem. Soc.*, **99**, 1903 (1977)
27. С.М.Брайловский, О.Н.Темкин, А.С.Костюшин, К.Ю.Одинцов. *Кинетика и катализ*, **31**, 1371 (1990)
28. Л.В.Щельцын, С.М.Брайловский, О.Н.Темкин. *Кинетика и катализ*, **31**, 1361 (1990)
29. I.Ugi, J.Bauer, K.Bley, A.Dengler, A.Dietz, E.Fontain, B.Gruber, R.Herges, M.Knauer, K.Reitsam, N.Stein. *Angew. Chem., Int. Ed. Engl.*, **32**, 201 (1993)
30. I.Ugi, J.Bauer, C.Blomberg, J.Brandt, A.Dietz, E.Fontain, B.Gruber, A. v. Scholley-Pfab, A.Senff, N.Stein. *J. Chem. Inf. Comput. Sci.*, **34**, 3 (1994)
31. Н.С.Зефирова, Е.В.Гордеева. *Успехи химии*, **56**, 1753 (1987)
32. R.Barone, M.Chanon. In *Computer Aids to Chemistry*. (Eds G.Vernin, M.Chanon). Ellis-Horwood, Chichester, 1986. P.19
33. M.Bersohn, A.Esack. *Chem. Rev.*, **76**, 269 (1976)
34. E.J.Corey, W.T.Wipke, R.D.Crammer, W.J.Howe. *J. Am. Chem. Soc.*, **94**, 431 (1972)
35. W.T.Wipke, H.Braun, G.Smith, F.Choplin, W.Sieber. *ACS Symp. Ser.*, **61**, 97 (1977)
36. E.J.Corey, A.K.Long, S.D.Rubenstein. *Science*, **228**, 408 (1985)
37. M.Bersohn. *Bull. Chem. Soc. Jpn.*, **45**, 1897 (1972)
38. J.Dugunji, I.Ugi. *Top. Curr. Chem.*, **39**, 19 (1973)
39. C.Jochum, J.Gasteiger, I.Ugi. *Angew. Chem., Int. Ed. Engl.*, **19**, 495 (1980)
40. C.Jochum, J.Gasteiger, I.Ugi. *Z. Naturforsch., B Chem. Sci.*, **37**, 1205 (1982)
41. I.Ugi, M.Wochner. *J. Mol. Struct. THEOCHEM*, **42**, 229 (1988)
42. M.Wochner, J.Brandt, A.v.Scholley, I.Ugi. *Chimia*, **42**, 217 (1988)
43. I.Ugi, M.Wochner, E.Fontain, J.Bauer, B.Gruber, R.Karl. In *Concepts and Applications of Molecular Similarity*. (Eds M.A.Johnson, G.M.Maggiora). Wiley, New York, 1990. P.239
44. J.Koča, M.Kratochvíl, V.Kvasnička, L.Matyska, J.Pospíchal. *Synthon Model of Organic Chemistry and Synthesis Design*. (Lect. Notes Chem. V.51). Springer-Verlag, Berlin, 1989
45. V.Kvasnička, M.Kratochvíl, J.Koča. *Collect. Czech. Chem. Commun.*, **48**, 2284 (1983)
46. J.Koča. *Collect. Czech. Chem. Commun.*, **53**, 1007 (1988)
47. J.Koča. *Collect. Czech. Chem. Commun.*, **53**, 3108 (1988)
48. J.Koča. *Collect. Czech. Chem. Commun.*, **53**, 3119 (1988)
49. J.Koča. *J. Math. Chem.*, **3**, 73 (1989)
50. J.Koča. *J. Math. Chem.*, **3**, 91 (1989)
51. V.Kvasnička, J.Pospíchal. *J. Math. Chem.*, **5**, 309 (1990)
52. V.Kvasnička, J.Pospíchal. *J. Math. Chem.*, **3**, 161 (1989)
53. V.Kvasnička, J.Pospíchal. *Int. J. Quantum Chem.*, **38**, 253 (1990)
54. J.B.Hendrickson, C.A.Parks. *J. Chem. Inf. Comput. Sci.*, **32**, 209 (1992)
55. С.В.Дозморков. *React. Kinet. Catal. Lett.*, **19**, 289 (1982)
56. S.H.Bertz. In *Chemical Applications of Topology and Graph Theory*. (Ed. R.B.King). Elsevier, Amsterdam, 1983. P.206
57. S.H.Bertz, W.C.Herndon. *ACS Symp. Ser.*, **306**, 169 (1986)
58. J.Gasteiger, M.G.Hutchings, P.Löw, H.Saller. *ACS Symp. Ser.*, **306**, 258 (1986)
59. В.А.Лихолобов, Л.И.Майер, Н.Н.Булгаков, А.В.Федотов, Л.Г.Штейнгауэр. В кн. *Материалы V Междунар. симп. по связи между гомогенным и гетерогенным катализом*. Новосибирск, 1986. С.231

60. L.I.Shtokolo, L.G.Shteingauer, V.A.Likholobov, A.V.Fedotov. *React. Kinet. Catal. Lett.*, **26**, 227 (1984)
61. V.A.Likholobov, L.I.Mayer, N.N.Bulgakov, A.V.Fedotov, L.G.Shteingauer. In *Homogeneous and Heterogeneous Catalysis*. (Eds Yu.Yermakov, V.Likholobov). 1986. P.229
62. L.G.Shteingauer, L.I.Mayer, N.N.Bulgakov, A.V.Fedotov, V.A.Likholobov. *React. Kinet. Catal. Lett.*, **36**, 139 (1988)
63. Л.Г.Заболотная. Дис. канд. физ.-мат. наук. ИК СО РАН, Новосибирск, 1989
64. Л.И.Майер. Дис. канд. хим. наук. ИК СО РАН, Новосибирск, 1990
65. J.Gasteiger. *Tetrahedron*, **35**, 1419 (1979)
66. J.Gasteiger, P.Jacob, U.Strauss. *Tetrahedron*, **35**, 139 (1979)
67. R.Barone, M.Chanon, M.L.H.Green. *J. Organomet. Chem.*, **185**, 85 (1980)
68. I.Theodosiou, R.Barone, M.Chanon. *J. Mol. Catal.*, **32**, 27 (1985)
69. I.Theodosiou, R.Barone, M.Chanon. *Adv. Organomet. Chem.*, **26**, 165 (1986)
70. D.Forster. *Adv. Organomet. Chem.*, **17**, 255 (1979)
71. J.F.Roth, J.H.Craddock, A.Hersham, F.E.Paulik. *Chem. Technol.*, **1**, 600 (1971)
72. P.J.Collman, L.S.Hegedus, J.R.Norton, P.Finke. *Principles and Applications of Organotransition Metal Chemistry*. University Sci. Books, Mill Valley, CA, 1987
73. E.Fontain, J.Bauer, I.Ugi. *Chem. Lett.*, 37 (1987)
74. E.Fontain, K.Reitsam. *J. Chem. Inf. Comput. Sci.*, **31**, 97 (1991)
75. E.Fontain. *Tetrahedron Comput. Method.*, **3**, 469 (1990)
76. I.Ugi, J.Bauer, K.Bley, A.Dengler, E.Fontain, M.Knauer, S.Lohberger. *J. Mol. Struct.*, **230**, 73 (1991)
77. I.Ugi, E.Fontain, J.Bauer. *Anal. Chim. Acta*, **235**, 155 (1990)
78. I.Ugi, J.Bauer, R.Baumgartner, E.Fontain, D.Forstmeyer, S.Lohberger. *Pure Appl. Chem.*, **60**, 1573 (1988)
79. E.J.Corey, W.T.Wipke. *Science*, **166**, 178 (1969)
80. E.J.Corey, X.-M.Cheng. *The Logic of Computer Synthesis*. Wiley, New York, 1989
81. Y.Yoneda. *Bull. Chem. Soc. Jpn.*, **52**, 8 (1979)
82. D.Bonchev. *Information Theoretic Indices for Characterization of Chemical Structures*. Research Studies Press, Chichester, 1983
83. R.E.Valdés-Pérez. *PhD Theses, Carnegie Mellon University; School of Computer Science, CMU-CS-90-191*, 1990
84. R.E.Valdés-Pérez. *Tetrahedron Comput. Method.*, **3**, 277 (1990)
85. R.E.Valdés-Pérez. *J. Comput. Chem.*, **13**, 1079 (1992)
86. R.E.Valdés-Pérez. *J. Phys. Chem.*, **96**, 2394 (1992)
87. R.E.Valdés-Pérez. *J. Comput. Chem.*, **14**, 1454 (1993)
88. R.E.Valdés-Pérez. *Catal. Lett.*, **28**, 79 (1994)
89. R.E.Valdés-Pérez. *J. Chem. Inf. Comput. Sci.*, **34**, 976 (1994)
90. R.E.Valdés-Pérez. *J. Comput. Chem.*, **15**, 1266 (1994)
91. R.E.Valdés-Pérez. *J. Chem. Inf. Comput. Sci.*, **31**, 554 (1991)
92. R.E.Valdés-Pérez. *Artif. Intell.*, **65**, 247 (1994)
93. R.E.Valdés-Pérez. Private communication
94. D.Bonchev, O.N.Temkin, D.Kamenski. *React. Kinet. Catal. Lett.*, **15**, 119 (1980)
95. D.Bonchev, D.Kamenski, O.N.Temkin. *J. Math. Chem.*, **1**, 345 (1987)
96. E.V.Gordeeva, D.Bonchev, D.Kamenski, O.N.Temkin. *Chem. Inf. Comput. Sci.*, **34**, 436 (1994)
97. А.В.Зейгарник, О.Н.Темкин. *Кинетика и катализ*, в печати
98. Г.М.Островский, А.Г.Зыскин, Ю.С.Снаговский. В кн. *Физическая химия: современные проблемы*. (Под ред. Я.М.Колотыркина). Химия, Москва, 1986. С.84
99. Ю.С.Снаговский, Г.М.Островский. *Моделирование кинетики гетерогенных каталитических процессов*. Химия, Москва, 1976
100. T.Morikawa. *Match*, **13**, 7 (1982)
101. D.I.Kaminski, O.N.Temkin, D.G.Bonchev. *Appl. Catal.*, **88**, 1 (1992)
102. J.Happel, P.H.Sellers. *Ind. Eng. Chem., Fundam.*, **21**, 67 (1982)
103. J.Happel, P.H.Sellers. *Chem. Eng. Commun.*, **83**, 221 (1989)
104. P.H.Sellers. In *Chemical Applications of Topology and Graph Theory*. (Ed. R.B.King). Elsevier, Amsterdam, 1983. P.420
105. В.И.Димитров. *Простая кинетика*. Наука, Новосибирск, 1982
106. R.H.S.Mah, R.Agis. *Ind. Eng. Chem., Fundam.*, **2**, 90 (1963)
107. В.В.Кафаров, В.Н.Писаренко, А.В.Солохин. *Журн. физ. химии*, **52**, 3143 (1978)
108. M.L.Mavrovouniotis, G.Stephanopoulos. *Ind. Eng. Chem. Res.*, **31**, 1625 (1992)
109. M.L.Mavrovouniotis. *Ind. Eng. Chem. Res.*, **31**, 1637 (1992)
110. M.L.Mavrovouniotis, G.Stephanopoulos, G.Stephanopoulos. *Biotechnol. Bioeng.*, **36**, 1119 (1990)
111. M.L.Mavrovouniotis, G.Stephanopoulos, G.Stephanopoulos. *Comput. Chem. Eng.*, **16**, 605 (1992)
112. M.L.Mavrovouniotis. In *Artificial Intelligence Approaches in Product and Process Engineering*. Academic Press, New York, 1995. Ch.3
113. T.D.Salatin, W.L.Jorgensen. *J. Org. Chem.*, **45**, 2043 (1980)
114. W.L.Jorgensen, E.R.Laird, A.J.Gushurts, J.M.Fleischer, S.A.Gothe, H.E.Helson, G.D.Paderes, S.Sinclair. *Pure Appl. Chem.*, **62**, 1921 (1990)
115. M.A.Baltanas, G.F.Froment. *Comput. Chem. Eng.*, **9**, 71 (1985)
116. Б.О.Польгейм, Л.Г.Брук, С.И.Шалгунов, О.Н.Темкин. *Тез. докл. VII Всесоюз. конф. «Математические методы в химии»*. Казань, 1991. С.220
117. С.И.Шалгунов, Л.Г.Брук, О.Н.Темкин. Неопубликованные результаты
118. С.И.Шалгунов. Неопубликованные результаты
119. C.A.Tolman. *Chem. Soc. Rev.*, **1**, 337 (1972)

COMPUTER-ASSISTED MECHANISTIC STUDIES

A.V.Zeigarnik, L.G.Bruk, O.N.Temkin, V.A.Likholobov, L.I.Maier

Lomonosov Academy of Fine Chemical Technology

86, Prosp. Vernadskogo, 117571 Moscow, Russian Federation, Fax +7(095)430-7983

Boriskov Institute of Catalysis, Siberian Branch of the Russian Academy of Sciences

5, Prosp. Akademika Lavrent'eva, 630090 Novosibirsk, Russian Federation, Fax +7(383-2)35-5766

Basic stages of the rational method for mechanistic studies are discussed. Advancing of mechanistic hypothesis is considered to be an initial key stage of the method. Computer programs that assist to advance mechanistic hypotheses are discussed. Empirical and formal-logical computer programs are compared and different aspects of their application are discussed. The difference between the computer programs for hypothesis generation and those for the assistance in organic synthesis design, which are similar in structure and algorithms, is revealed.

Bibliography — 119 references.

Received 10th May 1995