

УДК 519.688

МЕТОД АНАЛИЗА ХИМИЧЕСКОГО СОСТАВА В МОЛЕКУЛЯРНО-ДИНАМИЧЕСКИХ РАСЧЕТАХ С ПОТЕНЦИАЛОМ ВЗАИМОДЕЙСТВИЯ ReaxFF

О. В. Сергеев, А. В. Янилкин
(ВНИИА им. Н. Л. Духова, г. Москва)

Дано описание алгоритма и особенностей реализации программы для выделения молекул по данным молекулярно-динамического расчета. Обсуждается зависимость результатов от порогового значения порядка связи.

Ключевые слова: молекулярная динамика, ReaxFF, химический состав.

Введение

Классическая молекулярная динамика (МД) успешно применяется для решения разнообразных задач физики конденсированного состояния, термодинамики, физической кинетики. Этот подход основан на решении уравнений движения для системы взаимодействующих частиц. Взаимодействие описывается посредством потенциалов, которые полностью определяют свойства системы.

Первоначально потенциалы учитывали только двухчастичные взаимодействия, затем появились модели, позволяющие описывать системы со связями. Однако эти модели предполагали заранее определенную топологию системы, которая остается неизменной на протяжении всего расчета, и поэтому не позволяли описывать процессы с изменением химического состава. В дальнейшем были предложены потенциалы, которые позволили учитывать возможность образования и разрыва химических связей. Современной реализацией такого потенциала является ReaxFF, предложенный в 2001 г. [1] и используемый здесь в формулировке 2008 г. [2]. Применяемый в ReaxFF подход базируется на понятии порядка химической связи.

IUPAC определяет порядок связи как электронную населенность между атомами A и B в молекуле, образованную за счет уменьшения электронной плотности в непосредственной окрестности центров этих атомов. Такое определение не является однозначным, и разные спо-

собы расчета могут давать различные значения порядка связи [3]. В подходе Малликена порядок связи выражается через элементы $P_{\mu\nu}$ матрицы плотности и $S_{\mu\nu}$ матрицы перекрытия:

$$q_{AB} = 2 \sum_{\mu}^A \sum_{\nu}^B P_{\mu\nu} S_{\mu\nu}.$$

Физический смысл этого определения состоит в том, что порядок связи равен разности чисел пар связывающих электронов и антисвязывающих электронов. В теории валентных связей порядок связи определяется как взвешенное среднее формальных порядков связи (числа электронных пар в льюисовской структуре) в резонансных структурах [3].

Высокая электронная плотность в пространстве между атомами означает сильную химическую связь. Это позволяет использовать понятие порядка связи в моделях, которые не описывают явно электронную подсистему, например, в ReaxFF. В этом случае порядок связи можно рассматривать как условную величину, отражающую силу связи между атомами.

Подход, используемый в ReaxFF, основан на предположении, что порядок связи между двумя атомами в первом приближении определяется расстоянием между ними. Для каждой пары типов атомов потенциал определяет собственную зависимость порядка связи от расстояния. В молекулярно-динамических расчетах тип атома — это идентификатор, который позволяет по-разному описывать атомы, чем-либо

отличающиеся друг от друга. В рассматриваемых здесь задачах тип атомов — то же самое, что химический элемент, к которому они принадлежат. Возможна ситуация, когда в ходе расчета для какого-то атома сумма порядков образованных им связей превышает максимальную валентность соответствующего элемента; тогда производится корректировка порядков связей. Например, для водорода суммарный порядок связей ограничен 1, для кислорода — 2 и т. п.

Таким образом, для каждого данного расположения атомов в расчетной ячейке получается информация о существующих в системе связях и их порядке. Энергия системы рассчитывается на основе этой информации схожим образом с другими потенциалами, учитывающими химические связи (есть слагаемые, отвечающие длинам связей, плоским и двугранным углам и т. д.). Координаты атомов меняются на каждом шаге расчета, соответственно на каждом шаге связи определяются заново. Таким образом, образование и разрыв связей и соответствующие изменения потенциальной энергии оказываются непосредственно включенными в динамический расчет. Это позволяет не задавать заранее возможные пути химических превращений; кинетика химических реакций описывается адекватно в той степени, в которой ReaxFF воспроизводит поверхности потенциальной энергии.

Одной из наиболее популярных программ для классической молекулярной динамики является LAMMPS [4] — пакет с открытым исходным кодом, разрабатываемый главным образом в Сандийской национальной лаборатории. Для проведения расчетов с ReaxFF в этом пакете реализован на языке C++ модуль USER-REAXC [5], который имеет высокую производительность и

масштабируемость. Среди прочих возможностей, которые предоставляет модуль, доступен вывод порядка всех связей, определенных на заданных шагах расчета, что позволяет после завершения расчета использовать эту информацию для анализа изменений химического состава в расчетной ячейке. Здесь описывается программа moldetect, реализующая такой анализ для CHON-систем. Под CHON-системами понимаются системы, образованные молекулярными соединениями, в состав которых входят атомы углерода, водорода, азота и кислорода (такими соединениями являются, например, многие взрывчатые вещества: ТНТ, ТЭН, нитроглицерин, гексоген, октоген и др., а также продукты их превращений).

Код модуля USER-REAXC был модифицирован для того, чтобы выводить координаты атомов, образующих связь. Это позволяет отслеживать химические превращения не только во всей расчетной ячейке, но и в отдельных ее частях. При использовании измененного таким образом модуля USER-REAXC на выходе МД-расчета получаем файл, начало которого имеет вид, представленный на рис. 1.

Программа moldetect принимает следующие параметры:

- bonds_filename — имя файла с описанием связей (например, bonds.file);
- out_filename — основа имен выходных файлов;
- b_cutoff — пороговое значение порядка связи;
- x_min — минимальное значение x -координаты для расчета;
- x_max — максимальное значение x -координаты для расчета;
- Nparts — число частей, на которые делится интервал от x_{\min} до x_{\max} ;

```
# Timestep 0
#
# Number of particles 928
#
# Max number of bonds per atom 4 with coarse bond order cutoff 0.000}
# Particle connection table and bond orders}
# id type x y z nb id_1...id_nb mol bo_1...bo_nb abo nlp q
1 4 1.732 4.630 0.892 3 2 3 4 0 0.594 1.438 1.546 3.578 0.139 0.815
2 3 2.970 4.193 1.865 2 1 28 0 0.594 1.000 1.604 2.000 -0.437
3 3 1.373 3.778 0.096 1 1 0 1.438 1.494 2.000 -0.420
```

Рис. 1. Начало файла, получаемого на выходе МД-расчета с использованием модифицированного модуля USER-REAXC

`cell_lowboundary` — нижняя граница расчетной ячейки по оси x ;

`cell_uprboundary` — верхняя граница расчетной ячейки по оси x .

В результате выполнения `moldetect` получается набор файлов вида

```
x_min1_x_max1_out_filename,
...
x_minNparts_x_maxNparts_out_filename.
```

Для каждого момента времени выводится информация о количестве молекул всех химических соединений, присутствующих в данной части расчетной ячейки. Список соединений при этом сортируется по числу в молекуле атомов, последовательно углерода, азота, кислорода и водорода. При дальнейшей обработке можно выделить данные по некоторому отдельному соединению для разных координат в ячейке и шагов расчета.

Описание алгоритма

Область между минимальной и максимальной x -координатами, заданными при запуске программы, разбивается на также заданное при запуске количество частей; для каждой части определяются локальные минимальное и максимальное значения x . Описанная ниже процедура повторяется отдельно для каждой части.

Программа использует массивы структур данных, хранящих информацию об атомах и молекулах. Первый тип структур будем называть просто *атомами* (массив атомов), второй — *молекулами* (массив молекул).

Выделение молекул проводится в несколько шагов.

Первоначально считывается файл с информацией о связях. В массив для каждого атома заносятся его координаты и `id` (номера в МД-расчете) его соседей, порядок связи с которыми выше, чем `b_cutoff`, указанный при запуске программы. Таким образом, можно, например, выделить количество групп с двойными или тройными связями, а также отбросить связи низких порядков, такие как водородные.

Далее создается буфер, реализованный как двусторонняя очередь целых чисел. Начинается цикл по всем атомам; на каждом шаге цикла в очередь помещается номер текущего атома (`id`) и для него вызывается функция определения принадлежности к молекуле (функция `denote` в

исходном коде). Для каждого атома проверяется, включен ли он уже в какую-то молекулу. Если да, программа переходит к следующему атому, а номер текущего выталкивается из очереди. Если нет, создается новая структура с информацией о молекуле, и эта молекула "прикрепляется" к текущему атому. В структуре-молекуле сохраняется x -координата этого (ее первого) атома — впоследствии она используется для корректного учета периодических граничных условий (ПГУ).

Далее программа работает рекуррентно. Для каждого из соседних атомов проверяется, отнесен ли он уже к какой-либо молекуле. Номер текущего атома выталкивается из буфера, в буфер поочередно помещаются номера тех соседей текущего атома, для которых молекула еще не указана, и затем для них вызывается та же функция `denote` до тех пор, пока буфер не опустеет. Таким образом, если на каком-то шаге цикла попадает атом, еще не отнесенный к молекуле, в ходе последовательности рекуррентных вызовов функции `denote` отмечаются все другие атомы, принадлежащие молекуле, к которой должен быть отнесен рассматриваемый атом.

В той же функции `denote` фиксируется переход через ПГУ. Порядок связи в `ReaxFF` для СНОН-систем обнуляется на расстоянии около $2,5 \text{ \AA}$. Поэтому если два атома помечены как связанные, а расстояние между ними по x (разбиение ячейки делается только по x) превышает 3 \AA , молекула отмечается как проходящая через ПГУ.

Структура данных *молекула* содержит поля для хранения числа атомов каждого типа и координат центра молекулы. Под центром молекулы понимается среднее арифметическое координат ее атомов. При каждом добавлении атома в ходе выполнения `denote` его координаты добавляются в соответствующие поля молекулы. При этом, если молекула проходит через ПГУ, x -координата текущего атома сравнивается с x -координатой первого атома молекулы (см. выше). Если их отличие составляет больше половины длины расчетной ячейки, x -координата текущего атома изменяется на длину ячейки в большую или меньшую сторону (в зависимости от знака различия).

По такому алгоритму может быть неправильно определено положение атомов в молекулах, длина которых превышает длину расчетной ячейки минус 3 \AA . Однако это не вносит существенных ошибок, так как, если в ходе реакции

происходит образование молекулы, размеры которой сравнимы с размерами ячейки, конкретное положение центра масс такой молекулы в ячейке неважно.

После завершения цикла по атомам для каждого атома указано, к какой молекуле он принадлежит. При этом сами структуры-молекулы прикреплены посредством указателей к тем своим атомам, которые при проходе встретились первыми.

Далее выполняется еще один проход по атомам. В молекуле, к которой относится текущий атом, увеличивается счетчик, соответствующий его типу. Таким образом, после этого цикла каждая молекула содержит количество атомов каждого типа в своем составе.

Для молекул определяются логические операторы равенства и сравнения. Молекулы будут *равными*, если у них одинаковый состав. В этом случае считается, что молекулы относятся к одному типу. При разном составе *больше* будет та молекула, в которой больше углерода, либо при равных количествах углерода — больше азота, затем последовательно сравниваются количества кислорода и водорода.

Далее создается список типов молекул. Еще раз выполняется цикл по атомам. Для каждой отнесенной к атомам молекулы определяется положение ее центра. Если центр молекулы лежит между минимальным и максимальным x для данной части расчетной ячейки, молекула добавляется в список (если молекул данного типа еще не было обнаружено) либо увеличивается счетчик числа молекул такого типа (если они уже встречались). При этом список поддерживается отсортированным по размеру молекул.

Список молекул и их число в текущий момент времени выводятся в выходной файл. После этого программа переходит к данным о связях на следующем шаге МД-расчета.

Детали реализации

Программа написана на C++.

Для хранения информации об атомах и молекулах определяются классы `AtomInfo` и `Molecule`, не задающие никаких методов и по сути представляющие собой структуры данных.

Класс `AtomInfo` включает: массив `int attachedTo[MAXBONDS]`, где `MAXBONDS == 32` — значение, взятое как максимально возможное число связей у каждого атома в отдельный мо-

мент времени (в некоторых случаях атомы могут образовывать большое число связей низкого порядка); поля `type` и `mol` типа `int` (`type` содержит тип атома, `mol` — идентификатор молекулы, к которой он относится); три поля `float` для координат; указатель на `Molecule`, которая прикрепляется к данному атому в ходе выполнения программы.

Класс `Molecule` содержит: целочисленные счетчики для каждого типа атомов; целочисленную переменную `numberOf`, обозначающую число молекул данного типа в списке; четыре поля `float` для расчета координат центра молекулы и хранения x -координаты ее первого атома; логическую переменную, отмечающую переход через ПГУ; указатель на `Molecule`, используемый при построении списка.

Буфер номеров атомов, для которых должна быть вызвана функция `denote`, реализован с помощью двусторонней очереди `deque` из стандартной библиотеки шаблонов.

При включении молекулы в список производится последовательное сравнение ее с каждой уже имеющейся молекулой. При этом хранятся указатели на текущий (`currpos`) и предыдущий (`prevpos`) элементы списка типов молекул. Если молекула больше текущего элемента, она включается в список перед ним. Если она меньше текущего элемента, проход по списку продолжается. Если молекула равна текущему элементу, у него увеличивается счетчик `numberOf`. Наконец, если `currpos == NULL`, это означает, что пройден весь список, и молекула добавляется после `prevpos`, в конец списка.

Результаты выполнения программы

Выделением из выходных файлов информации о некотором отдельном соединении получается зависимость концентрации этого соединения от времени в разных частях расчетной ячейки. Эта зависимость может использоваться, например, для определения скорости распространения некоторой реакции в веществе или диффузии одного вещества в другом.

В качестве примера приводится расчет скорости распространения реакции горения в монокристалле пентаэритриттетранитрата (ТЭНа). МД-модель представляет собой ячейку в ПГУ размером 256 элементарных ячеек ТЭНа вдоль направления $\langle 100 \rangle$ и 4 ячейки в перпендикулярных направлениях. Каждая элементарная ячейка ТЭНа содержит 58 атомов, соответственно

полное число атомов в системе равно $58 \times 256 \times 4 \times 4 = 237568$ атомов. Размер системы зависит от давления и при нормальных условиях примерно равен $2400 \times 35 \times 30 \text{ \AA}$.

Расчетная ячейка приводится к заданным значениям температуры и давления (в рассматриваемом примере это 300 К и 3 ГПа) с помощью метода Нозе—Гувера. Затем в средней части ячейки выбирается слой толщиной около пяти элементарных ячеек, в котором искусственно задается температура 2000 К. Это инициирует начало химических реакций. Выделяющаяся при перестройке связей энергия переходит в кинетическую энергию образующихся частиц, они взаимодействуют с соседними слоями вещества, и таким образом реакция распространяется по кристаллу.

Ячейка в этом расчете делится на 128 слоев, для каждого из которых отдельно рассчитывается количество молекул каждого типа. Таким образом, каждый слой первоначально содержит 64 молекулы ТЭНа. Приведенные на рис. 2 (см. также цветную вкладку) изолинии концентраций некоторых соединений соответствуют уровню 32 молекулы на ячейку для ТЭНа и воды,

10 молекул на ячейку для NO_2 и 20 молекул на ячейку для CH_2O . Вода является одним из продуктов горения ТЭНа, тогда как NO_2 и CH_2O — промежуточные соединения. Цветом выделены данные для разных пороговых порядков связи.

Некорректное задание порогового порядка связи (`b_cutoff`) может привести к тому, что концентрация одних соединений будет искусственно завышаться, а других — занижаться. Желательно поэтому выбирать пороговое значение в том диапазоне, в котором интересующие величины слабо от него зависят. Для рассматриваемого примера существенно, чтобы определяемая скорость распространения реакции правильно определялась для выбранного порогового порядка связи.

Как видно из рис. 2 (см. также цветную вкладку), для взятых значений `b_cutoff` это требование выполняется (скорость не зависит от конкретного значения в рассмотренном диапазоне). Для трех соединений из четырех изолинии оказываются очень близки. Для NO_2 определяемое значение концентрации оказывается сильно зависящим от выбранного значения `b_cutoff`. Это объясняется тем, что отделение NO_2 — преобладающая реакция на первой стадии горения

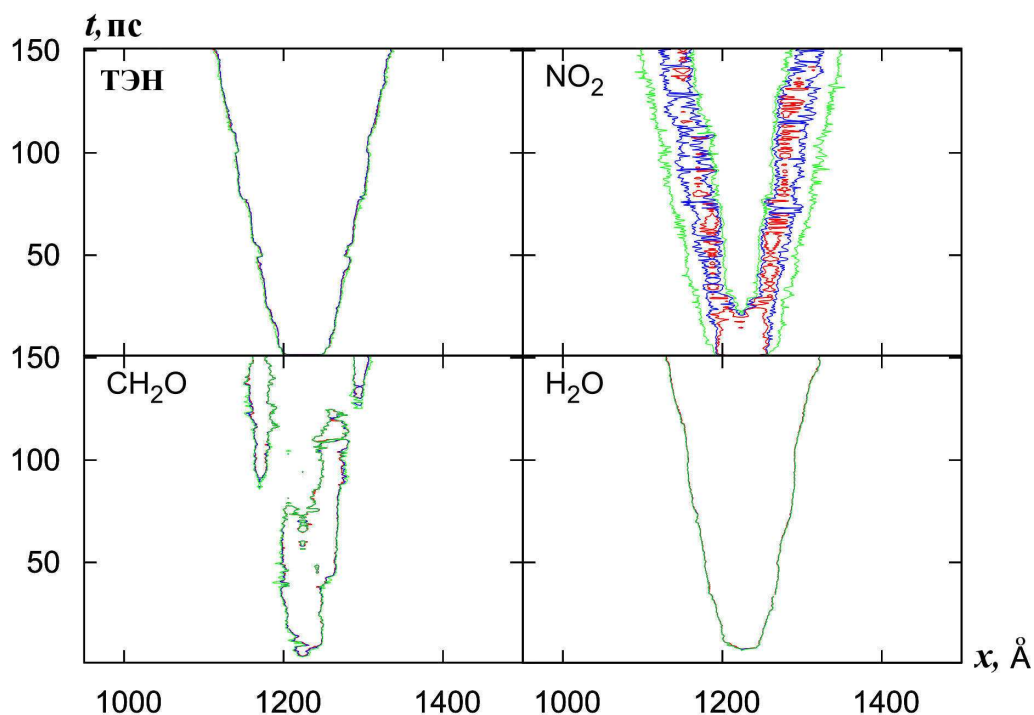


Рис. 2. Изолинии концентраций некоторых веществ в расчетной ячейке для разных пороговых значений порядка связи

ТЭНа, а связь этой группы с остатком — самая слабая в молекуле [6–8]. Поэтому при нагреве вещества средний порядок этой связи снижается. Как такового разрыва связи при этом не происходит, но при задании высокого значения порогового порядка связи часть групп может определяться как отдельная молекула.

Рассматривая некоторый слой в отдельности, можно получать информацию о кинетике происходящих при нагреве реакций. Зависимости концентрации от времени для тех же веществ, что на рис. 2, приведены на рис. 3 (см. также цветную вкладку). Показаны данные для слоя, первоначально находящегося при нормальной температуре (x -координата центров молекул между 1142 и 1160 Å).

Для приведенных на рис. 3 сглаженных зависимостей цветовые обозначения (см. цветную вкладку) те же, что на рис. 2. Серым цветом показаны несглаженные данные для порогового порядка связи 0,4. Для разных пороговых значений порядка связи последовательность образования различных веществ остается постоянной, так что наблюдаемые механизмы реакций не меняются. Вместе с тем абсолютные значения концентраций, в особенности для промежуточных

соединений, могут сильно зависеть от порогового значения.

В целом определение подходящего порогового порядка связи — существенная часть в изучении кинетики химических реакций по данным МД-расчетов. Характер реакций, моделируемые соединения, температура и давление в расчетной ячейке и другие факторы могут существенно влиять на результаты определения молекулярного состава. В данной работе показан пример такого определения на основании сходимости интересных величин по отношению к изменению порогового значения.

Возможны и другие подходы, как, например, в [9], где авторы исследовали химические реакции при прохождении сильной ударной волны в монокристалле ТЭНа. Для определения концентраций они использовали различные порядки связи для разных пар типов атомов. Набор пороговых порядков связи получался из условия, что типичные молекулы, участвующие в реакциях, такие как собственно ТЭН, вода, азот и углекислый газ, правильно определялись при нормальной плотности несжатого ТЭНа. Полученный таким образом пороговый порядок связи N—O равен 0,4.

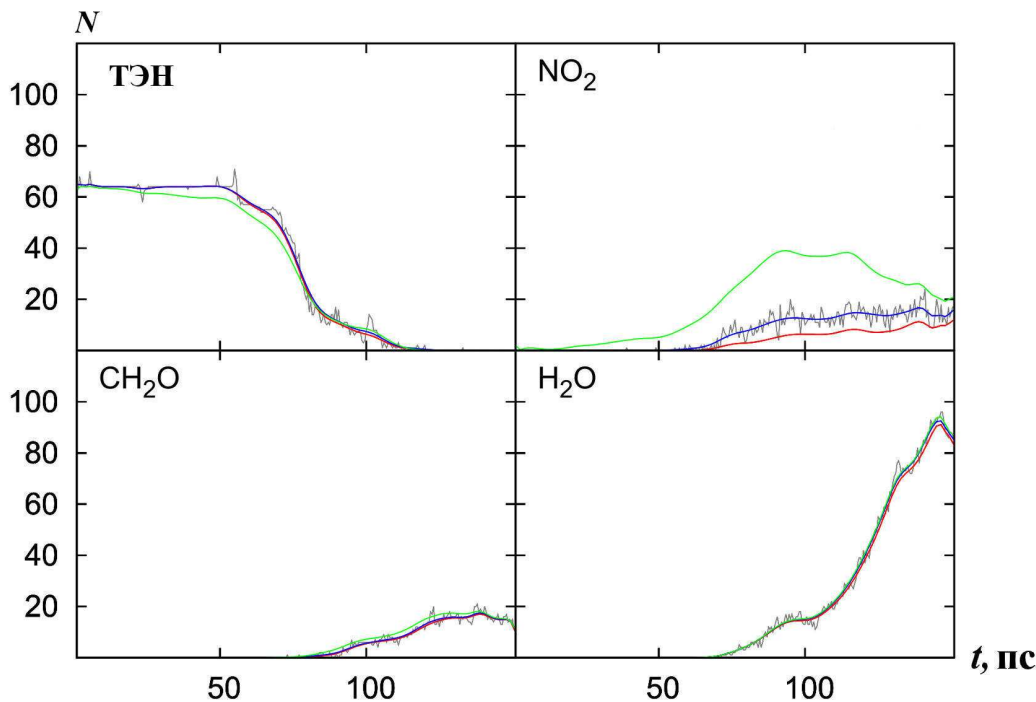


Рис. 3. Зависимости концентрации N от времени в первоначально неразогретой области

В расчетах авторов настоящей статьи определяемые значения концентрации NO₂ (сильно зависящие от порогового порядка связи N—O) для значения 0,4 близки к результатам для более низкого значения 0,2 и существенно отличаются от значений для порогового порядка связи 0,5 (см. рис. 2). Более высокие значения приводят к фрагментации определяемых молекул.

Таким образом, можно говорить о согласованности значений порогового порядка связи, получаемых с использованием различных подходов. Тем не менее в каждой задаче применимость тех или иных значений следует тестировать отдельно.

Расчеты проведены на вычислительных кластерах ВНИИА им. Н. Л. Духова и РФЯЦ-ВНИИЭФ.

Заключение

В работе описан алгоритм выделения молекул в расчетной ячейке по данным МД-расчета с потенциалом взаимодействия ReaxFF, а также особенности его реализации и входные параметры программы.

Обсуждается проблема определения порогового порядка связи, начиная с которого атомы считаются входящими в состав одной молекулы. На примере распространения горения в монокристалле ТЭНа показано, что определяемая скорость распространения реакции в этой задаче не зависит от порогового порядка связи в диапазоне 0,2—0,5, однако изменение этого параметра может влиять на наблюдаемую кинетику происходящих процессов.

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (грант 12-02-13511-офи-м-ра).

Список литературы

1. *Duin A. C. van, Dasgupta S., Lorant F., Goddard III W. A.* ReaxFF: a reactive force field for hydrocarbons // *J. Phys. Chem. A.* 2001. Vol. 105, No 41. P. 9396—9409.
2. *Chenoweth K., Duin A. C. van, Persson P. et al.* Development and application of a ReaxFF reactive force field for oxidative dehydrogenation on vanadium oxide catalysts // *J. Phys. Chem. C.* 2008. Vol. 112, No 37. P. 14645—14654.
3. *Minkin V. I.* Glossary of terms used in theoretical organic chemistry // *Pure Appl. Chem.* 1999. Vol. 71, No 10. P. 1919—1981.
4. *Plimpton S.* Fast parallel algorithms for short-range molecular dynamics // *J. Comp. Phys.* 1995. Vol. 117. P. 1—19.
5. *Aktulga H., Fogarty J., Pandit S., Grama A.* Parallel reactive molecular dynamics: Numerical methods and algorithmic techniques // *Parallel Computing.* 2012. Vol. 38, No 4—5. P. 245—259.
6. *Wu C. J., Ree F. H., Yoo C.-S.* A quantum mechanical molecular dynamics study of binary collisions of pentaerythritol tetranitrate (petn): its correlation to shock sensitivity // *Propellants, Explosives, Pyrotechnics.* 2004. Vol. 29, No 5. P. 296—303.
7. *Wu C. J., Manaa M. R., Fried L. E.* A molecular dynamics study of chemical reactions of solid pentaerythritol tetranitrate at extreme conditions. Livermore National Laboratory, 2006.
8. *Landerville A. C., Oleynik I. I., White C. T.* Reactive molecular dynamics of hypervelocity collisions of petn molecules // *J. Phys. Chem. A.* 2009. Vol. 113. P. 12094—12104.
9. *Budzien J., Thompson A. P., Zybin S. V.* Reactive molecular dynamics simulations of shock through a single crystal of pentaerythritol tetranitrate // *Ibid. B.* 2009. Vol. 113, No 40. P. 13142—13151.

Статья поступила в редакцию 08.10.13.