ПРОГРАММНЫЕ СРЕДСТВА ДЛЯ БЫСТРЫХ КВАНТОВОХИМИЧЕСКИХ РАСЧЕТОВ ДОКИНГ-КОМПЛЕКСОВ

Мендкович А.С., Аникин Н.А., Андреев А.М., Бобриков В.В., Кузьминский М.Б., Мускатин А.Ю.

Институт органической химии РАН, г.Москва

Мендкович А.С.
Аникин Н.А.
Андреев А.М.
Бобриков В.В.
Кузьминский М.Б.
Мускатин А.Ю.
Институт органической химии РАН

В рамках проекта РФФИ 05-07-08031-офи разработан базовый комплекс программных средств для массовых квантовохимических расчетов докинг-комплексов. Данные программные средства ориентированы на расчеты в гетерогенной распределенной вычислительной среде. Они включают собственно программы для квантовохимических расчетов (как протеина, включая распараллеленный алгоритм и программу сборки фокиана в технологии разреженных матриц, так и лигандов по предложенной ранее [1] и доработанной авторами методике). Эти программы являются мультиплатформенными; они работают в 32- и 64-разрядных x86-совместимых средах в ОС Linux и Windows2000/XP.

В процессе расчетов используется БД лигандов и Web-интерфейсы, обеспечивающие поиск в БД, подготовку и запуск программ на счет. Особое внимание уделено обеспечению отказоустойчивости при обработке потока из большого числа пакетных заданий с тем, чтобы при возможных сбоях гарантировать автоматическое повторное выполнение расчета соответствущего комплекса.

Начальная апробация метода выполнена с использованием полуэмпирического гамильтониана AM1. В качестве объекта исследования использовались комплексы человеческого иммуноглобулина с иммунным супрессором IMMUNOPHILIN FKBP-12 ("FKB", 1663 атома) и фермента панкреатическая рибонуклеаза ("QHC", 1852 атома), а также протеины меньшего размера. Среди исследованных лигандов были представлены, в частности, динуклеотиды оснований ДНК/РНК с соединением углеводных частей оснований через дифосфатный фрагмент, а также мононуклеозиды. Данные о структуре протеинов брались из БД Protein Data Bank (PDB) [2]. Параметры структуры лигандов (органические молекулы средних размеров, от 28 до 144 атомов) докинг-комплексов были взяты из БД образцов PDBBIND [3]

Было проведено сравнение быстродействия разработанной программы со стандартной процедурой диагонализации на итерациях ССП (см. Таблицу). На основании литературных данных сделаны также ориентировочные оценки относительного быстродействия по сравнению с методами (CG-DMS, DivCon [4]), созданными именно для быстрых расчетов больших био-молекул (см. Табл.). Расчеты производились на сервере с AthlonXP 3000+ (2,2 ГГц).

Как видно из таблицы, время расчетов по сравнению со стандартной диагонализацией сокращается как правило более чем 300 раз, а по сравнению с CG-DMS предложенный метод выигрывает, как правило, от 5 до 30 раз, по сравнению с методом DivCon от 18 до 120; причем выигрыш в быстродействии увеличивается с ростом размера протеина

Результаты для комплексов протеинов FKB (1663 атома) и QHC (1852 атома).

Протеин	Лиганд	Число атомов	Ускорение по сравнению с другими методами (раз)		
			Стандартная диагонализация	CG-DMS	DivCon
FKB	$C_{51}H_{79}NO_{13}$ (5 различных структур)	144	392-396	5-7	18-25
FKB	$C_{10}H_{16}N_2O_4S_2$	34	972	14	51
FKB	$C_{12}H_{14}N_2O_5Br_2$	35	914	13	48
FKB	$C_{11}H_{11}N_4OCl$	28	1027	14	54
QHC	$C_{19}H_{27}N_7O_{20}P_4$	77	1148	15	56

QHC	$C_{19}H_{29}N_8O_{19}P_4$	79	1168	11	41
QHC	$C_{19}H_{29}N_8O_{20}P_4$	80	1135	17	64
QHC	$C_{20}H_{29}N_7O_{20}P_4$	80	1136	14	54
QHC	$C_{20}H_{29}N_7O_{21}P_4$	81	1097	11	41
QHC	$C_{23}H_{23}N_3O_9S$	59	1512	21	78
QHC	$C_{10}H_{24}P_2S_2$	38	3541	48	184
QHC	$C_{14}H_{14}N_5O_4Cl_3$	40	2365	32	123

Литература

1. *N. A. Anikin* et al, 227th ACS National Meeting (Anaheim)/COMP / Beyond Michael Dewar's Legacy: Modern Semiempirical MO Theory. 03/31/2004, Paper 247 : (Oral) “Novel techniques for QM modeling of ultra-large bio-molecular systems. Evolution of quantum-mechanical and force field methods”.
2. http://www.pdb.org/pdb/Welcome.do
3. An Extensive Test of 14 Scoring Functions Using the PDBbind Refined Set of 800 Protein-Ligand Complexes// J. Chem. Inf. Comput. Sci., 44(2004)б P.2114-2125 .
4. *A.van-der Vaart, D.Suarez, and K.M.Merz*, Critical assessment of the performance of the semiempirical Divide and Conquer method for single point calculations and geometry optimizations of large chemical systems, J.Chem.Phys., 2000, 113, P. 10512-10523.