J Mol Model (2001) 7:422–432
DOI 10.1007/s00894-001-0055-8

ORIGINAL PAPER

Aurelien Dejoux · Piotr Cieplak · Nicolas Hannick
Guillermo Moyna · François-Yves Dupradeau

AmberFFC, a flexible program to convert AMBER and GLYCAM force fields for use with commercial molecular modeling packages

Received: 9 July 2001 / Accepted: 1 October 2001 / Published online: 23 November 2001

Abstract A program for converting the different existing AMBER and GLYCAM force fields for use with commercial molecular modeling packages is presented, using the Molecular Simulations Inc. (MSI or Accelrys) software package as a case model. Called AmberFFC, the program creates AMBER and GLYCAM force field files suitable for use with the Accelrys molecular mechanics modules by transforming the amino acid, nucleotide, and monosaccharide topology databases and force field parameter files to the Accelrys file format. It is intended for any modeler who is interested in using the current AMBER and GLYCAM force fields with the Accelrys FDiscover and CDiscover programs. AmberFFC has been written entirely with the Perl programming language, making it highly flexible and portable. In order to compare the implementation of the force fields converted by AmberFFC in the Accelrys package with their corresponding execution in the AMBER software, and also to verify the efficiency of the AmberFFC program, results from single point energy calculations for 13 model molecules were obtained with the two programs. It is demonstrated that results obtained with the CDiscover and FDiscover modules compare well to those found using Sander_classic, thus showing that AmberFFC is a highly efficient program. Some energy differences between the AMBER and Accelrys software have been observed, and their origin has been characterized and discussed.

A. Dejoux · N. Hannick · F.-Y. Dupradeau (✉)
Biomolécules et Pathologies Dégénératives, UPRES EA 2629,
Faculté de Pharmacie et de Médecine, 1–3 rue des Louvels,
80037 Amiens Cedex 1, France
e-mail: fyd@u-picardie.fr
Tel.: +33-(0)3-2282-7494, Fax: +33-(0)3-2282-7469

P. Cieplak
Chemistry Department, Warsaw University, Pasteur 1,
02-934 Warsaw, Poland

G. Moyna
Department of Chemistry and Biochemistry,
University of the Sciences in Philadelphia, 600 South 43rd Street,
Philadelphia, PA 19104-4495, USA

Keywords Accelrys/MSI · AMBER · Format conversion · Force fields · GLYCAM

Introduction

Computer modeling of complex biomolecules such as proteins, nucleic acids, and carbohydrates has been crucial in the study of their structure, function, and dynamics. [1, 2, 3] Due to the size of biomolecular systems, parameter-independent quantum mechanical methods are inapplicable to their analysis, and the majority of these studies have been carried out using parameterized molecular mechanics methods. Many empirical force fields have appeared in the literature for this purpose, such as the AMBER, [4, 5, 6, 7] CHARMM, [8, 9, 10, 11] CFF91, [12, 13, 14] GROMOS, [15] MMFF, [16, 17] and MM2/MM3 force fields. [18, 19, 20, 21] Each of these force fields has been parameterized against small representative fragments taken from the molecules it is intended to model, and therefore their accuracy varies according to the model molecules used in the parameterization process.

The Weiner et al., [4, 5] Cornell et al., [6] and Wang et al. [7] force fields, all developed by the Kollman group (http://www.amber.ucsf.edu/), and representing the different force fields available in the AMBER molecular modeling software, [22, 23] have been established as one of the standards in the study of proteins and nucleic acids. Their potential energy form is given in Eq. (1):

$$
\begin{aligned}
E_{total} = & \sum_{Bonds} K_r (r - r_0)^2 + \sum_{Angles} K_\theta (\theta - \theta_0)^2 \\
& + \sum_{\substack{Dihedrals \\ n=1\ to\ 4}} \frac{Vn}{2} [1 + \cos(n\phi - \phi_0)] \\
& + \sum_{i<j} \left[\left(\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} \right) + \frac{q_i q_j}{\varepsilon R_{ij}} \right] \\
& + \frac{1}{vdW_{Sc.f.}} \sum_{\substack{i<j \\ 1-4\ terms}} \frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} + \frac{1}{EE_{Sc.f.}} \sum_{\substack{i<j \\ 1-4\ terms}} \frac{q_i q_j}{\varepsilon R_{ij}} \\
& + \sum_{h-bonds} \left(\frac{C_{ij}}{R_{ij}^{12}} - \frac{D_{ij}}{R_{ij}^{10}} \right)
\end{aligned}
\tag{1}
$$

The first three terms deal with bonded interactions. Hooke's law is used to represent the harmonic bond stretching and angle bending potentials. K_r and K_θ are the force constants of the bond and angle, and r_{eq} and θ_{eq} their equilibrium values, respectively. The torsional potential is represented by Fourier series, where V_n, n, ϕ, and ϕ_0 are the dihedral energy barrier, the periodicity, the dihedral, and the phase, respectively. The remaining four terms account for the non-bonded interactions, and include van der Waals, electrostatic, and hydrogen bonding potentials. A_{ij}, B_{ij}, C_{ij}, and D_{ij} are Lennard-Jones and hydrogen bonding parameters that describe through-space attractive and repulsive interactions, and R_{ij}, ε, q_i, and q_j are the distance between the atoms i and j, the dielectric constant, and the electrostatic atom-centered charges, respectively. Finally, $1/vdW_{sc.f.}$ and $1/EE_{sc.f.}$ are scaling factors for the 1–4 van der Waals and electrostatic interactions. In addition to the standard parameters for the study of proteins and nucleic acids, the Woods group has developed the GLYCAM force field parameter sets, [24, 25] which employ the same AMBER potential energy function. The GLYCAM force fields are consistent with other AMBER parameter sets, and make these force fields suitable for the study of carbohydrates and glycoproteins.

One of the problems faced by users of commercial molecular modeling software is the delay experienced from the time new force fields are made available to the scientific community to the time they are implemented in their particular modeling package. This lag time can be in some cases more than 1 year. Furthermore, if the newer force fields are implemented by commercial software vendors, they are usually supplied as part of costly upgrades. In some cases, new force fields are not covered by the software vendor at all. This means in many cases that new software has to be acquired, which represents a considerable expense. In this report, we wish to present a program, named AmberFFC, that is able to convert, in its current version (number 1.2), the different AMBER and GLYCAM force fields freely available in the public domain for use with the Accelrys molecular modeling package (formerly Molecular Simulations, Inc., San Diego, Calif., http://www.accelrys.com/). We present the design and theory of operation of AmberFFC, as well as results from single point energy calculations for 13 molecules obtained with the AMBER force fields in the AMBER and Accelrys molecular mechanics packages. As discussed below, our comparisons not only show that AmberFFC is a highly effective force field converter, but also allowed us to discover errors, or bugs, present in the CDiscover module of the Accelrys software package.

Methods

Criteria of the programming language

The Perl (Practical Extraction and Report Language) programming language, written by Larry Wall, was selected to write AmberFFC. [26, 27, 28, 29] Thanks to its flexibility and efficiency, Perl is particularly well adapted to modify, extract, and re-format ASCII (or "text") files. Thus, AmberFFC is easily modifiable to incorporate new program capabilities. Furthermore, Perl is an interpreted programming language, which presents the advantage that the sources do not need to be compiled. Finally, the language follows the Open Source community philosophy, and is therefore freely available on the Internet. [30] Although it was originally envisioned and developed for Unix, it has been ported to all major operating systems, and this makes AmberFFC a highly portable software. Indeed, AmberFFC runs on different Unix computers such as Hewlett-Packard, SGI, IBM, Linux-PC, and SUN, as well as on Win32 (Windows NT, 95, 98, 2000) workstations.

Presentation of AmberFFC

AmberFFC contains 1600 lines of Perl code. Functions have been used whenever practical to give the code high flexibility. As the options of AmberFFC are simple and straightforward (Table 1), no graphical interface has been developed. Colors have been added to the program menus for better clarity. AmberFFC gives the user the possibility of converting nine force fields versions to the Accelrys file format: The Weiner et al. force field [4, 5] (PLEP and LEaP versions; cases "A" and "B" [31], respectively), the Cornell et al. force field [6] with two adaptations [32, 33] (cases "C", "D", and "E" [34]), the Wang et al. force field [7] (case "F" [34]), the GLYCAM_93 force field [24] (case "G" [35]), and two new GLYCAM force field versions [25] (cases "H" and "I"; provided by Dr Woods) are possible options. These new force field files are all suitable for use with the CDiscover molecular mechanics module of Accelrys. [36, 37] On the other hand, FDiscover can only use the Weiner et al. and the GLYCAM_93 force fields because it does not handle torsion terms with a periodicity of n=4. [38, 39]

Features of AmberFFC

Since the CDiscover and FDiscover Accelrys modules run only on SGI and IBM RS/6000 platforms, [40, 41] AmberFFC tests which operating system is installed on the workstation at start-up. Indeed, once all the force field ASCII files have been generated, two Accelrys scripts are automatically executed by AmberFFC to build the binary version of the force field files if the appropriate Unix operating system is found. [38, 42] In contrast, only the force field ASCII files are generated if another operating system is detected. Then, AmberFFC tests whether the AMBER and Accelrys software are installed. The user has the opportunity to run AmberFFC in two different modes. (i) In the automatic or default mode, both AMBER and Accelrys software packages are installed in the workstation. The user just runs AmberFFC and selects the AMBER or GLYCAM force field to be transformed to the Accelrys format. (ii) In the second mode, the AM-

Table 1 Proposed choices once AmberFFC has been run

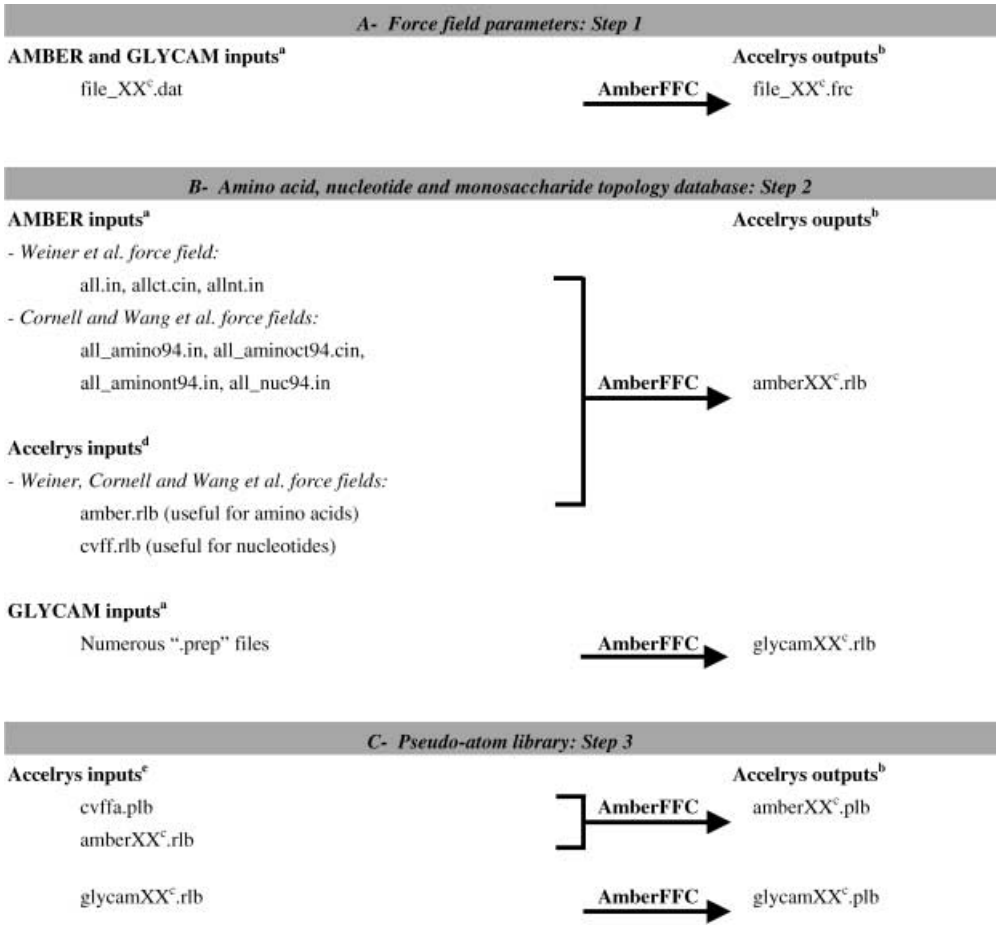
Which force field do you want to build for F&CDiscover? (X to exit)

(A) Amber 91 (PLEP)
(B) Amber 91X (LEAP)
(C) Amber 94
(D) Amber 96
(E) Amber 98
(F) Amber 99
(G) Glycam 93
(H) Glycam 2000
(I) Glycam 2010

(J) All

Your choice:

Table 2 Strategy used to build the AMBER and GLYCAM force fields for FDiscover and CDiscover (Accelrys software) (*A* building of the force field parameter file; *B* building of the amino acid, nucleotide, and monosaccharide topology database; *C* building of the pseudo-atom library)



[a] AMBER and GLYCAM force field files used as inputs by AmberFFC
[b] AMBER and GLYCAM force field files obtained to the Accelrys format using AmberFFC
[c] XX=year of the force field version
[d] Topology database files found in Accelrys software and used as "format donor" inputs by AmberFFC
[e] Accelrys input files used by AmberFFC to build the pseudo-atom library

BER and/or Accelrys software packages are not installed on the system, and the user has to specify in the main section of the code where the input force field files used by AmberFFC are located in the file system tree.

AmberFFC creates Accelrys force field files in five successive steps. First, it converts the force field parameter file to the Accelrys format. The Accelrys amberXX.frc or glycamXX.frc (where XX is the year of the force field version) file [14, 43] is obtained using the corresponding parmXX.dat AMBER or glycamXX.dat GLYCAM file as input. [4, 5, 6, 7, 24, 25, 31, 32, 33, 34, 35] Then, AmberFFC creates the Accelrys topology dat-

Table 3 Comparison of the implementation of the AMBER force fields in the AMBER and Accelrys molecular mechanics modules

Software[a]	Accelrys/CDiscover						AMBER/Sander_classic					
Interactions E[b]	Bond				Non-bond		Bond				Non-bond	
	Bond	Angle	Dih[c]	vdW[d]	Coul[e]	HB[f]	Bond	Angle	Dih[c]	vdW[d]	Coul[e]	HB[f]
Model **1**:	Cyclohexane (chair conformation)[g]											
AMBER force fields												
91X	0.0881	0.8932	0.2136	0.8290	0.6724	0.0000	0.0881	0.8932	0.2136	0.8290	0.6724	0.0000
94→98	0.0887	0.9204	0.2300	2.4028	0.0179		0.0887	0.9204	0.2300	2.4028	0.0179	
99	0.0887	0.9204	2.7450	2.4028	0.0179		0.0887	0.9204	2.7450	2.4028	0.0179	
Model **2**:	Cyclohexane (boat conformation)[g]											
AMBER force fields												
91X	0.3889	1.0933	4.8771	1.5228	0.4300	0.0000	0.3889	1.0933	4.8771	1.5227	0.4300	0.0000
94→98	0.3902	1.1348	5.2523	4.5302	0.0180		0.3902	1.1348	5.2523	4.5303	0.0180	
99	0.3902	1.1348	6.9889	4.5302	0.0180		0.3902	1.1348	6.9889	4.5303	0.0180	
Model **3**:	Methylethylether (*trans* conformation)[g]											
AMBER force fields												
91X	0.1845	1.0446	2.3014	0.8396	–6.0400	0.0000	0.1845	1.0445	2.3014	0.8396	–6.0399	0.0000
94→98	0.1860	1.0968	2.3014	0.7287	1.2500		0.1860	1.0968	2.3014	0.7287	1.2500	
99	0.1860	1.0968	3.0528	0.7287	1.2500		0.1860	1.0968	3.0528	0.7287	1.2500	
Model **4**:	Methylethylether (*gauche* conformation)[g]											
AMBER force fields												
91X	0.2027	1.4952	0.8527	0.6839	–1.3315	0.0000	0.2027	1.4952	0.8527	0.6839	–1.3315	0.0000
94→98	0.2046	1.6737	0.6632	0.7148	1.3096		0.2046	1.6736	0.6632	0.7148	1.3095	
99	0.2046	1.6737	1.3944	0.7148	1.3096		0.2046	1.6736	1.3944	0.7148	1.3095	
Model **5**:	*N*-Methylacetamide (*trans* conformation)[g]											
AMBER force fields												
91X	0.0003	0.0005	0.0000	1.6199	–20.5733[h]	0.0000	0.0003	0.0005	0.0000	1.6199	–20.5731	0.0000
94→98	0.0003	0.0015	0.0000	1.9383	–10.6862		0.0003	0.0015	0.0000	1.9383	–10.6862	
99	0.0003	0.0022	2.3998	1.9383	–10.6862		0.0003	0.0022	2.3998	1.9383	–10.6862	
Model **6**:	*N*-Methylacetamide (*cis* conformation)[g]											
AMBER force fields												
91X	0.0006	0.0006	1.3000	6.2685	–21.7679[h]	0.0000	0.0006	0.0006	1.3000	6.2685	–21.7677	0.0000
94→98	0.0006	0.0017	4.0000	7.8676	–12.0422		0.0006	0.0017	4.0000	7.8676	–12.0421	
99	0.0006	0.0025	6.3998	7.8676	–12.0422		0.0006	0.0025	6.3998	7.8676	–12.0421	
Model **7**:	Methanol[g]											
AMBER force fields												
91X	0.1702	0.2427	0.0007	–0.0137	–0.2181	0.0000	0.1702	0.2427	0.0007	–0.0137	–0.2181	0.0000
94→99	0.1710	0.3261	0.0007	0.0000	1.4501		0.1710	0.3261	0.0007	0.0000	1.4501	
Model **8**:	Deoxyadenosine (C2'endo conformation)[g]											
AMBER force fields												
91X	1.8987	4.1845	9.4785[ia]	3.6281	–67.2036[h]	–0.0588	1.8987	4.1845	9.4692	3.6281	–67.2033	–0.0588
94/96	1.9052	4.4535	13.3019[ia,ic]	5.7277	–65.4021[h]		1.9052	4.4535	14.8081	5.7278	–65.4017	
98	1.9052	4.4535	14.8112[ia,ic]	5.7277	–65.4021[h]		1.9052	4.4535	15.7632	5.7278	–65.4017	
99	1.9052	4.5523	16.0910[ia,ic]	5.7277	–65.4021[h]		1.9052	4.5523	17.3525	5.7278	–65.4017	
Model **9**:	Deoxyadenosine (C3'endo conformation)[g]											
AMBER force fields												
91X	1.6546	4.6881	9.5147[ia]	3.8180	–67.3614[h]	–0.0579	1.6546	4.6881	9.5062	3.8180	–67.3610	–0.0579
94/96	1.6616	4.9949	14.5119[ia,ic]	5.4462	–66.4872[h]		1.6616	4.9949	15.9455	5.4463	–66.4869	
98	1.6616	4.9949	16.6764[ia,ic]	5.4462	–66.4872[h]		1.6616	4.9949	17.3483	5.4463	–66.4869	
99	1.6616	5.0892	17.8089[ia,ic]	5.4462	–66.4872[h]		1.6616	5.0892	18.8743	5.4463	–66.4869	
Model **10**:	Alanine dipeptide (C5 conformation)[g]											
AMBER force fields												
91X	1.2952	0.8717	0.6344[ia,ib]	1.4879	–39.3059[h]	–0.3016	1.2952	0.8717	0.6338	1.4879	–39.3057	–0.3016
94/98	1.2994	0.9819	3.4960[ia–ic]	3.1057	–31.3712[h]		1.2994	0.9819	3.5007	3.1056	–31.3711	
96	1.2994	0.9819	1.5319[ia–ic]	3.1057	–31.3712[h]		1.2994	0.9819	1.5365	3.1056	–31.3711	
99	1.2994	1.0256	8.1899[ia–ic]	3.1057	–31.3712[h]		1.2994	1.0256	8.5140	3.1056	–31.3711	

Table 3 (continued)

Software[a]	Accelrys/CDiscover						AMBER/Sander_classic					
Interactions E[b]	Bond				Non-bond		Bond				Non-bond	
	Bond	Angle	Dih[c]	vdW[d]	Coul[e]	HB[f]	Bond	Angle	Dih[c]	vdW[d]	Coul[e]	HB[f]
Model **11**:	Alanine dipeptide (C7eq conformation)[g]											
AMBER force fields												
91X	1.2683	0.7800	1.0783[ia,ib]	0.8714	–42.5467[h]	–0.3001	1.2683	0.7800	1.0816	0.8714	–42.5465	–0.3001
94/98	1.2728	0.9615	5.3766[ia–ic]	2.4683	–34.4268[h]		1.2728	0.9615	5.5077	2.4682	–34.4266	
96	1.2728	0.9615	4.7803[ia–ic]	2.4683	–34.4268[h]		1.2728	0.9615	4.9115	2.4682	–34.4266	
99	1.2728	1.0040	12.1947[ia–ic]	2.4683	–34.4268[h]		1.2728	1.0040	12.6452	2.4682	–34.4266	
Model **12**:	Alanine dipeptide (C7ax conformation)[g]											
AMBER force fields												
91X	1.2402	2.3598	0.6058	0.4106	–42.3923[h]	–0.4590	1.2402	2.3598	0.6058	0.4106	–42.3921	–0.4590
94/98	1.2452	2.7583	5.7035[ia–ic]	2.2269	–35.0120[h]		1.2452	2.7583	5.7045	2.2269	–35.0118	
96	1.2452	2.7583	5.7192[ia–ic]	2.2269	–35.0120[h]		1.2452	2.7583	5.7202	2.2269	–35.0118	
99	1.2452	2.8708	12.9007[ia–ic]	2.2269	–35.0120[h]		1.2452	2.8708	13.2212	2.2269	–35.0118	
Model **13**:	PDB code: 1HO0, model_**4**											
AMBER force fields												
91X[j]	93.976	176.572	27.653[ia,ib]	–57.976	–657.835[h]	–2.641	93.976	176.572	27.693	–57.977	–657.832	–2.641
94/98	95.274	184.037	66.268[ia–ic]	–63.519	–433.842[h]		95.274	184.037	144.844	–63.519	–433.839	
96	95.274	184.037	89.945[ia–ic]	–63.519	–433.842[h]		95.274	184.037	132.388	–63.519	–433.839	
99	95.274	198.338	166.662[ia–ic]	–63.519	–433.842[h]		95.274	198.338	277.876	–63.519	–433.839	

[a] Software=Sander_classic (AMBER, version 6) and CDiscover (Accelrys, version 2000)
[b] E=Single point energy values (in kcal mol^{-1}) using the different AMBER force fields (no cutoff, and in vacuo simulation)
[c] Dih=(torsion+improper torsion) interactions
[d] vdW=(scaled 1–4 van der Waals+van der Waals) interactions. 1–4 van der Waals interactions are scaled by 0.5 in all the AMBER force fields
[e] Coul=(scaled 1–4 electrostatic+electrostatic) interactions. To calculate electrostatic energy of the seven organic models, ESP and RESP charges were calculated from STO-3G (Amber 91 and Amber 91X force field versions) and 6-31G* (Amber 94→Amber 99 force field versions) molecular electrostatic potentials using the GAMESS and Gaussian 98 software, respectively. On the contrary, for the two nucleosides, the three alanine dipeptide conformations and the 1HO0 PDB structure (model **4**), the charges were taken from the software topology databases. 1–4 electrostatic interactions are scaled by 0.5 in the Weiner et al. force field (Amber 91 and Amber 91X force field versions) and by 1/1.2 in the Cornell et al. force field (Amber 94→Amber 98 force field versions) and Wang et al. force field (Amber 99 force field version)
[f] HB=hydrogen bond interaction (only calculated in the Weiner et al. force field)
[g] The nucleoside conformations and the organic molecules (cyclohexane, methanol, ethylmethylether and *N*-methylacetamide) were optimized using the 6-31G* basis set (with Gaussian 98) while the alanine dipeptide conformations were minimized using the 6-31G** basis set (with PC-GAMESS)
[h] Different round-offs in the electrostatic energy calculations leads to small energy value differences in the two pieces of software
[ia] Difference between AMBER and Accelrys software to measure improper torsion geometry
[ib] CDiscover improper torsion bug
[ic] CDiscover torsion bug
[j] The CT–C–OH and O–C–OH angle and CT–O–'C'–OH (AMBER naming convention) improper torsion force field parameters were taken from the amber94.frc/parm94.dat force field files and manually added in the amber91X.frc/parm91X.dat ones to allow the energy calculation

abase, or residue library (".rlb" file), [44, 45] using two different approaches for the AMBER and GLYCAM force fields. The Accelrys residue topology database amberXX.rlb files are obtained using Accelrys "format donor" input files and AMBER files [31, 34] containing target values. Actually, for each amino acid or nucleotide found in the AMBER topology database, the atom types, partial charges, and out of plane centers are selected and added at their corresponding place in the Accelrys input file. Since no monosaccharide topology database is available in the Accelrys software package, the glycamXX.rlb files were obtained directly from the numerous ".prep" files provided in each GLYCAM force field version. [24, 25, 35] Third, the Accelrys pseudo atom library (".plb" file) is created using Accelrys input files. [44, 46] For each unit present in the amberXX.rlb or glycamXX.rlb file previously created, pseudo atoms are generated using the CVFF force field pseudo atom library as template. [14, 47] Explanations describing these first three steps are presented in Table 2. In the last two steps, AmberFFC generates an empty_templates.dat file, [48, 49] and, if appropriate, automatically executes Accelrys scripts to build binary versions of force field files. [36, 38, 42]

Computational methods

Single point energy values for 13 structures computed with the Accelrys and AMBER software were obtained on the same workstation, an SGI Indigo2 R4400, to compare the implementation of the studied force fields in the two packages and also to verify the validity of the force fields converted to the Accelrys format. These 13 mole-

Table 4 Comparison of the non-bond cutoff effect in the AMBER and Accelrys molecular mechanics modules

CV[a]	ABC[b]		GBC[c]		RBC[d]	
(Å)	vdW[e]	Coulomb[f]	vdW[e]	Coulomb[f]	vdW[e]	Coulomb[f]
	(kcal mol^{-1})		(kcal mol^{-1})		(kcal mol^{-1})	
Deoxyadenosine (C2'endo conformation)[g]						
4	8.5032	–108.7056	8.2760	–74.3178	5.7278[h]	–65.4017[h]
5	6.4658	–9.0756	8.2760	–74.3178		
6	5.9321	–101.8417	5.7277[h]	–65.4021[h]		
7	5.7909	–40.7249				
8	5.7395	–41.1492				
9	5.7282	–54.7445				
10	5.7277[h]	–66.4024				
11	5.7277[h]	–65.4021[h]				
Alanine dipeptide (C5 conformation)[g]						
4	4.0859	–27.6407	3.7681	–34.0743	3.2588	–30.8068
5	3.2573	–45.7158	3.2588	–30.8069	3.1056[h]	–31.3711[h]
6	3.1475	–26.7872	3.2588	–30.8069		
7	3.1143	–31.3734	3.2588	–30.8069		
8	3.1059	–33.1066	3.1057[h]	–31.3712[h]		
9	3.1057[h]	–31.3712[h]				

[a] Cutoff value
[b] Atom based cutoff (Accelrys/CDiscover)
[c] Group based cutoff (Accelrys/CDiscover)
[d] Residue based cutoff (AMBER/Sander_classic)
[e] vdW=(scaled 1–4 van der Waals+van der Waals) interactions. 1–4 van der Waals interactions are scaled by 0.5 in all the AMBER force fields
[f] Coulomb=(scaled 1–4 electrostatic+electrostatic) interactions. The 1–4 electrostatic interactions are scaled by 1/1.2 in the Wang et al. force field. The charges were taken from the software topology databases
[g] Single point energy value using the Wang et al. force field (Amber 99 force field version), in vacuo simulation
[h] Nonbond energy values corresponding to the ones obtained without cutoff

cules were studied using the AMBER force fields as models within the Sander_classic (AMBER version 6 [22, 23]) and CDiscover (Accelrys package, version 2000 [36, 37]) molecular mechanics programs. As FDiscover does not handle torsion terms with periodicities of n=4, [38, 39] it was only used to confirm CDiscover results obtained with the Weiner et al. force field. Three families of molecules were studied: (i) seven organic structures (models **1**–**7**, Table 3: cyclohexane chair and boat conformations, ethylmethylether *trans* and *gauche* conformations, *N*-methylacetamide *trans* and *cis* conformations, and methanol), (ii) two nucleoside structures (models **8** and **9**, Table 3: deoxyadenosine C2'endo and C3'endo conformations), and (iii) four peptide molecules (models **10**–**13**, Table 3: alanine dipeptide C5, C7eq, and C7ax conformations, and the 1HO0 protein data bank (PDB) structure [50, 51]).

Ab initio geometry optimizations were performed to locate the target energy minima for models **1**–**12**. The organic and nucleoside molecules were minimized using the 6-31G* basis set as implemented in Gaussian 98, [52] while the alanine dipeptide conformations were minimized using the 6-31G** basis with PC-GAMESS. [53, 54] The 1HO0 PDB structure was determined by a nuclear magnetic resonance (NMR) and simulated annealing (SA) study carried out earlier in our laboratory. [51]

ESP [55] and RESP [56, 57, 58] charges for the organic models were calculated from molecular electrostatic potentials based on either the STO-3G (GAMESS [53, 59]) or 6-31G* (Gaussian 98) basis sets. On the other hand, charges for the nucleosides, alanine dipeptide conformers, and the 1HO0 PDB structure were taken from the force field topology databases. When force field parameters were missing in a force field, they were incorporated by hand in the corresponding ".frc" (Accelrys software) and in the ".dat" (AMBER software) files. Thus, to obtain energy values for the 1HO0 PDB structure using the Weiner et al. force field (Amber 91X version), the CT–C–OH and O–C–OH angle and CT–O–'C'–OH improper torsion (AMBER naming convention [60]) force field parameters were taken from the amber94.frc and parm94.dat files and manually added in the amber91X.frc and parm91X.dat files, respectively. Moreover, for non-peptide and non-nucleotide models such as organic molecules (Table 3, models **1**–**7**), new residue library tables [44, 45] (Accelrys) and new LEaP OFF libraries [61] (AMBER software) were built by matching the residue and atom names, atom types, charges, and Cartesian coordinates of each structure.

Molecular mechanics energy calculations were performed in vacuo and without cutoffs using the Amber 91X, Amber 94, Amber 96, Amber 98, and Amber 99

force field versions in AMBER and Accelrys software. 1–4 van der Waals interactions were scaled by 0.5 in all calculations, while 1–4 electrostatic interactions were scaled by either 0.5 or 1/1.2, depending on the force field used in the calculations. Bond, angle, dihedral, van der Waals, electrostatic, and hydrogen bond interaction energies were obtained with an accuracy of 10^{-4} in AMBER and Accelrys software. When an energy difference between the two programs was observed, the corresponding energy values were decomposed into elementary components. Thus, for the dihedral interaction, Anal (AMBER software) was recompiled to study all the elementary torsional energies with a 10^{-4} accuracy, and a BTCL script (CDiscover/Accelrys) was used to characterize, [36, 62] with the same accuracy, the dihedrals presenting energy differences. Additionally, atom based and group based cutoff (CDiscover) and residue based cutoff (Sander_classic) simulations were compared to evaluate their effect on the electrostatic energy reported in the two programs (Table 4).

Results and discussion

AMBER and GLYCAM force field versions

About 15 years ago, Weiner et al. developed a molecular mechanics force field suitable for the study of proteins and nucleic acids. [4, 5] This force field had two versions, one that used united atoms and implicit inclusion of hydrogens on carbon, and a second one in which all atoms were explicitly represented. These force fields reproduce reasonably well the vibrational frequencies and energies of small model molecules. The empirical expression of the Weiner et al. force field was presented above (Eq. 1). Atom centered charges fitted to electrostatic potentials were derived from ab initio calculations at the STO-3G level, [55] and a weak 10–12 term for hydrogen bonds was used. The 1–4 electrostatic and van der Waals interactions are reduced by the application of scale factors, which are set to 1/2. This force field is still used for vacuum simulations using a distance-dependent dielectric constant. Ten years later, Cornell et al. presented a second generation force field for the simulation of proteins, nucleic acids, and organic molecules. [6] This force field has been especially well suited for the study of solvated systems. It uses a similar potential model to the Weiner et al. force field with a number of modifications: (i) the hydrogen bond potential term has been omitted; (ii) a new restrained electrostatic potential charge approach, derived from ab initio calculations at the 6-31G* level of theory, is used [56, 57, 58]; (iii) new van der Waals parameters have been adopted to reproduce liquid properties better; (iv) scale factors of 1/1.2 and 1/2 are used for the 1–4 electrostatic and van der Waals interactions, respectively; and (v) a periodicity of n=4 has been included in the torsion potential for some dihedrals. The Cornell et al. force field was updated twice since its original release. Kollman et al. [32] adjusted the peptide $\langle\phi,\psi\rangle$ dihedrals based on the work of Beachy et al., [63] and Cheatham et al. modified the glycosidic dihedral potential of nucleic acids. [33] Finally, Wang et al. published the last AMBER force field version available. [7] It presents improved torsional potentials for hydrocarbon molecules, and new force field parameters were developed for new chemical groups.

In 1993, Woods et al. reported the first version of the GLYCAM force field, named GLYCAM_93. [24] In this work, the authors concentrated their efforts on the torsion energy profiles and 6-31G* ESP partial atomic charges development. Van der Waals parameters were taken from the Weiner et al. force field, while bond and angle force field parameters were obtained from a CHARMM-type force field. [64] In 2000, Pathiaseril et al. published two updated versions of the GLYCAM force field. [25] In the first one, GLYCAM_98, (or glycam_2000 in the AmberFFC naming convention), the force field parameter set has been modified to be consistent with the Cornell et al. force field [6] and in the second one, GLYCAM_98R (or glycam_2010 in the AmberFFC naming convention), new RESP charges were developed employing a single stage fitting and a restraint weight of 0.01.

Force fields converted by AmberFFC and comparison between AMBER and Accelrys force fields

Accelrys only provides the AMBER Weiner et al. force field as an extra in the software distribution package. [14] Homans's force field parameters [65] have been added in this force field version allowing the program to model glycoconjugates. However, this Accelrys Weiner et al. force field shows several differences from the original Weiner et al. force field. The partial charges of the charged amino acid termini, some force field parameters, and some out of plane centers are different, and some amino acids are missing in the topology database, making its comparison to the standard force field difficult. Furthermore, no monosaccharide residue library file is provided in the Accelrys software package for Homans's carbohydrate force field. Thus, AmberFFC has been written to carefully transfer the properties of all the AMBER and GLYCAM force fields available in the public domain to the Accelrys format. The program allows the modeler to employ the AMBER and GLYCAM force fields with the Accelrys package, as well as to compare the results obtained in the AMBER and Accelrys software. Finally, we have decided not to convert the Weiner et al. united atom force field, as this force field version is becoming somewhat old and inaccurate when compared to the more recent AMBER force field versions.

Molecular mechanics results

AMBER force fields were used as models to compare the implementation in the Accelrys software package of

the force fields converted by AmberFFC with their corresponding execution in the AMBER software, and to verify the efficiency of the AmberFFC program. Thirteen structures were chosen for a single point energy comparative study where AMBER force fields were used in the Sander_classic and CDiscover molecular mechanics programs. Molecules **1**–**12** (Table 3) were taken as models because they have been widely used as key structures in the development of AMBER force field parameters, [4, 5, 6, 7, 32, 33] and polypeptide **13**, (Table 3), [51] available from our laboratory, was selected because of its ability to test a large number of different force field parameters simultaneously.

The energy calculations based on the Weiner et al., [4, 5] Cornell et al. [6] (with its two modifications [32, 33]), and Wang et al. [7] force fields (Amber 91X, Amber 94, Amber 96, Amber 98, and Amber 99 force field versions) were obtained in vacuo without the use of cutoffs. The energy values of each molecular interaction obtained with the two programs are reported in Table 3. The results were as follows. First, the bond, angle, van der Waals, and hydrogen bond interactions give equivalent energy values with an accuracy of 10^{-4} in all cases. Secondly, the electrostatic interaction shows small energy differences of $\pm 4\times10^{-4}$ for models **1**–**12**. Electrostatic energy differences were also reported when the Cornell et al. force field was manually converted to the CHARMM format. [33] Indeed, the authors have demonstrated that to achieve an accuracy better than 10^{-4} for the electrostatic interactions, the $4\pi\varepsilon_0$ constant needed to be modified from the CHARMM value (332.0716) to the AMBER value (332.0522173). As in the case of the AMBER/CHARMM comparison, negligible AMBER/Accelrys electrostatic energy differences could arise from small round-off differences. For instance, the use of 0.8333, instead of 1/1.2, as the electrostatic scaling factor for the Cornell et al. and Wang et al. force fields in CDiscover increases this difference by a factor of 2 for the models **10**–**12**. Finally, the dihedral interaction reveals inconsistent results. Indeed, models **1**–**7** show equivalent dihedral energy values while models **8**–**12** present non-negligible dihedral energy differences. To study these differences between the AMBER and Accelrys software, the dihedral energies of models **8** and **10** were decomposed into elementary values using the Weiner et al. (Amber 91X version) and Wang et al. force fields. Two important differences were observed. (i) The improper torsions are treated differently, leading in some cases to out of plane energy differences between the two programs. In AMBER, the convention for an improper torsion named A–B–'C'–D is that the out of plane center is listed in the third position and the order of the other three is determined alphabetically by atom type, and by atom number (i.e., their order in the molecule) when atom types are identical. Such a convention guarantees that the program always gives the same energies for a certain improper torsion. [60] On the contrary, in CDiscover and FDiscover the out of plane center comes second in the list and the improper term is asymmetric with respect to the three outer atoms. [14, 66] Thus, for each out of plane center, six geometry values can be used to calculate the improper torsion energy. CDiscover and FDiscover choose one geometry value from the other five by the order the atoms, appearing in the Cartesian coordinate (.car) file. [44, 67] Although the approaches of the two programs are different, both are correct, but lead in some cases to minor energy differences. Using the Amber 91X force field version, total dihedral energy differences between the AMBER and Accelrys software for models **8** and **10** are 0.0092 and 0.0006 kcal mol^{-1}, respectively. For model **8**, the energy difference, comes from a unique CA–H2–'N2'–H2 improper torsion (AMBER naming convention), which corresponds to an AMBER energy value of 0.3985 kcal mol^{-1} (out of plane geometry value=206.51°), while CDiscover and FDiscover report a value 0.4077 kcal mol^{-1} (out of plane geometry value=153.16°). For alanine dipeptide (model **10**), out of four improper torsions calculated, a difference has been observed for only one of the two C–CT–'N'–H out of plane centers. Surprisingly, for this improper torsion, Anal and FDiscover calculate the same energy value (0.0048 kcal mol^{-1}) while CDiscover reports a value of 0.0053 kcal mol^{-1}. Thus, these energy differences between the two programs originate from a bug in the CDiscover package. (ii) Another difference between Accelrys and AMBER software concerns the torsion energy. Indeed, although it has been described that the program can handle torsion terms with a periodicity n=4, [14, 68] errors have been observed for dihedral energies calculated with CDiscover. As examples, using the Wang et al. force field, the dihedral energy differences between the AMBER and Accelrys software for models **8** and **10** are 1.2615 and 0.3241 kcal mol^{-1}, respectively. Out of 86 torsions and five improper torsions calculated in model **8**, five torsions and one improper torsion show different energy values. For two CT–CT–OH–HO, one CT–OS–CT–CT, one OS–CT–CT–OH, and one CT–CT–CT–CT torsions, the energies are miscalculated by CDiscover. For each of these five dihedrals, the potential is deconvoluted into two or three Fourier series in the Wang et al. force field. In CDiscover, only the cosine function with the smallest periodicity is calculated, leading to underestimated torsion energy values. Thus, energy differences of –0.0054, –0.0027, –0.766, –0.1953, and –0.3013 kcal mol^{-1} were obtained for the dihedrals mentioned above. Moreover, the CA–H–'N2'–H improper torsion energy value observed in CDiscover has an energy difference of +0.0092 kcal mol^{-1} with respect to the value calculated by Sander_classic. The sum of these six differences gives the exact dihedral energy difference observed between the two programs. Similar errors have been characterized for model **10**: out of 41 torsions and four improper torsions calculated for this model, two HC–CT–C–O, one O–C–N–H torsions, and one C–CT–'N'–H improper torsion show energy differences of –0.1595, –0.1600, –0.0053, and +0.0007 kcal mol^{-1}, respectively. Once again, the sum of the errors corresponds to the total di-

hedral energy difference observed between the two programs. Surprisingly, the CT–CT–N–C and CT–CT–C–N dihedrals that present torsion terms with a periodicity n=4 are well calculated by CDiscover for models **10**–**12**. Thus, a second CDiscover bug has been detected for the case where several Fourier series are used for a dihedral when a Vn (n=1 to 4) dihedral table is built in an amberXX.frc or glycamXX.frc Accelrys file.

In the AMBER force field parameter development procedure, a minimalist approach has been applied. [33, 60] Indeed, the philosophy of the Kollman group is to focus mainly on the determination of atomic charges, van der Waals, and dihedral parameters, because these three interactions are the primary contributors to the relative conformational energy of a molecule. Thus, and since the CDiscover bug could have an important effect on the conformational energy of a large system, the dihedral energy difference between the Sander_classic and CDiscover programs for model **13** was studied. The energy values of each molecular interaction are presented in Table 3. It is clear that for a larger structure the total dihedral energy is totally different, and thus could have a tremendous effect on the conformational energy and, more importantly, on the conformational behavior of molecules studied with CDiscover. On the contrary, this model proved that all other interaction energies computed by the two programs are identical, and that the electrostatic energy shows the same energy difference values as the ones observed for the models **1**–**12**. Thus, our work showed not only that AmberFFC is an efficient AMBER force field converter, but also that there are two bugs in the Accelrys software that need to be corrected if the most recent AMBER and GLYCAM force field parameters are to be used with CDiscover.

A fundamental difference in the implementation of force fields between AMBER and Accelrys packages is in relation to the handling of non-bonded distance cutoffs. As calculating all non-bond interactions is computationally expensive for large molecules, molecular mechanics programs use cutoffs to neglect non-bonded interactions for pairs of atoms separated by distances greater than a certain value. Such cutoff distances are software dependent: AMBER package uses residue based cutoffs, [23] while Accelrys molecular mechanics modules use either atom or group based cutoffs. [14] As the van der Waals potential falls off as $1/r^6$, using a cutoff distance of 10 Å in our comparisons is reasonable. On the other hand, electrostatic interaction energies decrease as 1/r, leading to non-negligible effects even at long distances. However, most molecules are composed of neutral fragments with dipoles and quadrupoles. Then, dipole–dipole interactions, which decrease as $1/r^3$, are the main contributors to the electrostatic term. Using the atom based cutoffs in molecular mechanics simulations would break molecular dipoles into monopoles responsible for artificially high electrostatic interaction. Thus, it is advised to employ group based cutoffs in CDiscover and FDiscover, [14] while Sander_classic uses residue based cutoffs by default. To evaluate the cutoff effects on the electrostatic energy, atom and group based cutoffs in CDiscover and residue based cutoffs in Sander_classic were studied using structures **8** and **10** as models with the Wang et al. force field. Single point energy values with and without cutoffs are reported in Table 4. It is clear that using short distances in atom based cutoffs results in erroneous electrostatic energies due to monopole–monopole interaction artifacts. On the other hand, using short distances with group based cutoffs in CDiscover results in non-bonded energy values which are similar to those obtained using residue based cutoffs in Sander_classic. Although a cutoff distance between 10 and 12 Å is reasonable for both programs, non-bonded energy differences may subsist for larger structures such as 1HO0 PDB. Indeed, the divergent treatment of the non-bonded distance cutoff may lead to some electrostatic energy differences. Thus, each user should be aware of this problem before using the converted AMBER and GLYCAM force fields with Accelrys modules. However, modifying the switching atoms and the charge groups in the Accelrys residue topology database may allow users to obtain more "AMBER-like" electrostatic energy values. [44, 45] For instance, AmberFFC generates two switching atoms and two charge groups for each amino acid, one for the backbone and one for the side chain. Replacing these two switching atoms by a single one in the middle of the residue may lead to a "residue-like" based cutoff treatment in Accelrys molecular mechanics modules. Work to address these issues is currently in progress.

Conclusion

AmberFFC is a force field format converter which in the current version allows the modeler to use most of the AMBER and GLYCAM force field versions available in the public domain with the CDiscover and FDiscover modules of the Accelrys molecular modeling package. Easy to use, flexible, and highly portable to different operating systems and platforms, AmberFFC is freely available for academic laboratories through the world-wide web at the site http://www.u-picardie.fr/UPIC/UPJV/recherche/labos/bpd/AmberFFC.htm after signing a license.

The comparison of results from single point energy calculations performed using Accelrys programs with force field files converted with AmberFFC against those obtained using AMBER software indicates that the conversion program is highly efficient. We found no discrepancies for bond stretching or angle bending energy terms. For torsion and improper torsion energy terms, differences were observed originating from the presence of two CDiscover bugs that we were able to unambiguously identify through our study. Finally, the non-bonded distance cutoff effect on the electrostatic interaction has been studied in the two programs, the differences have been characterized, and their implications discussed.

We are currently working on the expansion of the capabilities of AmberFFC. New versions of AmberFFC will be able to convert AMBER and GLYCAM force fields for

use with Sybyl, [69] Spartan, [70] and Hyperchem [71] software. In some of these cases, AmberFFC will have to be supplemented with C++ function calls to create the necessary binary files to be employed by these packages, which should be straightforward once the structure of the binary force field files is elucidated. These additional capabilities will make AmberFFC a "pseudo-BABEL" program for force field conversion, [72, 73] which will enable scientists to employ newly developed force fields with currently available molecular modeling packages.

Acknowledgments We would like to thanks Dr R. J. Woods (Complex Carbohydrate Research Center, University of Georgia, Athens, Ga.) for providing us the new versions of the GLYCAM force fields, Dr Jodi Shaulski (Accelrys support team) for helpful discussion regarding the Accelrys force fields, Dr M.-A. Delsuc (Université de Montpellier, France) for his advice regarding licensing and copyright issues, I. Cannoux-Dupradeau for her assistance in writing the manuscript, and Professor J. Rochette for insightful comments. We would also like to acknowledge L. Fin (Laboratoire de Biophysique et Traitement d'Images Médicales, Amiens) and C. Poix (Direction des Services Techniques Centraux, Amiens) for allowing us to test AmberFFC under SUN, Hewlett-Packard and IBM RS/6000 workstations. F.-Y. D. is grateful to the Université de Picardie Jules Verne for financial support.

References

1. Eliel E, Allinger NL, Angyal SJ, Morrison GA (1965) In: Conformational analysis. Wiley, New York
2. Brooks CL, Karplus M, Pettit BM (1988) In: Proteins: a theoretical perspective of dynamics, structures, and thermodynamic, vol. LXXI. Wiley, New York
3. French AD, Brady JW (1990) In: Computer modeling of carbohydrate molecules. American Chemical Society, Washington, D.C.
4. Weiner SJ, Kollman PA, Case DA, Singh UC, Ghio C, Alagona G, Profeta S Jr, Weiner PA (1984) J Am Chem Soc 106:765–784
5. Weiner SJ, Kollman PA, Nguyen DT, Case DA (1986) J Comput Chem 7:230–252
6. Cornell WD, Cieplak P, Bayly CI, Gould IR, Mertz Jr KM, Ferguson DM, Spellmeyer CD, Fox T, Caldwell JW, Kollman PA (1995) J Am Chem Soc 117:5179–5197
7. Wang J, Cieplak P, Kollman PA (2000) J Comput Chem 21:1049–1074
8. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M (1983) J Comput Chem 4:187–217
9. Smith JC, Karplus M (1992) J Am Chem Soc 114:801–812
10. Schmidt JM, Brüschweiler R, Ernst RR, Dunbrack RL Jr, Joseph D, Karplus M (1993) J Am Chem Soc 115:8747–8756 and references cited therein
11. MacKerell Jr AD, Brooks B, Brooks III CL, Nilsson L, Roux B, Won Y, Karplus M (1998) The energy function and its parameterization with an overview of the program. In: Schleyer PvR, et al (eds) The encyclopedia of computational chemistry, vol. 1. Wiley, Chichester, pp 271–277
12. Mapple JR, Dinur U, Hagler AT (1988) Proc Natl Acad Sci USA 85:5350–5354
13. Mapple JR, Hwang MJ, Stockfisch TP, Dinur U, Waldman M, Ewig CS, Hagler AT (1994) J Comput Chem 15:162–182
14. Discover 2.9.7/95.0/3.0.0 (1995) Force field simulations, user guide, part 1. Biosym/Molecular Simulations, San Diego, Calif.
15. van Gunsteren WF, Berendsen HJC (1987) Groningen molecular simulations (GROMOS) library manual. Biomos, Groningen
16. Halgren TA (1992) J Am Chem Soc 114:7827–7843
17. Halgren TA, Nachbar RB (1996) J Comput Chem 19:587–615
18. Allinger NL (1977) J Am Chem Soc 99:8127–8134
19. Allinger NL, Yuh YH, Lii J-H (1989) J Am Chem Soc 111: 8551–8566
20. Lii J-H, Allinger NL (1989) J Am Chem Soc 111:8566–8575
21. Lii J-H, Allinger NL (1989) J Am Chem Soc 111:8576–8582
22. Pearlman DA, Case DA, Caldwell JW, Ross WS, Cheatham III TE, DeBolt S, Ferguson D, Seibel G, Kollman PA (1995) Comput Phys Commun 91:1–41
23. Case DA, Pearlman DA, Caldwell JW, Cheatham III TE, Ross WS, Simmerling CL, Darden TA, Merz KM, Stanton RV, Cheng AL, Vincent JJ, Crowley M, Tsui V, Radmer RJ, Duan Y, Pitera J, Massova I, Seibel GL, Singh UC, Weiner PK, Kollman PA (1999) AMBER 6. University of California, San Francisco, Calif.
24. Woods RJ, Dwek RA, Edge CJ, Fraser-Reid B (1995) J Phys Chem 99:3832–3846
25. Pathiaseril A, Woods RJ (2000) J Am Chem Soc 122:331–338 and references cited therein
26. Wall L, Christiansen T, Schwartz RL (1996) Programming perl, 2nd edn. O'Reilly, Cambridge
27. Schwartz RL, Christiansen T (1997) Learning perl, 2nd edn. O'Reilly, Cambridge
28. Christiansen T, Torkington N (1998) Perl cookbook. O'Reilly, Cambridge
29. http://www.perl.org
30. http://www.gnu.org
31. http://www.amber.ucsf.edu/amber/ff91.html
32. Kollman PA, Dixon R, Cornell W, Fox T, Chipot C, Pohorille A (1997) The development/application of a "minimalist" organic/biochemical molecular mechanics force field using a combination of ab initio calculations and experimental data. In: Wilkinson A, Weiner P, van Gunsteren W (eds) Computer simulation of biomolecular systems, vol 3. Kluwer, Dordrecht, pp 83–96
33. Cheatham TE III, Cieplak P, Kollman PA (1999) J Biomol Struct Dyn 16:845–861
34. http://www.amber.ucsf.edu/amber/ff94.html
35. http://www.amber.ucsf.edu/amber/ff94_glycam.html
36. Discover 95.0/3.0.0 (1995) Force field simulations, user guide, part 3. Biosym/Molecular Simulations, San Diego, Calif.
37. http://www.accelrys.com/doc/life/insight2K/cdiscover/CDisTOC.doc.html
38. Discover 2.9.7 (1995) Force field simulations, user guide, part 2. Biosym/Molecular Simulations, San Diego, Calif.
39. http://www.accelrys.com/doc/life/insight2K/fdiscover/FDisTOC.doc.html
40. System guide 2000 (2000) for InsightII products. Molecular Simulations, San Diego, Calif.
41. http://www.accelrys.com/doc/life/insight2K/sysguide/requirements.html#911584
42. http://www.accelrys.com/doc/life/insight2K/fdiscover/9_Application_Standalone.html#73592
43. http://www.accelrys.com/doc/life/insight2K/cdiscover/D_Files.html#925808
44. Common file format 95.0 (1995) for Biosym/MSI release 95.0 products, reference guide. Biosym/Molecular Simulations, San Diego, Calif.
45. http://www.accelrys.com/doc/life/insight2K/formats980/File_Formats_1998.html#144453
46. http://www.accelrys.com/doc/life/insight2K/formats980/File_Formats_1998.html#779673
47. Dauber-Osguthorpe P, Roberts VA, Osguthorpe DJ, Wolff J, Genest M, Hagler AT (1988) Proteins: Structure, Function and Genetics 4:31–47
48. Insight II 95.0 (1995) Molecular modeling system, user guide. Biosym/Molecular Simulations, San Diego, Calif.
49. http://www.accelrys.com/doc/life/insight2K/insight/a_PTR.html#309226
50. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) Nucleic Acids Res 28:235–242

51. Dupradeau FY, Richard T, Le Flem G, Oulyadi H, Prigent Y, Monti JP (2001) J Pept Res (2001) submitted
52. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Zakrzewski VG, Montgomery JA, Stratmann RE, Burant JC, Dapprich S, Millam JM, Daniels AD, Kudin KN, Strain MC, Farkas O, Tomasi J, Barone V, Cossi M, Cammi R, Mennucci B, Pomelli C, Adamo C, Clifford S, Ochterski J, Petersson GA, Ayala PY, Cui Q, Morokuma K, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Cioslowski J, Ortiz JV, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Gomperts R, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Gonzalez C, Challacombe M, Gill PMW, Johnson BG, Chen W, Wong MW, Andres JL, Head-Gordon M, Replogle ES, Pople JA (1998) Gaussian 98. Gaussian, Pittsburgh Pa.
53. Schmidt MW, Baldridge KK, Boatz JA, Elbert ST, Gordon MS, Jensen JH, Koseki S, Matsunaga N, Nguyen KA, Su SJ, Windus TL, Dupuis M, Montgomery J (1993) J Comput Chem 14:1347–1363
54. Granovsky AA http://classic.chem.msu.su/gran/gamess/
55. Singh UC, Kollman PA (1984) J Comput Chem 5:129–145
56. Cornell WD, Cieplak P, Bayly CI, Kollman PA (1993) J Am Chem Soc 115:9620–9631
57. Bayly CI, Cieplak P, Cornell WD, Kollman PA (1993) J Chem Phys 97:10269–10280
58. Cieplak P, Cornell WD, Bayly C, Kollman PA (1995) J Comput Chem 16:1357–1377
59. http://www.msg.ameslab.gov/GAMESS/GAMESS.html
60. http://www.amber.ucsf.edu/amber/newparams.html
61. Schafmeister CEAF, Ross WS, Romanovski V (1995) LeaP. University of California, San Francisco, Calif.
62. http://www.accelrys.com/doc/life/insight2K/cdiscover/4_DiscAlone.html#309934
63. Beachy MD, Chasman D, Murphy RB, Halgren TA, Friesner RA (1997) J Am Chem Soc 119:5908–5920
64. Ha SN, Giammonia A, Field M (1988) Carbohydr Res 180: 207–221
65. Homans SW (1990) Biochemistry 29:9110–9118
66. http://www.accelrys.com/doc/life/insight2K/cdiscover/D_Files.html#926044
67. http://www.accelrys.com/doc/life/insight2K/formats980/File_Formats_1998.html#781840
68. http://www.accelrys.com/doc/life/insight2K/cdiscover/D_Files.html#931125
69. Tripos Inc, 1699 South Hanley Road, St. Louis, MO 63144, USA, http://www.tripos.com
70. Hypercube Inc, 115 NW 4th Street, Gainesville, FL 32601, USA, http://www.hyper.com
71. Wavefunction Inc, 18401 Von Karman Avenue, Suite 370, Irvine, CA 92612, USA, http://www.wavefunction.com
72. Shah AV, Walters WP, Shah R, Dolata DP (1994) Babel, A tool for converting between molecular structural data formats. In: Lysakowski R, Gragg CE (eds) Computerized chemical data standards: databases, data interchange, and information systems. American Society for Testing and Materials, Philadelphia, Pa. pp 45–53
73. http://www.eyesopen.com/babel.html