Journal of Molecular Modeling

An Evaluation of the MM+ Force Field

Alexandre Hocquet[1] and Morten Langgård[2,#]

[1] Facultad de Ciencias Fisicas, Universidad de Chile, Santiago de Chile, Chile. E-mail: hocquet@rip.ens-cachan.fr

[2] Department of Life Sciences and Chemistry, Roskilde University, Post box 260, Dk-4000 Roskilde, Denmark.
E-mail: ml@virgil.ruc.dk

Abstract Hyperchem's MM+ force field, based on Allinger's MM2 is described and evaluated with respect to other MM2 variants, in terms of rotation barriers, conformational energy differences, and conjugation. Its ability to take missing parameters into account is also evaluated with respect to the Dreiding force field. This evaluation also intends to clearly separate the two different force fields MM+(91) and MM+(**) hiding under the MM+ denomination. It is shown that, whereas MM+ proves to be generally robust, caution must be the rule when dealing with conjugated molecules, particularly with heteroaromatics.

Keywords Molecular mechanics, Force Field Evaluation, MM+

Introduction

1997 represents the twentieth anniversary of MM2's first publication [1]. Since then, MM2 has become the most popularly used force field, and even though newer force fields may perform better, the success of its wide utilization makes it a reference. MM2 still remains the force field with the more diverse parameter set [2], and parameterization optimization of MM2 is still an up to date research interest in inorganic [3] and even in organic [4] chemistry (Figure 1).

The excellent performance of MM2 and the public domain status of the '77 version, has caused a variety of MM2 variants to be developed from this source. Despite the fact that the MM2 successor, MM3, was released eight years ago some of the MM2 variants has been published very recently [5]. Besides being MM2 clones, many of the variants has their own characteristics such as extensions for handling metal-complexes, molecular dynamics and open-shell systems.

The object of this study is to evaluate one of the MM2 variants, called MM+. MM+ is implemented in the molecular modeling package HyperChem [6]. Because of the commercial success of HyperChem, this force field is now being widely used. To our knowledge no publication has yet been dealing with the performance of MM+. Here we wish to provide a test of how MM+ compares to other force fields for subsequent proper utilization of the force field.

Correspondence to: A. Hocquet, 7bis rue Alexandre Parodi, F-75010 Paris, France
 # Present address: H. Lundbeck A/S, Ottiliavej 9, Dk-2500 Valby, Denmark. E-mail: mol@lundbeck.com

MM+ brief description

MM+ is thoroughly described in HyperChem's manual «Computational Chemistry» [6]. In short, MM+ is based on the functional form of MM2(77) but uses the more recent MM2(91) parameter set. The MM+ implementation does however present some differences and extensions compared to the original MM2. The MM+ code has for example, in addition to a "missing parameter generator" also been modified to allow molecular dynamics calculations.

Like for the vast majority of MM2 variants, there is no "MMP like" SCF procedure for treatment of conjugated pi systems implemented in MM+. To overcome some of the difficulties this omission causes, MM+ has a simple bond order scheme that can be utilized for conjugated systems.

For the purpose of molecular dynamics, the stretch energy term has been modified from the original to prevent the "cubic stretch" catastrophe, where the energy falls to negative infinity when distance increases too much. Furthermore, the non-bonded terms are modified to allow a cut off for proper treatment of periodic boundary conditions. Whereas an option for replacing the bond dipole interactions term with a point charges interaction term has been included in MM+, the correction terms for electronegativity differences and the anomeric effect have not.

The parameter estimator implemented in MM+ serves to overcome the well known "missing" parameter problem often encountered with MM2. Although a default parameter scheme has been published from Allinger's group [7] and implemented in MM3, this has not been done for MM2. The concept for this estimator is to supply parameters according to a defined wild card approach, whenever a lack in the MM2(91) parameter set is encountered. The parameters are generated from a generic approach that considers atom hybridization, bond orders and standard covalent radii for all elements in the periodic table. This approach is to a large extent derived from the Dreiding force field published by Mayo, Olafson and Goddard [8], but is changed in order to accommodate the MM2 functional form. Note that, apart from this, Hyperchem's parameter set can easily be extended with new parameters e.g. from the literature, by editing the relevant text files.

The implemented Dreiding scheme is not limited to act as a missing parameters generator, but can also be used as a complete stand-alone force field. Even though this "second" force field is called "The Default force field" in Hyperchem's manual, it is not a part of the default operation mode of HyperChem. This can be quite misleading. To invoke "The Default force field", the user needs to manually enforce all atoms in a molecule to be of the so called unknown type. In HyperChem language, unknown types are marked with a
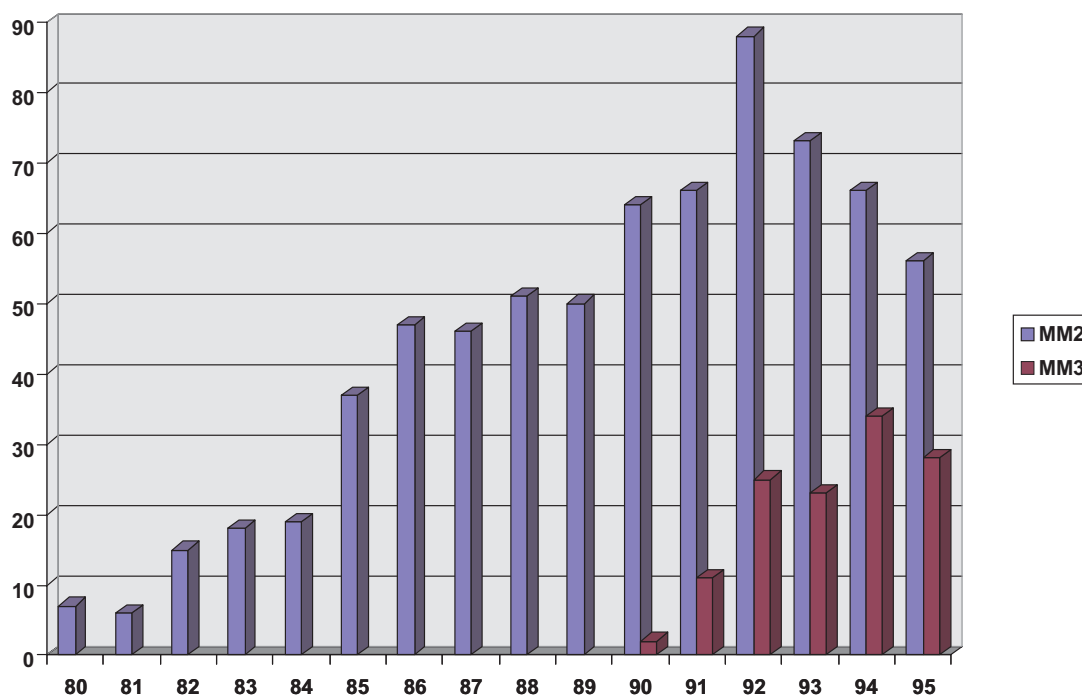

Figure 1 *Beilstein's Netfire results of bibliographic search. The number of hits is plotted for each year of search. The keywords are MM2 and MM3 in titles or abstracts. The journals covered represent grossly the area of organic chemistry.*

Note that neither the Journal of Computational Chemistry nor Theochem are covered. Also, the units squared and cubic millimeters and the crystallographic point group mm2 may interfere but only to a couple of units per year

Table 1 *Rotation barreers in kcal·mol⁻¹ of organic molecules, according to publication 13*

	Exp.	MM2*	«MM2»	MM2(91)	MMX	MM2(85)	MM+(91)	MM+(**)
Ethane	2.9	2.7	2.7	2.7	2.7	na	2.7	2.3
Propene , methyl rotation.	2.0	2.1	2.1	2.1	2.1	na	2.1	1.0
Isoprene	2.7	2.9	2.9	1.0	1.0	2.6	1.2	1.1
Ethylbenzene	1.7	1.5	1.5	1.0	1.1	1.1	1.0	2.3
2,4,6-Trimethyl-isopropylbenzene	12.8	11.0	11.0	9.5	7.2	9.7	10.4	12.4
Styrene	1.8	1.7	1.7	3.3	2.9	3.4	32.0	2.0

double asterisk (**). In this context it should be mentioned that the bond order scheme for conjugated systems is only active when the involved atoms are marked as unknown types.

For this evaluation we will test both modes of the MM+ force field. The default mode will be referred to as MM+(91), which means MM+ utilizing the MM2(91) parameter set in combination with the parameter generator turned on, for every parameter the force field does not encounter in the MM2(91) parameter set. The other mode "The default force field", with all atoms as unknown types, will be referred to as MM+(**). In some special cases we will also use MM+ in a mixed mode, which means that some particular atoms types are forced to be unknown (**). The distinction between the two force fields is not always clear in the HyperChem software presentation. Some of the tests presented here can therefore serve as a guide-line to help the user to mix the two force fields to obtain the best possible results with this package.

Material and methods

All MM+ calculations in this study were performed on IBM compatible Pentium PC's running Window95/NT and version 4.5 or 5.0 of the HyperChem software with the Chemplus add-on [9]. Note that the MM+ code in all versions before 5.0 contains a bug that causes troubles in some rare situations. The bug may be fixed by a patch provided by Hypercube [10]. All geometries have been optimized within a gradient limit of 0,1 kcal·mol⁻¹Å⁻¹ for the steric energy.
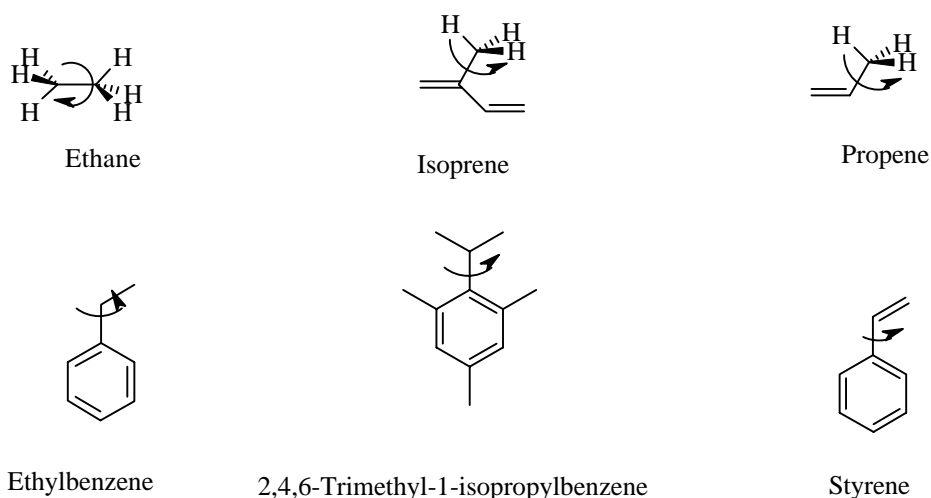
Figure 2a *Molecules involved in the Rotation barriers test (see Table 1)*

Gundertofte test : Rotation Barriers

Ethane

Isoprene

Propene

Ethylbenzene

2,4,6-Trimethyl-1-isopropylbenzene

Styrene

Table 2 *Conformational energy differences in kcal·mol⁻¹, according to publication 13. (fe) means the experimental value represents a free energy and not an enthalpy. The conformers of molecules with an * are shown in Figure 2*

	Exp.	MM2 (91)	abs. error	MM+ (91)	abs. error	MM+(**)	abs. error
Butane, g-a	*0.97*	0.9	0.07	0.87	0.10	0.70	0.27
2,3-Dimethylbutane, g-a *	*0.05*	0.15	0.15	0.19	0.14	-0.21	0.26
1,3,5-Trineopentyl- benzene, twosyn-allsyn *	*1.04 (fe)*	0.83	0.24	0.82	0.22	0.57	0.47
Methyle acetate, *E-Z* *	*8*	4.00	4	3.98	4.02	0.49	7.51
2-Butanone, skew-ecl	*2.0*	1.6	0.4	1.59	0.41	-0.50	2.50
Methoxyethane, g-a *	*1.5*	1.75	0.25	1.75	0.25	1.48	0.02
2-Methoxy tetrahydropyrane, eq-ax *	*1.0*	1.16	0.2	1.07	0.07	-0.42	1.42
Ethanol (C-O), g-a	*0.7*	0.6	0.1	0.61	0.09	0.39	0.31
Propanol (C-C, g-a	*-0.3*	0.3	0.6	0.31	0.61	0.13	0.43
Ethyl amine (C-N), g-a	*0.7*	-0.1	0.8	-0.13	0.57	0.38	0.32
N-methylacetamide, *E-Z*	*2.4 (fe)*	1.91	0.5	2.19	0.42	-0.02	2.42
N-Methylpiperidine, ax-eq	*3.2*	2.5	0.7	2.53	0.67	1.91	1.29
2-Methylpiperidine, ax-eq	*2.5*	2.1	0.4	2.12	0.38	1.75	0.75
3-Methylpiperidine, ax-eq	*1.6*	1.6	0	1.63	0.02	0.77	0.83
4-Methylpiperidine, ax-eq	*1.93*	1.7	0.23	1.74	0.19	1.17	0.76
Butadiene *	*2.5*	2.33	0.2	2.54	0.04	1.86	0.64
Acroleine *	*1.7*	1.70	0	1.71	0.01	0.52	1.18
Cyclohexane, twb-cd	*5.5*	5.4	0.1	5.36	0.14	5.75	0.25
Phenylcyclohexane, ax-eq	*2.87 (fe)*	3.62	0.73	3.40	0.53	3.92	1.05
Methylcyclohexane, ax-eq	*1.75*	1.8	0.05	1.78	0.03	1.19	0.56
Aminocyclohexane, ax-eq	*1.49*	1.4	0.09	1.41	0.08	0.38	1.11
N,N-Dimethylamino- cyclohexane, ax-eq	*1.31*	1.0	0.31	0.97	0.34	0.20	1.11
trans-1,2-Dimethyl- cyclohexane, ax,ax-eq-eq	*2.58*	2.4	0.18	2.44	0.14	1.33	1.25
cis-1,3-Dimethyl- cyclohexane, ax,ax-eq,eq	*5.5*	5.3	0.2	5.34	0.16	4.09	1.41
1,2-Difluoroethane, g-a	*-0.8*	-0.59	0.2	-0.63	0.17	0.02	0.82
Chloropropane, g-a	*-0.36*	0.2	0.56	0.24	0.60	0.48	0.84
1,2-Dichloroethane, g-a	*1.05*	1.65	0.55	1.63	0.58	0.35	0.70
1,3-Dichloropropane, g,a-g,g *	*1.1*	0.0	1.1	0.00	1.10	-0.46	1.56
1,3-Dichloropropane, a,a-g,g *	*1.5*	0.1	1.4	0.13	1.37	-0.91	2.41
Fluorocyclohexane, ax-eq	*0.16*	0.2	0.04	0.15	0.01	0.20	0.04
Chlorocyclohexane, ax-eq	*0.5*	0.4	0.1	0.42	0.08	0.78	0.28
Bromocyclohexane, ax-eq	*0.7*	0.5	0.2	0.46	0.24	0.55	0.15
Trans 1,2-Difluoro- cyclohexane, ax,ax-eq-eq *	*0.59*	0.85	0.31	0.82	0.23	0.38	0.21
Trans 1,2-Dichloro- cyclohexane, ax,ax-eq-eq *	*-0.93*	-0.86	0.03	-0.88	0.05	1.12	2.05
Trans 1,2-Dibromo- cyclohexane, ax,ax-eq-eq *	*-1.5*	-0.89	0.6	-0.91	0.59	0.91	2.41
Trans 1,4-Difluoro- cyclohexane, ax,ax-eq-eq *	*-1.14 (fe)*	-0.40	0.74	-0.43	0.71	0.39	1.53
Trans 1,4-Dichloro- cyclohexane, ax,ax-eq-eq *	*-0.8*	0.51	1.3	0.47	1.27	1.55	2.35
Trans 1,4-Dibromo- cyclohexane, ax,ax-eq-eq *	*-0.88*	0.77	1.68	0.71	1.59	1.08	1.96
Sum of abs. errors			0.50		0.47		1.17

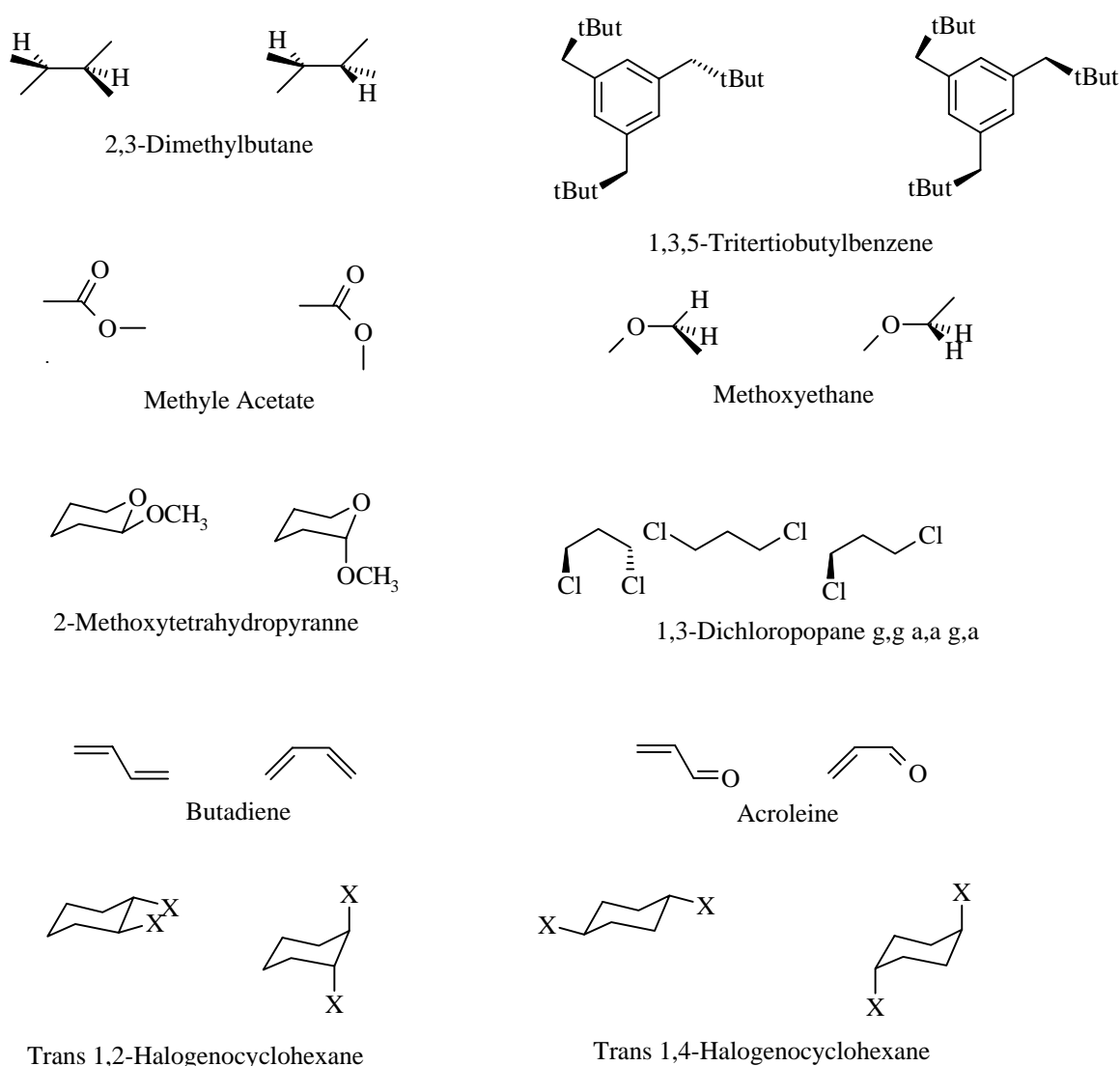Gundertofte test : Conformational differences

Figure 2b *Molecules involved in the conformational energies test (see Table 2)*

The MM3 calculations were performed by the MM3(94) Windows/DOS version from Tripos, Inc [11]. MM3(94) was used in its default mode including automatic atom typing, termination conditions, missing parameter estimation and conjugated pi-systems calculations.

The DFT BPW91/6-31G* calculation was performed by Gaussian 94W [12] utilizing default termination criteria.

The MM+ force field is evaluated with respect to three tests:

The first test is concerned with conformational energy differences and rotation barriers for some common organic compounds derived from the test bed provided by Gundertofte et al. [13,14]. The reader is referred to these papers for a detailed discussion and useful general comments on the mo-

lecular "test set". In these publications various force fields were compared and evaluated with respect to experimental energy values.

Among the involved force fields there were five variants of MM2: The MM2* version of Macromodel, the «MM2» version of Chem3D and TINKER, the "original" version MM2(91) as implemented in MacMimic, the MMX version of PC Model and a variant of MM2(85) [14]. Table 1 lists results of rotational barriers between two conformers and Table 2 list differences between two conformational energies (Figure 2). Besides the results from the MM+ calculation we have included the results from the above mentioned force fields for comparison, in Table 1. In Tables 1 and 2, the experimental values are enthalpies unless otherwise noted.

Table 3 *Structural parameters of conjugated molecules, according to publication 15. Bond lengths are given in Ångstrøms and bond angles in degrees. Bonds and angles are defined in Figure 3*

	Exp.	**MMP2**	**MM+(91)**	**MM+(**)**
Butadiene a	*1.47*	1.47	1.34	1.47
Benzene a	*1.40*	1.40	1.40	1.40
Naphtalene a	*1.41*	1.42	1.40	1.41
Naphtalene b	*1.37*	1.38	1.40	1.41
Naphtalene c	*1.42*	1.43	1.40	1.40
Naphtalene d	*1.42*	1.41	1.40	1.40
Biphenyle a	*1.49*	1.49	1.41	1.49
Biphenyle bab'	*42.*	37	14	30
trans Stilbene a	*1.34*	1.36	1.35	1.35
trans Stilbene b	*1.47*	1.47	1.35	1.48
trans Stilbene abc	*5.*	0	0	29
cis Stilbene a	*1.33*	1.35	1.35	1.35
cis Stilbene b	*1.49*	1.48	1.35	1.47
cis Stilbene abc	*43*	34	15	51
Cyclopentadiene a	*1.51*	1.51	1.51	1.51
Cyclopentadiene b	*1.34*	1.35	1.34	1.34
Cyclopentadiene c	*1.47*	1.46	1.34	1.46
Cyclopentadiene aa'	*103*	98	96	99
1,3-Cyclohexadiene a	*1.35*	1.35	1.34	1.34
1,3-Cyclohexadiene b	*1.47*	1.46	1.34	1.47
1,3-Cyclohexadiene c	*1.52*	1.51	1.51	1.51
1,3-Cyclohexadiene d	*1.53*	1.54	1.53	1.54
1,3-Cyclohexadiene aba'	*18*	15	8	14

The use of lone pairs is a very important part of the MM2(91) force field and is therefore also believed to be equally relevant to MM+(91). However, the use of lone pairs is not mentioned in the HyperChem documentation and is neither a part of the automatic atom typing rules. However, it is possible to manually attach lone pairs to atoms since lone pair parameters are provided by the MM+(91) parameter set. We have therefore chosen to use lone pairs in the "conformational energy" test for MM+(91), but not for MM+(**) since no rules are described. For MM+(91), we have strictly followed the lone pair rules recommended in the MM2(91) manual.

The second test is focused on structural parameters for a group of highly conjugated systems originally presented in two MM2 publications [15,16], involving the MMP2 procedure for aromatics. The structural parameters calculated by MM2 and MM+ are presented in table 3 for aromatics and in Table 4 for heteroaromatics. The structures of these molecules

are shown in Figure 3. Table 5 shows a comparison of MM+, MM2, crystallographic and DFT structural results for the adenine molecule.

The third test is concerned with the reproduction of experimental geometries and is based on the assumption that force field optimized structures should be similar to the crystal-structure. Even though that information from crystal structures can be used to derive force field parameters, the empirical force fields are in general designed to mimic gas phase structures rather than condensed phase structures. Crystal packing forces may have a dramatic impact on gas phase structures especially on highly flexible molecules. It should therefore be pointed out that a crystal structure comparison test is by no means an absolute test of force field performance but rather an evaluation of what makes a force field fail and in which conditions.

To allow for some kind of comparison with other force fields we have chosen to use the same set of structures from

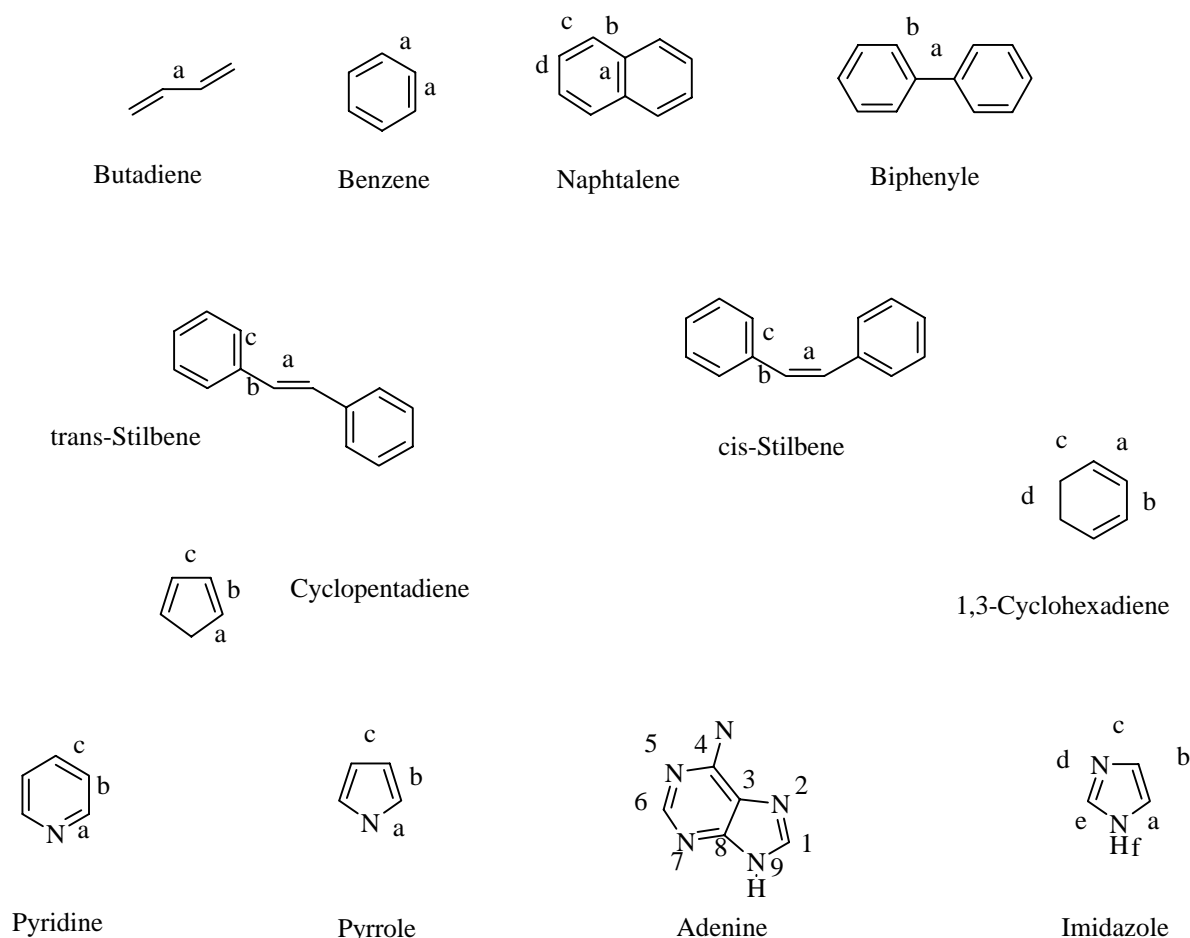Conjugation test : Definitions of distances, angles and torsions

Figure 3 *Definitions of structural parameters of molecules involved in the conjugation test (see Table 3 and 4)*

the Cambridge Structural Database (CSD) which has been use in similar studies previously [8, 17 – 19] (Figure 5). As noted in some of these works there are numerous pitfalls to consider when applying this test. The test is usually performed by relaxing the imported CSD structure by the force field in question until a termination criteria is reached (0,1 kcal·mol⁻¹Å⁻¹ for the energy gradient in our study). The RMS fit derivation for all heavy atoms (all atoms except hydrogens) in the optimized model and the CSD structure are then used as a measure for the impact of the force field (Table 6). In the following this will be referred to as the "atomic coordinates" test. The result from such a test is not only dependent on the force field functional form and its parameterization, but also on the selected geometry optimization algorithm. A thoughtful and detailed presentation of these pitfalls is given by Halgren [19]. A comparison test based on the "atomic coordinates" test alone is hampered by the fact that even minor changes in a single torsion angle may lead to a large RMS

deviation. One way to deal with this difficulty is to include the RMS for bonds, angles and dihedrals in the evaluation. Such a comprehensive comparison is not easily preformed by standard programs like Hyperchem, since they normally only provide a RMS fitting algorithm for atomic coordinates. In addition to the "atomic coordinates" test (Table 7) we have therefore selected an eight molecules subset (Figure 6) for closer inspection including a bond length RMS test (Table 8).

Another problem this test poses is the question about how to deduce the bond order information from the imported CSD files, which is essential for proper force field atom typing. We have chosen to use the same bond definitions as given in reference 8 and 17 with the only general exception of using HyperChem's aromatic bond definitions for aromatic ring systems. It should pointed out that in some cases, results could be dramatically improved with proper bond orders assigned according to a visual evaluation of C-C distances in the X-ray structure. Except for the removal of counter-ions and coordi-

Table 4 *Structural parameters of heteroaromatics, according to publication 16. Bond lengths are given in Ångstrøms and bond angles in degrees. Bonds, angles and types 1 to 4 are defined in Figure 4*

	Exp.	**MM2**	**MM+ type 1**	**MM+ type 2**	**MM+ type 3**	**MM+ type 4**
Pyridine a	*1.34*	1.34	1.26	1.35	1.27	1.32
Pyridine b	*1.39*	1.40	1.40	1.40	1.40	1.40
Pyridine c	*1.39*	1.40	1.40	1.40	1.40	1.40
Pyridine aa'	*117*	117	120	121	119	122
Pyrrole a	*1.37*	1.37	1.27	1.41	1.27	1.32
Pyrrole b	*1.38*	1.38	1.34	1.34	1.40	1.40
Pyrrole c	*1.42*	1.41	1.34	1.47	1.40	1.40
Pyrrole aa'	*110*	110	112	108	113	111
Pyrrole ab	*108*	108	108	108	109	109
Pyrrole bc	*107*	108	106	108	105	106
Imidazole a	*1.38*	1.38	1.32	1.40	1.32	1.34
Imidazole b	*1.37*	1.38	1.34	1.34	1.40	1.39
Imidazole c	*1.38*	1.38	1.39	1.42	1.36	1.35
Imidazole d	*1.33*	1.32	1.32	1.29	1.36	1.35
Imidazole e	*1.36*	1.35	1.32	1.41	1.32	1.34
Imidazole f	*1.05*	1.05	1.01	0.99	0.99	0.99

nating water molecules we have used the structures exactly as they were extracted from the CSD.

Results and discussions

Conformational energies

The data for the conformational energies are presented in Tables 1 and 2. Table 1 shows the rotational barriers for six bond rotations. The results for ethane, propene, isoprene and ethylbenzene are very similar for both MM2(91) and MM+(91), whereas the MM+(**) results differs slightly. For 2,4,6-trimethyl isopropyl benzene, the best agreement with experiment are found by MM+(**), but both force fields provide reasonably high barriers. In the case of styrene, MM+(91) predicts a barrier that is about 30 kcal·mol^{-1} too high. This fatal failure can completely be ascribed to the assignment of atom types in combination with the omission of a correction term for conjugated pi systems.

The results of the 39 molecules "Gundertofte test set" are combined in Table 2, which also includes the MM2(91) results given in reference 14 for comparison. The MM2(91) and the MM+(91) results are virtually identical (within 0.1

kcal·mol^{-1}) for 35 of the 39 compounds. The largest discrepancies between these two force fields are found for N-methylacetamide (0.3 kcal·mol^{-1}), butadiene (0.2 kcal·mol^{-1}), phenylcyclohexane (0.2 kcal·mol^{-1}) and 2-methoxy-tetrahydropyrane (0.1 kcal·mol^{-1}). The omission of the MM2(91) conjugation and anomeric correction terms in MM+(91) can explain these differences (except for N-methylacetamide). It is noteworthy how small these differences are and that the deviation in all four cases is in favor of the less sophisticated MM+(91) with respect to the experimental result.

The "default" MM+(**) force field performs significantly differently from MM+(91) and MM2(91). The average sum of absolute errors from the experimental value is 0.47 kcal·mol^{-1} for MM+(91), 0.50 kcal/mol for MM2(91) and 1.17 kcal/mol for MM+(**). It is not possible to point out a particular group of compounds, which has the main responsibility for the poor performance of MM+(**). Propanol, ethyl methyl ether, ethyl amine and methoxyethane are the only cases where MM+(**) shows a better result than MM+(91).

More seriously, MM+(**) predicts the incorrect "most stable" conformer in 13 out of the 39 molecules, whereas the number is 5 for MM+(91) and MM2(91). Ethylamine is the only case where MM+(**) gives the right conformer com-
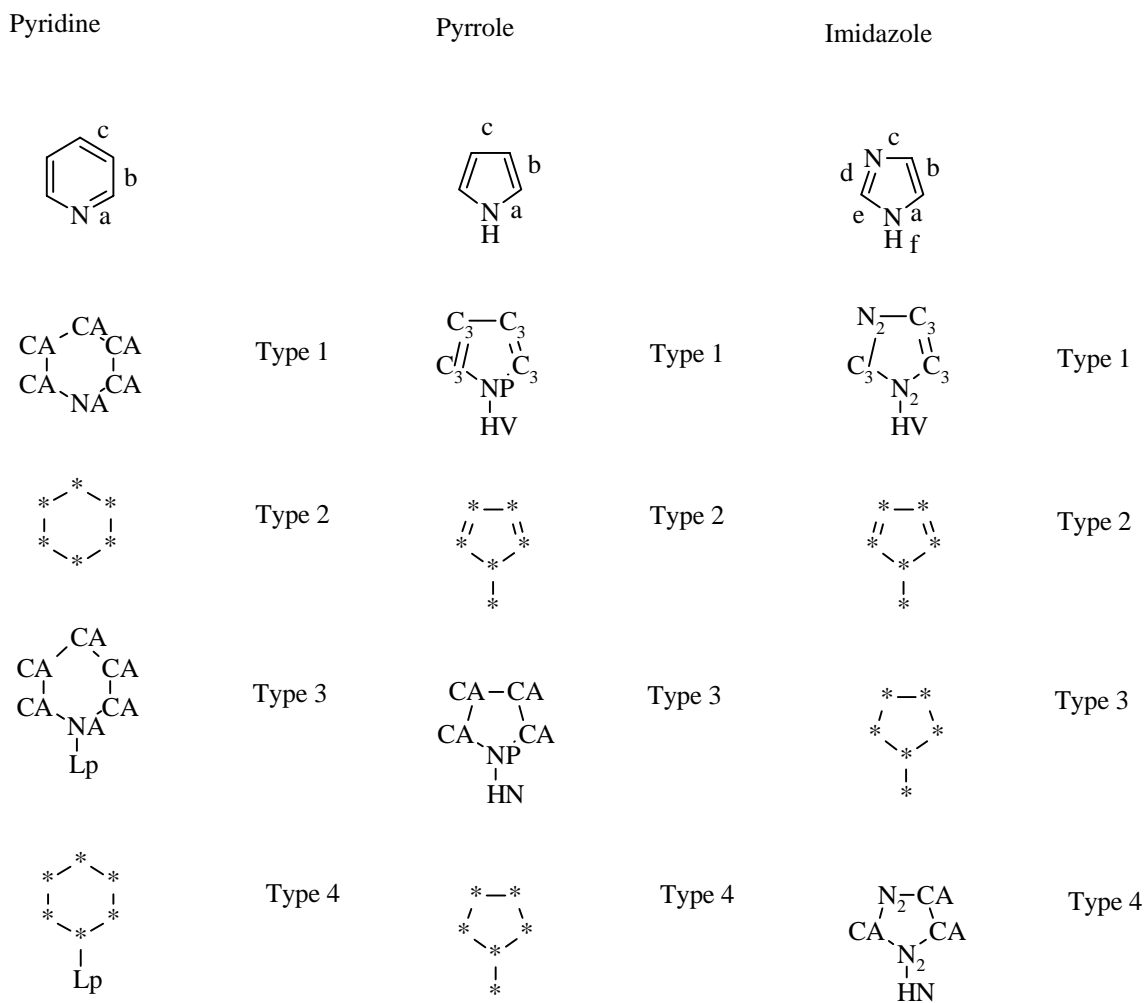
Pyridine Pyrrole Imidazole

Figure 4 *MM+ atom types probed in heteroaromatics conjugation test (see Table 4)*

pared to MM+(91), but the absolute energy difference is small (0.4 kcal·mol⁻¹).

Conjugation test

As mentioned in the introduction there is no VESCF method implemented in MM+ to account for conjugated molecules. The pure MM+(91) force field will therefore not be able to provide any advanced treatment of conjugated systems. This deficiency can clearly be seen in the examples from Table 3 where all C-C bond length in molecules like butadiene and stilbene are calculated to be 1.34 Å. For cases like naphthalene and benzene it is possible to use an aromatic atom typing (CA), which will give a more correct value of 1.34 Å for these bond types. However, this is a very static approach that only can provide an average aromatic bond length, which is insufficient for describing the more complicated cases like stilbene and biphenyl.

The "default" MM+(**) force field uses the bond order scheme mentioned in the introduction to obtain a better treat-

ment of strongly conjugated systems. This scheme seems to work quite successfully for all the compounds in Table 3. Butadiene like systems with alternating double bonds are also treated more correctly by MM+(**) than by MM+(91).

For hetoroaromatic systems the situation is less clear. If we consider the heteroaromatics in Table 4, the number of possibilities of mixed atom types / default types increases for this kind of molecules. The choice is not only between MM+(91) and MM+(**) but within a greater number of combinations. Four combinations were tried for each molecule (Figure 4). Pyridine performs well when all default is used without lone pair (type 2 in Figure 4), solution which corresponds to the original MM+(**). But Pyrrole and even more Imidazole structures are too complicated (half way between pyridine-like delocalization and cyclopentadiene-like localization) for any combination to succeed in describing bond lengths.

Many of the structures in the following CSD test have an adenine group or an adenine-like moiety attached. We have therefore selected adenine as a secondary test example of conjugated aromatic heterocycles. Table 5 shows the bond

Table 5 *«Bond lengths» and «Atomic coordinates» RMS in Ångstrøms of different structures of the Adenine molecule, according to publications 20 (X-ray structure) and 21(MM2(91)np structure). Adenine bond lengths are defined in Figure 3*

Bond	X-ray	MM+(91)	MM+(**)	MM2(91)np	MM3(94)	BPW91
1 2	1.316	1.348	1.298	1.311	1.316	1.323
2-3	1.395	1.350	1.414	1.383	1.405	1.390
3-4	1.411	1.341	1.341	1.432	1.415	1.419
4-5	1.347	1.266	1.423	1.352	1.341	1.355
5-6	1.343	1.267	1.296	1.366	1.368	1.350
6-7	1.324	1.267	1.425	1.328	1.333	1.347
7-8	1.350	1.261	1.284	1.377	1.359	1.346
8-9	1.371	1.333	1.398	1.371	1.371	1.386
1-9	1.358	1.337	1.420	1.369	1.378	1.385
3-8	1.381	1.338	1.451	1.397	1.396	1.409
4-10	1.334	1.340	1.409	1.341	1.373	1.358
Average error (absolute)		0.051	0.057	0.012	0.013	0.014
Average error		-0.044	0.021	0.009	0.011	0.013
Max error		0.089	0.101	0.027	0.039	0.028
RMS (bond lengths)		0.057	0.063	0.015	0.017	0.017
RMS (atomic coordinates)		0.086	0.067	NA	0.066	0.026

length details calculated by various methods and a crystallographic reference [20]. As a secondary reference structure to the x-ray structure, we have included the result from a DFT BPW91(6-31G*) electron structure model. We should note however that experimental (for X-ray diffraction) or theoretical (for quantum mechanics) are not strictly equivalent and that a little difference between a bond distance measured in the crystalline state and calculated in vacuo is normal. Attempts to quantify this difference do exist [7], and the order of magnitude involved should not alter the interpretation of our results. In this case, the DFT electronic structure provides the same RMS error order for the force fields as the X-ray reference. Adenine could not be studied by standard MM2(91) due to the lack of genuine parameters. Instead we have included the result from a previous recent study by Berg and Bladh [21], in which they have developed a new set of parameters for conjugated sp² nitrogen.

The results from this test shows that both MM3(94) and the special MM2(91)np results fall within an acceptable range from the given reference. The two MM+ force fields shows considerably larger errors in all parameters. Bond lengths calculated by MM+(91) are all too short, except for bond 1-2, whereas the other methods are more random in their deviation. It is predictable that MM+(91) will fail due to the

lack of correction terms for conjugations. The attempts to account for conjugation by the bond order considerations in MM+(**) is not convincing in this case and shows the maximum error.

Crystal structures comparison

The CSD test structures were all extracted as the first 76 structures in the Cambridge Structure Database having *R* factors below 0.05 (sorted by alphabetic record). Of the original set of 76 organic molecules extracted from the Cambridge Structural Database, 72 have been retained (Figure 5). Four were left out for non-concordant information from the Cambridge Structural Database and publication 8. The results from the RMS fit comparison for the two MM+ force fields as well as the previous results for Dreiding, CHARMm, Tripos and MMFF94 are listed in Table 6.

The averaged performance of the two MM+ force fields is in this test not notably different from the previous tested. This result could indicate that the two MM+ force fields are just as good as the others, but it could also indicate that a crystal structure comparison test, in average, is more or less insensitive towards the underlying force field. To judge from

AAXTHP ABAXES ABBUMO10 ABINOR02

ABINOS01 ABTOET ABZTCX ACADOS

ACAFLR ACANIL01 ACARAP ACBNZA01

ACBUOL ACCITR10 ACDXUR

ACENAP03 ACFPCH ACFUCN
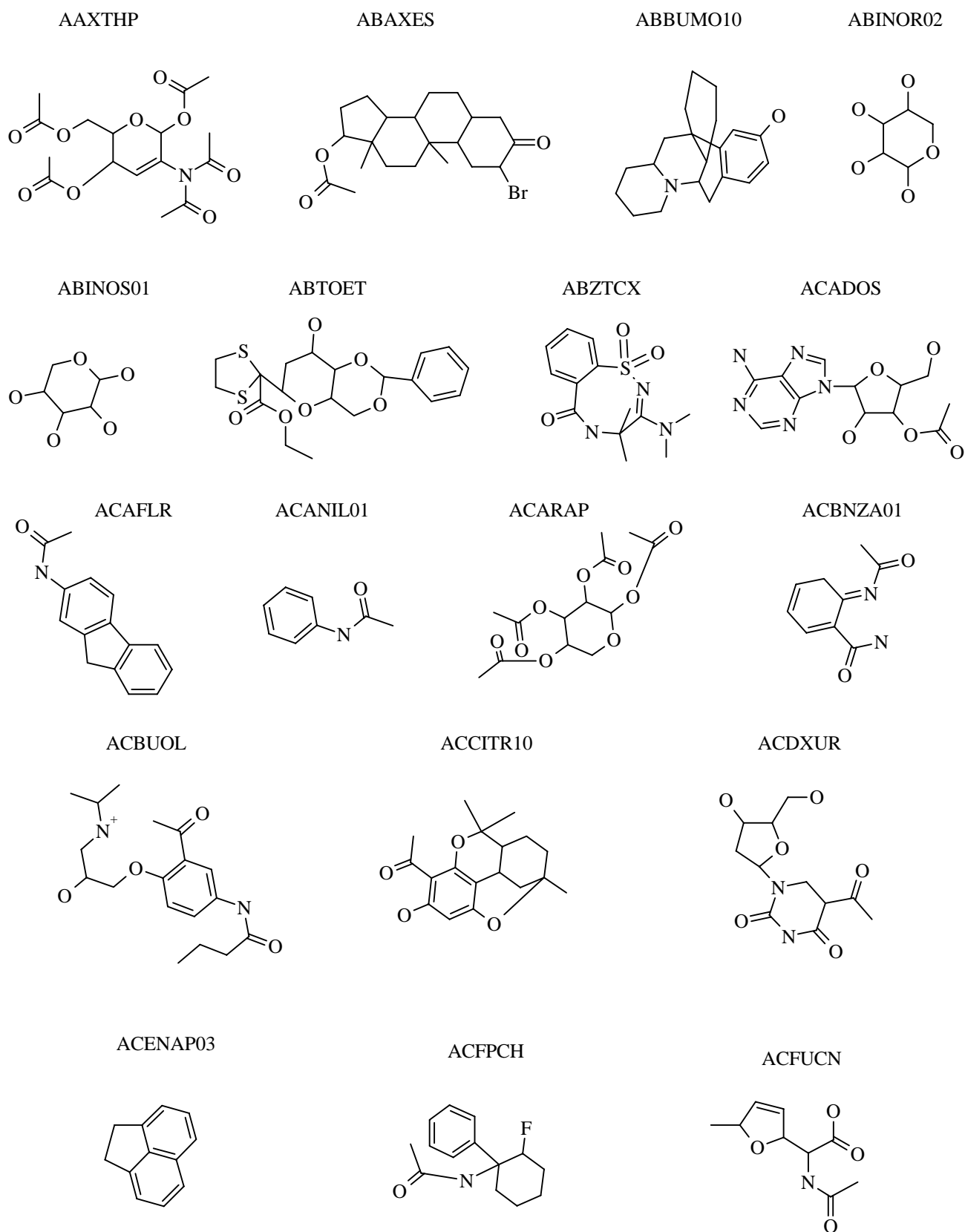
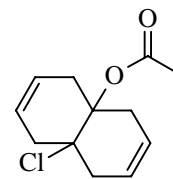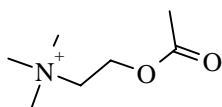**Figure 5** *72 molecules involved in the CSD test (see Table 6) (continued next pages)*
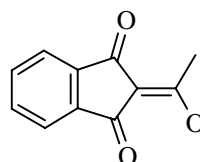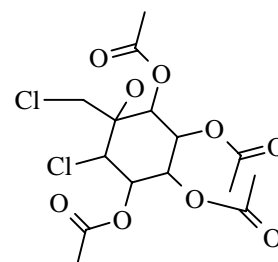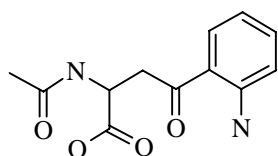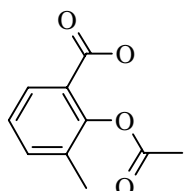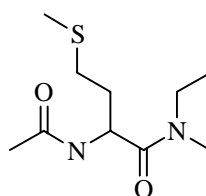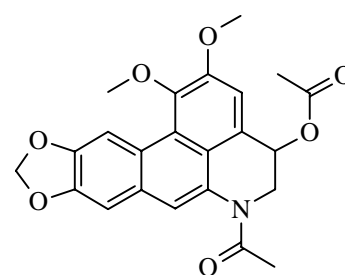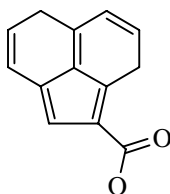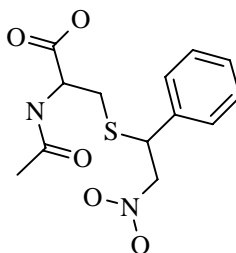
ACGLSP
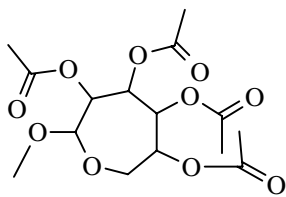
ACGLUA11
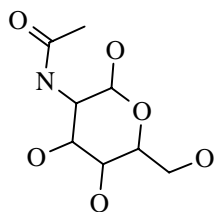
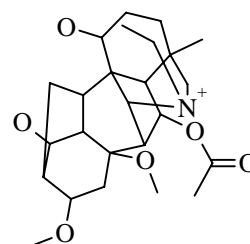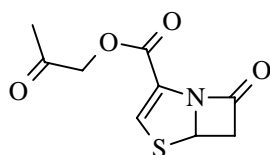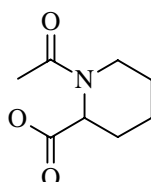ACHGAL
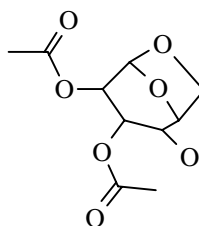
ACHNAP10

ACHTAR10

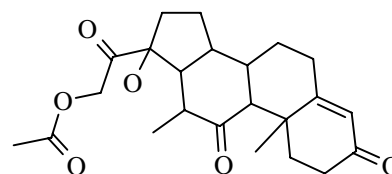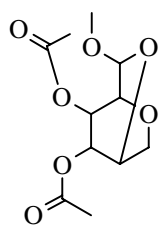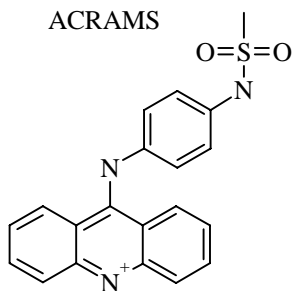ACIMDC

ACINDN

ACINST

ACYNU

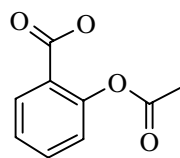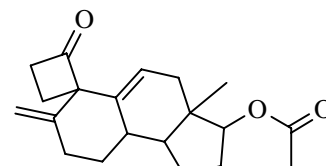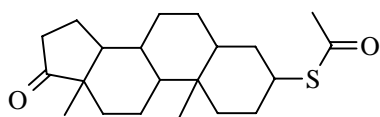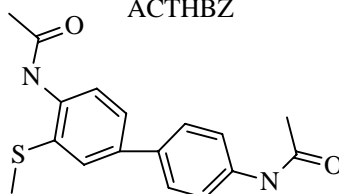ACMEBZ
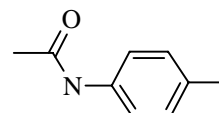
ACMTDE
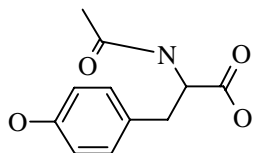
ACNORT

ACNPAC10
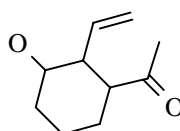
ACNPEC

ACONTN10

ACPENC10

ACPPCA
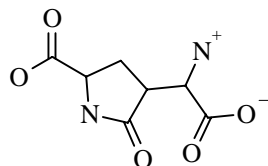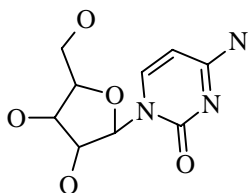
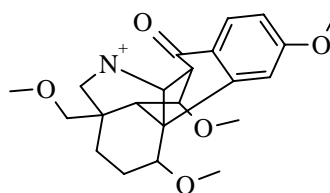ACPRET03

ACPYNS

ACRAMS

ACSALA01
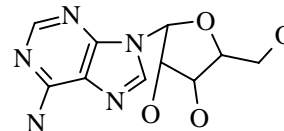
ACSESO10
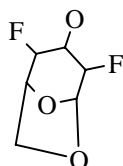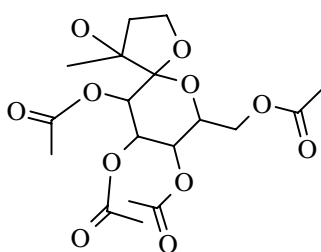
ACTAND

ACTHBZ

ACTOLD

ACTYSN

ACURID
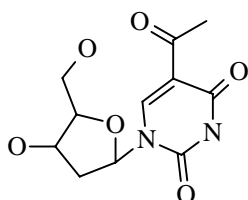
ACVCHO

ACXMOL
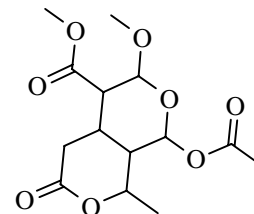
ACXMPR

ACYTID

ADELOX10

ADENOS10

ADFGLP

ADGSMH

ADHELA10

ADMANN

ADMHEP

ADMINA

ADMOPM

ADRTAR

ADYPNL

AEBDOD10

AENLAN10

AFCYDP

AFMSCY

AFURPO10

AFUTDZ10

AFUTHU

AGALAM10

AGLUAM10

AHARFU

AHDCLA

AHDITX
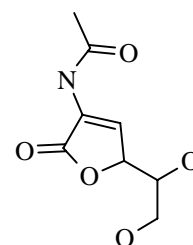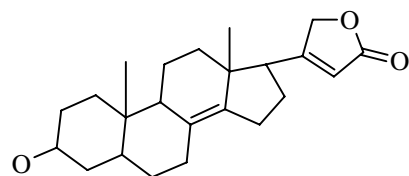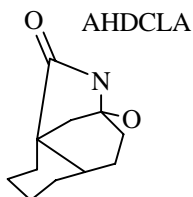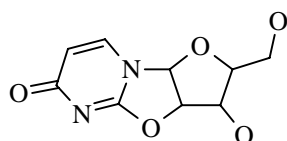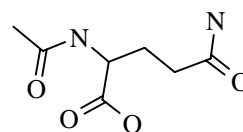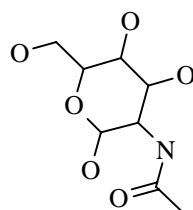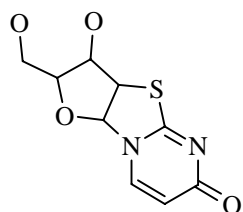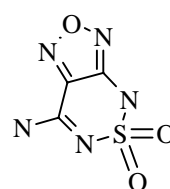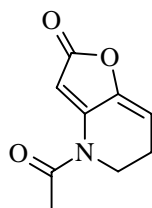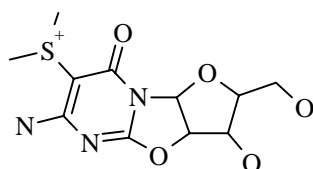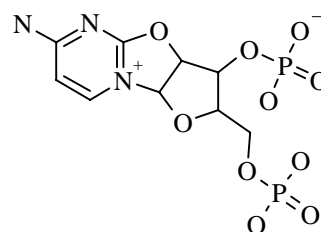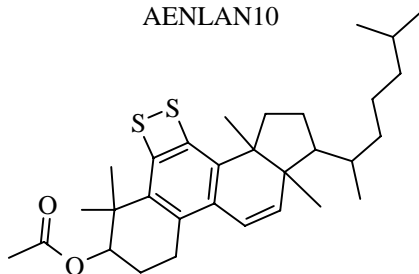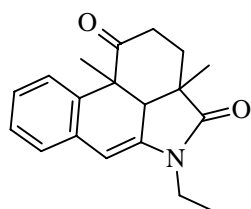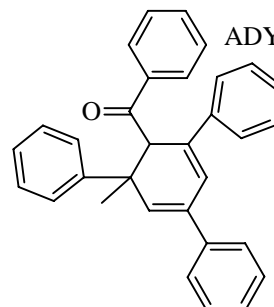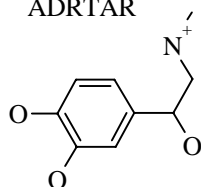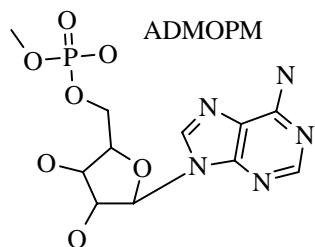
Table 6a *"Atomic coordinates" RMS in Ångstrøms of 72 molecules from the CSD, according to publications 8 (Dreiding results), 17 (Tripos results), 18 (CHARMm results) and 19 (MMF94 results).*

Molecule	Atoms	MM+(91)	MM+(**)	Dreiding	MMFF94	CHARMm	Tripos 5.2
AAXTHP	26	0.390	0.297	0.334	1.317	0.401	0.500
ABAXES	25	0.128	0.137	0.112	0.156	0.312	0.123
ABBUMO10	22	0.152	0.068	0.113	0.053	0.089	0.124
ABINOR02	10	0.092	0.078	0.090	0.062	0.106	0.125
ABINOS01	10	0.066	0.058	0.075	0.070	na	0.036
ABTOET	27	0.194	0.419	0.414	0.400	0.568	0.351
ABZTCX	20	0.557	0.465	0.351	0.766	0.268	0.480
ACADOS	22	0.197	0.440	0.148	0.160	0.262	0.264
ACAFLR	17	0.287	0.086	0.147	0.305	0.310	0.144
ACANIL01	10	0.166	0.290	0.070	0.127	0.133	0.250
ACARAP	22	0.551	0.467	0.340	0.561	0.611	0.463
ACBNZA01	13	0.312	0.097	0.146	0.194	0.307	0.155
ACBUOL	24	0.429	0.823	0.449	1.173	1.450	0.422
ACCITR10	22	0.094	0.358	0.265	0.143	0.305	0.091
ACDXUR	19	0.130	0.194	0.141	0.291	0.344	0.246
ACENAP03	12	0.021	0.017	0.023	0.027	0.027	0.020
ACFPCH	17	0.135	0.144	0.073	0.475	0.821	0.277
ACFUCN	14	0.204	0.394	0.275	0.163	0.330	0.425
ACGLSP	25	0.503	0.374	0.341	0.475	0.573	0.448
ACGLUA11	15	0.100	0.245	0.190	0.221	0.216	0.092
ACHGAL	17	0.151	0.150	0.180	0.236	0.240	0.297
ACHIST20	14	na	na	0.336	1.205	1.227	0.117
ACHNAP10	15	0.035	0.067	0.078	0.060	0.069	0.055
ACHTAR10	10	0.202	0.245	0.148	0.197	0.229	0.139
ACIMDC	4	0.031	0.051	0.023	0.018	0.028	0.019
ACINDN	14	0.114	0.135	0.072	0.050	0.067	0.072
ACINST	26	0.392	0.245	0.244	0.352	0.349	0.548
ACKYNU	18	0.254	0.683	0.391	0.487	1.138	0.420
ACMBPN	24	na	na	0.259	0.230	0.448	0.151
ACMEBZ	14	0.155	0.181	0.175	0.217	0.354	0.175
ACMTDE	16	0.158	0.325	0.254	0.158	0.564	0.175
ACNORT	31	0.189	0.237	0.266	0.195	0.295	0.215
ACNPAC10	15	0.092	0.048	0.050	0.057	0.061	0.047

Table 6b *"Atomic coordinates" RMS in Ångstrøms of 72 molecules from the CSD, according to publications 8 (Dreiding results), 17 (Tripos results), 18 (CHARMm results) and 19 (MMF94 results)*
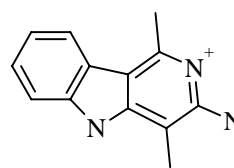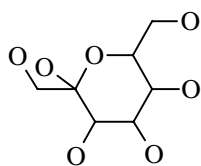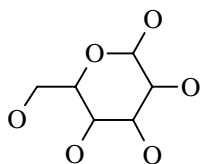
Molecule	Atoms	MM+(91)	MM+(**)	Dreiding	MMFF94	CHARMm	Tripos 5.2
ACNPEC	21	0.928	0.792	0.844	0.832	1.178	0.931
ACONTN10	32	0.177	0.134	0.186	0.175	na	0.270
ACPENC10	15	0.322	0.244	0.288	0.449	1.192	0.511
ACPPCA	12	0.072	0.087	0.225	0.178	0.219	0.140
ACPRET03	29	0.359	0.200	0.290	0.292	0.379	0.202
ACPYNS	18	0.215	0.159	0.263	0.370	0.429	0.280
ACRAMS	26	1.194	0.356	0.867	1.071	na	0.515
ACSALA01	13	0.473	0.461	0.097	0.075	0.312	0.075
ACSESO10	23	0.233	0.201	0.211	0.171	0.294	0.208
ACTAND	2	0.214	0.234	0.379	0.174	0.317	0.257
ACTHBZ	22	0.496	0.418	0.297	0.560	0.439	0.466
ACTHCP	10	na	na	0.093	0.085	0.117	0.093
ACTOLD	11	0.166	0.262	0.045	0.140	0.146	0.231
ACTYSN	16	0.728	0.902	0.774	0.381	0.355	0.448
ACURID	19	0.161	0.315	0.126	0.414	0.595	0.268
ACVCHO	12	0.052	0.104	0.109	0.048	0.050	0.071
ACXMOL	22	0.224	0.323	0.408	0.532	0.800	0.412
ACXMPR	14	0.207	0.473	0.671	0.972	0.247	0.203
ACYGLY11	8	na	na	0.095	0.071	na	0.048
ACYTID	17	0.087	0.119	0.202	0.171	na	0.120
ADELOX10	28	0.142	0.114	0.127	0.168	na	0.165
ADENOS10	19	0.074	0.686	0.088	0.201	0.637	0.109
ADFGLP	11	0.033	0.075	0.059	0.679	na	0.163
ADGSMH	29	0.482	0.338	0.340	0.440	na	0.382
ADHELA10	14	0.080	0.178	0.092	0.098	na	0.107
ADMANN	12	0.149	0.111	0.122	0.170	na	0.103
ADMHEP	14	0.057	0.095	0.088	0.083	na	0.113
ADMINA	16	0.083	0.073	0.061	0.054	na	0.066
ADMOPM	24	0.856	0.712	0.677	0.989	na	1.088
ADRTAR	13	0.156	0.786	0.139	0.747	na	0.094
ADYPNL	33	0.447	0.289	0.354	0.499	na	0.314
AEBDOD10	22	0.162	0.164	0.256	0.173	na	0.076
AENLAN10	35	0.535	0.493	0.420	0.531	na	0.466
AFCYDP	24	0.069	0.460	0.413	0.319	na	0.527
AFMSCY	20	0.346	0.224	0.306	na	na	0.209
AFURPO10	13	0.222	0.303	0.140	0.095	na	0.094
AFUTDZ10	12	0.229	0.233	0.247	0.052	na	0.209
AFUTHU	16	0.124	0.094	0.067	0.353	na	0.210
AGALAM10	15	0.414	0.250	0.109	0.216	na	0.071
AGLUAM10	13	0.349	0.634	0.328	0.334	na	0.297
AHARFU	16	0.144	0.198	0.120	0.215	na	0.375
AHCDLA	14	0.048	0.076	0.075	0.049	na	0.064
AHDITX	26	0.073	0.100	0.185	0.146	na	0.153
Average		0.251	0.282	0.235	0.325	0.412	0.246
Max		1.194	0.902	0.867	1.317	1.450	1.088

Figure 6 *molecules involved in the bond lengths RMS test (see Table 7 and 8)*

ACAFLR

ACANIL01

ACENAP03

ACINDN

ACNPAC10

AFURPO10

AFUTHU

AHARFU
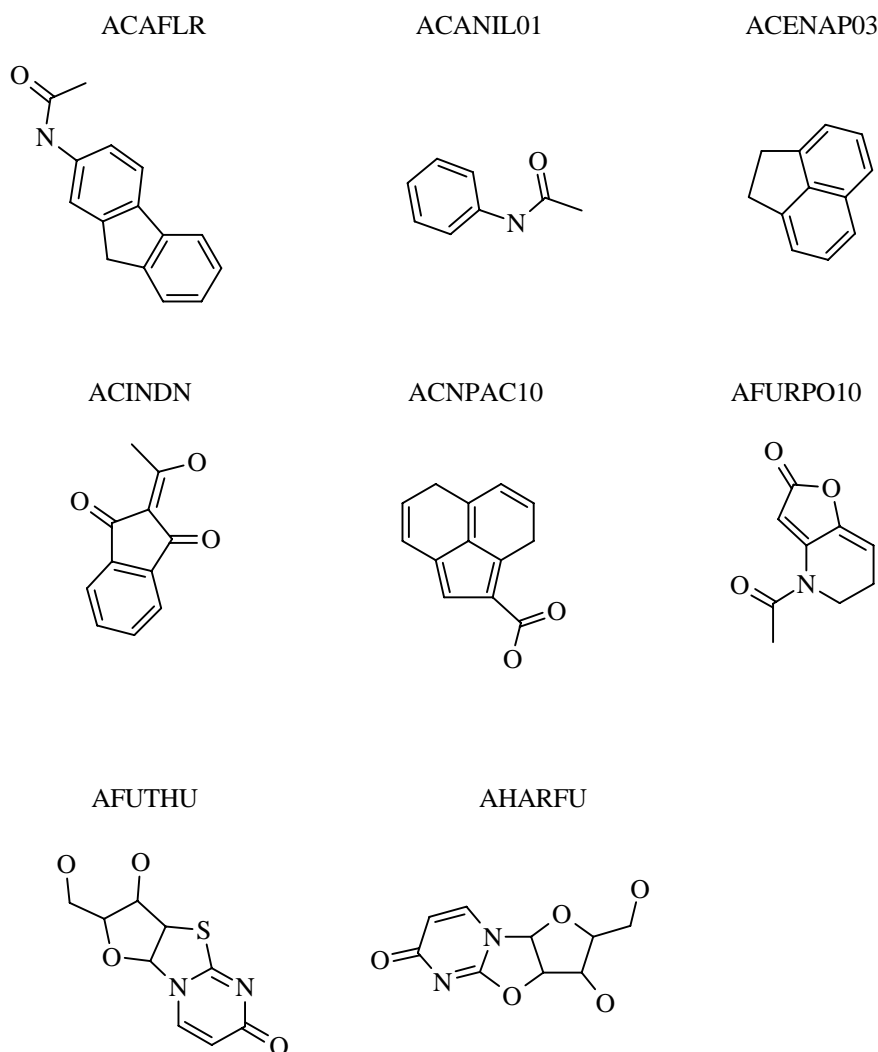
the discussions in the previous studies and our own results, the later seems most likely to be the case.

The two MM+ force fields are tested under exactly the same conditions, just utilizing different parameter sets. Therefore it should be possible to make a more direct comparison between these. In average MM+(91) performs slightly better than MM+(**), which could indicate that the more dedicated MM2(91) parameter set does a better job than the universal MM+(**) parameter generator. It should be noted that the MM+(91) results not always are based on the 100% pure MM2(91) parameter set. Only 25 of the 72 molecules could be handled with MM2(91) parameters alone. The rest of the molecules hold various amounts of generated parameters.

The average RMS deviation for the 25 "pure MM2(91)" results is 0.192, whereas the value raises to 0.290 for the remaining 47 molecules. This might indicate that the pure MM+(91) method provides the best quality, but when we look at the corresponding MM+(**) results (0.203 and 0.327), we find approximately the same ratio. The explanation is therefore more likely that the MM2(91) pure subset in general

involves less difficult cases with respect to size, composition or flexibility, compared to the rest.

A subset of eight structures (Figure 6) was selected for a more complete test including the RMS for all predicted bond lengths (table 8). Bond length is regarded as the parameter that is less sensitive to crystal packing effects. The selection is based on structures possessing a high degree of conjugations, heteroatoms and low sensitivity towards flexibility in side chains. The most striking result from this subset is the total lack of correlation between "atomic coordinates" (Table 7) and "bond lengths" (Table 8) RMS errors. MM3 and MMFF94, the most recent force fields, give the best result in the "bond lengths" test. In the atomic coordinates test MM3 performs equally well, whereas MMFF94 fails completely. For Dreiding the situation is opposite. The two MM+ force fields results are almost identical for the "atomic coordinates" test, but the difficulties for MM+(91) in dealing with conjugated systems are clearly demonstrated by the "bond length" test. These findings indicate that a "bond length" test is more sensitive and seems to provide a better test for the quality of a force field than the widely used "atomic coordinates" test.

Table 7 *«Atomic coordinates» RMS in Ånstrøms of 8 molecules from the CSD (Figure 6), according to publications 8 (Dreiding results), and 19 (MMF94 results)*

Molecule	Atoms	Dreiding	MM+(**)	MM+(91)	MM3(94)	MMFF94
ACAFLR	17	0.147	0.086	0.287	0.056	0.305
ACANIL01	10	0.070	0.290	0.166	0.259	0.127
ACENAP03	12	0.023	0.017	0.021	0.016	0.027
ACINDN	14	0.072	0.135	0.114	0.062	0.050
ACNPAC10	15	0.050	0.048	0.092	0.054	0.057
AFURPO10	13	0.140	0.303	0.222	0.083	0.095
AFUTHU	16	0.067	0.094	0.124	0.103	0.353
AHARFU	16	0.120	0.198	0.144	0.073	0.215
Total		0.689	1.171	1.169	0.706	1.229
Average		0.086	0.146	0.146	0.088	0.154
Max		0.147	0.303	0.287	0.259	0.353

It should be noted that differences in optimization algorithm and termination criteria used by the various methods may affect the optimized model slightly. Thus, only the results for MM+(91) and MM+(**) can be rigorously compared.

Conclusion

The object of this study was to evaluate the MM+ force field with particular attention towards its deviation from the parent MM2 force field. If we select the conformational energies from the "Gundertofte test" as a starting point for this discussion, it is possible to make a ranking of the two MM+ force fields between the 15 force fields tested in this paper [13]. In this context, MM+(91) performs slightly better than the original MM2(91) and will be placed in top five among MMFF(93), MM3(92) and the other "MM2 like" force fields. The poor performance of the MM+(**) force field in this test will place it in the bottom of the table among force fields like Tripos 5.2, Dreiding 2.21 and UFF 1.01. The default mode of MM+ operation in the HyperChem package is MM+(91) in combination with MM+(**) for every unknown parameter. This means that MM+ in practical terms will provide reasonable accurate conformational energies for the vast majority of organic compounds, even for the inexperienced user.

If the question is about how well MM+ reproduces structural parameters the answer is not equally promising. The huge problem here is naturally the question of conjugated systems. It is a curiosity that MM+(91) correctly reproduce the energy difference between s-cis and s-trans butadiene,

even though that the structures are wrong (all three carbon-carbon bonds are calculated to be 1.34 Å).

The bond order scheme used by the MM+(**) method treats both the butadiene case and some of the other conjugated system correctly, but the adenine case and some of the heteroaromatics are obviously too complicated. In general, MM+(91) will always fail with conjugated system and reasonable aromatic structures can only be obtained if aromatic (CA) atoms types are properly defined. A correct atom typing is however not always a guarantee for the best answer. In the case of the far too high (30 kcal) rotational barrier for styrene the crucial bond has the correct definition (CA-CA-C3-C3). However, if the two MM+ force fields are mixed by assigning the torsion to be CA–**–**–C3, then the barrier will drop to the more reasonable value of 2 kcal.

The good performance of MM+(91) on conformational energies is clearly a heritage from the original MM2 force field. Whereas the difficulties towards conjugated systems likewise is caused by the omission of the MM2 SCF correction term. The relatively poor performance of MM+(**) on conformational energies, means that it should be avoided as a stand-alone force field. On the other hand, the quality of MM+(**) is clearly acceptable for providing parameters to MM+(91), where they are needed for practical purposes. As demonstrated above, MM+(**) is also needed when working with conjugated systems, but the results are not always satisfactory, especially not for heteroaromatics. The combinatorial use of the two MM+ force fields demands careful intervention by the user and is unfortunately not a part of the default operation mode of HyperChem.

Table 8 *«Bond lengths» RMS in Ångstrøms of 8 molecules from the CSD (Figure 6), according to publications 8 (Dreiding results), and 19 (MMF94 results)*

Molecule	Bonds	Dreiding	MM+(**)	MM+(91)	MM3	MMFF94
ACAFLR	19	0.033	0.023	0.028	0.017	0.024
ACANIL01	10	0.026	0.024	0.026	0.013	0.011
ACENAP03	14	0.026	0.020	0.018	0.013	0.024
ACINDN	15	0.049	0.023	0.064	0.020	0.021
ACNPAC10	17	0.040	0.025	0.053	0.023	0.025
AFURPO10	14	0.040	0.023	0.042	0.019	0.016
AFUTHU	18	0.032	0.034	0.037	0.027	0.020
AHARFU	18	0.036	0.024	0.037	0.016	0.020
Total		0.282	0.195	0.306	0.148	0.161
Average		0.035	0.024	0.038	0.018	0.020
Max		0.049	0.034	0.064	0.027	0.025

References

1. Allinger, N. *J. Am. Chem. Soc.* **1977,** *99*, 8127-8134.
2. Lipkowitz, K.; Osawa, K. In *Reviews in Computational Chemistry;* Boyd, D., Eds; Wiley: New York, 1994, Vol. 6.
3. Hay, B. *Coor. Chem. Rev.* **1993,** *126*, 177-236.
4. Strassner, T. *J. Mol. Model.* **1996,** *2*, 217-226.
5. Holloway, K.; Wai, J.; Halgren, T.; Fitzgerald, P.; Vacca, J.; Dorsey, B.; Levin, R.; Thompson, W.; Chen, J.; deSolms, J.; Gaffin, N.; Ghosh, A.; Giuliani, E.; Graham, S.; Guare, J.; Hungate, R.; Lyle, T.; Sanders, W.; Tucker, T.; Wiggins, M.; Wiscount, C.; Woltersdorf, O.; Young, S.; Darke, P.; Zugay, J. *J. Med. Chem.* **1995,** *38*, 305-317.
6. *Computational chemistry*, HyperChem Manual, Hypercube Inc., Waterloo, Ontario, Canada: 1994.
7. Allinger, N.; Zhou, X.; Bergsma, J. *J. Mol. Struct. (Theochem)* **1994,** *312*, 69-83.
8. Mayo, S.; Olafon, B.; Goddard III, W. *J. Phys. Chem.* **1990,** *94*, 8897-8909.
9. Hyperchem available from Hypercube Inc., Gainesville, Florida, USA.
10. Polowin, J. *Hypernews* **1996,** *1*, Oct 96, 4.
11. MM3 (94), Tripos Inc., St Louis, MI, USA, 1995.
12. Gaussian94, Revision E.1, Frisch, M.J.; Trucks, G.W.; Schlegel, H.B.; Gill, P.M.W.; Johnson, B.G.; Robb, M.A.; Cheeseman, J.R.; Keith, T.; Petersson, G.A.; Montgomery, J.A.; Raghavachari, K.; Al-Laham, M.A.; Zakrzewski, V.G.; Ortiz, J.V.; Foresman, J.B.; Cioslowski, J.; Stefanov, B.B.; Nanayakkara, A.; Challacombe, M.; Peng, C.Y.; Ayala, P.Y.; Chen, W.; Wong, M.W.; Andres, J.L.; Replogle, E.S.; Gomperts, R.; Martin, R.L.; Fox, D.J.; Binkley, J.S.; Defrees, D.J.; Baker, J.; Stewart, J.P.; Head-Gordon, M.; Gonzalez, C. and Pople, J.A. Gaussian Inc., Pittsburgh, PA, USA, 1995
13. Gundertofte, K.; Liljefors, T.; Norrby, P.; Pettersson, I. *J. Comput. Chem.* **1996,** *17*, 429-449.
14. Gundertofte, K.; Palm, J.; Pettersson, I.; Stamvik, A. *J. Comput. Chem.* **1991,** *12*, 200-208.
15. Sprague, J.; Tai, J.; Yuh, Y.; Allinger, N. *J. Comput. Chem.* **1987,** *8*, 581-603.
16. Tai, J.; Allinger, N. *J. Am. Chem. Soc.* **1988,** *110*, 2050-2055.
17. Clark, M.; Cramer III, R.; Van Opdenbosch, N. *J. Comput. Chem.* **1989,** *10*, 982-1012.
18. Momany, F.; Rone, R. *J. Comput. Chem.* **1992,** *13*, 888-900.
19. Halgren, T. *J. Comput. Chem.* **1996,** *17*, 634-641.
20. Serra, M.A. *Acta Cryst., C* **1992,** *48*, 1957.
21. Berg, U.; Bladh, N. *J. Comput. Chem.* **1996,** *17*, 396-408.