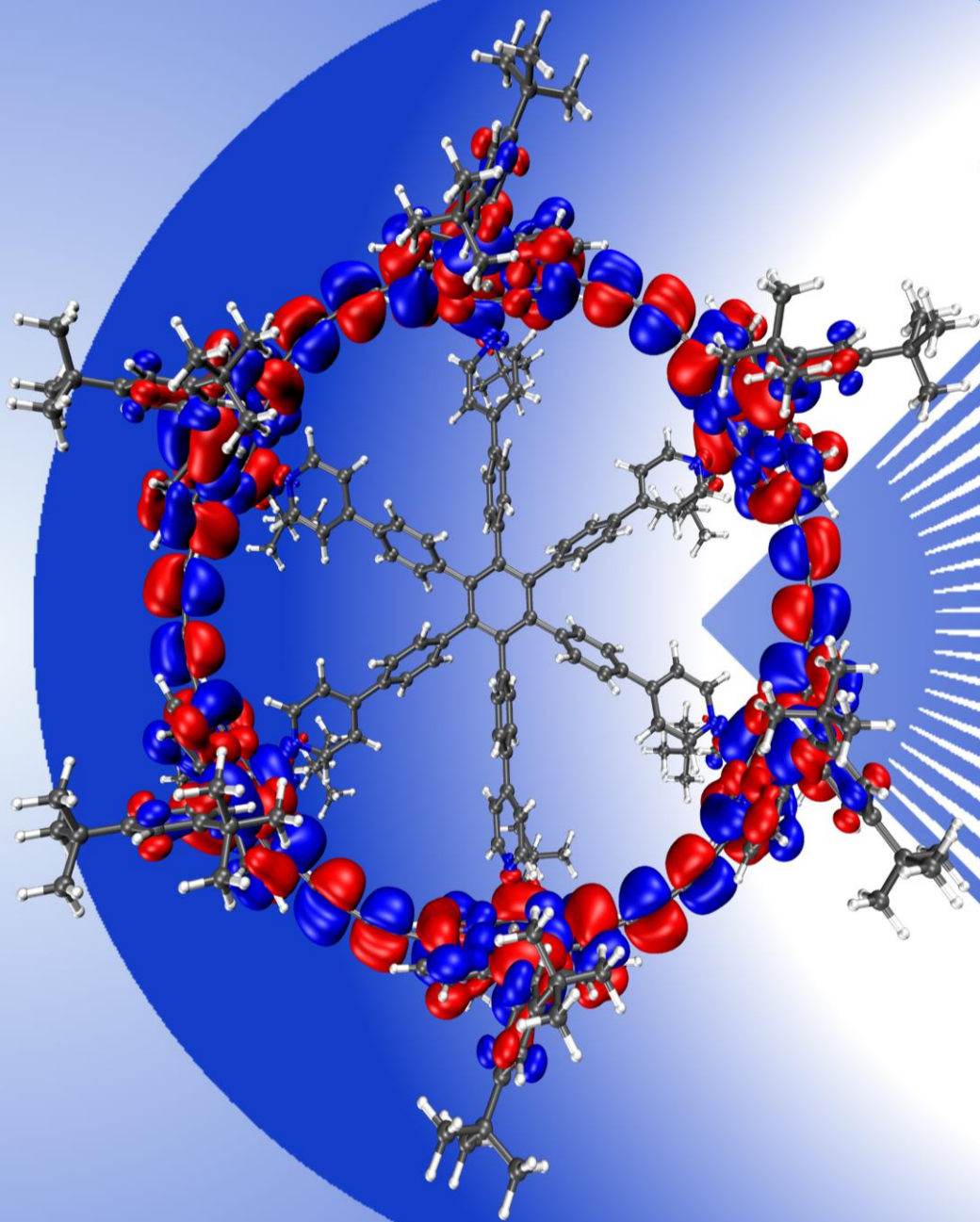


14



EMPIRETM

Impressum

Copyright

© 2015 by CEPOS InSilico Ltd.

The Old Vicarage
132 Bedford Road
Kempston
BEDFORD, MK42 8BQ
United Kingdom
www.ceposinsilico.com

Manual

Timothy Clark

Layout

www.eh-bitartist.de





EMPIRE™: a Parallel Semiempirical Molecular Orbital Program

TABLE OF CONTENTS

1 INTRODUCTION	4
2 INPUT FORMATS	5
2.1 .xyz and .xmol files	5
2.2 .sdf files	6
2.3 .dat files	7
3 PROGRAM OPTIONS	9
4 PERIODIC EMPIRE CALCULATIONS	12
5 PROGRAM OUTPUT	14
6 ADDITIONAL FILE FORMATS	19
6.1 EMPIRE™ archive file	19
6.2 HDF file format	20
6.3 Binary wavefunction file	21
6.4 SDF output file	22
7 GLOSSARY ENTRIES	25
8 SUPPORT	27
8.1 Contact	27
8.2 CEPOS InSilicoGmbH.	27
9 LIST OF TABLES	28
10 LIST OF FIGURES	29
11 REFERENCES	30



1 INTRODUCTION

EMPIRE™ is a completely new **NDDO**-based semiempirical molecular orbital (MO) program designed for moderately (the single-node version) or massively (the cluster version) parallel single-point calculations or geometry optimizations for closed-shell molecules within the **Restricted and Unrestricted Hartree-Fock** (RHF and UHF) **self-consistent field** (SCF) formalisms. Because of its radical new design, EMPIRE™ uses different algorithms for many steps in the calculation to the established traditional programs and can therefore provide impressive parallel performance, robustness and reliability and simplicity of operation. A further important advantage of EMPIRE™ is that it is able to calculate large molecules with ease (the size of the calculations is only restricted by the hardware available). For instance, one of the development tests for EMPIRE™ was to perform a single-point calculation on a molecule containing more than 76,000 atoms on 1,024 CPUs with high parallel efficiency. The current version of EMPIRE™ does not use **linear-scaling technologies**. It performs the full NDDO-based SCF calculations without approximations. It thus provides a “gold standard” against which all other programs can be measured.

EMPIRE™ currently only performs full geometry optimizations using Cartesian coordinates with a modified version of **Baker's EF-algorithm**.^[1] Later versions will provide other optimization options.

Citing EMPIRE:

EMPIRE14, M. Hennemann, J. T. Margraf, B. Meyer and T. Clark, Cepos InSilico Ltd, Bedford, 2015.

EMPIRE: A highly parallel semiempirical molecular orbital program: 1: Self-Consistent Field Calculations, M. Hennemann and T. Clark, J. Mol. Model. **2014**, 20, 2331 (11 pages, DOI: 10.1007/s00894-014-2331-4).

EMPIRE: A highly parallel semiempirical molecular orbital program: 2: Periodic Boundary Conditions, J. T. Margraf, M. Hennemann, B. Meyer and T. Clark, J. Mol. Model., **2015**, 21, 144 (7 pages, DOI: 10.1007/s00894-015-2692-3).



2 INPUT FORMATS

EMPIRE™ can use the following input file formats:

2.1 .xyz and .xmol files

Input files with the extension **.xyz** or **.xmol** are the simplest form of EMPIRE™ input. An example for adamantane is shown in **Figure 1**:

```
      26
Adamantane
C      3.495362    4.534200    3.532811    0.000000
C      2.257200    4.743420    4.412929    0.000000
H      3.639522    5.451926    2.872056    0.000000
H      2.379325    3.445733    1.981417    0.000000
H      1.339856    4.898190    3.754046    0.000000
H      2.412223    5.657913    5.075702    0.000000
C      3.104638    2.065800    3.532811    0.000000
C      4.342800    1.856580    4.412929    0.000000
H      4.218729    3.150494    1.979543    0.000000
H      5.260143    1.701809    3.754046    0.000000
H      4.187776    0.942086    5.075702    0.000000
C      2.065800    3.495362    5.277189    0.000000
C      1.856580    2.257200    4.397071    0.000000
H      1.148073    3.639522    5.937943    0.000000
H      3.154266    2.379325    6.828582    0.000000
H      0.946552    2.411425    3.727993    0.000000
H      1.695546    1.343203    5.059097    0.000000
C      4.534200    3.104638    5.277189    0.000000
C      4.743420    4.342800    4.397071    0.000000
H      5.451926    2.960477    5.937943    0.000000
H      3.441545    4.219989    6.830456    0.000000
H      5.653447    4.188574    3.727993    0.000000
H      4.904453    5.256796    5.059097    0.000000
C      3.300000    3.300000    2.637710    0.000000
C      3.300000    3.300000    6.172290    0.000000
H      2.960477    1.148073    2.872056    0.000000
```

Figure 1 .xyz (.xmol) input for adamantane

The input consists simply of the first line, which contains only the number of atoms in the molecule, a second, title line and then one line per atom containing the element symbol, x-, y- and z-coordinates and an optional (in this case fictitious) net atomic charge. The input is format-free. There may or may not be empty lines after the specification of the atoms.

Note that this file format does not allow the user to specify a molecular charge, which must then be specified (if it is not zero) via the command-line argument **charge=n** (see below).



2.2 .sdf files

The **structure-data (SD) file format version V2000** can be used as an input file with the extension **.sdf**. **Figure 2** shows an example, again for adamantane.

```
adamantane
Cepos2D3D 02041311173D 1 0.00000 0.00000 0
EMPIRE input file created by EMPIRE-caddle
26 28 0 0 0 0 1 V2000
-0.7632 -0.7127 1.2463 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0.7626 -0.7166 1.2484 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
1.2726 0.7213 1.2450 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0.7680 1.4401 -0.0025 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-0.7578 1.4420 -0.0031 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-1.2713 0.0054 -0.0002 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
1.2692 -1.4378 0.0030 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0.7632 -0.7212 -1.2452 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
1.2740 0.7164 -1.2465 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-0.7625 -0.7172 -1.2438 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-1.1421 -0.1945 2.1639 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-1.1447 -1.7653 1.2602 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
1.1351 -1.2449 2.1657 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
2.3922 0.7272 1.2602 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0.9105 1.2535 2.1614 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
1.1445 2.4973 -0.0044 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-1.1358 1.9782 -0.9106 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-1.1366 1.9819 0.9019 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-2.3936 0.0081 -0.0006 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0.9068 -2.4973 0.0049 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
2.3888 -1.4545 0.0033 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
1.1354 -1.2528 -2.1607 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0.9144 1.2454 -2.1657 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
2.3936 0.7214 -1.2596 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-1.1442 -1.7698 -1.2533 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
-1.1416 -0.2028 -2.1634 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
1 2 1 0 0 0
2 3 1 0 0 0
3 4 1 0 0 0
4 5 1 0 0 0
1 6 1 0 0 0
5 6 1 0 0 0
2 7 1 0 0 0
7 8 1 0 0 0
4 9 1 0 0 0
8 9 1 0 0 0
6 10 1 0 0 0
8 10 1 0 0 0
1 11 1 0 0 0
1 12 1 0 0 0
2 13 1 0 0 0
3 14 1 0 0 0
3 15 1 0 0 0
4 16 1 0 0 0
5 17 1 0 0 0
5 18 1 0 0 0
6 19 1 0 0 0
7 20 1 0 0 0
7 21 1 0 0 0
8 22 1 0 0 0
9 23 1 0 0 0
9 24 1 0 0 0
10 25 1 0 0 0
10 26 1 0 0 0
M END
$$$$
```

Figure 2 .sdf input for adamantane. The format is **structure-data file V2000**



If the **.sdf** format is used, EMPIRE™ can calculate the molecular charge from the **formal charges** of the atoms. The charge can be derived (in order of preference) from:

1. The **<TOTAL COULSON CHARGE>** floating point number
2. The **<TOTAL CHARGE>** integer
3. Summed atomic charges from **M CHG**
4. Summed atomic charges from the atoms block

The bond definitions have no effect on EMPIRE™ except that they are echoed to the output unmodified. The molecule is terminated by the **\$\$\$\$** line. EMPIRE™ can read **.sdf** files with multiple molecules. The next input simply starts after the **\$\$\$\$** line.

2.3 .dat files

The **.dat** file format is similar to that used by **Mopac™** and **Vamp™** and is shown in **Figure 3**.

```
Hamil=AM1 charge=0 calc=opt
adamantane
EMPIRE input file
C -2.81969539 1 -0.42849875 1 -0.75975579 1
C -2.18365402 1 0.47148130 1 -1.83477804 1
C -0.67864902 1 0.15981271 1 -1.94047398 1
C 0.00064190 1 0.43403192 1 -0.58652920 1
C -0.63342139 1 -0.46715388 1 0.48888876 1
C -2.13878008 1 -0.15942829 1 0.59514446 1
C -2.37988782 1 1.94829175 1 -1.44582725 1
C -1.70312377 1 2.21849294 1 -0.08896003 1
C -0.19846430 1 1.91055875 1 -0.19423207 1
C -2.33962326 1 1.31635509 1 0.98489736 1
H -2.67775461 1 -1.52266577 1 -1.04655070 1
H -3.93313548 1 -0.19703042 1 -0.68055019 1
H -2.68385018 1 0.27386442 1 -2.83994038 1
H -0.20925217 1 0.82113255 1 -2.74167472 1
H -0.53628795 1 -0.93588502 1 -2.22115130 1
H 1.11503337 1 0.20722999 1 -0.66585440 1
H -0.13323818 1 -0.26884374 1 1.49392100 1
H -0.48872273 1 -1.56086722 1 0.20174181 1
H -2.60574060 1 -0.82763989 1 1.39203911 1
H -3.49424775 1 2.17514548 1 -1.36620759 1
H -1.91233800 1 2.61641532 1 -2.24245014 1
H -1.84860379 1 3.31220683 1 0.19778976 1
H 0.30042996 1 2.10661307 1 0.81188282 1
H 0.27181706 1 2.58013872 1 -0.98801924 1
H -3.45462600 1 1.54188055 1 1.05910740 1
H -1.84331880 1 1.51579233 1 1.99162802 1
```

Figure 3 .dat input for an AM1 optimization of adamantane

The first line contains the program options, which are defined in **Section 3** below. Note that the standard Mopac™ or Vamp™ options will not be recognized, only the correct EMPIRE™ options. The



next two lines contain the molecular title and a comment, followed by one line per atom, which consists of the atomic symbol, x-, y- and z-coordinates, each followed by a **1** or **0** to indicate whether the coordinate should be optimized. The **1** or **0** optimization flags are necessary but have no effect on the EMPIRE™ optimization. The molecule is terminated by a blank line.



3 PROGRAM OPTIONS

EMPIRE™ options are either specified in the first line of the **.dat** file input or in the command line when the program is executed. The following table details the available options.

Table 1 EMPIRE™ Program options

Option	Details	
-h	Do not do a calculation, but print the list of options as a “help”	
Charge=<i>	Set the total molecular charge to <i>	
Mult=<n>	Set the Multiplicity to <n>	
Hamil=<s>	Set the semiempirical Hamiltonian to be used. The Hamiltonian files provided with EMPIRE™ allow the following values for <s>:	
	AM1	Use the AM1[2] Hamiltonian (this is the default if no Hamiltonian is defined)
	PM3	Use the PM3[3] Hamiltonian
	MNDO	Use the MNDO[4] Hamiltonian
	AM1S	Use the AM1*[5] Hamiltonian
MNDOD	Use the MNDO/d[6] Hamiltonian	
Method=<s>	Defines the method	
	RHF	Restricted Hartree-Fock (default)
UHF	Unrestricted Hartree-Fock	
Calc=<s>	Defines the type of calculation to be performed	
	SPT	Perform a single-point calculation (this is the default if no calculation is defined)
	OPT	Perform a full geometry optimization in Cartesian coordinates
	MD	Perform a molecular dynamics simulation
Initial Guess		
Guess=<s>	Initial guess Eigenvectors	
	Hückel	Extended Hückel calculation (Default)
	*.vwf	Read Eigenvectors from the specified .vwf file
	*.aux	Read Eigenvectors from the specified .aux file
InitDiag	Do an initial full diagonalization	
MixMOs [= -h, +l]	Mix the HOMO-h and LUMO+l orbitals of an UHF initial guess (Default: 0,0 => HOMO,LUMO ; for "Mult=1")	
Embedded Point Charges		
EmbedPCs	Calculate embedded point charges The point charges are specified in the .dat or .xyz files after the structure input, separated by a blank line, and terminated by a blank line, using the following format: X-Coord. Y-Coord. Z-Coord. Charge	
Periodic Boundary Conditions		
Periodic	Do a periodic calculation The 3 lattice vectors and optimization flags are specified after the molecular structure in the .dat or .xyz files, separated by a blank line. If a lattice optimization is requested, the first and second vector need to be aligned	



	with the x-axis and the xy-plane, respectively: Vector1-X 0/1 0.0 0 0.0 0 Vector2-X 0/1 Vector2-Y 0/1 0.0 0 Vector3-X 0/1 Vector3-Y 0/1 Vector3-Z 0/1					
ScreeningR=<f>	Simple screening radius beta in Ångström (Default = 30.0)					
DampingR=<f>	Integral damping radius in Bohr (Default = 10.0)					
DampingC=<f>	Integral damping coefficient (Default = 0.25)					
Output Options						
HDF	Write a HDF5 wavefunction file: <Basename>_e.h5					
VWF	Write a vamp wavefunction file: <Basename>_e.vwf (deprecated; please use HDF instead)					
XYZ [=<n>]	Write a XYZ trajectory file: <Basename>_e.xyz [add a structure every <n> OPT/MD steps (Default = 1)]					
DAT	Force output of a restart DAT file: <Basename>_e.dat					
SDF	Write a MDL structure-data file: <Basename>_e.sdf (SDF output is restricted to less than 1000 atoms; Atom type symbols are replaced by element symbols)					
SDFBonds=<s>	Type of bonds written to the output SDF					
	KEEP	do not alter bonds and bond types				
	CHECK	check if the connectivity changed (default)				
	ARYL	check connectivity and adjust aryl bond types to 4				
	CALC	calculate bond types from bond orders (Bond type 4 is used for aryl bonds)				
Center	Center the molecular structure for output after performing the calculation					
Calculation Details						
MaxTime=<n>	Define the maximum CPU-time (default = 365 days). Default units are seconds (e.g. 365 is interpreted as 365s), times can also be given in minutes (e.g. 30m), hours (e.g. 24h) or days (e.g. 7d)					
FullDiag=<n>	Do a full diagonalization every <n> OPT/MD steps (Default <n> = 1)					
Geometry Optimization						
MaxOpt=<n>	Define the maximum number of optimization cycles (default = 10,000)					
ConvergeG=<f>	Set the gradient norm criterion for the geometry optimization (default <f> = 0.4 kcal mol ⁻¹ Å ⁻¹)					
Molecular Dynamics						
MDsteps=<n>	Number of MD steps					
MDsteps=<n1,n2>	Use two numbers <n1,n2> to specify the starting and ending step (for restarting a MD calculation) (Default = 0,100)					
TimeStep=<f>	MD time step in fs (Default = 0.5)					
Temp=<f>	MD target temperature in K (Default = 300.0)					
Ensemble=<s>	MD Ensemble					
	NVT	Constant number of particles, volume and temperature (default)				
	NVE	Constant number of particles, volume and energy				
RelaxTime=<f>	Berendsen thermostat relaxation time in fs (Default = 100.0)					
ReadVelocities	Read initial velocities from the DAT file					



	(for restarting a MD calculation)	
Self Consistent Field		
MaxSCF=<n>	Define the maximum number of SCF cycles (default <n> = 1,000)	
ConvergeE=<f>	Set the energy convergence limit for the SCF (default <f>= 10 ⁻⁴ kcal mol ⁻¹)	
ConvergeR=<f>	Set the maximum off-diagonal CFC element for SCF convergence (default <f>= 10 ⁻⁴)	
CheckP	Use the convergence on the density matrix (RSS(Density)) to test for SCF convergence, instead of max(CFC). This option is not the default in EMPIRE™ and is not suitable for very large molecules. The default technique in EMPIRE™ is to test for the maximum rotation angle in the pseudorotation procedure . ^[7] This is equivalent to testing the density directly and far more suitable for large molecules.	
ConvergeP=<f>	Set the convergence limit for RSS(Density) (default <f>= 10 ⁻⁶). This option implies CheckP .	
IMemory=<n>	Use up to <n> MB/thread of memory for storing integrals (Default = 1024; Unlimited for periodic calculations)	
RecalcH	Calculate the one-electron matrix on-the-fly. This option saves memory and may be necessary for very large calculations. The default is to save the one-electron matrix, which is faster.	
FullRot	Do pseudorotations between all combinations of orbitals, including occupied-occupied and virtual-virtual rotations, rather than just virtual-occupied (This can help if the pseudodiagonalization fails).	
MinBandGap=<f>	Use dynamic level shifting to ensure that the band gap used for the SCF does not fall below <f> (default is to use level shifting dynamically and only if the SCF energy rises)	
AutoShift=<f>	Set automatic level shifting increment (Default = 1.0; Disabled when set to <= 0.0)	
PDcutoff=<f>	Defines the cutoff value for CFC at which rotations are included in the pseudodiagonalization (default = 0.04 × the maximum CFC value)	
EiCutoff=<f>	Defines the maximum value of CFC below which the Eigenvalues will not be recalculated in the SCF cycle. The default is <f>= 0.0, which means that the Eigenvalues are always calculated. Setting EiCutoff to higher values will make calculations faster but may slow down or prevent SCF convergence.	
Eigensolver=<*>	Preferred Eigensolver	
	DSYEVR	DSYEVR or PDSYEVR (Default)
	DSYEVD	DSYEVD or PDSYEVD
	ELPA1	ELPA one-stage
	ELPA2	ELPA two-stage

<n> positive number, <i> integer, <f> floating point number, <s> text string.



4 PERIODIC EMPIRE CALCULATIONS

EMPIRE performs calculations with periodic boundary conditions, if the keyword `periodic` is specified. The user has to supply three lattice vectors in line format at the bottom of the input file (.xyz and .dat formats are supported). The atom coordinates and lattice vectors must be separated by a blank line. In .dat files, optimization flags (1 or 0) must be given for each lattice coordinate, just as for the atomic coordinates (see examples below). If the calculation should only be periodic in 1 or 2 dimensions (e.g. for a polymer or a surface) simply supply the non-periodic dimensions as 0.000 0.000 0.000.

We recommend performing geometry optimizations of atomic coordinates and lattice vectors separately, by using the optimization flags in the .dat input format. A robust procedure is to first optimize the atomic coordinates and then the lattice vectors. This can be repeated, as necessary, to obtain a fully relaxed structure. In troublesome cases, it may be helpful to increase the SCF convergence criterion during the lattice optimization runs.

Example. xyz:

72

```
ZnO EMPIRE Input
Zn  -0.00000  1.87595  0.00000
O   -0.00000  1.87595  1.99161
Zn   1.62462  0.93797  2.60273
O    1.62462  0.93797  4.59433
[...]
Zn   3.24928  7.50387  5.20545
O    3.24928  7.50387  7.19705
Zn   4.87390  6.56590  7.80818
O    4.87390  6.56590  9.79978

    9.70780    0.00000    0.00000
    0.00000    8.44184    0.00000
    0.00000    0.00000   10.41090
```

Example .dat:

```
periodic calc=opt
Optimization of lattice vectors at fixed atomic coordinates
ZnO EMPIRE Input
Zn   0.00000  0  1.87595  0  0.00000  0
O    0.00000  0  1.87595  0  1.99161  0
Zn   1.62462  0  0.93797  0  2.60273  0
O    1.62462  0  0.93797  0  4.59433  0
[...]
Zn   3.24928  0  7.50387  0  5.20545  0
O    3.24928  0  7.50387  0  7.19705  0
```



```

Zn   4.87390  0  6.56590  0  7.80818  0
O    4.87390  0  6.56590  0  9.79978  0

    9.70780  1    0.00000  0    0.00000  0
    0.00000  0    8.44184  1    0.00000  0
    0.00000  0    0.00000  0   10.41090  1

```

Born-Oppenheimer Molecular Dynamics

The keyword `CALC=MD` specifies that a Born-Oppenheimer molecular dynamics simulation is performed. Currently non-periodic and periodic simulations can be performed in the NVT and NVE ensembles. In NVT simulations, the temperature is regulated via the Berendsen thermostat. The simulation parameters are set via the keywords `TimeStep`, `MDsteps`, `Temp`, `RelaxTime` and `Ensemble=[NVT,NVE]`. Times are given in fs, temperatures in K.

If `ReadVelocities` is specified, the initial atomic velocities are read from the input file, where they should be specified after each coordinate (see example). This feature is especially useful for restarting simulations. If `ReadVelocities` is not specified, the velocities are randomly initialized to the requested temperature.

Please note that the optimization flags also work in MD simulations. This can be used to partially constrain the geometry of the system.

```

Calc=MD TimeStep=0.5 RelaxTime=400 MDsteps=1000 ReadVelocities
Molecular Dynamics Simulation
EMPIRE Input
H 8.9125 1 -10.6523 1 -7.2520 1 -0.0023 0.0023 0.0095
H 7.5429 1 -11.3054 1 -7.0816 1 0.0235 -0.0240 -0.0017
O 8.4326 1 -11.3652 1 -6.8007 1 0.0027 -0.0001 -0.0070

```



5 PROGRAM OUTPUT

The EMPIRE™ output for an adamantane single point (using the input shown in **Figure 1** and the command `empire.exe adamantane.xmol hamil=AM1`) is shown in **Figure 4**.

```

EEE M M PPP III RRR EEE
E MM MM P P I R R E
EE M M M PPP I RRR EE
E M M M P I R R E
EEE M M P III R R EEE

EMPIRE'14 OMP (Revision 1950)

(c) Matthias Hennemann, Johannes Margraf, David Whitley, Tim Clark*

References:
=====

1) Matthias Hennemann, Timothy Clark,
"EMPIRE: a highly parallel semiempirical molecular orbital program:
1: self-consistent field calculations",
J Mol Model (2014) 20:2331 (11 pages).
DOI 10.1007/s00894-014-2331-4

2) Johannes T. Margraf, Matthias Hennemann, Bernd Meyer, Timothy Clark,
"EMPIRE: a highly parallel semiempirical molecular orbital program:
2: periodic boundary conditions",
J Mol Model (2015) 21:144 (7 pages).
DOI 10.1007/s00894-015-2692-3

<> Computing resources

Number of cores = 1
MPI tasks = 1
OMP threads = 1
OMP max threads = 1
OMP num procs = 16

<> Options

Input file = adamantane.xyz

Molecular Charge Charge = 0
Multiplicity Mult = 1

Hamiltonian Hamil = AM1
Method Method = RHF
Calculation Type Calc = SPT
Initial guess Guess =
Initial diagonalization InitDiag = F
Mix initial UHF orbitals MixMOs = F

Maximum computation time MaxTime = 31536000 seconds
Maximum optimization steps MaxOPT = 10000
Maximum SCF Iterations MaxSCF = 1000

Converge gradient norm ConvergeG = 4.0E-01
Converge delta(HoF) ConvergeE = 1.0E-04 kcal/mol
Converge max(CFC) ConvergeR = 1.0E-04

Pseudodiag. Cutoff PDCutoff = 4.0E-02
Eigenvalues Cutoff EiCutoff = 0.0E+00

Do full diagonalizations FullDiag = 1
Memory reserved for integrals IMemory = 1024 MB/thread
Recalculate Hcore RecalcH = F
Full rotations FullRot = F

Preferred Eigensolver Eigensolver = DSYEVR

<> Parameters

/ccc/SOFT/Cepos/cepos12/etc//AM1.par

```

[illegible]

Number of atoms	=	26
Number of electrons	=	56
Number of orbitals	=	56
Number of occupied orbitals	=	28

<> SCF

Step	Time Used	Heat of Formation	Delta HoF	max(CFC)	E (HOMO)	E (LUMO)
0	0.013	107.477013	0.000E+00	2.021E+00	-12.372	1.091
1	0.014	-24.119452	-1.316E+02	5.465E-01	-10.857	3.476
2	0.015	-36.829600	-1.271E+01	2.090E-01	-10.579	3.791
3	0.016	-38.920810	-2.091E+00	8.371E-02	-10.469	3.907
4	0.016	-39.296860	-3.761E-01	3.409E-02	-10.425	3.952
5	0.017	-39.366772	-6.991E-02	1.410E-02	-10.407	3.971
6	0.018	-39.380615	-1.384E-02	5.916E-03	-10.400	3.978
7	0.019	-39.383513	-2.898E-03	2.485E-03	-10.397	3.981
8	0.020	-39.384162	-6.491E-04	1.151E-03	-10.395	3.983
9	0.020	-39.384317	-1.549E-04	5.764E-04	-10.395	3.983
10	0.021	-39.384357	-3.944E-05	2.958E-04	-10.394	3.984
11	0.022	-39.384367	-1.065E-05	1.654E-04	-10.394	3.984
12	0.023	-39.384370	-2.984E-06	9.229E-05	-10.394	3.984

Heat of formation	=	-39.384370 kcal/mol
Electronic energy	=	-9210.208280 eV
Core-core repulsion	=	7707.744785 eV
Ionization potential	=	10.394216 eV (approx.)
Electron affinity	=	-3.983803 eV (approx.)
Computation time	=	0.023 seconds

Heat of formation	=	-39.384371 kcal/mol
Electronic energy	=	-9210.208280 eV
Core-core repulsion	=	7707.744785 eV
Ionization potential	=	10.343278 eV

© CEPOS InSilico2015



```
DOI 10.1007/s00894-014-2331-4

2) Johannes T. Margraf, Matthias Hennemann, Bernd Meyer, Timothy Clark,
"EMPIRE: a highly parallel semiempirical molecular orbital program:
2: periodic boundary conditions",
J Mol Model (2015) 21:144 (7 pages).
DOI 10.1007/s00894-015-2692-3

<> Computing resources

Number of cores =      1
MPI tasks       =      1
OMP threads     =      1
OMP max threads =      1
OMP num procs   =     16

<> Options

Input file                      = adamantane.xyz

Molecular Charge                Charge = 0
Multiplicity                    Mult = 1

Hamiltonian                     Hamil = PM3
Method                          Method = RHF
Calculation Type                Calc = OPT

Initial guess                   Guess =
Initial diagonalization         InitDiag = F
Mix initial UHF orbitals        MixMOs = F

Maximum computation time        MaxTime = 31536000 seconds
Maximum optimization steps      MaxOPT = 10000
Maximum SCF Iterations          MaxSCF = 1000

Converge gradient norm         ConvergeG = 4.0E-01
Converge delta(HoF)            ConvergeE = 1.0E-04 kcal/mol
Converge max(CFC)              ConvergeR = 1.0E-04

Pseudodiag. Cutoff             PDCutoff = 4.0E-02
Eigenvalues Cutoff             EiCutoff = 0.0E+00

Do full diagonalizations        FullDiag = 1
Memory reserved for integrals   IMemory = 1024 MB/thread
Recalculate Hcore               RecalcH = F
Full rotations                  FullRot = F

Preferred Eigensolver           Eigensolver = DSYEVR

<> Parameters

/ccc/SOFT/Cepos/cepos12/etc//PM3.par

Hamiltonian = PM3

H      J.J.P. Stewart, J. Comp. Chem., 10, 209-220 (1989).
Li     E. Anders, R. Koch, P. Freunscht, J. Comput. Chem., 14, 1301-1312 (1993).
Be     J.J.P. Stewart, J. Comp. Chem., 12, 320-341 (1991).
B      J.J.P. Stewart, J. Mol. Model., 10, 155-164 (2004).
C      J.J.P. Stewart, J. Comp. Chem., 10, 209-220 (1989).
N      J.J.P. Stewart, J. Comp. Chem., 10, 209-220 (1989).
O      J.J.P. Stewart, J. Comp. Chem., 10, 209-220 (1989).
F      J.J.P. Stewart, J. Comp. Chem., 10, 209-220 (1989).
Na     J.J.P. Stewart, J. Mol. Model., 10, 155-164 (2004).
Mg     J.J.P. Stewart, J. Comp. Chem., 12, 320-341 (1991).
Al     J.J.P. Stewart, J. Comp. Chem., 10, 209-220 (1989).
Si     J.J.P. Stewart, J. Comp. Chem., 10, 209-220 (1989).
P      J.J.P. Stewart, J. Comp. Chem., 10, 209-220 (1989).
S      J.J.P. Stewart, J. Comp. Chem., 10, 209-220 (1989).
Cl     J.J.P. Stewart, J. Comp. Chem., 10, 209-220 (1989).
K      J.J.P. Stewart, J. Mol. Model., 10, 155-164 (2004).
Ca     J.J.P. Stewart, J. Mol. Model., 10, 155-164 (2004).
Zn     J.J.P. Stewart, J. Comp. Chem., 12, 320-341 (1991).
Ga     J.J.P. Stewart, J. Comp. Chem., 12, 320-341 (1991).
Ge     J.J.P. Stewart, J. Comp. Chem., 12, 320-341 (1991).
As     J.J.P. Stewart, J. Comp. Chem., 12, 320-341 (1991).
Se     J.J.P. Stewart, J. Comp. Chem., 12, 320-341 (1991).
Br     J.J.P. Stewart, J. Comp. Chem., 10, 209-220 (1989).
Rb     J.J.P. Stewart, J. Mol. Model., 10, 155-164 (2004).
Sr     J.J.P. Stewart, J. Mol. Model., 10, 155-164 (2004).
Br     J.J.P. Stewart, J. Comp. Chem., 10, 209-220 (1989).
Rb     J.J.P. Stewart, J. Mol. Model., 10, 155-164 (2004).
Sr     J.J.P. Stewart, J. Mol. Model., 10, 155-164 (2004).
```

<> Time used : 0.640 s

© CEPOS InSilico2015

© CEPOS InSilico2015



6.2 HDF file format

HDF5 is a binary file format supported by the HDF Group (www.hdfgroup.org). It was designed for storing hierarchical data structures containing large numerical arrays. The overall structure of an HDF5 file is a tree with nodes called *Groups*. Each group may contain other groups, *Attributes* and *Datasets*. Attributes are intended for small data items and datasets store n-dimensional arrays. The hierarchy is based at the *root* group, denoted “/”, and objects are located via paths, e.g. “/group1/group2/datasetD”. This mimics the structure of a Unix file system, with groups in place of directories, and attributes and datasets in place of files. The conventional file extension for HDF5 files is “.h5”.

A cross-platform visual tool, HDFView, for browsing and editing HDF5 files, is available from the HDF Group site. Alternatively, the HDF5 distribution contains a command-line utility program, h5dump, which converts a binary HDF5 file into a readable Data Definition Language (DDL) format. The usage is simply “h5dump <file>.h5” and the output is written to the standard output channel. The “-H” option restricts the output to the header information (without the data), allowing the overall structure of the file to be seen more clearly.

Internally, HDF5 stores arrays of rank > 1 in row-major (C-style) order. This has implications for files that will be accessed by programs written in both Fortran and C-style languages (i.e. C/C++ and Java). When reading and writing matrices from Fortran, the conversion from the column-major memory storage used by Fortran to the row-major storage in the HDF5 file is transparent and can be ignored. However, a C/C++ or Java program will see the transpose of a matrix written by a Fortran program, and vice-versa. This applies, in particular, to h5dump (written in C) and HDFView (written in Java), which display the transpose of all matrices in HDF5 files written by EMPIRE (written in Fortran).

The overall structure of an HDF5 Wavefunction file is shown in Table XX. All HDF5 object names are in uppercase with no spaces, with multiple words broken by underscores where appropriate.

Table 2 HDF5 Wavefunction file structure

Group	/ (root)		
	Attribute CH5TYPE		
	Attribute CH5VERSION		
	Attribute PROGRAM		
	Group PROGRAM_OPTIONS		
	Group MOLECULES		
		Group MOLECULE_ID	
			Group IDENTITY
			Group STRUCTURE
			Group WAVEFUNCTION
			Group HAMILTONIAN

The root group contains three attributes:

- CH5TYPE A string defining the file format.



- **CH5VERSION** A compound data type comprising three integers representing a triplet version number (majnum, minnum, relnum).
- **PROGRAM** A string describing the program used to generate the file.

The **PROGRAM_OPTIONS** group contains a set of attributes recording the command-line options used to create the file.

The **MOLECULES** group contains a list of molecule groups labelled by a **MOLECULE_ID** "1", "2", "3", etc. The data for each molecule is then contained in the following groups:

- **IDENTITY** Molecule identifiers (Name, ID etc.).
- **STRUCTURE** Molecular structure (atomic numbers, charges and coordinates).
- **WAVEFUNCTION** The wavefunction eigenvalues and eigenvectors.
- **HAMILTONIAN** The Hamiltonian parameters for the atom types in the molecule.

Precise details of the entries in the HDF5 Wavefunction file are available on request.

6.3 Binary wavefunction file

The binary wavefunction file contains all the information necessary to process the results of the EMPIRE™ calculation further. Its contents are:

Table 3 Contents of binary wavefunction file

Definition	Type
First comment line from the input	character(len=80)
Title line from the input	character(len=80)
Hamiltonian	character(len=6) "AM1 ", "AM1* ", "MNDO ", "MNDO/c", "MNDO/d", "PM3 " or "PM6 "
Formalism	character(len=3) "RHF"
Number of atoms	integer
Number of orbitals (Norbs)	integer
Number of doubly occupied orbitals	integer
Number of singly occupied orbitals	integer = 0
Charge on the molecule	integer
Heat of formation in kcal mol ⁻¹	double precision
Energy of the HOMO (eV)	double precision
Energy of the LUMO (eV)	double precision
x, y and z-components of the dipole moment (Debye)	double precision(1:3)



For each atom:		
Atomic number		integer
Cartesian coordinates		double precision(1:3)
Number of atomic orbitals		integer
Principal quantum number (s and p)		integer
Principal quantum number (d)		integer
Slater exponents (ζ_s , ζ_p , ζ_d)		double precision(1:3)
Overlap integrals (β_s , β_p , β_d)		double precision(1:3)
Multipole parameters (dd and qq)		double precision(1:2)
End atoms		
Eigenvalues (eV)		Double precision(1:Norbs)
Eigenvectors		Double precision(1:Norbs2)

6.4 SDF output file

The output file for a PM3 optimization of adamantane using an **sdf** file as input is shown in **Figure 7**.

adamantane XXEMPIRE1408191511433D 1 1.00000 0.00000 0 EMPIRE input file created by EMPIRE-caddie 26 28 0 0 0 0 1 V2000 -0.7658 -0.7157 1.2505 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.7648 -0.7190 1.2524 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1.2766 0.7235 1.2488 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.7697 1.4446 -0.0025 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -0.7609 1.4456 -0.0031 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1.2756 0.0041 -0.0003 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1.2732 -1.4426 0.0031 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.7655 -0.7237 -1.2491 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1.2779 0.7186 -1.2504 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -0.7651 -0.7203 -1.2480 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1.1491 -0.2176 2.1620 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1.1512 -1.7532 1.2797 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1.1354 -1.2438 2.1635 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 2.3833 0.7376 1.2777 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.9367 1.2528 2.1598 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1.1433 2.4949 -0.0043 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1.1433 1.9870 -0.8899 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1.1441 1.9910 0.8810 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -2.3905 0.0064 -0.0005 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.9319 -2.4958 0.0050 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 2.3799 -1.4763 0.0034 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1.1363 -1.2516 -2.1583 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.9396 1.2446 -2.1640 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 2.3846 0.7321 -1.2778 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1.1505 -1.7580 -1.2735 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1.1479 -0.2259 -2.1616 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 2 1 0 0 0 0 2 3 1 0 0 0 0 3 4 1 0 0 0 0 4 5 1 0 0 0 0 1 6 1 0 0 0 0 5 6 1 0 0 0 0 2 7 1 0 0 0 0 7 8 1 0 0 0 0 4 9 1 0 0 0 0 8 9 1 0 0 0 0 6 10 1 0 0 0 0 8 10 1 0 0 0 0 1 11 1 0 0 0 0 1 12 1 0 0 0 0 2 13 1 0 0 0 0 3 14 1 0 0 0 0 4 5 1 0 0 0 0 1 6 1 0 0 0 0 5 6 1 0 0 0 0 2 7 1 0 0 0 0 7 8 1 0 0 0 0 4 9 1 0 0 0 0 8 9 1 0 0 0 0 6 10 1 0 0 0 0 8 10 1 0 0 0 0 1 11 1 0 0 0 0 1 12 1 0 0 0 0 2 13 1 0 0 0 0 3 14 1 0 0 0 0 3 15 1 0 0 0 0 4 16 1 0 0 0 0 5 17 1 0 0 0 0 5 18 1 0 0 0 0	> <COULSON CHARGES> [EMPIRE'14] -0.0939611467 -0.0727413007 -0.0939264079 -0.0727515856 -0.0940058157 -0.0727714670 -0.0939661993 -0.0727582705 -0.0939156611 -0.0939621862 0.0509602786 0.0509760227 0.0608050135 0.0509637460 0.0509712395 0.0608177347 0.0509336465 0.0509440071 0.0608217440 0.0509341364 0.0509643598 0.0608056022 0.0509610890 0.0509706991 0.0509747162 0.0509560053 > <TOTAL COULSON CHARGE> [EMPIRE'14] 0.0000000000 > <DENSITY MATRIX ELEMENTS> [EMPIRE'14] 1.1649094435 -0.0072558726 0.9550774770 -0.0051293933 0.0033643040 1.0006975938 0.0089123957 -0.0056772235 0.0235811659 0.9732766324 1.1777117473 0.0060953867 0.9579384570 -0.0086088333 -0.0049571773 0.9614616214 0.0149846768 0.0086077314 -0.0122010105 0.9756294750 1.1648976049 0.0072836407 0.9976328720 0.0051305238 -0.0118798198 0.9581308318 0.0089944466 -0.0202972985 0.0133141551 0.9732650993 1.1777059425 0.0061932331 0.9580043784 0.0172577496 0.0099646162 0.9826236834 -0.0000342533 -0.0000088691 -0.0000502211 0.9544175814 1.1649538858 -0.0071906554 0.9550368413 0.0102514997 -0.0066106345 0.9597529735 -0.0000092374 -0.0000005910 0.0001076138 1.0142621151 1.1776992704 -0.0183404250 0.9861541801 0.0000264780 -0.0000831807 0.9544593993 0.0000099717 0.0000157797 0.0000069952 0.9544586172 1.1649158283 0.0072451918 0.9977311689 -0.0102421912 0.0234759861 0.9809450208 0.0000173605 -0.0000330480 -0.0000548042 0.9503741813 1.1777024169 0.0060873591 0.9579490279 -0.0086615453 -0.0050218211 0.9615861725 -0.0149579901 -0.0086032500 0.0122404643 0.9755206531 1.1648894691 0.0072914110 0.9975837500 0.0051034369 -0.0118118204 0.9580497855 -0.0089019424 0.0203651962 -0.0132805567 0.9733926566 1.1649025110 -0.0072651024 0.9550677656 -0.0051523681 0.0033859515 1.0005347996 -0.0088807965 0.0056611576 -0.0236777799 0.9734571101 > <CHARGE ON HYDROGENS> [EMPIRE'14] 0.0509602786 0.0509760227 0.0608050135 0.0509637460 0.0509712395 0.0608177347 0.0509336465 0.0509440071 0.0608217440 0.0509341364 0.0509643598 0.0608056022 0.0509610890 0.0509706991 0.0509747162 0.0509560053 > <ORBITAL VECTORS> [EMPIRE'14] RHF calculation 28 56 28 28 Nhom, Nao, Nocc, Nunocc Nr. of atomic orbitals of each atom 4 4 4 4 4 4 4 4 4 4 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 Orbital information from occ. to unocc. -44.608 -0.266 -0.037 -0.026 0.045 -0.301 0.020 -0.028 0.049 -0.266 0.037 0.026 0.045 -0.301 0.020 0.056 -0.000 -0.266 -0.037 0.052 -0.000 -0.301 -0.060 0.000 -0.000 -0.266 0.037 -0.052 0.000 -0.301 0.020 -0.028 -0.049 -0.266 0.037 0.026 0.045 -0.266 -0.037 -0.026 -0.045 -0.104 -0.104 -0.105 -0.104 -0.104 -0.105 -0.104 -0.104 -0.105 -0.104 -0.104 -0.105 -0.104 -0.104 -0.104 -0.104 -32.559 Further orbitals not shown here 5.335 -0.018 -0.123 0.182 0.082 0.027 -0.108 -0.285 0.038 0.093 -0.153 -0.154 -0.201 0.046 -0.044 0.251 -0.173
--	--



6 19 1 0 0 0 0	0.007 -0.104 -0.101 0.150 -0.040 0.195 -0.164 -0.172
7 20 1 0 0 0 0	-0.007 0.192 0.108 -0.099 -0.032 -0.296 -0.018 0.081
7 21 1 0 0 0 0	0.018 0.106 -0.186 0.059 -0.093 -0.181 -0.074 -0.225
8 22 1 0 0 0 0	-0.134 0.100 -0.105 0.078 0.101 -0.180 0.095 -0.081
9 23 1 0 0 0 0	0.158 0.123 -0.136 0.126 0.117 -0.084 -0.061 -0.118
9 24 1 0 0 0 0	5.335
10 25 1 0 0 0 0	-0.094 -0.150 -0.155 0.204 -0.044 -0.278 0.096 -0.090
10 26 1 0 0 0 0	-0.018 0.181 -0.137 0.010 0.030 -0.043 0.169 0.253
M END	0.001 -0.079 -0.059 -0.220 -0.029 0.141 -0.101 0.253
> <CALCULATION> [EMPIRE'14]	-0.001 0.124 0.085 0.144 0.043 -0.003 -0.296 -0.084
9	0.094 -0.200 -0.090 0.201 0.018 -0.066 0.219 0.006
Geometry optimized using EF	-0.070 -0.109 0.171 -0.125 0.091 -0.117 -0.129 0.131
> <HAMILTONIAN> [EMPIRE'14]	0.114 0.088 -0.089 -0.168 0.068 0.111 0.127 -0.093
PM3	5.486
> <MOPACBASICS> [EMPIRE'14]	-0.025 -0.150 -0.105 0.183 0.089 -0.099 0.141 -0.244
-34.582993 -10.851084 3.573030 -0.000226 0.000023	-0.025 0.150 0.106 0.182 0.089 -0.100 -0.281 0.000
0.000015	-0.025 -0.149 0.211 -0.000 0.089 0.299 -0.001 -0.000
> <ENERGY> [EMPIRE'14]	-0.025 0.149 -0.211 0.001 0.089 -0.099 0.141 0.244
-9099.535505 7666.308684 -1433.226821	-0.025 0.150 0.105 -0.183 -0.025 -0.149 -0.106 -0.183
	-0.089 -0.089 0.168 -0.089 -0.088 0.168 -0.088 -0.089
	0.168 -0.088 -0.089 0.168 -0.088 -0.089 -0.088 -0.089
	> <CALCULATION STATUS> [EMPIRE'14]
	SUCCESS
	\$\$\$\$

Figure 7 The EMPIRE™ output .sdf file for a PM3 geometry optimization of adamantane

All EMPIRE™-specific tags use the suffix (**EMPIRE'12**). These are:

Table 4 EMPIRE™ specific tags

Tag	Contents
><CALCULATION>	1 Single point calculation <n> Geometry optimized using EF <n> is the number of optimization cycles (always 1 for a single point)
><HAMILTONIAN>	AM1, PM3, MNDO, AM1* or MNDOD
><MOPACBASICS>	1. Heat of formation 2. HOMO energy (eV) 3. LUMO energy (eV) x-, y-, and z-components of the dipole moment (Debye) Format: 6F13.6
><CHARGE>	Molecular charge as integer
><ENERGY>	1. Electronic energy 2. Core-core repulsion 3. Total energy All in eV, format 3F20.6
><COULSON CHARGES>	Coulson net atomic charges (one floating point per atom) Format: 5F14.10
><TOTAL COULSON CHARGE>	Sum of the Coulson charges (floating point) Format: F14.10
><DENSITY MATRIX ELEMENTS>	The one-atom blocks of the density matrix printed as lower triangles for each atom (10 elements (2 lines) for s,p-atoms and 45 elements (9 lines) for s,p,d-atoms. Hydrogen blocks are not included. Format: 5F14.10
><CHARGE ON HYDROGENS>	The Coulson charges of the hydrogen atoms (this block is redundant as the charges are also contained in the "><COULSON CHARGES>" block). Format: 5F14.10



><ORBITAL VECTORS>		Block for the molecular orbitals Eigenvectors	
	RHF calculation	1. Number of the HOMO (Nhomo) 2. Number of orbitals (Nao) 3. Number of occupied orbitals (Nocc) 4. Number of virtual orbitals (Nunocc)	Format: 4(i4,1x),
	Nr. of atomic orbitals of each atom	The number of basis functions (atomic orbitals) per atom; s-only = 1; s,p = 4, s,p,d = 9.	Format: 20i4
	Orbital information from occ. to unocc.	For each molecular orbital:	
		The Eigenvalue (eV)	Format: F9.3
		The MO coefficients	Format: F6.3,7F7.3
><CALCULATION STATUS>		SUCCESS	



7 GLOSSARY ENTRIES

CFC elements

“CFC elements” are the calculated interaction matrix elements between virtual and occupied orbitals that must be made zero in the SCF procedure. They are the basis for the pseudodiagonalization procedure.^[7]

Convergence on the density matrix

At SCF convergence, the density matrix should not change between iterations. This can be tested by calculating the maximum difference of any density-matrix element between cycles. This is the usual procedure in SCF calculations but is impractical for very large systems, so that by default, EMPIRE™ uses the CFC criterion, which is essentially equivalent to checking the convergence on the density matrix.

Coulson charge

Adapted from T.Clark, in Chemoinformatics – From Data to Knowledge (Ed.: J. Gasteiger), WILEY-VCH, Weinheim, **2003**, pp. 947-976.

The number of electrons assigned to an atom is expressed as the sum of the populations of the atomic orbitals (AOs) centered at its nucleus. The Coulson analysis assumes that the orbitals are orthogonal, which leads to the very simple expression that the electronic population of P_i of atom i is given by the following equation:

$$P_i = \sum_{j=1}^{N_{occ}} \sum_{k=i_{first}}^{i_{last}} n_j c_{j,k}^2$$

Where N_{occ} is the number of occupied molecular orbitals (MOs), i_{first} and i_{last} are the first and last atomic orbitals centered on atom i , respectively, n_j is the occupancy number of the molecular orbital j , and $c_{i,k}$ is the coefficient of AO k in MO j . The net atomic charge is simply the sum of the electronic population of the atom and its nuclear charge.

Dynamic level shifting

When solving the Hartree-Fock-Roothan equations problems concerning the convergence of the calculations sometimes arise. As a rule, this is connected with two problems: the slow convergence of iteration process and the occurrence of oscillations. These oscillations can be suppressed and convergence obtained using the dynamic level shift method, in which the energies of virtual orbitals are artificially increased.

The dynamic “level shift” method for improving the convergence of the SCF procedure, A. V. Mitin, Journal of Computational Chemistry, 1988, 9, 107-110.

**Gradient norm criterion**

The gradient norm is the square root of the sum of the squares of the energy gradients with respect to geometric distortions. The gradient norm should be zero at a minimum or transition state. In practice, the optimization stops when a sufficiently small value is reached.

Linear-scaling techniques

These are techniques for which the computational effort scales linearly with the size (number of atoms) of the system.

One-electron matrix

The one-electron matrix contains the electron-nucleus interactions and is used to form the Fock matrix, which is diagonalized in the SCF procedure. The one-electron matrix does not contain electron-electron interactions.

Self-consistent field

An iterative procedure used to obtain the converged wavefunction starting from guessed molecular orbitals (the initial guess). The procedure is judged to be converged when neither the energy nor the wavefunction change significantly from iteration to iteration.

8 SUPPORT

8.1 Contact

Questions regarding EMPIRE™ should be sent directly to:

support@ceposinsilico.com

8.2 CEPOS InSilicoGmbH.

Computer-Chemie-Centrum (CCC)
Nägelsbachstr. 25
91052 Erlangen
Germany

support@ceposinsilico.com

Tel. +49 (0)9131 9704910

Fax. +49 (0)91319704911

www.ceposinsilico.com/contact



9 LIST OF TABLES

Table 1	EMPIRE™ Program options	9
Table 2	HDF5 Wavefunction file structure	20
Table 3	Contents of binary wavefunction file	21
Table 4	EMPIRE™ specific tags	23



10 LIST OF FIGURES

Figure 1	.xyz (.xmol) input for adamantane.....	5
Figure 2	.sdf input for adamantane. The format is standard structural data file V2000	6
Figure 3	.dat input for an AM1 optimization of adamantane	7
Figure 4	The EMPIRE™ output for an AM1 single-point calculation on adamantane	16
Figure 5	The EMPIRE™ output for a PM3 geometry optimization of adamantane	18
Figure 6	The EMPIRE™ archive file for a PM3 geometry optimization of adamantane	19
Figure 7	The EMPIRE™ output .sdf file for a PM3 geometry optimization of adamantane.....	23



11 REFERENCES

- [1] J. Baker. An algorithm for the location of transition states, *Journal of Computational Chemistry* **1986**, 7, 385-395. doi:10.1002/jcc.540070402
- [2] a) M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, J. J. P. Stewart. Development and use of quantum mechanical molecular models. 76. AM1: a new general purpose quantum mechanical molecular model, *Journal of the American Chemical Society* **1985**, 107, 3902-3909. doi:10.1021/ja00299a024;b) M. J. S. Dewar, C. Jie, J. Yu. SAM1; The first of a new series of general purpose quantum mechanical molecular models, *Tetrahedron* **1993**, 49, 5003-5038. doi:10.1016/S0040-4020(01)81868-8;c) R. O. Freire, G. B. Rocha, A. M. Simas. Sparkle Model for the Calculation of Lanthanide Complexes: AM1 Parameters for Eu(III), Gd(III), and Tb(III)†, *Inorganic Chemistry* **2005**, 44, 3299-3310. doi:10.1021/ic048530+;d) G. B. Rocha, R. O. Freire, A. M. Simas, J. J. P. Stewart. RM1: A reparameterization of AM1 for H, C, N, O, P, S, F, Cl, Br, and I, *Journal of Computational Chemistry* **2006**, 27, 1101-1111. doi:10.1002/jcc.20425;e) A. R. Leach, *Molecular Modelling: Principles and Applications*, Pearson Education Limited, **2001**.
- [3] a) J. J. P. Stewart. Optimization of parameters for semiempirical methods I. Method, *Journal of Computational Chemistry* **1989**, 10, 209-220. doi:10.1002/jcc.540100208;b) J. J. P. Stewart. Optimization of parameters for semiempirical methods II. Applications, *Journal of Computational Chemistry* **1989**, 10, 221-264. doi:10.1002/jcc.540100209;c) J. J. P. Stewart. Optimization of parameters for semiempirical methods. III Extension of PM3 to Be, Mg, Zn, Ga, Ge, As, Se, Cd, In, Sn, Sb, Te, Hg, Tl, Pb, and Bi, *Journal of Computational Chemistry* **1991**, 12, 320-341. doi:10.1002/jcc.540120306;d) J. P. Stewart. Optimization of parameters for semiempirical methods IV: extension of MNDO, AM1, and PM3 to more main group elements, *Journal of Molecular Modeling* **2004**, 10, 155-164. doi:10.1007/s00894-004-0183-z
- [4] M. J. S. Dewar, W. Thiel. Ground states of molecules. 38. The MNDO method. Approximations and parameters, *Journal of the American Chemical Society* **1977**, 99, 4899-4907. doi:10.1021/ja00457a004
- [5] a) P. Winget, A. C. Horn, C. Selçuki, B. Martin, T. Clark. AM1* parameters for phosphorus, sulfur and chlorine, *Journal of Molecular Modeling* **2003**, 9, 408-414. doi:10.1007/s00894-003-0156-7;b) P. Winget, T. Clark. AM1* parameters for aluminum, silicon, titanium and zirconium, *Journal of Molecular Modeling* **2005**, 11, 439-456. doi:10.1007/s00894-005-0236-y;c) H. Kayi, T. Clark. AM1* parameters for copper and zinc, *Journal of Molecular Modeling* **2007**, 13, 965-979. doi:10.1007/s00894-007-0214-7;d) H. Kayi, T. Clark. AM1* parameters for vanadium and chromium, *Journal of Molecular Modeling* **2009**, 15, 1253-1269. doi:10.1007/s00894-009-0489-y;e) H. Kayi, T. Clark. AM1* parameters for bromine and iodine, *Journal of Molecular Modeling* **2009**, 15, 295-308. doi:10.1007/s00894-008-0419-4;f) H. Kayi, T. Clark. AM1* parameters for manganese and iron, *Journal of Molecular Modeling* **2010**, 16, 1109-1126. doi:10.1007/s00894-009-0614-y;g) H. Kayi, T. Clark. AM1* parameters for cobalt and nickel, *Journal of Molecular Modeling* **2010**, 16, 29-47. doi:10.1007/s00894-009-0503-4;h) H. Kayi, T. Clark. AM1* parameters for palladium and silver, *Journal of Molecular Modeling* **2011**, 17, 2585-2600. doi:10.1007/s00894-010-0940-0
- [6] a) W. Thiel, A. A. Voityuk. Extension of MNDO to d Orbitals: Parameters and Results for the Second-Row Elements and for the Zinc Group, *The Journal of Physical Chemistry* **1996**, 100, 616-626. doi:10.1021/jp952148o;b) W. Thiel. Perspectives on Semiempirical Molecular Orbital Theory, in *Advances in Chemical Physics*, John Wiley & Sons, Inc., **2007**, pp. 703-757. doi:10.1002/9780470141526.ch10
- [7] J. J. P. Stewart, P. Császár, P. Pulay. Fast semiempirical calculations, *Journal of Computational Chemistry* **1982**, 3, 227-228. doi:10.1002/jcc.540030214