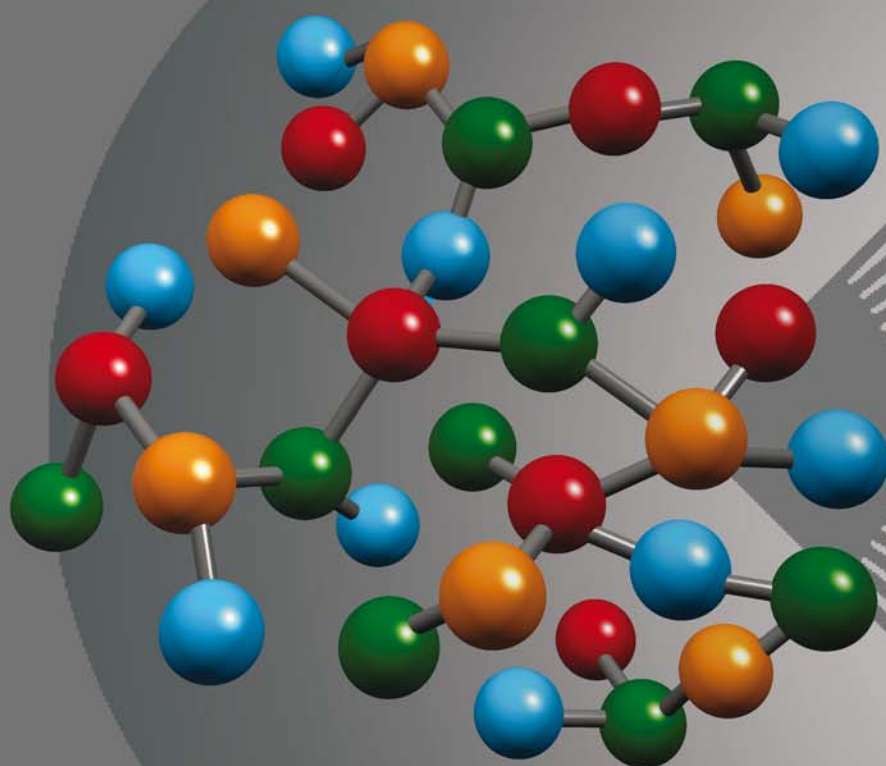


12

USER MANUAL



EMPIRE™

Impressum

Copyright

© 2013 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 PROGRAM OUTPUT	11
5 ADDITIONAL FILE FORMATS	15
5.1 EMPIRE™ archive file	15
5.2 Binary wavefunction file	16
5.3 SDF output file	17
6 GLOSSARY ENTRIES	20
7 SUPPORT	22
7.1 Contact	22
7.2 CEPOS InSilicoGmbH.	22
8 LIST OF TABLES	23
9 LIST OF FIGURES	24
10 REFERENCES	25



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 Hartree-Fock** (RHF) **self-consistent field** (SCF) formalism. 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.

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 **structural data (SD) file format inits V2000 version** 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 **standard structural 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™** or **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										
<code>-h</code>	Do not do a calculation, but print the list of options as a “help”										
<code>charge=<n></code>	Set the total molecular charge to <code><n></code>										
<code>hamil=<text></code>	Set the semiempirical Hamiltonian to be used. The Hamiltonian files provided with EMPIRE™ allow the following values for <code><text></code> : <table border="1" data-bbox="470 786 1407 1010"> <tbody> <tr> <td>AM1</td> <td>Use the AM1 [2] Hamiltonian (this is default if no Hamiltonian is defined)</td> </tr> <tr> <td>PM3</td> <td>Use the PM3 [3] Hamiltonian</td> </tr> <tr> <td>MNDO</td> <td>Use the MNDO [4] Hamiltonian</td> </tr> <tr> <td>AM1S</td> <td>Use the AM1* [5] Hamiltonian</td> </tr> <tr> <td>MNDOD</td> <td>Use the MNDO/d [6] Hamiltonian</td> </tr> </tbody> </table>	AM1	Use the AM1 [2] Hamiltonian (this is 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
AM1	Use the AM1 [2] Hamiltonian (this is 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										
<code>calc=<text></code>	Defines the type of calculation to be performed as follows: <table border="1" data-bbox="470 1055 1407 1167"> <tbody> <tr> <td>SPT</td> <td>Perform a single-point calculation (this is the default if no <code>calc=</code> option is defined)</td> </tr> <tr> <td>OPT</td> <td>Perform a full geometry optimization in Cartesian coordinates</td> </tr> </tbody> </table>	SPT	Perform a single-point calculation (this is the default if no <code>calc=</code> option is defined)	OPT	Perform a full geometry optimization in Cartesian coordinates						
SPT	Perform a single-point calculation (this is the default if no <code>calc=</code> option is defined)										
OPT	Perform a full geometry optimization in Cartesian coordinates										
<code>MaxTime=<n></code>	Define the maximum CPU-time (default = 24 hours). Default units are seconds, times can be given in minutes (e.g. 30m), hours (e.g. 24h) or days (e.g. 7d)										
<code>MaxOpt=<n></code>	Define the maximum number of optimization cycles (default=10,000)										
<code>MaxSCF=<n></code>	Define the maximum number of SCF cycles (default=1,000)										
<code>ConvergeR=<n></code>	Set the maximum off-diagonal CFC element for SCF convergence (default <code><n></code> = 10^{-4})										
<code>ConvergeE=<n></code>	Set the energy convergence limit for the SCF (default <code><n></code> = 10^{-4} kcal mol ⁻¹)										
<code>CheckP</code>	Use the convergence on the density matrix to test for SCF convergence. This option is not 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.										
<code>ConvergeP=<n></code>	Set the convergence limit if CheckP is used (default <code><n></code> = 10^{-6}). This option implies CheckP										
<code>ConvergeG=<n></code>	Set the gradient norm criterion for the geometry optimization (default <code><n></code> = 0.4 kcal mol ⁻¹ Å ⁻¹)										
<code>recalch</code>	Calculate the one-electron matrix on-the-fly. This option saves memory and may be necessary for very large calculations. Default is to save the one-electron matrix, which is faster.										
<code>fullRot</code>	Do full pseudorotations, rather than just virtual-occupied										
<code>MinBandGap=<n></code>	Use dynamic level shifting to ensure that the band gap used for the SCF does not fall below <code><n></code> (default is to use level shifting dynamically and only if the SCF energy rises)										



PDcutoff=<n>	Defines the cutoff value for CFC at which rotations are included in the pseudodiagonalization (default = 0.04 x the maximum CFC value)
EiCutoff=<n>	Defines the maximum value of CFC below which the Eigenvalues will not be recalculated in the SCF cycle. The default is <n>=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.
NoCenter	Do not center the molecule for output after performing the calculation (<i>i.e.</i> use the input coordinates directly)
Benchmark	Print extra timing information
Print	Print extra output
VWF	Write binary wavefunction file (default is not to write the file)
WFin=<filename>	Read the initial-guess wavefunction from the file <filename>



```

Core-core repulsion      =          7666.158888 eV
Ionization potential     =          10.851412 eV
Electron affinity        =           -3.573031 eV
Computation time        =           0.530 seconds

<> Dipole moment and components

Dipole      Total      X      Y      Z
Point-Chg.  0.002      0.002  0.000  0.000
Hybrid      0.001     -0.000 -0.000 -0.000
Sum         0.001      0.001  0.000 -0.000

<><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><><>

<> Time used :           0.546 s

```

Figure 5 The EMPIRE™ output for a PM3 geometry optimization of adamantane

The output is analogous to that for the single-point calculation except that, rather than the SCF iteration table, the course of the geometry optimization is printed. In this case, the gradient norm decreases monotonically to a value below the cutoff criterion of $0.4 \text{ kcal mol}^{-1} \text{ \AA}^{-1}$ in 14 cycles. The results are summarized as before.



H	0.139893	0.874006	2.431927	0.051001
H	2.336257	0.918539	-0.626700	0.050975
H	1.602278	1.933582	0.625213	0.050954
C	-0.000253	0.000031	-1.766927	-0.093913
C	0.000053	-0.000116	1.766984	-0.093892
H	-0.340326	-2.134024	-1.527902	0.060795

Figure 6 The EMPIRE™ archive file for a PM3 geometry optimization of adamantane

The calculation results are summarized followed by a table of the Cartesian coordinates of the optimized geometry and the Coulson net atomic charges in **.xyz** format.

5.2 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 2 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)

5.3 SDF output file

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

<pre>adamantane EMPIRE201202051309393D 1 1.00000 0.00000 0 VAMP input file created by Materials Studio 26 28 0 0 0 0 1 VZ000 -1.0218 -0.7172 1.2504 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.5088 -0.7205 1.2522 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1.0206 0.7219 1.2487 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.5137 1.4430 -0.0026 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1.0170 1.4441 -0.0032 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1.5317 0.0026 -0.0004 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1.0172 -1.4441 0.0029 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.5095 -0.7252 -1.2493 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1.0219 0.7171 -1.2505 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1.0211 -0.7218 -1.2482 C 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1.4051 -0.2191 2.1618 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1.4072 -1.7547 1.2796 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.8793 -1.2453 2.1633 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 2.1273 0.7361 1.2776 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.6807 1.2513 2.1597 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.8873 2.4934 -0.0045 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1.3993 1.9855 -0.8901 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1.4001 1.9895 0.8809 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -2.6465 0.0049 -0.0007 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.6759 -2.4973 0.0048 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 2.1239 -1.4778 0.0033 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.8803 -1.2532 -2.1584 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0.6836 1.2431 -2.1641 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 2.1286 0.7306 -1.2779 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1.4065 -1.7595 -1.2736 H 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 -1.4039 -0.2274 -2.1618 H 0 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 ><CALCULATION> [EMPIRE 2012] 9 Geometry optimized using EF ><HAMILTONIAN> [EMPIRE 2012] PM3 ><MOPACBASICS> [EMPIRE 2012] -34.582795 -10.851084 3.573031 -0.000227 0.000024 0.000014 ><CHARGE> [EMPIRE 2012] 0 ><ENERGY> [EMPIRE 2012] -9099.534965 7666.308123 -1433.226841 ><COULSON CHARGES> [EMPIRE 2012] -0.0939609684 -0.0727411291 -0.0939261180 -0.0727512305 -0.0940054870 -0.0727712777 -0.0939659309 -0.0727583817 -0.0939161115 -0.0939618762 0.0509598380 0.0509758459 0.0608048298 0.0509634965 0.0509713263 0.0608175389 0.0509335586 0.0509438811 0.0608215374 0.0509339016 0.0509643709 0.0608056725 0.0509608083 0.0509708477 0.0509748162 0.0509562412 ><TOTAL COULSON CHARGE> [EMPIRE 2012] -0.0000000000</pre>	<pre>><ORBITAL VECTORS> [EMPIRE 2012] RHF calculation 28 56 28 28 Nhomu, Nao, Nocc, Nunocc Nr. of atomic orbitals of each atom 4 4 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 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 -0.294 -0.057 -0.025 -0.037 -0.343 0.034 -0.061 -0.054 -0.329 0.041 0.041 -0.030 -0.063 0.010 -0.005 -0.126 -0.066 -0.009 -0.002 -0.063 -0.019 0.000 -0.020 -0.126 0.066 -0.005 0.001 -0.137 0.424 -0.031 0.031 -0.038 0.294 -0.032 -0.075 -0.045 0.329 0.076 0.005 -0.038 -0.147 -0.102 -0.154 -0.119 -0.160 -0.028 0.008 -0.064 -0.008 0.030 0.026 0.191 0.140 0.109 0.126 0.153 Further orbitals not shown here 5.334 -0.002 -0.113 -0.115 -0.148 0.039 -0.069 0.062 0.293 0.007 0.144 0.039 -0.172 -0.034 -0.300 -0.065 -0.011 -0.095 -0.181 0.237 0.009 -0.041 0.196 0.237 -0.012 0.095 -0.157 0.255 -0.007 0.035 -0.081 0.078 -0.286 0.002 0.160 0.038 0.163 -0.007 -0.116 -0.131 0.139 0.100 -0.105 -0.152 -0.106 0.120 0.131 -0.085 -0.097 0.159 0.102 0.080 -0.137 0.121 -0.117 -0.115 0.102 5.335 -0.020 -0.128 0.177 0.088 0.025 -0.116 -0.282 0.034 0.093 -0.148 -0.158 -0.201 0.047 -0.045 0.256 -0.165 0.007 -0.106 -0.103 0.143 -0.041 0.198 -0.168 -0.164 -0.007 0.196 0.110 -0.094 -0.031 -0.296 -0.027 0.079 0.020 0.100 -0.188 0.064 -0.093 -0.182 -0.067 -0.225 -0.136 0.097 -0.099 0.074 0.104 -0.183 0.091 -0.077 0.161 0.125 -0.139 0.121 0.119 -0.080 -0.057 -0.121 5.335 0.093 0.146 0.160 -0.201 0.045 0.275 -0.104 0.091 0.021 -0.186 0.133 -0.016 -0.029 0.042 -0.162 -0.258 -0.001 0.076 0.056 0.224 0.028 -0.135 0.097 -0.258 0.001 -0.118 -0.081 -0.147 -0.044 -0.006 0.296 0.086 -0.093 0.203 0.084 -0.199 -0.020 0.061 -0.221 -0.013 0.066 0.112 -0.174 0.127 -0.088 0.112 0.132 -0.133 -0.109 -0.084 0.085 0.172 -0.065 -0.114 -0.129 0.090 5.486</pre>
--	--



```

><DENSITY MATRIX ELEMENTS> [EMPIRE 2012]
1.1649095663 -0.0072555405 0.9550773709 -0.0051294175 0.0033642859 -0.025 -0.150 -0.105 0.183 0.089 -0.099 0.141 -0.244
1.0006974353 0.0089122218 -0.0056770871 0.0235811180 0.9732765959 -0.025 0.150 0.106 0.182 0.089 -0.100 -0.281 0.000
1.1777121871 0.0060953488 0.9579382967 -0.0086085470 -0.0049577219 -0.025 -0.149 0.211 -0.000 0.089 0.299 -0.001 -0.000
0.9614614629 0.0149847416 0.0086075416 -0.0122010956 0.9756291824 -0.025 0.149 -0.211 0.001 0.089 -0.099 0.141 0.244
1.1648977951 0.0072834877 0.9976326720 0.0051307025 -0.0118797255 -0.025 0.150 0.105 -0.183 -0.025 -0.149 -0.106 -0.183
0.9581306757 0.0088945547 -0.0202972573 0.0133139758 0.9732649752 -0.089 -0.089 0.168 -0.089 -0.088 0.168 -0.088 -0.089
1.1777058211 0.0061931503 0.9580042544 0.0172577505 0.0099645561 0.168 -0.088 -0.089 0.168 -0.088 -0.089 -0.088 -0.089
0.9826235384 -0.0000342903 -0.0000088556 -0.0000502811 0.9544176166
1.1649538144 -0.0071905464 0.9550367708 0.0102514719 -0.0066106369
0.9597528136 -0.0000091922 -0.0000004837 0.0001075910 1.0142620882
1.1776992224 -0.0183403425 0.9861540842 0.0000266234 -0.0000832152
0.9544593581 0.0000095708 0.0000156801 0.0000069503 0.9544586131
1.1649155754 0.0072453524 0.9977310715 -0.0102420839 0.0234759148
0.9809450812 0.0000173954 -0.0000333336 -0.0000549092 0.9503742029
1.1777029835 0.0060869876 0.9579489076 -0.0086615615 -0.0050219603
0.9615859158 -0.0149579422 -0.0086034066 0.0122403060 0.9755205748
1.1648899548 0.0072913045 0.9975837525 0.0051027701 -0.0118118041
0.9580497390 -0.0089018097 0.0203652109 -0.0132806653 0.9733926652
1.1649024652 -0.0072653766 0.9550677179 -0.0051523661 0.0033861316
1.0005345638 -0.0088808824 0.0056613159 -0.0236776560 0.9734571293

><CALCULATION STATUS> [EMPIRE 2012]
SUCCESS
$$$

><CHARGE ON HYDROGENS> [EMPIRE 2012]
0.0509598380 0.0509758459 0.0608048298 0.0509634965 0.0509713263
0.0608175389 0.0509335586 0.0509438811 0.0608215374 0.0509339016
0.0509643709 0.0608056725 0.0509608083 0.0509708477 0.0509748162
0.0509562412

```

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

All EMPIRE™-specific tags use the suffix (**EMPIRE 2012**). These are:

Table 3 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	<ol style="list-style-type: none">1. Number of the HOMO2. Number of orbitals3. Number of occupied orbitals4. Number of virtual orbitals	Format: 4(i4,1x), ' Nhomo , Nao , Nocc , Nunocc '
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	

6 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

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

Is 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.

7 SUPPORT

7.1 Contact

Questions regarding EMPIRE™ should be sent directly to:

support@ceposinsilico.com

7.2 CEPOS InSilico GmbH.

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

support@ceposinsilico.com

Tel. +49 (0)9131 970 4910

Fax. +49 (0)9131 970 4911

www.ceposinsilico.com/contact



8 LIST OF TABLES

Table 1	EMPIRE™ Program options	9
Table 2	Contents of binary wavefunction file	16
Table 3	EMPIRE™ specific tags	18



9 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	12
Figure 5	The EMPIRE™ output for a PM3 geometry optimization of adamantane	14
Figure 6	The EMPIRE™ archive file for a PM3 geometry optimization of adamantane	16
Figure 7	The EMPIRE™ output .sdf file for a PM3 geometry optimization of adamantane.....	18



10 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