Digichem: Computational Chemistry For Everyone

Supporting Information

Oliver S. Lee[a,b], *Malte C. Gather*[b,c*] *and Eli Zysman-Colman*[a*]

[a]Organic Semiconductor Centre, EaStCHEM School of Chemistry, University of St Andrews, St Andrews, UK, KY16 9ST

[b]Organic Semiconductor Centre, SUPA School of Physics and Astronomy, University of St Andrews, St Andrews, UK, KY16 9SS.

[c]Humboldt Centre for Nano- and Biophotonics, Department of Chemistry, University of Cologne, Greinstr. 4-6, 50939 Köln, Germany.

Contents

Benchmarking

Method

Optimisation calculations were performed on five different benzene oligomers (**Figure S1**), starting from benzene and including *para*-terphenyl (**3Phenyl**), *para*-pentaphenyl (**5Phenyl**), *para*-heptaphenyl (**7Phenyl**), and *para*-nonaphenyl (**9Phenyl**). The geometry of the larger oligomers take (on average) more optimisation cycles to converge to a minimum, and thus result in a larger log file (more program output). This permits the impact of the log file size to be evaluated in the parsing benchmarks.

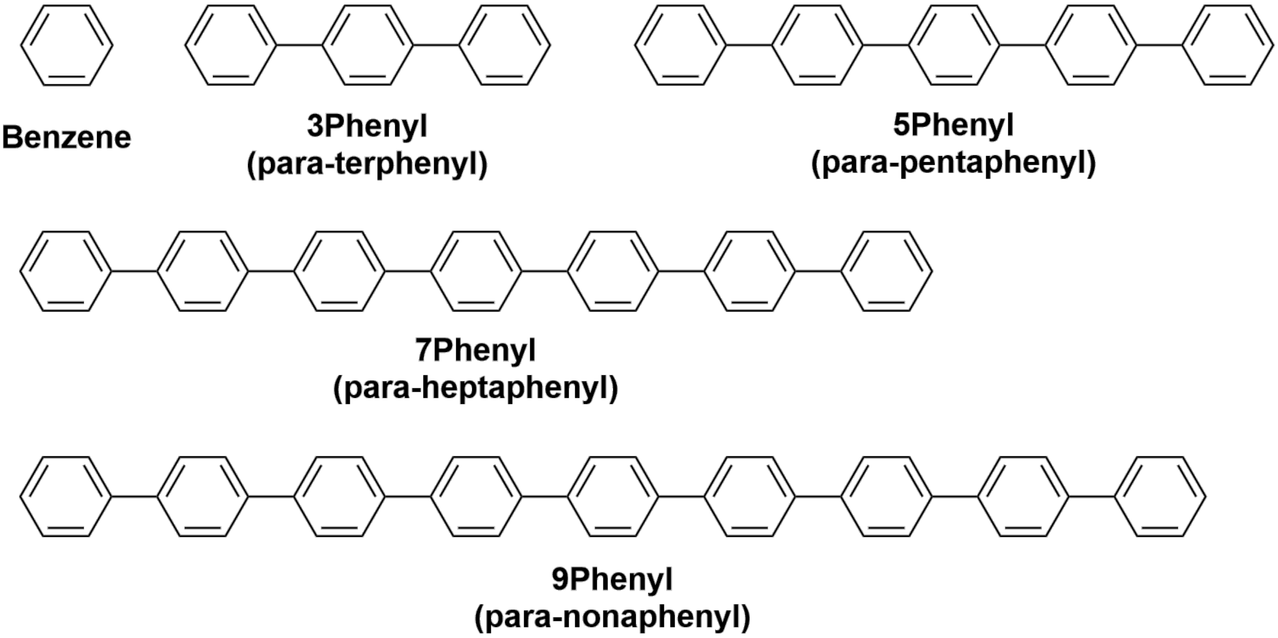


Figure S1. Names and structures of the five molecules benchmarked.

Identical calculations were carried out on all five molecules using Gaussian 16, revision C.01,[1] using the PBE0 functional[2–4] with the D3[5] version of Grimme's dispersion (including Beck-Johnson damping[6]) and the 6-31G** basis set.[7–9] The input files used for each calculation are available in the supporting information. All the calculations were performed on the same compute node of the University of St Andrews high-performance computing (HPC) cluster (Kennedy) and were submitted using the same version of Digichem (7.0.0-pre.3).[10] All 3D images were rendered using VMD 1.9.3[11] and the included version of Tachyon.[12] The number of threads used by Tachyon to render in parallel was left as the default, which uses the same number of threads as available CPU cores (32 in these tests) .For each calculation, Digichem automatically recorded the time taken to render each 3D image, as well as the total time spent rendering the entire PDF report (including all images together). The time taken to parse the log file and to generate just the PDF file (excluding image generation) were recorded separately using the main login node of the cluster.

We compared the rendering times for three different types of images for each molecule, to evaluate how the complexity of each render (*i.e.,* the number of surfaces in the scene) impacts the rendering time. The first render (structure) contains only the molecular geometry and does not include any isosurfaces. The second render (HOMO) additionally contains a single isosurface (the HOMO of each molecule), while the third render (HOMO/LUMO) contains two isosurfaces (the HOMO and LUMO of each molecule). The images were rendered in this order, which may be important when evaluating the impact of file caching. Each 'image' was rendered four times from four different angles, which is the same behaviour encountered in the PDF reports. The reported render time is the total time to render each of the four angles for each image. The four angles where the same for each molecule and each image type. The images rendered for Benzene are shown in **Figure S2**, and the complete set of rendered images is available in the supporting information.

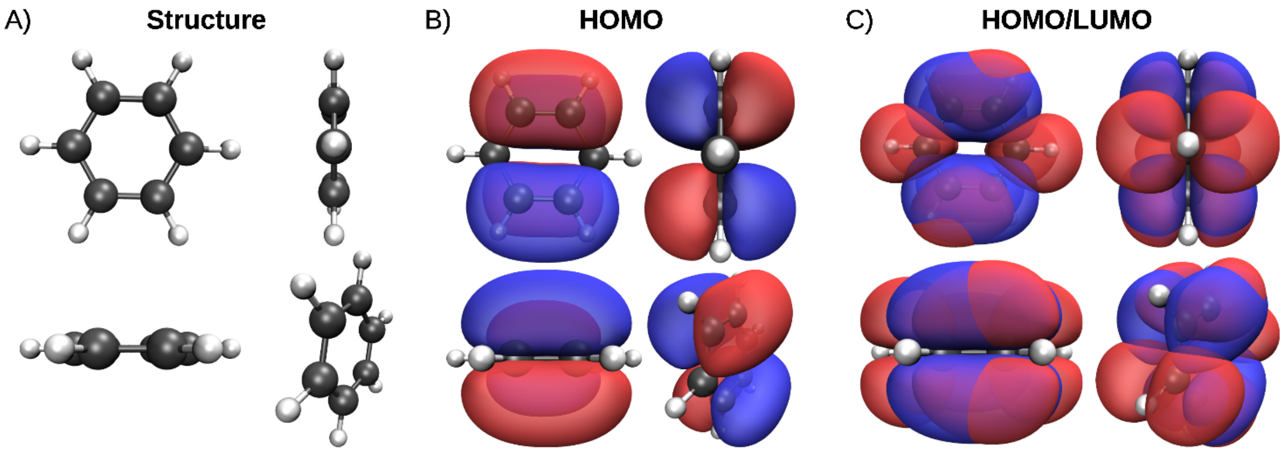


Figure S2. The three image types rendered for Benzene. a) the structure image (no isosurface). b) the HOMO image (1 isosurface). c) the HOMO/LUMO image (2 isosurfaces).

The Kennedy HPC system uses a network distributed file system (the General Parallel File System (GPFS)). This system permits each node in the cluster to seamlessly access the same file system but can be significantly slower than a traditional locally-mounted filesystem, especially when the system is under load (experiencing many simultaneous reads and/or writes). Additionally, GPFS utilises a form of local file caching that is beyond the control of the end-user. This file caching can result in inconsistent program timings because the first access to a file (which cannot make use of the cache as it does not exist yet) is often much slower than subsequent accesses to the same file. To combat this, each of the manually recorded operations (log file parsing and PDF report writing) was repeated three times. The first run (which typically does not make use of caching) produced inconsistent timings and was discarded. The second and third runs were averaged to give the reported times. All timings (included the discarded values) are reported for consistency in **Table S1**, **Table S2**, and **Table S3**.

Results

Log file parsing

As expected, the size of the log file produced by each calculation increases with the size of the molecule (**Figure S3a**), as the larger oligomers both require more optimisation steps and produce more output per optimisation step. The time taken by Digichem to parse each log file increased proportionally to the size of the log file (**Figure S3b**), likely indicating that the parsing process is IO-limited (*i.e.*, the bottleneck is reading from the log file rather than processing the resulting data), although this wasn't investigated directly. Even for the largest log file (**9Phenyl**, 29 MB, 420,000 lines), the parsing was complete in under 12 s.

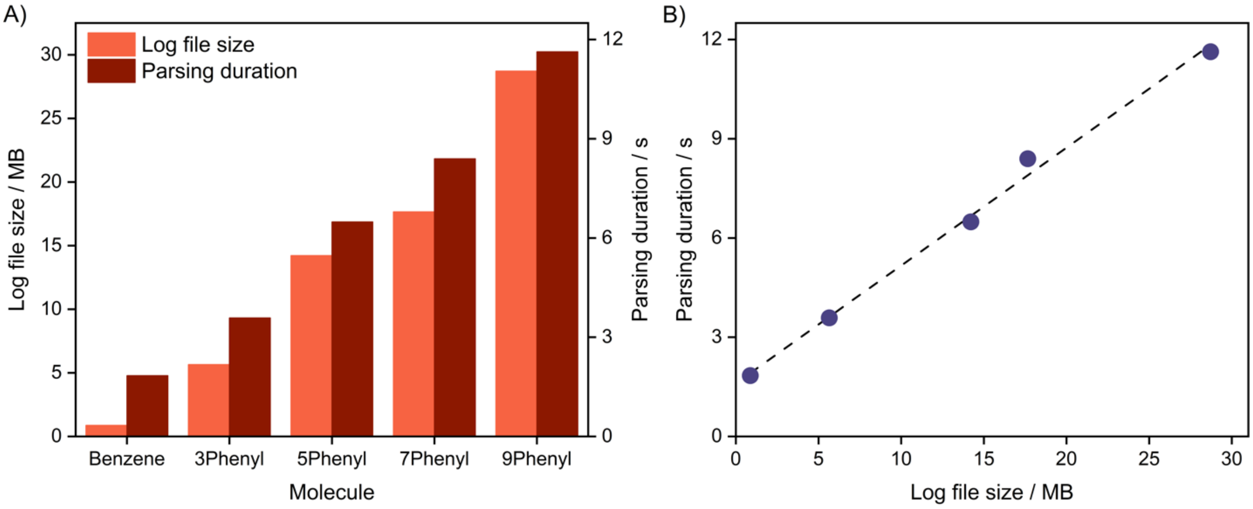


Figure S3. Log file size and parsing duration for each of the five molecules. The file size is reported in megabytes, where 1 MB = 1024^2 bytes.

Image rendering

There is little trend between the rending time of each image and the size of the molecule and/or the image complexity (**Figure S4a**). In two cases (**Benzene** and **7Phenyl**), the first image rendered (structure**)** took significantly longer than either the HOMO or HOMO/LUMO images, despite being the least complex (having no isosurfaces), again suggesting that the speed of the filesystem may by limiting the overall rendering speed. Except for **Benzene**, the render duration for the HOMO/LUMO image increased slightly with the larger oligomers, but no such trend can be observed for the other two types of images. Except for the outliers of the structure images for **Benzene** and **7Phenyl**, each image took between 10 – 30 s to render. Even for the slowest image, rendering was complete in under a minute (**Benzene**, structure).

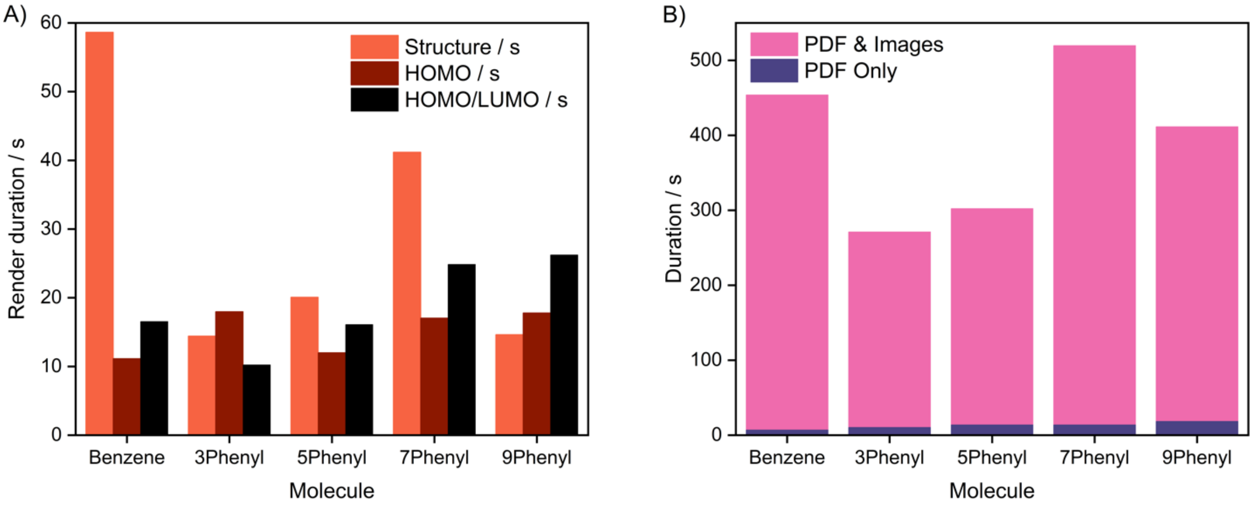


Figure S4. Duration of each image render (a) and PDF report generation (b).

Report generation

The report generation process consists of two broad steps. First, each of the required 3D images is rendered, followed by the PDF writing itself. The number of images included in each PDF is customizable by the user (depending on how many orbitals they wish to visualise etc.) and the type of calculation. Generally, excited-state calculations (which include natural-transition orbitals and/or difference density plots) include more images. In the calculations tested here, each report contained 6 different types of images (the total SCF electron density, the structure (no isosurfaces), the structure with the permanent dipole moment vector, the HOMO, the LUMO, and the HOMO/LUMO simultaneously). The time required to generate each report in its entirety (**Figure S4b,** pink bars) varied without much pattern from molecule to molecule and took on average 3.5 min (±1.7 min). Between **3Phenyl**, **5Phenyl**, and **9Phenyl** there is some suggestion that the larger oligomers resulted in longer report writing, but the pattern is not observable in **Benzene** or **7Phenyl**. In all cases, the time required to write the PDF file itself was only a small fraction of the overall process, taking 13 s (±4 s) on average (**Figure S4b,** purple bars). This indicates that the number of images in the report is the main determining factor for the overall report writing duration, rather than the log file size.

Table S1. Log file parsing duration.

Molecule	Log file size / MB	Parsing duration /s			
		First run (discounted)	Second run	Third run	Average
Benzene	0.9	33	2	2	2
3Phenyl	5.7	4	4	4	4

5Phenyl	14.2	7	7	6	6
7Phenyl	17.7	95	9	8	8
9Phenyl	28.7	114	12	12	12
Average					6
Deviation					4

Table S2. PDF file writing duration.

Molecule	**Log file size / MB**	**PDF Generation / s**			
		First run (discounted)	**Second run**	**Third run**	**Average**
Benzene	0.9	7	7	7	7
3Phenyl	5.7	13	11	10	10
5Phenyl	14.2	15	14	14	14
7Phenyl	17.7	77	14	14	14
9Phenyl	28.7	105	18	19	18
Average					13
Deviation					4

Table S3. Image rendering and total report generation duration.

Molecule	**Image rendering duration / s**			**Total report generation duration / s**
	Structure	**HOMO**	**HOMO/LUMO**	
Benzene	59	11	17	454
3Phenyl	14	18	10	271
5Phenyl	20	12	16	302
7Phenyl	41	17	25	520
9Phenyl	15	18	26	412
Average	30	15	19	392
Deviation	20	3	7	104

Molecular alignment procedures

Digichem supports three different molecular alignment procedures to re-orientate the molecular geometry, called **Symmetry** (**SYM**), **Average angle** (**AA**), and **Furthest atom pair** (**FAP**). **FAP** is *ab initio* and is entirely implemented in Digichem, while **SYM** and **AA** rely on the symmetry detection algorithm of the computational engine. The performance of these methods has been previously benchmarked elsewhere,[13,14] but in summary **SYM** typically performs the best so long as the computational engine implements a robust symmetry detection algorithm. For computational programs where symmetry is not routinely used (Orca, for example), **FAP** offers a reasonable fallback method. The performance of **FAP** and **AA** is normally similar, but **FAP** is more widely available as it does not rely on symmetry detection. The implementation of each method is briefly described below, see ref.[13] for more information.

In all cases, the longest identified molecular axis is rotated to coincide with the x-axis, while the second longest axis (confined to be at 90° to the long axis) is rotated to coincide with the y-axis.

Symmetry (SYM)

The molecule is aligned by the computational engine according to its detected symmetry. Exactly how this is done depends on the computational program, but typically the principal axis of symmetry (having the highest order) is rotated to coincide with one of the cartesian axes (often the y axis, although which does not matter). The molecule is then rotated by Digichem about its origin so that the greatest maximum difference in coordinates is in the x-axis, and the second most in the y-axis. The position of the origin is determined by the computational engine, and typically it is the molecule's centre of mass.

Average Angle (AA)

The molecule is first fully aligned using the **SYM** procedure before being translated so that the molecular origin is the centre of coordinates (**Equation S1**).

$$(\overline{x}, \overline{y}, \overline{z}) = \frac{1}{n} \sum_{j=1}^{n} (x_j, y_j, z_j)$$

Equation S1

Where (x_j, y_j, z_j) are the coordinates for atom j and n is the total number of atoms.

$$\overline{C} = \frac{1}{n} \sum_{j=1}^{n} cos(\theta_j)$$

Equation S2

$$\overline{S} = \frac{1}{n}\sum_{j=1}^{n} sin(\theta_j) \qquad \textbf{Equation S3}$$

$$\overline{\theta} = \begin{cases} tan^{-1}(\frac{\overline{S}}{\overline{C}}) & \overline{C} > 0, \overline{S} > 0 \\ tan^{-1}\left(\frac{\overline{S}}{\overline{C}}\right) + \pi & \overline{C} > 0 \\ tan^{-1}\left(\frac{\overline{S}}{\overline{C}}\right) + 2\pi & otherwise \end{cases} \qquad \textbf{Equation S4}$$

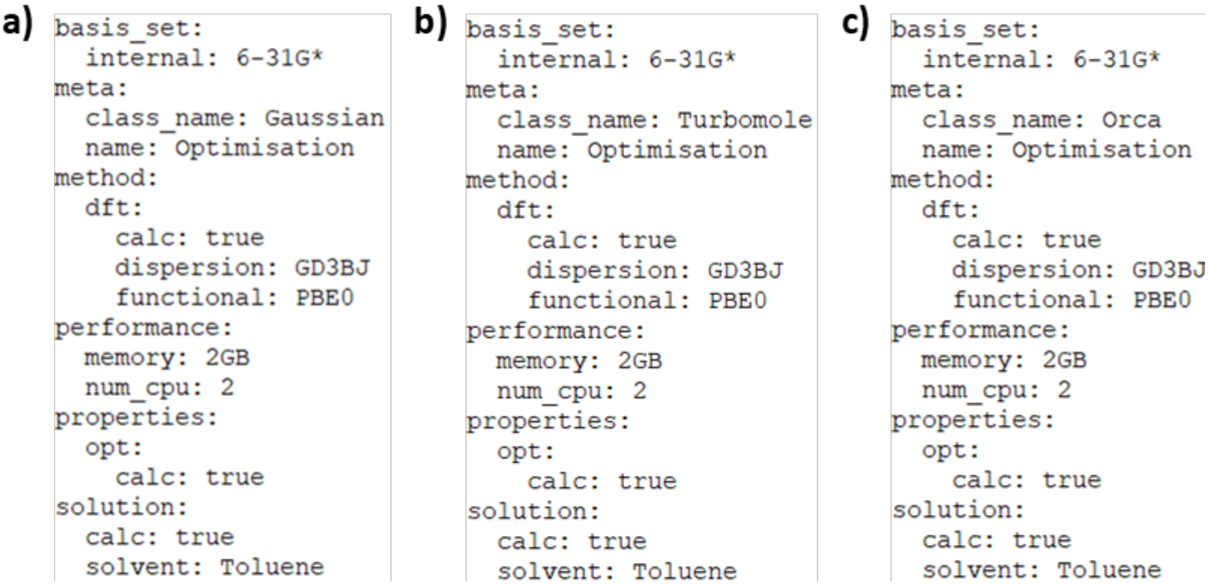
Where θ_j is the angle of the coordinates of atom j in a specific plane.

The molecule is then rotated in the XY-plane by the average angle (calculated according to Mardia and Jupp,[15] **Equation S4**) of each atom (in the same plane). This process is then repeated for the XZ- and YZ-planes, before the molecule is finally rotated about its origin so that the greatest maximum difference in coordinates is in the x-axis, and the second most in the y-axis.

Furthest Atom Paint (FAP)

Every unique pair of atoms in the molecule is iterated over to determine the pair with the greatest linear distance between them. The molecule is then translated so that the point equidistant between these two atoms becomes the origin, and rotated about this new origin so that the vector defined by this furthest atom pair coincides with the x-axis. The pair of atoms with the greatest linear distance in the newly defined YZ-plane is then found, and the molecule is rotated so the vector defined by these two points is parallel to the y-axis.

Additional graphs and figures



a)

```
basis_set:
  internal: 6-31G*
meta:
  class_name: Gaussian
  name: Optimisation
method:
  dft:
    calc: true
    dispersion: GD3BJ
    functional: PBE0
performance:
  memory: 2GB
  num_cpu: 2
properties:
  opt:
    calc: true
solution:
  calc: true
  solvent: Toluene
```

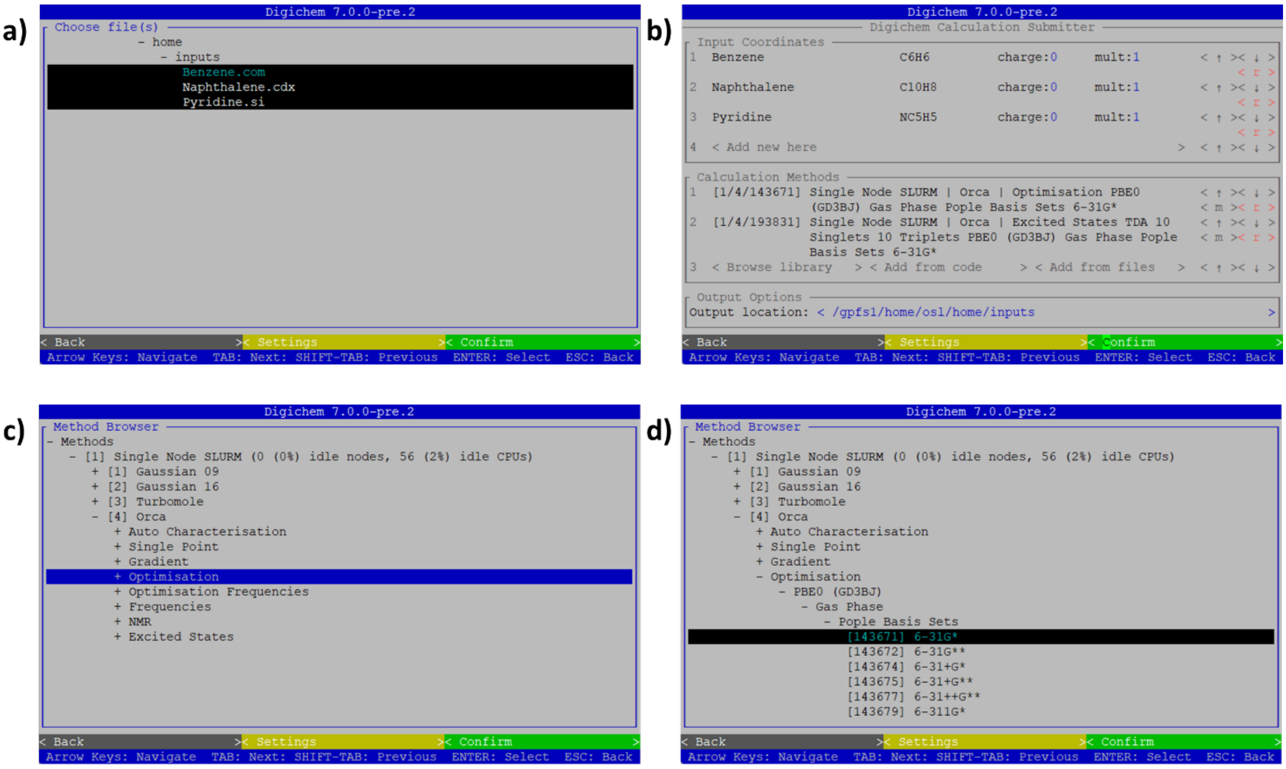
b)

```
basis_set:
  internal: 6-31G*
meta:
  class_name: Turbomole
  name: Optimisation
method:
  dft:
    calc: true
    dispersion: GD3BJ
    functional: PBE0
performance:
  memory: 2GB
  num_cpu: 2
properties:
  opt:
    calc: true
solution:
  calc: true
  solvent: Toluene
```

c)

```
basis_set:
  internal: 6-31G*
meta:
  class_name: Orca
  name: Optimisation
method:
  dft:
    calc: true
    dispersion: GD3BJ
    functional: PBE0
performance:
  memory: 2GB
  num_cpu: 2
properties:
  opt:
    calc: true
solution:
  calc: true
  solvent: Toluene
```

Figure S5. Comparison of the method file format for three equivalent calculations for the calculation engines a) Gaussian, b) Turbomole, and c) Orca.


a)
b)
c)
d)
< Back >< Settings >< Confirm >
Arrow Keys: Navigate TAB: Next: SHIFT-TAB: Previous ENTER: Select ESC: Back

a)

```
                         Digichem 7.0.0-pre.2
┌ Choose file(s) ──────────────────────────────────────────────────
           - home
             - inputs
                 Benzene.com
                 Naphthalene.cdx
                 Pyridine.si
```

b)

```
                         Digichem 7.0.0-pre.2
──────────────── Digichem Calculation Submitter ───────────────────
┌ Input Coordinates ───────────────────────────────────────────────
1  Benzene                 C6H6      charge:0   mult:1    < ↑ >< ↓ >
                                                               < r >
2  Naphthalene             C10H8     charge:0   mult:1    < ↑ >< ↓ >
                                                               < r >
3  Pyridine                NC5H5     charge:0   mult:1    < ↑ >< ↓ >
                                                               < r >
4  < Add new here                                      >  < ↑ >< ↓ >

┌ Calculation Methods ─────────────────────────────────────────────
1  [1/4/143671] Single Node SLURM | Orca | Optimisation PBE0  < ↑ >< ↓ >
                (GD3BJ) Gas Phase Pople Basis Sets 6-31G*     < m >< r >
2  [1/4/193831] Single Node SLURM | Orca | Excited States TDA 10  < ↑ >< ↓ >
                Singlets 10 Triplets PBE0 (GD3BJ) Gas Phase Pople < m >< r >
                Basis Sets 6-31G*
3  < Browse library  > < Add from code    > < Add from files  >  < ↑ >< ↓ >

┌ Output Options ──────────────────────────────────────────────────
Output location: < /gpfs1/home/osl/home/inputs                     >
```

c)

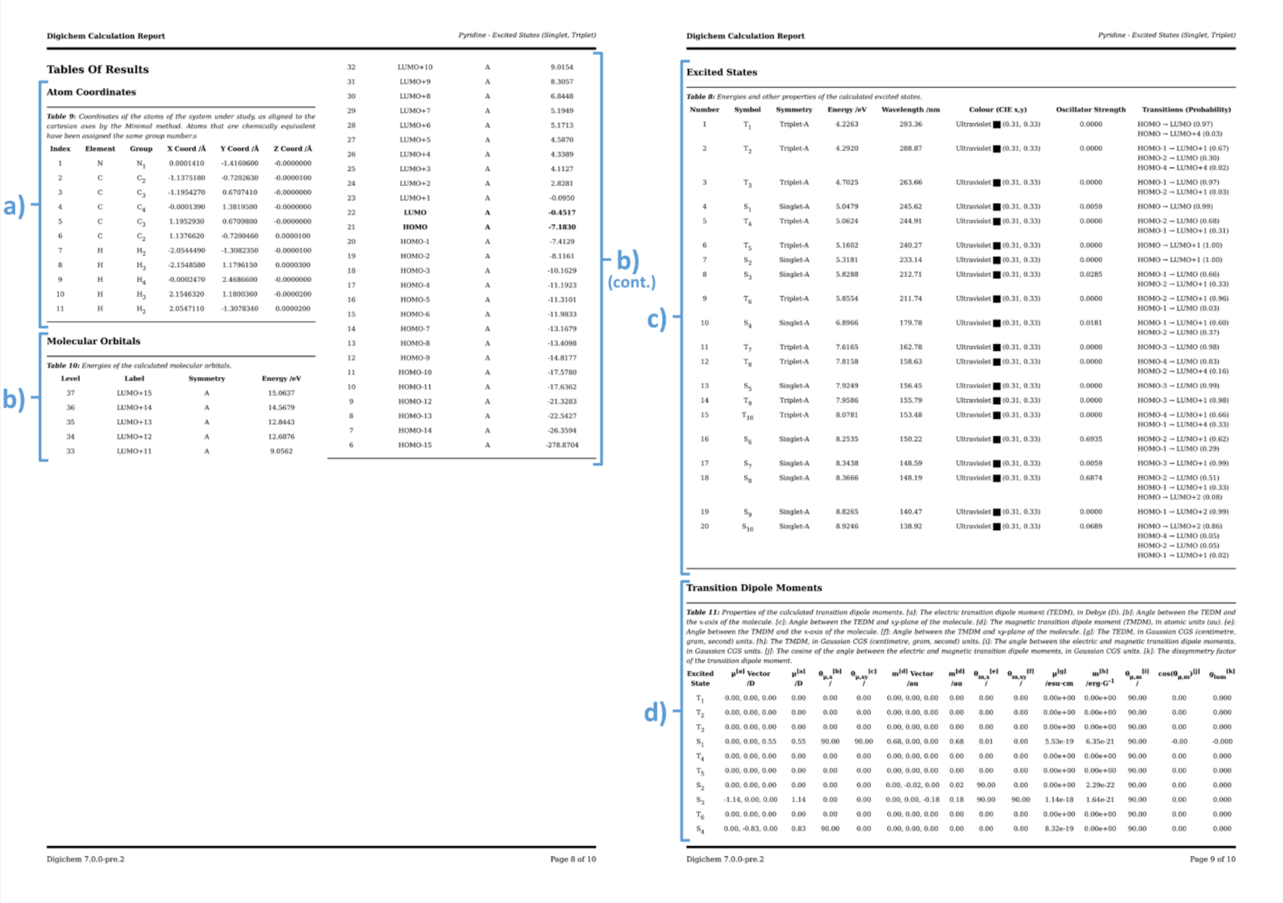
```
                         Digichem 7.0.0-pre.2
┌ Method Browser ──────────────────────────────────────────────────
- Methods
  - [1] Single Node SLURM (0 (0%) idle nodes, 56 (2%) idle CPUs)
    + [1] Gaussian 09
    + [2] Gaussian 16
    + [3] Turbomole
    - [4] Orca
      + Auto Characterisation
      + Single Point
      + Gradient
      + Optimisation
      + Optimisation Frequencies
      + Frequencies
      + NMR
      + Excited States
```

d)

```
                         Digichem 7.0.0-pre.2
┌ Method Browser ──────────────────────────────────────────────────
- Methods
  - [1] Single Node SLURM (0 (0%) idle nodes, 56 (2%) idle CPUs)
    + [1] Gaussian 09
    + [2] Gaussian 16
    + [3] Turbomole
    - [4] Orca
      + Auto Characterisation
      + Single Point
      + Gradient
      - Optimisation
        - PBE0 (GD3BJ)
          - Gas Phase
            - Pople Basis Sets
                [143671] 6-31G*
                [143672] 6-31G**
                [143674] 6-31+G*
                [143675] 6-31+G**
                [143677] 6-31++G**
                [143679] 6-311G*
```

Figure S6. Screenshots of the submission sub-module showing simultaneous set-up of three molecules (benzene, naphthalene, and pyridine) to be performed in parallel and two calculations (a geometry optimisation followed by TD-DFT excited-states calculation at the PBE0/6-31G** level of theory) to be performed in series. a) input coordinate file-picker. b) main submission interface.

c) internal method library, from which the calculation can be chosen. d) the same method library, but further expanded to show more options.



Digichem Calculation Report

Pyridine - Excited States (Singlet, Triplet)

Tables Of Results

Atom Coordinates

Table 9: Coordinates of the atoms of the system under study, as aligned to the cartesian axes by the Minimal method. Atoms that are chemically equivalent have been assigned the same group number.s

Index	Element	Group	X Coord /Å	Y Coord /Å	Z Coord /Å
1	N	N_1	0.0001410	-1.4160600	-0.0000000
2	C	C_2	-1.1375180	-0.7202630	-0.0000100
3	C	C_3	-1.1954270	0.6707410	-0.0000000
4	C	C_4	-0.0001390	1.3819500	-0.0000000
5	C	C_3	1.1952930	0.6709800	-0.0000000
6	C	C_2	1.1376620	-0.7200460	0.0000100
7	H	H_2	-2.0544490	-1.3082350	-0.0000100
8	H	H_3	-2.1548580	1.1796150	0.0000300
9	H	H_4	-0.0002470	2.4686600	-0.0000000
10	H	H_3	2.1546320	1.1800360	-0.0000200
11	H	H_2	2.0547110	-1.3078340	0.0000200

Molecular Orbitals

Table 10: Energies of the calculated molecular orbitals.

Level	Label	Symmetry	Energy /eV
37	LUMO+15	A	15.0637
36	LUMO+14	A	14.5679
35	LUMO+13	A	12.8443
34	LUMO+12	A	12.6876
33	LUMO+11	A	9.0562
32	LUMO+10	A	9.0154
31	LUMO+9	A	8.3057
30	LUMO+8	A	6.8448
29	LUMO+7	A	5.1949
28	LUMO+6	A	5.1713
27	LUMO+5	A	4.5870
26	LUMO+4	A	4.3389
25	LUMO+3	A	4.1127
24	LUMO+2	A	2.8281
23	LUMO+1	A	-0.0950
22	**LUMO**	**A**	**-0.4517**
21	**HOMO**	**A**	**-7.1830**
20	HOMO-1	A	-7.4129
19	HOMO-2	A	-8.1161
18	HOMO-3	A	-10.1629
17	HOMO-4	A	-11.1923
16	HOMO-5	A	-11.3101
15	HOMO-6	A	-11.9833
14	HOMO-7	A	-13.1679
13	HOMO-8	A	-13.4098
12	HOMO-9	A	-14.8177
11	HOMO-10	A	-17.5780
10	HOMO-11	A	-17.6362
9	HOMO-12	A	-21.3283
8	HOMO-13	A	-22.5427
7	HOMO-14	A	-26.3594
6	HOMO-15	A	-278.8704

Digichem 7.0.0-pre.2

Page 8 of 10

Digichem Calculation Report

Pyridine - Excited States (Singlet, Triplet)

Excited States

Table 8: Energies and other properties of the calculated excited states.

Number	Symbol	Symmetry	Energy /eV	Wavelength /nm	Colour (CIE x,y)	Oscillator Strength	Transitions (Probability)
1	T_1	Triplet-A	4.2263	293.36	Ultraviolet ■ (0.31, 0.33)	0.0000	HOMO → LUMO (0.97) HOMO → LUMO+4 (0.03)
2	T_2	Triplet-A	4.2920	288.87	Ultraviolet ■ (0.31, 0.33)	0.0000	HOMO-1 → LUMO+1 (0.67) HOMO-2 → LUMO (0.30) HOMO-4 → LUMO+4 (0.02)
3	T_3	Triplet-A	4.7025	263.66	Ultraviolet ■ (0.31, 0.33)	0.0000	HOMO-1 → LUMO (0.97) HOMO-2 → LUMO+1 (0.03)
4	S_1	Singlet-A	5.0479	245.62	Ultraviolet ■ (0.31, 0.33)	0.0059	HOMO → LUMO (0.99)
5	T_4	Triplet-A	5.0624	244.91	Ultraviolet ■ (0.31, 0.33)	0.0000	HOMO-2 → LUMO (0.68) HOMO-1 → LUMO+1 (0.31)
6	T_5	Triplet-A	5.1602	240.27	Ultraviolet ■ (0.31, 0.33)	0.0000	HOMO → LUMO+1 (1.00)
7	S_2	Singlet-A	5.3181	233.14	Ultraviolet ■ (0.31, 0.33)	0.0000	HOMO → LUMO+1 (1.00)
8	S_3	Singlet-A	5.8288	212.71	Ultraviolet ■ (0.31, 0.33)	0.0285	HOMO-1 → LUMO (0.66) HOMO-2 → LUMO+1 (0.33)
9	T_6	Triplet-A	5.8554	211.74	Ultraviolet ■ (0.31, 0.33)	0.0000	HOMO-2 → LUMO+1 (0.96) HOMO-1 → LUMO (0.03)
10	S_4	Singlet-A	6.8966	179.78	Ultraviolet ■ (0.31, 0.33)	0.0181	HOMO-1 → LUMO+1 (0.60) HOMO-2 → LUMO (0.37)
11	T_7	Triplet-A	7.6165	162.78	Ultraviolet ■ (0.31, 0.33)	0.0000	HOMO-3 → LUMO (0.98)
12	T_8	Triplet-A	7.8158	158.63	Ultraviolet ■ (0.31, 0.33)	0.0000	HOMO-4 → LUMO (0.83) HOMO-2 → LUMO+4 (0.16)
13	S_5	Singlet-A	7.9249	156.45	Ultraviolet ■ (0.31, 0.33)	0.0000	HOMO-3 → LUMO (0.99)
14	T_9	Triplet-A	7.9586	155.79	Ultraviolet ■ (0.31, 0.33)	0.0000	HOMO-3 → LUMO+1 (0.98)
15	T_{10}	Triplet-A	8.0781	153.48	Ultraviolet ■ (0.31, 0.33)	0.0000	HOMO-4 → LUMO+1 (0.66) HOMO-1 → LUMO+4 (0.33)
16	S_6	Singlet-A	8.2535	150.22	Ultraviolet ■ (0.31, 0.33)	0.6935	HOMO-2 → LUMO+1 (0.62) HOMO-1 → LUMO (0.29)
17	S_7	Singlet-A	8.3438	148.59	Ultraviolet ■ (0.31, 0.33)	0.0059	HOMO-3 → LUMO+1 (0.99)
18	S_8	Singlet-A	8.3666	148.19	Ultraviolet ■ (0.31, 0.33)	0.6874	HOMO-2 → LUMO (0.51) HOMO-1 → LUMO+1 (0.33) HOMO → LUMO+2 (0.08)
19	S_9	Singlet-A	8.8265	140.47	Ultraviolet ■ (0.31, 0.33)	0.0000	HOMO-1 → LUMO+2 (0.99)
20	S_{10}	Singlet-A	8.9246	138.92	Ultraviolet ■ (0.31, 0.33)	0.0689	HOMO → LUMO+2 (0.86) HOMO-4 → LUMO (0.05) HOMO-2 → LUMO (0.05) HOMO-1 → LUMO+1 (0.02)

Transition Dipole Moments

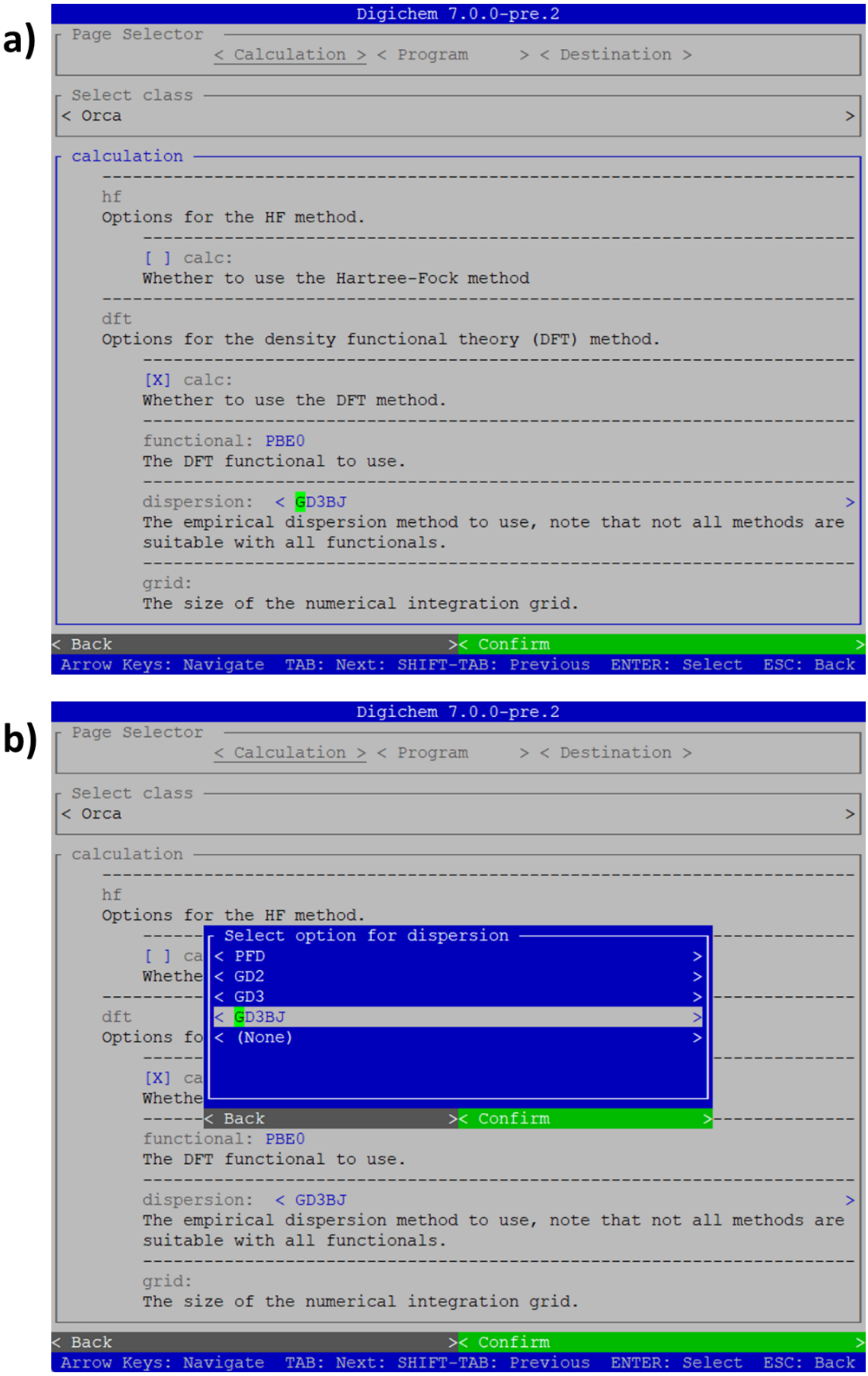
Table 11: Properties of the calculated transition dipole moments. [a]: The electric transition dipole moment (TEDM), in Debye (D). [b]: Angle between the TEDM and the x-axis of the molecule. [c]: Angle between the TEDM and xy-plane of the molecule. [d]: The magnetic transition dipole moment (TMDM), in atomic units (au). [e]: Angle between the TMDM and the x-axis of the molecule. [f]: Angle between the TMDM and xy-plane of the molecule. [g]: The TEDM, in Gaussian CGS (centimetre, gram, second) units. [h]: The TMDM, in Gaussian CGS (centimetre, gram, second) units. [i]: The angle between the electric and magnetic transition dipole moments, in Gaussian CGS units. [j]: The cosine of the angle between the electric and magnetic transition dipole moments, in Gaussian CGS units. [k]: The dissymmetry factor of the transition dipole moment.

Excited State	$\mu^{[a]}$ Vector /D	$\mu^{[a]}$ /D	$\theta_{\mu,x}{}^{[b]}$ /	$\theta_{\mu,xy}{}^{[c]}$ /	$m^{[d]}$ Vector /au	$m^{[d]}$ /au	$\theta_{m,x}{}^{[e]}$ /	$\theta_{m,xy}{}^{[f]}$ /	$\mu^{[g]}$ /esu·cm	$m^{[h]}$ /erg·G^{-1}	$\theta_{\mu,m}{}^{[i]}$ /	$\cos(\theta_{\mu,m})^{[j]}$	$g_{lum}{}^{[k]}$
T_1	0.00, 0.00, 0.00	0.00	0.00	0.00	0.00, 0.00, 0.00	0.00	0.00	0.00	0.00e+00	0.00e+00	90.00	0.00	0.000
T_2	0.00, 0.00, 0.00	0.00	0.00	0.00	0.00, 0.00, 0.00	0.00	0.00	0.00	0.00e+00	0.00e+00	90.00	0.00	0.000
T_3	0.00, 0.00, 0.00	0.00	0.00	0.00	0.00, 0.00, 0.00	0.00	0.00	0.00	0.00e+00	0.00e+00	90.00	0.00	0.000
S_1	0.00, 0.00, 0.55	0.55	90.00	90.00	0.68, 0.00, 0.00	0.68	0.01	0.00	5.53e-19	6.35e-21	90.00	-0.00	-0.000
T_4	0.00, 0.00, 0.00	0.00	0.00	0.00	0.00, 0.00, 0.00	0.00	0.00	0.00	0.00e+00	0.00e+00	90.00	0.00	0.000
T_5	0.00, 0.00, 0.00	0.00	0.00	0.00	0.00, 0.00, 0.00	0.00	0.00	0.00	0.00e+00	0.00e+00	90.00	0.00	0.000
S_2	0.00, 0.00, 0.00	0.00	0.00	0.00	0.00, -0.02, 0.00	0.02	90.00	0.00	0.00e+00	2.29e-22	90.00	0.00	0.000
S_3	-1.14, 0.00, 0.00	1.14	0.00	0.00	0.00, 0.00, -0.18	0.18	90.00	90.00	1.14e-18	1.64e-21	90.00	0.00	0.000
T_6	0.00, 0.00, 0.00	0.00	0.00	0.00	0.00, 0.00, 0.00	0.00	0.00	0.00	0.00e+00	0.00e+00	90.00	0.00	0.000
S_4	0.00, -0.83, 0.00	0.83	90.00	0.00	0.00, 0.00, 0.00	0.00	0.00	0.00	8.32e-19	0.00e+00	90.00	0.00	0.000

Digichem 7.0.0-pre.2

Page 9 of 10

Figure S7. Excerpts from an example calculation report generated by Digichem, demonstrating tabulated data. The excited states of pyridine at the PBE0/6-31G* level of theory using the Tamm-Dancoff approximation (TDA) to time-dependent DFT (TD-DFT) were calculated. a) table of molecular geometry, b) table of selected molecular orbitals, c) table of electronic excited states, d) table of transition dipole moments.



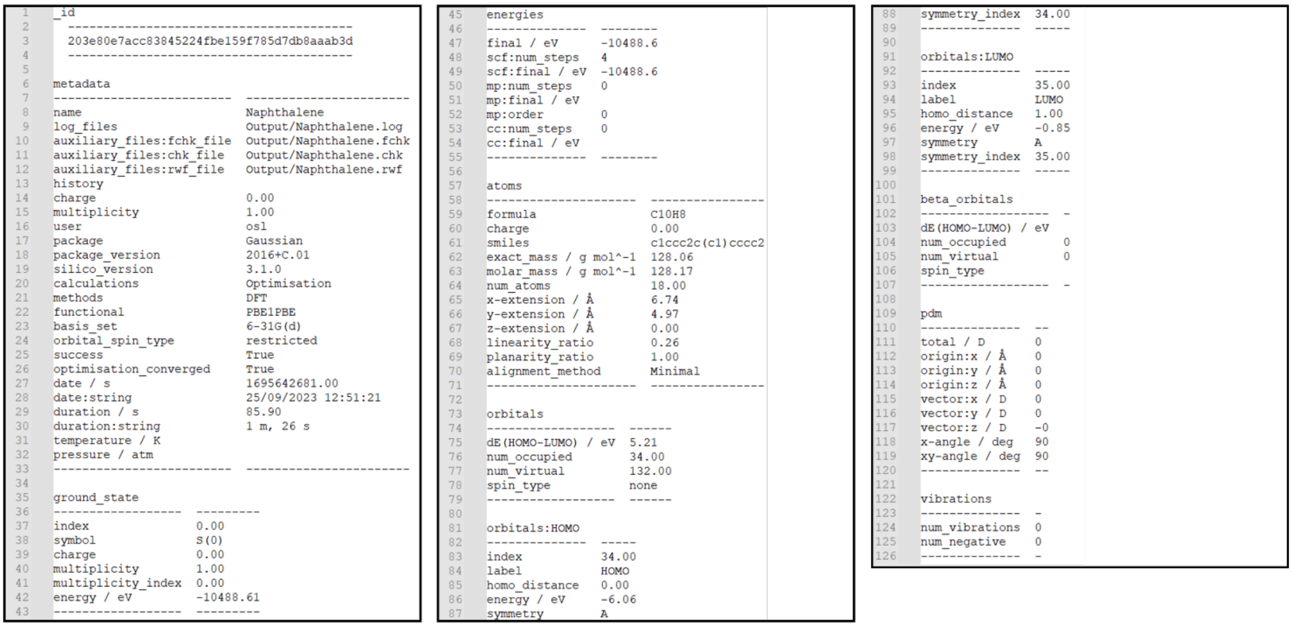
a)

```
                         Digichem 7.0.0-pre.2
┌ Page Selector ─────────────────────────────────────────────────────────┐
│                 < Calculation > < Program      > < Destination >        │
└─────────────────────────────────────────────────────────────────────────┘
┌ Select class ───────────────────────────────────────────────────────────┐
│< Orca                                                                  >│
└─────────────────────────────────────────────────────────────────────────┘
┌ calculation ────────────────────────────────────────────────────────────┐
│    ---------------------------------------------------------------------│
│    hf                                                                   │
│    Options for the HF method.                                           │
│        -----------------------------------------------------------------│
│        [ ] calc:                                                        │
│        Whether to use the Hartree-Fock method                           │
│    ---------------------------------------------------------------------│
│    dft                                                                  │
│    Options for the density functional theory (DFT) method.              │
│        -----------------------------------------------------------------│
│        [X] calc:                                                        │
│        Whether to use the DFT method.                                   │
│        -----------------------------------------------------------------│
│        functional: PBE0                                                 │
│        The DFT functional to use.                                       │
│        -----------------------------------------------------------------│
│        dispersion:  < GD3BJ                                            >│
│        The empirical dispersion method to use, note that not all methods are
│        suitable with all functionals.                                   │
│        -----------------------------------------------------------------│
│        grid:                                                            │
│        The size of the numerical integration grid.                      │
└─────────────────────────────────────────────────────────────────────────┘
< Back                                   >< Confirm                        >
 Arrow Keys: Navigate  TAB: Next: SHIFT-TAB: Previous  ENTER: Select  ESC: Back
```

b)

```
                         Digichem 7.0.0-pre.2
┌ Page Selector ─────────────────────────────────────────────────────────┐
│                 < Calculation > < Program      > < Destination >        │
└─────────────────────────────────────────────────────────────────────────┘
┌ Select class ───────────────────────────────────────────────────────────┐
│< Orca                                                                  >│
└─────────────────────────────────────────────────────────────────────────┘
┌ calculation ────────────────────────────────────────────────────────────┐
│    ---------------------------------------------------------------------│
│    hf                                                                   │
│    Options for the HF method.                                           │
│        ------┌ Select option for dispersion ─────────────┐--------------│
│        [ ] ca│< PFD                                     >│              │
│        Whethe│< GD2                                     >│              │
│    ----------│< GD3                                     >│--------------│
│    dft       │< GD3BJ                                   >│              │
│    Options fo│< (None)                                  >│              │
│        ------│                                          │--------------│
│        [X] ca│                                          │              │
│        Whethe└──────────────────────────────────────────┘              │
│        ------< Back                   >< Confirm                >--------------│
│        functional: PBE0                                                 │
│        The DFT functional to use.                                       │
│        -----------------------------------------------------------------│
│        dispersion:  < GD3BJ                                            >│
│        The empirical dispersion method to use, note that not all methods are
│        suitable with all functionals.                                   │
│        -----------------------------------------------------------------│
│        grid:                                                            │
│        The size of the numerical integration grid.                      │
└─────────────────────────────────────────────────────────────────────────┘
< Back                                   >< Confirm                        >
 Arrow Keys: Navigate  TAB: Next: SHIFT-TAB: Previous  ENTER: Select  ESC: Back
```

Figure S8. Screenshots of the calculation method editor interface. a) general overview, showing example options for the calculation level of theory. b) example validation for the DFT dispersion correction option.

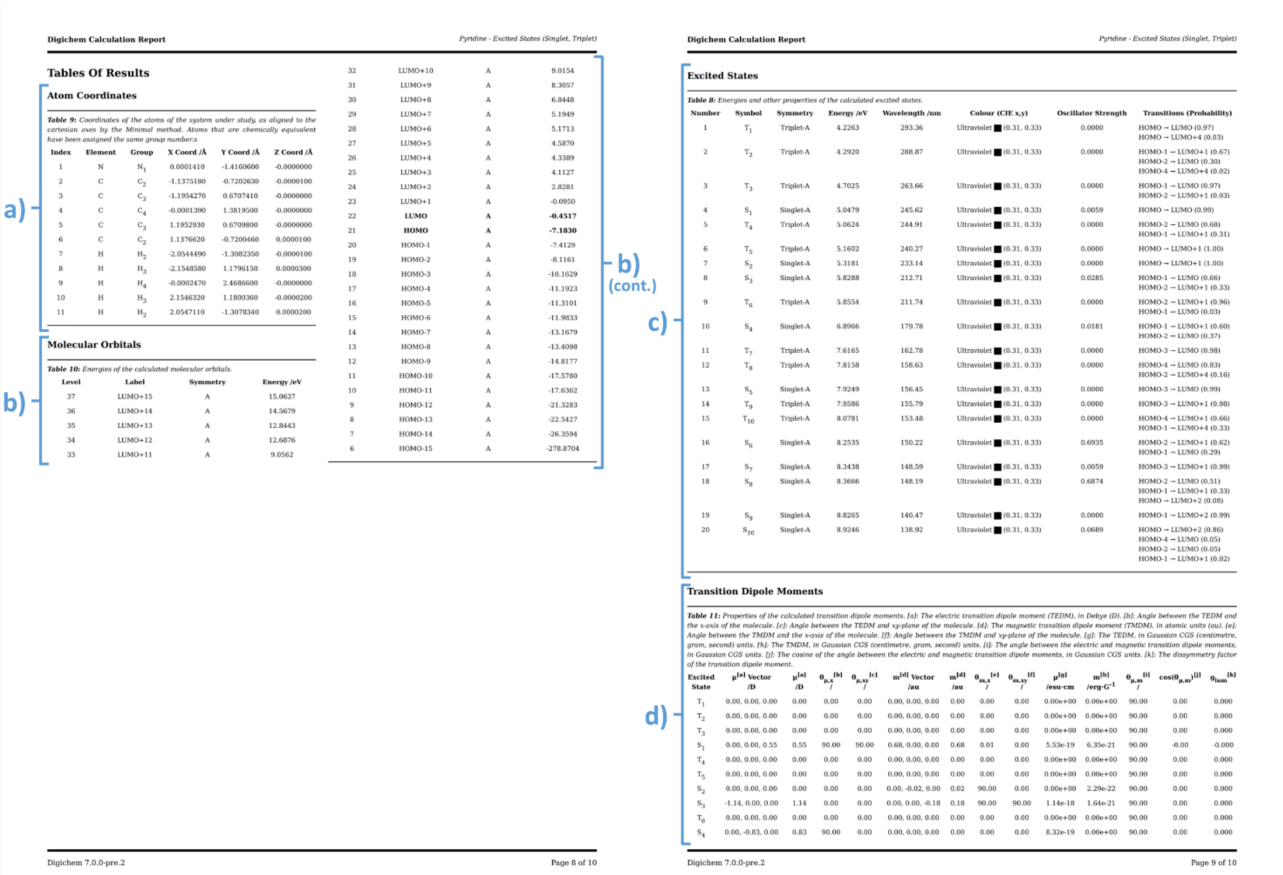


```
_id
  ----------------------------------------
  203e80e7acc83845224fbe159f785d7db8aaab3d
  ----------------------------------------

metadata
-------------------------   -----------------------
name                        Naphthalene
log_files                   Output/Naphthalene.log
auxiliary_files:fchk_file   Output/Naphthalene.fchk
auxiliary_files:chk_file    Output/Naphthalene.chk
auxiliary_files:rwf_file    Output/Naphthalene.rwf
history
charge                      0.00
multiplicity                1.00
user                        osl
package                     Gaussian
package_version             2016+C.01
silico_version              3.1.0
calculations                Optimisation
methods                     DFT
functional                  PBE1PBE
basis_set                   6-31G(d)
orbital_spin_type           restricted
success                     True
optimisation_converged      True
date / s                    1695642681.00
date:string                 25/09/2023 12:51:21
duration / s                85.90
duration:string             1 m, 26 s
temperature / K
pressure / atm
-------------------------   -----------------------

ground_state
------------------   ---------
index                0.00
symbol               S(0)
charge               0.00
multiplicity         1.00
multiplicity_index   0.00
energy / eV          -10488.61
------------------   ---------
```

```
energies
--------------   --------
final / eV       -10488.6
scf:num_steps    4
scf:final / eV   -10488.6
mp:num_steps     0
mp:final / eV
mp:order         0
cc:num_steps     0
cc:final / eV
--------------   --------

atoms
---------------------   ----------------
formula                 C10H8
charge                  0.00
smiles                  c1ccc2c(c1)cccc2
exact_mass / g mol^-1   128.06
molar_mass / g mol^-1   128.17
num_atoms               18.00
x-extension / Å         6.74
y-extension / Å         4.97
z-extension / Å         0.00
linearity_ratio         0.26
planarity_ratio         1.00
alignment_method        Minimal
---------------------   ----------------

orbitals
------------------   ------
dE(HOMO-LUMO) / eV   5.21
num_occupied         34.00
num_virtual          132.00
spin_type            none
------------------   ------

orbitals:HOMO
--------------   -----
index            34.00
label            HOMO
homo_distance    0.00
energy / eV      -6.06
symmetry         A
```

```
symmetry_index   34.00
--------------   -----

orbitals:LUMO
--------------   -----
index            35.00
label            LUMO
homo_distance    1.00
energy / eV      -0.85
symmetry         A
symmetry_index   35.00
--------------   -----

beta_orbitals
------------------   -
dE(HOMO-LUMO) / eV
num_occupied         0
num_virtual          0
spin_type
------------------   -

pdm
--------------   --
total / D        0
origin:x / Å     0
origin:y / Å     0
origin:z / Å     0
vector:x / D     0
vector:y / D     0
vector:z / D     -0
x-angle / deg    90
xy-angle / deg   90
--------------   --

vibrations
--------------   -
num_vibrations   0
num_negative     0
--------------   -
```

Figure S9. Excerpts from an example result summary output file written by Digichem.



Digichem Calculation Report

Pyridine - Excited States (Singlet, Triplet)

Tables Of Results

Atom Coordinates

Table 9: *Coordinates of the atoms of the system under study, as aligned to the cartesian axes by the Minimal method. Atoms that are chemically equivalent have been assigned the same group number.s*

Index	Element	Group	X Coord /Å	Y Coord /Å	Z Coord /Å
1	N	N_1	0.0001410	-1.4160600	-0.0000000
2	C	C_2	-1.1375180	-0.7202630	-0.0000100
3	C	C_3	-1.1954270	0.6707410	-0.0000000
4	C	C_4	-0.0001390	1.3819500	-0.0000000
5	C	C_3	1.1952930	0.6709800	-0.0000000
6	C	C_2	1.1376620	-0.7200460	0.0000100
7	H	H_2	-2.0544490	-1.3082350	-0.0000100
8	H	H_3	-2.1548580	1.1796150	0.0000300
9	H	H_4	-0.0002470	2.4686600	-0.0000000
10	H	H_3	2.1546320	1.1800360	-0.0000200
11	H	H_2	2.0547110	-1.3078340	0.0000200

Molecular Orbitals

Table 10: *Energies of the calculated molecular orbitals.*

Level	Label	Symmetry	Energy /eV
37	LUMO+15	A	15.0637
36	LUMO+14	A	14.5679
35	LUMO+13	A	12.8443
34	LUMO+12	A	12.6876
33	LUMO+11	A	9.0562
32	LUMO+10	A	9.0154
31	LUMO+9	A	8.3057
30	LUMO+8	A	6.8448
29	LUMO+7	A	5.1949
28	LUMO+6	A	5.1713
27	LUMO+5	A	4.5870
26	LUMO+4	A	4.3389
25	LUMO+3	A	4.1127
24	LUMO+2	A	2.8281
23	LUMO+1	A	-0.0950
22	**LUMO**	**A**	**-0.4517**
21	**HOMO**	**A**	**-7.1830**
20	HOMO-1	A	-7.4129
19	HOMO-2	A	-8.1161
18	HOMO-3	A	-10.1629
17	HOMO-4	A	-11.1923
16	HOMO-5	A	-11.3101
15	HOMO-6	A	-11.9833
14	HOMO-7	A	-13.1679
13	HOMO-8	A	-13.4098
12	HOMO-9	A	-14.8177
11	HOMO-10	A	-17.5780
10	HOMO-11	A	-17.6362
9	HOMO-12	A	-21.3283
8	HOMO-13	A	-22.5427
7	HOMO-14	A	-26.3594
6	HOMO-15	A	-278.8704

Digichem 7.0.0-pre.2

Page 8 of 10

Digichem Calculation Report

Pyridine - Excited States (Singlet, Triplet)

Excited States

Table 8: *Energies and other properties of the calculated excited states.*

Number	Symbol	Symmetry	Energy /eV	Wavelength /nm	Colour (CIE x,y)	Oscillator Strength	Transitions (Probability)
1	T_1	Triplet-A	4.2263	293.36	Ultraviolet ■ (0.31, 0.33)	0.0000	HOMO → LUMO (0.97) HOMO → LUMO+4 (0.03)
2	T_2	Triplet-A	4.2920	288.87	Ultraviolet ■ (0.31, 0.33)	0.0000	HOMO-1 → LUMO+1 (0.67) HOMO-2 → LUMO (0.30) HOMO-4 → LUMO+4 (0.02)
3	T_3	Triplet-A	4.7025	263.66	Ultraviolet ■ (0.31, 0.33)	0.0000	HOMO-1 → LUMO (0.97) HOMO-2 → LUMO+1 (0.03)
4	S_1	Singlet-A	5.0479	245.62	Ultraviolet ■ (0.31, 0.33)	0.0059	HOMO → LUMO (0.99)
5	T_4	Triplet-A	5.0624	244.91	Ultraviolet ■ (0.31, 0.33)	0.0000	HOMO-2 → LUMO (0.68) HOMO-1 → LUMO+1 (0.31)
6	T_5	Triplet-A	5.1602	240.27	Ultraviolet ■ (0.31, 0.33)	0.0000	HOMO → LUMO+1 (1.00)
7	S_2	Singlet-A	5.3181	233.14	Ultraviolet ■ (0.31, 0.33)	0.0000	HOMO → LUMO+1 (1.00)
8	S_3	Singlet-A	5.8288	212.71	Ultraviolet ■ (0.31, 0.33)	0.0285	HOMO-1 → LUMO (0.66) HOMO-2 → LUMO+1 (0.33)
9	T_6	Triplet-A	5.8554	211.74	Ultraviolet ■ (0.31, 0.33)	0.0000	HOMO-2 → LUMO+1 (0.96) HOMO-1 → LUMO (0.03)
10	S_4	Singlet-A	6.8966	179.78	Ultraviolet ■ (0.31, 0.33)	0.0181	HOMO-1 → LUMO+1 (0.60) HOMO-2 → LUMO (0.37)
11	T_7	Triplet-A	7.6165	162.78	Ultraviolet ■ (0.31, 0.33)	0.0000	HOMO-3 → LUMO (0.98)
12	T_8	Triplet-A	7.8158	158.63	Ultraviolet ■ (0.31, 0.33)	0.0000	HOMO-4 → LUMO (0.83) HOMO-2 → LUMO+4 (0.16)
13	S_5	Singlet-A	7.9249	156.45	Ultraviolet ■ (0.31, 0.33)	0.0000	HOMO-3 → LUMO (0.99)
14	T_9	Triplet-A	7.9586	155.79	Ultraviolet ■ (0.31, 0.33)	0.0000	HOMO-3 → LUMO+1 (0.98)
15	T_{10}	Triplet-A	8.0781	153.48	Ultraviolet ■ (0.31, 0.33)	0.0000	HOMO-4 → LUMO+1 (0.66) HOMO-1 → LUMO+4 (0.33)
16	S_6	Singlet-A	8.2535	150.22	Ultraviolet ■ (0.31, 0.33)	0.6935	HOMO-2 → LUMO+1 (0.62) HOMO-1 → LUMO (0.29)
17	S_7	Singlet-A	8.3438	148.59	Ultraviolet ■ (0.31, 0.33)	0.0059	HOMO-3 → LUMO+1 (0.99)
18	S_8	Singlet-A	8.3666	148.19	Ultraviolet ■ (0.31, 0.33)	0.6874	HOMO-2 → LUMO (0.51) HOMO-1 → LUMO+1 (0.33) HOMO → LUMO+2 (0.08)
19	S_9	Singlet-A	8.8265	140.47	Ultraviolet ■ (0.31, 0.33)	0.0000	HOMO-1 → LUMO+2 (0.99)
20	S_{10}	Singlet-A	8.9246	138.92	Ultraviolet ■ (0.31, 0.33)	0.0689	HOMO → LUMO+2 (0.86) HOMO-4 → LUMO (0.05) HOMO-2 → LUMO (0.05) HOMO-1 → LUMO+1 (0.02)

Transition Dipole Moments

Table 11: *Properties of the calculated transition dipole moments. [a]: The electric transition dipole moment (TEDM), in Debye (D). [b]: Angle between the TEDM and the x-axis of the molecule. [c]: Angle between the TEDM and xy-plane of the molecule. [d]: The magnetic transition dipole moment (TMDM), in atomic units (au). [e]: Angle between the TMDM and the x-axis of the molecule. [f]: Angle between the TMDM and xy-plane of the molecule. [g]: The TEDM, in Gaussian CGS (centimetre, gram, second) units. [h]: The TMDM, in Gaussian CGS (centimetre, gram, second) units. [i]: The angle between the electric and magnetic transition dipole moments, in Gaussian CGS units. [j]: The cosine of the angle between the electric and magnetic transition dipole moments, in Gaussian CGS units. [k]: The dissymmetry factor of the transition dipole moment.*

Excited State	μ[a] Vector /D	μ[a] /D	$θ_{μ,x}$[b] /	$θ_{μ,xy}$[c] /	m[d] Vector /au	m[d] /au	$θ_{m,x}$[e] /	$θ_{m,xy}$[f] /	μ[g] /esu·cm	m[h] /erg·G^{-1}	$θ_{μ,m}$[i] /	cos($θ_{μ,m}$)[j]	g_{lum}[k]
T_1	0.00, 0.00, 0.00	0.00	0.00	0.00	0.00, 0.00, 0.00	0.00	0.00	0.00	0.00e+00	0.00e+00	90.00	0.00	0.000
T_2	0.00, 0.00, 0.00	0.00	0.00	0.00	0.00, 0.00, 0.00	0.00	0.00	0.00	0.00e+00	0.00e+00	90.00	0.00	0.000
T_3	0.00, 0.00, 0.00	0.00	0.00	0.00	0.00, 0.00, 0.00	0.00	0.00	0.00	0.00e+00	0.00e+00	90.00	0.00	0.000
S_1	0.00, 0.00, 0.55	0.55	90.00	90.00	0.68, 0.00, 0.00	0.68	0.01	0.00	5.53e-19	6.35e-21	90.00	-0.00	-0.000
T_4	0.00, 0.00, 0.00	0.00	0.00	0.00	0.00, 0.00, 0.00	0.00	0.00	0.00	0.00e+00	0.00e+00	90.00	0.00	0.000
T_5	0.00, 0.00, 0.00	0.00	0.00	0.00	0.00, 0.00, 0.00	0.00	0.00	0.00	0.00e+00	0.00e+00	90.00	0.00	0.000
S_2	0.00, 0.00, 0.00	0.00	0.00	0.00	0.00, -0.02, 0.00	0.02	90.00	0.00	0.00e+00	2.29e-22	90.00	0.00	0.000
S_3	-1.14, 0.00, 0.00	1.14	0.00	0.00	0.00, 0.00, -0.18	0.18	90.00	90.00	1.14e-18	1.64e-21	90.00	0.00	0.000
T_6	0.00, 0.00, 0.00	0.00	0.00	0.00	0.00, 0.00, 0.00	0.00	0.00	0.00	0.00e+00	0.00e+00	90.00	0.00	0.000
S_4	0.00, -0.83, 0.00	0.83	90.00	0.00	0.00, 0.00, 0.00	0.00	0.00	0.00	8.32e-19	0.00e+00	90.00	0.00	0.000

Digichem 7.0.0-pre.2

Page 9 of 10

Figure S10. Excerpts from an example calculation report generated by Digichem, demonstrating tabulated data. The excited states of pyridine at the PBE0/6-31G* level of theory using the Tamm-Dancoff approximation (TDA) to time-dependent DFT (TD-DFT) were calculated. a) table of molecular geometry, b) table of selected molecular orbitals, c) table of electronic excited states, d) table of transition dipole moments.

Table S4. Table of supported input file types.

Code	Description	Read	Write	C&M
abinit	ABINIT Output Format	✅	❌	❌
acesin	ACES input format	❌	✅	❌
acesout	ACES output format	✅	❌	❌
acr	ACR format	✅	❌	❌
adf	ADF cartesian input format	❌	✅	❌
adfband	ADF Band output format	✅	❌	❌
adfdftb	ADF DFTB output format	✅	❌	❌
adfout	ADF output format	✅	❌	❌
alc	Alchemy format	✅	✅	❌
aoforce	Turbomole AOFORCE output format	✅	❌	❌
arc	Accelrys/MSI Biosym/Insight II CAR format	✅	❌	❌
ascii	ASCII format	❌	✅	❌
axsf	XCrySDen Structure Format	✅	❌	❌
bgf	MSI BGF format	✅	✅	❌
box	Dock 3.5 Box format	✅	✅	❌
bs	Ball and Stick format	✅	✅	❌
c09out	Crystal 09 output format	✅	❌	❌
c3d1	Chem3D Cartesian 1 format	✅	✅	❌
c3d2	Chem3D Cartesian 2 format	✅	✅	❌
cac	CAChe MolStruct format	❌	✅	❌
caccrt	Cacao Cartesian format	✅	✅	❌
cache	CAChe MolStruct format	❌	✅	❌
cacint	Cacao Internal format	❌	✅	❌
can	Canonical SMILES format	✅	✅	❌
car	Accelrys/MSI Biosym/Insight II CAR format	✅	❌	❌
castep	CASTEP format	✅	❌	❌
ccc	CCC format	✅	❌	❌
cdjson	ChemDoodle JSON	✅	✅	❌
cdx	ChemDraw binary format	✅	❌	❌
cdxml	ChemDraw CDXML format	✅	✅	❌
cht	Chemtool format	❌	✅	❌
cif	Crystallographic Information File	✅	✅	❌
ck	ChemKin format	✅	✅	❌
cml	Chemical Markup Language	✅	✅	❌
cmlr	CML Reaction format	✅	✅	❌

Table S4. Table of supported input file types.

Code	Description	Read	Write	C&M
cof	Culgi object file format	✅	✅	❌
com	Gaussian Input	✅	✅	✅
confabreport	Confab report format	❌	✅	❌
CONFIG	DL-POLY CONFIG	✅	✅	❌
CONTCAR	VASP format	✅	✅	❌
CONTFF	MDFF format	✅	✅	❌
crk2d	Chemical Resource Kit diagram(2D)	✅	✅	❌
crk3d	Chemical Resource Kit 3D format	✅	✅	❌
csr	Accelrys/MSI Quanta CSR format	❌	✅	❌
cssr	CSD CSSR format	❌	✅	❌
ct	ChemDraw Connection Table format	✅	✅	❌
cub	Gaussian cube format	✅	✅	❌
cube	Gaussian cube format	✅	✅	❌
dallog	DALTON output format	✅	❌	❌
dalmol	DALTON input format	✅	✅	✅
dat	Generic Output file format	✅	❌	❌
dmol	DMol3 coordinates format	✅	✅	❌
dx	OpenDX cube format for APBS	✅	✅	❌
ent	Protein Data Bank format	✅	✅	❌
exyz	Extended XYZ cartesian coordinates format	✅	✅	❌
fa	FASTA format	✅	✅	❌
fasta	FASTA format	✅	✅	❌
fch	Gaussian formatted checkpoint file format	✅	❌	❌
fchk	Gaussian formatted checkpoint file format	✅	❌	❌
fck	Gaussian formatted checkpoint file format	✅	❌	❌
feat	Feature format	✅	✅	❌
fh	Fenske-Hall Z-Matrix format	❌	✅	❌
fhiaims	FHIaims XYZ format	✅	✅	❌
fix	SMILES FIX format	❌	✅	❌
fps	FPS text fingerprint format (Dalke)	❌	✅	❌
fpt	Fingerprint format	❌	✅	❌
fract	Free Form Fractional format	✅	✅	❌
fs	Fastsearch format	✅	✅	❌
fsa	FASTA format	✅	✅	❌
g03	Gaussian Output	✅	❌	❌
g09	Gaussian Output	✅	❌	❌
g16	Gaussian Output	✅	❌	❌
g92	Gaussian Output	✅	❌	❌
g94	Gaussian Output	✅	❌	❌
g98	Gaussian Output	✅	❌	❌

Table S4. Table of supported input file types.

Code	Description	Read	Write	C&M
gal	Gaussian Output	✅	❌	❌
gam	GAMESS Output	✅	❌	❌
gamess	GAMESS Output	✅	❌	❌
gamin	GAMESS Input	✅	✅	❌
gamout	GAMESS Output	✅	❌	❌
gau	Gaussian Input	✅	✅	✅
gjc	Gaussian Input	✅	✅	✅
gjf	Gaussian Input	✅	✅	✅
got	GULP format	✅	❌	❌
gpr	Ghemical format	✅	✅	❌
gr96	GROMOS96 format	❌	✅	❌
gro	GRO format	✅	✅	❌
gukin	GAMESS-UK Input	✅	✅	❌
gukout	GAMESS-UK Output	✅	✅	❌
gzmat	Gaussian Z-Matrix Input	✅	✅	✅
hin	HyperChem HIN format	✅	✅	❌
HISTORY	DL-POLY HISTORY	✅	❌	❌
inchi	InChI format	✅	✅	❌
inchikey	InChIKey	❌	✅	❌
inp	GAMESS Input	✅	✅	❌
ins	ShelX format	✅	❌	❌
jin	Jaguar input format	✅	✅	❌
jout	Jaguar output format	✅	❌	❌
k	Compare molecules using InChI	❌	✅	❌
lmpdat	The LAMMPS data format	❌	✅	❌
log	Generic Output file format	✅	❌	❌
lpmd	LPMD format	✅	✅	❌
mcdl	MCDL format	✅	✅	❌
mcif	Macromolecular Crystallographic Info	✅	✅	❌
MDFF	MDFF format	✅	✅	❌
mdl	MDL MOL format	✅	✅	❌
ml2	Sybyl Mol2 format	✅	✅	❌
mmcif	Macromolecular Crystallographic Info	✅	✅	❌
mmd	MacroModel format	✅	✅	❌
mmod	MacroModel format	✅	✅	❌
mna	Multilevel Neighborhoods of Atoms (MNA)	❌	✅	❌
mol	MDL MOL format	✅	✅	❌
mol2	Sybyl Mol2 format	✅	✅	❌
mold	Molden format	✅	✅	❌
molden	Molden format	✅	✅	❌

Table S4. Table of supported input file types.

Code	Description	Read	Write	C&M
molf	Molden format	✅	✅	❌
molreport	Open Babel molecule report	❌	✅	❌
moo	MOPAC Output format	✅	❌	❌
mop	MOPAC Cartesian format	✅	✅	❌
mopcrt	MOPAC Cartesian format	✅	✅	❌
mopin	MOPAC Internal	✅	✅	❌
mopout	MOPAC Output format	✅	❌	❌
mp	Molpro input format	❌	✅	❌
mpc	MOPAC Cartesian format	✅	✅	❌
mpd	MolPrint2D format	❌	✅	❌
mpo	Molpro output format	✅	❌	❌
mpqc	MPQC output format	✅	❌	❌
mpqcin	MPQC simplified input format	❌	✅	❌
mrv	Chemical Markup Language	✅	✅	❌
msi	Accelrys/MSI Cerius II MSI format	✅	❌	❌
msms	M.F. Sanner's MSMS input format	❌	✅	❌
nw	NWChem input format	❌	✅	❌
nwo	NWChem output format	✅	❌	❌
orca	ORCA output format	✅	❌	❌
orcainp	ORCA input format	❌	✅	❌
out	Generic Output file format	✅	❌	❌
outmol	DMol3 coordinates format	✅	✅	❌
output	Generic Output file format	✅	❌	❌
paint	Painter format	❌	✅	❌
pc	PubChem format	✅	❌	❌
pcjson	PubChem JSON	✅	✅	❌
pcm	PCModel Format	✅	✅	❌
pdb	Protein Data Bank format	✅	✅	❌
pdbqt	AutoDock PDBQT format	✅	✅	❌
png	PNG 2D depiction	✅	✅	❌
pointcloud	Point cloud on VDW surface	❌	✅	❌
pos	POS cartesian coordinates format	✅	❌	❌
POSCAR	VASP format	✅	✅	❌
POSFF	MDFF format	✅	✅	❌
pov	POV-Ray input format	❌	✅	❌
pqr	PQR format	✅	✅	❌
pqs	Parallel Quantum Solutions format	✅	✅	❌
prep	Amber Prep format	✅	❌	❌
pwscf	PWscf format	✅	❌	❌
qcin	Q-Chem input format	❌	✅	❌

Table S4. Table of supported input file types.

Code	Description	Read	Write	C&M
qcout	Q-Chem output format	✅	❌	❌
report	Open Babel report format	❌	✅	❌
res	ShelX format	✅	❌	❌
rinchi	RInChI	❌	✅	❌
rsmi	Reaction SMILES format	✅	✅	❌
rxn	MDL RXN format	✅	✅	❌
sd	MDL MOL format	✅	✅	❌
sdf	MDL MOL format	✅	✅	❌
si	Silico Input Format	✅	✅	✅
siesta	SIESTA format	✅	❌	❌
smi	SMILES format	✅	✅	❌
smiles	SMILES format	✅	✅	❌
smy	SMILES format using Smiley parser	✅	❌	❌
stl	STL 3D-printing format	❌	✅	❌
svg	SVG 2D depiction	❌	✅	❌
sy2	Sybyl Mol2 format	✅	✅	❌
t41	ADF TAPE41 format	✅	❌	❌
tdd	Thermo format	✅	✅	❌
therm	Thermo format	✅	✅	❌
tmol	TurboMole Coordinate format	✅	✅	❌
txyz	Tinker XYZ format	✅	✅	❌
unixyz	UniChem XYZ format	✅	✅	❌
VASP	VASP format	✅	✅	❌
vmol	ViewMol format	✅	✅	❌
wln	Wiswesser Line Notation	✅	❌	❌
xed	XED format	❌	✅	❌
xml	General XML format	✅	❌	❌
xsf	XCrySDen Structure Format	✅	❌	❌
xtc	XTC format	✅	❌	❌
xyz	XYZ cartesian coordinates format	✅	✅	❌
yob	YASARA.org YOB format	✅	✅	❌
zin	ZINDO input format	❌	✅	❌

Recreated from ref.[16] **Read**: Indicates this format can be read. **Write**: indicates this format can be written. **C&M**: Indicates this format supports charge and multiplicity. The 'com'/'gjf', 'si', and 'xyz' formats are parsed internally. The 'log' and related formats are parsed with cclib.[17] The remaining formats are parsed with Open Babel.[18,19]

References

1 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman and D. J. Fox, Gaussian 16, Revision C.01 2016.

2 J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1997, **78**, 1396–1396.

3 J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865–3868.

4 C. Adamo and V. Barone, *J. Chem. Phys.*, 1999, **110**, 6158–6170.

5 S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *J. Chem. Phys.*, 2010, **132**, 154104.

6 S. Grimme, S. Ehrlich and L. Goerigk, *J. Comput. Chem.*, 2011, **32**, 1456–1465.

7 R. Ditchfield, W. J. Hehre and J. A. Pople, *J. Chem. Phys.*, 1971, **54**, 724–728.

8 W. J. Hehre, R. Ditchfield and J. A. Pople, *J. Chem. Phys.*, 1972, **56**, 2257–2261.

9 M. M. Francl, W. J. Pietro, W. J. Hehre, J. S. Binkley, M. S. Gordon, D. J. DeFrees and J. A. Pople, *J. Chem. Phys.*, 1982, **77**, 3654–3665.

10 O. S. Lee and E. Zysman-Colman, Digichem (version 7.0.0-pre.3) Digichem Project, St Andrews, Scotland, 2024.

11 W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graph.*, 1996, **14**, 33–38.

12 J. Stone, PhD thesis, Computer Science Department, University of Missouri-Rolla, 1998.

13 O. S. Lee, PhD thesis, University of St Andrews, 2024.

14 F. Tenopala-Carmona, O. S. Lee, E. Crovini, A. M. Neferu, C. Murawski, Y. Olivier, E. Zysman-Colman and M. C. Gather, *Adv. Mater.*, 2021, **33**, 2100677.

15 K. V. Mardia and P. E. Jupp, in *Encyclopedia of Statistical Sciences*, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2006, vol. 25.

16 Digichem Convert — Digichem 6.1.0 documentation, https://doc.digichem.co.uk/reference/digichem-convert.html#supported-file-formats, (accessed 22 July 2024).

17 E. Berquist, A. Dumi, S. Upadhyay, O. D. Abarbanel, M. Cho, S. Gaur, V. H. C. Gil, G. R. Hutchison, O. S. Lee, A. S. Rosen, S. Schamnad, F. S. S. Schneider, C. Steinmann, M. Stolyarchuk, J. E. Vandezande, W. Zak and K. M. Langner, .

18 N. M. O'Boyle, M. Banck, C. A. James, C. Morley, T. Vandermeersch and G. R. Hutchison, *J. Cheminformatics*, 2011, **3**, 33.

19 N. M. O'Boyle, C. Morley and G. R. Hutchison, *Chem. Cent. J.*, 2008, **2**, 5.