PILGRIM v2.0
A thermal rate constant calculator and a chemical kinetics simulator

David Ferro-Costas[1]
Donald G. Truhlar[2]
Antonio Fernández-Ramos[1]

[1] Universidade de Santiago de Compostela (SPAIN)
[2] University of Minnesota (USA)
About Pilgrim License

MIT License

Copyright (c) 2020 David Ferro Costas and Antonio Fernández Ramos

Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the “Software”), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software.

THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.

February 2020
## Part One - Introduction to Pilgrim

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Introduction</td>
<td>9</td>
</tr>
<tr>
<td>1.2 Cathedral Package</td>
<td>10</td>
</tr>
<tr>
<td>1.3 How to obtain <em>Pilgrim</em>?</td>
<td>10</td>
</tr>
<tr>
<td>1.4 Terms of use</td>
<td>10</td>
</tr>
<tr>
<td>1.5 About this manual</td>
<td>10</td>
</tr>
<tr>
<td>1.6 <em>Pilgrim</em> software contents</td>
<td>11</td>
</tr>
<tr>
<td>1.7 Software requirements</td>
<td>11</td>
</tr>
<tr>
<td>1.7.1 The electronic structure software</td>
<td>11</td>
</tr>
<tr>
<td>1.8 <em>Pilgrim</em> Facebook web page</td>
<td>12</td>
</tr>
<tr>
<td>1.9 How to cite <em>Pilgrim</em>?</td>
<td>12</td>
</tr>
</tbody>
</table>
1.10 Version history

1.10.1 List of changes in version 2.0

2 Getting started with Pilgrim

2.1 Setup

2.1.1 Organization of the electronic structure files

2.1.2 The --gather option

2.1.3 About labels

2.2 Input files preparation

2.3 Calculations

2.3.1 Options

2.3.2 Targets

2.3.3 Output files

2.3.4 Plotting results

2.3.5 Kinetic Isotopic Effects

II Part Two - Worked examples

3 Isomerization of formic acid

3.1 About this system

3.2 WE01: TST

3.3 WE02: CVT/SCT (I)

3.4 WE03: CVT/SCT (II)

3.5 WE04: CVT/SCT with an isotopic substitution
4  The inversion of ammonia ............................................. 33
   4.1 About this system .............................................. 33
   4.2 WE05: LL and DL CVT/SCT .................................... 34
   4.3 WE06: LL and DL CVT/SCT for root and deuterated species 35

5  Hydrogen abstraction reactions ................................... 37
   5.1 About this system .............................................. 37
   5.2 WE07: MS-TST and KMC ....................................... 39
   5.3 WE08: MP-CVT/SCT with torsional anharmonicity and KMC 40
   5.4 WE09: KMC simulation using analytical expressions ............ 41
   5.5 WE10: MP-CVT/SCT, isotopic substitutions and anharmonicity 42

6  Solvated HA with ONIOM ........................................... 43
   6.1 About this system .............................................. 43
   6.2 WE11: MP-CVT/SCT rate constants with ONIOM ................. 44

7  Checking results .......................................................... 47

III  Part Three - Pilgrim input files

8  Pilgrim input files ......................................................... 51
   8.1 The pit.struc file ............................................... 52
       8.1.1 Conformers and weights .................................. 53
       8.1.2 Electronic states .......................................... 54
       8.1.3 Internal coordinates ....................................... 54
       8.1.4 Isotopic substitutions ..................................... 55
8.2 The *pif.temp* file 57
8.3 The *pif.chem* file 58
8.4 The *pif.path* file 59
8.4.1 Basic keywords .............................................................. 59
8.4.2 Advanced keywords ........................................................... 61
8.5 The *pif.calcs* file 64
8.6 The *pif.kmc* file 66
8.7 The *pif.dlevel* file 69

IV  Part Four - Bibliography

9  Bibliography ................................................................. 73
9.1 General references about methods in *Pilgrim.* 73
9.2 Specific references for specific methods in *Pilgrim.* 78
Part One - Introduction to Pilgrim

1 About Pilgrim ....................... 9
1.1 Introduction
1.2 Cathedral Package
1.3 How to obtain Pilgrim?
1.4 Terms of use
1.5 About this manual
1.6 Pilgrim software contents
1.7 Software requirements
1.8 Pilgrim Facebook web page
1.9 How to cite Pilgrim?
1.10 Version history

2 Getting started with Pilgrim ...... 15
2.1 Setup
2.2 Input files preparation
2.3 Calculations
1. About Pilgrim

1.1 Introduction

Pilgrim is an user-friendly program written in Python 3 and designed to use direct-dynamics to calculate thermal rate constants of chemical reactions and to simulate chemical kinetics mechanisms.

For reaction processes with many elementary steps, each of these steps can be calculated using conventional transition state theory (TST) or variational transition state theory (VTST). In this version, Pilgrim can calculate thermal rate constants with the canonical version of the variational transition state theory (CVT), which requires the calculation of the minimum energy path (MEP) associated to each elementary step. Moreover, multi-dimensional quantum effects can be incorporated through the small-curvature tunneling approximation (SCT). The above methodologies are available for reactions involving a single structure and for reactions involving flexible molecules with multiple conformations. Specifically, for systems with many conformers the program can evaluate each elementary reaction by multi-structural canonical variational transition state theory (MS-CVT) or multi-path VTST (MP-VTST). Torsional anharmonicity can be also incorporated through the MSTor and Q2DTor programs.

Pilgrim also performs dual-level calculations automatically. First, low-level calculations are carried out for the reaction of interest and second, single-point energy calculations of the reactants, transition state, points along the MEP and products are performed at a higher level. Low-level calculations are corrected with high-level single point energies using the interpolated single-point energies (ISPE) algorithm.

Once all the rate constants of the chemical processes of interest are known, by means of their calculation using Pilgrim or by using an analytical expression, it is possible to simulate the whole process using kinetic Monte Carlo (KMC). This algorithm allows performing a kinetics simulation and monitoring the time evolution of each chemical species, as well as providing its chemical yield.
1.2 Cathedral Package

*Pilgrim* v2.0 belongs to a bigger repository named **The Cathedral Package**:

https://github.com/cathedralpkg/

The following codes are part of this repository:

- *Pilgrim* v.2.0
- Q2DTor v2.0
- TorsiFlex (to be released)

1.3 How to obtain *Pilgrim*?

*Pilgrim* can be downloaded from GitHub at:

https://github.com/cathedralpkg/pilgrim/releases

or from Don Truhlar's group webpage:

https://comp.chem.umn.edu/pilgrim/

1.4 Terms of use

*Pilgrim* is free software under MIT license. We refer to the GitHub webpage for more details about the license:

https://github.com/cathedralpkg/pilgrim

1.5 About this manual

This manual is organized as follows:

- **Chapter 1** introduces *Pilgrim* and gives important information about its installation (the current chapter).
- **Chapter 2** presents an overview about how to use *Pilgrim*. For users who are familiar with VTST and with electronic structure programs, this chapter contains most of the information they need to run the program. Occasionally, they may need to consult other chapters to find a more detailed explanation about a given particular aspect.
- **Chapters 3-7** describe some worked examples (WEs) that allow the user to explore some capabilities of the program.
1.6 **Pilgrim software contents**

*Pilgrim* is provided in a tar.gz file which can be uncompressed using:

```
tar -zxvf Pilgrim-2.0.tar.gz
```

A new directory with the name `Pilgrim-2.0/` is created. Inside the directory the user can find the README file and three subdirectories:

- **src/**, which contains the source code;
- **docs/**, which includes this manual.
- **tests/**, which contains two directories:
  - **esfils/**, with the electronic structure files (ESFILs) of the stationary points needed to run each of the worked examples (WEs) using *Pilgrim* (see Chapters 3-6);[i]
  - **outputs/** with the ‘already run’ WEs, so they can be compared with those obtained by running the tests.
  - **cpc_examples/** with the examples described in *Comput. Phys. Commun.*

1.7 **Software requirements**

Before running *Pilgrim* the user should have installed the following software:

- Python 3 together with the following Python libraries:

  ```
  cmath  mathlib  random
  fcntl  multiprocessing  scipy
  glob  numpy  sys
  math  os  time
  ```

  If any of these libraries is missing, *Pilgrim* displays an error message indicating which module needs to be installed.

- Electronic structure software (ESSO), in particular, *Gaussian* and/or Orca

1.7.1 **The electronic structure software**

In order to interact with *Gaussian* and/or Orca, *Pilgrim* needs to know the location of some executable files. Such information is obtained from the following environment variables, which have to be defined and exported by the user in their `.bashrc` file:

[i] The ESFILs are the starting point to run the tests. All the tests were designed to be run at low levels (HF/STO-3G, HF/3-21G and PM6//HF/STO-3G dual level) so they can be performed in a short period of time.
• Environment variables for Gaussian software:
  – GauExe: with the path to the Gaussian executable
  – GauFchk: with the path to the formchk tool
• Environment variable for Orca software:
  – OrcaExe: with the path to the Orca executable

For example:

```
.bashrc: exporting environment variables
export GauExe="/home/programs/g09/g09"
export GauFchk="/home/programs/g09/formchk"
export OrcaExe="/home/programs/orca_4_0_1_2/orca"
```

### 1.8 Pilgrim Facebook web page

We recommend to follow our Cathedral Package Facebook web page for updates, questions and other program information: https://www.facebook.com/cathedralpkg/

### 1.9 How to cite Pilgrim?

If you are using Pilgrim, please cite it using the following reference:

- D. Ferro-Costas, D.G. Truhlar, A. Fernández-Ramos, Pilgrim - version 2.0 (University of Minneapolis, Minnesota, MN, and Universidade de Santiago de Compostela, Spain, 2020). https://github.com/cathedralpkg/Pilgrim

We refer to Chapter 9 for references about key general methods and about specific methods used in Pilgrim.

### 1.10 Version history

#### 1.10.1 List of changes in version 2.0

- Pilgrim is now part of a bigger project, The Cathedral Package;
- Pilgrim has been ported to Python 3;
- The reading function of Gaussian output files has been improved.
- Linear molecules are automatically orientated along X, Y or Z axis in order to avoid problems in the calculation of vibrational frequencies.
• ONIOM layers (without link atoms) can be defined in the calculation of the MEP through the \texttt{oniomX} keyword ($X = l, m$ and $h$) in \texttt{pif.path} (only available with \textit{Gaussian} calculations).
• Keyword \texttt{keeptmp} is available in \texttt{pif.path}.
• If $E_0 > V_{a}^{AG}$, the transmission coefficients are set to unity.
• The \texttt{sctmns} keyword has been removed.
• Multi-structural (MS) rate constants are also calculated.
• Several mechanisms may be defined inside \texttt{pif.kmc}; now this file contains \texttt{kmc} blocks.
• New analytic expression (\texttt{analytic5}) has been included.
• The calculation of kinetic isotopic effects (KIEs) is now available through the new execution option --\texttt{kies}.
• If external anharmonicity files are specified, \textit{Pilgrim} assumes that the user is only interested in anharmonic rate constants, so harmonic ones are not calculated.
• Other minor changes.
2. Getting started with Pilgrim

The execution of Pilgrim can be divided into three different parts:

1. Setup.
2. Input files preparation.
3. Calculations.

In this chapter, we explain how to proceed at each stage.

2.1 Setup

This section describes how to organize the electronic structure files (ESFILs) for the reaction mechanism object of study. At the moment Pilgrim can work with the following electronic structure software (ESSO):

- Gaussian 03, 09 and 16
- Orca 3.x and Orca 4.x

2.1.1 Organization of the electronic structure files

Prior to work with Pilgrim, the user should have at hand the ESFILs of all the stationary points relevant to the reaction mechanism. The reaction mechanism may consist of one or several elementary chemical reactions. Each ESFIL should contain the optimized geometry and Hessian of a given stationary point, and they have to be organized as indicated in Figure 2.1 (blue boxes) and outlined below:

- Create a directory (hereafter MDIR/) to study your reaction mechanism using Pilgrim.
- Create a subdirectory called UDATA/ inside MDIR/.
- Create a subdirectory inside UDATA/ for each of the species that participate...
in the chemical mechanism. Hereafter the subdirectories inside UDATA/ will have the generic name spname/ and the spname label is the name by which Pilgrim recognizes that stationary point species.\textsuperscript{[i]} Therefore use meaningful names for those directories.\textsuperscript{[ii]}

- Place the ESFILs in the corresponding spname/ directory. The allowed formats are listed in the green boxes of Figure 2.1. If a given species has several conformational isomers\textsuperscript{[iii]} they should be included in the same spname/ directory.
  
  - **Which conformers should be included?** All the conformers that are distinguishable except in the case of conformational enantiomers that only one of them should be included.\textsuperscript{[iv]}

- Pilgrim can incorporate torsional anharmonicity calculated with MSTor or Q2DTor programs. The output files from these programs should be located in a directory hanging from MDIR/ and called ANHAR/. All the torsional anharmonicity files should be placed inside this directory.\textsuperscript{[v]}

### 2.1.2 The \texttt{--gather} option

Once the ESFILs are organized as previously described, execute Pilgrim from MDIR/ using:

```bash
pilgrim.py --gather
```

Pilgrim creates two directories inside MDIR/: (i) 1-GTS/ contains the .gts files (for internal use by Pilgrim) of all electronic structure files of the stationary points included in the subdirectories of UDATA/; (ii) 5-MOLDEN/ includes files for visualizing with Molden. The program also creates the files tracking and pif.struc. The former correlates the ESFILs of UDATA/ with the .gts files of 1-GTS/ and the latter provides information about each structure. The .gts files are labeled as: spname.idx.gts, where spname indicates the name of the species and idx is the numbering of the torsional conformation for that species (starting at 001) and sorted by increasing electronic energy. If there is only one conformation the .gts file inside 1-GTS/ will be called spname.001.gts

File pif.struc is the only one that may need some editing by the user as it contains,

\begin{itemize}
\item \textsuperscript{[i]}For species with a single stationary point the ESFIL may hang directly from UDATA/, but for species with several conformations the creation of a directory is compulsory. For the former the spname label is the name of the ESFIL without extension.
\item \textsuperscript{[ii]}Names should start with a letter; the rest of characters may be underscores (_), letters or numbers.
\item \textsuperscript{[iii]}Also called rotamers, torsional conformers or simply conformers. Conformational isomers are species that can interconvert among them by internal rotations.
\item \textsuperscript{[iv]}For instance, ethanol has three conformations, two gauche and one anti; however, the two gauche conformations are conformational enantiomers. Notice that this couple of conformational enantiomers can be described by a unique ESFIL; therefore, only one of them has to be specified and the other is easily incorporated later on when the input files of Pilgrim are created.
\item \textsuperscript{[v]}Anharmonicity output files are not recognized if placed somewhere else.
\end{itemize}
2.1 Setup

Figure 2.1: This scheme shows the directories tree that should be created by the user (blue boxes) and the directories that are created by the program (red boxes) after execution with \texttt{--gather}. At this stage the user can also incorporate the output files of the MSTor or Q2DTor programs into the \texttt{ANHAR/} directory.

for instance, the scaling factor for vibrational frequencies, the weight of a given conformation,\textsuperscript{[vi]} etc. For more information, see Chapter 8.

2.1.3 About labels

The \texttt{pif.struc} file is formed by a set of blocks (see Chapter 8), i.e., one block for each of the species defined in \texttt{UDATA/} and labeled by its corresponding \texttt{spname} (which was previously defined by the user). However, with the idea of accommodating isotopic substituted species, the generic name \texttt{ctcsp} will be used instead of \texttt{spname} for the labels hereinafter. In the same manner, the generic label \texttt{ctcts} will be used for blocks associated with transition state species.\textsuperscript{[vii]} Notice that the \texttt{ctcsp} and \texttt{ctcts} labels are also used by \texttt{Pilgrim} as ‘targets’ during the execution stage.

\textit{Pilgrim} automatically assigns all the \texttt{spname} labels as \texttt{ctcsp} labels, but intervention by the user is required to study isotopically substituted reactions. In that case the ESFILs associated with a given \texttt{spname} are valid for the root species, as well as for any isotopically modified species, but new blocks with their corresponding labels have to be added to incorporate the new species. In this case the \texttt{ctcsp} labels include those of the root species plus the ones of the isotopically substituted species.

For example, to study the reaction \texttt{EtOH + H $\rightarrow$ EtO + H\textsubscript{2}} the \texttt{spname} labels were

\textsuperscript{[vi]} Conformational isomers that have another enantiomer, should be added at this stage

\textsuperscript{[vii]} Therefore the label of a transition state is at the same time a \texttt{ctcsp} label and a \texttt{ctcts} label.
named as: EtOH for ethanol, H for hydrogen atom, EtO for the EtO radical, H2 for the H2 molecule and TSc for the hydrogen abstraction transition state. The ctctsp labels used by Pilgrim in pif.struc coincided with the spname labels. Within the ctctsp labels, the only ctctts label is TSc.

However, to study also the isotopically substituted reaction EtOD + D → EtO + D2, new blocks and labels should be added to incorporate the isotopic species. Notice that the spname labels are the same as for the previous reaction. There is also a new ctctts label: TScOD. With the exception of pif.struc which was generated in the previous stage. However, the interactive menu can modify it.

2.2 Input files preparation

The generation of Pilgrim input files can be carried out using:

```
pilgrim.py --input
```

which displays an interactive and self-explanatory menu. This menu allows the user to generate all the input files. The variables inside the menu, their meaning and the generated Pilgrim input files (pifs) are displayed in Table 2.1. Although this menu can generate each of the files needed to run Pilgrim, some of them may need further editing outside the menu to adjust some default options. For a detailed description of each input file, the user should read Chapter 8.

Table 2.1: Description of the variables of Pilgrim inside the input menu and the input files modified or generated by each option.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Information</th>
<th>Pilgrim Input File</th>
</tr>
</thead>
<tbody>
<tr>
<td>struc</td>
<td>Structures</td>
<td>pif.struc</td>
</tr>
<tr>
<td>isomass</td>
<td>Isotopic masses</td>
<td>pif.struc</td>
</tr>
<tr>
<td>temp</td>
<td>Temperatures (in K)</td>
<td>pif.temp</td>
</tr>
<tr>
<td>chem</td>
<td>Chemical reactions</td>
<td>pif.chem</td>
</tr>
<tr>
<td>path</td>
<td>MEP parameters</td>
<td>pif.path &amp; pif.calcs</td>
</tr>
<tr>
<td>kmc</td>
<td>Chemical kinetics simulation</td>
<td>pif.kmc</td>
</tr>
<tr>
<td>dlevel</td>
<td>Dual-level calculations</td>
<td>pif.dlevel &amp; pif.calcs</td>
</tr>
</tbody>
</table>

When the interactive menu is initialized, it displays the status of every possible input file. The first time that we enter the menu it looks like:
Status of input files

<table>
<thead>
<tr>
<th>input file</th>
<th>status</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1 : pif.struc</td>
<td>1</td>
</tr>
<tr>
<td>#2 : pif.temp</td>
<td>-1</td>
</tr>
<tr>
<td>#3 : pif.path</td>
<td>-1</td>
</tr>
<tr>
<td>#4 : pif.calcs</td>
<td>-1</td>
</tr>
<tr>
<td>#5 : pif.chem</td>
<td>-1</td>
</tr>
<tr>
<td>#6 : pif.kmc</td>
<td>-1</td>
</tr>
<tr>
<td>#7 : pif.dlevel</td>
<td>-1</td>
</tr>
</tbody>
</table>

status = -1 ==> file does not exist
status = 0 ==> file exists but it is empty
status = 1 ==> file exists and contains data

Notice that none of the files exists with the exception of pif.struc. The menu also displays:

There are several variables ($var) and commands ($cmd) available in this interactive menu. The command line should have the following syntax:

> $cmd $var [$values]

where the square brackets indicate that the $cmd-$var combination may require the specification of values ($values). For more information, use the 'help' command on each variable inside the interactive menu.

List of commands ($cmd) and variables ($var):

<table>
<thead>
<tr>
<th>$cmd$var</th>
<th>struc</th>
<th>isomass</th>
<th>temp</th>
<th>chem</th>
<th>path</th>
<th>kmc</th>
<th>dlevel</th>
</tr>
</thead>
<tbody>
<tr>
<td>help</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>ls</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>add</td>
<td></td>
<td>x+</td>
<td>x+</td>
<td>x+</td>
<td>x+</td>
<td>x+</td>
<td>x</td>
</tr>
<tr>
<td>mod</td>
<td>x+</td>
<td></td>
<td></td>
<td></td>
<td>x+</td>
<td>x+</td>
<td></td>
</tr>
<tr>
<td>rm</td>
<td>x+</td>
<td>x</td>
<td>x+</td>
<td>x+</td>
<td>x+</td>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>

x: the combination $cmd $var is available
+: the combination $cmd $var requires $values

Information about variables ($var):

<table>
<thead>
<tr>
<th>$var</th>
<th>addresses...</th>
<th>which contains...</th>
</tr>
</thead>
<tbody>
<tr>
<td>struc</td>
<td>pif.struc</td>
<td>structures &amp; isot. masses</td>
</tr>
</tbody>
</table>
Chapter 2. Getting started with Pilgrim

To go back to an upper level in the menu or to exit, use one of the next strings: end / .. / exit

The information about a given variable can be displayed using the `help` command. We highly recommend to read the help message of each of the variables to learn how to proceed.

### 2.3 Calculations

`Pilgrim` was designed to be executed in steps, and some of the options of the program follow a given order. Within this step-by-step philosophy, each execution carries out a different task, whose result can be checked by the user before proceeding to another step. The general command line for executing `Pilgrim` is:

```
pilgrim --option [target[.idx]] [--dlevel] [--software ESSO]
```

#### 2.3.1 Options

The available options in `Pilgrim` are shown in Table 2.2. The `--software` option can be only combined with `--path` (calculation of the minimum energy path) and with `--hlcalc` (high-level single point calculation). By default, the ESSO is set to `Gaussian` (`--software gaussian`). Therefore, this keyword is only required when using Orca (`--software orca`). The `--dlevel` option should be only specified when performing dual-level calculations (previous high-level calculations are required).

#### 2.3.2 Targets

When `target` is omitted from the command line, `--option` is applied to all possible targets. Different options have different targets as indicated in Table 2.3. For instance, option `--path` can have as target only `ctcts` labels. If a `ctcts` contains more than one conformation, it is possible to calculate the MEP of an individual TS structure using `ctcts.idx`, where `idx` is the numerical index of the conformation (from 001 to 999) inside the block in `pif.struc`. In the case of the calculation of rate constants, the option `--rcons` may have as target the name of a elementary chemical reaction that was defined in the input with the variable `chem`, and with
Table 2.2: Description of the options of Pilgrim; the order of execution of the main options is also indicated.

<table>
<thead>
<tr>
<th>Order of execution, --option and explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 --pfn Calculates the partition functions at the stationary points</td>
</tr>
<tr>
<td>2 --path Calculates the MEP &amp; VTST coefficients</td>
</tr>
<tr>
<td>3 --rcons Calculates the thermal rate constants</td>
</tr>
<tr>
<td>4 --kmc Performs a kinetic Monte Carlo simulation</td>
</tr>
</tbody>
</table>

Options for special actions

--dlevel Dual-level calculations and interpolation via ISPE
--software Selects ESSO [gaussian (default) or orca];
            it goes together with --path or --hlcalc;
            if omitted, gaussian is selected by default;

Additional options and tools

--ics Generates internal coordinates
--hlcalc Performs high-level calculations needed for dual-level
--fit Fits the thermal rate constants to prescribed analytical expressions
--plot Generates diverse plots
--kies Calculates total KIEs and its contributions

Information options

--help Displays help messages; --help option can also be used
--ls Lists the species defined in pif.struc
--version Displays the version of the program

the generic label chemname. For --kmc the target is a mechanism defined in the pif.kmc file and referred as kinmec.

Table 2.3: List of the targets allowed by the different options. The default (the target is omitted) is also indicated.

<table>
<thead>
<tr>
<th>--option</th>
<th>default</th>
<th>allowed target</th>
</tr>
</thead>
<tbody>
<tr>
<td>--pfn</td>
<td>all ctcsp</td>
<td>ctcsp</td>
</tr>
<tr>
<td>--path</td>
<td>all ctcts</td>
<td>ctcts[.idx]</td>
</tr>
<tr>
<td>--rcons</td>
<td>all chemname</td>
<td>chemname</td>
</tr>
<tr>
<td>--kmc</td>
<td>all kinmec</td>
<td>kinmec</td>
</tr>
<tr>
<td>--ics</td>
<td>all ctcts</td>
<td>ctcts[.idx]</td>
</tr>
<tr>
<td>--hlcalc</td>
<td>all ctcsp</td>
<td>ctcsp[.idx]</td>
</tr>
<tr>
<td>--fit</td>
<td>all chemname</td>
<td>chemname</td>
</tr>
<tr>
<td>--plot</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>--kies</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>
2.3.3 Output files

A list of all the directories generated by Pilgrim is given in Table 2.4. The output files generated after the execution with a given option are indicated in Table 2.5.

Table 2.4: Directories created by Pilgrim during the execution of the program

<table>
<thead>
<tr>
<th>Directory</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-GTS/</td>
<td>The .gts files generated by Pilgrim from the ESSO output files</td>
</tr>
<tr>
<td>2-PLG_DATA/</td>
<td>Data files (Pilgrim’s internal use)</td>
</tr>
<tr>
<td>3-PLG_OUTPUT/</td>
<td>Output files</td>
</tr>
<tr>
<td>4-PLG_RST/</td>
<td>Restart files for MEP calculations (Pilgrim’s internal use)</td>
</tr>
<tr>
<td>5-MOLDEN/</td>
<td>Files in Molden and xyz format</td>
</tr>
<tr>
<td>6-PLOTFILES/</td>
<td>Data for plotting</td>
</tr>
</tbody>
</table>

Table 2.5: Output files generated in the 3-PLG_OUTPUT/ directory when executing Pilgrim with different options according to the selected target. If the options are run together with --dlevel, the word slevel is replaced by dlevel in the output file name.

<table>
<thead>
<tr>
<th>--option</th>
<th>target</th>
<th>Output files</th>
</tr>
</thead>
<tbody>
<tr>
<td>--pfns</td>
<td>ctcsp</td>
<td>pfns.ctcs.slevel.txt</td>
</tr>
<tr>
<td>--path</td>
<td>ctcts.idx</td>
<td>path.ctcts.idx.slevel.txt</td>
</tr>
<tr>
<td>--rcons</td>
<td>chemname</td>
<td>rcons.chemname.slevel.txt</td>
</tr>
<tr>
<td>--kmc</td>
<td>mech</td>
<td>kmc.mech.slevel.txt</td>
</tr>
</tbody>
</table>

2.3.4 Plotting results

During the execution, Pilgrim stores data for plotting in:

- 6-PLOTFILES/plots.slevel.txt or
- 6-PLOTFILES/plots.dlevel.txt (when --dlevel is activated)

In order to plot this data, use the --plot option (together with --dlevel if necessary). For more information, we refer to the help message associated to this option, which can be displayed with:

```
pilgrim --help plot
```

2.3.5 Kinetic Isotopic Effects

If root and isotopic reactions are defined in the same working directory, the --kies tool can be used to calculate the total Kinetic Isotopic Effect (KIE) as well as to split its value into different contributions. When executed...
Pilgrim asks for both the root and isotopic reactions. Once introduced, Pilgrim calculates and prints the corresponding KIEs.
Part Two - Worked examples

3
Isomerization of formic acid .......................... 27

3.1 About this system
3.2 WE01: TST
3.3 WE02: CVT/SCT (I)
3.4 WE03: CVT/SCT (II)
3.5 WE04: CVT/SCT with an isotopic substitution

4
The inversion of ammonia .................................. 33

4.1 About this system
4.2 WE05: LL and DL CVT/SCT
4.3 WE06: LL and DL CVT/SCT for root and deuterated species
4.4 WE07: MS-TST and KMC
4.5 WE08: MP-CVT/SCT with torsional anharmonicity
4.6 WE09: KMC simulation using analytical expressions
4.7 WE10: MP-CVT/SCT, isotopic substitutions and anharmonicity
4.8 WE11: MP-CVT/SCT rate constants with ONIOM

5
Hydrogen abstraction reactions ........................... 37

5.1 About this system
5.2 WE0: MS-TST and KMC
5.3 WE08: MP-CVT/SCT with torsional anharmonicity
5.4 WE09: KMC simulation using analytical expressions
5.5 WE10: MP-CVT/SCT, isotopic substitutions and anharmonicity

6
Solvated HA with ONIOM ................................. 43

6.1 About this system
6.2 WE11: MP-CVT/SCT rate constants with ONIOM

7
Checking results ........................................... 47

(Continued on the next page)
3. Isomerization of formic acid

3.1 About this system

This Chapter describes the isomerization process between the two isomers of formic acid (FA): $Z$-FA and $E$-FA. The concerned reaction is:

$$Z - \text{HCOOH} \rightleftharpoons E - \text{HCOOH}$$

and we want to estimate the rate constant of the process at 300, 400 and 500 K.

The four WEs designed for this reaction include the following methodologies:

- WE01: TST
- WE02: CVT/SCT (I)\(^{[i]}\)
- WE03: CVT/SCT (II)\(^{[ii]}\)
- WE04: CVT/SCT with an isotopic substitution
- ESSO: Gaussian 09
- Level of calculation: HF/STO-3G
- Working temperatures (in K): 300, 400 and 500

\(^{[i]}\)The projection of the frequencies along the path is carried out in Cartesian coordinates.

\(^{[ii]}\)The projection of the frequencies along the path is carried out in internal coordinates.
Chapter 3. Isomerization of formic acid

3.2 WE01: TST

Study the isomerization reaction using conventional TST.

Steps:

- Create the directory WE01/. Hereafter, the working directory for this system.
- Copy the directory UDATA_FA/ from tests/esfils/ to WE01/ as UDATA/. Note that no subdirectories are needed inside UDATA/ because none of the structures has additional conformational isomers (in fact the reaction is a conformational isomerization itself). The ESFILs are: Z.fchk, E.fchk and ZE_TS.fchk.
- Run ‘pilgrim --gather’. This creates the files tracking and pif.struc, as well as the directories 1-GTS/ and 5-MOLDEN/ that contain the gts and molden files, respectively. The latter can be visualized with Molden.
- Use the interactive menu of Pilgrim by typing ‘pilgrim --input’ and define the temperatures and the chemical reaction with:
  > add temp 300 400 500
  > add chem FA_ISO : Z --> ZE_TS --> E
After exiting the menu, the files pif.temp and pif.chem are created. In the input menu we have labeled the reaction as FA_ISO. Notice that the chemical species have the same name as the ESFILs of UDATA/ without the extension. The labels E, Z and ZE_TS are the spname labels of the species.
- Run ‘pilgrim --pfn’ to calculate the partition functions. Once the program is executed the directory 3-PLG_OUTPUT/ is created. Inside that directory the user can find the files pfn.E.slevel.txt, pfn.Z.slevel.txt and pfn.ZE_TS.slevel.txt, which contain information about the energetics and partition functions of the species labeled as E, Z and ZE_TS, respectively.
- Run ‘pilgrim --rcons FA_ISO’ or simply ‘pilgrim --rcons’. The forward and backward TST rate constants, as well as the equilibrium constants, are calculated and printed in rcons.FA_ISO.slevel.txt, inside 3-PLG_OUTPUT/.
Calculate the CVT/SCT thermal rate constants for the isomerization process. The projection of the frequencies along the path is carried out in Cartesian coordinates. The MEP limits are defined as $-2.0 \, a_0$ and $2.0 \, a_0$ with a stepsize of $0.01 \, a_0$ and Hessian updates every 10 steps. We require a convergence of 0.1% for the SCT transmission coefficient at the lowest temperature.

Steps:

- All the information obtained in WE01 can be used to obtain the CVT/SCT rate constants, so the first step is to copy the directory WE01/ as WE02/.
- Run ‘pilgrim --input’ inside WE02/ to create the pif.path. Change the values of sbw and sfw to $-2.0 \, a_0$ and $2.0 \, a_0$, respectively. Include also the keyword scterr 0.1, which is needed to obtain converged results within 0.1% for the SCT transmission coefficients.
  
  ```
  > add path *
  >> sbw = -2.0
  >> sfw = +2.0
  >> scterr = 0.1
  >> ..
  > ..
  ```

  After exiting the menu, the pif.path and pif.calcs files are created.
- The file pif.calcs does not need further editing because the level by default is HF/STO-3G.
- Run ‘pilgrim --path ZE.TS’ or simply ‘pilgrim --path’. The program launches Gaussian using the information contained in the pif.calcs and pif.path files. Pilgrim uses the Page-McIver algorithm to evaluate the MEP, and from that information calculates the variational $\Gamma^{\text{CVT}}$ and tunneling transmission $\kappa^{\text{SCT}}$ coefficients as well as the threshold correction coefficient, $\kappa^{\text{CVT/CAG}}$. The file path.ZE.TS.001.slevel.txt is written in 3-PLG/OUTPUT/. A restart file called ZE.TS.001.rst containing the information along the MEP is stored in the 4-PLG/RST/ directory. Notice that Pilgrim searches in the 4-PLG/RST/ directory prior launching Gaussian. Therefore, if the ZE.TS.001.rst file exists and contains the MEP, Pilgrim uses that information.
- Notice that due to the scterr the MEP is only calculated between $[-1.5 \, a_0, 1.4 \, a_0]$. 
- Run ‘pilgrim --rcons FA_ISO’ or simply ‘pilgrim --rcons’ to calculate the CVT/SCT thermal rate constants.

---

[iii] Notice that the keyword scterr prevents the MEP to be calculated in the whole defined domain. This keyword stops the MEP calculation if the SCT transmission coefficient converges within a smaller MEP domain. If convergence is not achieved, sbw and sfw keywords should be modified.
Chapter 3. Isomerization of formic acid

3.4 **WE03: CVT/SCT (II)**

Same as WE02 but in this case the frequencies along the MEP are projected using redundant internal coordinates.

**Steps:**

- Copy the directory `WE02/` as `WE03/`, so the restart file `ZE_TS.001.rst` inside the `4-PLG_RST/` directory can be used in this new run. Notice that the MEP does not depend on the coordinates defined to project the normal-mode frequencies along the path.
- The redundant internal coordinates are generated automatically by running `pilgrim --ics` inside the `WE03/` folder. The program looks for a suitable set of internal coordinates that correctly describes the saddle point frequencies. The resulting coordinates are incorporated into the `pif.struc` file.
- By default the program always use internal coordinates if a set of them is provided in the `pif.struc` file. Consequently, `pif.path` does not have to be edited.
- Run `pilgrim --path ZE_TS` or simply `pilgrim --path`. The program restarts the MEP stored in the `4-PLG_RST/` directory. This time the vibrationally adiabatic potential is calculated using projected frequencies obtained from the previously obtained internal coordinates set.
- Run `pilgrim --rcons FA_ISO` to obtain the CVT/SCT thermal rate constants. They are printed in the file `rcons-FA_ISO.slevel.txt`. 
3.5 WE04: CVT/SCT with an isotopic substitution

Same as WE03, but replacing by deuterium the hydrogen atom that moves from $Z$ to $E$ during the isomerization. The ESFILs are the same as for previous runs and the user can reuse the files saved previously (for instance the ones from directory WE01/), but the program has to be executed from the beginning to include the isotopic substitution. Notice that the MEP stored in WE02/ or in WE03/ cannot be restarted because the MEP is mass-dependent.

Steps:

• Copy the directory WE01/ as WE04/.
• Initialize the interactive menu with 'pilgrim --input' and include the isotopic substitutions by typing
  
  ```
  > mod struc Z
  >> iso = D(4)
  >> ..
  ```

  where we are indicating that the index of the hydrogen atom to be replaced is 4. Do the same for $E$ and ZE_TS.
• Use the interactive menu to create the files pif.path and pif.calcs. Set the maximum length of the MEP from $-2.0$ $a_0$ to $2.00$ $a_0$. Include also the keyword scterr 0.1. See WE02 if needed.
• Run 'pilgrim --ics' to incorporate the internal coordinates to pif.struc.
• Run 'pilgrim --pfn' to calculate the partition functions. The following files pfn.E.slevel.txt, pfn.Z.slevel.txt and pfn.ZE_TS.txt are created in 3-PLG_OUTPUT/.
• Run 'pilgrim --path ZE_TS' to evaluate the variational $\Gamma^{CVT}$ and tunneling transmission $\kappa^{SCT}$ coefficients, as well as the threshold correction coefficient, $\kappa^{CVT/CAG}$. The file path.ZE_TS.001.txt is written in 3-PLG_OUTPUT/ and the restart file ZE_TS.001.rst containing the information along the MEP is stored in the 4-PLG_RST/ directory.
• Notice that $\kappa^{SCT}$ is smaller with the isotopic substitution.
• Run 'pilgrim --rcons' to obtain the CVT/SCT thermal rate constants. They are printed in the file rcons.FA_ISO.slevel.txt.

Notice that we have used the same labels for the deuterated species as for the root species. This is not important in this case because the execution with the isotopic substituted species was carried out in a different directory than the one containing the root species. If we desire to run everything in the same directory, new labels for the isotopic species should be provided.
4. The inversion of ammonia

4.1 About this system

In the inversion of ammonia, the pyramidal structure of the molecule passes through a planar transition state structure by an umbrella-type of motion reaching at the end the same pyramidal structure. This symmetric reaction is studied at 100, 200 and 300 K for the root and fully deuterated species. Dual-level (DL) calculations are performed using PM6 as the low-level (LL) and HF/STO-3G as the high level calculations. A frequency scale factor of 1.078 is used for the frequencies obtained at the PM6 level. The MEP limits are defined as $-0.90 \, \text{a}_0$ and $+0.90 \, \text{a}_0$ with a stepsize of 0.005 $\text{a}_0$, and with Hessian updates every 10 steps. The MEP is extended till the SCT transmission coefficient at the lowest temperature is converged within a 1% The initial step along the path is taken using the normal mode eigenvector of the imaginary frequency at the transition state (default). The frequencies along the MEP are projected using redundant internal coordinates.

The WEs for this system are:

- **WE05**: LL and DL CVT/SCT
- **WE06**: LL and DL CVT/SCT calculations for root and deuterated species
- **ESSO**: *Gaussian 09*
- **Level of calculation**: PM6//HF/STO-3G
- **Working temperatures (in K)**: 100, 200 and 300
Chapter 4. The inversion of ammonia

4.2 WE05: LL and DL CVT/SCT

Study the inversion reaction using conventional CVT/SCT.

Steps:

- Create the directory WE05/ (hereafter the working directory for this system).
- Copy the directory UDATA_NH3/ from tests/esfils/ to WE05/ as UDATA/. Inside UDATA/, there are two subdirectories with the names nh3_pyram/ and nh3_planar/.
- Run `pilgrim --gather`. This will create the directories 1-GTS/ and 5-MOLDEN/ with the gts and molden files, respectively. The latter can be visualized with Molden.
- Run `pilgrim --input` and
  - incorporate the frequency scale factor by typing:
    
    ```plaintext
    > mod struc *
    >> freqscal = 1.078
    >> ..
    ```
  - define the chemical reaction with:
    ```plaintext
    > add chem nh3inv : nh3_pyram --> nh3_planar --> nh3_pyram
    ```
  - generate the files required for dual-level calculations:
    ```plaintext
    > add dlevel
    ```
- Run `pilgrim --ics` to incorporate the internal coordinates to the pif.struc file.
- Run `pilgrim --pfn` to calculate the partition functions of the stationary points with the energies calculated at the low level.
- Edit pif.calcs to set PM6 for the low-level MEP calculations (meppoint block).
- Run `pilgrim --path` to build the MEP at the low level.
- Run `pilgrim --rcons` to calculate the CVT/SCT at the low level.
- Before performing the high-level calculations, edit again pif.calcs to set the HF/STO-3G in the two highlevel blocks (the one for nh3_pyram and the one for nh3_planar).
- Run `pilgrim --hlcalc` to perform the high-level calculations. By default the high-level calculations are preformed at the stationary points plus 3 equidistant points at each side of the MEP. The high-level energies are stored in the highlevel.txt file, inside 2-PLG_DATA/.
- Run `pilgrim --pfn --dlevel` to generate the partition functions of the stationary points with the energies calculated at the high level. Notice that for molecules without torsional conformers, the high-level partition functions coincide with the low-level partition functions. In spite of that, this step is mandatory.
- Run `pilgrim --path --dlevel` to build the MEP at the dual-level using the ISPE algorithm.
- Run `pilgrim --rcons --dlevel` to calculate the dual-level CVT/SCT thermal rate constants.
4.3 WE06: LL and DL CVT/SCT for root and deuterated species

Same as WE05 but with the substitution of all hydrogen atoms by deuterium. In this case, we would like to have the results of the root and deuterated species in the same directory, so new labels have to be defined for the deuterated reaction. Because the electronic structure information is the same for the reaction inversion of NH$_3$ and ND$_3$ no more ESFILs are needed. However, new additions to the pif.struc, pif.chem, pif.path and pif.dlevel are needed.

Steps:

- Copy the directory WE05/ as WE06/ (hereafter the working directory).
- Run `pilgrim --input` and
  - add the deuterated species, i.e., nd3.pyram and nd3.planar:
    > mod struc nh3.pyram
    >> copywith D(all H) as nd3.pyram
    > mod struc nd3.planar
    >> copywith D(all H) as nd3.planar
  - define the new chemical reaction:
    > add chem nd3inv : nd3.pyram --> nd3.planar --> nd3.pyram
  - add the new MEP with
    > add path nd3.planar
    and modify the variables as for the case of nh3.planar.
  - consider the isotopic reaction for high-level calculations with
    > add dlevel
- Edit pif.calcs to set PM6 for the low-level MEP calculations (meppoint block).
- Run `pilgrim --pfn` to calculate the partition functions.
- Run `pilgrim --path nd3.planar` to build the MEP for the deuterated reaction at the low level.
- Notice that $r^{\text{SCT}}$ does not converge with the limits defined for the MEP. Increase the MEP limits (sbw=-1.5 and sfw=+1.5) and execute Pilgrim again with --path.
- Run `pilgrim --rcons nd3inv` to calculate the thermal rate constants for the deuterated reaction at the low level.
- Before performing the high-level calculations, edit again pif.calcs to set the HF/STO-3G in the two highlevel blocks (the one for nd3.pyram and the one for nd3.planar).
- Run `pilgrim --hlcalc` to perform the calculations at the high level. The calculations for the root species are already stored, so they are not repeated.
- Run `pilgrim --pfn --dlevel` to calculate the partition functions at the dual level.
- Run `pilgrim --path nd3.planar --dlevel` to build the MEP for the deuterated reaction at the dual level.
- Run `pilgrim --rcons nd3inv --dlevel` to calculate the thermal rate constants for the deuterated reaction at the dual level.
5. Hydrogen abstraction reactions

5.1 About this system

This chapter presents several tests of hydrogen abstraction reaction from ethanol by atomic hydrogen. This process is more complex than previous examples due to the fact that it involves three different reactions, as the hydrogen can be abstracted from ethanol at three different positions:

\[
\begin{align*}
\text{CH}_3\text{CH}_2\text{OH} + H & \xrightarrow{TSA} H_2 + \text{CH}_3\text{CHOH} \quad \text{(Ra)} \\
\text{CH}_3\text{CH}_2\text{OH} + H & \xrightarrow{TSB} H_2 + \text{CH}_2\text{CH}_2\text{OH} \quad \text{(Rb)} \\
\text{CH}_3\text{CH}_2\text{OH} + H & \xrightarrow{TSC} H_2 + \text{CH}_3\text{CH}_2\text{O} \quad \text{(Rc)}
\end{align*}
\]

Ethanol presents some conformational flexibility. It has an anti conformer and two gauche conformers (which are conformational enantiomers). Therefore the anti and one of the gauche conformations have to be supplied, so the program can calculate the MS-HO partition function. All the torsional conformers of the TS for the hydrogen abstraction from the \(\alpha\)-C (TSA), methyl group (TSB) and hydroxyl group (TSC) are supplied in tests/esfils/. There are 3, 9 and 3 torsional conformers of the transition states corresponding to the reactions Ra, Rb and Rc at the HF/STO-3G level of calculation. The rate constant are calculated using MS-TST and MP-CVT/SCT. In all cases the MEP was calculated with a stepsize of 0.010 \(a_o\) and with Hessian update every 10 steps. The MEP limits are set to \(-2.5\ a_o\) and \(+2.5\ a_o\) for all cases and the SCT transmission coefficients are calculated using a convergence criterion of 0.50%. In this case the electronic structure calculations are performed with Orca. Once the rate constants are available a KMC simulation is carried out to obtain the products yield.

In some of the WEs, hindered rotor anharmonicity is also included (calculated with Q2DTor and MSTor programs). The output files of this software are included in the directories ANHAR_Q2DTOR and ANHAR_MSTOR inside tests/esfils/.
Chapter 5. Hydrogen abstraction reactions

The following isotopic substitutions are also considered in these tests:

- **First isotopic substitution**
  \[
  \begin{align*}
  &\text{CD}_3\text{CD}_2\text{OH} + \text{H} \xrightarrow{\text{TSAD}_1} \text{HD} + \text{CD}_3\text{CDOH} \quad \text{(RaD1)} \\
  &\text{CD}_3\text{CD}_2\text{OH} + \text{H} \xrightarrow{\text{TSBD}_1} \text{HD} + \text{CD}_2\text{CD}_2\text{OH} \quad \text{(RbD1)} \\
  &\text{CD}_3\text{CD}_2\text{OH} + \text{H} \xrightarrow{\text{TSCD}_1} \text{H}_2 + \text{CD}_3\text{CD}_2\text{O} \quad \text{(RcD1)}
  \end{align*}
  \]

- **Second isotopic substitution**
  \[
  \begin{align*}
  &\text{CH}_3\text{CH}_2\text{OD} + \text{D} \xrightarrow{\text{TSAD}_2} \text{HD} + \text{CH}_3\text{CHOD} \quad \text{(RaD2)} \\
  &\text{CH}_3\text{CH}_2\text{OD} + \text{D} \xrightarrow{\text{TSBD}_2} \text{HD} + \text{CH}_2\text{CH}_2\text{OD} \quad \text{(RbD2)} \\
  &\text{CH}_3\text{CH}_2\text{OD} + \text{D} \xrightarrow{\text{TSCD}_2} \text{D}_2 + \text{CH}_3\text{CH}_2\text{O} \quad \text{(RcD2)}
  \end{align*}
  \]

- **Third isotopic substitution**
  \[
  \begin{align*}
  &\text{CD}_3\text{CD}_2\text{OD} + \text{D} \xrightarrow{\text{TSAD}_3} \text{D}_2 + \text{CD}_3\text{CDO}_2 \quad \text{(RaD3)} \\
  &\text{CD}_3\text{CD}_2\text{OD} + \text{D} \xrightarrow{\text{TSBD}_3} \text{D}_2 + \text{CD}_2\text{CD}_2\text{OD} \quad \text{(RbD3)} \\
  &\text{CD}_3\text{CD}_2\text{OD} + \text{D} \xrightarrow{\text{TSCD}_3} \text{D}_2 + \text{CD}_3\text{CD}_2\text{O} \quad \text{(RcD3)}
  \end{align*}
  \]

The WEs for this system are:

- **WE07**: MS-TST and KMC
- **WE08**: MP-CVT/SCT with torsional anharmonicity and KMC
- **WE09**: KMC simulation using analytical expressions
- **WE10**: MP-CVT/SCT with isotopic substitutions and anharmonicity

- **ESSO**: Orca
- **Level of calculation**: HF/STO-3G
- **Working temperatures** (in K): 250, 300, 350, 400, 450, 500, 700, 1000, 1500, 2000 and 2500
5.2 WE07: MS-TST and KMC

Steps:

- Create the directory WE07/ (hereafter the working directory for this system).
- Copy the directory UDATA_ETOH/ inside tests/esfils/ to WE07/ as UDATA/.
  Notice that reaction products are not provided. TST rate constants are
  independent of the product species and, even lacking this information about
  products, it is possible to calculate their yield using KMC.
- Run 'pilgrim --gather'. This will create the directories 1-GTS/ and 5-MOLDEN/
  with the gts and the molden files, respectively.
- Run 'pilgrim --input' and add the proper weights for the conformers with
  enantiomers. In the case of TSB and TSC there are conformational enantiomers,
  but in the case of TSA the conformers are configurational and should be
  included when the KMC simulation is carried out:

  > mod struc EtOH
  >> weight(all)=2
  >> ..

  > mod struc TSB
  >> weight(all)=2
  >> ..

  > mod struc TSC
  >> weight(all)=2
  >> ..

  The keyword weight(all) = 2 changes all the weights of structures with $C_1$
  point group symmetry to 2.
- Add the temperatures and the chemical reactions. Products can be defined
  even if their ESFILs are not available. For example:

  > add chem Ra : EtOH + H --> TSA --> H2 + PA

  - Initialize the KMC variables with:

    > add kmc Habstr
    >> pop(Etoh)_0 = 1e14
    >> pop(H)_0 = 1e6
    >> ..

  - Run 'pilgrim --pfn' to generate the MS-HO partition functions for ethanol,
    TSA, TSB and TSC.
  - Run 'pilgrim --rcons' to calculate the MS-TST thermal rate constants.
  - Edit pif.kmc by hand and multiply Ra by two:

    \[ k(Ra) \times 2 \ \text{tst} \]

    This coefficient indicates that there is another channel identical to Ra with
    TSs that are configurational isomers to the ones that are being considered.
  - Run 'pilgrim --kmc Habstr' to obtain the products distribution and the time
    scale of the reaction for the consumption of the limiting reagent. The output
    file is stored as 3-PLG_OUTPUT/kmc.Habstr.slevel.txt.
Chapter 5. Hydrogen abstraction reactions

5.3 WE08: MP-CVT/SCT with torsional anharmonicity and KMC

Steps:

- Copy the directory WE07/ as WE08/ (hereafter the working directory).
- Copy tests/esfiles/ANHAR_Q2DTOR/ in WE08/ as ANHAR/. It contains the Q2DTor output files for EtOH, TSA, TSB and TSC.
- Specify the output files of Q2DTor using the interactive menu. For example:
  ```
  > mod struc Etoh
  >> anharfile = ethanol.q2dtor
  >> ..
  ```
- Use the interactive menu to generate the pif.path and pif.calcs files. As product files are not included, it is important to assert the MEP direction. Check the fwdir keyword in pif.path to be:
  - for TSA: fwdir 8-10 --
  - for TSB: fwdir 4-10 --
  - for TSC: fwdir 9-10 --
  The coordinate corresponds to the H···H bond formation. The -- indicates that this coordinate should decrease in the forward direction.
- Run `pilgrim --ics` to generate the internal coordinates for each of the TS species.
- Run `pilgrim --pfn`. Pilgrim compares the MS-HO partition functions with the one calculated by Q2DTor. If there is an important discrepancy, Pilgrim gives a warning and anharmonicity is not included for that species. If this is the case, the weights of the conformers may be incorrect. Check them!
- Run `pilgrim --path --software orca` to start the calculation of the MEP from each of the TSs in a sequential manner. The calculation can be sped up by running several paths at the same time. For instance:
  ```
  pilgrim --path TSA --software orca
  ```
  runs sequentially the MEPs from each of the conformers of TSA, whereas:
  ```
  pilgrim --path TSA.001 --software orca
  ```
  runs the MEP for conformer 001 of TSA. You can run each MEP in different UNIX shells at the same time.
- Run `pilgrim --rcons` to calculate the anharmonic MP-CVT/SCT thermal rate constants.
- Modify the pif.kmc file to include anharmonic on the MP-CVT/SCT rate constants. The interactive menu can be used:
  ```
  > mod kmc Habstr
  >> k(Ra) * 2 = mpcvtsct
  >> k(Rb) = mpcvtsct
  >> k(Rc) = mpcvtsct
  ```
- Run `pilgrim --kmc` to obtain the products distribution and the time scale of the reaction for the consumption of the limiting reagent.
5.4 WE09: KMC simulation using analytical expressions

Steps:

- Run ‘pilgrim --fit’ in WE08/ to fit the rate constants to the different analytical expressions. Copy the the anharmonic MP-CVT/SCT forward rate constants with the parameters corresponding to the analytic form 4.
- Create the directory WE09/ and copy the files pif.temp, pif.chem and pif.kmc of WE08/ into WE09/.
- In WE09/, modify by hand the pif.kmc file by introducing the parameters for the analytical expressions (just paste the corresponding lines from the --fit execution). Do not forget to multiply the rate constant of the Ra reaction by 2.
- Run ‘pilgrim --kmc’ to obtain the products distribution and the time scale of the reaction for the consumption of the limiting reagent. Results should be very similar to those obtained in WE08.

### WE09: The pif.kmc file

```
start_kmc Habstr
# KMC Parameters
psteps 1000        # print data each nstp steps
volume 1.00E+00    # simulation volume (mL)
timeunits ps        # units for time variable

# Initial (non-zero) populations (number of molecules)
pop(EtOH)_0 1.00e+14
pop(H)_0 1.00e+06

# Selection of the rate constant to use
k(Ra.fw)*2 analytic4 7.9693E-15 3.1652E+03 3.5434E+00 3.0000E+02 2.5697E+02
k(Rb.fw) analytic4 9.3849E-15 4.3467E+03 3.9031E+00 3.0000E+02 2.3777E+02
k(Rc.fw) analytic4 2.1161E-13 1.2487E+03 2.4590E+00 3.0000E+02 1.9332E+02
end_kmc
```
Chapter 5. Hydrogen abstraction reactions

5.5 WE10: MP-CVT/SCT, isotopic substitutions and anharmonicity

Steps:

- Copy the directory WE08/ as WE10/ (hereafter the working directory).
- Remove the ANHAR/ folder inside WE10/. Copy tests/esfils/ANHAR_MSTOR/ in WE10/ as ANHAR/. It contains the MsTor output files.
- Run ‘pilgrim --input’ to create the deuterated species. The atoms to be deuterated are:

<table>
<thead>
<tr>
<th>H</th>
<th>EthOH</th>
<th>TSA</th>
<th>TSB</th>
<th>TSC</th>
</tr>
</thead>
<tbody>
<tr>
<td>iso1</td>
<td>3-7</td>
<td>4-8</td>
<td>4-8</td>
<td>4-8</td>
</tr>
<tr>
<td>iso2</td>
<td>1</td>
<td>9,10</td>
<td>9,10</td>
<td>9,10</td>
</tr>
<tr>
<td>iso3</td>
<td>all H (3-7,9)</td>
<td>all H (4-10)</td>
<td>all H (4-10)</td>
<td>all H (4-10)</td>
</tr>
</tbody>
</table>

For example:

```
> mod struc TSA
>> copywith D(4-8) as TSA_iso1
> mod struc TSA
>> copywith D(all_H) as TSA_iso3
```

- Specify the output files of MSTor to include torsional anharmonicity. For example:

  ```
  > mod struc TSA_iso1
  >> anharfile = tsa_iso1.mstor
  >
  ```

- Calculate the partition functions for all the species.
- Consider the isotopic substitutions inside pif.chem, pif.kmc, pif.path and pif.calcs. To do so, execute Pilgrim with --input. See WE08 if needed.
- Calculate the MEPs for the isotopic transition structures.
- Calculate all the rate constants.
- Obtain the products distribution for each isotopic situation.
6. Solvated HA with ONIOM

6.1 About this system

This Chapter presents an example with the ONIOM algorithm as implemented in Gaussian. We have selected one hydrogen abstraction (HA) reaction by atomic H in ethanol. For the sake of simplicity we will consider the HA at the CH₂ moiety (i.e. Rₐ in the previous examples). Additionally, six explicit water molecules surround the hydroxyl group of ethanol. In this manner, we divide the system into two layers:

- the ‘high layer’, with ethanol and the H atom (atoms 1 to 10) [HF/3-21G];
- the ‘low layer’ with the six water molecules (atoms 11 to 28) [HF/STO-3G].

The WE for this system is:

- WE11: MP-CVT/SCT rate constants with ONIOM
- ESSO: Gaussian
- Level of calculation: ONIOM(HF/3-21G:HF/STO-3G)
- Working temperatures (in K): 250
6.2 **WE11: MP-CVT/SCT rate constants with ONIOM**

As a consequence of the water cluster, the two gauche conformers do not present the same energy. Consequently, both gauche conformers will be considered explicitly. Notice that at the ONIOM(HF/3-21G:HF/STO-3G) level of calculation, the HA only presents two transition state structures. The arrangement with the hydroxyl H in anti to the hydrogen abstraction does not correspond to a transition state structure.

The MEP is calculated with a stepsize of $0.010 \, \text{a}_0$ and with Hessian update every 10 steps. The MEP limits is set to $-2.0 \, \text{a}_0$ and $+2.0 \, \text{a}_0$. The SCT transmission coefficients are calculated using a convergence criterion of 0.5%.

**Steps:**

- Create the directory `WE11/` (hereafter the working directory for this system).
- Copy the directory `UDATA_ONIOM/` inside `tests/esfils/` to `WE11/` as `UDATA/`.
- Run `pilgrim --gather`. This will create the directories `1-GTS/` and `5-MOLDEN/` with the `gts` and the `molden` files, respectively.
- Generate the input files using `pilgrim --input`. Notice that conformer weights should not be modified.
  ```
  > add temp 250
  > add chem HA : EtOH_6W + H --> TS_6W --> H2 + EtO_6W
  > add path TS_6W
  >> sbw = -2.0
  >> sfw = +2.0
  >> scterr = 0.5
  >> ..
  > ..
  ```
- Run `pilgrim --pfn` to generate the MS-HO partition functions.
- Run `pilgrim --ics` to generate a set of internal coordinates.
- Modify `pif.path` to define the ONIOM layers using `oniomh` and `onioml` keywords. Notice that:
  - the `high layer` includes atoms 1 to 10
  - the `low layer` includes atoms 11 to 28
- Modify also `pif.calcs` to perform ONIOM calculations.
- Below, you can find both `pif.path` and `pif.calcs` files. Assert you have modified them correctly.
- Run `pilgrim --path` to calculate the MEP, as well as the VTST transmission coefficients.
- Run `pilgrim --rcons` to calculate the thermal rate constants.
WE11: The **pif.calcs** file

```
start_meppoint TS_6W gaussian
%nproc=1
%mem=1GB
%chk=[Pilgrim_name].chk
%p ONIOM(hf/3-21g:hf/sto-3g)
scf=verytight
NoSymm
[Pilgrim_gradhess]

Input file for MEP calculation

0 2 0 2
[Pilgrim_geometry]

end_meppoint
```

WE11: The **pif.path** file

```
start_mep TS_6W
sbw -2.0000
sfw +2.0000
ds 0.01000
hsteps 10
fwdir 9-10 --
oniomh 1-10
onioml 11-28
end_mep
```
7. Checking results

In this chapter we recollect the value of some magnitudes calculated in the WEs. We encourage the user to compare their results with those shown in the tables, in order to assert that Pilgrim was executed properly.

Notice that:

- $\kappa^{SCT}$ is the SCT transmission coefficient. This value can be found in the output file generated when Pilgrim is executed with the --path option.
- $k^X$ is the rate constant calculated with the $X$ method. This value can be found in the output file generated when Pilgrim is executed with the --rcons option.
- $\chi_P$ is the yield of a given product. This value can be found in the output file generated when Pilgrim is executed with the --kmc option.
- The units for the rate constants shown in the tables are:
  - for WE01 to WE06 $\rightarrow s^{-1}$
  - for WE07 to WE11 $\rightarrow \text{cm}^3 \text{ molecule}^{-1} s^{-1}$
- Some values may differ slightly with regard to those in the tables.

<table>
<thead>
<tr>
<th>Test</th>
<th>Magnitude</th>
<th>T (K)</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>WE01</td>
<td>$k^{TST}$ (forward)</td>
<td>300.0</td>
<td>5.144E+06</td>
</tr>
<tr>
<td></td>
<td>$k^{TST}$ (backward)</td>
<td>300.0</td>
<td>4.236E+09</td>
</tr>
<tr>
<td>WE02</td>
<td>$\kappa^{SCT}$</td>
<td>300.0</td>
<td>2.659E+00</td>
</tr>
<tr>
<td></td>
<td>$k^{CVT/SCT}$ (forward)</td>
<td>300.0</td>
<td>1.368E+07</td>
</tr>
<tr>
<td>WE03</td>
<td>$k^{CVT/SCT}$ (forward)</td>
<td>300.0</td>
<td>1.378E+07</td>
</tr>
<tr>
<td>WE04</td>
<td>$\kappa^{SCT}$</td>
<td>300.0</td>
<td>1.489E+00</td>
</tr>
<tr>
<td></td>
<td>$k^{CVT/SCT}$ (forward)</td>
<td>300.0</td>
<td>4.812E+06</td>
</tr>
</tbody>
</table>
## The inversion of ammonia

<table>
<thead>
<tr>
<th>Test</th>
<th>Magnitude</th>
<th>T (K)</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>WE05</td>
<td>$\kappa_{\text{SCT}}$ (slevel)</td>
<td>100.0</td>
<td>1.001E+05</td>
</tr>
<tr>
<td></td>
<td>$\kappa_{\text{SCT}}$ (dlevel)</td>
<td>100.0</td>
<td>3.228E+12</td>
</tr>
<tr>
<td></td>
<td>$k_{\text{CVT/SCT}}$ (slevel)</td>
<td>100.0</td>
<td>1.438E+08</td>
</tr>
<tr>
<td></td>
<td>$k_{\text{CVT/SCT}}$ (dlevel)</td>
<td>100.0</td>
<td>5.783E+04</td>
</tr>
<tr>
<td>WE06</td>
<td>$\kappa_{\text{SCT}}$ for ND$_3$ (slevel)</td>
<td>100.0</td>
<td>1.441E+04</td>
</tr>
<tr>
<td></td>
<td>$\kappa_{\text{SCT}}$ for ND$_3$ (dlevel)</td>
<td>100.0</td>
<td>3.190E+10</td>
</tr>
<tr>
<td></td>
<td>$k_{\text{CVT/SCT}}$ for ND$_3$ (slevel)</td>
<td>100.0</td>
<td>1.438E+08</td>
</tr>
<tr>
<td></td>
<td>$k_{\text{CVT/SCT}}$ for ND$_3$ (dlevel)</td>
<td>100.0</td>
<td>5.783E+04</td>
</tr>
</tbody>
</table>

## Hydrogen abstraction reactions

<table>
<thead>
<tr>
<th>Test</th>
<th>Magnitude</th>
<th>T (K)</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>WE07</td>
<td>$k_{\text{MS-TST}}$ for Ra</td>
<td>250.0</td>
<td>5.192E-26</td>
</tr>
<tr>
<td></td>
<td>$k_{\text{MS-TST}}$ for Rb</td>
<td>250.0</td>
<td>6.376E-29</td>
</tr>
<tr>
<td></td>
<td>$k_{\text{MS-TST}}$ for Rc</td>
<td>250.0</td>
<td>1.092E-17</td>
</tr>
<tr>
<td></td>
<td>$\chi_{P}$ for PC</td>
<td>2000.0</td>
<td>0.766</td>
</tr>
<tr>
<td>WE08</td>
<td>$\kappa_{\text{SCT}}$ for TSA.001</td>
<td>250.0</td>
<td>2.823E+05</td>
</tr>
<tr>
<td></td>
<td>$\kappa_{\text{SCT}}$ for TSB.001</td>
<td>250.0</td>
<td>1.214E+06</td>
</tr>
<tr>
<td></td>
<td>$\kappa_{\text{SCT}}$ for TSC.001</td>
<td>250.0</td>
<td>1.060E+02</td>
</tr>
<tr>
<td></td>
<td>$k_{\text{MP-CVT/SCT}}$ for Ra</td>
<td>250.0</td>
<td>1.530E-20</td>
</tr>
<tr>
<td></td>
<td>$k_{\text{MP-CVT/SCT}}$ for Rb</td>
<td>250.0</td>
<td>8.075E-23</td>
</tr>
<tr>
<td></td>
<td>$k_{\text{MP-CVT/SCT}}$ for Rc</td>
<td>250.0</td>
<td>5.299E-16</td>
</tr>
<tr>
<td></td>
<td>$\chi_{P}$ for PA</td>
<td>250.0</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>$\chi_{P}$ for PB</td>
<td>250.0</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>$\chi_{P}$ for PC</td>
<td>250.0</td>
<td>0.000</td>
</tr>
<tr>
<td>WE09</td>
<td>$\chi_{P}$ for PA</td>
<td>2500.0</td>
<td>0.206</td>
</tr>
<tr>
<td></td>
<td>$\chi_{P}$ for PB</td>
<td>2500.0</td>
<td>0.157</td>
</tr>
<tr>
<td></td>
<td>$\chi_{P}$ for PC</td>
<td>2500.0</td>
<td>0.637</td>
</tr>
<tr>
<td>WE10</td>
<td>$\kappa_{\text{SCT}}$ for TSA_iso1.001</td>
<td>250.0</td>
<td>9.514E+04</td>
</tr>
<tr>
<td></td>
<td>$\kappa_{\text{SCT}}$ for TSB_iso2.001</td>
<td>250.0</td>
<td>8.244E+04</td>
</tr>
<tr>
<td></td>
<td>$\kappa_{\text{SCT}}$ for TSC_iso3.001</td>
<td>250.0</td>
<td>2.491E+01</td>
</tr>
<tr>
<td></td>
<td>$\chi_{P}$ for PA_iso1</td>
<td>2000.0</td>
<td>0.107</td>
</tr>
<tr>
<td></td>
<td>$\chi_{P}$ for PA_iso2</td>
<td>2000.0</td>
<td>0.189</td>
</tr>
<tr>
<td></td>
<td>$\chi_{P}$ for PA_iso3</td>
<td>2000.0</td>
<td>0.142</td>
</tr>
</tbody>
</table>

## Solvated HA with ONIOM

<table>
<thead>
<tr>
<th>Test</th>
<th>Magnitude</th>
<th>T (K)</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>WE11</td>
<td>$\kappa_{\text{SCT}}$ for TS_6W.001</td>
<td>250.0</td>
<td>1.753E+04</td>
</tr>
<tr>
<td></td>
<td>$\kappa_{\text{SCT}}$ for TS_6W.002</td>
<td>250.0</td>
<td>1.778E+04</td>
</tr>
<tr>
<td></td>
<td>$k_{\text{MS-TST}}$</td>
<td>250.0</td>
<td>1.114E-26</td>
</tr>
<tr>
<td></td>
<td>$k_{\text{MP-CVT/ZCT}}$</td>
<td>250.0</td>
<td>3.618E-24</td>
</tr>
<tr>
<td></td>
<td>$k_{\text{MP-CVT/SCT}}$</td>
<td>250.0</td>
<td>1.955E-22</td>
</tr>
</tbody>
</table>
8 \textit{Pilgrim input files} \hspace{2cm} 51

8.1 The \textit{pif.struc} file
8.2 The \textit{pif.temp} file
8.3 The \textit{pif.chem} file
8.4 The \textit{pif.path} file
8.5 The \textit{pif.calcs} file
8.6 The \textit{pif.kmc} file
8.7 The \textit{pif.dlevel} file
8. *Pilgrim* input files

This Chapter explains in some detail the format and keywords of each of the seven *Pilgrim* input files (*pifs*). Along the description of the *pifs*, some extracts from the WEs described in the previous chapters are included.

Before proceeding, we highlight that the first input file, *pif.struc*, is generated when *Pilgrim* is run with the --gather option. Although the other six input files can be generated from scratch, we highly recommend to create and modify them using the interactive menu. This menu can also be used to modify *pif.struc*.

Notice that in all the input files, with exception of *pif.calcs*, the # symbol precedes a comment.
8.1 The *pif.struc* file

The *pif.struc* files gathers all the structural information of the stationary points (*ctcsp*) in blocks. This file is generated when executing:

```
pilgrim.py --gather
```

Each block, which has a *ctcsp* label, may contain one chemical structure or a cluster of torsional conformers (CTCs). The blocks are limited by the `start_ctc` and `end_ctc` keywords:

```
start_ctc ctcsp
  root  spname
  $keyword $value
  ...
  $keyword $value
end_ctc
```

where *ctcsp* is the name given to the block. Notice that the keyword *root* points towards the corresponding ESFILs and that, under isotopic substitution, the *spname* label should remain unchanged, whereas the *ctcsp* label should be different from the one of the root species. The available keywords are listed in Table 8.1.

Table 8.1: Keywords and their allowed values inside the CTC blocks of the *pif.struc* file. Keywords between [ ] do not have default values and these should be provided by the user.

<table>
<thead>
<tr>
<th>Keyword</th>
<th>Format</th>
<th>Default</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>root</td>
<td>str</td>
<td><em>spname</em></td>
<td>Label of the species inside UDATA/(or inside 1-GTS/); compulsory with iso</td>
</tr>
<tr>
<td>conformer</td>
<td>int * int</td>
<td>001 * 1</td>
<td>1st integer between 001-999, 2nd integer (1 or 2), the conformer weight</td>
</tr>
<tr>
<td>mformu</td>
<td>str</td>
<td>Provided by Pilgrim</td>
<td>The molecular formula</td>
</tr>
<tr>
<td>ch</td>
<td>int</td>
<td>Provided by Pilgrim</td>
<td>The molecular charge</td>
</tr>
<tr>
<td>mtp</td>
<td>int</td>
<td>Provided by Pilgrim</td>
<td>The spin multiplicity</td>
</tr>
<tr>
<td>type</td>
<td>int</td>
<td>Provided by Pilgrim</td>
<td>0 for minimum, 1 for saddle point</td>
</tr>
<tr>
<td>freqscal</td>
<td>float</td>
<td>1.000</td>
<td>Frequency scale factor (also used along the MEP)</td>
</tr>
<tr>
<td>elestate</td>
<td>int float</td>
<td>Provided by Pilgrim</td>
<td>Degeneracy and relative E in hartree</td>
</tr>
<tr>
<td>[ics[.idx]]</td>
<td>str(s)</td>
<td>Provided by the user or by --ics</td>
<td>Definition of internal coordinates (see Table 8.2)</td>
</tr>
<tr>
<td>[iso[.idx]]</td>
<td>str(s)</td>
<td>Provided by the user</td>
<td>Isotopic substitution</td>
</tr>
<tr>
<td>[anharfile]</td>
<td>str</td>
<td>Provided by the user</td>
<td>The name of the MSTor/Q2DTor output file placed inside ANHAR/</td>
</tr>
</tbody>
</table>

The program creates as many blocks as species inside UDATA/ (or 1-GTS/). When dealing with isotopic substitutions, it is possible to add additional blocks, and therefore to add additional species pointing to the same *spname* label:
8.1 The \textit{pif.struc} file

\begin{verbatim}
WE12: Extract of pif.struc

start_ctc H2
  root   H2
  conformer 001 * 1

... end_ctc

start_ctc HD
  root   H2
  conformer 001 * 1
  iso     D(1)

end_ctc
\end{verbatim}

In the example, the molecule of H$_2$ has H$_2$ as the \texttt{spname}. When one of the hydrogen atoms is substituted by deuterium, a new block with the species labeled HD is added. The \texttt{root} keyword indicates that HD is an isotopic derivative of H$_2$.

8.1.1 Conformers and weights

Each of the conformers inside a block is specified using the syntax:

\begin{verbatim}
conformer idx * wgtconf
\end{verbatim}

where \texttt{idx} refers to the index of the conformation in the .gts file inside the 1-GTS/ directory, and \texttt{wgtconf} is an integer number: 1 if the structure does not have a conformational enantiomer, and 2 otherwise. Therefore, \texttt{wgtconf} avoids electronic structure calculations for species that are conformational enantiomers. The program also include additional information (given as comments) as the relative energy of a given conformer and the point group symmetry. The latter can help us to decide if a given structure has conformational enantiomers.

In the case of a molecule with just one equilibrium structure, only one \texttt{conformer} line is used:

\begin{verbatim}
WE05: Extract of pif.struc

start_ctc nh3_pyram
  # conformers
  conformer 001 * 1 # 0.00 kcal/mol, C3v

... end_ctc
\end{verbatim}

For a compound with several conformations (for instance ethanol) the block contains one line per conformation:
For the case of ethanol, `conformer 001` corresponds to one of the gauche conformers and, therefore, the weight of that structure is 2, whereas `conformer 002` is the anti conformer of ethanol without conformational enantiomers.

8.1.2 Electronic states

The keyword `elestate` can be used to define the electronic state to be considered in the electronic partition function. This keyword is followed by the degenerancy of the state and its relative energy with regard to the electronic ground state (in hartree). By default, only the ground state is considered. However, more states can be included just by adding more lines, which may be particularly important for atomic species. For example, for the fluorine atom two states should be specified as:

```
  elestate 4 0.00000000E+00
  elestate 2 0.00184124E+00
```

8.1.3 Internal coordinates

For the case of transition states\textsuperscript{[i]} a set of non-redundant or redundant internal coordinates can be specified with the keyword `ics`. These coordinates can be generated with the option `--ics` (run `pilgrim --help ics` for more information) or can be introduced directly by the user. This option is compulsory if the projection of the frequencies along the MEP is performed in internal coordinates. The keyword accepts as many coordinates as desired, as long as they are defined according to the format described in Table 8.2. For instance, for the transition structure of the ammonia inversion the internal coordinates were defined as:

```
  ics 1 -2 1 -3 1 -4
  ics 2 -1 -3 2 -1 -4 3 -1 -4
  ics _3_4_1

\textsuperscript{[i]}Identified by a `ctcts` label; i.e., a `ctcsp` label with `type = 1`.  
```
In cases with many conformations, it is possible to specify different sets of internal coordinates for different saddle point conformers by adding the index of the conformer to the ics keyword:

Example of two sets of internal coordinates in pif.struc

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ics</td>
<td>1-2</td>
<td>1-4</td>
<td>1-5</td>
<td>1-6</td>
<td>2-3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ics</td>
<td>1.3</td>
<td>8</td>
<td>2</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>ics</td>
<td>1-2</td>
<td>3-9</td>
<td>5-1</td>
<td>2-7</td>
<td></td>
</tr>
<tr>
<td>ics</td>
<td>.002</td>
<td>1-2</td>
<td>1-4</td>
<td>1-5</td>
<td>1-6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ics</td>
<td>.002</td>
<td>1.7</td>
<td>8</td>
<td>2</td>
<td>2.4</td>
</tr>
<tr>
<td>ics</td>
<td>.002</td>
<td>3-2</td>
<td>1-6</td>
<td>7-2</td>
<td>3-9</td>
</tr>
</tbody>
</table>

Table 8.2: Format for the definition of internal coordinates. Each atom, specified by the integers m, n, o, or p is referred using its numbering in the molecule (starting at 1).

<table>
<thead>
<tr>
<th>Type of internal coordinate</th>
<th>Format</th>
<th>Example</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stretch</td>
<td>m-n</td>
<td>1-2</td>
<td></td>
</tr>
<tr>
<td>Linear bend</td>
<td>m-n=o</td>
<td>1-2=3</td>
<td></td>
</tr>
<tr>
<td>Angular bend</td>
<td>m-n-o</td>
<td>1-2-3</td>
<td></td>
</tr>
<tr>
<td>Proper torsion</td>
<td>m-n-o-p</td>
<td>1-2-3-4</td>
<td>n-o is the central bond</td>
</tr>
<tr>
<td>Improper torsion</td>
<td>m-n-o-p</td>
<td>1-2-3-4</td>
<td>p is the central atom</td>
</tr>
</tbody>
</table>

8.1.4 Isotopic substitutions

The pif.struc file also includes at the beginning a list of the most common masses (in amu) for isotopic substitutions (imass block) where additional atomic masses can be included:

Extract of pif.struc

```
start_isomass
D    =    2.0141018
T    =    3.0160493
...  
O18   =    17.9991600
end_isomass
```

It is possible to specify different isotopic substitutions with iso[idx]. If idx is not specified, the program assumes that the numbering of the atoms is the same for all
the conformers. The syntax of the iso line is as follows:

\[
\text{iso} [\text{. idx}] \ \text{imass(atidx(s))} \ ...
\]

where:

- \text{imass} is the name of the isotopic mass defined in the isomass block,
- \text{atidx(s)} are the numerical indexes of the affected atoms; if all the atoms of a type are going to be substituted, all_\text{X} can be used, \text{X} being the atomic symbol.

For example, all the hydrogen atoms of ammonia can be replaced by:

![WE06: Extract of pif.struc](image)

Equivalent to ‘iso D(all_H)’ is:\textsuperscript{[ii]}

\[
\text{iso}.001 \ D(\text{all}\_H) \\
\text{iso} \ D(2,3,4) \\
\text{iso}.001 \ D(2,3,4)
\]

\textsuperscript{[ii]}Notice that labels 2, 3 and 4 refer to the hydrogen atoms.
8.2 The *pif.temp* file

This file contains the list of working temperatures in Kelvin. For example:

<table>
<thead>
<tr>
<th>WE08: <em>pif.temp</em> file</th>
</tr>
</thead>
<tbody>
<tr>
<td>250.00  300.00  350.00  400.00  450.00  500.00  700.00</td>
</tr>
<tr>
<td>1000.00  1500.00  2000.00  2500.00</td>
</tr>
</tbody>
</table>
8.3 The \textit{pif.chem} file

In this file, each (elementary) reaction of interest is defined using the following scheme:

\begin{verbatim}
chemname : reaction_equation
\end{verbatim}

where

- \texttt{chemname} is a label given by the user to identify the reaction, and
- \texttt{reaction_equation} contains the name of reactants, transition state and products separated by arrows (\texttt{-->}).

The names should coincide with the \texttt{ctcsp} labels of the blocks defined in \textit{pif.struc}.\textsuperscript{[iii]}

The available species can be listed with the \texttt{--ls} option. In the case of two reactants (or products), the “+” symbol should be used to separate their names.

For example, in the inversion of ammonia we have:\textsuperscript{[iv]}

\begin{verbatim}
WE05: pif.chem

nh3_inv : nh3_pyram --> nh3_planar --> nh3_pyram
\end{verbatim}

whereas for the hydrogen abstraction from the ethanol molecule by atomic hydrogen:

\begin{verbatim}
WE08: pif.chem

Ra : Etoh + H --> TSA --> PA + H2
Rb : Etoh + H --> TSB --> PB + H2
Rc : Etoh + H --> TSC --> PC + H2
\end{verbatim}

\textsuperscript{[iii]}Actually, if we are not interested in the backward process (from products to reactants), the products do not have to be defined inside \textit{pif.struc}. In such a case, the user should give a name to each product in the definition of the chemical reaction, even if the ESFILs associated with it does not exist. In this manner, the product(s) can still be considered in a further study by kinetics Monte Carlo.

\textsuperscript{[iv]}The calculation of the thermal rate constants of the reactions defined in \textit{pif.chem} already include the rotational symmetry numbers (internal and external) and the weight of conformations (included in the \textit{pif.struc} file). However, there are special cases, as the the ammonia inversion (WE05), in which the calculated rate constant is half the one that ignores the symmetry numbers. This spurious factor of 1/2 is always present in reactions for which the reactants and products are indistinguishable. \textit{Pilgrim} corrects that, by multiplying the calculated rate constants by two. \textit{Pilgrim} also automatically ‘corrects’ the rate constants for bimolecular reactions with one reagent.
8.4 The *pif.path* file

The file *pif.path* contains all the keywords needed to control the calculation of the MEP. The file is organized in `mep` blocks, one for each `ctcts` defined by the user. Each block has the following format:

```
start_mep ctcts
  keyword[.idx] value
  ...
end_mep
```

The available keywords are listed in Table 8.3. Every keyword inside a block is common to all torsional conformers of `ctcts`. However, the user can add the index of a given conformer to a keyword to treat it differently. For instance, in the example below, the MEP is extended till $s = -1.00\ a_0$ instead of to $s = -0.50\ a_0$ in the backward direction for conformer 002.

```
Example of the *pif.path* file

<table>
<thead>
<tr>
<th>start_mep</th>
<th>TASA</th>
</tr>
</thead>
<tbody>
<tr>
<td>sbw</td>
<td>-0.50</td>
</tr>
<tr>
<td>sfw</td>
<td>0.50</td>
</tr>
<tr>
<td>ds</td>
<td>0.01</td>
</tr>
<tr>
<td>hsteps</td>
<td>10</td>
</tr>
<tr>
<td>sbw.002</td>
<td>-1.00</td>
</tr>
<tr>
<td>end_mep</td>
<td></td>
</tr>
</tbody>
</table>
```

By default, *Pilgrim* calculates the MEP using the Page-McIver algorithm, but the Euler algorithm is also available. Notice that the Page-McIver algorithm makes use of Hessian matrices to calculate the MEP and its curvature, whereas the Euler algorithm is exclusively based on the direction of the gradient. However, Hessians along the path are needed anyway to obtain the VTST coefficients. In consequence, the computational effort associated with the Euler and Page-McIver algorithms is similar, but the latter is more accurate. For this reason, we highly recommend using the default algorithm.

8.4.1 Basic keywords

The interactive menu allows modifying the following keywords:

- **sbw**: the maximum extension of the MEP in the backward direction (in bohr).
- **sfw**: the maximum extension of the MEP in the forward direction (in bohr).
- **ds**: the step in the intrinsic reaction coordinate in bohr.
- **hsteps**: the number of steps for the Hessian matrix update.
- **scterr**: defines the convergence criterium for the SCT coefficient (in percentage). If the relative variation in the coefficient is smaller than the specified percentage, the MEP stops before reaching its limits, `(sbw,sfw)`. The keywords
Table 8.3: Available keywords in the `pif.path` input file. Some keywords are explained with more detail in the text. The ‘advanced’ keywords cannot be modified in the input menu and the user has to edit the `pif.path` file.

<table>
<thead>
<tr>
<th>Keyword</th>
<th>Accepted values</th>
<th>Default</th>
<th>Brief description</th>
</tr>
</thead>
<tbody>
<tr>
<td>sbw</td>
<td>float</td>
<td>-0.50</td>
<td>Minimum value of $s$ (in $a_0$)</td>
</tr>
<tr>
<td>sfw</td>
<td>float</td>
<td>+0.50</td>
<td>Maximum value of $s$ (in $a_0$)</td>
</tr>
<tr>
<td>ds</td>
<td>float</td>
<td>0.01</td>
<td>Stepsize along the path (in $a_0$)</td>
</tr>
<tr>
<td>hsteps</td>
<td>integer</td>
<td>10</td>
<td>Number of steps for Hessian update</td>
</tr>
<tr>
<td>paral</td>
<td>no/yes</td>
<td>no</td>
<td>yes to get both sides of the MEP simultaneously</td>
</tr>
<tr>
<td>scterr</td>
<td>float</td>
<td>1.0</td>
<td>See text</td>
</tr>
</tbody>
</table>

**Advanced keywords related to the MEP**

<table>
<thead>
<tr>
<th>Keyword</th>
<th>Accepted values</th>
<th>Default</th>
<th>Brief description</th>
</tr>
</thead>
<tbody>
<tr>
<td>fmdir</td>
<td>ic (++/--)</td>
<td>Provided by Pilgrim</td>
<td>The internal coordinate increases (++); decreases (--) in the forward direction</td>
</tr>
<tr>
<td>cubic</td>
<td>no/yes/float</td>
<td>no</td>
<td>Use cubic first step</td>
</tr>
<tr>
<td>mtype</td>
<td>es/pm</td>
<td>pm</td>
<td>Euler or Page-McIver algorithms to calculate the path</td>
</tr>
<tr>
<td>mu</td>
<td>float</td>
<td>1.0</td>
<td>The scaling mass (in $amu$)</td>
</tr>
<tr>
<td>epse</td>
<td>float</td>
<td>1e-8</td>
<td>See text (in $E_h$)</td>
</tr>
<tr>
<td>epsg</td>
<td>float</td>
<td>1e-4</td>
<td>See text (in $E_h/a_0$)</td>
</tr>
<tr>
<td>eref</td>
<td>float / auto</td>
<td>auto</td>
<td>Defines the ref. energy for the MEP (in $E_h$)</td>
</tr>
<tr>
<td>keeptmp</td>
<td>list of integers</td>
<td></td>
<td>If specified, Pilgrim does not delete previous data in the corresponding temporal directory</td>
</tr>
<tr>
<td>oniomh</td>
<td>list of integers</td>
<td></td>
<td>Defines atoms in ONIOM high layer</td>
</tr>
<tr>
<td>oniomm</td>
<td>list of integers</td>
<td></td>
<td>Defines atoms in ONIOM medium layer</td>
</tr>
<tr>
<td>onioml</td>
<td>list of integers</td>
<td></td>
<td>Defines atoms in ONIOM low layer</td>
</tr>
</tbody>
</table>

**Advanced keywords related to the VTST coefficients**

<table>
<thead>
<tr>
<th>Keyword</th>
<th>Accepted values</th>
<th>Default</th>
<th>Brief description</th>
</tr>
</thead>
<tbody>
<tr>
<td>lowfq</td>
<td>integer/float</td>
<td>[++/--/+-]</td>
<td>Integer: the index of the imaginary frequency; float: the new value for the frequency in cm$^{-1}$; $++$: $s &gt; 0$; $--$: $s &lt; 0$; $+-$: $s &lt; 0 &amp; s &gt; 0$</td>
</tr>
<tr>
<td>cvt</td>
<td>no/yes</td>
<td>yes</td>
<td>Calculates the $\Gamma^{CVT}$ coefficient</td>
</tr>
<tr>
<td>sct</td>
<td>no/yes</td>
<td>yes</td>
<td>Calculates the $\kappa^{CVT/SCT}$ transmission coefficient</td>
</tr>
<tr>
<td>qrc</td>
<td>integer</td>
<td>1 1000</td>
<td>See text</td>
</tr>
<tr>
<td>muintrpl</td>
<td>linear/cubic int</td>
<td>linear 0</td>
<td>Type of interpolation for $\mu_{eff}$ at the TS and at the first points along the MEP</td>
</tr>
<tr>
<td>e0</td>
<td>float / auto</td>
<td>auto</td>
<td>Specifies the lowest energy for the tunneling calculations (in $E_h$)</td>
</tr>
<tr>
<td>useics</td>
<td>no/yes</td>
<td>yes</td>
<td>The projection of the Hessians along the MEP is performed in ics (specified in pif.struct)</td>
</tr>
<tr>
<td>v1mode</td>
<td>grad/hess</td>
<td>grad</td>
<td>Indicates how the $B_{m,F}$ terms are calculated</td>
</tr>
</tbody>
</table>

$\text{epse}$ and $\text{epsg}$ may override $\text{scterr}$, because the program may consider that the MEP is converged.  
- $\text{paral}$: to activate the calculation of both sides of the MEP at the same time (yes). By default this option is deactivated (no).
8.4.2 Advanced keywords

In order to modify or include these keywords, the user should edit the pif.path file.

- **fwdir**: Defines a forward direction based on the increase (++) or decrease (--) of a given internal coordinate. The internal coordinate can be a distance, a bond angle or a dihedral angle. For instance, `fwdir 3-1-2-4 ++` means that the forward direction of the MEP is toward the increase of the dihedral angle containing the 3-1-2-4 atoms. *Pilgrim* provides a value for this variable based on the eigenvector of the mode with the imaginary frequency.
- **mu**: Defines the scaling mass that transforms Cartesian coordinates to mass-scaled coordinates. Its default value is 1.0 amu.
- **cubic**: if set to `no` (the default), the initial step of the path is taken using the eigenvector associated to the imaginary frequency. If set to `yes`, the program calculates two Hessian matrices at both sides of the path and with a stepsize away from the saddle point of $s = \pm 10^{-4}a_o$. This stepsize is the default in `cubic yes`, but it can be changed (for instance `cubic 2E-4`). This two additional Hessian calculations allow the use of a quadratic Taylor expansion for the first step of the MEP.
- **epse**: Stops the MEP if the variation in energy between two consecutive steps is smaller than a given threshold. The default value for this threshold is $1 \times 10^{-8}E_h$.
- **epsg**: Stops the MEP if the module of the gradient in Cartesian coordinates ($|\vec{g}|$) is smaller than the value of this keyword. The default value is $1 \times 10^{-4}E_h/a_o$.
- **eref**: Defines the reference energy for the MEP (in $E_h$). It can be set to a float number or to `auto`. In the last case, the reactants energy is set as the reference if the saddle point of this path has been defined in pif.chem and *Pilgrim* was previously executed with the `--pfn` option. Otherwise, the energy at sbw is used as the reference.
- **keeptmp**: If the MEP calculation is repeated, *Pilgrim* removes the corresponding temporal directory. This keyword avoids this action.
- **oniomh**: Defines the atoms included in the ONIOM high layer. This keyword can only be used when the MEP is calculated with the *Gaussian* package. As an example, in order to include atoms 1, 3 and 5 in the high layer, this keyword should be set to:

```
oniomh 1 3 5```

To consider atoms 1 to 10, it should be:

```
oniomh 1-10```

Notice that both formats can be combined. Thus:

```
oniomh 1-5 10 11```

would include atoms 1 to 5, 10 and 11 in the layer. The user must modify the *Gaussian* template (pif.calcs) properly, so it is compatible with the ONIOM methodology. This keyword is only used by *Pilgrim* to include the
layer identifier when the Cartesian coordinates are written in the *Gaussian*
input file.

- **oniomm**: Defines the atoms included in the ONIOM *medium* layer. See oniomh
  keyword for more information.
- **onioml**: Defines the atoms included in the ONIOM *low* layer. See oniomh
  keyword for more information.
- **lowfq**: Sometimes, the lowest vibrational modes along the MEP may present
  really small (or even imaginary) frequencies. With this keyword, those
  frequencies may be substituted by fixed values. It requires two values: (1) the
  frequency index \( i \) and (2) the frequency limit in \( \text{cm}^{-1} \), \( \omega_L \). It only switches on
  when the \( i \)-th frequency along the MEP, \( \omega_i \), is smaller than \( \omega_L \), i.e.:

  \[
  \text{if } \omega_i < \omega_L \Rightarrow \omega_i \leftarrow \omega_L
  \]

  The user can define as many lowfq lines as needed. The symbols ++, -- and
  +- indicate that the correction is performed for \( s > 0 \), \( s < 0 \) and for any sign
  of \( s \), respectively.
- **qrc**: Considers quantized-reactant-states tunneling calculations. The energy
  levels are determined using the harmonic approximation. It can only be used
  with unimolecular reactions. This keyword should be included when studying
  tunneling effects at very low temperatures. It requires two values: (i) an
  integer that specifies the index of the normal mode to be quantized (notice
  that in *Pilgrim* the normal-mode frequencies are sorted in increasing value
  of frequency and the default value is one, that is the normal-mode with the
  lowest frequency at reactants); (2) the maximum energy level at which this
  quantization is carried out; the default value is 1000.
- **muintrpl**: The effective mass of tunneling at the TS, \( \mu_{\text{eff}}(s = 0) \), cannot be
  calculated at the transition state. It has to be interpolated using values at
  the nearest points along the MEP. The default is linear 0, which means that
  only \( \mu_{\text{eff}}(s = 0) \) is interpolated, and the value is obtained as the average
  of \( \mu_{\text{eff}}(s = s_{-1}) \) and \( \mu_{\text{eff}}(s = s_1) \), being \( s = s_{-1} \) and \( s = s_1 \) the first points
  along the backward and forward direction of the MEP at which a Hessian is
  available. If the user specifies linear 1 the values of \( \mu_{\text{eff}} \) at \( s = s_{-1} \), \( s = 0 \)
  and \( s = s_1 \) are linearly interpolated using the values at \( s = s_{-2} \) and \( s = s_2 \)
  (the second points along the backward and forward direction of the MEP
  at which a Hessian is available. If cubic is used instead of linear, \( \mu_{\text{eff}} \) is
  interpolated using a cubic spline.
- **e0**: Defines the lowest energy at which the tunneling probability is evaluated.
  The user can either define this value (in \( E_0 \)) or can set it to auto. In the last
  case, this value is taken as

  \[
  E_0 = \max\{(E + \text{ZPE})_R, (E + \text{ZPE})_P\}
  \]

  where \( E \) is the total energy, ZPE is the vibrational zero-point energy, and \( R \)
  and \( P \) represent the reactants and the products, respectively. Notice that this
  requires the previous execution of *Pilgrim* with the --pfn option.
- **useics**: If this keyword is set to yes, the Hessian matrix at each point of
  the path is projected using internal coordinates. This option requires the
8.4 The *pif.path* file

definition of a set of internal coordinates in the *pif.struc* file. If set to no, Cartesian coordinates are used.

• *v1mode*: If set to *grad*, the curvature of the path at each point is calculated numerically from the gradient vectors; if set to *hess*, the curvature is obtained from the Hessian matrix. The default value is *grad*, which in general provides more stable results than *hess*. 
8.5 The \textit{pif.calcs} file

The \textit{pif.calcs} file contains the templates for the electronic structure calculations needed by \textit{Pilgrim} in the calculation of the MEP. It also contains a template to perform the high-level calculations. Unlike in other \textit{pif} files, ‘#’ cannot be used to start a comment.

The templates for the MEP calculations are defined inside \texttt{meppoint} blocks:

\begin{verbatim}
start_meppoint ctcts ESSO
   the template
end_meppoint
\end{verbatim}

An example, extracted from WE04, is shown below:

\begin{verbatim}
start_meppoint nh3_plane gaussian
  %nproc=1
  %mem=1GB
  %chk=[Pilgrim\_name].chk
  #p pm6
  scf=verytight
  NoSymm
  [Pilgrim\_gradhess]

  Input file for MEP calculation

  0 1
  [Pilgrim\_geometry]
end_meppoint
\end{verbatim}

Notice that the commands in brackets (in blue) are not \textit{Gaussian} commands, but are indicators to \textit{Pilgrim} to perform different actions when the ESSO is working with a given geometry:

- [Pilgrim\_name]: is substituted by the name of the file generated by \textit{Pilgrim}.
- [Pilgrim\_geometry]: is substituted by the Cartesian coordinates.
- [Pilgrim\_gradhess]: is substituted by the corresponding \textit{Gaussian}/Orca keyword for the calculation of the gradient and/or the Hessian matrix.

The templates for the high-level calculations (including single-point energy calculations for MEP points) are defined inside \texttt{highlevel} blocks:

\begin{verbatim}
start_highlevel ctcsp ESSO
   the template
end_highlevel
\end{verbatim}

An example is:
### WE04: pif.calcs file

```
start_highlevel nh3_plane gaussian
%nproc=1
%mem=1GB
%chk=[Pilgrim_name].chk
#p hf/sto-3g
scf=verytight
NoSymm

Input file for MEP calculation

0 1
[Pilgrim_geometry]

end_highlevel
```
8.6 The pif.kmc file

The pif.kmc file contains the information to perform a simulation of a chemical kinetics mechanism using the kinetic Monte Carlo algorithm. The file can be created with the aid of the interactive menu but it may need further editing by the user. Similar to other input files, pif.kmc consists of kmc blocks, which allow the consideration of different mechanisms to evaluate:

```
start_kmc kinmec1
  keyword  value
  ...
end_kmc

start_kmc kinmec2
  keyword  value
  ...
end_kmc
```

The general syntax for the selection of a chemical reaction inside each kmc block is:

```
k(chemname[.fw/.bw])[*wgtreac] labelrc
```

where each reaction is chosen using its label (chemname) and the values for labelrc are given in Table 8.4. Notice that the chemical reactions must be defined inside pif.chem:

```
WE08: pif.chem

Ra : Etoh + H --> TSA --> PA + H2
Rb : Etoh + H --> TSB --> PB + H2
Rc : Etoh + H --> TSC --> PC + H2
```

Reactions that lead to one product are considered reversible (forward and backward reactions), whereas reactions that lead to two products are considered irreversible (only forward reaction). It is possible to modify these defaults specifying both forward and backward reactions through chemname.fw and chemname.bw, respectively.

Besides of the selection of the thermal rate constants, Table 8.5 shows a list of additional keywords that can be specified. The thermal rate constant can be multiplied by a weight (integer), specified in wgtreac. In the example, the rate constant for the abstraction of one of the hydrogen atoms from the α-C of ethanol (reaction Ra) should be multiplied by two because it is possible to abstract another equivalent hydrogen from the α-C. The transition states for the abstraction of one of the hydrogen atoms are configurational isomers of the transition states resulting from the abstraction of the other hydrogen.\[^{[v]}\]

\[^{[v]}\]Notice that the weight of a given configuration in the pif.struc (wgt.conf) file only applies to conformational isomers, but not to configurational isomers. This situation always occurs when there
The following example shows a case in which KMC uses MP-CVT/SCT thermal rate constants that were previously calculated with Pilgrim:

**WE08: pif.kmc file**

```
start_kmc Habstr
# KMC Parameters
psteps 1000   # print data each nstp steps
volume 1.00E+00  # simulation volume (mL)
timeunits ps  # units for time variable

# Initial (non-zero) populations (number of molecules)
pop(EtOH)_0   1.00e+14
pop(H)_0   1.00e+06

# Selection of the rate constant to use
k(Ra)*2  mpcvtsct
k(Rb)  mpcvtsct
k(Rc)  mpcvtsct
end_kmc
```

It is also possible to specify rate constants that are given by analytical expressions, instead of being calculated by Pilgrim. In such a case, the user must choose the analytic equation and include the corresponding parameters (see Table 8.6). For the case of analytical expressions for thermal rate constants, the final part of the file should be similar to the one of WE09, i.e.,

**WE09: Extract of the pif.kmc file.**

```
k(Ra.fw)*2  analytic4  7.774E-15 3.151E+03 3.555E+00 3.000E+02 2.581E+02
k(Rb.fw)  analytic4  9.430E-15 4.349E+03 3.901E+00 3.000E+02 2.379E+02
k(Rc.fw)  analytic4  2.118E-13 1.247E+03 2.459E+00 3.000E+02 1.945E+02
```

is a hydrogen abstraction from a carbon atom of the type R1-CH2-R2, where the two hydrogen atoms are equivalent and R1 and R2 are two different non-hydrogen substituents. This is because the abstraction of each of the hydrogen atoms leads to two configurational isomers with two different asymmetric carbon atoms. None of the structures generated by the abstraction of one hydrogen atom can interconvert with the structures generated by the abstraction of the other hydrogen atoms, leading to two chiral reaction channels. The weight *wgtreac* is the unity for the hydrogen abstractions from the methyl and hydroxyl groups.
Table 8.4: Different options for labelrc.

<table>
<thead>
<tr>
<th>Rate constant</th>
<th>labelrc</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>TST</td>
<td>tst</td>
<td></td>
</tr>
<tr>
<td>MS-TST/ZCT</td>
<td>mststzct</td>
<td></td>
</tr>
<tr>
<td>MS-TST/SCT</td>
<td>mststsc</td>
<td></td>
</tr>
<tr>
<td>MS-CVT</td>
<td>mscvt</td>
<td></td>
</tr>
<tr>
<td>MS-CVT/ZCT</td>
<td>msctzct</td>
<td></td>
</tr>
<tr>
<td>MS-CVT/SCT</td>
<td>mscvtsct</td>
<td></td>
</tr>
<tr>
<td>MP-TST/ZCT</td>
<td>mptstzct</td>
<td></td>
</tr>
<tr>
<td>MP-TST/SCT</td>
<td>mptstsc</td>
<td></td>
</tr>
<tr>
<td>MP-CVT</td>
<td>mpct</td>
<td></td>
</tr>
<tr>
<td>MP-CVT/ZCT</td>
<td>mpctzct</td>
<td></td>
</tr>
<tr>
<td>MP-CVT/SCT</td>
<td>mpctsc</td>
<td></td>
</tr>
<tr>
<td>Arrhenius</td>
<td>analytic1</td>
<td>Requires two parameters</td>
</tr>
<tr>
<td>van’t Hoff type 1</td>
<td>analytic2</td>
<td>Requires three parameters</td>
</tr>
<tr>
<td>van’t Hoff type 2</td>
<td>analytic3</td>
<td>Requires four parameters</td>
</tr>
<tr>
<td>Truhlar and co.</td>
<td>analytic4</td>
<td>Requires five parameters</td>
</tr>
<tr>
<td>Truhlar and co.</td>
<td>analytic5</td>
<td>Requires five parameters</td>
</tr>
</tbody>
</table>

Table 8.5: Available keywords in the pif.kmc input file.

<table>
<thead>
<tr>
<th>Keyword</th>
<th>Type</th>
<th>Default</th>
<th>Brief description</th>
</tr>
</thead>
<tbody>
<tr>
<td>psteps</td>
<td>integer</td>
<td>1000</td>
<td>Prints the number of molecules of each species every given number of steps.</td>
</tr>
<tr>
<td>volume</td>
<td>float</td>
<td>1</td>
<td>Volume of the reaction vessel (in cm$^3$).</td>
</tr>
<tr>
<td>timeunits</td>
<td>str</td>
<td>ps</td>
<td>Time evolution units: fs, ps, mcs, ms, s, min, hr femtoseconds, picoseconds, microseconds, miliseconds, seconds, minutes and hours, respectively.</td>
</tr>
<tr>
<td>pop(ctcsp)_0</td>
<td>float</td>
<td>0</td>
<td>Initial number of particles of the ctcsp species.</td>
</tr>
</tbody>
</table>

Table 8.6: Available analytical expressions for the rate constants to be entered in the pif.kmc file. Pilgrim fits the expressions analytic3, analytic4 and analytic5 using $T_r = 300$ K

<table>
<thead>
<tr>
<th>label</th>
<th>Equation</th>
<th>Parameters (in input order)</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>analytic1</td>
<td>$k = A e^{-B/T}$</td>
<td>$A$, $B$</td>
<td>$A$: molecule cm$^{-3}$ s$^{-1}$ $B$: K</td>
</tr>
<tr>
<td>analytic2</td>
<td>$k = AT^n e^{-B/T}$</td>
<td>$A$, $B$, $n$</td>
<td>$AT^n$: molecule cm$^{-3}$ s$^{-1}$ $B$: K</td>
</tr>
<tr>
<td>analytic3</td>
<td>$k = A \left( \frac{T}{T_r} \right)^n e^{-B/T}$</td>
<td>$A$, $B$, $n$, $T_r$</td>
<td>$A$: molecule cm$^{-3}$ s$^{-1}$ $B$ and $T_r$: K</td>
</tr>
<tr>
<td>analytic4</td>
<td>$k = A \left( \frac{T+T_0}{T_r} \right)^n e^{-B(T+T_0)/(T^2+T_0^2)}$</td>
<td>$A$, $B$, $n$, $T_r$, $T_0$</td>
<td>$A$: molecule cm$^{-3}$ s$^{-1}$ $B$, $T_r$ and $T_0$: K</td>
</tr>
<tr>
<td>analytic5</td>
<td>$k = A \left( \frac{T+T_0}{T_r} \right)^n e^{-B(T+T_0)/(T^2+T_0^2)}$</td>
<td>$A$, $B$, $n$, $T_r$, $T_0$</td>
<td>$A$: molecule cm$^{-3}$ s$^{-1}$ $B$, $T_r$ and $T_0$: K</td>
</tr>
</tbody>
</table>
8.7 The *pif.dlevel* file

The *pif.dlevel* contains the target structures to be calculated at high-level. The syntax of the file is:

```
spname1
spname2
...
ctcts1 [ . idx ] { auto_nptbw_nptfw}
ctcts2 [ . idx ] { auto_nptbw_nptfw}
...
```

The high-level single-point energies of the stationary points do not depend upon the isotopic substitution, therefore the *spname* labels are used. However, the MEP is mass-dependent, and the points to be calculated along the MEP (given between curly brackets) are preceded by a *ctcts* label. This allows us to specify which MEP is being calculated when there are reactions with isotopic substitutions. The integers *nptbw* and *nptfw* indicate the number of (equally spaced) points to be calculated in the backward and forward directions of the MEP, respectively. For instance:

**WE06: pif.dlevel**

```
nh3_planar { auto_3_3 }
nh3_pyram
```

calculates single-point high energies for *nh3_pyram* (reactant), *nh3_planar* transition state and three equally spaced points along the backward and forward directions of the MEP. The first point to be calculated corresponds to the geometry with the smallest value of *s*, whereas the last point corresponds to the geometry with the highest value of *s*.

It is also possible to specify the values of *s* along the reaction path at which the high-level single-point calculations are going to be performed. The syntax is:

```
ctcts1 [ . idx ] { svalue1, svalue2, ..., svalueN}
```

Notice that the level of the electronic structure calculations is specified in the *pif.calcs* file. Only when the two files are properly modified (*pif.dlevel* and *pif.calcs*), the high-level calculations can be carried out. This is achieved by executing *Pilgrim* with the `--hlcalc` option. High-level energies will be stored in `2-PLG_DATA/highlevel.txt`.

Once the high-level calculations are performed, the dual-level methodology is applied using the interpolated single-point energies (ISPE) algorithm. The MEP is calculated at a low-level, but it is corrected with high-level single-point calculations at some locations along the path. The high-level energies are interpolated using a cubic spline. For a reasonably good interpolation, it is important to provide, apart from the stationary points, the high-level energies of the smallest and highest
Chapter 8. *Pilgrim* input files

values of $s$, and some additional points along the MEP. The program automatically incorporates those points with 2 additional points at each side of the MEP. Although the *pif.dlevel* file that is generated using the interactive menu can be edited, we recommend to use the default values.
9 Bibliography .......................... 73
9.1 General references about methods in Pilgrim.
9.2 Specific references for specific methods in Pilgrim.
9. Bibliography

9.1 General references about methods in Pilgrim.

In this section we list some key references related to the algorithms, procedures and methodologies implemented in Pilgrim.

- General Reviews about Variational Transition State Theory (VTST):

- Other Software for VTST calculations:

- Direct dynamics:
Chapter 9. Bibliography


- **Minimum Energy Path (MEP):**

- **Canonical Variational Transition State Theory (CVT):**

- **Small-Curvature Tunneling (SCT):**

- **Multi-Structural Variational Transition State Theory (MS-VTST):**
9.1 General references about methods in Pilgrim.


- *Multi-Path Variational Transition State Theory (MP-VTST):*

- *Kinetic Monte Carlo (KMC):*

- *Dual-level Calculations by Interpolated Single-Point Energies (ISPE):*

- *Internal Coordinates:*
Application to $^{12}$CO$_2$ and $^{13}$CO$_2$.

• Symmetry Numbers:

• Molecular Symmetry:

• Anharmonicity of Coupled Torsions:
9.1 General references about methods in Pilgrim.


- Electronic Structure Software (ESSO):
9.2 Specific references for specific methods in Pilgrim.

In this section we give specific suggestions of citations for methods used.

- Conventional Transition State Theory

- Canonical Variational Transition State Theory

- Minimum Energy Path (MEP) in Isoinertial Coordinates

- Generalized Normal Modes in Nonredundant Internal Coordinates Along a Reaction Path

- Generalized Normal Modes in Redundant Internal Coordinates Along a Reaction Path

- Quantized-reactant-states tunneling
  - S. E. Wonchoba, W.-P. Hu, D. G. Truhlar in *Theoretical and Compu*
9.2 Specific references for specific methods in *Pilgrim*.


- **Small Curvature Tunneling**

- **Multi-Structural VTST (MS-VTST)**

- **Multi-path VTST (MP-VTST)**

- **MS-T Method for Torsional Anharmonicity**

- **Q2DTor Method for Torsional Anharmonicity**