TimsPy: access timsTOF Pro data easily from Python

Mateusz K. Łącki¹, Sven Brehmer², Ute Distler¹, Stefan Tenzer¹
¹University Medical Center of the Johannes Gutenberg University Mainz, ²Bruker Daltonic GmbH

- timsTOF Pro is a novel instrument by Bruker
- uses trapped ion mobility spectrometry (TIMS) coupled with liquid chromatography (LC) and mass spectrometry (MS)
- collected data is stored in vendor's format accessible with freely available software development kit from many programming languages
- data dimensions include: retention time (RT)
  1/K₀ (inverse ion mobility IM)
  mass/charge (MZ)

TimsPy gives you simple access to this data, the way data-scientists love it:

```python
from timspy import TimsDIA
D = TimsDIA('path/to/data_folder.d')
D[frames, scans]
```

Above, `frames` select some frames and `scans` some scans, like on the right:

```text
<table>
<thead>
<tr>
<th>frames</th>
<th>scans</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>
```

Of course, we are talking of gigabytes of data. To minimize RAM usage, TimsPy offers iterators:

```python
it = D.iter[1:10001, 0:918]
```
and so on, until 10000th frame is reached.

frames and scans can also be very general expressions, covering broad use-cases

```python
D[1:5, [33, 50]]
D[[1,2,10],[1:1000]]
D[[1,2,10],[1,50]]
D[[rt > 10 and rt < 50], 1:599]
```

And if you want to directly use physical quantities in the query? We have you covered. Use:

```python
D.phys[RT, IM] or D.physIter[RT, IM]
```

We still experiment with usage of VAEX and HDF5 to optimize your daily experience with timsTOF data on your laptop. You don’t need a server or a supercomputer to study your data. Follow our github page!

- Give it a go now! It’s as simple as:
  ```bash
  pip install timspy
  ```