there is no document that describes the entire file format, but there is open-source code that reads or writes it. In principal the files contains these sections:
- header with version and row count
- column properties assigning properties to columns identified by names. These include
  - special types like idcode (encoded chemical structure), rxncode (encoded reaction), descriptors
  - parent-child relationships, e.g. to associate a descriptor column to a parent idcode column
  - detail type and access information in case the column contains references to details (images, html code)
  - others
- the TAB delimited data table itself
- optionally encoded embedded details
- optionally encoded row lists
- optionally embedded macros
- the datawarrior template, which includes information about open views and filters