**NMR-STAR d**ata model and BMRB entry format conventions

|  |  |
| --- | --- |
| **Table 2. NMR-STAR d**ata model and BMRB entry format conventions | |
|  | **Entry (file) level conventions** |
| **1** | NMR-STAR follows a subset of the STAR specification |
| **2** | Entry files consist of a single data block that begins with a data\_<ASCII string without white space> |
|  |  |
|  | **Save frame (object) conventions** |
| **1** | Save frames are roughly equivalent to an object |
| **2** | The category for each save frame is defined by the value given to the tag <category>.Sf\_category found in the save frame |
| **3** | Each save frame in an NMR-STAR file has a tag (<category>.ID) that specifies an integer ID value for the save frame starting at ‘1’ and incremented for each save frame of the same category |
| **4** | An additional special tag within the save frame captures the name (framecode) given to the save frame so that this information can be retained in a relational database derived from the NMR-STAR files and an NMR-STAR file for an entry can be regenerated from the relational database with the appropriate save frame names |
| **5** | For relational integrity and to be able to reconstruct a BMRB entry NMR-STAR file from the relational database every loop construct within a save frame contains a tag that is a foreign key pointing to the save frame primary ‘ID’ tag that is the ‘object’ ID. These tags are constructed with the loop category first followed by a period then followed by the category of the save frame tags followed by an underscore followed by ‘ID’. <loop\_category>.<saveframe ID tag category>\_ID |
|  |  |
|  | **Loop (table) conventions** |
| **1** | Nested loops are not allowed, although they are part of the full STAR specification |
| **2** | Loops are started with the keyword ‘loop\_’ and terminated with the keyword ‘stop\_’ |
| **3** | Loops that contain a single set of data values are expressed as loops and not as a list of tag value pairs outside of a loop construct |
| **4** | Tags used in a loop are all of the same category, which is unique within a save frame instance |
| **5** | All loops are located inside save frames |
|  |  |
|  | **Tag and data item conventions** |
| **1** | All tokens (tag plus data) are encapsulated within save frames |
| **2** | Tags may be either in or outside loops, but not both (single data row loops are still represented as loops) |
| **3** | All tags consist of two parts, ‘category’ and ‘item’ separated by a ‘.’. The ‘category’ translates to the name of a relation or table in a relational schema. The ‘item’ portion translates to the name of an attribute in a relational schema. |
| **4** | Every save frame tag category and every loop tag category contains an <category>.Entry\_ID tag that takes the BMRB entry accession code as its value for relational integrity in the relational database |
| **5** | Within a save frame, all tags not part of a loop are of the same category |
| **6** | For relational integrity, every NMR-STAR tag category (table) contains a primary key |
| **7** | Tags may or may not have enumerated values and the enumeration may be open or closed. Tags that have values not included in the dictionary enumeration list are valid if the tag enumeration is defined as open. |
| **8** | In almost all cases a tag that ends with the word ‘label’ or ‘Label’ takes as its value a STAR pointer to a save frame where these are strings starting with the ‘$’ character. |
|  |  |
|  | **Data value conventions** |
| **1** | In some cases, a value cannot be found for a tag defined as mandatory in the ontology. A value of ‘na’ meaning ‘not available’ or ‘not applicable’ is used as a placeholder. A ‘?’ is the STAR specified value. |
| **2** | Tags that are Boolean in nature take values of ‘yes’ or ‘no’ |
| **3** | The null value is represented using a single unquoted period character, as defined by the STAR specification |