Table 24.2.5.1. *Quick reports available from the NDB*

Report name	Contents
NDB status	Processing status information
Cell dimensions	Crystallographic cell constants
Primary citation	Primary bibliographic citations
Structure identifier	Identifiers, descriptor, coordinate availability
Sequence	Sequence
Nucleic acid sequence	Nucleic acid sequence only
Protein sequence	Protein sequence only
Refinement information	*R* factor, resolution and number of reflections used in refinement
Nucleic acid backbone torsions (NDB)	Sugar–phosphate backbone torsion angles using NDB residue numbers
Nucleic acid backbone torsions (PDB)	Sugar–phosphate backbone torsion angles using PDB residue numbers
Base-pair parameters (global)	Global base-pair parameters calculated using *Curves* 5.1 (Lavery & Sklenar, 1989)
Base-pair step parameters (local)	Local base-pair step parameters calculated using *Curves* 5.1
Groove dimensions	Groove dimensions using Stoffer & Lavery definitions from *Curves* 5.1

validation projects was the creation of the force constants and restraints that are now in common use for crystallographic refinement of nucleic acid structures (Parkinson *et al.*, 1996). The program *SFCHECK* (Vaguine *et al.*, 1999) is used to validate the model against the structure-factor data. The *R* factor and resolution are verified and the residue-based features are examined with this program. Once an entry has been processed satisfactorily, it is entered into the database.

24.2.4. The database

The core of the NDB project is a relational database in which all of the primary and derived data items are organized into tables. At present, there are over 90 tables in the NDB, with each table containing five to 20 data items. These tables contain both experimental and derived information. Example tables include: the citation table, which contains all the items that are present in literature references; the cell_dimension table, which contains all items related to crystal data; and the refine_parameters table, which contains the items that describe the refinement statistics.

Interaction with the database is a two-step process (Fig. 24.2.4.1). In the first step, the user defines the selection criteria by combining different database items. As an example, the user could select all B-DNA structures whose resolution is better than 2.0 Å, whose *R* factor is better than 0.17, and which were determined by the authors Dickerson, Kennard, or Rich. Once the structures that meet the constraint criteria have been selected, reports may be written using a combination of table items. For any set of chosen structures, a large variety of reports may be created. For the example set of structures given above, a crystal-data report or a backbone torsion-angle report can be easily generated, or the user could write a report that lists the twist values for all CG steps together with statistics, including mean, median and range of values. The constraints used for the reports do not have to be the same as those used to select the structures. Some examples of reports from the NDB are given in Fig. 24.2.4.2.

24.2.5. Data distribution

Data are made available *via* a variety of mechanisms, such as ftp and the World Wide Web. Coordinate files, reports, software programs and other resources are available *via* the ftp server (ndbserver.rutgers.edu). In addition to links to the ftp server, the web server provides a variety of methods for querying the NDB and accessing reports prepared from the database (http://ndbserver.rutgers.edu/).

24.2.5.1. *Archives*

The NDB archives, a section of the web site, contain a large variety of information and tables useful for researchers. Prepared reports about the structure identifiers, citations, cell dimensions and structure summaries are available and are sorted according to structure type. The dictionaries of standard geometries of nucleic acids as well as parameter files for *X-PLOR* (Brünger, 1992) are also available. The archives section links to the ftp server, providing coordinates for the asymmetric unit and biological units in PDB and mmCIF formats, structure-factor files, and coordinates for nucleic acid structures determined by NMR.

24.2.5.2. *Atlas*

A very popular and useful report is the NDB Atlas report page. An Atlas page contains summary, crystallographic and experimental information, a molecular view of the biological unit and a crystal-packing picture for a particular structure. Atlas pages are created directly from the NDB database (Fig. 24.2.5.1). The Atlas entries for all structures in the database are organized by structure type on the NDB web site.

24.2.5.3. *NDB searches*

A web interface was designed to make the query capabilities of the NDB as widely accessible as possible. To highlight the special features of NDB, the interface operates in two modes. In the quick search/quick report mode, several items, including structure ID, author, classification and special features, can be limited either by entering text in a box or by selecting an option from the pull-down menu. Any combination of these items may be used to constrain the structure selection. If none are used, the entire database will be selected. After selecting 'Execute Selection', the user will be presented with a list of structure IDs and descriptors that match the desired conditions. Several viewing options for each structure in this list are possible. These include retrieving the coordinate files in either mmCIF or PDB format, retrieving the coordinates for the biological unit, viewing the structure with *RasMol* (Sayle & Milner-White, 1995), or viewing an NDB Atlas page.

Preformatted quick reports can then be generated for the structures in this results list. The user selects a report from a list of 13 report options (Table 24.2.5.1), and the report is created

Copyright © 2006 International Union of Crystallography

references