```
<?xml version="1.0" encoding="UTF-8" ?>
<PDBx:datablock  datablockName="1XY2"
 xmlns:PDBx="http://deposit.pdb.org/pdbML/pdbx-v0.905.xsd"
 xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
 xsi:schemaLocation=
    "http://deposit.pdb.org/pdbML/pdbx-v0.905.xsd
         pdbx-v0.905.xsd">
 <PDBx:audit_authorCategory>
  <PDBx:audit_author name="Cooper, S."></PDBx:audit_author>
  <PDBx:audit_author name="Blundell, T.L.">
                                      </PDBx:audit_author>
  <PDBx:audit_author name="Pitts, J.E."></PDBx:audit_author>
  <PDBx:audit_author name="Wood, S.P."></PDBx:audit_author>
  <PDBx:audit_author name="Tickle, I.J."></PDBx:audit_author>
 </PDBx:audit_authorCategory>
 <PDBx:cellCategory>
  <PDBx:cell entry_id="1XY2">
   <PDBx:length_a>27.080</PDBx:length_a>
   <PDBx:length_b>9.060</PDBx:length_b>
   <PDBx:length_c>22.980</PDBx:length_c>
   <PDBx:angle_alpha>90.00</PDBx:angle_alpha>
   <PDBx:angle_beta>102.06</PDBx:angle_beta>
   <PDBx:angle_gamma>90.00</PDBx:angle_gamma>
   <PDBx:Z_PDB>4</PDBx:Z_PDB>
  </PDBx:cell>
 </PDBx:cellCategory>
 <PDBx:citationCategory>
  <PDBx:citation id="primary">
   <PDBx:title>Crystal structure analysis of
               deamino-oxytocin: conformational flexibility
               and receptor binding.</PDBx:title>
   <PDBx:journal_abbrev>Science</PDBx:journal_abbrev>
   <PDBx:journal_volume>232</PDBx:journal_volume>
   <PDBx:page_first>633</PDBx:page_first>
   <PDBx:page_last>636</PDBx:page_last>
   <PDBx:year>1986</PDBx:year>
   <PDBx:journal_id_ASTM>SCIEAS</PDBx:journal_id_ASTM>
   <PDBx:country>US</PDBx:country>
   <PDBx:journal_id_ISSN>0036-8075</PDBx:journal_id_ISSN>
   <PDBx:journal_id_CSD>0038</PDBx:journal_id_CSD>
  </PDBx:citation>
 </PDBx:citationCategory>
 <PDBx:computingCategory>
  <PDBx:computing entry_id="1XY2">
   <PDBx:structure_solution>SHELX</PDBx:structure_solution>
   <PDBx:structure_refinement>SHELX-76
                              </PDBx:structure_refinement>
  </PDBx:computing>
 </PDBx:computingCategory>
 <PDBx:database_2Category>
  <PDBx:database_2 database_id="PDB" database_code="1XY2">
                                         </PDBx:database_2>
 </PDBx:database_2Category>
 <PDBx:entityCategory>
  <PDBx:entity id="1">
   <PDBx:type>polymer</PDBx:type>
   <PDBx:src_method>man</PDBx:src_method>
   <PDBx:pdbx_description>OXYTOCIN</PDBx:pdbx_description>
   <PDBx:formula_weight>978.189</PDBx:formula_weight>
   <PDBx:pdbx_number_of_molecules>1
                           </PDBx:pdbx_number_of_molecules>
  </PDBx:entity>
  <PDBx:entity id="2">
   <PDBx:type>water</PDBx:type>
   <PDBx:src_method>nat</PDBx:src_method>
   <PDBx:pdbx_description>water</PDBx:pdbx_description>
   <PDBx:formula_weight>18.015</PDBx:formula_weight>
   <PDBx:pdbx_number_of_molecules>7
                           </PDBx:pdbx_number_of_molecules>
  </PDBx:entity>
 </PDBx:entityCategory>
```

Fig. 5.3.8.2. Sample XML output from the *OpenMMS* XML generator. Lines have been omitted or wrapped to fit the present column width.

addresses this by supplying a library of object-oriented routines implemented in Python (van Rossum, 1991) that are designed to integrate with existing or new applications in an easy way.

The objective of *mmLib* is to build a support platform to handle the increasingly rich data about macromolecular structure

Table 5.3.8.1. *The modules provided by the mmLib toolkit*

mmLib.mmCIF	mmCIF parser
mmLib.PDB	PDB format parser
mmLib.Library	Base chemical library
mmLib.Extensions.CCP4Library	Data retrieval from CCP4 monomer library
mmLib.Elements	Chemical data for elements
mmLib.AminoAcids	Chemical data for amino acids
mmLib.NucleicAcids	Chemical data for nucleic acids
mmLib.Structure	Macromolecular structure model
mmLib.GLViewer	OpenGL visualizer

```
import mmLib
from mmLib.FileLoader import LoadStructure, SaveStructure
    struct = LoadStructure(
                fil = cif,
                format = "PDB",
                build_properties = ("no_bonds",) )
    SaveStructure(
        fil = pdb,
        structure = struct,
        format = "CIF")
```

Fig. 5.3.8.3. A snippet of code illustrating mmCIF/PDB file format conversion with the *mmLib* toolkit.

available to structural biologists. Not only do applications need to be able to handle atomic positions and build appropriate three-dimensional structure representations; but links to and integration with information on sequence, homologous structures, and biochemical, genetic and medical form and function are also demanded from individual program systems. Since much of these data are available from external databases in a variety of formats, *mmLib* will not be restricted to the handling of files in a single format. Its initial release provides support for mmCIF, for the PDB format files that historically have been used for representation of macromolecular structures (Westbrook & Fitzgerald, 2003) and for the MTZ format used by the *CCP4* program suite (Collaborative Computational Project, Number 4, 1994).

Table 5.3.8.1 lists the main modules in the current release. *mmLib.mmCIF* and *mmLib.PDB* are read/write parsers for mmCIF and PDB format files, respectively, which handle file input and output in these formats, and provide support for inspection or modification of such file formats. They are typically used in conjunction with the *mmLib.FileLoader* component to populate the *mmLib.Structure* internal representation of the macromolecular structure. The high-level abstraction of such functionality allows for very succinct programmatic constructs. Fig. 5.3.8.3 illustrates this with a program snippet that (apart from the necessary system calls for file management) achieves the conversion of an mmCIF input file to a PDB format representation. This is sufficiently robust and lightweight to act as an input filter to software already designed for handling PDB format files.

mmLib.Structure represents the internal representation of a molecular structure and is implemented as an object hierarchy with four basic object classes: *Structure*, *Chain*, *Fragment* and *Atom*. The *Fragment* class has subclasses *AminoAcidResidue* and *NucleicAcidResidue*. In order to build a complete representation of a structure, the toolkit may need to load data from an input mmCIF or PDB format file, and also from standard data sets of properties of individual monomers and chemical elements; these standard libraries of chemical properties are provided by the *mmLib.Library* module. The core *mmLib* source includes a limited library of such chemical properties (accessible through the subclasses *mmLib.Elements*, *mmLib.AminoAcids* and *mmLib.NucleicAcids*)