Gromacs - Bug #725

Four letter residues and default index groups

03/28/2011 12:17 PM - Daniel Larsson

Status: Closed
Priority: Low
Assignee: Erik Lindahl
Category: preprocessing (pdb2gmx,grompp)
Target version: 5.0
Affected version - extra info: 4.5.3
Affected version: 4.6

Description
Some forcefields have residues with 4 letters, e.g. amber with N- and C-terminals prepended with an N or a C.

Pdb files written by gromacs tools only write 3 letters. These files are incorrectly parsed by make_ndx and other tools that construct the default index groups. The N- and C-terminal residues are not recognized as being part of a protein.

Either these residues should be written with all four letters (my suggestion, since column 21 in the PDB file format standard is not defined: http://www.wwpdb.org/documentation/format32/sect9.html) or the routine that generate the default groups should be made more intelligent to recognize these as being residues.

Related issues:
Related to Gromacs - Bug #779: PDB files use incorrect names
Closed 07/13/2011

Associated revisions
Revision 43a02a47 - 06/23/2014 05:49 PM - Erik Lindahl
Enable 4-letter resname in PDB output, keeps more pdbinfo.
This still fully adheres to the PDB standard since column 21 is not used by the standard. All common programs (PyMol, VMD, etc) understand the 4-letter format, and programs that only read three letters will still read the same filename as they used to. In particular, this conserves most residue names during pdb<->gro format conversions. We have also killed the non-standard wide pdb format to avoid writing broken PDB files.
Fixes #725. Refs #917.
Change-Id: I9b6b8f2e191acdfb65ca2b5d96f39249cd71ea98

History
#1 - 10/17/2012 11:02 PM - Roland Schulz
- Priority changed from Normal to Low

I think this is a feature request not a bug and thus I change the priority to low.

#2 - 05/22/2014 06:29 PM - Erik Lindahl
- Category set to preprocessing (pdb2gmx,grompp)
This is not really a bug but a limitation of the PDB format; we've never claimed to write more than 3 letters. Since Gromacs-4.6 uses our new force field setup (that does not rely on 4-letter amino acid names in PDB files) that particular problem is no longer a concern, but I'll try to add 4-letter residue names in PDB output for Gromacs-5.0.