**GROMACS - Bug #1928**

**GMXLIB can double list of forcefields in pdb2gmx, making -ff flag not work**

03/25/2016 09:51 PM - Chris Neale

<table>
<thead>
<tr>
<th>Status:</th>
<th>Closed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Priority:</td>
<td>Low</td>
</tr>
<tr>
<td>Assignee:</td>
<td></td>
</tr>
<tr>
<td>Category:</td>
<td>preprocessing (pdb2gmx,grompp)</td>
</tr>
<tr>
<td>Target version:</td>
<td></td>
</tr>
<tr>
<td>Affected version -</td>
<td>Extra info:</td>
</tr>
<tr>
<td>Affected version:</td>
<td>5.1.2</td>
</tr>
</tbody>
</table>

**Description**

In gromacs version 5.1.2, the effect of GMXLIB on the behaviour of pdb2gmx has changed since v4.6.7. Now, defining GMXLIB can cause the -ff flag not to work. Also, an empty GMXLIB gives an error, which I would expect to not be the case since gromacs should fall back to the compiled location (I presume GMXLIB is still a colon delimited list of paths).

```
####### 1: desired behaviour is obtained when GMXLIB is not set:
$ unset GMXLIB
$ gmx pdb2gmx ...

Select the Force Field:
From '/nh/nest/u/cneale/exe/GROMACS/exec/gromacs-5.1.2/serial/share/gromacs/top':
1: AMBER03 protein, nucleic AMBER94 (Duan et al., J. Comp. Chem. 24, 1999-2012, 2003)
2: AMBER94 force field (Cornell et al., JACS 117, 5179-5197, 1995)
3: AMBER96 protein, nucleic AMBER94 (Kollman et al., Acc. Chem. Res. 29, 461-469, 1996)
...

####### 2: explicitly setting GMXLIB to the compile directory gives two copies of everything:
$ echo $GMXLIB
/nh/nest/u/cneale/exe/GROMACS/exec/gromacs-5.1.2/serial/share/gromacs/top
$ gmx pdb2gmx ...

Select the Force Field:
From '/nh/nest/u/cneale/exe/GROMACS/exec/gromacs-5.1.2/serial/share/gromacs/top':
1: AMBER03 protein, nucleic AMBER94 (Duan et al., J. Comp. Chem. 24, 1999-2012, 2003)
2: AMBER03 protein, nucleic AMBER94 (Duan et al., J. Comp. Chem. 24, 1999-2012, 2003)
3: AMBER94 force field (Cornell et al., JACS 117, 5179-5197, 1995)
4: AMBER94 force field (Cornell et al., JACS 117, 5179-5197, 1995)
5: AMBER96 protein, nucleic AMBER94 (Kollman et al., Acc. Chem. Res. 29, 461-469, 1996)
6: AMBER96 protein, nucleic AMBER94 (Kollman et al., Acc. Chem. Res. 29, 461-469, 1996)
...

NOTE: this causes a problem when combined with -ff flag to pdb2gmx:
```
Program: gmx pdb2gmx, VERSION 5.1.2
Source file: src/gromacs/gmxpreprocess/pdb2top.cpp (line 199)
Function: void choose_ff_impl(const char*, char*, int, char*, int)

Inconsistency in user input:
Force field 'amber03' occurs in 2 places, but not in the current directory.
Run without the -ff switch and select the force field interactively.

```
####### 3: setting an empty string for GMXLIB gives an error:
$ GMXLIB=""
$ gmx pdb2gmx ...
```

Program: gmx pdb2gmx, VERSION 5.1.2
Source file: src/gromacs/utility/directoryenumerator.cpp (line 279)
Function: gmx::DirectoryEnumerator::DirectoryEnumerator(const char*, bool)

Assertion failed:
Condition: dirname != NULL && dirname[0] != '\0'
Attempted to open empty/null directory path

Associated revisions
Revision 5fd2b449 - 03/26/2016 06:45 AM - Teemu Murtola
Ignore empty GMXLIB

If the GMXLIB env.var. was set, but empty, it triggered an assert (at least) in pdb2gmx. Now this case is silently handled as if the variable wasn't set at all.

Part of #1928.
Change-Id: i51646ec659d119b20b7cb429f9292bfa5775aaa

Revision 2873d40f - 01/03/2018 12:14 PM - Erik Lindahl
Remove duplications in GMXLIB search paths
Remove entries that are duplicated, or identical to the default search path, to avoid e.g. listing identical force fields multiple times.

Fixes #1928.
Change-Id: Ib3caa116b4a6e1ce504244086832a96da52eba774

History
#1 - 03/26/2016 06:40 AM - Gerrit Code Review Bot
Gerrit received a related DRAFT patchset '1' for issue #1928.
Uploader: Teemu Murtola (teemu.murtola@gmail.com)
Change-Id: i51646ec659d119b20b7cb429f9292bfa5775aaa
Gerrit URL: https://gerrit.gromacs.org/5742

#2 - 03/26/2016 08:03 AM - Teemu Murtola
The error from an empty GMXLIB is addressed in the linked change.

The change in 5.1 compared to earlier releases is that GMXLIB adds search directories instead of overriding the default for everything. This allows you to just put your own forcefield somewhere and point GMXLIB to it, without duplicating the whole share/top/ directory everywhere where you want to have forcefields. The side effect of course is that you will always see the default force fields in pdb2gmx, which by itself shouldn't be any problem.

We could possibly make -ff prefer directories from GMXLIB, but other than that, I think the new behavior is more reasonable.

#3 - 03/28/2016 10:58 PM - Chris Neale
Or maybe throw an error if multiple versions are found with the same name? I discourage colleagues from piping numbers into pdb2gmx in their scripts since then one may get an undesired outcome if something has changed in the listing, but it remains a common usage and now maybe another source of related issues is caused?

#4 - 03/30/2016 06:43 PM - Mark Abraham
- Description updated

#5 - 03/30/2016 08:00 PM - Teemu Murtola
I just now took a look at the exact text in the description. For this particular case that you set GMXLIB to the path that is already the default, we could possibly add some filtering that ignores duplicates in GMXLIB and/or duplication with the default data directory, but the workaround for these is really easy: just remove the duplicates manually. So I'm not sure how much effort is warranted here, although this will be relatively easy. The question is just whether those duplicates should be ignored silently, with some noise, or with a fatal error...

The case where it is less clear what is the desired behavior if you set GMXLIB to a path where you have forcefields by the same name as in the Gromacs installation. Options:

1. Current behavior: -ff refuses to select any, interactive selection only allows you to select the ones from your custom folder.
2. Make this unconditionally a fatal error. This might be too much of a limitation. In particular, if people have earlier just copied a full forcefield and then did their own additions/modifications there, this would render those unusable before they are renamed.

02/19/2020
3. Make -ff prefer forcefields from GMXLIB, and allow it to be used if a unique match is found in GMXLIB. Leave interactive behavior as it is.
4. Add a -no-default-ff option that explicitly ignores the default force fields, so that -ff would work the same as above when this is specified, and the interactive behavior would also change.

Any opinions?

#6 - 12/31/2017 11:22 AM - Erik Lindahl
- Status changed from New to Closed

The present code now handles this, but only in the sense that we detect multiple identical names and refuse to proceed - unfortunately not until you have selected it, though. The main reason for this is another shortcoming in different library code where we would always try to read from the first such-names directory (of all the GMX library and GMXLIB ones), rather than carrying along a full directory path.

We will make this more flexible in a future replacement to pdb2gmx, but in the interest of not creating other bugs we won't do it here, so I'll close this.

#7 - 12/31/2017 11:25 AM - Erik Lindahl
- Status changed from Closed to In Progress

Or, on second thought - we can probably fix the problem where their paths are identical at least.

#8 - 12/31/2017 12:14 PM - Gerrit Code Review Bot
Gerrit received a related patchset '1' for Issue #1928.
Uploader: Erik Lindahl (erik.lindahl@gmail.com)
Change-Id: gromacs~release-2018~Ib3caa116b4a6e1ce50424408832a96da52eba774
Gerrit URL: https://gerrit.gromacs.org/7404

#9 - 12/31/2017 01:53 PM - Erik Lindahl
- Status changed from In Progress to Fix uploaded

#10 - 01/03/2018 01:16 PM - Erik Lindahl
- Status changed from Fix uploaded to Resolved

#11 - 01/03/2018 01:17 PM - Erik Lindahl
- Status changed from Resolved to Closed