

GROMACS - Bug #2268

v-site construction ignore water fails for charmm force filed

10/10/2017 04:35 PM - Flaviyan Jerome Irudayanathan

Status: Closed	
Priority: Normal	
Assignee:	
Category:	
Target version:	
Affected version - extra info:	Difficulty: simple
Affected version: 2016.4	
Description	
The ignore water logic in gen_vsites.cpp checks if the heavy atom name is "OW" and skips the waters in the input file. But this logic fails in charmm force field where the water's naming conventions are different.	

Associated revisions

Revision 24525608 - 11/06/2017 07:11 PM - David van der Spoel

Fixed check for water in gen_vsites.cpp

Pdb2gmx would break when generating virtual sites if water oxygens were not named OW. Now checking for the atomnumber instead.

Fixes #2268

Change-Id: I326f683e4940ad02351dcbe0c00e266a82b203f6

History

#1 - 10/10/2017 08:04 PM - Berk Hess

- Status changed from New to Feedback wanted

I don't understand your report.

Water is not handled by gen_vsites.cpp. Water models are never handled by v-site conversion code. A water model either has v-sites (like TIP4P) or not (like TIP3P).

#2 - 10/10/2017 08:41 PM - Flaviyan Jerome Irudayanathan

Sorry my report wasn't clear.

I agree with you water is not handled by the vsite code,

if I provide an input file with say TIP3 water, the gen_vsites.cpp should detect and skip the water molecules, this is done in the code as:

```
1809     else if ( /*(nrHatoms == 2) && (nrbonds == 2) && REMOVED this test
1810                DvdS 19-01-04 */
1811                (gmx_strncasecmp(*at->atomname[Heavy], "OW", 2) == 0) )
1812     {
1813         bAddVsiteParam = FALSE; /* this is water: skip these hydrogens */
1814         if (bFirstWater)
1815         {
1816             bFirstWater = FALSE;
1817             if (debug)
1818             {
1819                 fprintf(debug,
1820                        "Not converting hydrogens in water to virtual sites\n");
1821             }
1822         }
1823     }
```

But this fails for CHARMM force field as the TIP3 oxygen is named OH2.

This can be handled at itp level by changing atom names, but still wanted to report it.

#3 - 10/10/2017 09:29 PM - Berk Hess

With can not also check for the name OH2, because also CHARMM lipids have atoms named OH2.
I would think the #hydrogen=2, #bond=2 check is better. But the comment says David removed this. I hope David can explain why he did this.

#4 - 10/10/2017 09:35 PM - Flaviyan Jerome Irudayanathan

Yes in charmm27 lipids have the OH2 name. I can't find non-water OH2 in charmm36.

#5 - 10/10/2017 10:01 PM - David van der Spoel

Interesting how those old comments come back and bite you.

Anyway I do not recall why I did this a while ago, but we definitely need to get rid of hardcoded atomnames in the code wherever we can. How about
if (nrHatoms 2 && nrbonds 2 && at->atomnumber[Heavy] == 8)

#6 - 10/10/2017 10:27 PM - Berk Hess

Do all force fields with water models have atom numbers?

#7 - 10/10/2017 10:51 PM - Flaviyan Jerome Irudayanathan

Not sure if all force-fields have atom numbers
I compiled this:

```
(nrHatoms == 2) && (nrbonds == 2) && (at->atom[Heavy].atomnumber == 8)
```

and the logic works fine to detect water in charmm36 and 27.

#8 - 10/10/2017 10:52 PM - David van der Spoel

Official pdb files have elements in the last column but we are not fussy about reading it. We try to deduct it from the atom name otherwise in
pdbio.cpp.

There is a trade off in reading accuracy of pdb files and giving fatal errors, that is biased somewhat against fatal errors.

#9 - 10/10/2017 10:54 PM - David van der Spoel

Good work Flaviyan, maybe you can check what happens with those lipids?
Obviously you could have been "lucky" with your pdb file.

Should be easy to check other force fields too, but our regression tests will do some.

#10 - 10/10/2017 10:55 PM - Berk Hess

But we don't even know if the old check led to false positives and/or false negatives. Hydrogens can not be converted to vsites when the heavy atom has no bond to another heavy atom. That would be the case for water, but that is not exactly the check that is there. Most water models have no bonds at all, since they use constraints, so those might not end up at the point in the code at all?
So I would think a check that #Hatoms = #bonds would catch water and anything else that can't be converted to vsites. The question is then still if we need to warn or not.
But I must say I don't really understand what types of atom connections can end up in this code.

#11 - 10/11/2017 12:46 AM - Flaviyan Jerome Irudayanathan

Hi David and Berk,

This is my understanding so far kindly correct me if i am wrong:

while constructing virtual sites if water is given in the input it should be skipped and not processed for v-site

the previous logic fails to do this because the check used "OW" atom name and pdb2gmx fails with in different force filed water is named differently

```
Warning: cannot convert atom 12983 H1 (bound to a heavy atom OT with  
2 bonds and 2 bound hydrogens atoms) to virtual site
```

```
Software inconsistency error:  
Undetected error in setting up virtual sites
```

checking for the atomnumber fixed this i.e. started skipping water.

I checked this for a POPC + TIP3 system and seems to work for smoothly for charmm36, 27. (i.e. #v-sites is the same)

the lipid that used OH2 type atom in charmm27 will not be detected by this logic (i.e. a v-site will be created as OH2 is connected only to one hydrogen and therefore #Hatoms will fail)

i am not entirely sure about all other force field as POPC entry is not available in the default rtp for all force fields.

i agree with Berk that the condition is not fool proof, and i am not sure if constrains are processed before v-site construction ?

But i can't also think of cases where it might fail, i.e. satisfying 2-bonds, 2-connected hydrogen and atom-number = 8.

while on the point pdb2gmx process all atoms in the input file as one molecule and writes a very long topology,

say there is one protein, 100 lipids and 10,000 waters, pdb2gmx will process all as one molecule,

I totally understand if that is the intended behavior as claimed by the disclaimer:

```
"The program has limited intelligence"
```

#12 - 10/11/2017 07:09 AM - David van der Spoel

There could be the unlikely case of a CH3-OH2+ molecule, where #nbonds != #H, but we don't want to treat such unexpected things in an automated fashion anyway I would think.

Grepping for OW in the gmxpreprocess directory I find there is more hardcoded stuff in hizzie.cpp, that should be replaced.

#13 - 10/11/2017 07:10 AM - Gerrit Code Review Bot

Gerrit received a related patchset '1' for Issue [#2268](#).

Uploader: David van der Spoel (davidvanderspoel@gmail.com)

Change-Id: gromacs~release-2016~l9d9b26f93cde731b48a364c70116d52eb74eac5

Gerrit URL: <https://gerrit.gromacs.org/7025>

#14 - 10/12/2017 09:09 AM - Gerrit Code Review Bot

Gerrit received a related patchset '1' for Issue [#2268](#).

Uploader: David van der Spoel (davidvanderspoel@gmail.com)

Change-Id: gromacs~master~l326f683e4940ad02351dcbe0c00e266a82b203f6

Gerrit URL: <https://gerrit.gromacs.org/7029>

#15 - 11/07/2017 01:09 AM - David van der Spoel

- Status changed from *Feedback wanted* to *Resolved*

Applied in changeset [245256081501fa96ebba90fc2c0e1548714ae44b](#).

#16 - 12/11/2017 12:16 PM - Erik Lindahl

- Status changed from *Resolved* to *Closed*