**Request for modification of the InChI algorithm**

Due to the necessity to represent polymeric end groups, e.g. 

branched structured repeat units (SRUs), e.g. 

polymeric links, e.g. 

and other fragments, the InChI Trust member FDA requests that the InChI Trust provide experimental support for a pseudo element (\*) in InChI. Please consider the following:

1. Star atoms are already supported in the MDL MOLFILE format
2. FDA developers have successfully added a pseudo element “X” in the InChI table of chemical elements for the purpose of proving the concept
3. Pseudo element “Zz” is already used in InChI 1B to support polymers but it is implicit (not included in the InChI string)
4. There is no need to assign a role to the pseudo atoms in the sub-standard (InChI algorithm). The role (for example “connection point”, “R-group”) can be assigned outside the sub-standard by a super-standard (e.g. SPL) that utilizes the sub-standard
5. The InChI developer (Igor Pletnev) has been consulted and has confirmed that providing a support for one pseudo element is technically feasible; his comments are enclosed below.

The proposed modification should ensure that the following is fulfilled:

1. If the structure has no brackets, each star atom should be interpreted as pseudo element “Zz” that has valency 1 and is last in the table of elements.
2. If the structure has polymeric brackets, the star atoms should be interpreted the same way as it is currently implemented in InChI 1B.
3. Except for polymers, pseudo atoms should be indicated explicitly in the InChI string and numbered canonically using the standard InChI canonicalization algorithm.

**Comments from InChI developer**

I. I support the idea of adding pseudo element to InChI functionality.

In my opinion, this, most probably, will not cause significant interferences with other features and applications of InChI. However, unexpected side effects are possible as always, so extensive testing is desirable (I think, interest from FDA ensures this).

Pseudo element “Zz” is already used internally in experimental InChI (1B) to support polymers, but with a limited technical role and in implicit manner (“transparent” for user). It may have sense to expose it now.

II. Pseudo element atoms must be considered just placeholders, whose actual role(s) may be different and application dependent.

Introducing placeholders may be an important step towards integrating InChI into a number of contexts or modifying its role in existing applications.

One evident example is a “large molecules” context. Current support of regular single-strand polymers is of little hep in case of polynucleotides and polypeptides. However, InChI strings for 20 amino acid or 5 nucleobase CRU’s may easily be integrated into higher-level structured description like FDA’s SPL or HELM. InChI is good for small molecules, and may be good for small fragments of much larger assemblies (which themselves are far better described “outside” InChI... our recent experience with generating InChI strings for proteins of PDB and alike demonatrated some bottlenecks...).

III. Technically, symbol “Zz” seems to be appropriate for several reasons (unfortunately, “\*” is already used in InChI to denote segment repeatitions). Th “Zz” would be considered an univalent pseudo element having the least possible InChI seniority (so the related atoms will always have maximal possible canonical number(s)).

IV. I agree that for regular single-strand polymers (as detected by the presence of polymeric brackets crossing the bonds to pseudo-atom), the “Zz” must be treated as current “star atoms”.

Would “Zz” ends be exposed in polymer InChI strings or not, requires additional special analysis.

V. Upcoming release of InChI Software v. 1.051 seems to be a good moment to introduce pseudo atoms.

This version significantly modifies the details of polymers treatment; fixing previously reported issues concerning selection of preferred SRU, etc., will inevitably change many “polymeric” InChIs, so additional changes associated with appearance of pseudo atoms will be not so dramatic.

VI. Please note that polymer support is still experimental. If pseudo element extension is accepted, it of course should be placed under “InChI=1B/” umbrella.