Adaptations of proposal based on Fred’s suggestions

- Size is now specified by specifically defining the coordinate system and a starting and end value connected by “:” (e.g. Ns0:2rxyz-9) for a 2nm spherical particle

- Alphabetical order of components is kept: As Gerd mentioned, there would be the possibility to define different groups of components (core, shells, ligands,…). However, as shown in the examples of the nanofiber and the nano-sized plane (see slides below), keeping everything in the same group allows extensions to other materials than nanoparticles

Adaptations of proposal based on Fred's suggestions

- Connectivity: The differentiation between molecular species and nano-sized objects (solids) makes sense.
- However, this is implicitly expressed by the use of one length per dimension for molecular species and a starting and end value for nano objects (solids) (but see issue with the nanopores on the crazy slide).
- Do we need to be more specific (see also point on chirality and surface coverage below)? Having a InChI for solid SiO2 and one for the molecular form is contradicting the InChI specification (at least to my knowledge) and the version proposed by Fred is actually wrong (the connection table show binding to atom 4, which does not exist in the molecular formular and would therefore break every software reading in InChIs).

Adaptations of proposal based on Fred's suggestions

- Crystallinity: fcc and other simple forms to define crystallinity as proposed by Fred are probably better for many of us not being experts in crystallography. However, providing space groups provides more information and can be applied to many more structures. Coding these as ascii strings can be done according to https://pscf.readthedocs.io/en/latest/groups.html
- For molecular species but also carbon nanotubes, the definition of crystallinity doesn't make sense. Therefore, it would probably be better to replace this characteristics by surface coverage (ligands) or chirality (tubes). The later was already defined in the prototype and could be just reintroduced (even if I argued to keep it out in the previous proposal). The first can be integrated by a new NInChI layer Nc (coverage, better names are welcome, Nd = decoration for our nanomedicine friends). Units, in what this coverage can be specified need to be agreed on and coded. At the moment, sc = surface coverage (e.g. Nc(100sc) for 100% surface coverage) and m% = mass percentage are used in the examples but only to show how this should work.

Adaptations of proposal based on Fred's suggestions

- Modifiers of numerical values can also be introduced:

Ns0:~2rxyz-9 might be used for 2nm partials with a very broad distribution or when the diameter is assumed based on the manufacturing conditions (protocol, information from provider) but not actually measured.

Ns0:2rxy-9,0:>100lz-9 could be used for fibres where the z dimension is not in the nano range. 0:<100lz-9 could then be used to specify that all dimentsions are in the nano range even if the correct size is not known.

or we even have ranges: Ns0:2rxy-9,0:50<<100lz-9

or ratios: Ns0:2rxy-9,0:(3*rxy)lz-9 or Ns0:2rxy-9,0:(rxy/3)lz-9

- Style guide: Many decisions on the styles are just what come up in my mind first. Some alternatives are also possible:

Ns(rxy0E-9->2E-9),(lz0E-9->50E-9), Ns(0E-9->2E-9)rxy,(0E-9->50E-9)lz

- Please not that the NInChI has to be a pure ASCII string. Therefore, things like subscript or "fancy" writings of space groups are not possible. I hope that the examples show that this is actually not needed (even if it would increase the human readability).

NInChI: silica (20 nm)

Prototype: *NInChI*=0.00.1A/O2Si/c1-3-2/msp/s20d-9/k000/y1

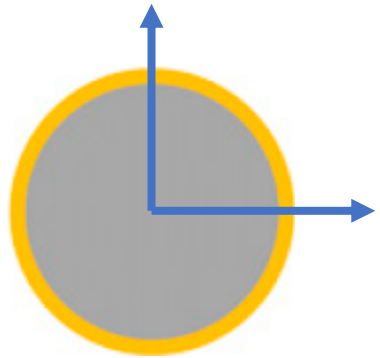
Proposal for standard: *NInChI*=0.00.2A/O2Si/c1-3-2/Nmsp/Ns0:10rxyz-9/Nk(P 32 2 1)/Ny1

Space groups according to https://pscf.readthedocs.io/en/latest/groups.html

NInChI: gold (2 nm) on silica (20 nm)

Prototype: *NInChI*=0.00.1A/Au/msh/s2t10r1-9;12r2-9! /O2Si/c1-3-2/msp/s20d-9/k000/y2&1

Proposal for standard: *NInChI*=0.00.2A/
Au/Nmsp/Ns10:12rxyz-9/Nk(F m -3 m)!
O2Si/c1-3-2/Nmsp/Ns0:10rxyz-9/Nk(P 32 2 1)/Ny2>>1



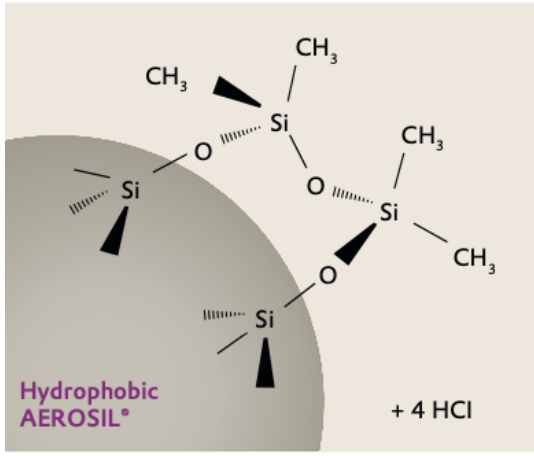
NInChI: silica (20 nm) with explicit hydroxy groups on the surface

Proposal for standard: *NInChI*=0.00.2A/HOSi/c1-2/h1H/Nmsp/Ns10rxyz-9/Nc100sc!
O2Si/c1-3-2/Nmsp/Ns0:10rxyz-9/Nk(P 32 2 1)/Ny2>>1

- Chemistry on the surface is defined by defining a pseudo-molecule with the hydroxy group bound to a pseudo-atom Si
- Molecular nature is indicated by just one radius
- Surface coverage Nc??sc is assumed to be 100%
- Nc100sc could also be replaced by Nc??mr-2 specifying the surface coverage as mass ratio

No guarantee for correctness of InChIs for ligands (created manually)!

NInChI: silica (20 nm) after surface reaction



Proposal for standard: *NInChI*=0.00.2A/
C4H12O2Si4/c1-9(2,5-7)10(3,4)6-8/h1-4H3/Nmsp/Ns10rxyz-9/Nc20sc!
O2Si/c1-3-2/Nmsp/Ns0:10rxyz-9/Nk(P 32 2 1)/Ny2>>1

Proposal for standard: *NInChI*=0.00.2A/
HOSi/c1-2/h1H/Nmsp/Ns10rxyz-9)/Nc80sc!
C4H12O2Si4/c1-9(2,5-7)10(3,4)6-8/h1-4H3/Nmsp/Ns10rxyz-9/Nc20sc!
O2Si/c1-3-2/Nmsp/Ns0E:10rxyz-9/Nk(P 32 2 1)/Ny3>>(1|2)

- Assuming that the reaction leads to 20% surface coverage
- The second version makes it explicit that the rest of the surface is covered by OH.

No guarantee for correctness of
InChIs for ligands
(created manually)!

NInChI: gold (2 nm) hollow shell

Proposal for standard: *NInChI*=0.00.2A//Nmsp/Ns0:10rxyz-9/Nk(0)!
Au/Nmsp/Ns10:12rxyz-9/Nk(F m -3 m)/Ny1>>2

Or to specify that it is actually not empty but filled by something undefined ("un" as InChI but needs checking with the other InChI groups):

NInChI=0.00.2A/un/Nmsp/Ns0:10rxyz-9/Nk(0)!
Au/Nmsp/Ns10:12rxyz-9/Nk(F m -3 m)/Ny1>>2

And then we can put something into the void (even if this probably doesn't make sense in this case):

NInChI=0.00.2A//Nmsp/Ns0:10rxyz-9/Nk(0)!
Au/Nmsp/Ns10:12rxyz-9/Nk(F m -3 m)!
C8H9NO2/c1-6(10)9-7-2-4-8(11)5-3-7/h2-5,11H,1H3,(H,9,10)/Nmmol/Ns0rxyz-9/Ny(1|3)>2

Questions:

- How should we indicate that the paracetamol is floating in the void (mol for molecular is used right now as morphology)?
- Should we give an radius to indicate that it is the innermost nanomaterial layer? Even if it can move around up to were the gold started, giving a second radius should probably be avoided since it makes it look like a nano object (solid)

NInChI Example: carbon nanotube

Prototype: *NInChI*=0.00.1A/C/mtu/s4d-10/w(3,1)/y1

Proposal for standard : *NInChI*=0.00.2/C/Nmfib/Ns20rxy-9,0:100lz-9/Nw(3,1)/y1

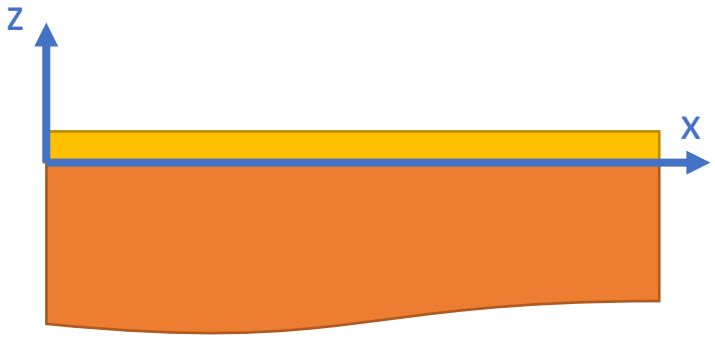
However, we would probably need to specify the void to be consistent:

NInChI=0.00.2/

/Nmfib/Ns0:20rxy-9,0:100lz-9!

C/Nmfib/Ns20rxy-9,0:100lz-9/Nw(3,1)/y1>2

Some crazy ideas:



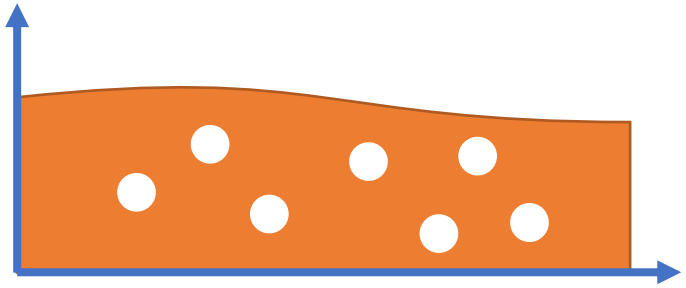
Gold nanolayer on macroscopic silica:

NInChI=0.00.2A/

Au/Nmpl/Ns0:∞lx-9,0:∞ly-9,0:2lz-9/Nk(F m -3 m)!

O2Si/c1-3-2/Nmbu/Ns0:∞lx-9,0:∞ly-9,-∞:0lz-9/Nk(P 32 2 1)/Ny2>>1

Please note the minus sign in the definition of the length in the z direction. This puts the coordination origin at the interface between the silica and gold material layers.



Nanosized holes in gold:

NInChI=0.00.2A/

/Nmbu/Ns2rxyz-9/Nc2vr-1!

Au/Nmpl/Ns0:∞lx-9,0:∞ly-9,0: ∞lz-9/Nk(F m -3 m)/Ny(1|2)

Unfortunately, this is not yet completely consistent with the other NInChIs with ligands. In case here, the single radius means the maximum radius while it is the radius at which the ligands are placed in the other examples (assuming that the size of the ligands is so small that the increase in the radius caused by the ligands is neglectable)

Nc could then give the volume ration to show how many nanoholes are there (20% in this case).