Electronic Supplementary Material (ESI) for RSC Advances.

Electronic Supplementary Information

The Effects of Solvent Composition on the Affinity of a Peptide towards Hair Keratin: Experimental and Molecular Dynamics Data.

*Egipto Antunes, Célia F. Cruz, Nuno G. Azoia, Artur Cavaco-Paulo**

CEB – Centre of Biological Engineering, University of Minho, 4710-057 Braga, Portugal

Keratin protofibril model building.

The majority of the hair weight corresponds to the keratin fiber. However a complete macrofibril or the smaller microfibril reach sizes far beyond the usual systems sizes simulated in molecular dynamics, even with coarse-grained force fields like MARTINI. Considering this, we build a computational model of a truncated protofibril assembling 8 truncated α keratin. The truncated zone corresponds of α-keratins 2B domain from the L2 link to C tail, and the assembling of the 4 dimers is based in the A_{22} model, with dimer to dimer anti-parallel pairing. The size of the model, approximately 20nm of length, was expected to be big enough to reproduce hair keratins properties at a reasonable simulation time. Each of the four dimer are composed for one K34 (type I) and one K86 (type II), since, according the literature, these keratins belong to the group of most expressed keratins in hair, and are presented in the cortex, the most representative zone of the hair. As stated in the Introduction section of main document, there are no atomistic models of entire keratins, so we build one computational model of each of the two chosen keratins with Pymol software[1], from the L2 domain to the C terminal of the chain. The chains were defined, in the software, as having α-helix structure with the exception of the L2 and C terminal domains that have no defined secondary structure. The aminoacid sequence of the K34 and K86 was accessed from UniProt database entry O76011 (http://www.uniprot.org/uniprot/O76011) and O43790 (http://www.uniprot.org/uniprot/O43790), respectively. These all-atoms models were converted to coarse-grained models by the Martinize script (available at http://md.chem.rug.nl/cgmartini), and for each keratin type six replicas were made, and each one was simulated for 65ns (three replicas in water and three in vacuum of each K34 and K86). The short simulation allows the relaxation of the aminoacid sequence, and a more favorable spatial conformation of the keratins chains. All the replicas showed the same behavior,

namely the compaction of the L2 and C tail, forming globular zones at the terminal zones of keratins monomer models, and the 2B domains presented lesser compaction maintaining a coiled structure as expected. In the next step we manually put one K34 and K86 side to side pairing its domains, in water and vacuum to simulate its assembling. Since the behavior of the monomers in the first simulations was similar in all replicas, were tested different combinations, using the resulting conformations of the previous replicas, resulting in three systems of K34/K86 dimer in water (DW1, DW2 and DW3), and three in vacuum (DV1, DV2, and DV3) all with small conformations variations. In this case we choose to build six systems that are very similar to be equivalent to experiment replicas, but with small differences in the spatial conformation that allows a larger sampling of this property. The systems were simulated for 150ns and in this case there are few differences between systems at the end. The systems with the monomers in vacuum, DV1, DV2, and DV3 were excluded from the next simulations due to the structures presented in the end of simulations, in which the coiled domains of the two monomers are not fully aligned, leading to conformations incompatibles with the A_{22} assembling model. The DW1, DW2 and DW3 systems presented similar structure, at the end of the 150ns simulated. The L2 and C domains of each keratin in all systems interact with the same domain of the adjacent keratin, forming globular zones in the truncated dimer terminations. The globular zones do not presented a tendency to any particular structure, and differ between the three systems. In both systems the 2B domain maintained its coiled structure, and interacted with each other, approximating both keratin chains. In the D_W3 system the assembly of the C domains was lightly different, resulting in a separation of C termination domains comparing with the DW1 and DW2 systems. In the available molecular models of other IFs like vimentin[2], or of skin keratins[3], the 2B domains have two points where the helix chains of each keratin turn around the other, and in

the middle the chains are more separated (see Figure 2 in the article). In our non-excluded simulations there are also to points of more close contact and the separation in the middle zone, however due to a lesser turn of one keratin around the other the contacts and separation are lesser evident, and in the DW3 system there was also a separation between the C domains. These differences can be due the fixed secondary structure of the proteins in the MARTINI force field, (can lead to a lesser flexible keratin chain, avoiding bigger turns in the chain) or due to the truncation of the model (can add degrees off freedom). Since the evaluation of molecular structure of keratins protofibrils are out of the scope of this study we considered the dimers models reliable for the study propose.

After the dimer assembling simulations we joined together four dimers in order to build a truncated model of keratin protofibril (PT). We build 8 systems, (4 different systems each one with one replica), using the structure's from previous simulations in water. The systems PT1, PT2, PT3 and its replicas (PT1R, PT2R and PT3R) are composed for 4 equal dimers from the previous DW1, DW2 and DW3, respectively. The remaining system and its replica, PT4 and PT4R, have two dimers from the DW2 and two from DW3 systems. It is important to remember that although the systems are being treated as different, the difference is only relative to its spatial conformations, since all systems have the same constitution (dimers of K34 and K86). Once again in the next simulations are sample several starting points but with little differences to be possible compare the final conformations.

After joining the dimers, the systems were hydrated and its behavior in water was simulated by about 500ns. The simulations showed the approximation between the dimers, resulting, in most cases, in the interaction of the 4 dimers that composed the systems forming a core structure. In the systems PT1, PT1R, PT2, PT2R and PT4R the attachment of the four dimers was full, and

take between 45 to 70ns. However the systems PT3 and PT3R presented, at the end of 500ns, structures not fully compact with some of its keratins chains away of the core structure, in its central domains. The system PT4 also presented some separation between some chains, but at lower level. Although the compactness of the final structures was similar between the replicas, the spatial arrangement of the keratins chains, and the shape of the PT were different. The truncated octamer of the systems PT1, PT2R e PT4R presented a cylindrical shape, and the systems PT1R, PT2 and PT4, in turn, arranged its dimers in an elliptical cylinder shape, as well the systems PT3 and PT3R but these with a broader center. The evaluation of these simulations results is a challenge task due to the poor knowledge at molecular level of the IFs assembling. However we chose the PT structure from the system PT4R to proceed with the interaction simulations, due to more compactness and cylindrical shape presented, similar to the models described in literature for vimentin and intermediate filaments[2–4].

Some considerations:

The PT structural conformations presented in the several simulated systems were different, even between replicas. The time evolution of the dimers attachments and posterior rearrangement of its chains showed that the way the dimers are placed in the systems and the dimer structure itself (that resulted from the dimer assembling simulations), are factors that affect the final structure of the PT model. This is because all the PT systems, in which its keratins chains are not full joined together, had the dimer for the DW3 system that presented a separation between the C domains of the two keratins chains, which seems difficult the full join of the keratins chains in the PT assembling. However this is not the principal influence in the assembling of the dimers, once even between replicas the final structures are different, and the

PT4R system, for example, have DW3 dimers but do not present separation in the C domain zone. Take this is consideration, the system PT4R was considered reliable to model the hair keratin behavior in MD simulations.

References

(1) Schrodinger LLC The PyMOL Molecular Graphics System, Version 1.1, **2010**.

(2) Strelkov, S. V; Herrmann, H.; Geisler, N.; Wedig, T.; Zimbelmann, R.; Aebi, U.; Burkhard, P. *EMBO J.* **2002**, *21*, 1255–66.

(3) Lee, C.-H.; Kim, M.-S.; Chung, B. M.; Leahy, D. J.; Coulombe, P. a *Nat. Struct. Mol. Biol.* **2012**, *19*, 707–15.

(4) Qin, Z.; Kreplak, L.; Buehler, M. J. *PLoS One* **2009**, *4*, e7294.

(5) Marrink, S. J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. *J. Phys. Chem. B* **2007**, *111*, 7812–24.

Attachments

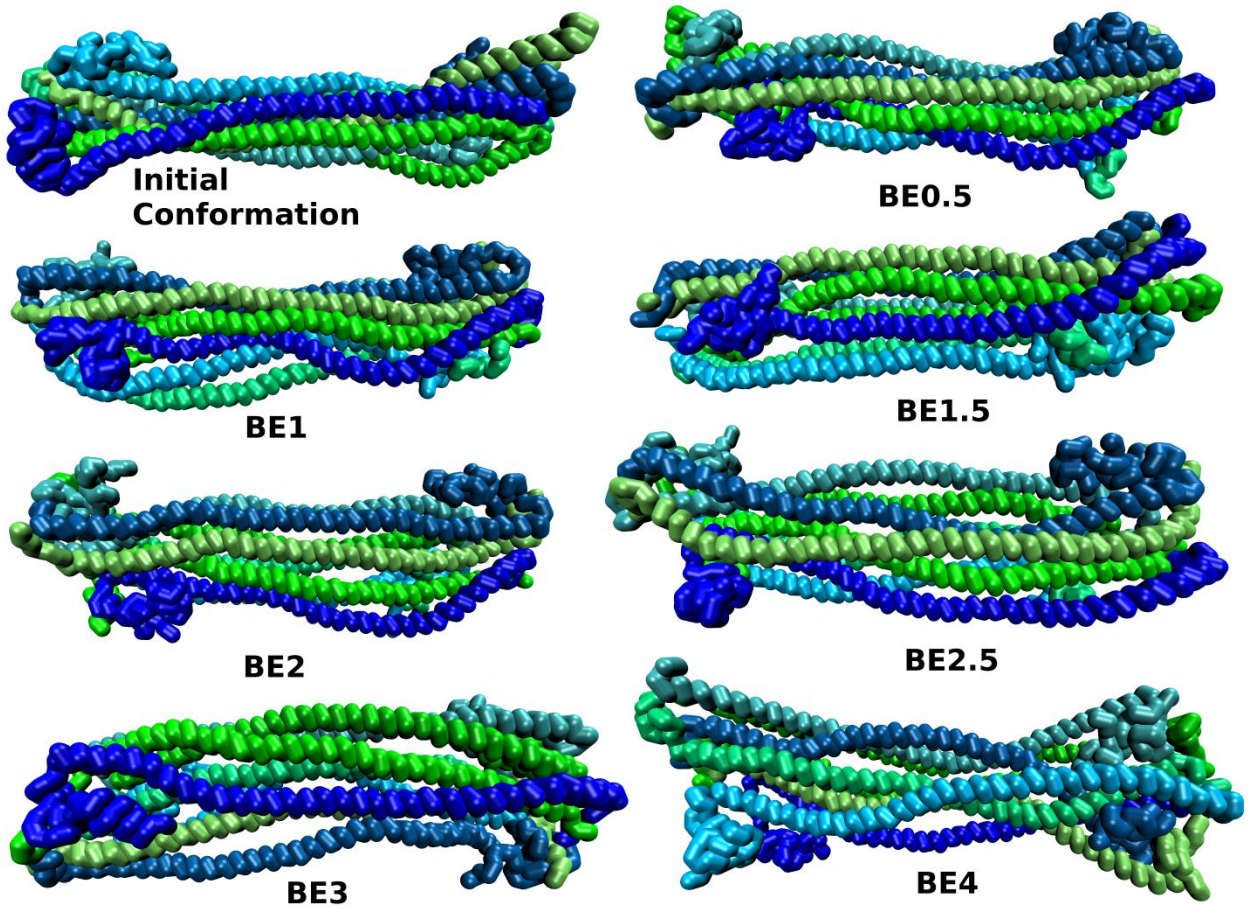


Figure 1 – Initial (up at left) and final (510ns) conformations of the systems in the simulations containing ethanol.

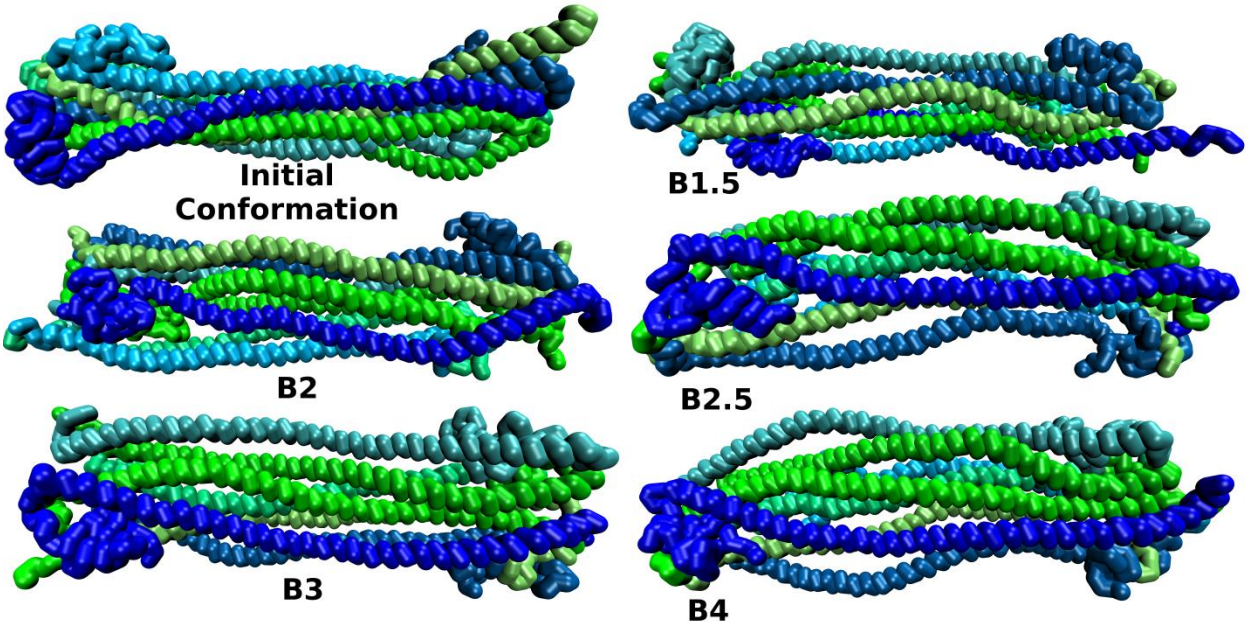


Figure 2 – Initial (up at left) and final (510ns) conformations of the systems in the simulations without ethanol.

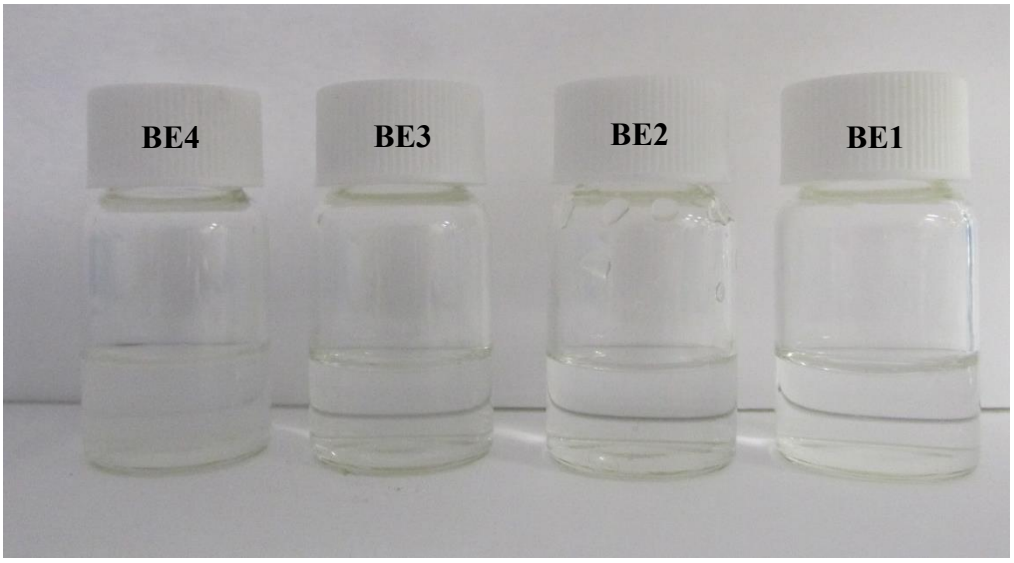


Figure 3 – Picture of the formulations containing 10% ethanol and 4, 3, 2 and 1% of benzyl alcohol (from left to right). To notice the turbidity in the formulation BE4.

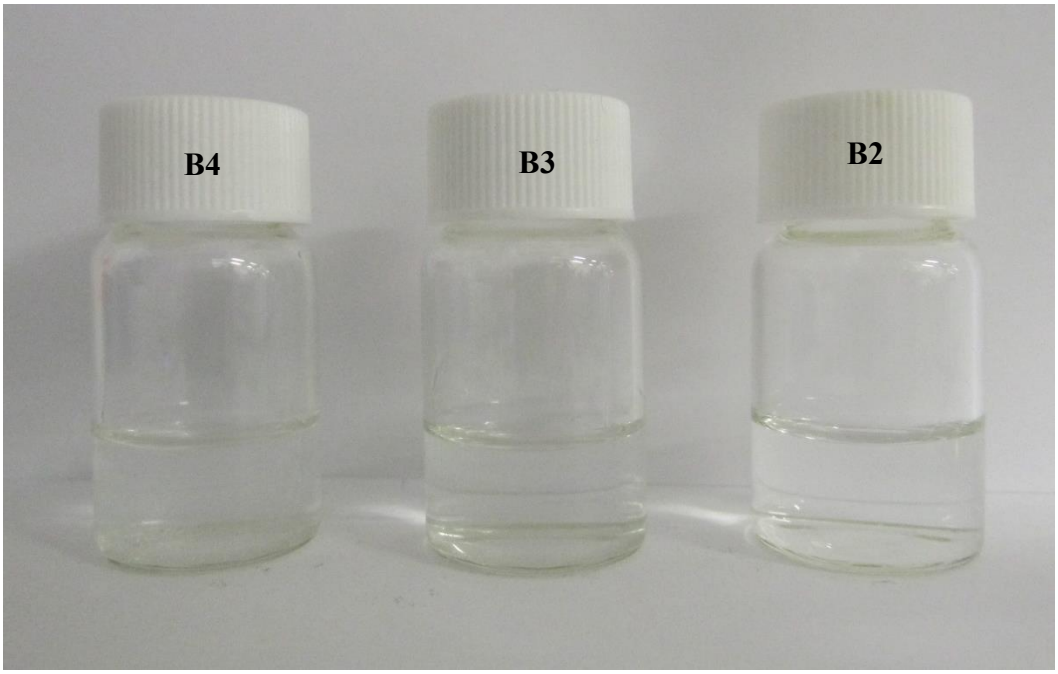


Figure 4 Picture of the formulations containing 4, 3 and 2% of benzyl alcohol (from left to right). To notice the turbidity in the formulation B4.

Notes: The simulation systems of BE4 and B4 (with 4 % (v/v) of benzyl alcohol and 10% of ethanol in the BE4 formulation) were not included in the work presented in the main document. As it is visible in the Figure 3 and 4 the solutions with this benzyl alcohol concentration present some turbidity, likely due to the fact that the 4% concentration is in the limit of its solubility in water (3.5 g/100mL at 20ºC and 4.29 g/mL at 25ºC). The turbidity had impact in the preliminary

fluorescence tests that we made, and we conclude that these formulations do not allow the achievement of reliable results. Because we cannot compare the simulations results of the systems with BE4 and B4 formulations with experimental data, we exclude these simulations from the work, although in the simulations there are no problem with the solubility, showing that the use of molecular dynamics simulations it may be advantageous to test cosmetic compositions, comparing with the use of aqueous solutions in the laboratory.

United Atoms simulations

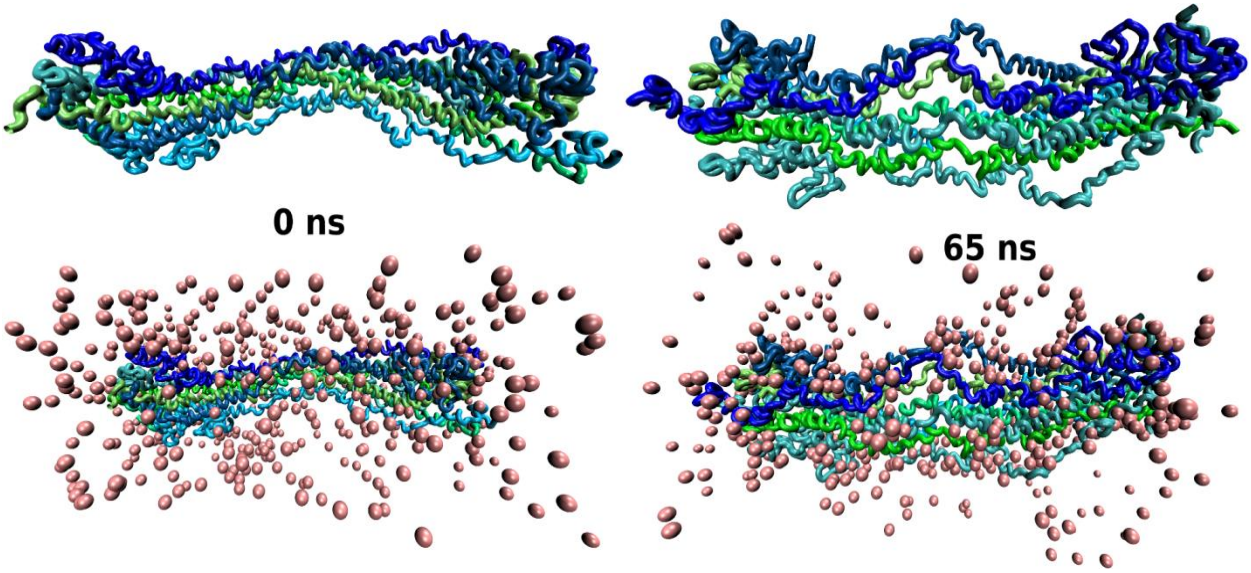
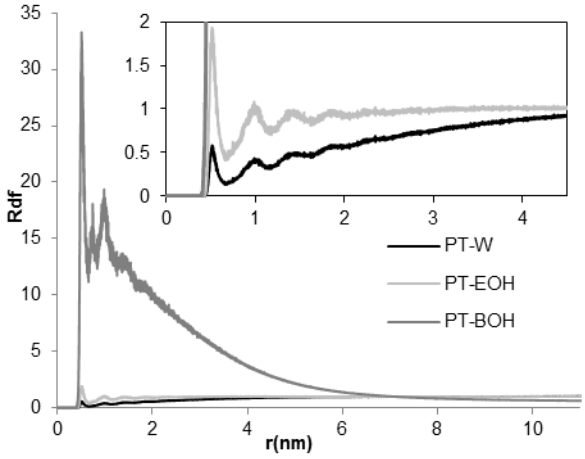


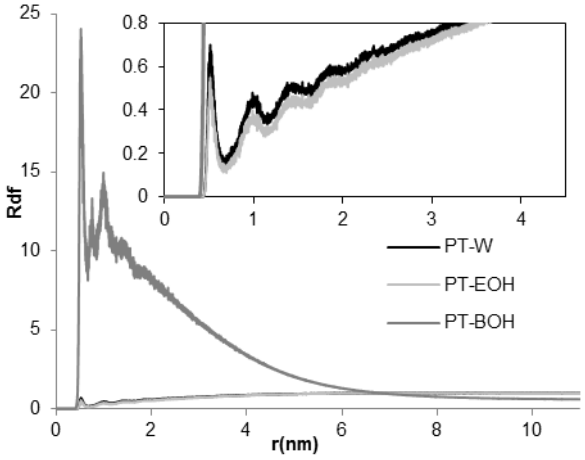
Figure 5 Graphic representations of united atom simulations. The system was composed of one protofibril (the keratins are represented as tubes, each one with different color) in the BE3 formulation. In the upper pictures it is represented the evolution of the protofibril structure over 65ns, and the bottom pictures represent the same time evolution of the protofibril and benzyl alcohol molecules (pink beads). The remaining solvent molecules were excluded from the pictures for clarity. In these simulations it is clear the lateral disassembling of the keratin chains, as well the tendency of benzyl alcohol molecules to be located near to the fiber, such as in the coarse grained simulations.

Radial Distribution function Graphics

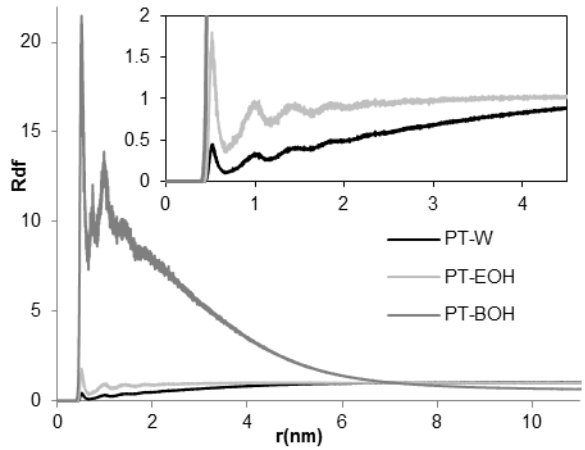
BE0.5



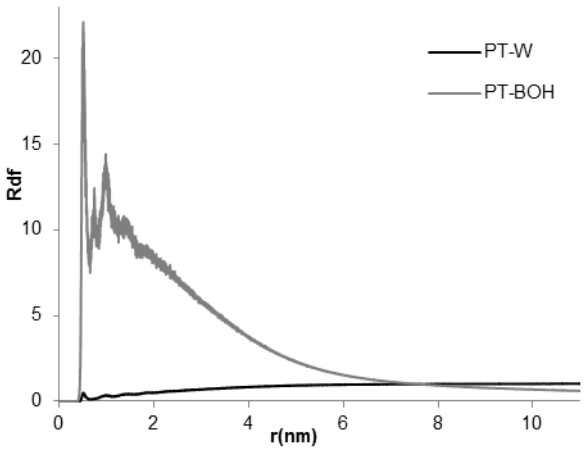
BE1



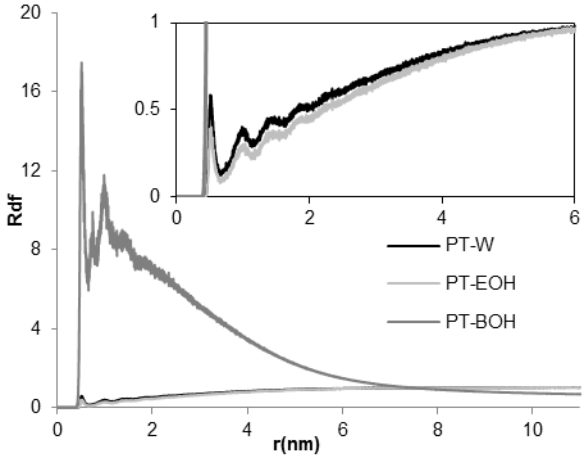
BE1.5



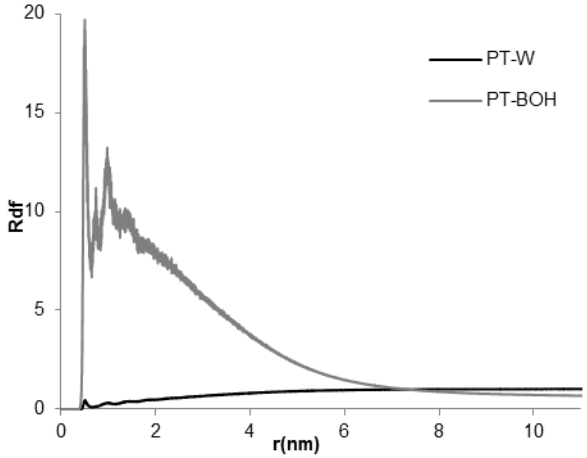
B1.5



BE2



B2



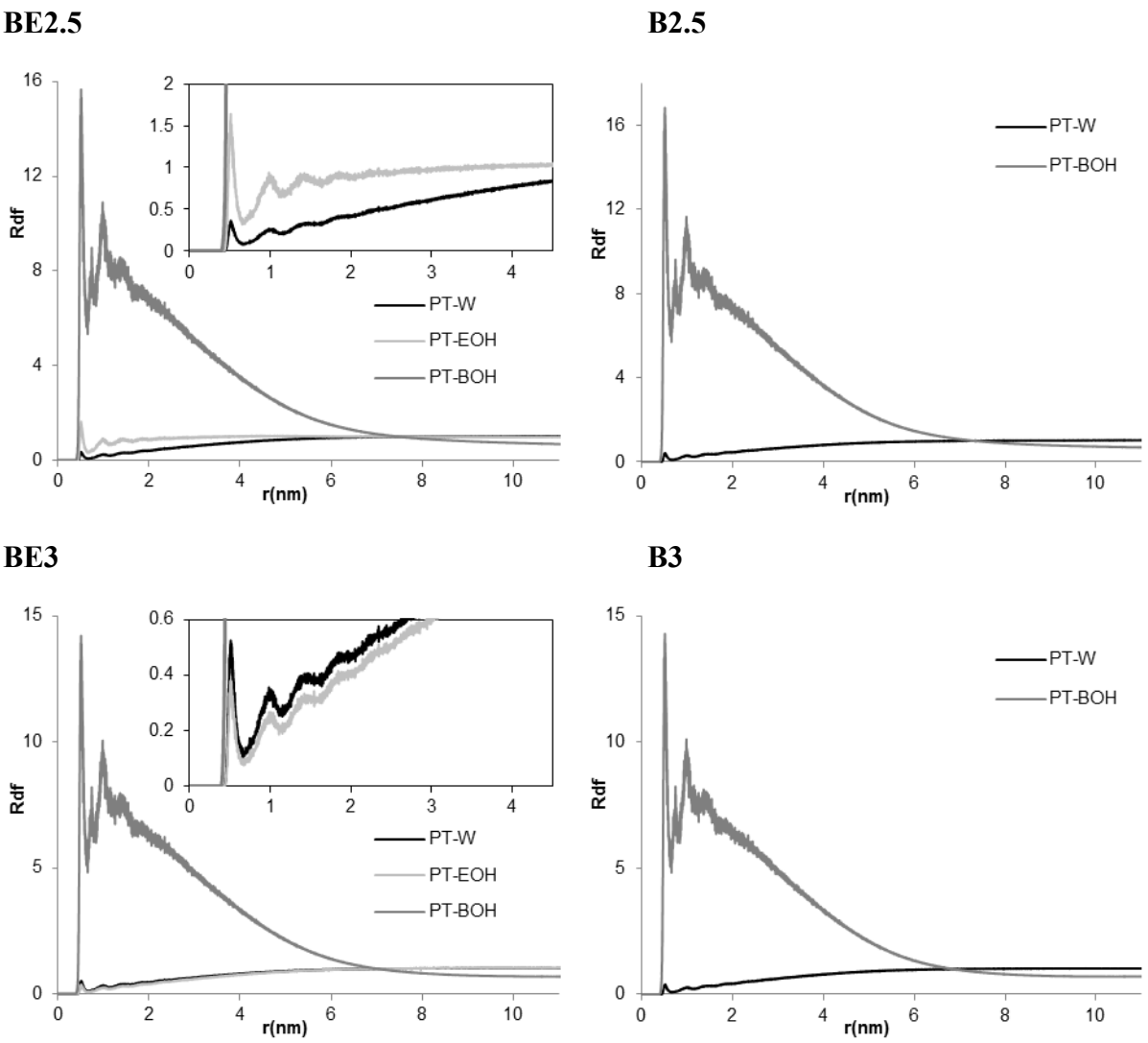


Figure 6 Radial distribution function of the systems simulated. It is clear the tendency of benzyl alcohol molecules to be located near to the protofibril chains.

BOH parametrization

Itp file:

```
[moleculetype]
; molname       nrexcl
  BNOH          1

[atoms]
; id    type    resnr   residu    atom    cgnr    charge    mass
  1     SC5     1       BNOH      R1      1       0         26.0378
  2     SC5     1       BNOH      R2      2       0         26.0378
  3     SC5     1       BNOH      R3      3       0         25.0299
  4     P1      1       BNOH      R4      4       0         31.0343

[constraints]
; i j       funct       length
  1 2       1           0.27
  2 3       1           0.27
  1 3       1           0.27
  3 4       1           0.27

[angles]
;   i     j     k       funct       angle       force.c.
    1     3     4       2           117         100
    2     3     4       2           175         100
```

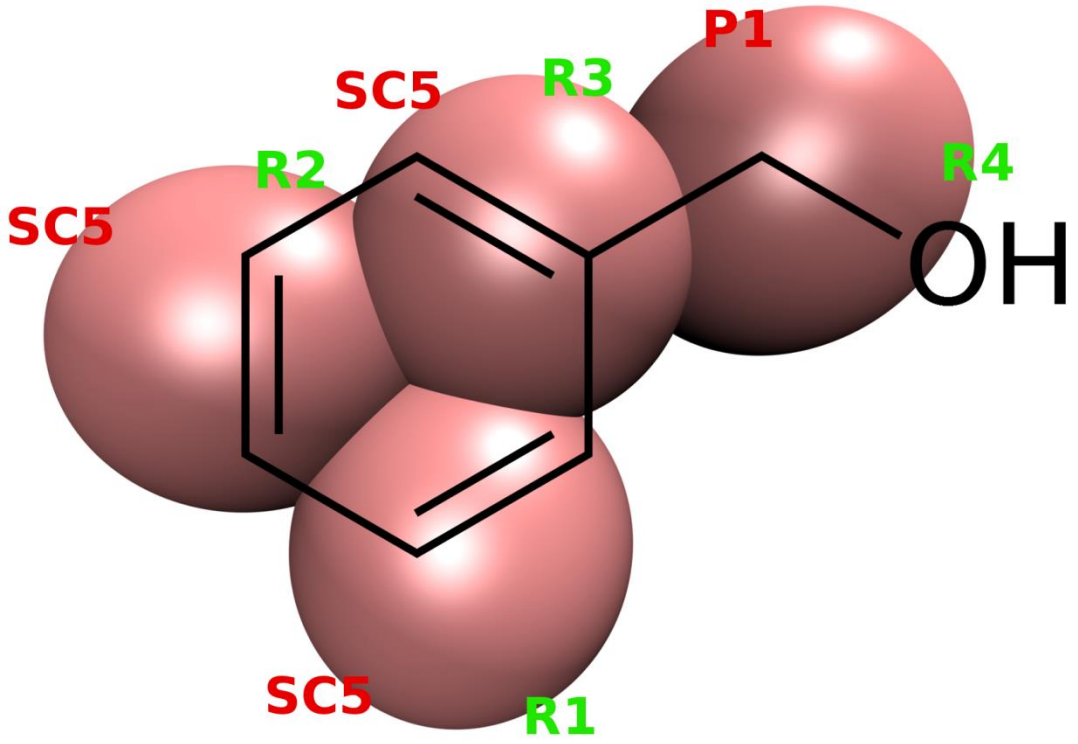


Figure 7 – Schematic representation of the BOH molecule parametrization. Each bead represents two atoms. The green labels indicates the "atom number" according with the itp file, and the red labels the type of bead according the Martini force field developers (see reference 5).