PCCP

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/pccp

Phosphorylation promotes Al(III) binding to proteins: GEGEGSGG as a case study

Rafael Grande-Aztatzi, Elena Formoso, Jon I. Mujika, Jesus M. Ugalde and Xabier Lopez.
Kimika Fakultatea, Euskal Herriko Unibertsitatea (UPV/EHU), and Donostia Internacional Physics Center (DIPC), P.K, 1072, 20080 Donostia, Euskadi, Spain.

Abstract

Aluminum, the third most abundant element in the earth crust and one key industrial components of our everyday life, has been associated with several neurodegenerative diseases due to its ability to promote neurofilament tangle and β-amyloid peptide aggregation. However, the experimental characterization of aluminum speciation in vivo is a difficult task. In the present paper, we develop a theoretical protocol that combines molecular dynamics simulations, clustering of structures, and Density Functional Theory, for the characterization of the binding of Aluminum to the synthetic neurofilament analogue octapeptide GEGEGSGG and its phosphorylated variant. Our protocol is tested with respect to previous NMR experimental data, which allows for a full interpretation of experimental information available and its relation with key thermodynamic quantities. Our results demonstrate the importance of phosphorylation in the ability of a peptide to bind aluminum. Thus, phosphorylation changes: *i)* the binding pattern of aluminum to GEGEGSGG, shifting the preferential binding site from C-terminal to S6(P); *ii)* increases the binding affinity by a factor of around 15 kcal/mol in free energy; and *iii)* may cause significant changes in secondary structure and stiffness of the polypeptide chain, specially in the case of bidentate binding modes. Our results shed light on the possibility of aluminum to induce aggregation of β-amyloid proteins and neurofilament tangles.

1. Introduction

Although aluminum is the third most abundant element in the Earth's crust, it is not an essential element in any living system and it has been excluded from the biotic cycles.[1, 2] However, over the last century and because of its various technological applications, human intervention has increased the bioavailability of Al(III) cation. Unfortunately, this is unlikely to be without consequences. There is increasing evidence of potential toxic effects of aluminum in biological systems,[3-5] linking the presence of aluminum in the human body with several diseases,[6, 7] such as dialysis encephalopathy, osteodystrophy and microcytic anemia derived from chronic renal failure where the aluminum accumulates in kidney, liver, bone and heart.[8, 9] The full molecular bases for aluminum toxicity are still unknown, but some potential mechanisms have been outlined. For instance, it has been well established that aluminum shows a significant pro-oxidant activity in biological systems,[10-12] it also inhibits the normal function of various enzymes involved in the glycolysis pathway[13-16] and in the production of glutamate[17, 18] affecting the TCA cycle.[19-21] Besides, several studies have shown that aluminum competes with magnesium as a metal-ATP cofactor,[22-27] and that can modify the electronic and structural properties of the chelated biological ligands.[28-30] In general, aluminum has a predominance for O donors species,[31] being the strongest aluminum binders biomolecules with negatively charged groups as carboxylates, phenolates, catecholates and phosphates; the last is particularly important because of the diverse range of molecules that contain this group and the cellular processes in which they participate.[32-34]

Aluminum is considered as a neurotoxic element associated to the Alzheimer disease (AD), but is not clear the role that it plays in neurodegeneration.[35-39] In a seminal work, Exley *et al.* demonstrated the ability of aluminum to bind β-amyloids,[40] which led to a switch between α-helix to β-sheet structure. Besides, in vitro experiments determined that Al(III) promotes more efficiently than other metals the aggregation of β-amyloid peptides,[41-43] the main constituent of insoluble amyloid plaques known as senile plaques, and in fact, aluminum has been detected in senile plaques extracted from the brains of patients with AD.[44] In addition, this metal also induces the abnormal neurofilament tangle (NFT) aggregation and promotes the hyperphosphorylation of normal proteins.[45-47] As a matter of fact, phosphorylation of peptides is thought to increase affinity for aluminum and could be a key aspect in the understanding of aluminum promotion of NFT aggregation process. Therefore, there is a high interest to understand the changes that phosphorylation causes in the interaction of polypeptides with aluminum. As a paradigmatic case, Hollender et al.

performed a full study of the structure of the synthetic neurofilament analogue octapeptide GEGEGSGG and its Ser(P) derivative, both in the absence and presence of aluminum.[48] They used potentiometry, CD and multinuclear NMR spectroscopies, and the results were complemented with MD simulations, albeit of limited accuracy, due to the lack at the time of proper aluminum parameters and the inherent limitations of a classical force field. Quite interestingly, they observed that both non-phosphorylated and phosphorylated octapeptide were able to bind Al(III). However, the NMR chemical shifts and the broadening of specific signals suggested a significant change in Al-binding pattern upon phosphorylation. Thus, in the non-phosphorylated polypeptide, coordination of aluminum to the C-terminal carboxylate is preferred, with E4-sidechain as an additional possible chelation site. However, no evidence was found for coordination to the Ser alcoholic–OH group, while in the case of the phosphorylated derivative, the Ser6(P) behaved as the primary binding site, with the C-terminal carboxylate as a potential additional binder. Their results points clearly to an increase in aluminum affinity of a given residue upon phosphorylation, which could provoke significant conformational changes and a more pronounced ability to promote protein aggregation processes.

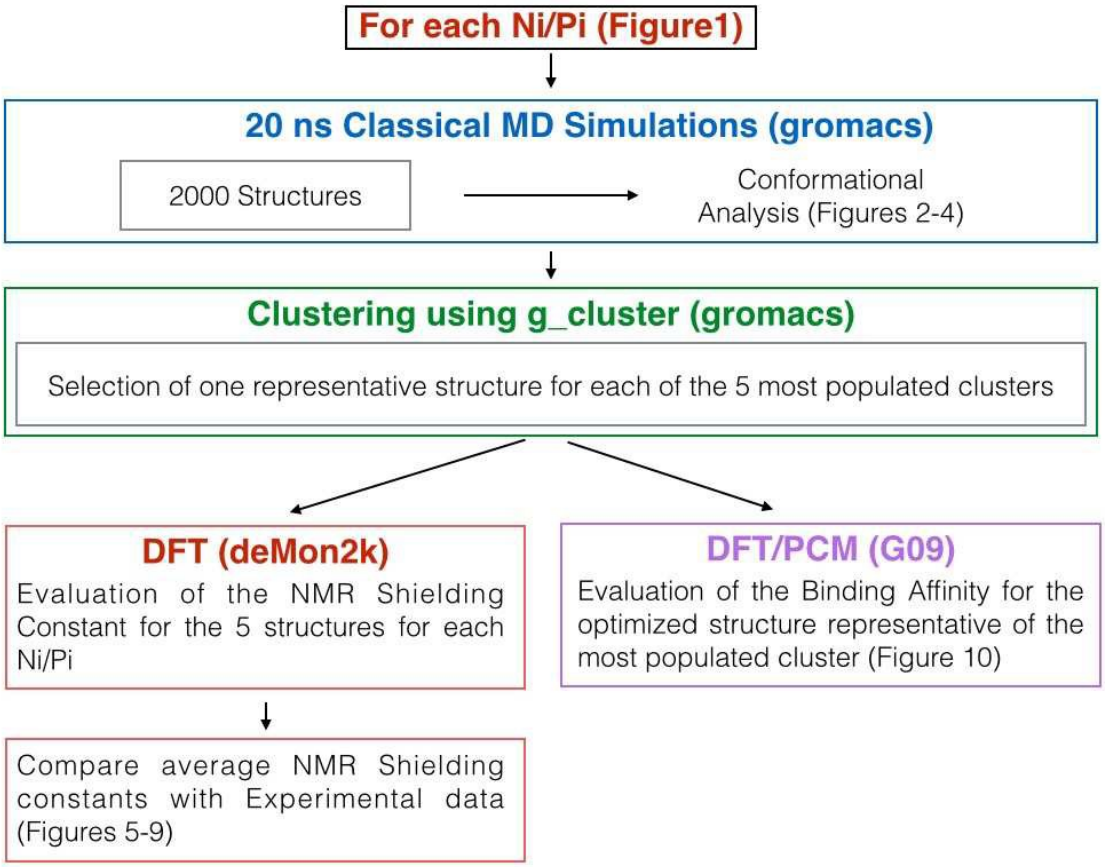
Given the complexity of the coordination chemistry of aluminum in biological systems, theoretical calculations can help to elucidate the structural and electronic aspects of the aluminum complexes in solution.[49-56] In the present paper, we apply a protocol that combines molecular dynamic simulations with properly optimized Al(III) parameters,[57] density functional theory (DFT) calculations of NMR shielding constants, and DFT estimations of binding affinities in the context of polarizable continuum models to consider bulk solvent effects, for both phosphorylated and non-phosphorylated GEGEGSGG octapeptide and in the presence/absence of Al(III). Different coordination modes were tested (see Fig. 1). The aim of this work is to understand the changes in polypeptide structure and electronic properties caused by the presence of Al(III) in combination with peptide phosphorylation. A careful comparison of our results with the experimental NMR chemical shifts showed a good agreement, and allowed to rationalize the interaction mode of aluminum towards phosphorylated amino acids. Finally, we report the calculation of key thermodynamic data for the complex formation, which allow us to estimate the increase in aluminum binding affinity expected by phosphorylation of Ser in the octapeptide, and we compare our estimates with previous studies using similar methodologies of well-known aluminum binders in biological media, such as citrate, 2,3-DPG or ATP. Our results points to a clear increase in binding affinity upon peptide phosphorylation, although its degree is not high enough as to be able to compete with strong aluminum chelators.

2. Methodology

The theoretical protocol used throughout this work is summarized in Scheme I. Each of the structures of Figure 1 were built from scratch with Molden[58] in a β-conformation with the N-terminal protonated and the C-terminal deprotonated as well as the glutamic acids. The MD trajectories were carried out using Gromacs package,[59, 60] version 4.5.7 and the CHARMM36 force field[61] including the nonbonded parameters of Al(III) from Faro *et al.*, used previously.[57, 62, 63] The peptides were solvated in a 3.75 nm^3 dodecahedral water box with the TIP3P model,[64] depending on the case, sodium and chloride ions were added to neutralize the charge. An initial energy minimization of the systems was performed using the steepest descent method for 10000 steps. The Particle Mesh Ewald (PME) method[65, 66] was used to calculate the electrostatic interactions with a cut-off of 0.8 nm for the short range Van der Waals interactions. The integration time step was of 2 fs and after 200 ps equilibration, trajectories of 20 ns were performed in the canonical ensemble at 300 K with the Bussi thermostat.[67] All MD trajectories were sampled every 10 ps and the structures were clustered by the Gromacs utility that use the algorithm described in Daura et al.[68] In our case, this clustering brings together more than 46% of the sampled structures into the first five groups (Table S1).

The NMR calculations were done using one representative structure of each group including only explicit water molecules located at ~3.0 Å from the peptide; the values were averaged to keep the dynamic effects and compared with the experimental data. The NMR chemical shifts respect sodium 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) were computed with DFT as implemented in the deMon2k code,[69-72] using the exchange-correlation functional PBE,[73] the orbital basis set aug-cc-pVTZ [74, 75] and an automatically generated GEN-A2 auxiliary basis set, used to avoid the evaluation of four center integrals by invoking the variational fitting of the Coulomb energy.[76, 77]

The binding affinities were computed with Gaussian09 program;[78] one structure from the most populated cluster of each coordination mode was optimized and characterized as local minima with a harmonic frequency analysis using the B3LYP[79, 80] functional and the 6-31+G(d) basis set, the zero-point vibrational energy (ZPVE) and the thermal vibrational corrections at 298 K to evaluate the enthalpies and Gibbs free energies were extracted from the frequency analysis; the entropy was computed using the standard statistical mechanics expressions for the partition function in the canonical ensemble. Solvent effects were included using the implicit solvation model IEFPCM.[81] Electronic energies were refined by single-point energy calculation at the B3LYP/6-311++G(d,p) level.



Scheme I. Theoretical Protocol used throughout this work for each of the structures of Figure 1, based on: i) MD simulations, ii) clustering of structures iii) DFT evaluation of NMR shielding constants for selected structures iv) DFT/PCM evaluation of binding free energy/enthalpy affinities.

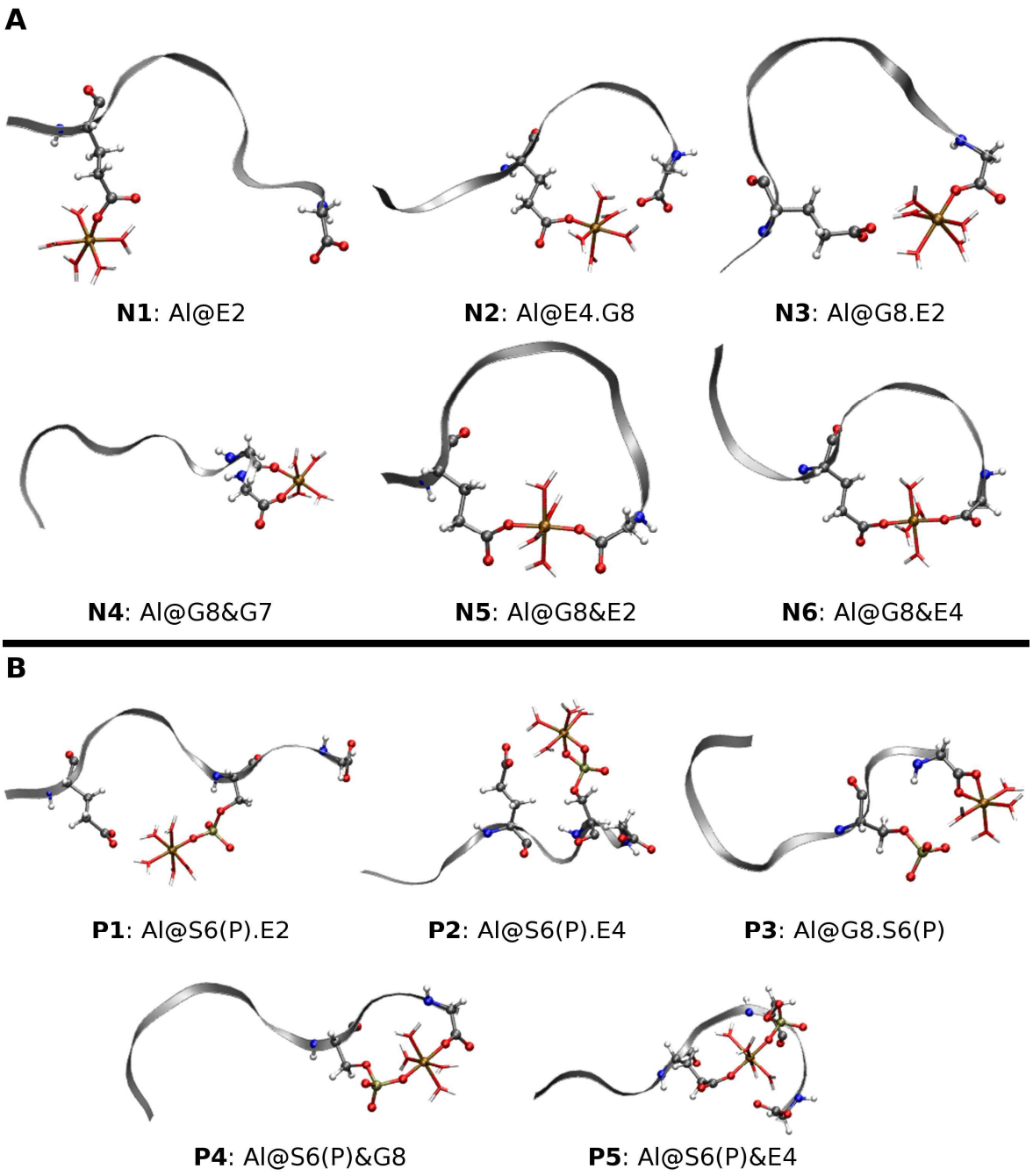


Figure 1. Structures considered for non-phosphorylated (**Ni**, i=1-6) and phosphorylated (**Pi**, i=1-5) peptides. In the non-phosphorylated peptide (**A**) the ion is bound to the glutamic acids E2 or E4 (**N1** and **N2**), to the C-terminal in mono and bidentate fashion (**N3** and **N4**) and to the C-terminal in combination with E2 or E4 (**N5** and **N6**). In the phosphorylated derivative (**B**) the ion is coordinated to the phosphate group mono and bidentately (**P1** and **P2**), to the C-terminal (**P3**) and to the phosphate group combined either with the C-terminal (**P4**) or with E4 (**P5**). The labels indicate the residue or residues at which the Al(III) is coordinated and the "dot" stands for the group placed in the second coordination sphere.

3. Results

3.1. Conformational Analysis from MD Simulations

The Ramachadran plots obtained from the MD simulations for all the structures can be found in Figure 2. In the absence of aluminum, there are well defined regions that belongs to β-sheet conformation (-135º, 135º), to the PPII (-75º, 150º) and α-helix regions both left- (70º, 25º) and right-handed (-70º, -30º), although the starting structures were created using a β-sheet conformation. This suggests that our 20 ns MD simulation time allows for a proper sampling of the conformational space. The phosphorylation of S6 has a sizeable effect in the α-helix regions, being less populated upon phosphorylation.

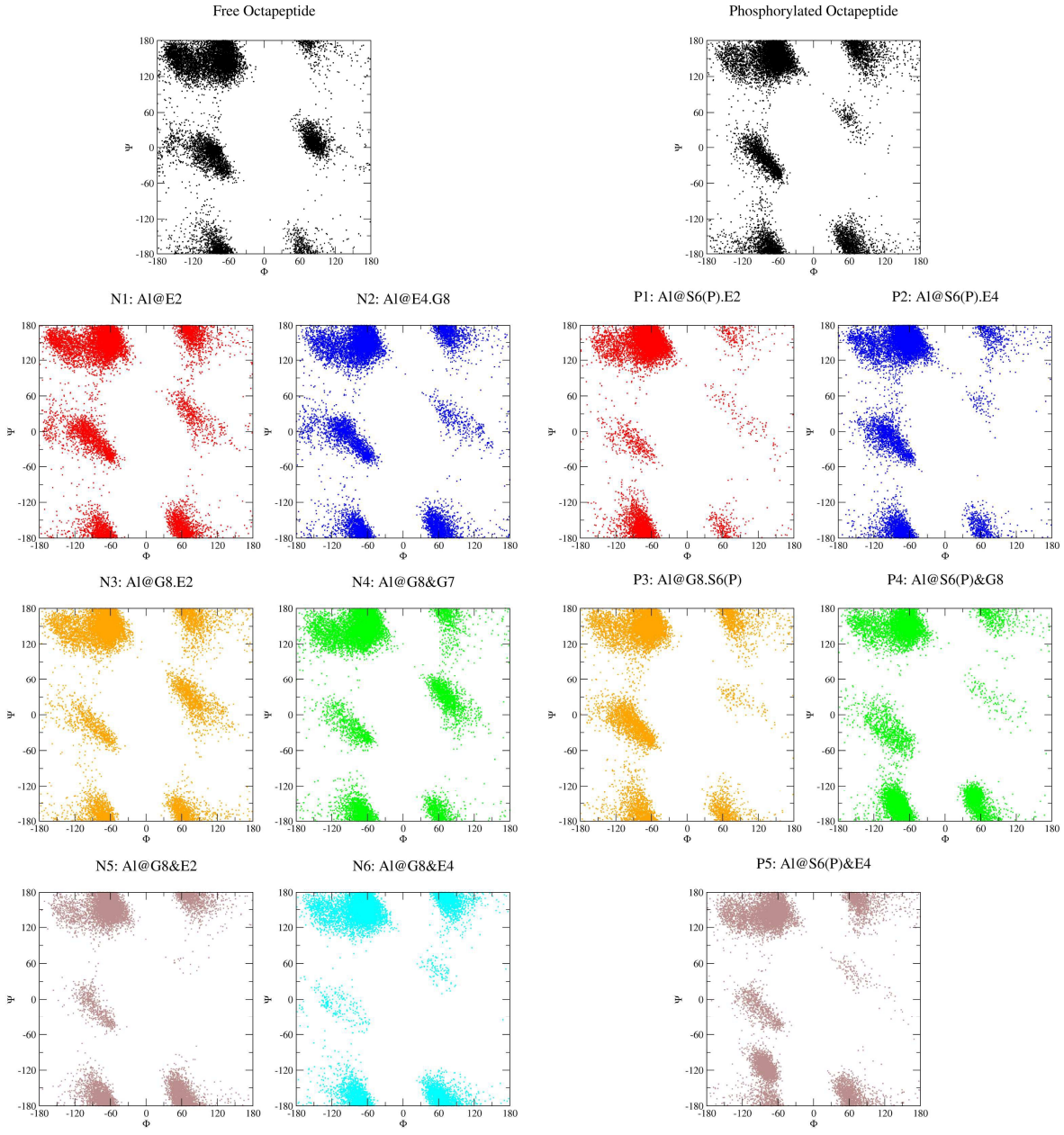


Figure 2. Ramachadran Plots from MD simulations for the non-phosphorylated and phosphorylated structures.

The presence of aluminum can also have a significant effect on the conformational preference depending on the coordination type. In the case of the non-phosphorylated peptides the largest effects are observed for the two bidentate structures (**N5** and **N6**), where the Lα region is very poorly populated, and correspondingly, there is a stabilization of the β-sheet conformation. The rest of possible aluminum coordination modes seem to have a lower effect in the β/α conformational preference. In the case of the phosphorylated structures, we observe a lower tendency to occupy the Lα region as well, although in this case the free phosphorylated octapeptide also shows a depletion of this zone.

The presence of aluminum can also lead to changes in the rigidity of the polypeptide. In Figure 3, the root mean square fluctuations (RMSF) of Cα's as a function of residue are depicted. In most of them we observe a lowering of RMSF's respect to the free peptide, specially for **N5**, **N6**, **P4** and **P5**, all of them, bidentate structures. This implies that the aluminum has the ability to increase the rigidity of the octapeptide by decreasing the mobility of the coordination site.

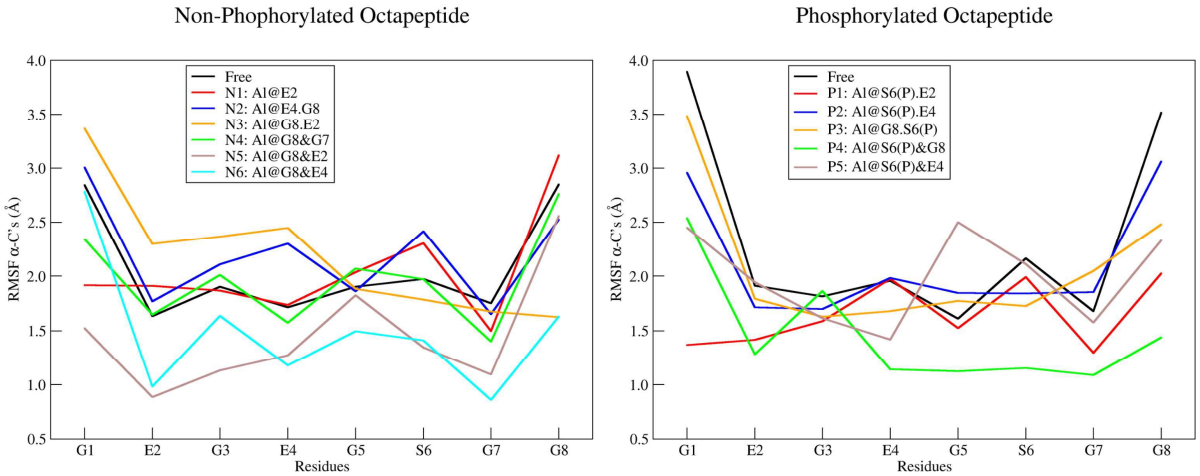


Figure 3. Root mean square fluctuations of Cα's in each of the non-phosphorylated and phosphorylated structures, calculated from MD simulations.

In order to further asses the change in structure upon aluminum binding, in Figure 4 we depict the average distance between pair of residues, comparing the results for aluminum complexes (y axis) versus the free octapeptides (x axis). Data above the diagonal line implies an elongation with respect to the distance in the free octapeptide. Conversely, the

points below the diagonal line imply a shortening of the distance between two residues upon aluminum binding. A variety of situations are found depending on the type of aluminum coordination, demonstrating that the secondary structure of the peptide is sensible to aluminum binding. For instance, in the case of non-phosphorylated octapeptide the bidentate binding to G8 and E2 (**N5**) leads to a clear shortening of the distances between residues. The same is true for the phosphorylated **P2** structure. However, the binding to S6(P) in **P1** leads to elongation of the structure of the polypeptide.

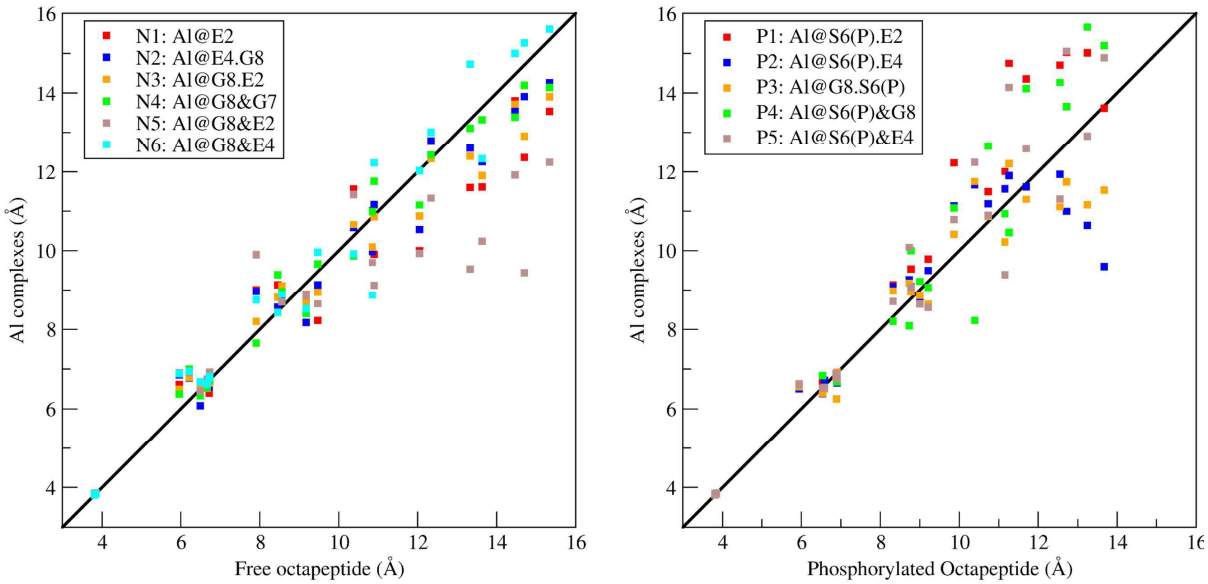


Figure 4. Coordinating effect of aluminum on the average distances between pairs of Cα's in each of the non-phosphorylated and phosphorylated structures, calculated from MD simulations.

3.2. NMR Chemical Shifts

In this part the NMR chemical shifts are discussed, it is worth to note that the computed data are averages of five structures from MD simulations of each coordination mode and of the free octapeptides both non-phosphorylated and phosphorylated. First we compare the data for the free octapeptides with the experimental values to have an idea of the differences with the methodology used (section 3.2.1); after that, the corresponding results for the non-phosphorylated complexes are discusses (section 3.2.2), to evaluate the effect of the aluminum coordination in the peptide, the signals of the complexes are compared with the free octapeptide to visualize the chemical shifting looked experimentally and finally the same is done for the phosphorylated complexes, besides the ^{31}P signals are also compared (section 3.2.3).

3.2.1. NMR Chemical shifts of the GEGEGSGG peptide and its Ser(P) derivative.

The computed chemical shifts of the free octapeptide have a good agreement with the experimental data. Respect to the ^{1}H experimental values, the largest deviations in the CH and CH_2 groups are 0.6 ppm, while the amidic protons have differences less than 3.0 ppm (Fig. 5 **A** and **B**). The ^{13}C values of the CH and CH_2 groups have maximum differences of 5.0 and 12.1 ppm, respectively and the CO and COO groups deviate from the experimental values up to 8.6 and 6.5 ppm (Fig. 5 **C** and **D**).

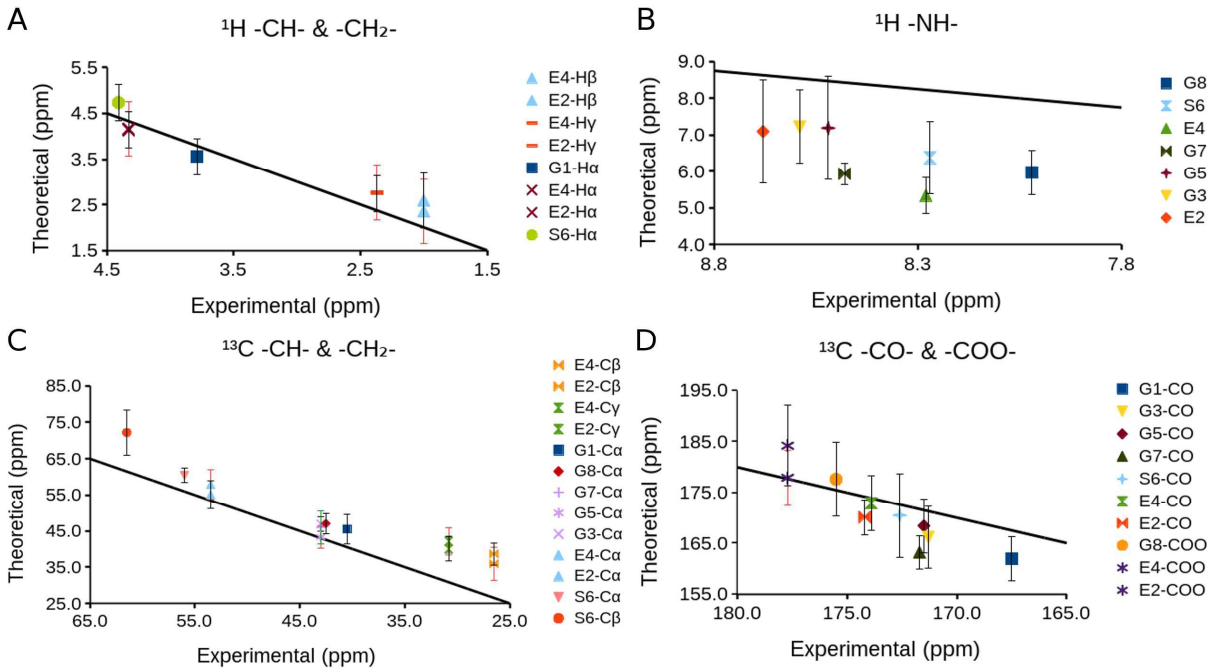


Figure 5. Average values of the NMR chemical shifts computed with the free octapeptide. Black line is the perfect agreement between theoretical and experimental values. Bars are the standard deviations.

The ^{1}H chemical shifts of the phosphorylated peptide are quite similar to the non-phosphorylated, as was also observed experimentally;[48] the variations between computed values are less than 1.0 ppm, while in the ^{13}C values, the general trend marks a good agreement with the experiment being the larger differences of 7.6 ppm in the CH, 15.0 ppm in the CH_2, 8.3 ppm in the CO and 7.1 ppm in the COO (Fig. 6). It is worth noting that the ^{13}C differences obtained with the selected methodology in most cases are twice of those expected when optimized structures are used.[72, 82-84]

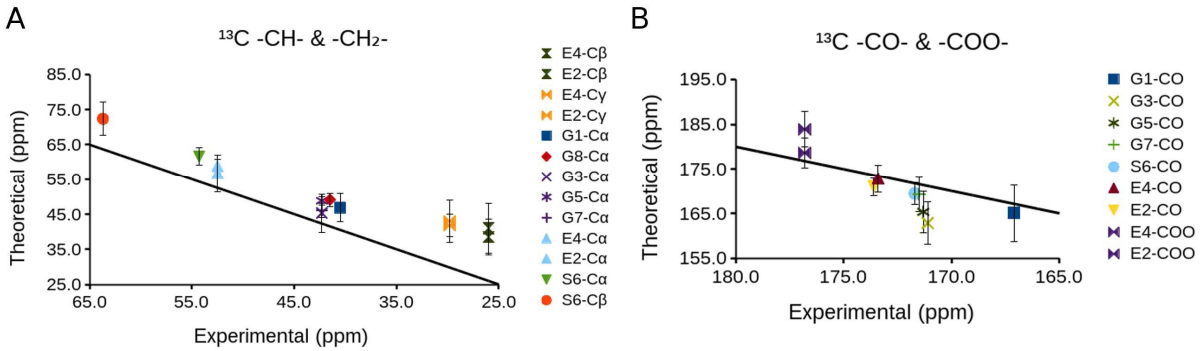


Figure 6. Average values of the ^{13}C NMR chemical shifts computed with the phosphorylated octapeptide. Black line is the perfect agreement between theoretical and experimental values. Bars are the standard deviations.

3.2.2. Al(III) binding to the GEGEGSGG peptide.

As was mentioned above, experimentally it is observed a line-broadening and loss of some ^{13}C signals in the NMR spectra upon complexation;[48] the former can not be reproduced, but the latter can be simulated to compare the computed chemical shifts of the aluminum complexes with the free peptide, specially those of the groups involved in the binding. The **N1**, **N2** and **N3** models have monodentate coordination to the carboxyl group and in **N4**, **N5** and **N6** are bidentate modes (Fig. 1).

The results show that the aluminum coordinated to E2 change their CO and COO values by 3.0 ppm (Fig. 7 **N1**), the other groups undergo changes the same extent. The coordination to E4 displays a similar trend with variations around 4.0 ppm for the CO and COO groups, the CO value of G7 deviates 7.3 ppm, but it is within the range of differences obtained for the free peptide (Fig 7 **N2**). The coordination to the C-terminal in a monodentate fashion displays similar variations as in the previous cases (Fig 7 **N3**), but in combination with the carbonyl the larger deviations are found respect to the free octapeptide values, 12.1 ppm in CO-G7 and 9.1 ppm in COO-terminal, the other groups have similar variation as in the other models (Fig. 7 **N4**). The coordination of the C-terminal combined with any glutamic acids (Fig. 7 **N5** and **N6**) show that the carbons involved in the coordination remains intact and do not reflected the experimental changes and therefore their existence is hardly conceivable, as was stated.[48]

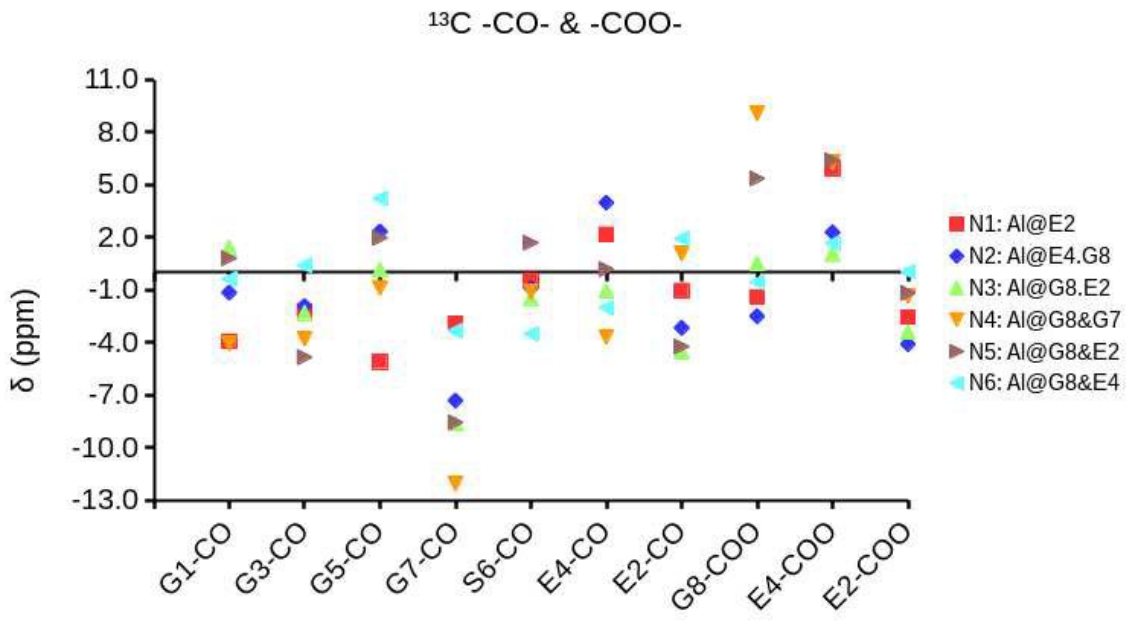


Figure 7. Differences of the computed ^{13}C NMR chemical shifts between the free octapeptide and its Al(III) complexes.

Taking into account the agreement with the experimental values in the free octapeptide and the variations in the aluminum complexes, the coordination with the glutamic acids does not reflect the changes in the experimental NMR spectrum and neither does the monodentate coordination with the C-terminal; only when the carbonyl of G7 is involved in the coordination the differences in the computed values of carboxyl groups on G8 and E4 are greater than 6.0 ppm and twice in the carbonyl on G7, such shifting on G7 could be hidden by the signal broadening.

3.2.3. Al(III) binding to the phosphorylated peptide.

In the phosphorylated spectra, the line-broadening of the ^{13}C signals is the substantial change, as well as the shifting of the CH and CH_2 signals from S6, which help us to distinguish between coordination modes. The models have the aluminum bound to the phosphate group mono- and bidentate (**P1** and **P2**), to the C-terminal as bidentate ligand (**P3**), and to the phosphate combined with the C-terminal or with E4 (**P4** and **P5**), the monodentate coordination was not tested since similar results are expected to those obtained with the non-phosphorylated octapeptide.

The results of the differences between the aluminum complexes and the phosphorylated peptide, show that the values of S6 have differences of about 3.0 ppm (Fig. 8 **P1**), indicating small structural changes and therefore small signal shifting, the bidentate

phosphate shows similar variations to the monodentate complex and is not possible discard this coordination mode (Fig 8 **P2**), but for the bidentate C-terminal it is, since modifies considerably the CO region with differences greater than 7.0 ppm and the CH and CH_2 signals of S6 are slightly affected (Fig 8 **P3**). The combined phosphate with the C-terminal change its carboxyl signal about 3.0 ppm, but the most interesting results are the changes by 8.8 and 1.2 ppm in the CH and CH_2 groups of S6, respectively (Fig. 8 **P4**), although the variation on CH_2 is small, this model displays the biggest change in the aliphatic region; while the combination with the carboxyl of E4 does not reflect significant changes in the aliphatic region nor in the signals of the carbons involved in the coordination (Fig 8. **P5**), the greatest variations correspond to G3 and G5 carbonyls both with 8.0 ppm and experimentally it is only observed line-broadening of that region. Therefore according to the ^{13}C chemical shifts, the best candidates to reproduce the experimental situation are **P1** and **P4**, where the aluminum coordinated to the phosphate group in a monodentate fashion does not alter the resonance pattern and the binding of C-terminal produce the observed signal shifting on the aliphatic carbons of serine.

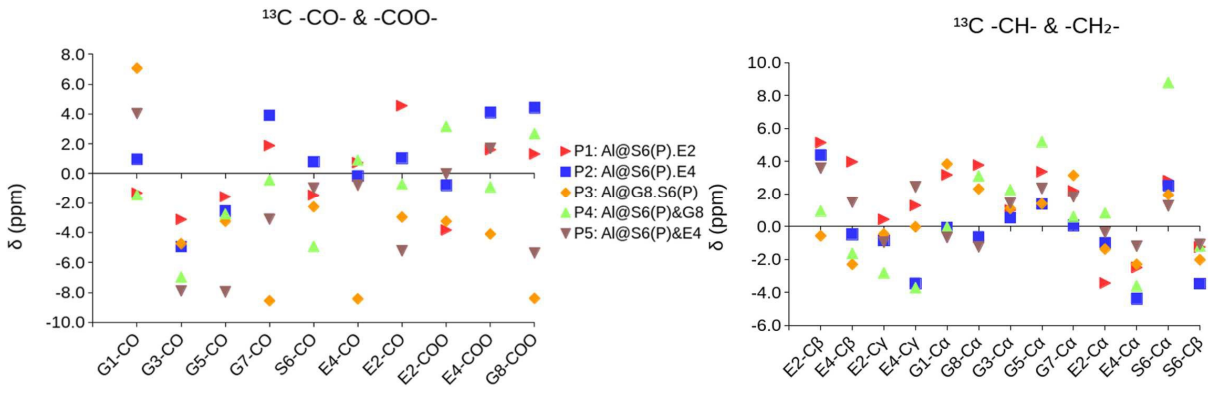


Figure 8. Differences of the computed ^{13}C NMR chemical shifts between the phosphorylated octapeptide and its Al(III) complexes.

Finally, the aluminum coordination to the phosphate group was proved experimentally comparing the ^{31}P chemical shifts of the phosphorylated peptide and its complexes; it was assigned the signal at 1.36 ppm at the free octapeptide and the range of 0.8 to -11 ppm to the aluminum complexes.[48]

The computed values of the phosphorylated peptide show positive and negative values (Fig. 9 Free), averaging them we have values of 2.6 and -7.0 ppm and a variation with the most intense experimental signal less than 2 ppm. For aluminum complexes, the computed values of the five models show that if the phosphate group does not participate

in the aluminum coordination the ^{31}P chemical shifts have the largest deviation form the experimental region (Fig. 9 **P3**), while the other cases are closer to the experimental trend and discard one of them straightforward is not possible; however the model **P1** would be the best candidate to display the experimental resonance pattern.

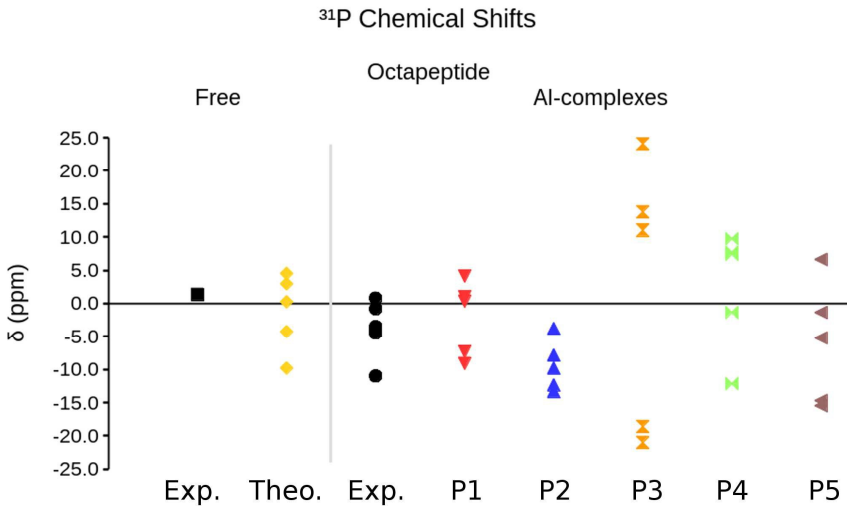


Figure 9. Experimental and theoretical ^{31}P chemical shifts for the phosphorylated peptide and its aluminum complexes. Experimental values taken from reference [48].

3.3 Binding free energies in solution.

The complex formation stability was studied following the next ligand substitution reaction:

$$[Al(H_2O)_6]^{3+}_{(aq,\,1M)} + L^{k-}_{(aq,\,1M)} \rightarrow [Al.L.(H_2O)_{6-m}]^{3-k}_{(aq,\,1M)} + mH_2O_{(aq,\,1M)}$$

where *L* is the octapeptide or the phosphorylated derivative with *k* = 2 and 4, respectively and *m* depends on the coordination mode, 1 for monodentate and 2 for bidentate. The free energy in solution corresponding to the aluminum binding to the peptides is calculated as follow:

$$\Delta G_{aq} = G_{aq}([Al.L.(H_2O)_{6-m}]^{3-k}) + m.G_{aq}(H_2O) - G_{aq}([Al(H_2O)_6]^{3+}) - G_{aq}(L^{k-}) + \Delta nRT\ln(24.26) + mRT\ln(55.34)$$

where the last two terms account for the volume change due to the transformation from 1 atm to 1M in solution, Δn indicates the change in the number of species in the reaction, and the concentration of 55.34 M of water in liquid water, respectively.[85] The enthalpy values were also computed excluding the last term for it.

The complexation energies of both non-phosphorylated and phosphorylated aluminum complexes are displayed in Figure 10. The aluminum binding energies results for the non-phosphorylated complexes (**N1** to **N6** in Fig. 1) show that the monodentate coordination with E2 (**N1**) would be unlikely (-33.70/-36.48 kcal/mol for $\Delta G_{aq}/\Delta H_{aq}$) and that there is a clear preference for coordination to E4 (**N2**, -60.98/-71.81 kcal/mol) and the C-terminal carboxylate (**N3**, -56.47/-67.19 kcal/mol), with both sites being competitive for Al binding.

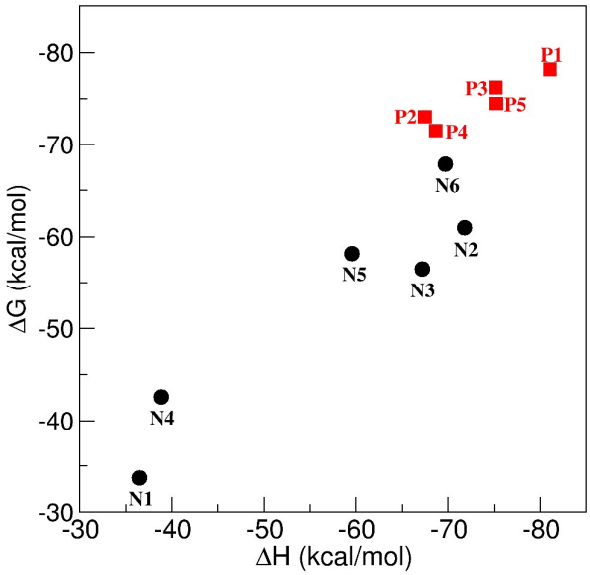


Figure 10. Complexation energies of the non-phosphorylated (black) and phosphorylated (red) aluminum complexes.

Next, we also estimate the binding affinities in bidentate binding modes with two residues (**N5** and **N6**). In both structures, one of the ligands is through the C-terminal of G8, whereas the other coordination is through the acetate group of E2 (**N5**) or E4 (**N6**). Again, we detect a preference for E4 coordination with a binding energy of -67.86/-69.71 kcal/mol for **N6** versus -58.17/-59.57 kcal/mol for **N5**.

In summary, our results for the non-phosphorylated octapeptide points out to a competition between E4 and the C-terminal for the binding to aluminum, with a variety of situations of similar energy: monodentate to one of the groups or bidentate to both of them. This also suggests that it is highly likely that the octapeptide coordinate two aluminum atoms, one in the C-terminal and other in E4 as was suggested by Hollender *et al*,[48] but the dinuclear species are beyond the scope of this work.

On the other hand, the binding energies for the phosphorylated octapeptide were done with five models (**P1** to **P5** in Fig. 1). In this case, it is observed that the preferred coordination mode is through the phosphate at the S6 (**P1**, -78.18/-81.04 kcal/mol), with

other groups binding to the waters of the first coordination shell of aluminum. There are also other competitive structures close in energy, like **P3** (-76.21/-75.16 kcal/mol) in which aluminum is interacting directly with C-terminal and indirectly with the phosphate of S6 through one of the coordinating waters, and **P5** (-74.47/-75.19 kcal/mol) in which aluminum is interacting bidentately with S6 and E4. It is worth to note that during the optimization some proton transfer were observed from the water molecules to the closer carboxylic group (Note S1), which are known to be due to limitations in the solvation models.[86]

Our calculations demonstrate a clear increase in aluminum binding affinity upon phosphorylation. Averaging over the three most favorable structures in each case (non-phosphorylated and phosphorylated octapeptide); we obtain an average binding free energy/enthalpy of -61.8/-69.6 kcal/mol for the non-phosphorylated octapeptide versus -76.3/-75.2 kcal/mol for the phosphorylated one. Thus, we can conclude that phosphorylation leads to a clear higher affinity for aluminum, and therefore, the presence of aluminum can favor thermodynamically the formation of phosphorylated peptides. Notice, however, that our calculations correspond to an acidic pH (as were the NMR experiments of Hollender et al.). Extension of the present results to more physiological pH values should be done with caution, since the affinity of the octapeptide towards aluminum could be altered by the presence of hydroxides in the first solvation layer around aluminum as we increase the pH.

4. Biological Implications

The interest of studying the structure and thermodynamic properties of formation of aluminum complexes with neurofilament analogue octapeptide GEGEGSGG and its Ser(P) derivative is of high interest by two main reasons. On one hand, the analysis of the structural effects that aluminum provokes upon binding to this polypeptide could shed light on the mechanism by which aluminum enhances aggregation of β-amyloids and NFTs. On the other hand, phosphorylation could be key to find a lead compound for the design of a selective peptide-based Al(III) chelators and our data could help in the rationalization of the effect expected in binding affinity by phosphorylation.

Based on the theoretical results of the present paper, we can conclude that phosphorylation has a profound effect in the binding pattern of a polypeptide to Al(III). Phosphorylation changes the structure of the binding pattern of Al(III) towards the octapeptide, shifting the preferable binding site from E4 to S6(P), with the C-terminal carboxylate being a favorable binding site in both non-phosphorylated and phosphorylated

peptides. In addition, phosphorylation increases significantly the affinity of the GEGEGSGG octapeptide towards Al(III) binding. An increase around 15 kcal/mol in binding free energy was estimated by our calculations upon phosphorylation. Our results are coherent with the experimental NMR information and points out to the preference of coordination of Al(III) to negatively charged groups, specially phosphate monoester functional groups bearing two negative charges. Our previous estimates for aluminum binding affinities of phosphate monoesters gave an estimate of -65.9/-67.3 kcal/mol.[31] Our results for the phosphorylated octapeptide **P1** is significantly higher -78.2/-81.0 kcal/mol, an indicative of the collaborative effects between the phosphate and carboxylate groups in binding to aluminum.

Is the increase in binding affinity upon phosphorylation enough to transform the present octapeptide to a strong Al(III) chelator in biological systems? To answer this question, we compare our binding affinities to the ones calculated in previous works for known strong biological chelators: citrate, the main low molecular mass chelator of aluminum in blood serum, 2,3-DPG, chelator of aluminum in red cells, and ATP-like triphosphates.[31, 86] The binding free energies/enthalpies for these chelators are: citrate (-133.0/-124.9 kcal/mol) > 2,3-DPG (-123.5/-118.9 kcal/mol) > ATP-like triphosphates (-109.2/-108.7 kcal/mol). Thus, these chelators show much larger binding affinities than the phosphorylated octapeptide. In this sense, it is clear that the phosphorylation of one residue is not enough to transform the octapeptide into a strong Al(III) chelator. In fact, Hollender *et al*. established that neither the non-phosphorylated nor the phosphorylated peptide prevented the precipitation of Al(III) complexes at around the physiological pH, keeping Al(III) in solution only in the weakly acidic pH range.[48]

On the other hand, based on the favorable formation of intramolecular bridges between functional groups through bidentate structures, we can expect that aluminum could induce aggregation of polypeptides through different type of intermolecular bridges. The most likely possibility at low pH would be through phosphate-Al-phosphate intermolecular bridges, with phosphates located in different polypeptides. However, other competitive structures would be carboxylic-Al-phosphate or carboxylic-Al-carboxylic bridges, which could be more likely at higher pH's. In fact, in our previous work with 2,3-DPG phosphate chelator,[86] we have demonstrated the propensity of aluminum to form 1:2 complexes, and therefore, the possibility to act as an aggregation agent for phosphates. Our results in the present paper point in the same direction. Moreover, we have observed that the formation of such bidendetate structures provokes changes in the secondary structure of the polypeptide, increasing its stiffness and favoring the conformational space region

associated to β-sheet structures.

5. Conclusions

In this work we have applied a theoretical protocol (*vide supra*) to study the coordination of aluminum to a synthetic neurofilament analogue GEGEGSGG octapeptide, both phosphorylated and non-phosphorylated variant, combining classical mechanics simulations, selection of structures based on clustering methods, DFT evaluation of NMR shielding constants and DFT/PCM calculation of binding affinities. Our results shed light on the important subject of interaction of aluminum to polypeptides, a key interaction at the center of the potential toxic effects of aluminum.

Our results points to a clear enhancement of the affinity of aluminum by phosphorylation of peptides in this model system, with changes that covers both structural and thermodynamic aspects. On one hand, phosphorylation changes the preferential binding site for aluminum, shifting the preference from C-terminal and/or G7/E4 in the non-phosphorylated octapeptide to S6(P) upon phosphorylation, either monodentately or in combination with the C-terminal or E4. Our calculations also show a variety of possible structures for aluminum binding, suggesting that there is a variety of competitive structures for aluminum binding which leads to a rich diversity of potential structures to be formed. Our data is in agreement with previous NMR estimates and allows therefore for a clarification of previous experiments. Whether this behavior can be extended to peptides of larger size and to proteins in general, will be the subject of future work.

On the other hand, the bind of the polypeptide to aluminum could have a sizable effect on its secondary structure, being sensible to the type of coordination to the metal. Thus, specially in the case of bidentate structures, in which aluminum is bound to two different groups of the octapeptide, we have observed an enhancement of the population of the regions associated with β-sheet structures in the Ramachadran plots, with an increase of the stiffness of the backbone.

Finally, the computation of the affinity binding energies shows clearly a promotion of aluminum binding upon peptide phosphorylation. This enhancement is not enough to transform GEGEGSGG peptide into a strong aluminum chelator. Our results clearly points to the possibility of promotion of protein aggregation induced by aluminum by the formation of intermolecular phosphate monoester bridges.

- **Acknowledgements**

The authors thanks for technical and human support provided by SGI/IZO (SGIker) of UPV/EHU. RG-A thanks to Consejo Nacional de Ciencia y Tecnología (CONACYT) for the postdoctoral fellowship. Financial support comes from Eusko Jaurlaritza and the Spanish Office for Scientific Research. XL would like to thank the Spanish Ministerio de Ciencia e Innovación (CTQ2012-38496-C05-04) and UPV/EHU (PES 14135) for funding.

- **References**

1. C. Exley, *J. Inorg. Biochem.*, 2003, **97**, 1-7.
2. C. Exley, *Trends Biochem. Sci*, 2009, **34**, 589-593.
3. C. Exley, *Coord. Chem. Rev.*, 2012, **256**, 2142-2146.
4. R. B. Martin, *Acc. Chem. Res.*, 1994, **27**, 204-210.
5. R. A. Yokel, *Current Inorganic Chemistry*, 2012, **2**, 54-63.
6. T. L. Macdonald and R. B. Martin, *Trends Biochem. Sci*, 1988, **13**, 15-19.
7. B. P. Vasudevaraju, M. Govindaraju, A. P. Palanisamy, K. Sambamurti and K. S. J. Rao, *Indian Journal of Medical Research*, 2008, **128**, 545-556.
8. F. B. Seibert and H. G. Wells, *Archives of Pathology*, 1929, **8**, 230-262.
9. C. A. Shaw and L. Tomljenovic, *Immunologic Research*, 2013, **56**, 304-316.
10. C. Exley, *Free Radical Biol. Med.*, 2004, **36**, 380-387.
11. J. I. Mujika, F. Ruiperez, I. Infante, J. M. Ugalde, C. Exley and X. Lopez, *J. Phys. Chem. A*, 2011, **115**, 6717-6723.
12. F. Ruiperez, J. J. Mujika, J. M. Ugalde, C. Exley and X. Lopez, *J. Inorg. Biochem.*, 2012, **117**, 118-123.
13. S. W. Cho and J. G. Joshi, *Toxicology*, 1988, **48**, 61-69.
14. C. Exley, N. C. Price and J. D. Birchall, *J. Inorg. Biochem.*, 1994, **54**, 297-304.
15. G. A. Trapp, *NeuroToxicology*, 1980, **1**, 89-100.
16. F. S. Xu, L. Melethil, L. Winberg and M. Badr, *J. Pharmacol. Exp. Ther.*, 1990, **254**, 301-305.
17. S. J. Yang, J. W. Huh, J. E. Lee, S. Y. Choi, T. U. Kim and S. W. Cho, *Cellular and Molecular Life Sciences*, 2003, **60**, 2538-2546.
18. P. Zatta, E. Lain and C. Cagnolini, *Eur. J. Biochem.*, 2000, **267**, 3049-3055.
19. J. Lemire, R. Mailloux, S. Puiseux-Dao and V. D. Appanna, *J. Neurosci. Res.*, 2009, **87**, 1474-1483.
20. R. J. Mailloux, R. Hamel and V. D. Appanna, *Journal of Biochemical and Molecular Toxicology*, 2006, **20**, 198-208.
21. J. I. Mujika, J. M. Ugalde and X. Lopez, *J. Phys. Chem. B*, 2014, **118**, 6680-6686.
22. T. Dudev and C. Lim, in *Annual Review of Biophysics*, 2008, vol. 37, pp. 97-116.
23. T. Kiss, *J. Inorg. Biochem.*, 2013, **128**, 156-163.
24. E. Rezabal, J. M. Mercero, X. Lopez and J. M. Ugalde, *J. Inorg. Biochem.*, 2006, **100**, 374-384.
25. E. Rezabal, J. M. Mercero, X. Lopez and J. M. Ugalde, *J. Inorg. Biochem.*, 2007, **101**, 1192-1200.
26. E. Rezabal, J. M. Mercero, X. Lopez and J. M. Ugalde, *Chemphyschem*, 2007, **8**, 2119-2124.
27. X. Yang, Q. Zhang, L. Li and R. Shen, *J. Inorg. Biochem.*, 2007, **101**, 1242-1250.
28. J. F. Fan, L. J. He, J. Liu and M. Tang, *J. Mol. Model.*, 2010, **16**, 1639-1650.

29. J. I. Mujika, J. M. Ugalde and X. Lopez, *Theor. Chem. Acc.*, 2011, **128**, 477-484.
30. A. L. Oliveira de Noronha, L. Guimaraes and H. A. Duarte, *Journal of Chemical Theory and Computation*, 2007, **3**, 930-937.
31. N. B. Luque, J. I. Mujika, E. Rezabal, J. M. Ugalde and X. Lopez, *PCCP*, 2014, **16**, 20107-20119.
32. K. Atkari, T. Kiss, R. Bertani and R. B. Martin, *Inorg. Chem.*, 1996, **35**, 7089-7094.
33. T. Kiss, P. Zatta and B. Corain, *Coord. Chem. Rev.*, 1996, **149**, 329-346.
34. D. Mazzuca, N. Russo, M. Toscano and A. Grand, *J. Phys. Chem. B*, 2006, **110**, 8815-8824.
35. C. Exley and O. V. Korchazhkina, *J. Inorg. Biochem.*, 2001, **84**, 215-224.
36. V. B. Gupta, S. Anitha, M. L. Hegde, L. Zecca, R. M. Garruto, R. Ravid, S. K. Shankar, R. Stein, P. Shanmugavelu and K. S. J. Rao, *Cellular and Molecular Life Sciences*, 2005, **62**, 143-158.
37. V. Kumar and K. D. Gill, *Arch. Toxicol.*, 2009, **83**, 965-978.
38. L. Tomljenovic, *Journal of Alzheimers Disease*, 2011, **23**, 567-598.
39. J. R. Walton, *NeuroToxicology*, 2006, **27**, 385-394.
40. C. Exley, N. C. Price, S. M. Kelly and J. D. Birchall, *FEBS Lett.*, 1993, **324**, 293-295.
41. M. Kawahara, M. Kato and Y. Kuroda, *Brain Research Bulletin*, 2001, **55**, 211-217.
42. M. Kawahara, K. Muramoto, K. Kobayashi, H. Mori and Y. Kuroda, *Biochem. Biophys. Res. Commun.*, 1994, **198**, 531-535.
43. P. W. Mantyh, J. R. Ghilardi, S. Rogers, E. Demaster, C. J. Allen, E. R. Stimson and J. E. Maggio, *J. Neurochem.*, 1993, **61**, 1171-1174.
44. S. Yumoto, S. Kakimi, A. Ohsaki and A. Ishikawa, *J. Inorg. Biochem.*, 2009, **103**, 1579-1584.
45. M. Abdelghany, A. K. Elsebae and D. Shalloway, *J. Biol. Chem.*, 1993, **268**, 11976-11981.
46. G. D. Fasman, *Coord. Chem. Rev.*, 1996, **149**, 125-165.
47. T. Kiss, *Archives of Gerontology and Geriatrics*, 1995, **21**, 99-112.
48. D. Hollender, A. Karoly-Lakatos, M. Forgo, T. Kortvelyesi, G. Dombi, Z. Majer, M. Holloi, T. Kiss and A. Odani, *J. Inorg. Biochem.*, 2006, **100**, 351-361.
49. A. J. A. Aquino, D. Tunega, G. Haberhauer, M. H. Gerzabek and H. Lischka, *PCCP*, 2001, **3**, 1979-1985.
50. J. Beardmore and C. Exley, *J. Inorg. Biochem.*, 2009, **103**, 205-209.
51. J. Beardmore, G. Rugg and C. Exley, *J. Inorg. Biochem.*, 2007, **101**, 1187-1191.
52. M. B. Hay and S. C. B. Myneni, *J. Phys. Chem. A*, 2008, **112**, 10595-10603.
53. B.-M. Lu, X.-Y. Jin, J. Tang and S.-P. Bi, *J. Mol. Struct.*, 2010, **982**, 9-15.
54. M. I. Lubin, E. J. Bylaska and J. H. Weare, *Chem. Phys. Lett.*, 2000, **322**, 447-453.
55. A. J. Sillanpaa, J. T. Paivarinta, M. J. Hotokka, J. B. Rosenholm and K. E. Laasonen, *J. Phys. Chem. A*, 2001, **105**, 10111-10122.
56. T. W. Swaddle, *Coord. Chem. Rev.*, 2001, **219**, 665-686.
57. T. M. C. Faro, G. P. Thim and M. S. Skaf, *J. Chem. Phys.*, 2010, **132**.
58. G. Schaftenaar and J. H. Noordik, *J. Comput. Aided Mol. Des.*, 2000, **14**, 123-134.
59. B. Hess, C. Kutzner, D. van der Spoel and E. Lindahl, *Journal of Chemical Theory and Computation*, 2008, **4**, 435-447.
60. S. Pronk, S. Pall, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoel, B. Hess and E. Lindahl, *Bioinformatics*, 2013, **29**, 845-854.
61. R. B. Best, X. Zhu and A. D. MacKerell Jr, *Journal of Chemical Theory and Computation*, 2012, **8**, 3257-3273.
62. E. Formoso, J. I. Mujika, J. Grabowski and X. Lopez, *J. Inorg. Biochem.*, 2015, **152**,

139-146.
63. J. I. Mujika, B. Escribano, E. Akhmatskaya, J. M. Ugalde and X. Lopez, *Biochemistry*, 2012, **51**, 7017-7027.
64. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926-935.
65. T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089-10092.
66. U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, *J. Chem. Phys.*, 1995, **103**, 8577-8593.
67. G. Bussi, D. Donadio and M. Parrinello, *J. Chem. Phys.*, 2007, **126**.
68. X. Daura, K. Gademann, B. Jaun, D. Seebach, W. F. van Gunsteren and A. E. Mark, *Angewandte Chemie-International Edition*, 1999, **38**, 236-240.
69. A Alvarez-Ibarra, P. Calaminici, A. Goursot, C. Z. Gómez-Castro, R. Grande-Aztatzi, T. Mineva, D. R. Salahub, J. M. Vásquez-Pérez, A. Vela, B. Zuniga-Gutierrez and A. M. Köster, in *Frontiers in Computational Chemistry*, eds. Zaheer Ul-Haq and J. D. Madura, Bentham Science, 2015, vol. 2, pp. 281-325.
70. A. M. Koster, P. Calaminici, M. E. Casida, V. D. Dominguez, R. Flores-Moreno, G. U. Gamboa, G. Geudtner, A. Goursot, T. Heine, A. Ipatov, F. Janetzko, J. M. del Campo, S. Patchkovskii, J. U. Reveles, A. Vela, B. Zuniga and D. R. Salahub, *Journal*, 2009.
71. A. Alvarez-Ibarra and A. M. Koester, *J. Chem. Phys.*, 2013, **139**, 024102-024101-.
72. B. Zuniga-Gutierrez, G. Geudtner and A. M. Koester, *J. Chem. Phys.*, 2011, **134**.
73. J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865-3868.
74. T. H. Dunning, *J. Chem. Phys.*, 1989, **90**, 1007-1023.
75. D. E. Woon and T. H. Dunning, *J. Chem. Phys.*, 1993, **98**, 1358-1371.
76. B. I. Dunlap, J. W. D. Connolly and J. R. Sabin, *J. Chem. Phys.*, 1979, **71**, 3396-3402.
77. A. M. Koster, J. M. del Campo, F. Janetzko and B. Zuniga-Gutierrez, *J. Chem. Phys.*, 2009, **130**.
78. G. W. T. M. J. Frisch, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, *Journal*, 2009.
79. A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648-5652.
80. C. T. Lee, W. T. Yang and R. G. Parr, *Physical Review B*, 1988, **37**, 785-789.
81. J. Tomasi, B. Mennucci and R. Cammi, *Chem. Rev.*, 2005, **105**, 2999-3093.
82. A. Buczek, M. Makowski, M. Jewginski, R. Latajka, T. Kupka and M. A. Broda, *Biopolymers*, 2013, **101**, 28-40.
83. X. P. Xu and D. A. Case, *J. Biomol. NMR*, 2001, **21**, 321-333.
84. S. Zaretsky, J. L. Hickey, M. A. St. Denis, C. C. G. Scully, A. L. Roughton, D. J. Tantillo, M. W. Lodewyk and A. K. Yudin, *Tetrahedron*, 2014, **70**, 7655-7663.
85. J. Ali-Torres, L. Rodriguez-Santiago and M. Sodupe, *PCCP*, 2011, **13**.
86. N. Luque, J. I. Mujika, E. Formoso and X. Lopez, *Rsc Advances*, 2015, **5**, 63874-63881.