Molecular Modeling of the RNA Binding N-terminal part of CCMV coat protein in Solution with Phosphate Ions.

David van der Spoel[‡] K. Anton Feenstra[†]
Marcus A. Hemminga[†] Herman J.C. Berendsen[‡*]

July 29, 1997

[‡]Bioson Research Institute and Laboratory of Biophysical Chemistry
University of Groningen
Nijenborgh 4
9747 AG Groningen
The Netherlands

[†]Department of Molecular Physics
Agricultural University
P.O. Box 8128
6700 ET Wageningen
The Netherlands

[*]To whom correspondence should be addressed

Keywords

protein-nucleic acid interaction
phosphate binding
protein conformation
phosphate force field
molecular dynamics
ab initio computations

Abbreviations

CCMV	cowpea chlorotic mottle virus
NMR	nuclear magnetic resonance
pep25	pentacosapeptide representing the N-terminus of the coat protein of CCMV
Pi	monophosphate molecule ($H_2PO_4^-$)
Pi4	tetraphosphate molecule ($H_2P_4O_{13}^{4-}$)
Pi8	octaphosphate molecule ($H_2P_8O_{25}^{8-}$)
RNA	ribonucleic acid
MD	Molecular Dynamics
NOE	Nuclear Overhauser Effect
CD	Circular Dichroism
RDF	Radial Distribution Function

Abstract

We have studied the RNA-binding N-terminal arm of the coat protein of cowpea chlorotic mottle virus (CCMV) using 5 Molecular Dynamics simulations of 2.0 ns each. This 25 residue peptide (pep25) is highly charged: it contains 6 Arg and 3 Lys residues. An α-helical fraction of the sequence is stabilized in vitro by salts. We have studied the interaction of mono phosphate (Pi) ions with pep25, and found that only 2 Pi ions are bound to pep25 on average, but water-mediated interactions between pep25 and Pi, providing electrostatic screening for intra-peptide interactions, are abundant. Shielding by the Pi ions of repulsive electrostatic interactions between Arg side chains increases the α-helicity of pep25. A hydrogen-bond at the N-terminal end of the α-helix renders extension of the α-helix in the N-terminal direction impossible, which is in agreement with evidence from Nuclear Magnetic Resonance (NMR) experiments.

1 Introduction

Cowpea chlorotic mottle virus (CCMV) is a spherical plant virus of the group of bromoviruses. It consists of RNA surrounded by 180 identical, icosahedrally arranged protein subunits (Dasgupta & Kaesberg, 1982). The virus particle can easily be dissociated and reassembled in vitro by changing pH and ionic strength (Bancroft & Hiebert, 1967). The highly positively charged N-terminal arm (six Arg and three Lys) consisting of the first 25 amino acids of the coat protein has been found to be essential for binding the encapsulated RNA (Vriend *et al.*, 1981). Based on NMR experiments and secondary structure predictions, Vriend *et al.* (Vriend *et al.*, 1982; Vriend *et al.*, 1986) proposed a "snatch-pull" model for the assembly of CCMV coat protein and RNA. In this model the N-terminal region of the coat protein changes from a random coil conformation into an α-helical conformation located between Arg10 and Asn20 upon binding viral RNA. Chemical synthesis of the pentacosapeptide pep25 that contains the first 25 N-terminal amino acids of CCMV coat protein (ten Kortenaar *et al.*, 1986), allowed detailed spectroscopic studies to test the snatch-pull model. Circular dichroism experiments showed that addition of inorganic salts to pep25 results in an increase of α-helix content, especially in the presence of octadecaphosphate (van der Graaf & Hemminga, 1991). A two-dimensional ^{1}H-NMR study on pep25 indicated the presence of a conformational ensemble consisting of helical structures rapidly converting into more extended states (van der Graaf *et al.*, 1991). The helical region is found to be situated between Thr9 and Ala17, which agrees remarkably well with the secondary structure predictions (Vriend *et al.*, 1986; Vriend *et al.*, 1982). This NMR work has been extended by performing distance geometry calculations of pep25 in the presence of tetraphosphate (Pi4) (van der Graaf *et al.*, 1992). This resulted in eight structures belonging to two structure families. The first family consists of five structures with an α-helix-like conformation in the middle of the peptide, and the second family consists of three structures with a more open conformation. Although the occurrence of two structure families suggests that even in the presence of Pi4 the peptide is flexible, its tendency to form an α-helical conformation upon binding of phosphate groups is suggested to play an essential role in RNA binding (van der Graaf *et al.*, 1992).

It is well known that changes in the environment of a polypeptide chain may induce changes in the conformation (Timasheff, 1992), for example, the unfolding of proteins induced by urea (Sijpkes *et al.*, 1993), or the promotion of α-helix content by alcohols (Brooks III & Nilsson, 1993) or phosphate ions (van der Graaf & Hemminga, 1991). The precise mechanism of action is not known in general, therefore it is interesting to study the binding of small molecules to polypeptides. Theoretical methods such as Molecular Dynamics (MD) simulations are very attractive tools for these studies since they can give atomic detail on a time scale (ns) that is relevant for the motion of these small molecules. MD simulations have been used previously to study the influence of trifluoroethanol and other alcohols on the structure of peptides (de Loof *et al.*, 1992; van Buuren & Berendsen, 1993; Brooks III & Nilsson, 1993; Kovacs *et al.*, 1995) and simulations of protein unfolding are quite common (Mark & van Gunsteren, 1992; Brooks III, 1993; Tirado-Rives & Jorgensen, 1993; Caflisch & Karplus, 1994). The effect of salts on the structure of peptides has

also been studied by simulation techniques (Smith & Pettitt, 1991b; Smith *et al.*, 1993). Refinement of protein structures based on NMR- or X-Ray data is a very important application of MD and many methodological improvements are still being developed (van Schaik *et al.*, 1993; Clarage & Philips Jr, 1994; Fennen *et al.*, 1995). In the context of structure determination, pep25 is a very challenging case: there is a large amount of experimental data from NMR and CD spectroscopy, but the structure is not known in atomic detail. The conformation of pep25 is influenced by the addition of phosphate ions; both concentration and chemical composition of the phosphate molecules are important for the structure of pep25 (van der Graaf *et al.*, 1991). It is important to note that α-helical structure is induced in pep25 not only by neutralization or shielding of the charged residues, but also by specific interaction of phosphate groups with the peptide. Experimental results have shown that in the presence of phosphates the α-helix content is slightly larger than with other salts, and that the use of longer phosphate chains, resembling RNA, drastically increased the α-helicity (van der Graaf *et al.*, 1991).

In this work we have attempted to model pep25 by Molecular Dynamics (MD) simulations in the presence of mono phosphate (Pi) ions. The questions we will address are:
1) where do phosphate ions bind to the peptide,
2) which part of the peptide is α-helical in the presence of phosphates,
3) how stable is the α-helix in the simulation, and
4) why does the α-helix not extend in the N-terminal direction from Arg10.
We will put emphasis on the comparison of simulation data with experimental data, such as $^1H_\alpha$ chemical shifts and α-helicity (corresponding to ellipticity from CD measurements). For comparison, we have also performed simulations of pep25 *without* counter ions, pep25 in the presence of a Pi4 molecule and in the presence of a Pi8 molecule, although these simulations suffer from methodological problems, in particular the treatment of electrostatic interactions (Smith & Pettitt, 1991a) and the limited sampling of phase space.

2 Methods

Starting Structure for pep25

The peptide pep25 resembles the N-terminal arm of the CCMV coat protein; its sequence is given in Table 1. As a starting structure for our simulations we used the distance geometry structure with the largest number of α-helical residues (structure 3 of (van der Graaf *et al.*, 1992)).

Docking phosphate on pep25

The usual procedure of placing ions around a polypeptide chain for an MD simulation, is first solvating the protein, and subsequently replacing water molecules by ions at the positions that are energetically most favorable. Because this does not work for large ions such as Pi, we have employed another method. The starting conformations for the MD

simulations with phosphate were obtained by performing vacuum MD simulations in which pep25 was surrounded by randomly distributed phosphate molecules. To avoid immediate clustering of the phosphate molecules on the charged side chains of pep25 due to the omission of explicit water molecules, a relative dielectric constant ϵ_r of 4 was used to reduce the strength of the coulomb interactions. The peptide was harmonically restrained to its starting conformation with a force constant of 1000 kJ $mol^{-1}nm^{-2}$ to conserve its structure. The simulations were performed at constant temperature using weak coupling to a temperature bath at 300 K with a coupling constant of 0.1 ps (Berendsen *et al.*, 1984). The time step was 1 fs in all vacuum simulations. All bonds were constrained to their equilibrium length using the SHAKE algorithm (Ryckaert *et al.*, 1977). The simulations were performed using the **GROMACS** software package (van der Spoel *et al.*, 1996; Berendsen *et al.*, 1995) on a Silicon Graphics Indy workstation.

A single vacuum run with p25 and nine randomly distributed Pi molecules was performed for 25 ps. The starting conformation for the monophosphate run is the last frame from this simulation, which seemed to have a reasonable distribution of Pi molecules around pep25 and in solution, as judged by visual inspection. The simulation from this starting conformation will be referred to as Sim 9Pi-D (D for Docked). To test what influence the starting conformation has on the distribution of the Pi ions during the simulation, we generated an extra starting conformation using the same peptide structure and nine randomly placed Pi ions *without* docking. The simulation from this starting conformation will be referred to as Sim 9Pi-R (R for Random).

Ten separate vacuum runs with p25 and 3 randomly distributed Pi4 molecules were performed for 100 ps. Because of the high charge on the Pi4 molecules, and the lower mobility of Pi4 relative to Pi because of the larger size, clustering is a more serious problem than with Pi, therefore multiple docking simulations had to be performed. The last time frame of all ten simulations were overlayed to see whether Pi4 clusters to the peptide in specific positions along the sequence. Then, one of these frames was selected having all three Pi4 molecules in or close to a different phosphate cluster. From this conformation, two Pi4 molecules were deleted, leaving the phosphate closest to Arg10, 13, 14 and 18. To keep things simple for Pi8, one Pi8 was manually docked at about the same position the Pi4 has in the starting conformation described above, using the Quanta software package (MSI, 1994). The simulations with these large phosphate chains will be referred to as Sim Pi4 and Sim Pi8 respectively.

Solvation of peptides

For the simulation of pep25 in aqueous solution the peptide was solvated in a box of simple point charge (SPC) water (Berendsen *et al.*, 1981) large enough to accommodate the peptide (approx. 4.8 x 4.8 x 4.8 nm). This water box was built by stacking cubic boxes containing 216 equilibrated water molecules. All water molecules within 0.23 nm of a peptide atom were removed. For the simulations of pep25 with phosphate ions, the phosphate molecules were placed in the box with pep25 as described above before

dissolving the peptide-ion complex. The solvation procedure was then the same as in the run without ions. The total number of atoms in each simulation is given in Table 2. All five starting structures were then subject to energy minimizations using the steepest descents method. During minimization the peptide was harmonically restrained with a force constant of 1000 kJ $mol^{-1}nm^{-2}$. A cut-off for electrostatic interactions of 1.8 nm was used in the energy minimizations. The final structures were used as starting conformations for the production MD runs.

Force field for phosphate ions

In the GROMOS87 force field (van Gunsteren & Berendsen, 1987) no complete parameter set for the phosphate ions is available. Therefore we decided to use the Lennard-Jones parameters, as well as the force constants for angle-bending and dihedral angle torsions from GROMOS87 and determine the charges and geometries from ab-initio calculations. A full geometry optimization of Pi ($H_2PO_4^-$) was done with the 6-31++G** basis-set using the Gaussian-92 package (Frisch *et al.*, 1993). Symmetry was not taken into account during the calculation. Subsequently, we derived the atomic charges for the Pi molecule. The Merz-Kollman population analysis which fits the charges to the molecular electrostatic potential (Besler *et al.*, 1990) was used with the included solvent reaction field as implemented in Gaussian-92. The computation was done using a dielectric constant of 78.5 and a cavity radius of 0.358 nm. Exact symmetry was imposed afterwards by assigning average charges, bond lengths and angles to symmetrically identical atoms or bonds. The resulting parameters are listed in Table 3. The charges are almost 50% higher than those of (Lee & Prohofsky, 1982), probably because of the inclusion of the solvent reaction field.

A full geometry optimization of Pi4 ($H_2P_4O_{13}^{4-}$) was done using the 3-21G basis-set, followed by a Merz-Kollman population analysis with cavity radius of 0.475 nm. A smaller basis set than for Pi was used because of the larger system size (19 atoms). The force constants for the angles OA-P-OA (Pi) and OA-P-OS (Pi4) are taken to be the same as those for all O-P-O angles, except OM-P-OM, consistent with the GROMOS87 force field. The parameters are listed in Table 4.

The geometry of Pi8 ($H_2P_8O_{25}^{8-}$) was not determined by ab-initio calculations, but instead it was copied from Pi4, because of the large system size (35 atoms) and because no large deviations from Pi4 are expected. In effect, the two middle PO_3 groups were each copied two times and inserted between their original position and the central oxygen atom. The charge distribution on Pi8 was again calculated using a Merz-Kollman population analysis with cavity radius of 0.579 nm. The charges are listed in Table 5. A plot of the phosphate molecules is given in Fig. 1

Molecular Dynamics

Five MD simulations of pep25 in water were performed. Sim W had no counter ions, Sim 9Pi-D had nine Pi molecules starting from a docking simulation in vacuum as described

above, Sim 9Pi-R also had nine Pi ions starting from random positions, Sim Pi4 had one Pi4 ion, and Sim Pi8 had one Pi8 ion. The GROMOS forcefield (van Gunsteren & Berendsen, 1987) was used, with increased repulsion between water oxygen and carbon atoms, as suggested by Van Buuren *et al.* (van Buuren *et al.*, 1993), the resulting parameter set is the one referred to as SW by (Daura *et al.*, 1996). Each simulation was 2.0 ns, with a time step of 2 fs. SHAKE (Ryckaert *et al.*, 1977) was used for all bonds. A twin-range cut-off for non-bonded interactions was employed with short-range cut-off for Van der Waals and coulomb interactions of 1.0 nm, and a long-range cut-off of 1.8 nm for coulomb interactions which were calculated during neighbor-list generation, every 20 fs. Sim W and Sim 9Pi-D were actually performed *twice*, the first time with a single cut-off of 1.0 nm, the second time as indicated above. Due to the 1.0 nm cut-off the first simulations were not reliable, as will be explained in the Results section. Therefore we have used the simulations with 1.8 nm long-range interactions for our analysis only. The temperature was controlled using weak coupling to a bath of 300 K (Berendsen *et al.*, 1984) with a time constant of 0.1 ps. Protein and solvent (including phosphate where appropriate) were coupled to the heat bath independently. The pressure was also controlled using weak coupling with a time constant of 1.0 ps. In contrast to the vacuum simulations no shielding of electrostatic interactions was performed, by using a relative dielectric constant ϵ_r of 1.

The MD simulations were carried out using the **GROMACS** software package (van der Spoel *et al.*, 1996) on custom-built parallel computers (Bekker *et al.*, 1993; Berendsen *et al.*, 1995) with Intel $i860$ CPU's and on a Silicon Graphics Power Challenge. On our fastest parallel computer with 28 processors the 2.0 ns simulation of pep25 in water (10,326 particles) with parameters as stated above took 8.6 days of computer time.

Calculation of Chemical Shifts

To be able to compare our simulation data to experimental data we need a method to compute chemical shifts from a structure. For this purpose we have used the empirical method of Williamson and Asakura (Williamson & Asakura, 1993). Recently there has been considerable progress in the computation of chemical shifts from structural data (Spera & Bax, 1991; Ösapay & Case, 1991; Williamson & Asakura, 1993; Ösapay & Case, 1994; Case, 1995). Although there is impressive progress in the ab initio based methods (Pearson *et al.*, 1995; de Dios *et al.*, 1995) they are not practical for our purpose; we want to average the chemical shift over our trajectory of 4000 frames, therefore the computation time must be short. The Williamson method actually calculates the δ value, thus we had to subtract random coil values from the experimental data to be able to compare. For this purpose we used both the recently published values of (Merutka *et al.*, 1995) and the old standard of (Bundi & Wüthrich, 1979) although the differences for Hα protons are not large.

3 Results

Long MD simulations, such as the ones presented in this work, produce enormous amounts of data to be analyzed. Some properties can be averaged over the trajectory or part of the trajectory, and compared to equilibrium properties from experiments. Examples of these are chemical shifts and radial distribution functions. Other properties are interesting when dynamic information is analyzed, such as the binding of phosphate molecules to the peptide or dynamic changes in conformation. Most properties however, can be viewed both ways: for example, the formation and breaking of a hydrogen-bond or salt bridge is a dynamic process; the radial distribution function is an average property. We will analyze a number of the properties that can be deduced from the simulations we have performed. It should be noted here, that average properties are meaningful only when the trajectory represents an equilibrium ensemble. Most peptides in solution do not have a well defined conformation, or, in other words, an ensemble of many configurations is present in a peptide solution. We do not pretend to find all possible conformations, or even the most important conformations (with lowest free energy) in a 2.0 ns simulation of a single peptide. Nevertheless, we can calculate a number of average properties from the simulation trajectories. Comparison to experimental data should then point out whether (part of) the simulation data is physically meaningful.

α-Helicity.

The first quantity we analyzed is the α-helicity, which corresponds to the ellipticity at 222 nm as measured by circular dichroism (CD) experiments. The α-helicity is computed from Ramachandran angles according to a formula of Hirst and Brooks (Hirst & Brooks III, 1994). We have plotted the fraction of time that each residue is in an α-helical conformation in Sim W, Sim 9Pi-D and Sim 9Pi-R (Fig. 2). The α-helix from residue 11-16 is present for more than 60 % of the time in Sim 9Pi-D. The smaller peak around Gly4 in Sim 9Pi-D and the single peak at Lys21 can be explained by single residues having α-helical ϕ/ψ angles rather than a local α-helical conformation of the peptide. In pure water (Sim W), the α-helix is not as pronounced as in Sim 9Pi-D, especially notable is the dip at Arg14. The peak at Lys21 is missing, but in the N-terminal part, around Thr5 a small peak is present. The curve for Sim 9Pi-R resembles the one of Sim W in the α-helical region with a notable dip at Arg14, and in contrast to Sim 9Pi-D and Sim W a peak at Arg22. When averaged over the whole peptide and over time, pep25 has 6 α-helical residues in Sim 9Pi-D versus 5 in Sim W and 4.3 in Sim 9Pi-R. It is important to note that the ellipticity at 222 nm as obtained from CD experiments, is an average over *all* ϕ/ψ angles, so it can be compared to our calculated averages. In Table 6 we have listed the averages. Although we can not prove that our simulation represents an equilibrium ensemble, the correspondence between experiment and Sim 9Pi-D is striking.

$^1H_\alpha$ Chemical Shifts.

$^1H_\alpha$ chemical shifts have been measured for pep25 in the presence of different phosphates (van der Graaf *et al.*, 1992). The most striking feature of the δ values (difference from random coil values) when plotted as a function of residue number is the sharp decrease at Arg10 followed by a slow increase to values around 0 ppm (Fig. 3). The sharp decrease at Arg10 has been interpreted (together with NOE data) as the start of the α-helical region of pep25. We have calculated the $^1H_\alpha$ chemical shifts from our trajectories of Sim 9Pi-D and Sim 9Pi-R, using the method of Willamson and Asakura (Williamson & Asakura, 1993), and averaged the shifts over time (Fig. 3). The shifts calculated as an average over Sim 9Pi-D closely follow the experimental pattern, the trend from Ala17 to Thr24, is qualitatively reproduced. The sharp decrease in δ around Arg10 is not as clear in Sim 9Pi-D as in the experimental data, and it seems to be shifted by one residue to Ala11. In Sim 9Pi-R the peak at Leu8 is reproduced quantitatively, but some large deviations around Ala15 are visible as well. In the N-terminal part of the peptide there is a significant difference between experiment and both Sim 9Pi-D and Sim 9Pi-R. Since the experimental shifts are very close to random coil values, this suggests that some residual structure is present in pep25 during the simulations, which is not observed experimentally.

Conformation of pep25 in the simulations.

Snapshots from the trajectories of Sim W, Sim 9Pi-D and Sim 9Pi-R are plotted in Fig. 4 (Sim Pi4 and Sim Pi8 are not shown). In all simulations the starting conformation of pep25 is the same. The starting positions of the Pi ions are clearly different in Sim 9Pi-D and Sim 9Pi-R, in the former most ions are indeed docked on the peptide. The Arg residues in the α-helical region are pointing out of the plane of the paper.

After 1000 ps in Sim W the N- and C-terminal ends of pep25 have begun to collapse on top of the α-helical region, a process which proceeds until the end of the simulation. After 2000 ps the backbone is curled up and all charged residues are pointing outward. Strangely, the α-helix seems not to be affected to a large degree and a fair amount of α-helix is present in the end structure of Sim W. In Sim 9Pi-D four of the Pi molecules are initially located close to the α-helical region. The peptide seems to behave more like one would intuitively expect: it stretches (at 1000 ps) and then contracts again (2000 ps). The Pi ions are swarming around pep25, they seem not at all tightly bound. Both in Sim 9Pi-D and Sim 9Pi-R the α-helix deforms somewhat, but the final structures are still more or less α-helical.

Secondary Structure

A revealing view of a protein or peptide simulation is a plot of the secondary structure as a function of time. In Fig. 5 we have plotted the secondary structure, calculated by the DSSP program (Kabsch & Sander, 1983). We see that in Sim W the peptide is α-helical

from residue 11-20 with a single residue in the middle (Ala15) that is not α-helical. After 250 ps the N-terminal part of the α-helix unfolds; in the rest of the simulations residues 11-14 alternate between α-helix, 3_{10}-helix and π-helix. From about 1800 ps, the α-helix is complete with a length of 9 residues (t=2000 ps). During the whole simulation the N-terminal part of the peptide forms many hydrogen-bonded turns. In Sim 9Pi-D the α-helical region of the peptide alternates between π-helix, 3_{10}-helix and α-helix, where the C-terminal end is less stable than the N-terminal end. After 500 ps however, the α-helix is rather stable, and it remains so for the rest of the simulation, with little structure in the rest of the peptide. From 0 through 1320 ps the α-helix starts at Ala11, but from 1340 to 2000 ps it starts at Arg10. The peptide in Sim 9Pi-R behaves similar to the one in Sim W with the C-terminal end being very stable and the N-terminal end being α-helical predominantly in the first half of the simulation.

Binding of Pi to pep25.

We have analyzed the binding of Pi molecules to pep25 in detail. In Fig. 6 we have plotted the distance from each charged residue to the closest Pi molecule, calculated as the smallest distance from any Pi atom to any atom of the charged groups in the peptide (NZ+protons for Lys and NE,CZ,NH1,NH2+protons for Arg). This way the black spots indicate direct hydrogen-bonding, and one shade of grey lighter corresponds to water mediated interactions. Initially the Pi molecules are distributed rather evenly around pep25 in Sim 9Pi-D, while they are far away in Sim 9Pi-R, except for a direct hydrogen-bond with Lys7. During the simulations the distribution of Pi molecules changes; in Sim 9Pi-D the Pi molecules dissolve after 500 ps only to come back later on in the simulation, only Lys7 is close to a Pi all the time. The Pi ions in Sim 9Pi-R find pep25 after 700 ps, and from that time on ions are constantly close to most charged residues.

At a more detailed level, a radial distribution function (RDF) can give information about the environment of individual atoms; it can be used to study hydrogen-bonding and salt bridges and liquid structure in general (Allen & Tildesley, 1987). The RDF is defined as the local density of a certain group of particles around a central particle, and it may be averaged over multiple central particles. In Fig. 7 we have plotted the radial distribution of phosphorus atoms around proton donor groups in pep25. For the central particles we have used the proton donor atoms rather than the hydrogen atoms to reduce noise from small fluctuations. The peak around the neutral groups of Ser/Thr and Asn/Gln is much higher than the one around any of the other groups in Sim 9Pi-D, but not so in Sim 9Pi-R. From this we can conclude two things: the Pi ions starting from random positions do not find the neutral groups, and once Pi binds to these groups, it binds tightly. The Arg and Lys RDF plots are very similar in Sim 9Pi-D and Sim 9Pi-R, indicating that the long-ranged electrostatic interactions are strong enough to attract the Pi ions from their random starting positions in Sim 9Pi-R. The graphs also show a pronounced second peak which may be due to neighboring residues, but also to water-mediated hydrogen-bonds.

Some special hydrogen-bond-configurations between pep25 and Pi molecules are plotted

in Fig. 8. An extended hydrogen-bonding network occurs between Arg10, Arg14 and a Pi molecule. Triple hydrogen-bonding occurs between Lys7, Arg14, Lys19 and Pi. Although double hydrogen-bonds are not particularly abundant, bifurcated hydrogen-bonds (one acceptor, two donors, or vice versa) can be found quite frequently. The average number of pep25-Pi hydrogen-bonds, including multiple hydrogen-bonds with one residue is 4.6 (Sim 9Pi-D) resp. 3.9 (Sim 9Pi-R); the average number of Pi ions hydrogen-bonded to pep25 is 2.3 (Sim 9Pi-D) resp 1.3 (Sim 9Pi-R). We see from this that every Pi that is bound to pep25 makes, on average, two hydrogen-bonds with it in Sim 9Pi-D and three hydrogen-bonds in Sim 9Pi-R.

Hydrogen-bonding involving Thr9.

The hydrogen-bonding network around Thr9 deserves special attention since Thr9 is one of the key residues in the 'snatch-pull' model for RNA binding (Vriend *et al.*, 1986). Hydrogen-bonds from backbone as well as side chain of Gln12 to the Thr9 side chain occur in Sim 9Pi-D. A simultaneous hydrogen-bond of the Gln12 side chain to the Lys7 backbone further stabilizes this hydrogen-bond network. Although this network is found only at the very end of the simulation (1900 to 2000 ps), we also did observe a hydrogen-bond between Gln12-NH and Thr9-Oγ (from 1300 to 2000 ps). For the rest of the simulation the amino group of the Gln12 side chain is hydrogen-bonded to it's own backbone atoms, or those of residues further back in the sequence (Leu8-O, Lys7-O, Thr5-O). The Gln12-Oϵ1 is hydrogen-bonded to its own backbone NH or to other backbone and sidechain atoms (Arg13-NH1, Arg13-NH2, Ala17-NH). Hydrogen-bonds are also observed between the backbone of Leu8 and Gln12 from 0-800 ps and of Thr9 and Arg13 from 1300-2000 ps. In effect, the latter hydrogen-bond implies an extension of the α-helical conformation towards the N-terminus.

Charge-Charge interactions.

In Sim 9Pi-D we do not only have the attractive interactions between Pi and peptide, but also repulsive interactions between Pi molecules mutually and between charged sidechains of Lys and Arg residues. These interactions can cause problems when simulated with a cut-off radius. What typically happens is that two charges of the same sign repel each other until the distance is larger than the cut-off. After that, both charges diffuse randomly until the distance becomes smaller than the cut-off again. When this happens a large number of times, an effective *accumulation of charge* at the cut-off is found. This can be checked quite easily by plotting radial distribution functions. In Fig. 9 we have plotted the RDF of the phosphorus atoms in Pi relative to each other for two simulations with different cut-off, and the RDF of Arg-CZ relative each other in Sim 9Pi-D and Sim W. In the RDF of P for the 1.0 nm cut-off run there is a peak at 0.43 nm corresponding to hydrogen-bonded Pi molecules, but also a broader and higher peak at 1.0 nm, which is exactly the cut-off distance. In the simulation with a 1.8 nm cut-off, the peak at 0.43 nm is higher and there is also a small second peak at 0.7 nm corresponding to water mediated hydrogen-bonds.

The peak at the cut-off distance is significantly lower than in the simulation with cut-off 1.0 nm, nevertheless, there is still an effect. These peaks at the cut-off distance are a well known artefact arising from the use of cut-offs and can only be rigorously avoided by using an appropriately chosen long-range electrostatic description (Berendsen, 1993). Therefore we have used the simulations with 1.8 nm long-range interactions for our analysis only. In the RDF of Arg-CZ around each other we see a different effect. Without Pi there is a clear cut-off effect at 1.8 nm, with Pi there is only a small effect, and the peak at 0.7 nm is significantly higher and broader. This means that the Pi ions supply an effective screening of electrostatic interactions for the Arg-Arg interactions. This occurs in both Sim 9Pi-D and Sim 9Pi-R as the RDF plots are very similar.

4 Discussion

The 25 residue N-terminal fragment (pep25) of the coat protein of cowpea chlorotic mottle virus (CCMV) is a heavily charged peptide (+9). It is not visible in the electron density of the crystal structure (Speir *et al.*, 1995) which was determined using the symmetry of the spherical virus coat that contains 180 icosahedrally arranged protein monomers. NMR work has shown that the first eight residues of pep25 as part of the intact coat protein are mobile in empty viral capsids (without RNA), while the remainder of the peptide shows a measurable deviation from random coil $^1H_\alpha$ chemical shifts, very similar to the isolated pep25 (van der Graaf *et al.*, 1991). NMR- and CD-spectroscopic work has revealed that the isolated peptide in solution has a tendency to form α-helices in the presence of phosphate ions (van der Graaf, 1992). In those experiments, different phosphate molecules were used as a model for RNA, to which pep25 is known to bind in the intact virus particle (Vriend *et al.*, 1986). We have performed molecular dynamics simulations to study the interaction of monophosphate (Pi) ions with pep25, as a step towards a deeper understanding of pep25-RNA binding inside the virus particle. Although our work is not intended as structure refinement, we did compare the MD trajectories to experimental data to verify whether our results are meaningful in the biological context of the pep25-RNA interaction.

Three of our simulations, Sim W, Sim Pi4 and Sim Pi8, suffer from methodological problems. In Sim W the absence of counterions induces large repulsive forces between the nine charged residues. This problem is slightly reduced in Sim Pi4, where the Pi4 molecule provides some electrostatic screening but there still is a net charge of +5 in the system. Both Sim Pi4 and Sim Pi8 suffer from another problem: the absence of electrostatic screening from other ions forces pep25 and the phosphate group to maximize their interaction. In the case of Sim Pi8 all secondary structure in pep25 is lost after 1000 ps due to this effect. In Sim Pi4 this results in partial loss of α-helical structure and the termini of pep25 fold around the Pi4 molecule. Although the conformations that occur in Sim W, Sim Pi4 and Sim Pi8 are not impossible, they do not seem very probable. The electrostatic screening problem is amplified by the use of a cut-off for electrostatic interactions. The cut-off effect is more serious for the Pi8 and Pi4 molecules than for either Pi or pep25, because they carry a higher charge than Pi and they are more rigid than pep25, while the charges are

also more concentrated in Pi4 and Pi8 than in pep25. Both rigidity and charge density of Pi4 and Pi8 will probably induce a large effect on the surrounding water molecules. Another, minor cut-off effect is observed in Sim 9Pi-D: when a Pi molecule is further than the cut-off distance from pep25, it no longer feels the attractive forces exerted by the positively charged residues. However, it may feel the repulsion of the Pi molecules which are still close to pep25. Therefore Pi ions that are far away from pep25, will have a low probability of approaching pep25. The extent of this effect can be assessed by calculating the percentage of Pi which is outside the cut-off from pep25. By integration of the radial distribution function of Pi around pep25 from the cut-off to inifinity, we estimate that on average 3.5 out of 9 Pi ions in Sim 9Pi-D do not see pep25. The result of this will be a reduced electrostatic screening of intra-peptide interactions.

Some methods are available to treat long-range electrostatic interactions in MD simulations, such as Ewald summation (Allen & Tildesley, 1987; Smith & Pettitt, 1991a) or particle-particle, particle-mesh (PPPM) (Hockney & Eastwood, 1981; Luty *et al.*, 1994), but none of these methods has, to our knowledge, been implemented on message-passing parallel computers (Bekker *et al.*, 1993; Berendsen *et al.*, 1995) which we need to perform these long simulations. Therefore we are currently working on an improved long-range electrostatics treatment based on a Poisson solver (Berendsen, 1993), that can be implemented on our parallel computers. Until then, we believe our simulations are state of the art in the field of peptide simulations, since they cover a relatively long time (2.0 ns) and were performed with a relatively long cut-off of 1.8 nm. Although there still is a cut-off effect in Sim 9Pi-D, the radial distribution function of Arg-CZ around Arg-CZ (Fig. 9) shows that the positively charged Arg sidechains can approach each other more often due to the electrostatic screening by the Pi ions. This graph also shows that the strength of the cut-off effect (the height of the peak at the cut-off length of 1.8 nm) is reduced by the addition of counterions in both Sim 9Pi-D and Sim 9Pi-R.

It should be noted that the pep25 concentration in our simulation is only a factor of 2.5 higher than in the NMR experiments that were done at a peptide concentration of 7 mM (corresponding to a peptide:water ratio of 1:7850, whereas we have 1:3264 in Sim 9Pi-D). This means that interactions between different pep25 molecules will often occur, and also that interactions between pep25 and multiple phosphate molecules are very likely. The latter point is supported by experiments which show that an increase of the Pi concentration leads to larger α-helicity of pep25 (van der Graaf & Hemminga, 1991), which can only be explained by more interactions. With pep25 and Pi in a test-tube, there will always be Pi molecules in every direction around a pep25 molecule. This was also the case in Sim 9Pi-D and Sim 9Pi-R where one pep25 molecule was surrounded by 9 Pi molecules. Hence, the sampling problem that is apparent in the other simulations does not affect Sim 9Pi-D. Therefore we conclude that these simulations are good enough for, at least, a qualitative interpretation of the data.

Correspondence with Experimental Data

We have compared our simulation data of pep25 in Sim 9Pi-D to the relevant experimental data. The correspondence between CD spectroscopy and α-helicity in Sim 9Pi-D (Table 6) is remarkable. It should be kept in mind however, that the result is biased by the starting structure which contains a rather long α-helical part. The same reasoning can also be applied to justify the results: since we remain relatively close to the starting structure (which was derived from experimental data) our simulation is sampling the configurational space close to a relevant structure.

A good qualitative agreement between measured and calculated $^1H_\alpha$ chemical shifts was found. The trend of the experimental data is reproduced (Fig. 3), although there is no quantitative agreement. Especially the ten N-terminal residues deviate from the experimental data; part of this effect may be caused by binding of Pi to the Thr2 and Thr5 sidechains and to interactions of Pi with backbone atoms (Fig. 6). Since the experimental data show that this region of the peptide is in a random coil conformation (van der Graaf, 1992), the larger deviation indicates that the first 10 residues have too much secondary structure in Sim 9Pi-D. From the secondary structure plot (Fig. 5) and from the Pi binding plot (Fig. 6) it can be seen that the amount of secondary structure in the first 10 residues is reduced when no Pi is bound to it (after 1000 ps). This suggests that a longer simulation may improve the correspondence of $^1H_\alpha$ chemical shifts between experiment and simulation for this region of the peptide. Since the Williamson & Asakura method for calculating chemical shifts (Williamson & Asakura, 1993) is an empirical method that is fitted on $^1H_\alpha$ atoms, we expect good correspondence between calculated and experimental data since we only consider $^1H_\alpha$ atoms. Furthermore, no aromatic residues are present in pep25, so no ring-current effects occur which may produce large $^1H_\alpha$ shifts. The only problem that may arise in the calculation of chemical shifts from structural data, is the contribution of the local electric field, which does not take the solvent into account. While it is hard to assess the strength of this effect, we think that differences in $^1H_\alpha$ chemical shifts should be attributed primarily to different local conformation or insufficient sampling. Nevertheless, the qualitative agreement of experimental chemical shift data with our simulation data, together with the good agreement of α-helicity demonstrates that the trajectory of Sim 9Pi-D is a physically meaningful data set, which is both valid and useful for further analysis. It seems that the measured $^1H_\alpha$ chemical shift delta values are closer to 0 using the random coil values of (Merutka *et al.*, 1995) than with those of (Bundi & Wüthrich, 1979). This may be caused by the difference in pH at which the random coil values were measured (pH 5.0 for Merutka *et al.* vs. pH 7.0 for Bundi & Wüthrich); pH 5.0 was also used for the experiments with pep25.

The results from Sim 9Pi-R are not nearly as good as those from Sim 9Pi-D. This means that the starting structure, including the positions of the ions relative to the peptides is very important for the final result. Although some features of the chemical shift curve (Fig. 3) are reproduced very well in Sim 9Pi-R, like the peak at Leu8, the overall correspondance with experiment is not good, in particular the strange peak at Ala15. This peak corresponds with a dip in the α-helicity curve at Arg14 (Fig. 2). When the peptide

is α-helical there is substantial repulsion between the Arg13/14 and Arg18 sidechains if there is no electrostatic screening, like in the beginning of Sim 9Pi-R when the Pi ions are far away from the Arg residues. Breaking of the α-helix may be the simplest way to increase the distance between the positively charged residues, thereby facilitating water insertion into the backbone (DiCapua *et al.*, 1990).

Interactions between Pi and pep25

All interactions between pep25 and Pi are largely electrostatic in nature. Nevertheless, it is useful to distinguish a number of different types of interactions: hydrogen-bonds between Pi and pep25 backbone, hydrogen-bonds with neutral hydrogen-bond forming sidechains, and ionic interaction between Pi and charged sidechains. The latter type is essentially different because of its long range nature. Direct hydrogen-bonds between Pi and uncharged sidechains (Ser, Thr, and Asn) seem to be relatively strong once formed (Fig. 7). However, in Sim 9Pi-R, where such interactions are not formed initially, the Pi ions do not "find" the uncharged sidechains in the 2.0 ns simulation. Thus we can not draw any conclusions on the amount of Pi around these sidechains, since our simulations have not converged. For Pi binding to charged Lys and Arg sidechains in contrast, we *do* find converged RDF plots (see Figs. 7 and 9). We also plotted RDFs for the first and last nanoseconds of our simulations (data not shown). In this case, there was no convergence between the simulations, and no convergence between first and last nanosecond of our simulations either. From these findings we conclude that the 2.0 ns simulation is the smallest simulation time long enough to sample charge-charge configurations. One can expect that the sampling will be worse when fewer ions are used, and vice-versa. To obtain a proper sampling of the interactions between Pi and the neutral sidechains probably requires a simulation time which is an order of magnitude larger than the 2.0 ns we have performed. Another feature of the RDFs is that both Arg and Lys have a broad distribution of Pi (Fig. 7) and a large secondary peak, corresponding to water mediated hydrogen-bonds, which is virtually absent in the other RDFs. The observation that Pi interacts with pep25 mostly through water mediated hydrogen bonds is similar to what was found in a simulation study of Enkaphalin in 1M Acetate solution (Smith *et al.*, 1993).

Binding of Pi to pep25 involves some kind of cooperativity: once a Pi ion is bound, to either a sidechain or backbone atom, it very often binds to more than one residue simultaneously or makes multiple hydrogen-bonds with the same residue. Despite the high charge of +9 on pep25, only a small number of Pi ions bind to pep25; apparently the desolvation of both Pi and pep25 is energetically unfavorable.

The promotion of α-helix content by Pi in Sim 9Pi-D is primarily an effect of electrostatic screening. The experimental results can also be explained this way: increasing the salt concentration increases the charge screening effect. The small difference between phosphates and other salts is that they sometimes, but certainly not often, bind through direct hydrogen-bonds. One effect of longer phosphate molecules is that the repulsion between negatively charged groups cannot increase their distance, and therefore screening

is increased even more. It is likely that the possibility of making several ionic contacts simultaneously, together with the higher rigidity of the longer phosphate chains, meaning less entropy loss upon binding, will make binding of oligophosphates more favorable than binding of Pi ions. It seems plausible that the same reasoning can be applied to binding of pep25 to RNA.

α-Helix initiation.

It was suggested that Gln12 and Thr9 are involved in the primary loop formation of the α-helical structure. This would involve a stable hydrogen-bond between Gln12 and Thr9 (van der Graaf *et al.*, 1992). We did indeed find hydrogen-bonds between the Gln12 sidechain and backbone and the Thr9 sidechain. These hydrogen-bonds were not present in our starting conformation but were formed during the simulation. Interestingly, a hydrogen-bond between Thr9-CO and Arg13-NH was found from 1300 to 2000 ps, simultaneously with the Thr9-Oγ to Gln12-NH hydrogen-bond, while neither the DSSP program (Kabsch & Sander, 1983) (Fig. 5), nor the Ramachandran angle based criterium of Hirst and Brooks (Hirst & Brooks III, 1994) (Fig. 2) determined Thr9 to be in an α-helical conformation. Arg10 however *is* in α-helical conformation from 1340 ps to the end of the simulation. The ϕ-angle of Arg10 changes from +60 to -60, entering the α-helical region of the Ramachandran plot shortly after the Thr9-CO to Arg13-NH hydrogen-bond is formed. In contrast, the Thr9 residue is found to be in the β-sheet region of the Ramachandran plot with a positive ψ angle. Therefore we conclude that indeed the Thr9 and Gln12 are involved in a start of the α-helix and that the α-helix can not extend in the N-terminal direction because of the "wrong" ϕ/ψ angles of residue Thr9. An NOE connectivity $d_{\alpha N}$ between Thr9 and Ala11 that was found by NMR experiments (van der Graaf, 1992), is also present in Sim 9Pi-D: Thr9-C_α and Ala11-H are never more than 0.4 nm apart.

Conclusions

Of the five MD simulation that we have performed, Sim 9Pi-D is the best one in methodological terms. We have shown that for this simulation, there is at least qualitative agreement with experimental $^1H_\alpha$ chemical shift data and α-helicity from CD experiments. The interaction between Thr9 and Gln12 that was suggested on the basis of chemical shifts was indeed found. Due to this interaction, the ϕ/ψ angles of residue Thr9 are not α-helical, thereby prohibiting α-helix extension in the N-terminal direction. The hydrogen-bond between Thr9 and Gln12 probably induces a transition of Arg10 to the α-helical region of the Ramachandran plot. In our simulation the region from residue Arg10 through Lys19 is the α-helical part, which corresponds very well to experimental results (van der Graaf, 1992). The control simulation starting from random Pi positions (Sim 9Pi-R) shows that in 2.0 ns the charge-charge interactions can be sampled adequately, although the overall equilibration of phosphate distribution is not achieved. The latter may imply that the interactions of Pi with neutral sidechains in Sim 9Pi-D are exaggerated due to the docking

procedure. The interaction of Pi with charged side chains is very often mediated by water molecules; apparently, the complete desolvation of two charged groups is energetically unfavorable.

Acknowledgements

We would like to thank Dr. Janez Mavri for assistance with the ab-initio calculations and Dr. Gert Vriend and Dr. Tjeerd Schaafsma for valuable discussions. DvdS acknowledges support from the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for Scientific Research (NWO).

References

Allen, M. P. & Tildesley, D. J. (1987). *Computer Simulations of Liquids.* Oxford: Oxford Science Publications.

Bacon, D. J. & Anderson, W. F. (1988). A fast algorithm for rendering space-filling molecule pictures. *J. Molec. Graph.* **6**, 219–220.

Bancroft, J. B. & Hiebert, E. (1967). Formation of an infectious nucleoprotein from protein and nucleic acid isolated from a small spherical virus. *Virology,* **32**, 354–356.

Bekker, H., Berendsen, H. J. C., Dijkstra, E. J., Achterop, S., van Drunen, R., van der Spoel, D., Sijbers, A., Keegstra, H., Reitsma, B., & Renardus, M. K. R. (1993). Gromacs: A parallel computer for molecular dynamics simulations. In: *Physics Computing 92*, (de Groot, R. A. & Nadrchal, J., eds) pp. 252–256, Singapore: World Scientific.

Berendsen, H. J. C. (1993). Electrostatic interactions. In: *Computer Simulation of Biomolecular Systems*, (van Gunsteren, W. F., Weiner, P. K., & Wilkinson, A. J., eds) pp. 161–181. ESCOM Leiden.

Berendsen, H. J. C., Postma, J. P. M., DiNola, A., & Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690.

Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., & Hermans, J. (1981). Interaction models for water in relation to protein hydration. In: *Intermolecular Forces* pp. 331–342. D. Reidel Publishing Company Dordrecht.

Berendsen, H. J. C., van der Spoel, D., & van Drunen, R. (1995). GROMACS: A message-passing parallel molecular dynamics implementation. *Comp. Phys. Comm.* **91**, 43–56.

Besler, B. H., Merz Jr., K. M., & Kollman, P. A. (1990). Atomic charges derived from semiempirical methods. *J. Comp. Chem.* **11**, 431–439.

Brooks III, C. L. (1993). Molecular simulations of peptide and protein unfolding: in quest of a molten globule. *Curr. Opin. Struct. Biol.* **3**, 92–98.

Brooks III, C. L. & Nilsson, L. (1993). Promotion of helix formation in peptides dissolved in alcohol and water-alcohol mixtures. *J. Am. Chem. Soc.* **115**, 11034–11035.

Bundi, A. & Wüthrich, K. (1979). ^{1}H-NMR parameters of the common amino acid residues measured in aqueous solutions of the linear tetrapeptides H-Gly-Gly-X-L-Ala-OH. *Biopolymers,* **18**, 285–297.

Caflisch, A. & Karplus, M. (1994). Molecular dynamics studies of protein and peptide folding and unfolding. In: *The Protein Folding Problem and Tertiary Structure Prediction*, (Merz Jr., K. M. & Le Grand, S. M., eds) pp. 193–230. Birkhäuser Boston.

Case, D. A. (1995). Calibration of ring-current effects in proteins and nucleic acids. *J. Biomol. NMR,* **6**, 341–346.

Clarage, J. B. & Philips Jr, G. N. (1994). Cross-validation tests of time-averaged molecular dynamics refinements for determination of protein structures by X-ray crystallography. *Act. Cryst. D.* **50**, 24–36.

Dasgupta, R. & Kaesberg, P. (1982). Complete nucleotide sequence of the coat protein messenger RNAs of brome mosaic virus and cowpea chlorotic mottle virus. *Nucl. Acid. Res.* **10**, 703–713.

Daura, X., Oliva, B., Querol, E., Avilés, F. X., & Tapia, O. (1996). On the sensitivity of MD trajectories to changes in water-protein interaction parameters: The potato carboxypeptidase inhibitor in water as a test case for the GROMOS force field. *PROTEINS: Struct. Funct. Gen.* **25**, 89–103.

de Dios, A. C., Pearson, J. G., & Oldfield, E. (1995). Secondary and tertiary structural effects on protein NMR chemical shifts: An ab initio approach. *Science,* **260**, 1491–1496.

de Loof, H., Nilsson, L., & Rigler, R. (1992). Molecular dynamics simulations of galanin in aqueous and nonaqueous solution. *J. Am. Chem. Soc.* **114**, 4028–4035.

DiCapua, F. M., Swaminathan, S., & Beveridge, D. L. (1990). Theoretical evidence for destabilization of an α helix by water insertion: Molecular dynamics of hydrated decaalanine. *J. Am. Chem. Soc.* **112**, 6768–6771.

Fennen, J., Torda, A. E., & van Gunsteren, W. F. (1995). Structure refinement with molecular dynamics and Boltzmann-weighted ensemble. *J. Biomol. NMR,* **6**, 163–170.

Frisch, M. J., Trucks, G. W., Schlegel, H. B., Gill, P. M. W., Johnson, B. G., Wong, M. W., Foresman, J. B., Robb, M. A., Head-Gordon, M., Replogle, E. S., Gomperts, R., Andres, J. L., Raghavachari, K., Binkley, J. S., Gonzalez, C., Martin, R. L., Fox, D. J., Defrees, D. J., Baker, J., Stewart, J. J. P., & Pople, J. A. (1993). *Gaussian 92/DFT, Revision G. 1.* Gaussian, Inc. Pittsburgh PA.

Hirst, J. D. & Brooks III, C. L. (1994). Helicity, circular dichroism and molecular dynamics of proteins. *J. Mol. Biol.* **243**, 173–178.

Hockney, R. W. & Eastwood, J. W. (1981). *Computer simulation using particles.* New York: McGraw-Hill.

Kabsch, W. & Sander, C. (1983). Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers,* **22**, 2577–2637.

Kovacs, H., Mark, A. E., Johansson, J., & van Gunsteren, W. F. (1995). The effect of environment on the stability of an integral membrane helix: Molecular dynamics simulations of surfactant protein C in chloroform, methanol and water. *J. Mol. Biol.* **247**, 808–822.

Kraulis, P. J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst.* **24**, 946–950.

Lee, W. K. & Prohofsky, E. W. (1982). A molecular dynamics study of the solvation of a sodium ion bound to dihydrogen phosphate ion. *Chem. Phys. Lett.* **85**, 98–102.

Luty, B. A., Davies, M. E., Tironi, I. G., & van Gunsteren, W. F. (1994). A comparison of particle-particle, particle-mesh and ewald methods for calculating electrostatic interactions in periodic molecular systems. *Mol. Sim.* **14**, 11–20.

Mark, A. E. & van Gunsteren, W. F. (1992). Simulation of the thermal denaturation of hen egg white lysozym: Trapping the molten globule state. *Biochemistry,* **31**, 7745–7748.

Merritt, E. A. & Murphy, M. E. P. (1994). Raster3D version 2.0: A program for photorealistic molecular graphics. *Act. Cryst. D.* **50**, 869–873.

Merutka, G., Dyson, H. J., & Wright, P. E. (1995). 'Random coil' ^{1}H chemical shifts obtained as a function of temperature and trifluoroethanol concentration for the peptide series GGXGG. *J. Biomol. NMR,* **5**, 14–24.

MSI (1994). *QUANTA 3.0.* Molecular Simulations Incorporated United Kingdom.

Ösapay, K. & Case, D. A. (1991). A new analysis of proton chemical shifts. *J. Am. Chem. Soc.* **113**, 9436–9444.

Ösapay, K. & Case, D. A. (1994). Analysis of proton chemical shifts in regular secondary structure of proteins. *J. Biomol. NMR,* **4**, 215–230.

Pearson, J. G., Wang, J. F., Markley, J. L., Le, H. B., & Oldfield, E. (1995). Protein structure refinement using carbon-13 nuclear magnetic resonance spectroscopic chemical shifts and quantum chemistry. *J. Am. Chem. Soc.* **117**, 8823–8829.

Ryckaert, J. P., Ciccotti, G., & Berendsen, H. J. C. (1977). Numerical integration of the cartesian equations of motion of a system with constraints; molecular dynamics of n-alkanes. *J. Comp. Phys.* **23**, 327–341.

Sijpkes, A. H., van der Kleut, G., & Gill, S. C. (1993). Urea-diketopiperazine interactions: A model for urea induced denaturation of proteins. *Bioph. Chem.* **46**, 171.

Smith, P. E., Marlow, G. E., & montgomery Pettitt, B. (1993). Peptides in ionic solutions: A simulation study of a bis(penicillamine) enkaphalin in sodium acetate solution. *J. Am. Chem. Soc.* **115**, 7493–7498.

Smith, P. E. & Pettitt, B. M. (1991a). Peptides in ionic solutions: A comparison of the ewald and switching function techniques. *J. Chem. Phys.* **95**, 8430–8441.

Smith, P. E. & Pettitt, B. M. (1991b). Effects of salt on the structure and dynamics of the bis(penicillamine) enkaphalin zwitterion: A simulation study. *J. Am. Chem. Soc.* **113**, 6029–6037.

Speir, J. A., Munshi, S., Wang, G., Baker, T. S., & Johnson, J. E. (1995). Structures of the native and swollen forms of cowpea chlorotic mottle virus determined by X-ray crystallography and cryo-electron microscopy. *Structure,* **3**, 63–78.

Spera, S. & Bax, A. (1991). Empirical correlation between protein backbone conformation and C_α and C_β ^{13}C nuclear magnetic resonance chemical shifts. *J. Am. Chem. Soc.* **113**, 5490–5492.

ten Kortenaar, P. B. W., Krüse, J., Hemminga, M. A., & Tesser, G. I. (1986). Synthesis of an arginine rich fragment of a viral coat protein using guanidium functions exclusively. *Int. J. Pept. Prot. Res.* **27**, 401–413.

Timasheff, S. N. (1992). Solvent effects on protein stability. *Curr. Opin. Struct. Biol.* **2**, 35–39.

Tirado-Rives, J. & Jorgensen, W. L. (1993). Molecular dynamics simulations of the unfolding of apomyoglobin in water. *Biochemistry,* **33**, 4175–4184.

van Buuren, A. R. & Berendsen, H. J. C. (1993). Molecular dynamics simulation of the stability of a 22 residue alpha-helix in water and 30 % trifluoroethanol. *Biopolymers,* **33**, 1159–1166.

van Buuren, A. R., Marrink, S. J., & Berendsen, H. J. C. (1993). A molecular dynamics study of the decane/water interface. *J. Phys. Chem.* **97**, 9206–9212.

van der Graaf, M. (1992). *Conformation of the RNA-binding N-terminus of the coat protein of cowpea chlorotic mottle virus.* PhD thesis Wageningen Agricultural University Wageningen, The Netherlands.

van der Graaf, M. & Hemminga, M. A. (1991). Conformational studies on a peptide representing the RNA-binding N-terminus of a viral caot protein using circular dichroism and NMR spectroscopy. *Eur. J. Biochem.* **201**, 489–494.

van der Graaf, M., Kroon, G. J. A., & Hemminga, M. A. (1991). Conformation and mobility of the RNA-binding N-terminal part of the intact coat protein of cowpea chlorotic mottle virus. A two dimensional Nuclear Magnetice Resonance study. *J. Mol. Biol.* **220**, 701–709.

van der Graaf, M., Scheek, R. M., van der Linden, C. C., & Hemminga, M. A. (1992). Conformation of a pentacosapeptide representing the RNA-binding N-terminus of cowpea chlorotic mottle virus coat protein in the presence of oligophosphates: A

two-dimensional proton Nuclear Magnetic Resonance and distance geometry study. *Biochemistry,* **31**, 9177–9182.

van der Spoel, D., van Buuren, A. R., Apol, E., Meulenhoff, P. J., Tieleman, D. P., Sijbers, A. L. T. M., van Drunen, R., & Berendsen, H. J. C. (1996). *Gromacs User Manual version 1.3.* Nijenborgh 4, 9747 AG Groningen, The Netherlands. Internet: `http://rugmd0.chem.rug.nl/~gmx`.

van Gunsteren, W. F. & Berendsen, H. J. C. (1987). *Gromos-87 manual.* Biomos BV Nijenborgh 4, 9747 AG Groningen, The Netherlands.

van Schaik, R. C., Berendsen, H. J. C., Torda, A. E., & van Gunsteren, W. F. (1993). A structure refinement method based on molecular dynamics in 4 spatial dimensions. *J. Mol. Biol.* **234**, 751–762.

Vriend, G., Hemminga, M. A., Verduin, B. J. M., de Wit, J. L., & Schaafsma, T. J. (1981). Segmental mobility involved in protein-RNA interaction in cowpea chlorotic mottle virus. *FEBS Lett.* **134**, 167–171.

Vriend, G., Verduin, B. J. M., & Hemminga, M. A. (1986). Role of the N-terminal part of the coat protein in the assembly of cowpea chlorotic mottle virus. *J. Mol. Biol.* **191**, 453–460.

Vriend, G., Verduin, B. J. M., Hemminga, M. A., & Schaafsma, T. J. (1982). Mobility involved in protein-RNA interaction in spherical plant viruses, studied by nuclear magnetic resonance spectroscopy. *FEBS Lett.* **145**, 49–52.

Williamson, M. P. & Asakura, T. (1993). Empirical comparisons of models for chemical-shift calculation in proteins. *J. Magn. Reson. Ser. B,* **101**, 63–71.

Table 1: Amino acid sequence of pep25, resembling the N-terminal arm of the CCMV coat protein (Dasgupta & Kaesberg, 1982). The Nac residue at the C-terminus is an $NHCH_3$ group present in the chemically synthesized peptide (ten Kortenaar et al., *1986). Indicated are charged groups (+), and hydrogen-bond forming sidechains, OH = OH group, ON = $CONH_2$ group.*

	OH	OH			OH		+		OH	+		ON	+	+	
Ac	Ser	Thr	Val	Gly	Thr	Gly	Lys	Leu	Thr	Arg	Ala	Gln	Arg	Arg	Ala
	1				5					10					15
			+	+	ON	+	+	ON	OH	+					
	Ala	Ala	Arg	Lys	Asn	Lys	Arg	Asn	Thr	Arg	Nac				
					20					25					

Table 2: Number of atoms in each simulation.

	Sim W	Sim 9Pi-D/R	Sim Pi4	Sim Pi8
pep25	276	276	276	276
Phosphate		63 (9 Pi)	19 (Pi4)	35 (Pi8)
H_2O	10050	9792	8682	9972
Total	10326	10131	8977	10283

Table 3: Charges and geometry of Pi. OA is the alcohol oxygen, OM is the "charged" oxygen (these names correspond to the GROMOS atomtypes). The two OM atoms have different angle to the OA atoms. Dihedral parameters (H-OA-P-OA) were taken from GROMOS. b_0 is the bond length, θ_0 is the bond angle, k_b is the force constant for bond stretching (in kJ mole^{-1} nm^{-2}), k_θ is the force constant for angle bending (in kJ mole^{-1} rad^{-2}).

atom	q (e)
P	1.596
OM	-0.943
OA	-0.777
HO	0.422

atoms	b_0 (nm)	k_b
P-OM	0.1478	376560
P-OA	0.1637	251040
OA-H	0.0943	313800

atoms	θ_0 (°)	k_θ
OM-P-OM	124.8	585.76
OM1-P-OA	105.9	397.48
OM2-P-OA	108.2	397.48
P-OA-H	108.2	397.48
OA-P-OA	101.5	400

Table 4: Charges and geometry of Pi4. OA is the alcohol oxygen, OM is the "charged" oxygen, OS is the ether oxygen (these names correspond to the GROMOS atomtypes). Counting from the central ether ogygen (OS), the OS-P-OS-P dihedral angles are 147 and 222 alternatingly. The parameter b_0 is the bond length, θ_0 is the bond angle, ϕ_0 is the dihedral angle, k_b is the force constant for bond stretching (in kJ mole^{-1} nm^{-2}), k_θ is the force constant for angle bending (in kJ mole^{-1} rad^{-2}), k_ϕ is the force constant for the torsion potential (in kJ mole^{-1}), the multiplicity n_ϕ in the dihedral potential is the number of minima in the potential on the interval of 0 to 360°.

atom	q (e)
P1	1.85
P2	1.95
OM	-1.0
OS1	-0.92
OS2	-0.88
OA	-1.02
HO	0.58

atoms	θ_0 (°)	k_θ
OA-P-OS	97	400
OS-P-OS	102	397.48
OS-P-OM	109	397.48
OA-P-OM	110	397.48
P-OA-H	113	397.48
OM-P-OM	118	585.76
P1-OS1-P2	137	397.48
P2-OS2-P2	147	397.48

atoms	b_0 (nm)	k_b
P-OM	0.155	376560
P-OS	0.165	251040
P-OA	0.164	251040
OA-H	0.101	313800

atoms	ϕ_0 (°)	k_ϕ	n_ϕ
H-OA-P-OS	34	3.0	3
OA-P-OS-P	-19	3.0	3
OS-P-OS-P	-222/-147	3.0	3

Table 5: Charges on the Pi8 molecule. OA is the alcohol oxygen, OM is the "charged" oxygen, OS is the ether oxygen, these names correspond to the GROMOS atomtypes. Bondlengths, angles, dihedrals and force constants are taken from the Pi4 geometry (see Table 4).

atom	q (e)
P1	1.85
P2	1.95
P3	1.79
OM	-1.00
OS1	-0.88
OS2	-0.82
OS3	-0.76
OA	-1.05
HO	0.58

Table 6: Average α-helicity in Sim 9Pi-D/R and Sim W compared to values from CD spectroscopy (van der Graaf & Hemminga, 1991).

MD		CD	
Sim W	20±3%	no salt	21.0%
Sim 9Pi-D	24±3%	50 mM Pi	23.6%
Sim 9Pi-R	18±3%		

List of Figures

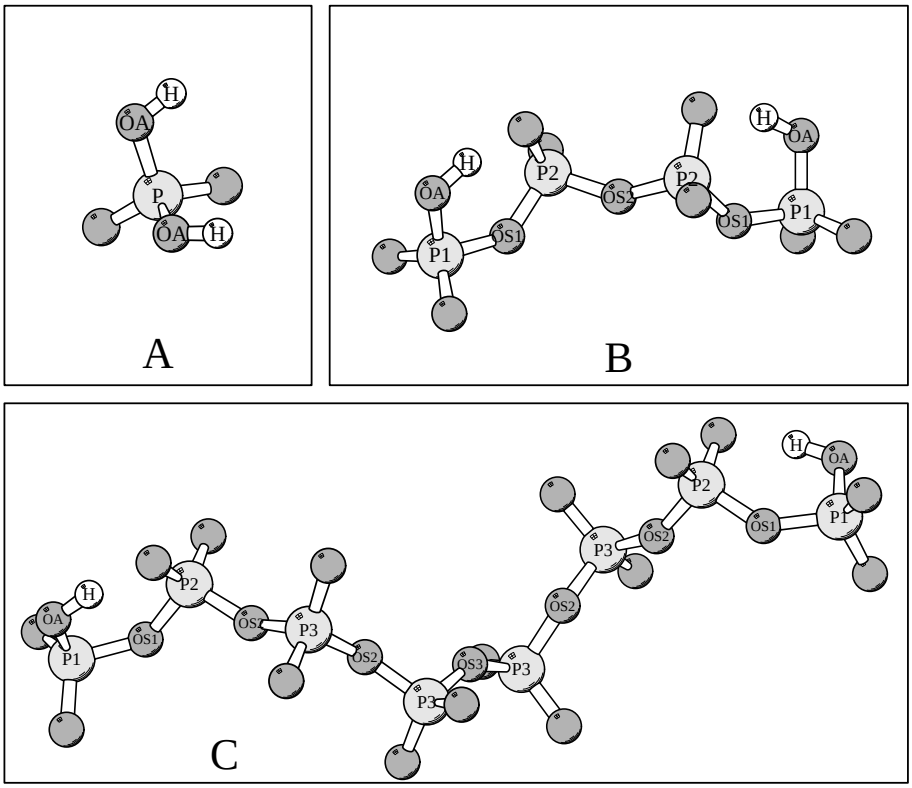

Figure 1: The Pi(A), Pi4(B), and Pi8(C) molecules. Plot was made with MOLSCRIPT (Kraulis, 1991).

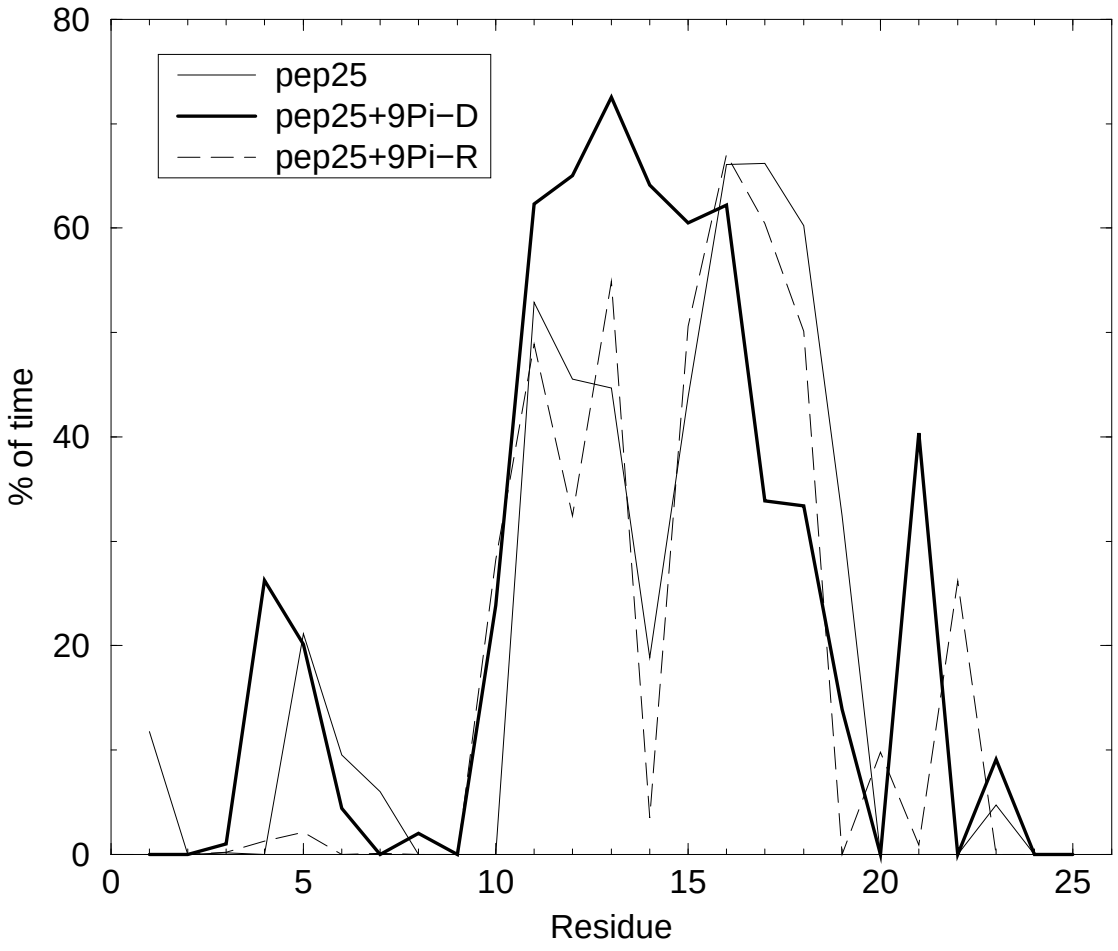

Figure 2: α-helicity of all residues of pep25 averaged over time in Sim W and Sim 9Pi-D, according to the criterion of Hirst and Brooks (Hirst & Brooks III, 1994).

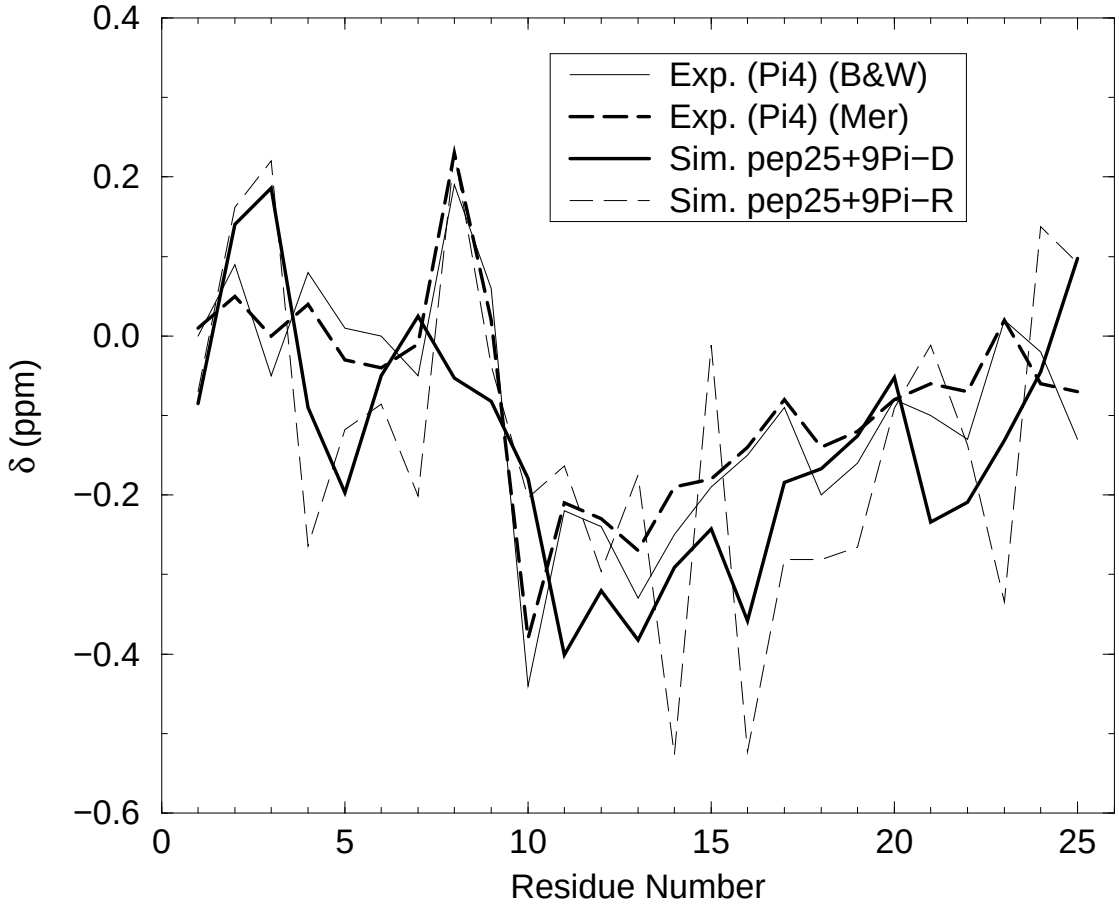

Figure 3: Conformational Cα-H chemical shifts measured by NMR in the presence of Pi4 (van der Graaf et al., *1992) and calculated from Sim 9Pi-D and Sim 9Pi-R using the empirical method of (Williamson & Asakura, 1993). (Mer) indicates random coil values of (Merutka* et al., *1995), (B&W) indicates random coil values of (Bundi & Wüthrich, 1979).*

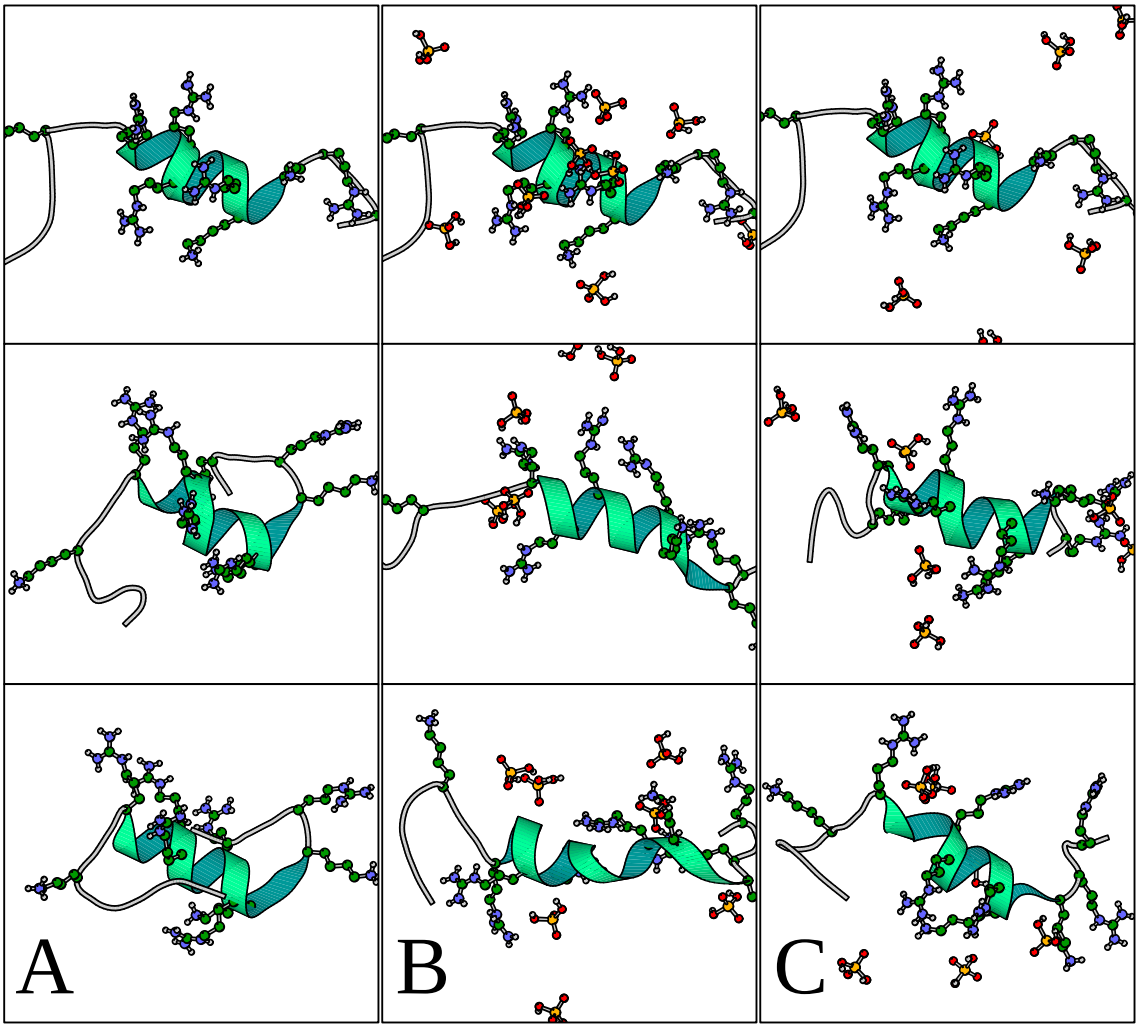

Figure 4: p25 molecules at (top to bottom) t=0, t=1000 ps and t=2000 ps in **A:** *Sim W,* **B:** *Sim 9Pi-D,* **C:** *Sim 9Pi-R. The peptides were rotated such that the Arg residues are pointing out of the plane. Plots were made with MOLSCRIPT (Kraulis, 1991) and Raster3D (Bacon & Anderson, 1988; Merritt & Murphy, 1994).*

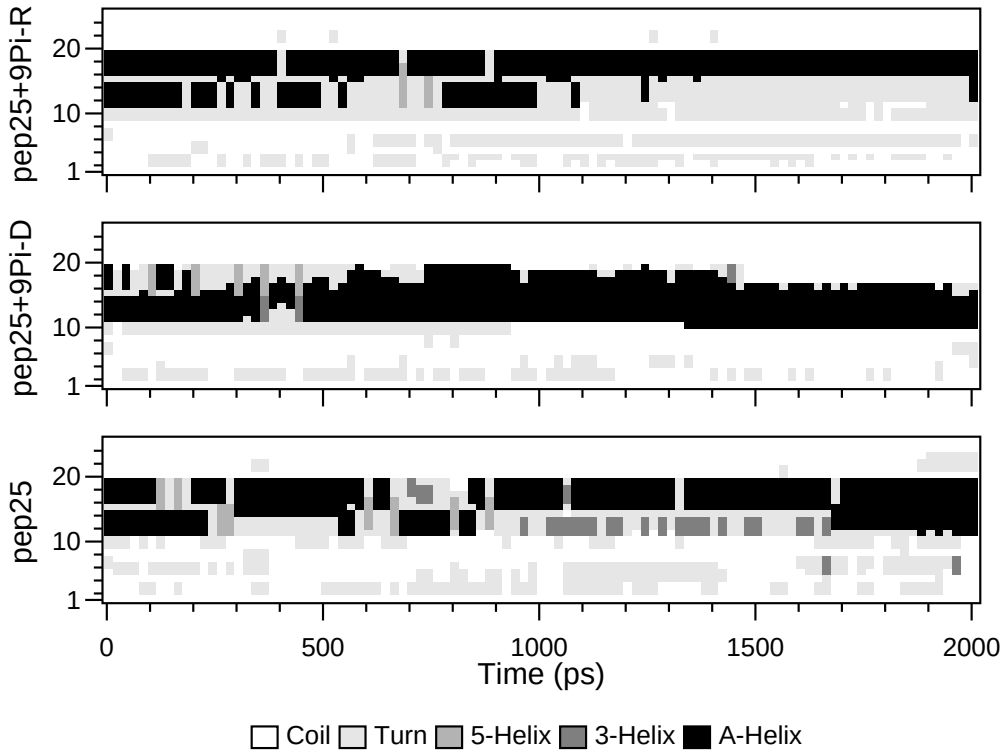

Figure 5: Secondary structure of pep25 in Sim W, Sim 9Pi-D and Sim 9Pi-R.

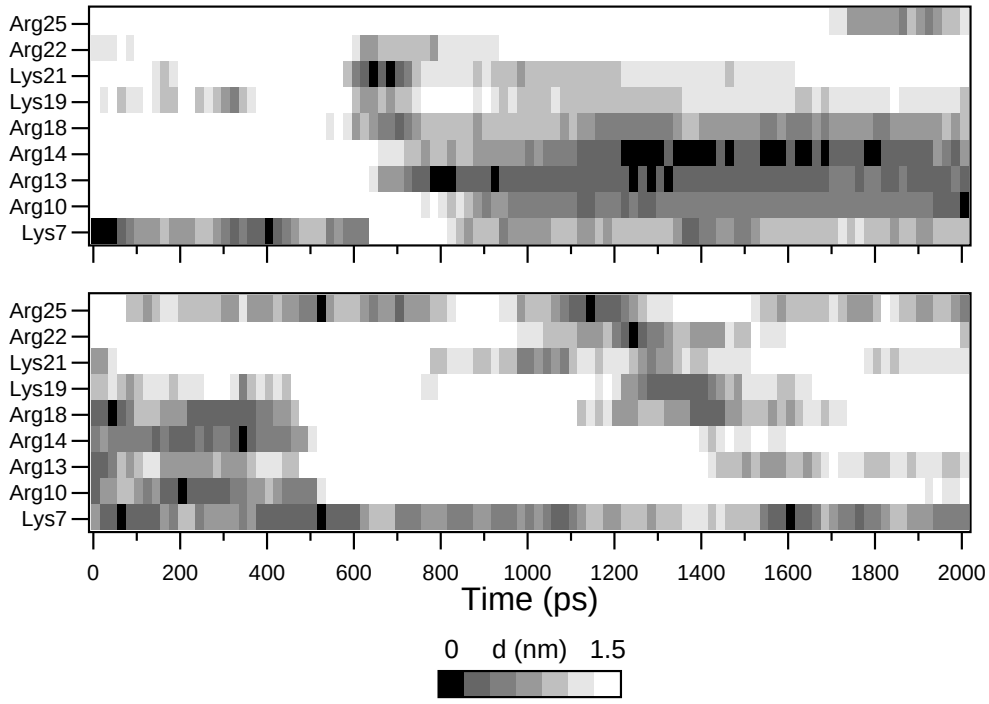

Figure 6: Smallest distance between any Pi atom to any atom of a charged sidechain of pep25 in Sim 9Pi-D (bottom) and Sim 9Pi-R (top).

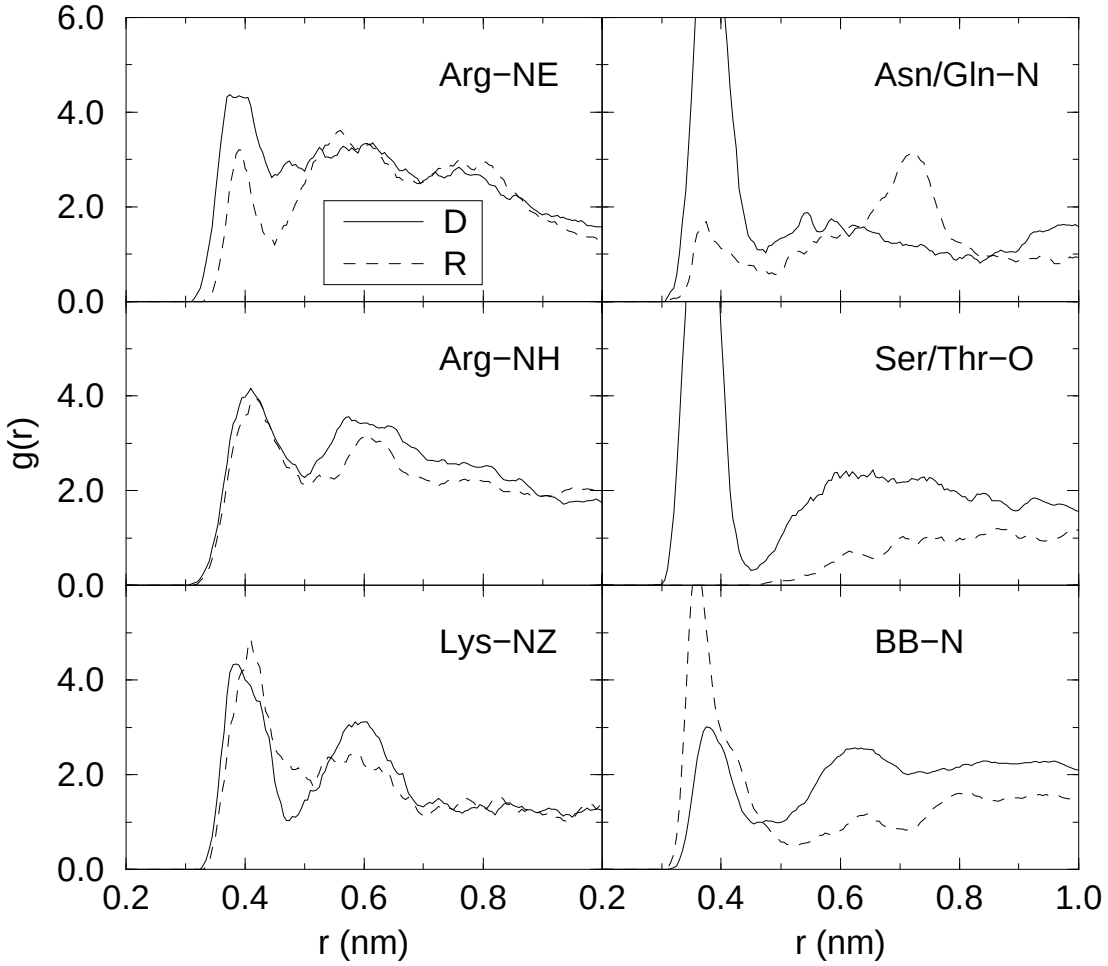

Figure 7: Radial distribution function of Pi-P atoms around sidechain and backbone proton donors. The central particles over which was averaged are denoted in the graphs. A running average over 0.025 nm was used to enhance clarity.

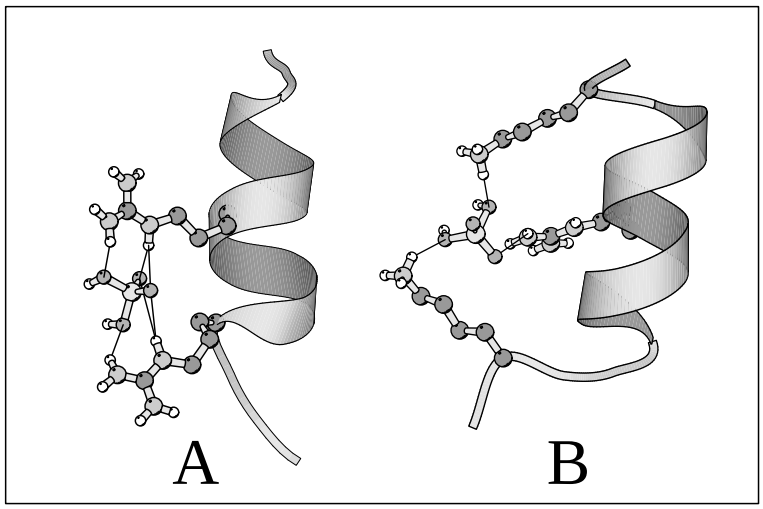

Figure 8: Hydrogen bonding between Pi and pep25 sidechains. **A***: Pi ion trapped between Arg10 and Arg14,* **B***: interaction between a Pi ion and Lys7, Arg14 and Lys19. Plots were made with MOLSCRIPT (Kraulis, 1991) and Raster3D (Bacon & Anderson, 1988; Merritt & Murphy, 1994).*

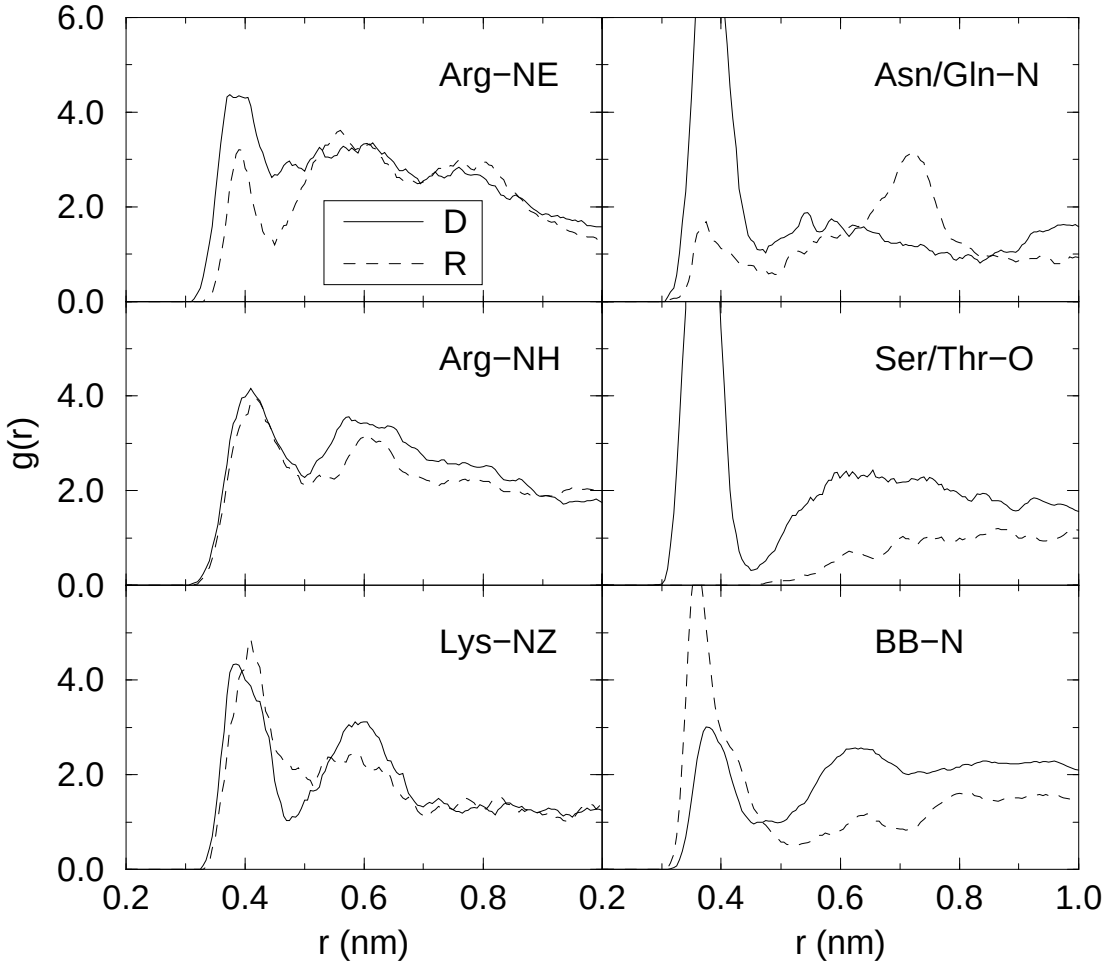

Figure 9: Radial distribution function of Pi-P around Pi-P in Sim 9Pi-D (g_{P-P}, bottom) and of Arg-CZ around Arg-CZ (g_{CZ-CZ}, top). Rc is the cut-off distance for Coulomb interactions. A running average over 0.015 nm was used to enhance clarity.