Journal of
Molecular
Modeling

FULL PAPER

Molecular Modelling of a Multiphosphorylated Sequence Motif Bound to Hydroxyapatite Surfaces

N. Laila Huq, Keith J. Cross, and Eric C. Reynolds

Biochemistry and Molecular Biology Unit, School of Dental Science, The University of Melbourne, 711 Elizabeth Street, Melbourne, 3000, Victoria, Australia. Tel: +61-3-9341-0270; Fax: +61-3-9341-0236; E-mail: e.reynolds@dent.unimelb.edu.au

Abstract Proteins and peptides containing the multiphosphorylated motif -Ser(*P*)-Ser(*P*)-Ser(*P*)--Glu-Glu- stabilise amorphous calcium phosphate (ACP) in body fluids and bind with high affinity to crystalline calcium phosphate phases such as hydroxyapatite (HA) regulating crystal growth. Binding of this motif to hydroxyapatite surfaces was investigated in this study using molecular modelling techniques. Using a three-step computational procedure, we have determined the relative binding energies of the motif Ser(*P*)-Ser(*P*)-Ser(*P*)-Glu-Glu to different crystalline surfaces of HA. This analysis revealed preferences of the motif for (100) and (010) surfaces of the crystal and preferences for particular orientations on a given surface. These preferences are principally governed by electrostatic interactions between the crystal lattice and the peptide with the most stable conformers adopting structures where alternate residues exhibit backbone angles characteristic of a β-strand and values of an α-helix or a distorted α-helix, allowing maximal interaction between the acidic side groups and surface calciums. The results of this study are consistent with experimentally-derived data on the interaction of multiphosphorylated proteins/peptides with HA and have implications for the role of these proteins/ peptides in calcium phosphate stabilisation and biomineralisation processes.

Keywords Hydroxyapatite, Phosphoserine, Protein-mineral interaction

Running title Phosphopeptide hydroxyapatite interactions

Introduction

Many organisms form mineralised structures by the process of biomineralisation [1]. This process may be uncontrolled or may be highly regulated by macromolecules as in the case of matrix-mediated biomineralisation. These macromolecules are usually acidic, either rich in aspartic acid or aspartic and glutamic acids. Some have polysaccharide moieties rich in carboxylate or sulphate groups and some are phosphorylated [2-5].

Several multiple phosphoseryl-containing proteins have been identified in mineralising tissue or associated *in vivo* with calcium phosphate phases. In mineralising tissue, they have been proposed to act as nucleators and/or regulators of

Correspondence to: E. C. Reynolds

biomineralisation. In body fluids, e.g. saliva and milk, these multiphosphorylated proteins play an important role in stabilising the ACP phase, preventing phase transformations and precipitation [6-8]. From the sequences of proteins that are so far available, and known to stabilise ACP, a common feature is apparent as depicted in Table 1. The sequences contain multiple phosphoseryl residues in clusters and in most cases are together with other acidic residues that direct phosphorylation by an appropriate kinase [9]. Given that these proteins with similar biological functions have either aspartyl or glutamyl residues flanking the phosphoseryl clusters, the nature of the acidic residues may not be critical for their role in biomineralisation as long as they direct phosphorylation.

During matrix-mediated biomineralisation, crystal growth does not occur equally at all crystal surfaces, but is controlled by the specific adsorption of regulatory macromolecules resulting in a specific crystal morphology. For example, during the biomineralisation of the tooth hard tissues, dentine and enamel, growth occurs predominantly at the (001) plane, that is along the c-axis [10]. Specific molecular recognition between an inorganic mineral and a regulatory macromolecule is therefore essential for the regulation of crystal growth during matrix-mediated biomineralisation. However, the conformational requirements for molecular recognition during protein-mineral interaction is unknown. Overall, the mechanisms of these interactions in crystalline calcium phosphate-containing tissues, such as bone, dentine and enamel are poorly understood.

Biochemical approaches to date have not been able to unravel the complex processes in biomineralisation or reveal the conformations of the proteins at the mineral surface. Consequently, in conjunction with biochemical studies, we have adopted a new approach to investigate the molecular interactions during biomineralisation. This approach involves the computer-based simulation of the interaction between the protein and mineral surface.

For this investigation, we chose as an example, the motif -Ser(P)-Ser(P)-Ser(P)-Glu-Glu- found in proteins and peptides that bind HA with high affinity and stabilise nanoclusters of ACP [11]. We have demonstrated using synthetic peptides that all three phosphorylated residues of the -Ser(P)-Ser(P)-Ser(P)-Glu-Glu- motif are required for maximal binding affinity to hydroxyapatite [11]. Furthermore, dephosphorylation or substitution of the O-phosphoseryl residues by aspartyl or glutamyl residues resulted in significantly reduced binding [11]. Consequently, we have developed a molecular modelling procedure to study the relative binding energies of this motif Ser(P)-Ser(P)-Ser(P)-Glu-Glu to different crystalline surfaces of HA.

Materials and methods

All simulations were performed on an Indigo^2 Impact Silicon Graphics Workstation. The software consisted of a suite of programs including Insight II[12], Discover [13] and Catalysis [14].

The 'Consistent Valence Force Field' (CVFF) [15] was used in all the calculations. Non-bond interactions were calculated using sums over all the atoms in the system. Crystal atom coordinates were fixed. Atomic charges were determined using the default charges in the CVFF residue libraries. The CVFF does not specify charges on phosphate groups. The charge on the phosphate ions in the hydroxyapatite was determined within the ESFF [13] and the calculated charges transferred to the hydroxyapatite. Similarly, charges on the phosphate side-chains of the phosphoseryl residues was determined with ESFF[13] and transferred to the CVFF. We used a full ionic charge of +2 on the calcium ions. Recent calculations [16] suggest that a charge of about +1.45 is appropriate to calcium ions in an octahedral environment. The reduced positive charge being compensated by changes in the charge of the surrounding anions. Because each calcium ion is surrounded by many anions, seven or nine in the body of the crystal and probably a similar number at the surface of the crystal, the change in the charge of the anions is relatively small. The effect of these changes would be to reduce the electrostatic interactions by about 25%. However, it should be noted that a more accurate treatment of the electrostatic interactions demands a more accurate treatment of solvation effects.

Pure hydroxyapatite (HA) crystallises in the monoclinic space group P2$_1$/b with the lattice parameters shown in Table 2. Rows of phosphate ions are positioned along the a-axis with calcium and hydroxide ions located between the phosphate groups [17, 18]. Two types of calcium ion sites are present: nine-coordinate Ca(1) sites and seven-coordinate Ca(2) sites. The Ca(2) sites form isosceles triangles in the *c*-plane that are centred on the six-fold screw axes that run parallel to the *c*-axis. Hydroxide ions alternate above and below the planes of the triangles defined by the Ca(2) sites with their long axes located on the screw axes. The Ca(1) sites form larger hexagons surrounding the Ca(2) sites. The asymmetry of the monoclinic HA unit cell is exaggerated by the values reported for the lattice parameters in Table 1; for example, the ratio 2a/b is 1.000021. Consequently, the spatial arrangement of atoms along the *a*- and *b*-axes is very similar.

Biological hydroxyapatites are disordered compared to the monoclinic form [18] and disorder of the hydroxide ions results in lattice strain that is relieved by ion substitution and vacancies. As a consequence, the lattice parameters in the *a*- and *b*-directions are slightly smaller than in the monoclinic crystal, while the *c* dimension does not change: the ratio of $a_{monoclinic}/a_{hexagonal}$ is 1.000403. The unit cell of biological hydroxyapatites is usually considered to have an effective hexagonal symmetry. However, the different symmetry designation does not indicate a significant difference in unit cell morphology.

Atomic coordinates were obtained from the Inorganic Crystal Structure Database, Germany (ICSD). Hydroxyapatites that are disordered and or randomly substituted are not suitable as input to the simulation program. Consequently, the unit cell coordinates of a synthetic monoclinic hydroxyapatite were used in this study (Table 2). While

Table 1 *Partial amino acid sequences of proteins involved in calcium phosphate stabilisation and/or regulation of biomineralisation*

Amino acid sequence	Protein
-Ser(*P*)-Ser(*P*)-Ser(*P*)-Ser(*P*)-Ser(*P*)-Ser(*P*)-	Phosvitin [35]
-Asp-Ser(*P*)-Ser(*P*)-Asp-Ser(*P*)- Ser(*P*)-	Phosphophoryn [33]
-Ser(*P*)-Ser(*P*)-Ser(*P*)-Glu-Glu-	Dentine matrix phosphoprotein [36]
-Ser(*P*)-Met-Ser(*P*)-Ser(*P*)-Ser(*P*)-Glu-Glu-	Riboflavin binding protein [37]
-Ser(*P*)-Ser(*P*)-Gly-Ser(*P*)-Ser(*P*)-Glu-Glu-	Osteopontin [38]
-Ser(*P*)-Ile-Ser(*P*)-Ser(*P*)-Ser(*P*)-Glu-Glu-	α_{S1}-Casein [39]
-Ser(*P*)-Leu-Ser(*P*)-Ser(*P*)-Ser(*P*)-Glu-Glu-	β-Casein [39]
-Asp-Ser(*P*)-Ser(*P*)-Glu-Ser(*P*)-	Matrix-Gla protein [40]
Asp-Ser(*P*)-Ser(*P*)-Glu-Glu-	Statherin [6]

Table 2 *Lattice parameters of monoclinic hydroxyapatite and a typical 'hexagonal' biological hydroxyapatite [13, 14]*

Symmetry	a [nm]	b [nm]	c [nm]	α [°]	β [°]	γ [°]
Hexagonal	9.4176	9.4176	6.8814	90	90	120
Monoclinic	9.4214	18.8424	6.8814	90	90	120

the crystal has monoclinic symmetry, the (100) and (010) surfaces are effectively equivalent as discussed above, as would be the case in a hexagonal point group. Consequently, we chose different cleavage planes to construct the (010) and (100) surfaces. The positions of atoms in the HA crystal were calculated from the asymmetric unit by applying the appropriate crystal symmetry operations and translations. The distributions of calcium ions on the three model surfaces (100), (010), and (001) are notably different.

The principle interaction responsible for the binding of the Ser(*P*)-Ser(*P*)-Ser(*P*)-Glu-Glu motif to the hydroxyapatite surface would be expected to be electrostatic in origin. Consequently, we took care to minimise extraneous factors that might influence the calculated energy of binding. The thickness of the crystal was chosen by docking $H_2PO_4^-$ or HCO_3^- against the surface, then increasing the thickness and repeating the docking procedure. The surface was considered sufficiently thick when the energy of two successive iterations differed by less than 0.4 kJ mol^{-1}. In practice, a crystal plane thickness of about 1.4 nm or thicker was found to be adequate. A dielectric constant $\varepsilon=1$ was used in these preliminary calculations; the use of a distance dependent dielectric in later calculations will further diminish the effect of finite crystal thickness. Model crystal planes were constructed to be at least three times larger than the longest dimension of the extended peptide. This is expected to be sufficiently large to minimise electrostatic edge effects that may influence the preferred calculated conformation of the peptide. Finally, the crystals were made electrically neutral by either deleting ions from the edges of the back surface or converting calcium ions on the back surface edges to sodium ions. Given the insensitivity of the calculated binding energy to the crystal thickness, these changes were expected to have minimal impact on the calculated binding energy.

The Ser(*P*)-Ser(*P*)-Ser(*P*)-Glu-Glu peptide was modelled using Insight 95. The terminal amide and carboxylate groups were capped by protons to mimic the electrostatically neutral peptide residues that flank the motif in biologically sig-

nificant proteins. In preliminary calculations we docked both α-helical and β-strand structures on the surfaces. However, unconstrained energy minimisation of the resulting structures always resulted in peptide structures that deviated significantly from the starting regular secondary structures. Thorough sampling of the ϕ/ψ space was necessary to ensure adequate coverage of the conformational space of the peptide motif, since the peptide ϕ and ψ angles relaxed significantly during docking.

A library of 100 random conformations of the Ser(*P*)-Ser(*P*)-Ser(*P*)-Glu-Glu peptide was generated in a 100 ps dynamics run at 1000 K, with structures sampled every 1 ps. The peptide was constrained to be in an all-*trans* conformation during the dynamics run. The charges on the molecule were ignored during the dynamics run to prevent extended conformations predominating in the library. A cluster analysis of backbone atom coordinates confirmed that there was little similarity between conformations within the library.

Structures were chosen at random from the library and docked against the hydroxyapatite surfaces using the Solids_Docking module of the Catalysis software. A thousand structures were sampled at random orientations on each crystal face. In preliminary calculations each structure was rigidly docked against a crystal face with the centre of mass coordinates and the orientation of the peptide being adjusted by a steepest-descents least-squares algorithm to give a best energy conformation. These structures were then refined in the final series of calculations.

In the final calculations a distance dependent dielectric constant of 2.0*r was used; this is expected to roughly model the effects of solvent and counter-ions in solution. Preliminary calculations using a constant dielectric constant of 1.0 gave substantially larger binding energies for individual surfaces, but had no affect on the principal conclusions of the study. The energy was minimised using a steepest-descents algorithm and allowing the peptide internal coordinates to change with a harmonic force between bonded atoms until

the maximum derivative was less than 21 kJ mol^{-1} nm^{-1}. The cross-terms between atom coordinates were then added and a further minimisation using the quadratically convergent Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm was performed until the maximum derivative was less than 0.21 kJ·mol^{-1}·nm^{-1}. Finally a Morse type force between atoms was introduced and the BFGS algorithm again applied until the maximum derivative was less than 0.0004 kJ·mol^{-1}·nm^{-1}.

In a separate series of calculations a constrained β-strand structure was docked against the (100) surface. A 63 kJ·mol^{-1} radian^{-2} constraining force was applied to restrict the φ/ψ angles to values appropriate to a β-strand (120 °, -140 °). A total of 85 orientations was sampled and energy minimised to a maximum derivative of 0.004 kJ mol^{-1}·nm^{-1}.

Results

For each of the HA (001), (010) and (100) surfaces a total of 1000 docked conformations of the motif Ser(*P*)-Ser(*P*)-Ser(*P*)-Glu-Glu were sampled. The binding energy was defined as the total energy of the system minus the internal energies of the crystal and the peptide; a negative value indicated that a stable bound system had formed. The interactions that contributed to the binding energy were the non-bond (non-covalent) interactions: that is the van der Waals' and electrostatic interactions.

Direct comparison of binding energies is confounded by the ability of the peptide to adopt a different conformation at each surface thus changing the internal energy of the peptide. Experimentally the enthalpy of binding is determined relative to some arbitrary reference state, for example, the average energy of an ensemble of peptide molecules in solution at 298 K. We chose the most strongly bound β-strand conformation on the (100) surface as the reference state for the peptide ligand, and the difference in internal energy between the bound peptide conformer and the reference state was added to the binding energy to give a relative binding energy. The relative binding energy includes contributions from the non-bond terms plus bond terms reflecting the differences in the peptide conformation at the different surfaces. The relative binding energies of the peptide docked onto different HA surfaces were then directly compared. Figure 1
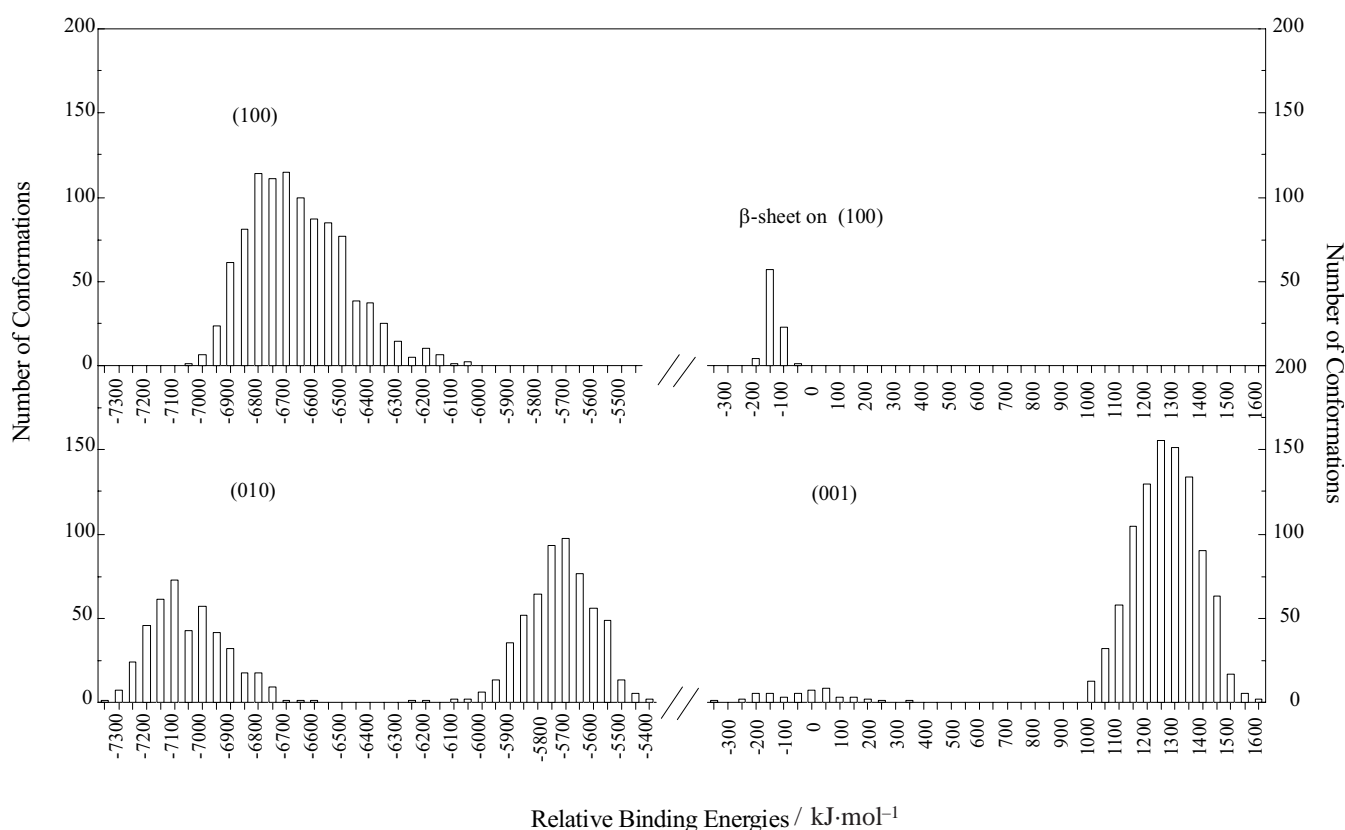
Figure 1 *Number of conformations of the motif Ser(P)-Ser(P)-Ser(P)-Glu-Glu bound to the three HA faces (100), (010) and (001) as a function of relative binding energy. The height of each bar indicates the number of conformations having a relative binding energy within ± 25 kJ·mol^{-1} of the bar's ordinate value. Note the break in the ordinate*

Table 3 *The relative binding energy and contributions of the various interactions (kJ·mol^{-1}) of the most strongly bound conformers at the three surfaces of HA and the β-strand conformer on the (100) surface*

Energy	Surface			
	β-strand on HA (100)	HA (001)	HA (010)	HA (100)
Total	**-222.266**	**-359.957**	**-7370.463**	**-7052.655**
Internal	0.000	72.844	12.527	-4.038
Bond	0.000	82.182	-15.431	-1.134
Angle	0.000	-22.614	-5.849	-4.628
Torsion	0.000	13.941	31.242	3.096
Out of Plane	0.000	-0.665	2.565	-1.372
Non-bond	-222.266	-432.801	-7382.990	-7048.617
VdW	-142.967	284.240	-99.905	50.179
Repulsive	103.847	794.583	546.895	581.802
Dispersive	-246.814	-510.343	-646.800	-531.623
Electrostatic	-79.299	-717.041	-7283.085	-7098.796

shows for each HA surface, a plot of the frequency of conformers as a function of relative binding energies. The smooth variation in the number of states as a function of relative binding energy suggests that the distribution of peptide conformations at each surface has been well sampled. The sharp decline in the number of strongly bound conformations at each surface further suggests that it is unlikely that significantly stronger bound conformations would exist on these surfaces. The interactions between the motif and the (100) and (010) surfaces were significantly stronger than that with the (001) surface. The major difference between the (010) and the (100) surface was the choice of cleavage plane. The relative binding energy of the most strongly bound, constrained β-strand conformer on the HA (100) surface was comparable to the relative binding energies of the most strongly bound conformers on the (001) surface.

The contributions to the total relative binding energy for the most strongly bound peptide conformation at each surface (001), (010) and (100) and for the constrained β-strand conformer docked on the (100) surface are summarised in Table 3.

Figure 2 shows the conformation of the most strongly bound, constrained β-strand on the HA (100) surface. The least strongly and most strongly bound conformations at each surface were also examined in detail. Figures. 3, 4, and 5 show the orientation of the most strongly bound conformer on the (001), (010), and (100) surfaces respectively. Ramachandran plots for the most strongly bound conformers on the HA (001), (010) and (100) surfaces are presented in Figure 6. To further characterise the interaction, the neighbouring atoms at the interface of the motif Ser(*P*)-Ser(*P*)-Ser(*P*)-Glu-Glu and all three HA crystal surfaces were examined. Generally the more strongly bound conformers had more charged oxygen atoms facing the HA crystal surface when compared with less strongly bound conformers. Similarly the more stable conformers were in proximity to a greater number of calcium ions in the HA surfaces.

Discussion

Analysis of relative binding energies

The relative binding energy consists of internal (bond) and non-bond contributions. The internal contributions reflect differences in the peptide conformation and represent only a small fraction of the relative binding energy in these structures. The non-bond energies (van der Waals' and electrostatic contributions) dominate the overall binding and are due largely to interactions between the ligand and crystal surface.

The binding of the constrained β-strand conformation to the HA (100) surface can be primarily attributed to the van der Waals' interaction between the peptide and the crystal (Table 3). The electrostatic interaction was smaller than the repulsive van der Waals' interaction. Maintaining a β-strand conformation prevented effective contact between the negatively charged amino acyl residues and the positively charged calcium ions of the crystal.

The most strongly bound conformer docked onto the HA (001) surface had a larger electrostatic contribution to the relative binding energy. However, inspection of Figure 3c showing the docking of this conformer onto the (001) surface, suggests that the peptide cannot approach the calcium ions closely without also coming close to the negatively charged ions in the surface. The large values of the internal bond energy and van der Waals' repulsive interaction relative to the β-strand conformation, suggest that the peptide had to adopt an energetically unfavourable conformation to gain a favourable electrostatic interaction with the HA (001) surface.

The relative binding energies of the most strongly bound conformers docked on the HA (010) and HA (100) surfaces were dominated by electrostatic contributions, a factor of more than ten greater than the other contributing terms. The internal energy contributions were relatively small, suggesting that

the peptide did not have to distort greatly to dock onto these surfaces. The large torsional contribution to the internal energy on the HA (010) surface is consistent with the Ramachandran plots that suggest the peptide was distorted away from the low-energy conformations usually obtained in these plots.

The size of the contributions to the net van der Waals' interaction was significantly larger for the low-energy conformers on the HA (001), (010), and (100) surfaces relative to the constrained β-strand conformation docked on HA (100). This reflects the stronger electrostatic interaction between the crystal and the ligand forcing the atoms together and thus

increasing the magnitude of these terms. The net van der Waals' interactions increase in the most strongly bound conformers on the HA (001), (010), and (100) surfaces, but are more than compensated for by the increased electrostatic contributions.

The most strongly bound conformations of the peptide docked onto the HA surfaces were those with the most charged atoms, including the phosphate groups of the phosphoseryl residues and the carboxylate groups of the glutamyl residues, facing the HA crystal surface. Weakly bound conformations were characterised by charged groups not making contact with the HA surface. The electrostatic energy gained by plac-
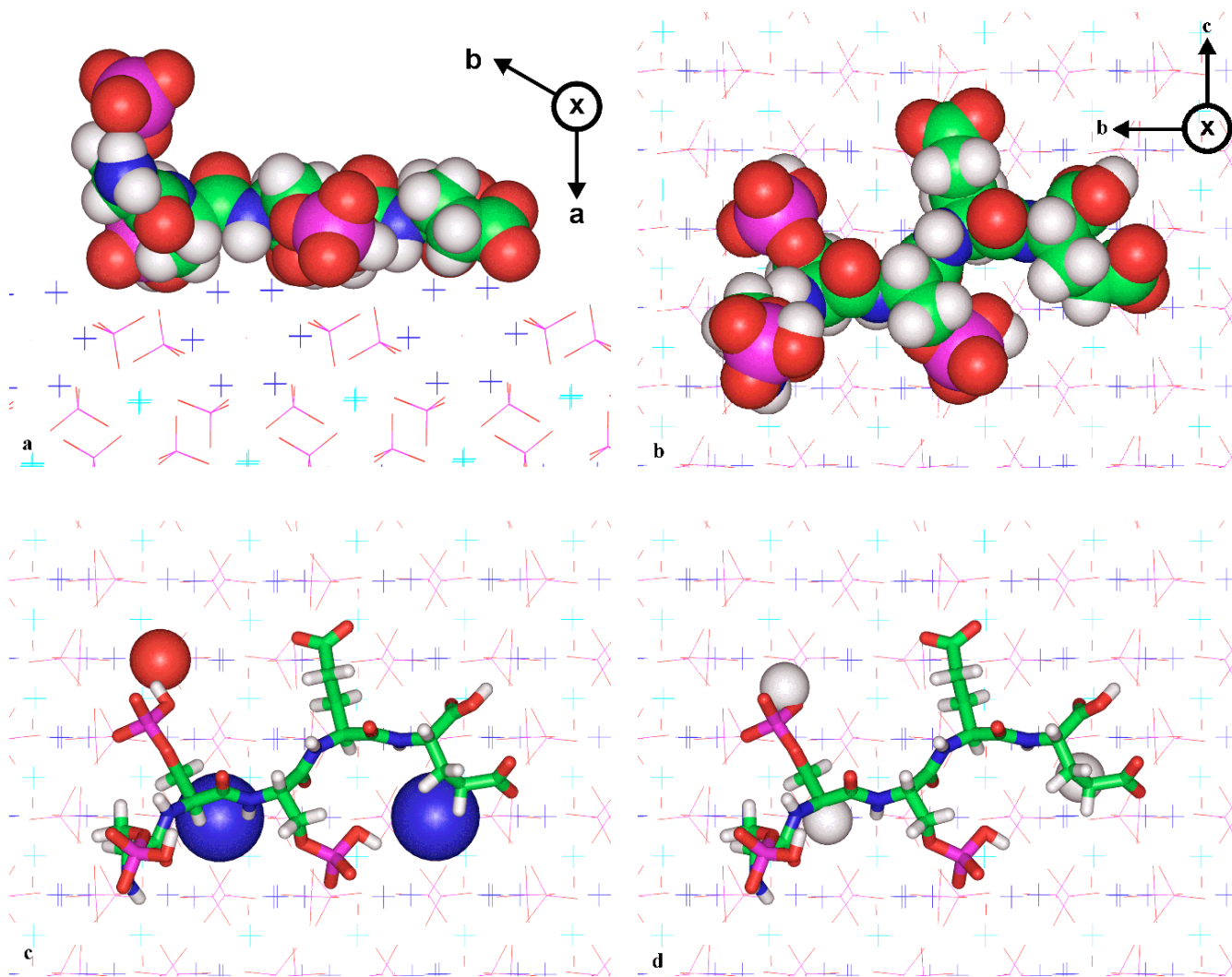
Figure 2 *The HA (100) crystal surface with the most strongly bound, constrained β-strand conformation of the motif Ser(P)-Ser(P)-Ser(P)-Glu-Glu. The atoms are colour coded as follows: calcium(1) atoms are light blue crosses; calcium(2) atoms are dark blue crosses; oxygen atoms are red; phosphorus atoms are magenta; carbon atoms are green; nitrogen atoms are blue, and hydrogen atoms are grey. The symbol ⊗ indicates a crystallographic axis projecting into the paper.*

Four views are presented: (a) a side view along the c-axis with the peptide rendered in CPK and the crystal atoms in 'line' form, (b) as in (a) but viewed from above looking down on the HA (100) face, (c) as in (b) with the peptide displayed in stick form and the atoms in the HA surface within 0.25 nm of the peptide rendered in CPK, (d) as in (b) with the peptide displayed in stick form and the atoms of the peptide within 0.25 nm of the HA surface rendered in CPK

ing the charged side-chains in contact with the surface more than compensated for any interaction energy required to distort the molecule.

The reliability of any molecular modelling exercise is determined by the quality of the force-field. Because the relative binding energies were dominated by the well-understood electrostatic interaction, the conclusions of our study are robust and will not be strongly dependent on the force-field chosen to represent the molecular system. The use of a distance dependent dielectric ($\varepsilon = 2r$) to mimic the effect of solvent and allow the Coulomb term summation to be truncated at 1.4 nm, will have the effect of weakening the electrostatic interaction in these model studies. Using a constant dielectric will have the largest effect on those conformations that have the largest electrostatic contributions to the relative binding energy: namely HA (010) and HA (100). Other terms in the force-field will play a role in the determination of the

structure of the peptide, but their effects are expected to be subtle and not to be major factors in determining the overall structure.

Comparison of motif conformations

Figure 2a shows the orientation of the most strongly bound, constrained β-strand conformer bound to the (100) surface. The cleavage plane of the (100) surface used was such that the Ca(2) atoms coloured dark blue were the most accessible. Despite the (100) and (010) planes being the best surfaces for this motif, Figure 2a shows that for this conformation not all the charged groups were in proximity to the surface. The surface atoms within 0.25 nm of the motif are highlighted in Figure 2c. Conversely, the motif atoms within 0.25 nm of the HA surface are highlighted in Figure 2d. Despite
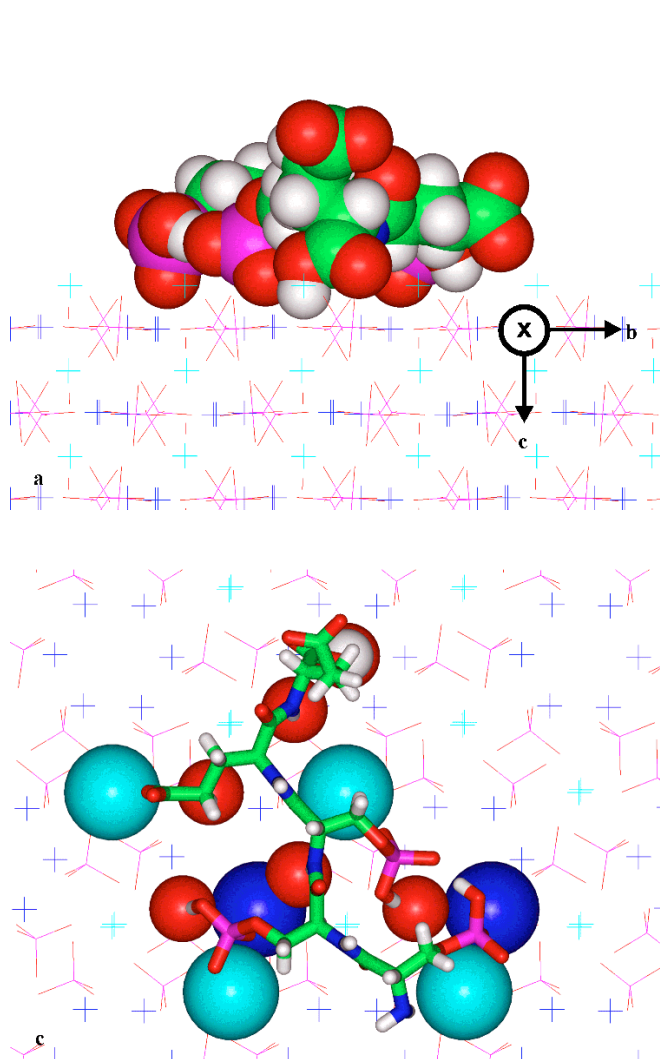
Figure 3 *The HA (001) crystal surface with the most strongly bound docked conformer. Atoms are colour coded as in Figure 2. Four views are presented with atoms rendered as in*
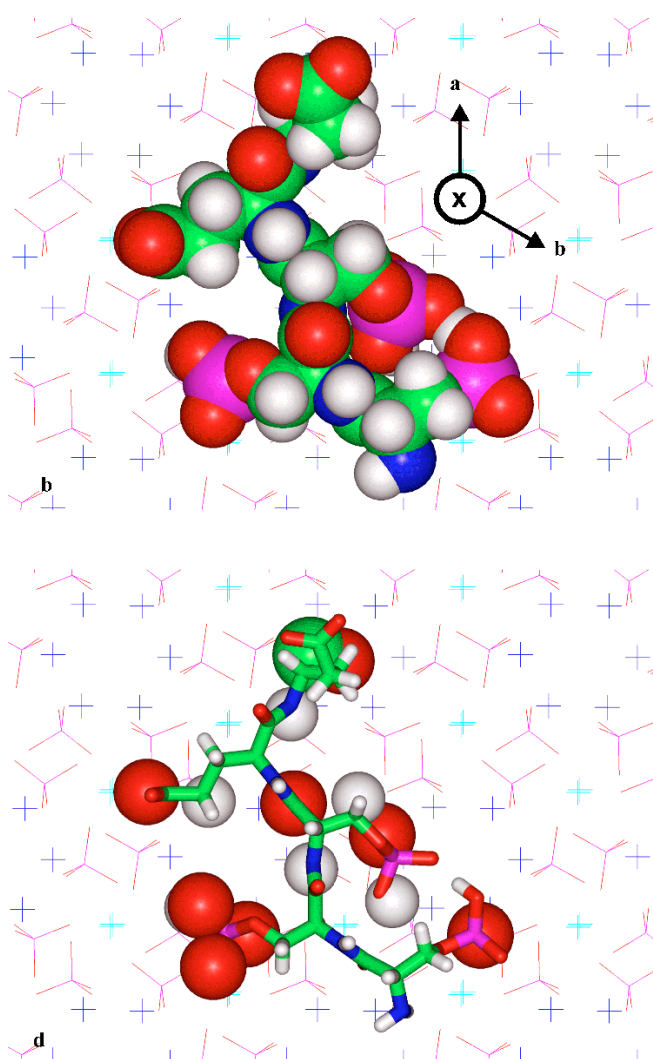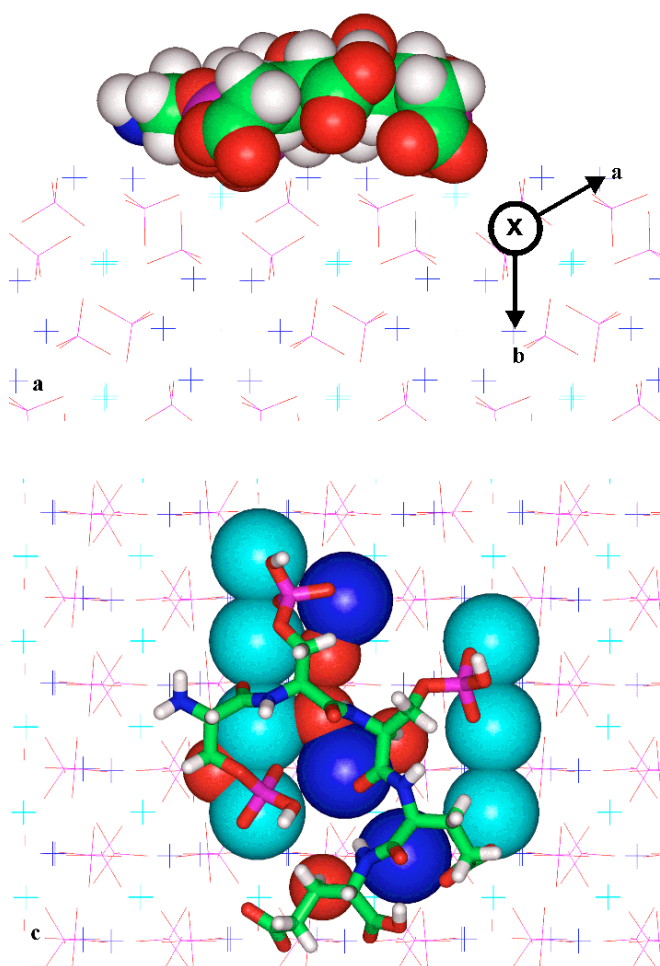
*Figure 2 (**a**) a side view along the a-axis, and (**b**), (**c**) and (**d**) viewed from above looking down on the HA (001) face*

the near coincidence of the HA *c*-axis dimension and the spacing of the residues in a β-strand, the preferred conformation of the β-strand docked on the HA (100) surface was found to be oriented almost perpendicular to the *c*-axis.

At the HA (001) surface, the Ca(1) atoms (light blue) were more accessible than the Ca(2) atoms (dark blue) as shown in Figure 3. This HA (001) surface was the least preferred by the motif. The most strongly bound conformation on the HA (001) face does not appear to have any particular orientation with regard to the crystal features. The highly symmetric locations of atoms on this face do not favour particular orientations of the linear motif. Figure 3c shows the atoms of the HA (001) crystal surface that are within 0.25 nm of the most strongly bound conformation of the motif. Figure 3d shows the atoms of the same conformer that were within 0.25 nm of the HA (001) crystal surface. The side-chain atoms of the

peptide came close to oxygen atoms of the crystal, resulting in an unfavourable electrostatic interaction.

The HA (010) surface had columns of Ca(1) type calciums between pairs of Ca(2) rows parallel to the *c*-axis. The Ca(2) type calciums were raised above the surface defined by the Ca(1) calciums. The peptide adopted an extended conformation (Figure 4a and 4b) on docking to the HA (010) surface with the acidic side-chains in contact predominantly with the Ca(1) type calciums (Figure 4c and 4d). The ϕ/ψ angles vary between values typical of a β-strand and a region having ϕ/ψ angles of 80° and -150° (Figure 6). This is not a region classically allowed in the Ramachandran plot. The ligand energy associated with this conformation is higher than that observed for the HA (100) surface, but this is more than compensated for by the increase in binding energy with the HA (010) surface. An explanation for this is that while the average dis-

Figure 4 *The HA (010) crystal surface with the most strongly bound docked conformer. Atoms are colour coded as in Figure 2. Four views are presented with atoms rendered as in*
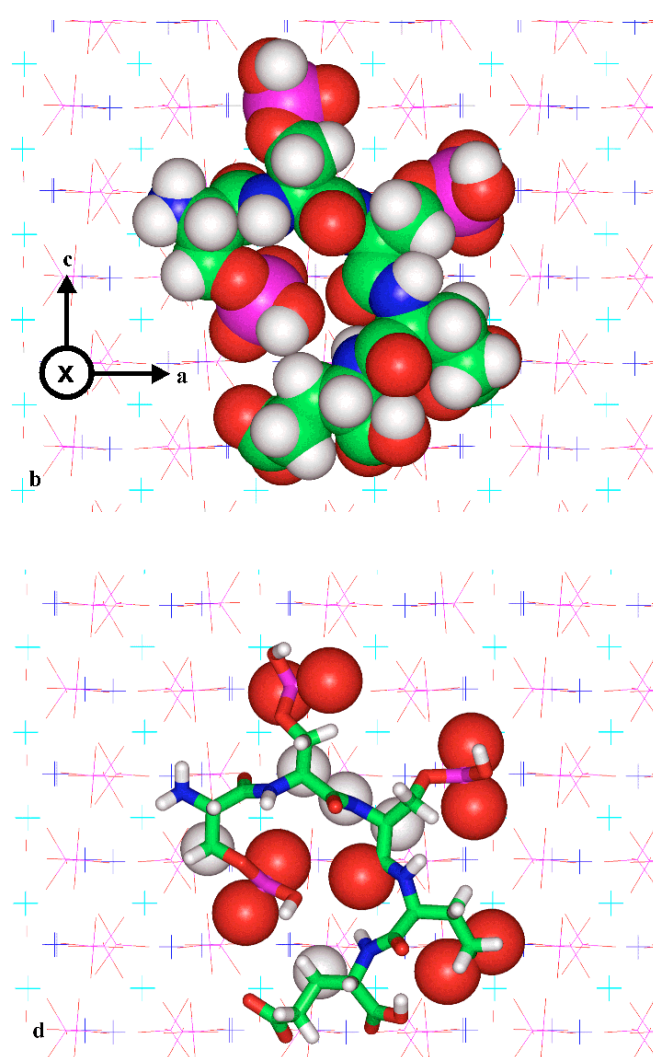
Figure 2 (a) a side view along the c-*axis, and (b), (c) and (d) viewed from above looking down on the HA (010) face*

tance between the columns of Ca(1) type calciums on this surface corresponds to the distance between the ridges of Ca(2), the calcium atoms in the type 1 sites do not zig-zag back and forth as do the type 2 calcium ion sites. Consequently the peptide backbone is forced to distort to make optimal contact with the type 1 calcium sites.

The HA (100) surface was similar to the HA (010) surface without the columns of Ca(1) sites. Figure 5 shows the most strongly bound conformation of the Ser(*P*)-Ser(*P*)-Ser(*P*)-Glu-Glu motif on the HA (100) crystal surface. Figures 5a and 5b reveal that the peptide adopted an extended conformation on the HA (100) surface, with the oxygen atoms of the acidic side-chains in close contact with the surface (Figure 5d). The peptide backbone lay approximately parallel to the c-axis of the crystal between two of the ridges of Ca(2) atoms. The acidic side-chains of the peptide extend on alternate sides of the peptide backbone and make close contact (<

0.25 nm) with these Ca(2) calcium atoms (Figure 5c). Strong electrostatic interactions between the oxygen atoms and the calcium atoms account for much of the stability of this peptide conformation. The ϕ/ψ angles for residues 2, 3 and 4 alternate between values typical of a β-strand, and a region having ϕ/ψ angles of typical α-helical structure (Figure 6). Alternation of the ϕ/ψ angles along the peptide backbone allows all the acidic amino acid residues to make contact with the crystal surface. This allows optimal electrostatic interactions to take place.

Significance of the β-strand conformation

A number of earlier biochemical studies of peptide interactions with hydroxyapatite and aragonite have noted that the atomic spacing of the calcium atoms on these crystal sur-
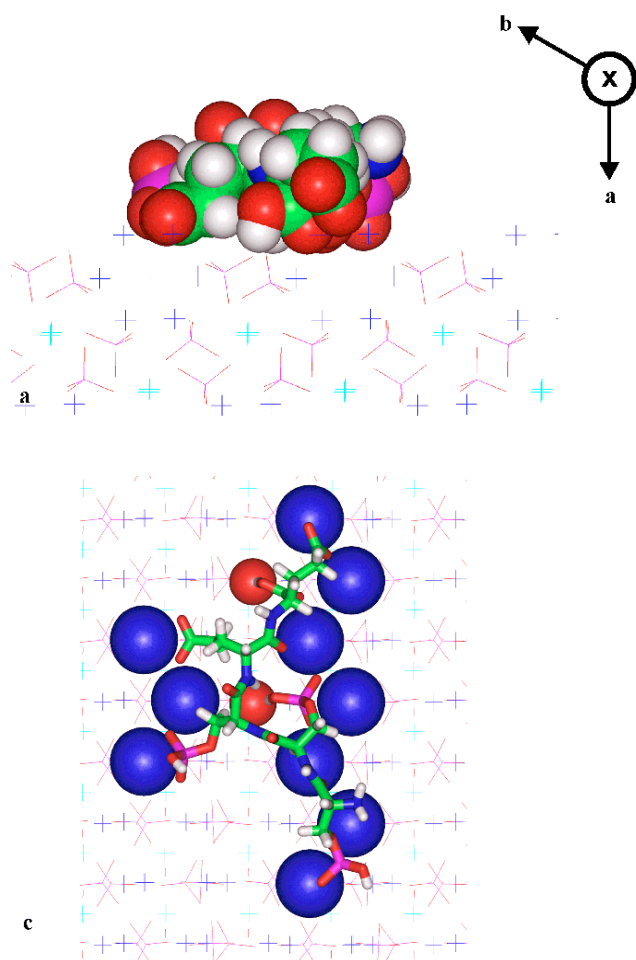
Figure 5 *The HA (100) crystal surface with the most strongly bound docked conformer. Atoms are colour coded as in Figure 2. Four views are presented with atoms rendered as in*
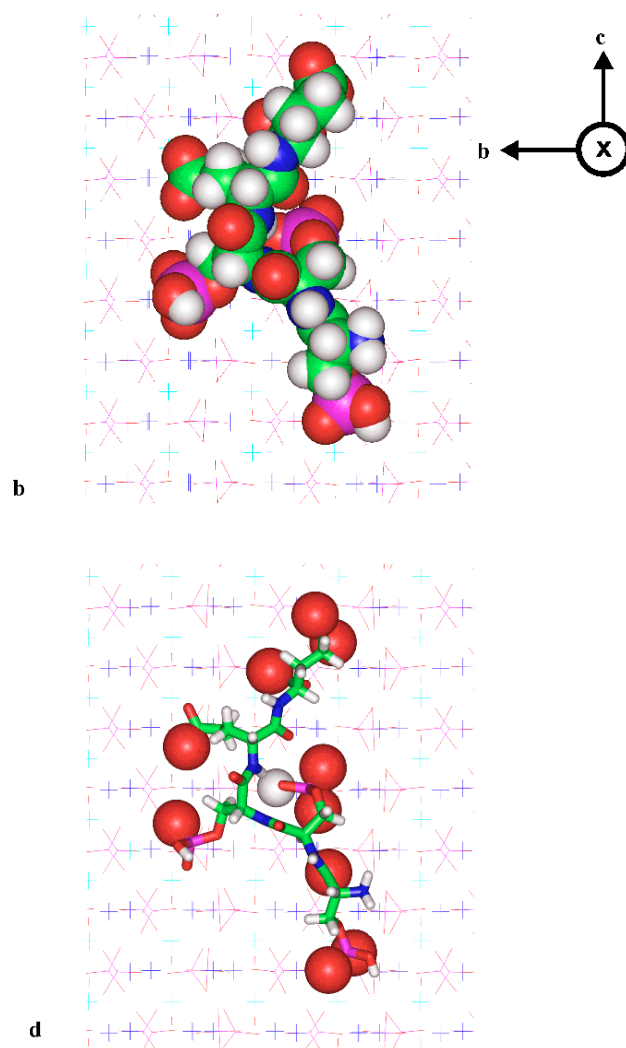
*Figure 2 (**a**) a side view along the* c-*axis, and (**b**), (**c**) and (**d**) viewed from above looking down on the HA (100) face*

faces is close to that observed between adjacent residues of a β-strand or β-sheet structure [1, 19-22]. It was suggested that the sequence poly-(-Asp-X-) (where X is a neutral residue) commonly found in EDTA-soluble proteins of both aragonite and calcite layers of mollusc shells [20, 22] adopts the β-strand configuration at the mineral surface. In the case of a poly-(Asp-X) sequence, a β-strand conformation allows alternate residues to make contact with the surface, thus maximising the electrostatic interaction between the aspartyl residues and the calcite or aragonite surface. The hydrophobic residues minimise their exposure to solvent by interacting with adjacent layers of the protein/mineral complex that forms the mollusc shell. Although this is plausible for alternating acidic residues, there is little published experimental evidence that a β-strand is the preferred conformation for proteins or peptides with continuous stretches of acidic residues that interact with mineral. The protein dentine phosphophoryn, which contains many clusters of *O*-phosphoseryl residues, is reported to have a large cross-sectional area per molecule when adsorbed onto HA [24]. Furthermore, the number of proteolytic sites in phosphophoryn increases when phosphophoryn is adsorbed onto HA surfaces [25]. These observations taken together may be consistent with an extended protein conformation [25] however, the interpretation of solid-state NMR chemical shifts [26] as supporting a β-strand conformation is open to the criticism that there are no reference chemical shifts for peptides in a β-strand conformation adsorbed on mineral surfaces and subject to the intense electric fields expected near such a surface. We are currently investigating the solution structures of three multiphosphorylated peptides α_{S1}-casein(59-79), β-casein (1-25) and α_{S2}-casein(1-21) that contain the -Ser(*P*)-Ser(*P*)-Ser(*P*)-Glu-Glu- motif. ^{1}H NMR analysis of these multiphosphorylated peptides in the presence of calcium ions yielded secondary chemical shifts of both Hα and amide protons suggestive of extensive β-structure. However the observed NOEs, that are more definitive of structure than the secondary chemical shifts, suggest a structure consisting of loops and turns [27-29].

Our modelling studies have demonstrated that in the case of the motif in the β-strand conformation the relative binding energies are very small. A β-strand conformation forces al-
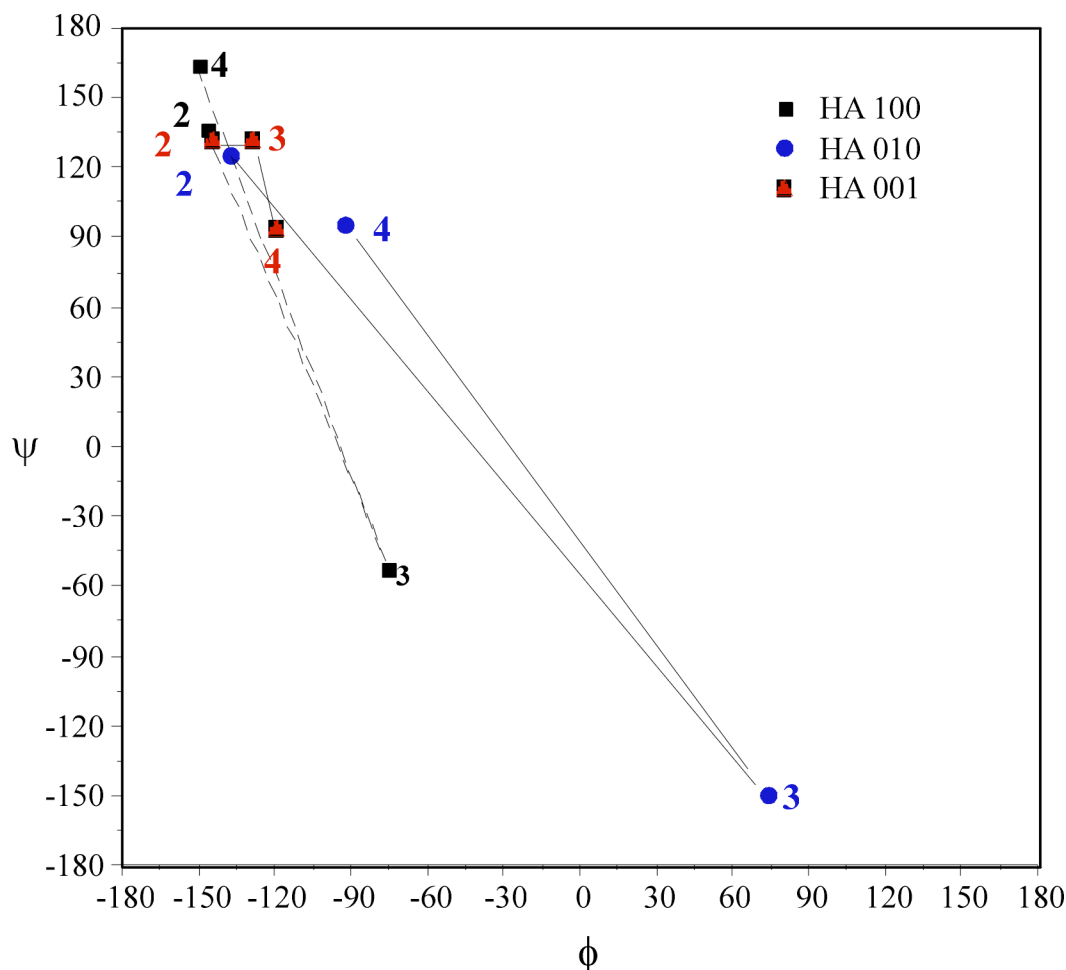
Figure 6 *Ramachadran plots for the most strongly bound conformers of the motif Ser(*P*)-Ser(*P*)²-Ser(*P*)³-Glu⁴-Glu on the HA (001)* ▲*, HA (010)* ●*, and HA (100)* ■ *faces*

ternate residues to make poor contact with the HA surface and reduces the net electrostatic interaction. The formation of a β-sheet with the concomitant formation of hydrogen bonds will not compensate for the reduced electrostatic interaction between alternate residues and the mineral surface. Hydrogen bonds are a fairly weak interaction with typical bond energies in the range of 4 to 40 kJ mol^{-1}; by comparison the electrostatic interaction between the five-residue motif and the HA (100) and (010) surfaces were calculated to be about 7100 kJ mol^{-1}.

Correlation with biochemical studies

The adsorption of the motif Ser(*P*)-Ser(*P*)-Ser(*P*)-Glu-Glu onto HA surfaces has been previously investigated using a series of synthetic peptides, homologues and analogues by experimentally measuring peptide binding and using the Langmuir adsorption model to determine maximal binding and affinity constants [11]. Replacement of any of the phosphoseryl residues by either an aspartyl or glutamyl residue resulted in lower affinity for the HA surface. Our modelling studies have demonstrated that the electrostatic contribution is the dominant factor in the relative binding energies. The partial negative charge of the oxygen atoms in the phosphate group is greater than that in a carboxylate group, therefore we would predict a stronger electrostatic interaction with the hydroxyapatite surface by the phosphoseryl containing sequence. Thus the conclusion of the previous experimental adsorption studies, that all three phosphate groups were necessary for maximal binding affinity to hydroxyapatite [11], is consistent with the results of this study that show that all the phosphoseryl groups of the most strongly bound conformers were in contact with the HA surface.

Our modelling results are also consistent with previous reports of the acidic proteins involved in biomineralisation having preferences for particular crystallographic surfaces. For example, the multiphosphorylated protein, phosphophoryn from dentine, and other acidic proteins from bone osteocalcin, osteonectin and bone small proteoglycan II labelled with fluorescein isothiocyanate were found to preferentially bind to the (100) face of HA [30]. Examination of the interaction of rat phosphophoryn and egg yolk phosvitin with octacalcium phosphate (OCP) revealed preferential binding with the OCP (010) surface [31]. OCP is a crystalline phase of calcium phosphate with crystals in the triclinic form P_{-1} (a = 1.9692 nm, b = 0.9523 nm, c = 0.6835 nm; α = 90.15°, β = 92.54°, γ = 108.65°). These crystals have a layered structure in which layers almost identical to hydroxyapatite (~1.1 nm thick) alternate with hydrated layers of water, calcium and phosphate ions (~0.8 nm thick) along the (100) axis. Consequently, the (010) plane of OCP has stripes of "(100) HA" type structure alternating with hydrated layers. It is therefore consistent with our modelling results that the multiphosphorylated proteins, rat phosphophoryn and egg yolk phosvitin preferred binding to the OCP (010) surface as previously reported [31].

Phosphophoryn is produced by dentine forming cells, odontoblasts, and is believed to both nucleate and regulate the growth of hydroxyapatite crystals during dentinogenesis. George et al. [32] have recently identified a clone from a rat odontoblast cDNA expression library that expressed a protein, designated Dmp2, which was recognised by antibodies generated to a dephosphorylated phosphophoryn. Sequence information was obtained for the carboxy-terminal region of the Dmp2 protein which revealed repeat motifs of -(Asp-Ser)$_m$- where m = 2-3 and -(Asp-Ser-Ser)$_n$- where n was up to 14. As the seryl residues in phosphophoryn are over 90% phosphorylated then the seryl residues of the repeat motifs are likely to be phosphorylated. We have recently determined the N-terminal 32 residues of bovine dentine phosphophoryn using chemical conversion of the phosphoseryl residues to *S*-propyl cysteinyl residues which are stable to Edman chemistry [33]. This sequence analysis revealed that the N-terminus of bovine dentine phosphophoryn contained the repeat motif -[Asp-Ser(*P*)-Ser(*P*)]$_n$-. Therefore, from the current modelling data and the presence of this repeat phosphorylated motif in phosphophoryn we would predict preferential binding of the protein at the HA (100) and (010) surfaces such that crystal growth would be allowed to continue only at the HA (001) plane or along the *c*-axis. This is consistent with the pattern of hydroxyapatite crystal growth observed *in vivo* during dentine development [10].

The experimental determination of peptide structure at solid interfaces is an extremely challenging task. However, recently Long et al.[34] reported solid-state NMR results that provide strong support for the model proposed in this paper. The authors studied the dipole induced relaxation of doubly ^{13}C labelled peptides bound to hydroxyapatite crystallites. The peptide chosen in this study, Asp-Ser(*P*)-Ser(*P*)-Glu-Glu-Lys, is the N-terminal sequence of the salivary peptide statherin. The ^{13}C labelled carbonyls of the two phosphoseryl residues were found to have a mean separation of 0.32 ± 0.01 nm; however, the data obtained at longer relaxation times were consistent with a broader distribution of distances. The best fit to the experimental data suggested a distribution that was 40% α-helix and 60% β-sheet. It is difficult to reconcile such different structures of the peptide with a stable peptide conformation at the hydroxyapatite surface. While the possibility of kinetic trapping of the peptide in the different conformations can not be ruled out, the results do agree very well with the model developed in this paper. In a long poly-Ser(*P*) peptide bound to hydroxyapatite, there would be no energetic reason for a residue near the centre of the peptide to adopt an α-conformation in preference to a β-conformation and one would expect the experiment to give a 50/50 mixture of α-helix and β-sheet based on our model. However, in the peptide studied the presence of an asparaginyl residue adjacent to the first Ser(*P*) in the sequence might account for the slight conformational preference observed in the NMR results.

Other solid state NMR experiments have been reported on similar systems [26], but these studies have been limited to characterising the chemical shifts of ^{13}C labelled peptides bound to hydroxyapatite. While the chemical shifts have been

interpreted as indicating a β-sheet type conformation these results need to be interpreted very carefully. The chemical shift is a very sensitive indicator of chemical environment and the ^{13}C reference shifts are determined from amino acyl residues in globular proteins. The strong electrostatic fields adjacent to the hydroxyapatite surface would be expected to have significant effects on the chemical shifts of peptides bound to the surface. Hence, the only conclusion that can be drawn from such studies is that the chemical shifts are similar to those of a β-sheet in proteins, not that the peptide has adopted a β-sheet conformation.

Importance of simulation methods

This is the first simulation study of the binding of a multiphosphorylated peptide sequence to HA crystal surfaces. Simulation techniques have previously been exploited to examine inorganic surface chemistry. These techniques have been adapted in this study to investigate the surface biochemistry of hydroxyapatite during biologically controlled mineralisation. In particular, for a thorough analysis of the relative binding energies and docked conformations, the use of a library of random conformations is more appropriate for small peptides that are likely to adopt a range of conformations in solution. This method allows the study of the structural features that are important in the interaction between a mineral surface and a negatively charged peptide or protein. Development of such simulation techniques is essential for the understanding of the molecular events during biomineralisation.

Biological implications

The interaction between multiphosphorylated proteins and calcium phosphate phases is an important feature of many biological systems. Examples include the directed growth of hydroxyapatite (HA) crystals during dentine and enamel formation, and the stabilisation of supersaturated solutions of calcium phosphate in milk and saliva.

A popular hypothesis for the interaction of these proteins with HA is founded on the observation that the *c*-axis repeat distance in HA of 0.688 nm is very close to the repeat distance of residues in a β-strand conformation. However, this hypothesis does not explain the observed preferential binding of these proteins to specific crystallographic faces of the HA crystals. There is no published experimentally-derived evidence that unambiguously shows that the β-strand conformation is the preferred conformation of these multiphosphorylated proteins on HA surfaces.

The molecular modelling study described in this report provides the basis for rationalising a wide variety of previously isolated biochemical observations into a coherent model of the interaction of this important class of proteins with mineral phases. Contrary to the popular hypothesis that these proteins interact with HA surfaces as β-strands or sheets, we have shown that the most stable conformers adopt structures in which alternate residues exhibit backbone angles characteristic of a β-strand and values of an α-helix or a distorted α-helix.

Conclusion

In summary, we have investigated, the binding of a multiphosphorylated motif Ser(*P*)-Ser(*P*)-Ser(*P*)-Glu-Glu to HA using a simulation strategy to probe the binding mechanisms. This method has the potential to be used to examine the interactions of other peptides with crystalline surfaces to understand the features of molecular recognition and molecular complementarity during biomineralisation.

Supplementary material available Atomic coordinates for the low energy conformations on each of the surfaces HA(100) and HA(001) are available in InsightII (car/mdf) and in Protein Data Base (pdb) format.

References

1. Mann, S. In *Biomineralization: Clinical and Biomedical Perspective*; Mann, S.; Webb, J.; Williams, R., Eds.; VCH Publishers: New York, 1989; pp 35-60.
2. Glimcher, M.; Reit, B.; Kossiva, D. *Calc. Tiss. Int.* **1981**, *33*, 185.
3. Stetler-Stevenson, W. G.; Veis, A. *Biochemistry* **1983**, *22*, 4326.
4. Weiner, S. *Biochemistry* **1983**, *22*, 4139.
5. Lowenstam, H. A.; Weiner, S. In *On Biomineralization*; Oxford University Press: New York, 1989; pp 25-49.
6. Holt, C.; Van Kemenade, M. J. J. M. In *Calcified Tissue*; Hukins, D. W. L., Ed.; CRC Press: Florida, 1989; pp 175-213.
7. Marsh, M. E. *Biochemistry* **1989**, *28*, 339.
8. Nagata, T.; Bellows, C. G.; Kasugai, S.; Butler, W. T.; Sodek, J. *Biochem. J.* **1991**, *274*, 513.
9. Roach, P. J. *J. Biol. Chem.* **1991**, *266*, 14139.
10. Simmer, J. P.; Fincham, A. G. *Crit. Rev. Oral Biol. Med.* **1995**, *6*, 84.
11. (a) Reynolds, E. C. *J. Spec. Care Dent.* **1998**, *18*, 8; (b) Reynolds, E. C. unpublished results.
12. *Insight II User Guide*, October **1995**. San Diego: Biosym/MSI.
13. *Discover 2.9.7/95.0/3.00 User Guide*, October **1995**. San Diego: Biosym/MSI.
14. *Catalysis 4.00 User Guide*, September **1996**. San Diego: Biosym/MSI.
15. Dauber-Osguthorpe, P.; Roberts, V. A.; Osguthorpe, D. J.; Wolff, J.; Genest, M.; Hagler, A. T. *Prot. Struct. Funct. Gen.* **1988**, *4*, 31.

16. Cross, K.J. unpublished results.
17. Posner, A. S.; Perloff, A.; Diorio, A. F. *Acta Cryst*. **1958**, *11*, 308.
18. Kay, M. I.; Young, R. A. *Nature* **1964**, *204*, 1050.
19. Reynolds, E. C.; Riley, P. F.; Storey, E. *Calc. Tiss. Int.* **1982**, *34*, 52.
20. Addadi, L.; Weiner, S. *Proc. Natl. Acad. Sci.* **1985**, *82*, 4110.
21. Addadi, L.; Moradian, J.; Shay, E.; Maroudas, N.; Weiner, S. *Proc. Natl. Acad. Sci.* **1987**, *84*, 2732.
22. Wierzbicki, A.; Sikes, C. S.; Madura, J. D.; Drake, B. *Calc. Tiss. Int.* **1994**, *54*, 133.
23. Weiner, S.; Hood, L. *Science* **1975**, *190*, 987.
24. Fujisawa, R.; Kuboki, Y.; Sasaki, S. *Calc. Tiss. Int.* **1986**, *39*, 248.
25. Fujisawa, R.; Kuboki, Y. *Bull. l'Inst. océanographique Monaco* **1994**, *14*, 159.
26. Fujisawa, R.; Kuboki, Y. *Eur. J. Oral Sci.* **1998**, *106*, 249.
27. Huq, N. L.; Cross, K. J.; Reynolds, E. C.*Biochim. Biophys. Acta* **1995**, *1247*, 201.
28. Huq, N. L.; Cross, K. J.; Reynolds, E. C. *J. Dent. Res.* **1995**, *74*, 579.
29. Huq, N. L.; Cross, K. J.; Stanton, D.; Reynolds, E. C. *J. Dent. Res.* **1996**, *75*, 601.
30. Fujisawa, R.; Kuboki, Y. *Biochim. Biophys.Acta* **1991**, *1075*, 56.
31. Füredi-Milhofer, H.; Moradian-Oldak, J.; Weiner, S.; Veis, A.; Mintz, K.; Addadi, L. *Conn. Tiss. Res.* **1994**, *30*, 251.
32. George, A.; Bannon, L.; Sabsay, B.; Dillon, J. W.; Malone, J.; Veis, A.; Jenkins, N. A.; Gilbert, D. J.; Copeland, N. G. *J. Biol. Chem.* **1996**, *271*, 32869.
33. a) Crossley, M. A.; Huq, N. L.; Kirszbaum, L.; Reynolds, E. C. *J. Dent. Res.* **1996**, *75*, 154. b) Reynolds, E. C. unpublished results.
34. Long, J. R.; Dindot, J. L.; Zebroski, H.; Kiihne S.; Clark, R. H.; Campbell, A. A.; Stayton, P. S.; Drobny, G. P. *Proc. Nat. Acad. Sci. USA* **1998**, *95*, 12083.
35. Prescott, B.; Renugopalakrishnan, V.; Glimcher, M. J.; Bhushan, A.; Thomas Jr., G. J. *Biochemistry* **1986**, *25*, 2792.
36. George, A.; Sabsay, B.; Simonian, P. A. L.; Veis, A. *J. Biol. Chem.* **1993**, *268*, 12624.
37. Aoki, T.; Yamao, Y.; Yonemasu, E.; Kumasaki, Y.; Kako Y. *Arch. Biophys. Biochem.* **1993**, *305*, 242.
38. Sørensen, E. S.; Petersen, T. S. *Biochem. Biophys. Res. Comm.* **1994**, *198*, 200.
39. Swaisgood, H. E. In *Developments In Dairy Chemistry – 1*. Fox, P. F, Ed.; Applied Science Publishers: London, 1982, pp 1-43.
40. Price, P. A.; Rice, J. S.; Williamson, M. K. *Prot. Sci.* **1994**, *3*, 822.