Electronic Supplementary Information

M1. A model of reaction

Here we present a simplified model of reaction (Scheme S1), to make the data analysis simpler while still retaining the critical aspects of the reaction.
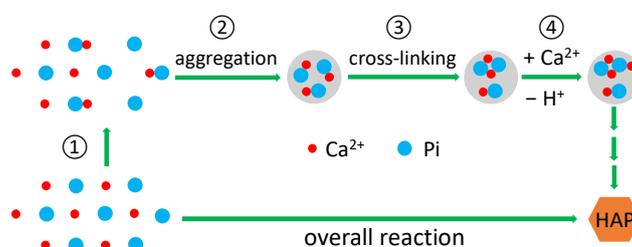
Scheme S1 Calcium phosphate precipitation via aggregation-facilitated cross-linking.

Inorganic phosphate (Pi) is present mainly as $H_2PO_4^-$ and HPO_4^{2-} in the near-neutral medium (eqn (1)). They bind Ca^{2+} ions, forming $CaH_2PO_4^+$ and $CaHPO_4$ (eqn (2) and (3), Step ① in Scheme S1), respectively; where x and $(1 - x)$ represent their proportions. Given the H_2O molecules in the coordination sphere of a Ca^{2+} ion,[1, 2] $CaH_2PO_4^+$ and $CaHPO_4$ could be called "calcium phosphate clusters".

These solution clusters form aggregate particles (eqn (5), Step ② in Scheme S1), known as ACP. Under the experimental conditions, the incipient ACP is dominantly composed of $CaHPO_4$ clusters, while the content of $CaH_2PO_4^+$ in ACP is much lower, as judged by their formation constants. Moreover, most of the $CaH_2PO_4^+$ clusters in the particles become $CaHPO_4$ clusters by releasing protons (eqn (4)). Because the clusters in an aggregate particle are closer to each other than those in solution, the cross-linking among them takes place more readily (eqn (6), Step ③ in Scheme S1), where m_0 denotes the number of phosphates in a cross-linked cluster. Some of the cross-linked clusters get additional Ca^{2+} ions, and the increased positive charges facilitate the proton release from a Ca^{2+}-bound HPO_4^{2-} ion, leading to the generation of a Ca^{2+}-bound PO_4^{3-} ion (eqn (7), Step ④ in Scheme S1).

The overall reaction is shown in eqn (8), where y and $(1 - y)$ represent the PO_4^{3-} and HPO_4^{2-} proportions in the precipitate formula $Ca_{(1-y)m_0}(HPO_4)_{(1-y)m_0}Ca_{1.5ym_0}(PO_4)_{ym_0}$, respectively. Using the numbers of Ca and P atoms in the formula, we relate y with the experimentally determined Ca/P molar ratio (r, eqn (9)). Using eqn (9), we express the overall reaction as eqn (10), where the formula $Ca_{rm_0}(HPO_4)_{(1-y)m_0}(PO_4)_{ym_0}$ represents a cross-linked cluster that exists as the structural unit of a precipitate particle. The formula is equivalent to $Ca_{rm_0}(H)_{(1-y)m_0}(PO_4)_{m_0}$, and the latter becomes $Ca_{10}(OH)_2(PO_4)_6$ if $m_0 = 6$, $r = 5/3$, and $(H)_{-2} = (OH)_2$. Following the convention of presenting precipitation reactions, we express the dissociation constant in eqn (11), where $(IP)_{m_0}$ and $[S_{m_0}]$ represent the ionic product and the precipitate concentration in terms of the structural unit, respectively. To indicate the multiple structural units in a particle, we rewrite eqn (10) as eqn (12). For now, however, we focus on the reaction concerning the structural unit (eqn (10)). But one should be aware that the structural unit is within an aggregate particle, rather than a discrete one dispersed in solution.

$$H_2PO_4^- = H^+ + HPO_4^{2-} \quad (K_{a2} = 10^{-7.21} \text{ at 25 °C }[3]) \tag{1}$$

$$xCa^{2+} + xH_2PO_4^- = xCaH_2PO_4^+ \quad (0 < x \leq 1,\ K_2 = 25.6 \text{ at 25 °C }[4]) \tag{2}$$

$$(1 - x)Ca^{2+} + (1 - x)HPO_4^{2-} = (1 - x)CaHPO_4 \quad (K_1 = 548 \text{ at 25 °C }[4]) \tag{3}$$
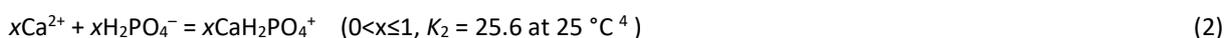
where K_2 and K_1 are the formation constants of $CaH_2PO_4^+$ and $CaHPO_4$, respectively.

$$CaH_2PO_4^+ = H^+ + CaHPO_4 \quad (K_a = K_1K_{a2}/K_2 = 10^{-5.88} \text{ at 25 °C}) \tag{4}$$
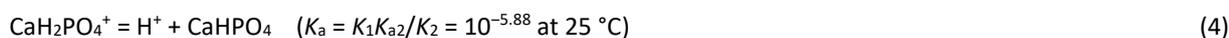
$$v\text{CaHPO}_4 = (\text{CaHPO}_4)_v \tag{5}$$

where v denotes the number of clusters in an aggregate particle.

$$(\text{CaHPO}_4)_{m_0} = \text{Ca}_{m_0}(\text{HPO}_4)_{m_0} \quad (m_0 \leq v) \tag{6}$$

$$0.5ym_0\text{Ca}^{2+} + \text{Ca}_{m_0}(\text{HPO}_4)_{m_0} = ym_0\text{H}^+ + \text{Ca}_{(1-y)m_0}(\text{HPO}_4)_{(1-y)m_0}\text{Ca}_{1.5ym_0}(\text{PO}_4)_{ym_0} \tag{7}$$

$$rm_0\text{Ca}^{2+} + (1-x)m_0\text{HPO}_4{}^{2-} + xm_0\text{H}_2\text{PO}_4{}^- = (x+y)m_0\text{H}^+ + \text{Ca}_{(1-y)m_0}(\text{HPO}_4)_{(1-y)m_0}\text{Ca}_{1.5ym_0}(\text{PO}_4)_{ym_0} \tag{8}$$

$$y = 2(r-1) \tag{9}$$

$$rm_0\text{Ca}^{2+} + (1-x)m_0\text{HPO}_4{}^{2-} + xm_0\text{H}_2\text{PO}_4{}^- = (x+y)m_0\text{H}^+ + \text{Ca}_{rm_0}(\text{HPO}_4)_{(1-y)m_0}(\text{PO}_4)_{ym_0} \tag{10}$$

$$K_{d,m_0} = \frac{(\text{IP})_{m_0}}{[S_{m_0}]} \tag{11}$$

$$rm_0\text{Ca}^{2+} + (1-x)m_0\text{HPO}_4{}^{2-} + xm_0\text{H}_2\text{PO}_4{}^- = (x+y)m_0\text{H}^+ + \frac{1}{u}\left\{\text{Ca}_{rm_0}(\text{HPO}_4)_{(1-y)m_0}(\text{PO}_4)_{ym_0}\right\}_u \tag{12}$$

The concentration $[S_{m_0}]$ in eqn (11) can be expressed in either the P or Ca content ($[S_\text{P}]$ and $[S_\text{Ca}]$, eqn (13)). The corresponding ionic product $(\text{IP})_{m_0}$ is defined in eqn (14), where the ion charges are eliminated for clarity and the exponential term m_0 is a constant to be determined from the solution chemistry data. The square brackets indicate (pseudo)equilibrium concentrations in molarity. In the presence of 150 mmol L^{-1} NaCl, the reaction-caused change in the ion strength of the solution could be neglected. Thus, we use the numerical values of molarity in place of the dimensionless activities.

$$[S_{m_0}] \equiv \left[\text{Ca}_{rm_0}(\text{HPO}_4)_{(1-y)m_0}(\text{PO}_4)_{ym_0}\right] \equiv \left[\text{Ca}_{rm_0}(\text{H})_{(1-y)m_0}(\text{PO}_4)_{m_0}\right] = \frac{[S_\text{P}]}{m_0} = \frac{[S_\text{Ca}]}{rm_0} \tag{13}$$

$$(\text{IP})_{m_0} = (\text{IP})_0{}^{m_0} \tag{14a}$$

$$(\text{IP})_0 = \frac{[\text{Ca}]^r[\text{HPO}_4]^{(1-x)}[\text{H}_2\text{PO}_4]^x}{[\text{H}]^{(x+y)}} \tag{14b}$$

By introducing eqn (13) and (14) into eqn (11), we obtain

$$[S_\text{P}] = (\text{IP})_0{}^{m_0} \cdot \frac{m_0}{K_{d,m_0}} \tag{15}$$

Taking the natural logarithm of eqn (15) gives

$$ln[S_\text{P}] = m_0 \cdot ln(\text{IP})_0 + ln\left(\frac{m_0}{K_{d,m_0}}\right) \tag{16}$$

Similarly, we obtain the equations where $[S_{m_0}]$ is expressed in terms of $[S_\text{Ca}]$:

$$[S_\text{Ca}] = (\text{IP})_0{}^{m_0} \cdot \frac{rm_0}{K_{d,m_0}} \tag{17}$$

$$ln[S_\text{Ca}] = m_0 \cdot ln(\text{IP})_0 + ln\left(\frac{rm_0}{K_{d,m_0}}\right) \tag{18}$$

To apply eqn (16) and (18) to the solution chemistry data, we need to calculate $(\text{IP})_0$ by eqn (14b), where the values of x and y are determined in different ways. The symbol y represents the proportion PO$_4{}^{3-}$/(PO$_4{}^{3-}$ + HPO$_4{}^{2-}$) in the precipitate formula. Its values from CRC onwards are calculated by eqn (9) from the experimentally determined r values. On the assumption that the five reaction systems were in a common pseudo-equilibrium state at a sampling time, the precipitate samples collected from these systems should be the same in composition and structure, only different in quantity. Hence, the average r value of the five systems is used in calculating the y value. As to the amorphous precipitate formed mainly from the CaHPO$_4$ clusters in the induction period, we assume $r = 1$ and thus $y = 0$.

The x values are dependent on the solution composition. Since a CaH$_2$PO$_4{}^+$ cluster and a CaHPO$_4$ cluster have roughly the same probability of forming aggregate particles via random collision, the CaHPO$_4$ proportions that are derived from

$H_2PO_4^-$ and HPO_4^{2-} in an aggregate particle are equal to their proportions in solution (x and $(1-x)$), respectively. Therefore, we calculate the x values for each of the five systems by eqn (19), using their respective pH values.

$$x = \frac{[\text{CaH}_2\text{PO}_4^+]}{[\text{CaH}_2\text{PO}_4^+]+[\text{CaHPO}_4]} = \frac{K_2[\text{H}^+]}{K_2[\text{H}^+]+K_1 K_{a2}} \tag{19}$$

It should be noted that the calcium phosphate clusters, freshly formed in the medium containing 150 mmol L^{-1} NaCl, interfered with the measurement of free ionic calcium ($[\text{Ca}^{2+}]$) by a Ca-selective electrode, and it was hard to estimate the magnitude of deviation in a continuous measurement. To obviate the direct measurement of $[\text{Ca}^{2+}]$, we calculate it from other measured quantities. Eqn (20) and (21) are the mass balances of the calcium and phosphate in the supernatant (LCa and LPi) from centrifugation, respectively. Here cCa and cPi denote the complexed Ca^{2+} ion and the Ca-bound phosphate, respectively, which are defined in eqn (22). To facilitate solving the unknown quantities, we construct eqn (23) from eqn (21) and (22). Then, we eliminate the quantity $[\text{Ca}^{2+}]$ by plugging eqn (20), which enables us to solve cCa from the equation and to determine its value using LCa, LPi, and pH. Finally, we calculate the $[\text{Ca}^{2+}]$ value by eqn (20).

$$LCa = cCa + [\text{Ca}^{2+}] \tag{20}$$

$$LPi = cPi + [\text{HPO}_4^{2-}] + [\text{H}_2\text{PO}_4^-] \tag{21}$$

$$cCa = cPi = [\text{CaHPO}_4] + [\text{CaH}_2\text{PO}_4^+] \tag{22}$$

$$\frac{cPi}{LPi - cPi} = \frac{(K_1 K_{a2}+K_2[\text{H}^+])\cdot[\text{Ca}^{2+}]}{K_{a2}+[\text{H}^+]} \tag{23}$$

References

1. J. R. Dorvee and A. Veis, *J. Struc. Biol.*, 2013, **183**, 278-303.
2. L.-W. Du, S. Bian, B.-D. Gou, Y. Jiang, J. Huang, Y.-X. Gao, Y.-D. Zhao, W. Wen, T.-L. Zhang and K. Wang, *Cryst. Growth Des.*, 2013, **13**, 3103-3109.
3. D. R. Lide, *CRC Handbook of Chemistry and Physics*, CRC Press, 2008.
4. A. Chughtai, R. Marshall and G. H. Nancollas, *J. Phys. Chem.*, 1968, **73**, 208-211.

Table S1. Linear fitting results of ln(S_P) *vs.* ln(IP)$_0$ and ln(S_{Ca}) *vs.* ln(IP)$_0$ plots*

(a) Series CPn (n = 3.0, 3.5, 4.0, 4.5, 5.0)

CPn	ln(S_P) *vs.* ln(IP)$_0$ plot				ln(S_{Ca}) *vs.* ln(IP)$_0$ plot				slope average
	Slope	Intercept	R-Square	Data	Slope	Intercept	R-Square	Data	
Ind-1	1.87 ± 0.15	15.45 ± 1.86	0.9874	del 4.0	1.97 ± 0.28	16.81 ± 3.55	0.9599	del 4.5	1.92 ± 0.05
Ind-2	1.66 ± 0.35	12.88 ± 4.33	0.8829	all	1.62 ± 0.41	12.64 ± 5.15	0.8352	all	1.64 ± 0.02
Ind-3	1.84 ± 0.27	15.36 ± 3.45	0.9589	del 4.0	2.09 ± 0.40	18.76 ± 4.98	0.9009	all	1.97 ± 0.13
CRC	1.06 ± 0.01	-1.59 ± 0.06	0.9999	[3.5, 4.5]	1.03 ± 0.16	-1.44 ± 0.76	0.9768	[3.5,4.5]	1.05 ± 0.02
1H	0.619 ± 0.10	-4.29 ± 0.35	0.9497	del 3.0	0.707 ± 0.078	-3.69 ± 0.27	0.9761	del 3.0	0.66 ± 0.04
3H	0.670 ± 0.04	-4.60 ± 0.12	0.9921	del 3.0	0.626 ± 0.022	-4.40 ± 0.06	0.9976	del 3.0	0.65 ± 0.03
6H	0.498 ± 0.08	-4.92 ± 0.24	0.9538	del 3.0	0.498 ± 0.03	-4.62 ± 0.09	0.9925	del 3.0	0.50 ± 0.00
12H	0.587 ± 0.07	-5.13 ± 0.16	0.9630	all	0.549 ± 0.08	-4.90 ± 0.20	0.9384	all	0.57 ± 0.02

(b) Series pH (pH = 7.60, 8.00, 8.50, 9.00, 9.50)

pH	ln(S_P) *vs.* ln(IP)$_0$				ln(S_{Ca}) *vs.* ln(IP)$_0$				slope average
	Slope	Intercept	R-Square	Data	Slope	Intercept	R-Square	Data	
Ind-1	-3.18 ± 0.13	-47.34 ± 1.60	0.9984	7.6,8.5,9.5	-2.72 ± 0.18	-41.26 ± 2.28	0.9955	8.0,8.5,9.5	2.95 ± 0.23
Ind-2	-2.63 ± 0.11	-40.43 ± 1.36	0.9965	del 9.0	-2.42 ± 0.30	-37.35 ± 3.74	0.9702	del 7.6	2.53 ± 0.11
Ind-3	-2.49 ± 0.24	-38.46 ± 2.94	0.9735	all	-1.13 ± 0.04	-21.17 ± 0.51	0.9987	[8.0,9.0]	1.81 ± 0.68
CRC	-0.588 ± 0.059	-8.45 ± 0.19	0.9801	del 7.6	-0.606 ± 0.077	-8.20 ± 0.25	0.9683	del 7.6	0.60 ± 0.01
1H	-0.463 ± 0.045	-8.11 ± 0.15	0.9727	all	-0.456 ± 0.048	-7.78 ± 0.16	0.9679	all	0.46 ± 0.004
3H	-0.368 ± 0.070	-7.67 ± 0.22	0.9024	all	-0.409 ± 0.027	-7.49 ± 0.09	0.9866	all	0.39 ± 0.02
6H	-0.262 ± 0.11	-7.32 ± 0.33	0.7400	del 9.5	-0.428 ± 0.062	-7.54 ± 0.19	0.9602	del 9.5	0.35 ± 0.08
12H	-0.412 ± 0.069	-7.27 ± 0.13	0.9474	del 9.0	-0.397 ± 0.11	-6.90 ± 0.20	0.8770	del 9.0	0.40 ± 0.01

* Significant digit = 3 for (IP)$_0$, (S_P), and (S_{Ca}). Curve-fitting was performed using Origin software.