Isolation of the intra-crystalline proteins and kinetic studies in Struthio camelus (ostrich) eggshell for amino acid geochronology

(Article begins on next page)

07 August 2024

New experimental evidence for in-chain amino acid racemization of serine in a model peptide

Beatrice Demarchi[1]*, Matthew Collins[1], Ed Bergström[2,4], Adam Dowle[3,4] , Kirsty Penkman[1],

Jane Thomas-Oates[2,4], Julie Wilson[2,5]

[1] BioArCh, Departments of Archaeology and Chemistry, University of York, York YO10 5DD,

UK

[2] Department of Chemistry, University of York, York YO10 5DD, UK

[3] Bioscience Technology Facility, Department of Biology, University of York, York, YO10

5DD, UK

[4] Centre of Excellence in Mass Spectrometry, University of York, York, YO10 5DD, UK

[5] Department of Mathematics, University of York, York, YO10 5DD, UK

* Corresponding author. E-mail: beatrice.demarchi@york.ac.uk; Tel: +44 (0) 1904 322574, Fax:

+44 (0) 1904 322516

Abstract

The facile racemization of protein-bound amino acids plays an important role in the ageing and pathologies of living tissues, and can be exploited for protein geochronological studies in sub-fossil biominerals. However, the in-chain degradation pathways of amino acids are complex and difficult to elucidate. Serine has proven to be particularly elusive and its ability to racemize as a peptide-bound residue (like asparagine and aspartic acid) has not been demonstrated. This study investigates the patterns of degradation of a model peptide (WNSVWAW) at elevated temperatures, quantifying the extent of racemization and peptide bond hydrolysis using RP-HPLC and tracking the presence of degradation products by MALDI-MS. We provide direct evidence that under these experimental conditions both serine and asparagine are able to undergo racemization as internally-bound residues, which shows their potential for initiating protein breakdown and provides an explanation for the presence of D-enantiomers in living mammalian tissues.

Introduction

The presence of D-amino acids in mammalian tissues is well recognized[1] and contradicts previous assumptions that only L-enantiomers may be found in living proteins (with the exception of bacteria)[2]. Both free and protein-bound D-amino acids (Asp, Ala, Ser) have been observed in mammalian tissues and have also been correlated to ageing and pathophysiological processes, such as Alzheimer's and Parkinson's disease, schizophrenia and renal disease[3]. Free D-Ser occurs naturally in mammalian brain and has been shown to act as a neurotransmitter and a gliotransmitter produced by the enzyme serine racemase[4-5].

The non-enzymatic (chemical) modification of protein-bound L-Ser to D-Ser plays a role in the ageing of "long-lived" proteins in cartilage, brain, dentine and eye lens[6]. This is a curious finding, because it is believed that most racemization (with the exception of peptide-bound Asn/Asp, where in-chain racemization is succinimide-mediated[7]) occurs at the N-terminus of proteins or in free amino acid molecules[8].

In-vivo Ser racemization occurs in at least two sites in the eye-lens protein alpha-crystallin and it has been proposed that Ser may promote cleavage of its N-terminal peptide bond, postulated to occur due to an N $\rightarrow$ O acyl shift [9-10]. Steinberg and colleagues[11] reported enhanced rates of Ser racemization in a poly-Ser peptide over that observed in free Ser, after heating to 100˚C under various pH conditions.

Protein diagenesis studies have focused on the chemical mechanisms of *post-mortem* Ser degradation, because of its potential use as a geochronometer in sub-fossil biominerals[12], especially those retaining a closed-system of intra-crystalline proteins[13-14]. These studies have revealed that Ser undergoes an atypical pattern of racemization, proceeding initially more rapidly than that of Asp/Asn (Asx), followed by a general decrease in the extent of racemization over time. There are two plausible explanations for the initial burst of racemization (Figure 1); either (i) hydrolytic cleavage at Ser takes precedence over that of other amino acids (as postulated by Lyons and collegues[9] and Su and colleagues[10]), generating an overwhelming number of N-terminal Ser residues, which rapidly racemize, and are then hydrolyzed to free Ser, or (ii) Ser can undergo in-chain racemization when internally bound, as suggested by a computational study[15], although experimental evidence to support this proposal is currently lacking. The ability to distinguish between these two scenarios would provide a better understanding of *in-vivo*

degradation mechanisms, with important consequences for medical research and the study of age-related pathologies.
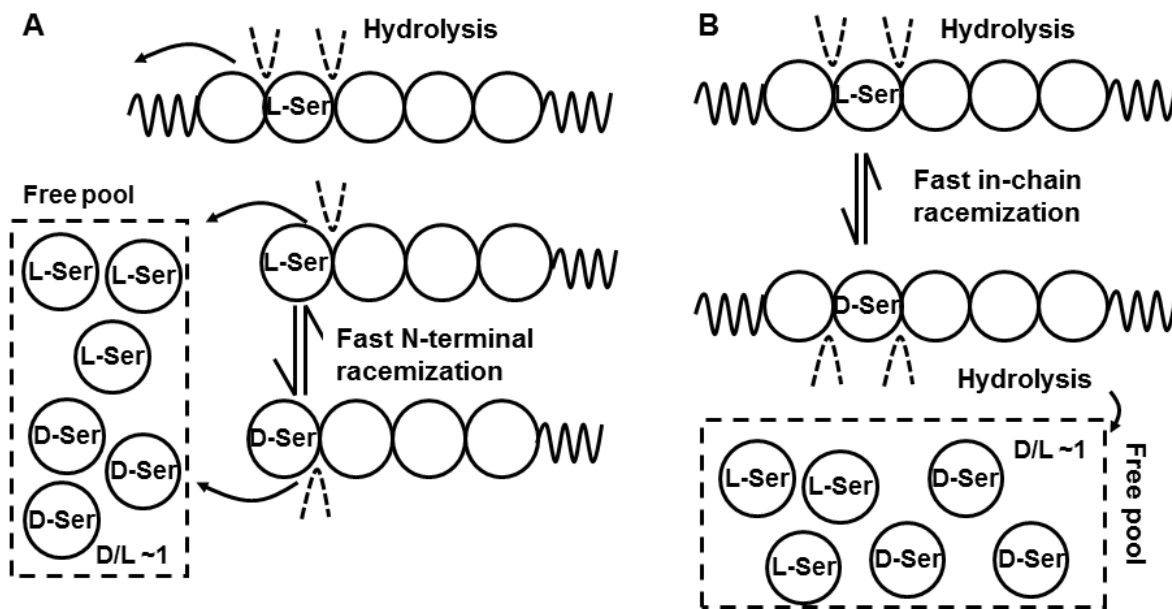
Figure 1. Schematic representation of two possible routes of Ser degradation: A. Preferential hydrolytic cleavage (dashed arrows) produces highly-racemized N-terminal Ser residues, which are released into the free amino acid pool; B. In-chain Ser racemization may occur anywhere in the peptide chain and on hydrolysis and these racemized Ser are released into the free pool.

In order to obtain direct experimental evidence of the pathways of Ser racemization, we have investigated a simple model peptide (sequence: Trp-Asn-Ser-Val-Trp -Ala-Trp, WNSVWAW), which was degraded at 140˚C in aqueous solution in sealed ampoules, to simulate a closed-system environment. Both bulk chiral amino acid analysis and mass spectrometry (MS) were applied to explore the pattern of peptide degradation. Quantification of the extent of degradation was obtained by examining two main diagenesis indicators: 1) the release of amino acids by peptide bond hydrolysis; 2) the extent of racemization, by measuring D/L values over time for

each amino acid. The distribution of mass fragments was assessed using TOF-MS. Collision induced dissociation (CID) product ion tandem mass spectrometry (MS/MS) was used to provide information on the amino acid sequence in the peptide products and on the possible modifications of amino acids within the sequence.

Experimental methods

Model peptide

Peptide WNSVWAW (theoretical pI/monoisotopic mass: 5.52 / 947.43) was purchased from Biomatik Corporation (Cambridge, Ontario, Canada). It corresponds to positions 75-81 of the sequence of struthiocalcin 1 (SCA-1), the main protein from ostrich (*Struthio camelus*) eggshell[16] (OES), a biomineral with importance as a substrate for protein diagenesis geochronology[17].

High temperature treatment

Artificial diagenesis of the peptide was induced by isothermal heating at 140°C. The dehydrated peptide was resuspended in ultrapure water (18.0 MΩ) at room temperature and 60 µL aliquots (~0.4 mM) were transferred into individual sterile glass ampoules, which were then flame-sealed and placed in an oven at 140˚C for various times (1, 2, 4, 6, 24, 48 hours). Three experimental replicates were prepared for each time point. After heating, two subsamples of solution (15-20 µL) were taken from each ampoule and transferred into sterile glass vials, one for the analysis of the total hydrolysable fraction (THAA) and one for the analysis of the free amino acid fraction (FAA). The remaining solution was transferred into plastic tubes and stored at -20°C, prior to MS analysis.

Chiral amino acid analysis

THAA subsamples were prepared by hydrolyzing the peptide solutions (previously flushed with N_2 to minimize oxidation) in 6 M HCl at 110°C for 24 h, followed by evaporation of the acid in a centrifugal evaporator and rehydration with a solution containing 0.01 mM HCl, 0.77 mM sodium azide at pH 2 and containing an internal standard (0.01 mM L-homo-arginine rehydration fluid). FAA subsamples were dried in a centrifugal evaporator and rehydrated with L-h-Arg rehydration fluid. The peptide solutions were analyzed for chiral amino acids on an automated RP-HPLC system equipped with a fluorescence detector, using a modification of the method of Kaufman and Manley[18]. In brief, 2 µL of rehydrated sample was mixed online with 2.2 µL of derivatizing reagent (260 mM *N*-isobutyryl-l-cysteine, 170 mM *o*-phthaldialdehyde, in 1 M potassium borate buffer, adjusted to pH 10.4 with KOH) immediately prior to injection. Separation of enantiomeric pairs was achieved on a C_{18} HyperSil BDS column (3 mm×250 mm) at 25 °C using a gradient elution of sodium acetate buffer (23 mM sodium acetate tri-hydrate, 1.5 mM sodium azide, 1.3 µM EDTA, pH 6.00±0.01), methanol, and acetonitrile (95% sodium acetate buffer and 5% methanol initially, grading to 50% methanol and 5% acetonitrile at 95 min).

This method allows accurate quantification of the L- and D-enantiomers of the amino acid residues present in the sample, except for Trp, which normally undergoes irreversible decomposition during acid hydrolysis[19]. However, both L-Trp and D-Trp were detected; the D/L values are therefore presented in this study (see Results and Discussion), although it should be noted that some degradation during preparation is likely to have occurred. Asn undergoes irreversible deamination to Asp during acid hydrolysis and is therefore, together with Asp, referred to as Asx[20].

Calibration curves were obtained by analyzing solutions of known concentrations of single-enantiomer standards for L-Asp, L-Ser, L-Val, L-Ala, L- and D-Trp, dissolved in L-h-Arg rehydration fluid. The lowest limit of the calibration is at 0.01 mM; therefore experimental concentration data lower than 0.01 mM are below the limit of quantification (LOQ). For the quantification of D-Asp, D-Ser, D-Val and D-Ala we derived amino acid-specific conversion factors from 100 standard solutions which are routinely analyzed in our laboratory, containing the enantiomeric pairs of Asp, Glu, Ser, Arg, Ala, Tyr, Val, Met, Phe, Ile, Leu at known D/L values (D/L = 0.2; D/L = 0.5; D/L = 1.0) and concentrations (40.78 µM). We then applied this conversion factor to calculate a "corrected" area for the D-enantiomer and calculated the concentration values using the calibration curve for the L-enantiomer. Since Trp is not included in the standard enantiomeric mixtures, we obtained calibration curves for both L- and D-Trp. Conversion factors and calibration curves are reported in Supporting Information (Table S-2)

MALDI-MS

1 µL aliquots of sample solutions (~0.4 µg/µL) from each experimental replicate were spotted onto a Bruker polished steel MALDI target plate, mixed together with 1 µL of α-cyano-4-hydroxycinnamic acid (CHCA) matrix solution (0.5% w:w in 50% ACN/H_2O v:v) and allowed to dry. External calibration was performed by analysing an adjacent spot containing six calibration peptides (des-Arg1-bradykinin, M+H$^+$ at *m/z* 904.681; angiotensin I, M+H$^+$ at *m/z* 1296.685; Glu1-fibrinopeptide B, M+H$^+$ at *m/z* 1750.677; ACTH (1-17 clip), M+H$^+$ at *m/z* 2093.086; ACTH (18-39 clip), M+H$^+$ at *m/z* 2465.198; ACTH (7-38 clip), M+H$^+$ at *m/z* 3657.929). Internal calibration was performed post-analysis by using a list of known CHCA matrix peaks and the known *m/z* of the intact peptide itself. MS analyses were performed using a Brukerultraflex III MALDI-TOF/TOF mass spectrometer equipped with a smartbeam^{TM} laser.

Mass spectra were acquired over the *m/z* range 0-1000. Each experimental replicate was spotted in duplicate and each spot analyzed in positive mode (number of laser shots: 800; laser power: 62%). A subset of the samples (a single spectrum for each of the experimental replicates for each time point) was analyzed using a Bruker solariX FTMS instrument with 9.4-Tesla superconducting magnet. The FTMS data were used for *de novo* sequencing of selected peptide fragments; this was carried out by manual interpretation of the product ion (MS/MS) spectra obtained on collision-induced dissociation (laser pulse frequency: 100-1000 Hz; laser power: 25-35%; number of shots: 50-400; CID voltage: 11-40 V).

Automated MS data analysis

In order to automate the analysis of the MS data, a list of "theoretical degradation products" was compiled, which included the $[M+H]^+$, and $[M+Na]^+$ *m/z* values for: a) the intact peptide, b) the degradation products produced by all combinations of the progressive losses of the N-terminal and C-terminal residues and c) the possible modifications occurring on each of these. These modifications were: loss of water (18 Da) by cyclisation of the peptide fragment occurring at any residue, dehydration of Ser and deamidation of Asn. For each mass spectrum, baseline correction was applied before considering the signal to noise ratios (SNR) of any peaks. No normalization was performed, as intensities were not compared between spectra; rather signal to noise levels were used to determine the presence or absence of a peptide signal. As noise levels can vary significantly across a spectrum, the local noise level was estimated from the intensities immediately before the peak in question, with the assumption that adjacent noise peaks should have similar intensities. At this stage, peaks with SNR < 3 were rejected. Each spectrum was searched against the list of theoretical degradation product *m/z* values. Any peak maximum within 0.1 *m/z* unit of a theoretical *m/z* value was investigated and only those corresponding to

the first peak of an isotope distribution considered further. More stringent cut-off points were applied at a second stage of the analysis of the data, in order to minimize the effect of low-intensity, low-m/z peaks, which interfere with the peptide signal in the spectra of samples generated using longer heating times; the results reported in Table 1 and Supporting Information (Table S-1) include only peaks with SNR > 7, which were observed for more than one time-point and in at least 50% of the replicate sample spectra.

Results and discussion

Extent of hydrolysis and amino acid racemization

Degradation of the peptide at high temperature and in an aqueous environment is expected to proceed via hydrolysis of the peptide bonds and racemization of the amino acid residues. The peptide investigated here contains hydrophobic amino acids (Val, Ala and three Trp residues) and hydrolysis of these peptide bonds is expected to be relatively slow. Indeed, the acid hydrolysis step performed for the quantification of the THAA concentrations was insufficient to break some of the more stable peptide bonds; the relative THAA composition of the unheated peptide shows that [THAA] Asx and Ala are over-represented (~ 25% instead of 14%, as would be expected from the composition of the peptide), whilst Val and Trp are under-represented (10% instead of 14%, 26% instead of 42%) (Figure 2A).

Upon heating, it is expected that hydrolysis would progressively release free amino acids (FAAs) until the [FAA] and [THAA] attain the same value (%FAA = 100). For all amino acids %FAA values remain <50% in samples heated for 0 to 48h (Figure 2B), and Ser shows the lowest %FAA values throughout the experiment. This suggests that during heating for between 0 and 48h, the peptide bonds involving Ser are stable and that only ~10% of the Ser residues are

released into the FAA pool (Figure 1), whilst at least ~90% of the serine present in the system at any given point is bound in the peptide chain, either as a terminal or internal residue.

A

B
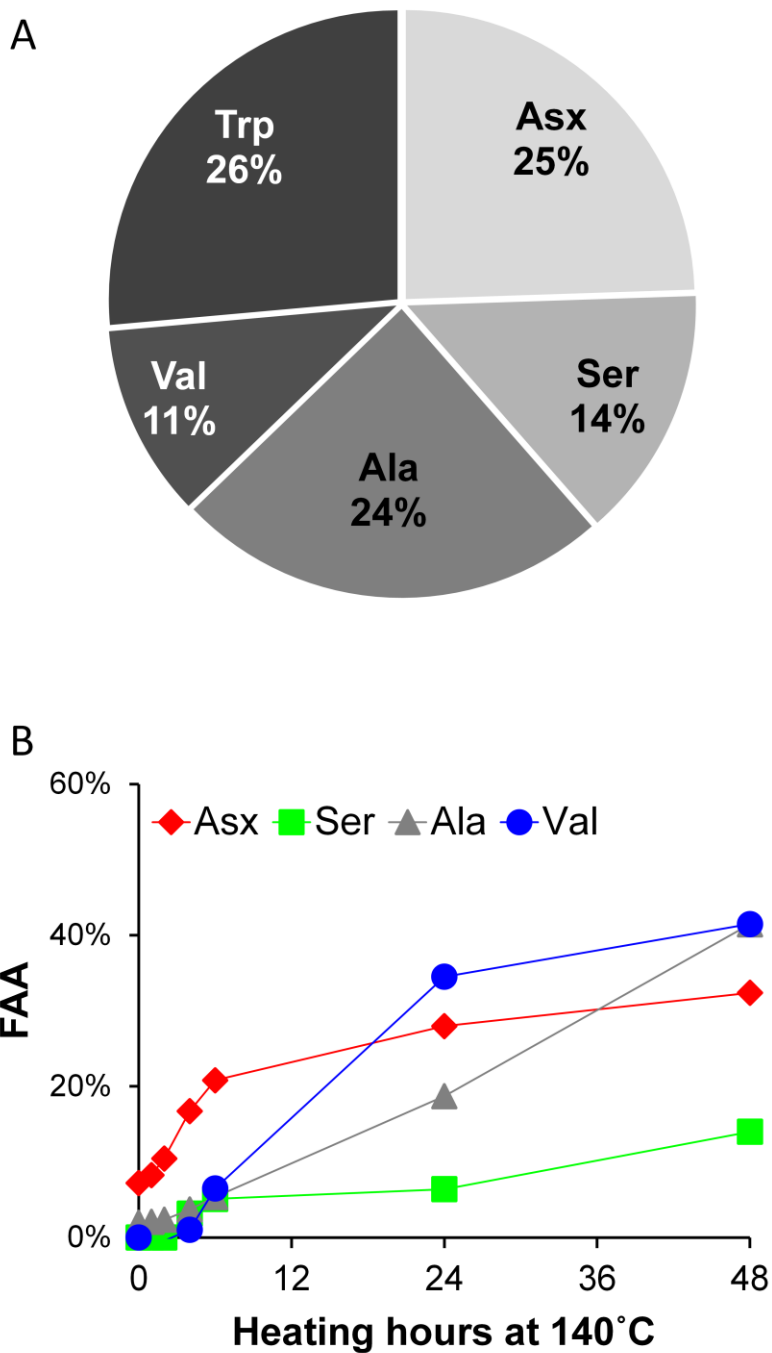
Figure 2. A.THAA composition of unheated peptide WNSVWAW. B. THAA and %FAA values for Asx, Ser, Ala and Val upon heating at 140°C. %FAA values were calculated for each time point as: %FAA = [FAA]/[THAA]x100.

As racemization during heating proceeds with time, it is expected that the D/L values for all amino acids will increase from 0 up to a maximum of 1 (racemic equilibrium) and that amino acids more prone to racemization will reach equilibrium earlier in the experiment. This was observed in this study (Figure 3) and the difference between the racemization rates of the five residues is extraordinary; Asx and Ser (THAA D/Ls ~ 0.9 after 6 h heating) displayed the fastest racemization rates. Trp, Ala and Val showed a completely different pattern of increase in the extent of racemization with heating time, and did not reach equilibrium by the end of the experiment (48 h). Interestingly, although Ser and Asx displayed a very similar pattern of racemization and they both plateau at near-equilibrium by 6 h heating, Ser THAA D/L values remained close to 0 after 1 hour heating, unlike Asx. For all amino acids except Asx the concentration of D-enantiomers was below the LOQ until 24 h heating; therefore FAA D/L values are not shown here.
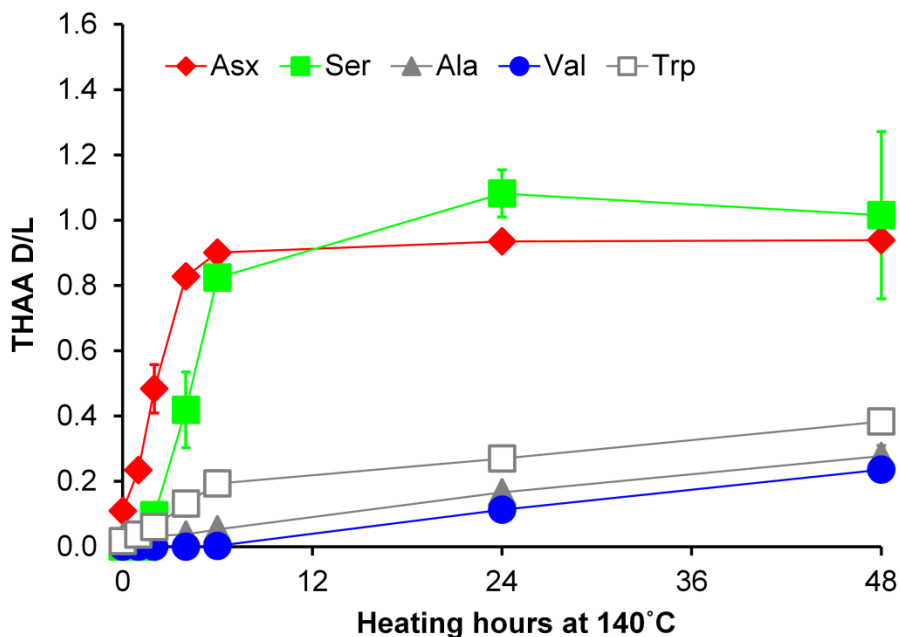
Figure 3. THAA D/L values for Asx, Ser, Ala, Val and Trp in peptide WNSVWAW upon heating at 140°C. Error bars represent one standard deviation around the mean of three experimental replicates. Note the initial burst of racemization displayed by Ser and Asx, while Trp, Ala and Val D/L values increase at a much slower rate and do not reach racemic equilibrium by 48h. Ser D/L values for heating over 24-48 h are slightly above unity but also more variable, presumably due to decomposition (dehydration) of Ser during the later stages of diagenesis, which might affect the two enantiomers differently.

Detection of degradation products by MALDI-MS

Mass spectrometry was used to examine the mass shifts in the peptide detectable during heating, by obtaining mass spectra for samples taken at each time point (Figures 4A and 4B). The degradation products were identified by matching *m/z* values with those for the expected hydrolytic breakdown products of the peptide and, where possible, by MS/MS and *de novo* sequencing (Table 1; see Supporting Information for full list of *m/z* values observed). MS data

obtained from the peptide samples were compared with procedural blanks (CHCA matrix, water blanks from the high-temperature experiments) and matrix peaks were excluded from the list of ions interpreted (Table 1).

Theoretical ion	m/z	time points	ID (MS/MS)
[WNSVWAW+H]$^+$	948.44	0-1-2 (D)	WNSVWAW
[WNSVWAW+Na]$^+$	970.42	0-1-2 (D)	WNSVWAW
[WNSVWAW-17+Na]$^+$	953.40	1-2-4-24 (D)	WNSVWAW
[WDSVWAW+H]$^+$	949.43	2-4-24 (D)	WDSVWAW
[NSVWAW-35 +H]$^+$/ [WNSVWA-35 +H]$^+$	727.32	1-2	Cyclic NSVWAW (-NH$_3$)
[NSVWAW-18 +H]$^+$/ [WNSVWA-18 +H]$^+$	744.35	0-1-2 (D)	MS/MS spectrum very similar to that obtained from the precursor at m/z 727.32 – i.e. cyclic NSVWAW
[NSVWAW-18 +Na]$^+$/ [WNSVWA-18 +Na]$^+$	766.33	1-2-4	As for spectrum of m/z 744.35 cyclic NSVWAW
[WDSVW+H]$^+$	692.31	2-4-24	Not recorded
[SVWAW+Na]$^+$	670.30	1-2-4-6-24-48	SVWAW
[SVWAW+H]$^+$	648.3	MS/MS only	SVWAW
[VWAW+Na]$^+$	583.27	24	Not recorded
[VWAW+H]$^+$	561.28	MS/MS only	VWAW
[SVWA+H]$^+$/[WAW+H]$^+$	462.24	0-1-2 (I)	WAW
[WNS-35 +H]$^+$/ [WDS-36 +H]$^+$	371.14	2-4-6 (D)	Not recorded
[WNS-18 +H]$^+$	388.16	0-1-2 (D)	Not recorded
[WNS-53 +H]$^+$/ [WDS-54 +H]$^+$/	353.12	1-2-4-6 (D)	Not recorded
[WN-18 +Na]$^+$	323.11	1-4	Not recorded

| [WN-18 +H]$^+$ | 301.13 | 0-1 | Not recorded |
| [SV-36 +H]$^+$ | 169.09 | 6-48 | Not recorded |

Table 1. Assignment of MALDI MS signals to possible degradation products of peptide WNSVWAW: *m/z*, heating times at which the ions are observed (hours), identification by MS/MS (product ion spectra are reported in Supporting Information). D = decrease in intensity over time; I= increase in intensity over time. Asn deamidation to Asp is indicated in the "ID" column by underlined "N" and "D".

Usually, both [M+Na]$^+$ and [M+H]$^+$ ions were detectable in the samples; however in some instances only the [M+Na]$^+$ species were consistently detected and could be identified, e.g. [SVWAW+Na]$^+$ and [VWAW+Na]$^+$. Using the [M+Na]$^+$ precursor complicates the product ion spectrum and therefore MS/MS analysis of the [M+H]$^+$ ions was performed to identify the amino acid sequence. Although these were only detected in a few analytical replicates, it was possible to isolate this precursor on the FT-ICR instrument and perform CID.

The peptide sequences NSVWAW and WNSVWA are isomeric and both could be formed by hydrolytic cleavage of the peptide chain; the product ion spectra of the dehydrated charged species at *m/z* 766.33 (M+Na$^+$-18), 744.35 (M+H$^+$-18) and 727.32 (M+H$^+$-18-NH$_3$) (see *m/z* 727.32 in Figure S-E, Supporting Information) show the presence of a cyclic peptide where the ion fragmentation series derives from the ring opening between Trp-Ala, Ala-Trp, Ser-Val, or Val-Trp but not between Asn-Trp or Asn-Ser. The stability of the Asn-Ser bond may be explained by the formation of a stable lactone between the two neighboring side chains upon Asn deamidation (observed as equivalent to the loss of the elements of ammonia from these residues – this results in the loss of 17 Da).

The precursor ions at m/z 670.3 and 648.3 correspond in mass to SVWAW (M+Na$^+$ and M+H$^+$ respectively) and although the y$_4$ ion is weak, the sequence can be identified and indicates presence of N-terminal Ser in this degradation product (Figure S-F, Supporting Information); the fact that the y ions deriving from the fragmentation at the Ser-Val bond are weak suggests that this bond is also stable in the gas phase.

Charged species at $m/z < 400$ fall within the region of the MALDI matrix clusters and did not yield interpretable product ion spectra, presumably because the spectra are "mixtures" of peptide and CHCA due to the instrument's isolation window for the precursor ion, and are therefore assigned based on matching of their m/z values with expected hydrolytic fragments.
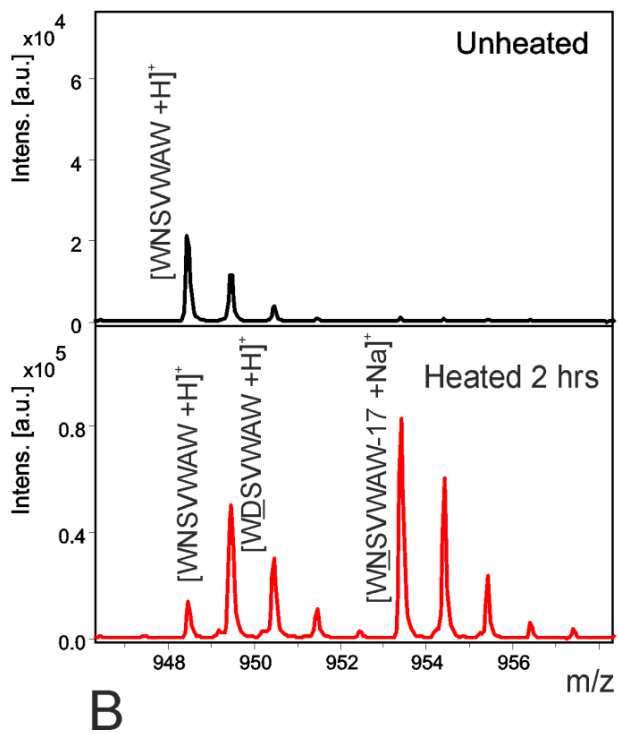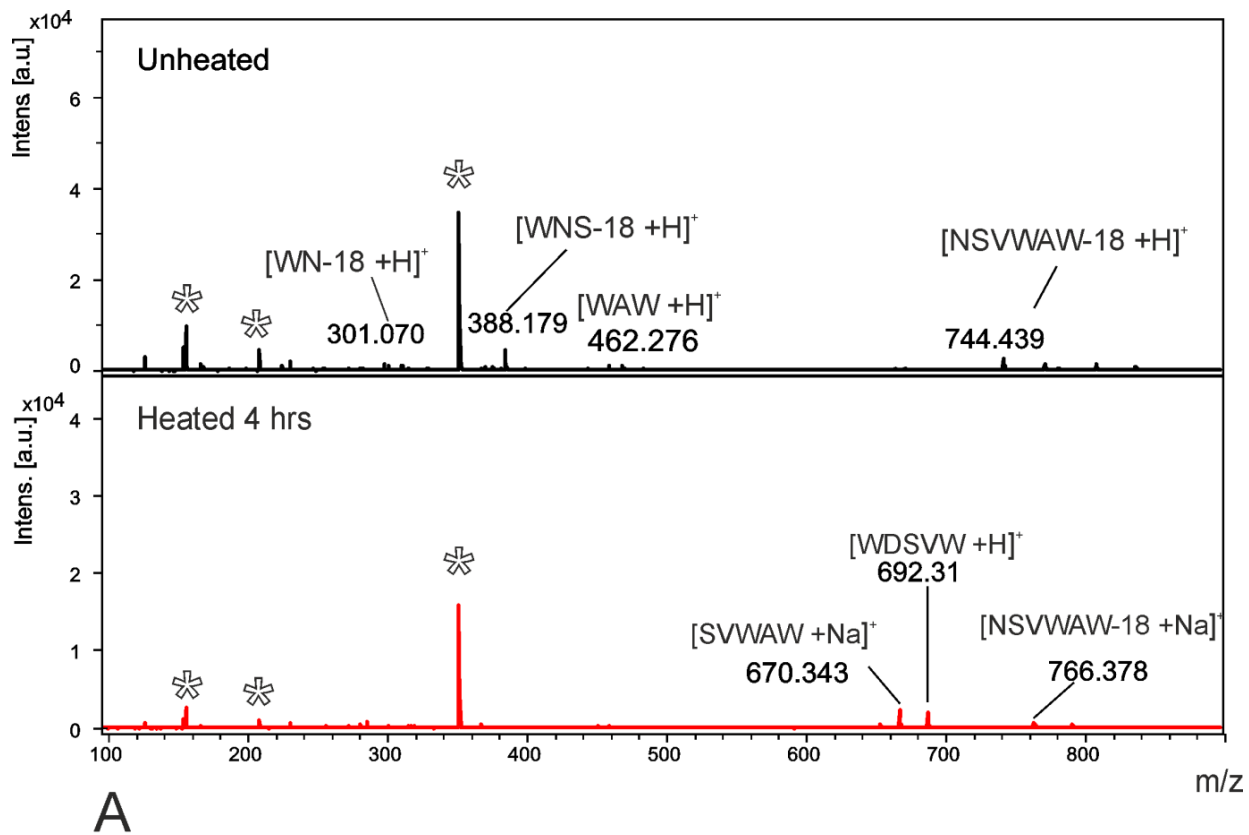
Figure 4: A. MS of peptide WNSVWAW unheated (top) and after 4 hours heating at 140˚C (bottom); an asterisk is used to indicate peaks observed in matrix and water blanks. B. Asn deamidation is shown by a +1 *m/z* unit increment on the native peptide (*m/z* 949.4) while loss of 17 *m/z* units from the [WNSVWAW+Na]$^+$ species (*m/z* 953.4), by 2 h heating can be accounted for by the formation of an intramolecular lactone between the Ser side chain and the newly generated Asp side chain.

Patterns of Ser diagenesis at high temperature

MS data were used to track the degradation products of peptide WNSVWAW formed upon heating, whilst at the same time the bulk chiral amino acid determinations monitored racemization. The focus was to determine whether in-chain racemization of Ser (and Asn) could be assessed.

MS is able to detect the presence of degradation products and their relative abundance, and to enable peptide sequencing of those degradation products which are *detectable* under the experimental conditions used; however, MS may not account for *all* the degradation products. Moreover, in the peptide under investigation three Trp residues are present and therefore degradation products with different amino acid sequence but identical mass can be formed. As a result, it is not possible to link the extent of racemization or hydrolysis of a specific residue directly to the position it occupies within the peptide chain at any one time.

However, by combining bulk chiral amino acid analysis and MS we can begin to link the extent of racemization of amino acids to their three alternative states – either bound internally within the peptide, cleaved to a terminal position or free in solution – and thus begin to explore the alternative hypotheses outlined in Figure 1.

17

[M+H]$^+$ and [M+Na]$^+$ for the intact peptide could be observed in the spectra of unheated

samples (Table 1 and Figure 4), and were also present in samples heated for 1 and 2 hours at

140°C, although after 2 h heating the SNR of the signal is reduced (on average) to ~20% of the

unheated signal. The spectra of the unheated samples, however, also show the presence of

fragments at lower *m/z* than the intact peptide (Figure 4A); this is likely to be due to incomplete

synthesis of the peptide during manufacture. This unfortunately complicates the interpretation of

the experimental data, because the bulk amino acid data will reflect the contribution to the extent

of racemization of all the residues originally present in the peptide sample. Although the relative

intensity of ionized species in mass spectra is not necessarily linked to their relative

concentration in the sample, the SNR for these incomplete synthesis products is usually 6-9

times less than that for the [M+H]$^+$ ion of the full length peptide. More importantly, Ser is

internally bound in all these incomplete synthesis products, except for the (tentatively identified)

tri-peptide WNS (*m/z* 388.16), which occurs in its doubly-dehydrated form (-36 Da) and is

therefore likely to be cyclic. Within 2 h of heating, the low mass peptides were no longer

detectable in the MS dataset (Figure 4A), probably due to hydrolysis to FAA; the [FAA] values

for all amino acids between 0-2 h heating was below the LOQ and this supports the hypothesis

that the relative concentration of the incomplete synthesis products in the sample is negligible.

Therefore, at the beginning of the experiment, Ser is bound, either within the peptide chain or in

the (low-concentration) incomplete synthesis products.

Between 1 and 2 hours heating, a range of degradation products becomes detectable, including

two species at *m/z* 949.4 and 953.4, which correspond to the intact peptide in which Asn

deamidation has occurred ([M+H]$^+$at *m/z* 949.4), and where formation of an intramolecular

lactone has taken place between the Ser side chain and the newly generated Asp side chain

([M+Na]$^+$ at *m/z* 953.4) (Figure 4B and Table S-1 in Supporting Information). Within this time interval, the extent of racemization (THAA D/L) remains close to zero for all amino acids except Asx, which increases rapidly from ~0.11 (in the unheated peptide) to ~0.48. Asn deamidation is known to proceed via intramolecular nucleophilic attack on the carbonyl carbon resulting in the formation of an unstable succinimide intermediate (Asu) which is able to racemize very rapidly [21-23]. This is supported here: after heating to 140°C between 0 and 24 h, Asx D/L values reached near-equilibrium and after 24 h, the ions at *m/z* 949.4 and 953.4, representing the intact peptide, are no longer detectable in the MS data.

Other degradation products are formed between 1-2 h heating, such as the cyclic peptide NSVWAW (Table 1); these are consistent with hydrolytic damage occurring at the peptide chain termini. Hydrolysis may occur at either terminus of the chain, although the hydrophobicity of the C-terminus is likely to hinder rapid hydrolysis. Indeed, limited hydrolysis during this time interval (and beyond) is confirmed by the low %FAA values detected up to 6 h heating (%FAA < 10% for all amino acids except Asx) (Figure 2B).

Both %FAA values (Figure 2B) and THAA D/L values (Figure 3) demonstrate very clearly that the extensive racemization of Ser, observed between 0-6 hours heating, occurs while Ser is peptide-bound (either terminally or internally). Since we only acquired hydrolytic data at 140˚C, it is not possible to calculate the activation energy for Ser racemization in the peptide WNSVWAW. However, we have calculated the racemization rate (*k*) between 2-6 h heating to be ~6 times faster than in ostrich eggshell heated at 140˚C, for a similar D/L range (*k* = 1.2x10^{-5} s^{-1} in OES vs 7.5x10^{-5}s^{-1} in peptide WNSVWAW; THAA D/L = 0.1-0.8). This rate is also similar to the deamidation-mediated racemization of the neighboring Asn between 0-6 h heating (*k* = 6.8x10^{-5}s^{-1}).

The data also provide information as to whether Ser is prevalently found in the N-terminal or internal position; between 2-6 h heating Ser is *internally* bound in almost all the degradation products observed by MALDI-MS, except for two: a) the tri-peptide WNS (-53 Da) and b) the [M+Na]$^+$ ion at *m/z* 670.3, identified as SVWAW (Table 1). The tripeptide WNS is present in a dehydrated form and therefore likely to be cyclic and so stable, but it was identified only based on its *m/z* value and its mass falls in the *m/z* region where the CHCA signal heavily interferes with the analytes, therefore it should be treated with caution. However, the degradation product SVWAW is clearly identifiable in samples generated by heating for 1-24 h ; this peptide's formation suggests that sequential N-terminal hydrolysis of the intact peptide occurs upon heating, and that *some* of the racemization observed will occur while Ser is bound at the N-terminus of the peptide fragment SVWAW. MS data do not allow quantification of the proportion of the intact peptide that had been converted to SVWAW due to cleavage of the Asn-Ser bond; SNR of the [SVWAW+Na]$^+$ ion is between 9 and 3.5 times lower (after 1 h and 6 h heating, respectively) than that of [SVWAW+Na]$^+$ in the unheated peptide. Bulk amino acid data are more informative than MS for assessing the proportion of N-terminal Ser: by 6 hours heating, a maximum of just 20% FAA Asn is detected in the system. As Asn is on the N-terminal side of Ser, and hydrolysis of the Asn-Ser bond has to occur before Ser may occupy the N-terminal position, this indicates that only a small proportion of Ser (<20%) is N-terminally bound. Given that %FAA Ser for this time interval is negligible (Figure 2B), at least 80% of the total Ser present in the system is *internally* or *C-terminally* bound. The low intensity of fragment ions generated by fragmentation of the Ser-Val bond upon CID (Figure S-F for the [M+H$^+$] ion) might suggest the stability of this bond in the gas phase, while its stabilityin the liquid phase is supported by the fact that the degradation product VWAW is detected only after 24 h heating and

that [FAA] concentrations of Val are below the LOQ up to 6 h heating. Therefore Ser is not likely to be exposed at the C-terminus, further demonstrating that Ser mainly occupies an internal position in the chain.

Therefore, both N-terminal and internal Ser are present in the THAA pool; if racemization was solely driven by N-terminal Ser, with no in-chain racemization, we would expect to find the overall THAA D/L Ser values significantly dampened by the "bound-Ser" pool (estimated to represent at least 80% of the total pool), where THAA D/L=0.

Consequently, it is unlikely that the observed rapid racemization of Ser in peptide WNSVWAW is linked to preferential hydrolysis of this residue to N-terminal Ser; indeed we observed that when the N-terminal flanking residue is Asn, the Asn-Ser bond is very stable, probably due to the formation of a lactone between the two side chains. The only satisfactory explanation for the data is that Ser is racemizing internally - as proposed by Takahashi and colleagues[15].

Whilst Asn racemization is deamidation-mediated, the mechanism for Ser racemization remains to be fully elucidated. However, the study of Takahashi et al., based on density-functional theory (DFT), suggested that Ser racemization may occur via a similar mechanism to the Asx succinimide intermediate, i.e. via two-H_2O-assisted enolization, involving the H(α)-atom and the CO(α) groups, where the enol intermediate is stabilized by the side-chain C-O bond[15] (Figure 5).

The calculated energy barrier reported by Takahashi and colleagues for Ser is 136 kJ/mol (or 119 kJ/mol if the values are corrected for the zero-point vibrational energy); the higher value is not dissimilar to that calculated experimentally for some biomineral proteins (ostrich eggshell: 130 kJ/mol[24]; limpet shells: 135 kJ/mol[25]). Our study provides independent evidence for in-chain

racemization of Ser, but whether the mechanism proposed by Takahashi and colleagues is responsible for the observations reported here will require further investigation.
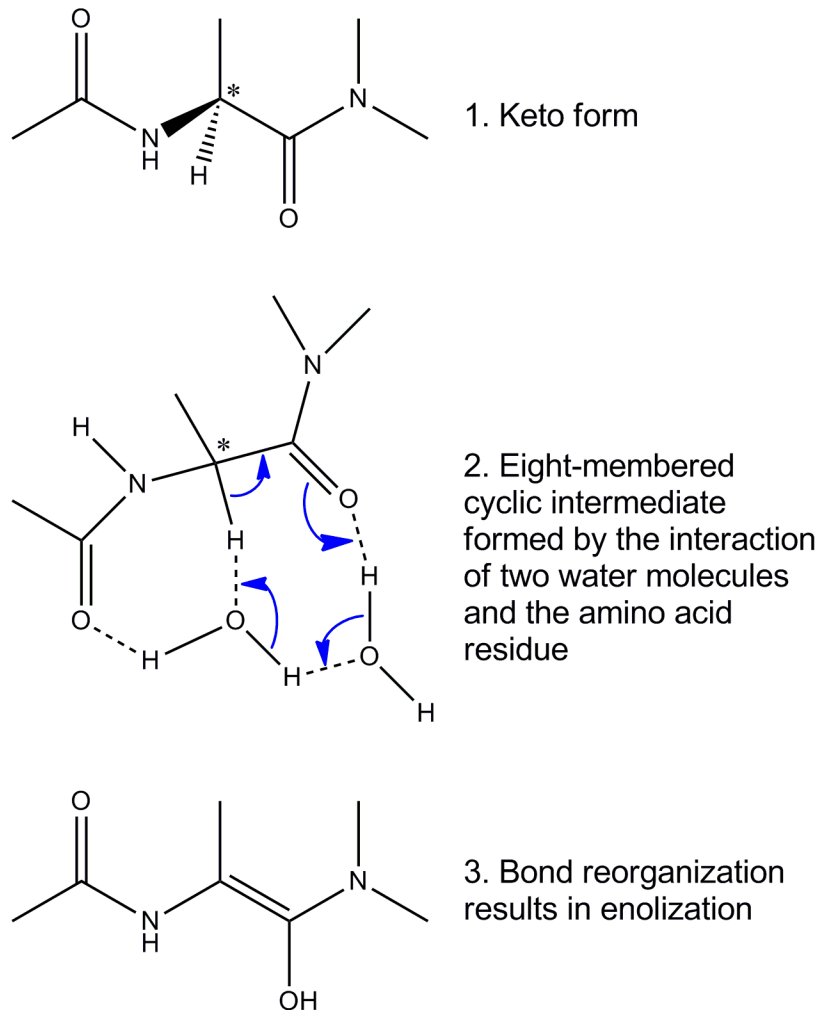
Figure 5. Mechanism of two-H_2O-assisted enolization proposed by Takahashi and colleagues for in-chain racemization of amino acids[15].

Conclusions

We have investigated the breakdown patterns of peptide WNSVWAW upon heating at 140°C by both bulk chiral amino acid (RP-HPLC) and MS analyses. These data provided direct

experimental evidence that Ser and Asn are able to undergo racemization as bound amino acid residues under these experimental conditions, and that their pathways of diagenesis are different from those followed by Val, Ala and Trp.

The model of racemization proposed by Mitterer and Kriausakul in 1984[8] postulates that in-chain racemization is generally not a common mechanism of racemization for most amino acids and that N-terminal racemization is the most likely pathway; our data support this, but we show that in the peptide examined here Ser is predominantly *internal* whilst undergoing very extensive racemization. Therefore, Ser represents an important exception to the model of Mitterer and Kriausakul and it is possible that other hydroxy-amino acids, such as threonine, may display similar degradation patterns. We suggest that this may be a common pathway for serine degradation *in vivo*, with implications for the understanding of the mechanisms of protein degradation in living tissues.
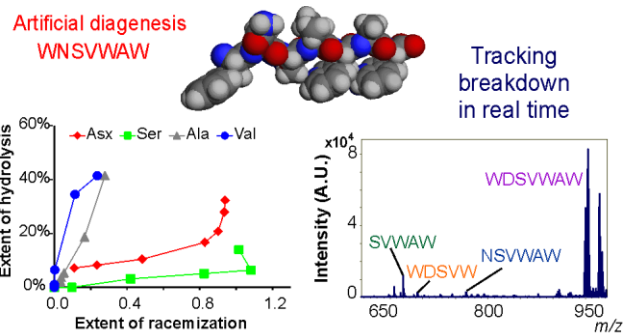
References

(1)   Fujii, N. *Bioland Pharm Bull* **2005**, *28*, 1585-1589.

(2)   Corrigan, J. J. *Science* **1969**, *164*, 142-149.

(3)   Hamase, K.; Morikawa, A.; Zaitsu, K. *J Chromatogr B* **2002**, *781*, 73-91.

(4)   Crow, J. P.; Marecki, J. C.; Thompson, M. *Neurol Res Int* **2012**, *2012*, ID: 625245.

(5)   Nishikawa, T. *J Chromatogr B* **2011**, *879*, 3169-3183.

(6)   Hooi, M. Y. S.; Truscott, R. J. W. *Age* **2011**, *33*, 131-141.

(7)   Stephenson, R. C.; Clarke, S. *J Biol Chem* **1989**, *264*, 6164-6170.

(8)   Mitterer, R. M.; Kriausakul, N. *Org Geochem* **1984**, *7*, 91–98.

(9)   Lyons, B.; Jamie, J.; Truscott, R. J. W. *Int J Pept Res Ther* **2011**, *17*, 131-135.

(10) Su, S. P.; Lyons, B.; Friedrich, M.; McArthur, J. D.; Song, X.; Xavier, D.; Truscott, R. J.; Aquilina, J. A. *Aging cell* **2012**, *11*, 1125-1127.

(11) Steinberg, S. M.; Masters, P. M.; Bada, J. L. *Bioorg Chem* **1984**, *12*, 349-355.

(12) Bada, J. L.; Shou, M. Y.; Man, E. H.; Schroeder, R. A. *Earth Planet Sc Lett* **1978**, *41*, 67-76.

(13) Demarchi, B.; Williams, M. G.; Milner, N.; Russell, N.; Bailey, G.; Penkman, K. *Quatern Int* **2011**, *239*, 114-124.

(14) Penkman, K. E. H.; Kaufman, D. S.; Maddy, D.; Collins, M. J. *Quat Geochronol* **2008**, *3*, 2-25.

(15) Takahashi, O.; Kohayashi, K.; Oda, A. *Chem Biodivers* **2010**, *7*, 1625-1629.

(16) Mann, K.; Siedler, F. *Bba-Proteins Proteom* **2004**, *1696*, 41-50.

(17) Miller, G. H.; Beaumont, P. B.; Deacon, H. J.; Brooks, A. S.; Hare, P. E.; Jull, A. J. T. *Quaternary Sci Rev* **1999**, *18*, 1537-1548.

(18) Kaufman, D. S.; Manley, W. F. *Quaternary Sci Rev (Quat Geochron)* **1998**, *17*, 987-1000.

(19) Lugg, J. W. H. *Biochem J* **1938**, *32*, 775-783.

(20) Hill, R. L. *Advances in Protein Chemistry* **1965**, *20*, 37-107.

(21) Brennan, T. V.; Clarke, *S. Protein Sci* **1993**, *2*, 331-338.

(22) Capasso, S.; Mazzarella, L.; Sica, F.; Zagari, A.; Salvadori, S. *J Chem Soc Perk T 2* **1993**, 679-682.

(23) Geiger, T.; Clarke, S. *J Biol Chem* **1987**, *262*, 785-794.

(24) Crisp, M.; Demarchi, B.; Collins, M.; Morgan-Williams, M.; Pilgrim, E.; Penkman, K. *Quat Geochronol*, in press, DOI: http://dx.doi.org/10.1016/j.quageo.2012.09.002.

(25) Demarchi, B.; Collins, M. J.; Tomiak, P. J.; Davies, B. J.; Penkman, K. E. H. *Quat Geochronol*, in press, DOI: http://dx.doi.org/10.1016/j.quageo.2012.08.001.

Artificial diagenesis
WNSVWAW

Tracking
breakdown
in real time

For Table of Contents Only