Crystal structure and silica condensing activities of silicatein α–cathepsin L chimeras†‡

Michael Fairhead,§[a] Kenneth A. Johnson,§[b] Thomas Kowatz,§[b] Stephen A. McMahon,[b] Lester G. Carter,[b] Muse Oke,[b] Huanting Liu,[b] James H. Naismith*[b] and Christopher F. van der Walle*[a]

Received (in Cambridge, UK) 28th November 2007, Accepted 25th January 2008
First published as an Advance Article on the web 11th February 2008
DOI: 10.1039/b718264c

Cathepsin L mutants with the ability to condense silica from solution have been generated and a 1.5 Å crystal structure of one of these chimeras allows us to rationalise the catalytic mechanism of silicic acid condensation.

Silica materials are of interest in biotechnology and drug delivery[1] but controllable silicate formation is very difficult. Yet a variety of sponge species make highly ordered specific glass structures called spicules.[2–5] Harnessing the processes that underlie this biosynthesis has considerable application. The enzyme silicatein α forms part of the organic filament found in spicules which *in situ* condenses silicate (Fig. 1a), although the exact form of the natural substrate is not known.[5] Both wild type and recombinant silicatein α have been shown to catalyse the condensation of siloxanes such as tetraethoxysilane[4,6] in solution and at surfaces.[7] Silicatein α has also been shown to have the ability to cause the deposition of other compounds at surfaces, including titanium phosphate, titanium oxide, zirconium oxide and l-lactide.[8,9]

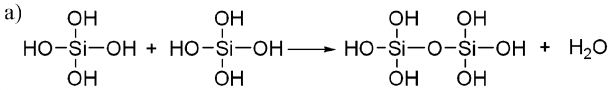
Above 100 ppm, condensation of silicic acid occurs spontaneously, presumably as the concentration of nucleophilic ionised molecules is high enough. There are two possible mechanisms for enzyme catalysis: (i) stabilise at the active site one molecule of deprotonated silicic acid (the nucleophile) which will then react with another molecule of silicic acid; or (ii) stabilise a protonated silicic acid (the electrophile) which will then react with another molecule of silicic acid. Neither the wild type nor recombinant silicatein α is amenable to biophysical study due to low levels of protein expression and inclusion body formation when recombinantly expressed in *E. coli*.[4,7,8]

Sequence comparison identifies a significant degree of similarity between silicatein α and the human cysteine protease cathepsin L[4] with an overall 52% identity and 65% similarity. The most notable differences between the two sequences are the presence of a loop of four amino acids in cathepsin L [10] that is absent in silicatein α, a large number of hydroxyl containing residues (serines, threonines and tyrosines) in silicatein α that are not present in cathepsin L and the substitution of the catalytic cysteine (C25) in cathepsin L [11] for serine (S25) in silicatein α. This serine residue has been shown by mutagenesis to be essential for the function of silicatein α.[12] The other residues in the catalytic triad of cathepsin L (H163 and N187) are similarly conserved in silicatein α. A mechanism has been proposed for polymerisation of $Si(OEt)_4$ in which a covalent protein silicate intermediate is formed; this active electrophile is then decomposed by water yielding a more active nucleophile.[12] Whether this mechanism operates *in vivo* is unresolved. Interestingly, there is also a change in the flanking residues either side of the catalytic serine and histidine and this difference is conserved across several different silicateins (Fig. 1b).

We have made a series of cathepsin L mutants that increasingly match the sequence features that are unique to silicatein α (Table 1). Mutants were made using a combination of the Quikchange™ mutagenesis technique and conventional PCR. The mutant proteins were then recombinantly expressed as the pro-protein and purified from *Pichia pastoris* as described previously.[13] Mature cathepsin L–silicatein α chimeras were then obtained *via* published procedures.[14]

The various mutant proteins were assayed for silica condensation activity (detected by precipitation of silicate) using

a)



b)

H. sapiens	^{21}QGQCGS**C**WAF^{28}	^{157}SSEDMD**H**GVL^{166}
T. aurantia	^{21}QGDCGA**S**YAF^{28}	^{158}SSSSLN**H**AMV^{167}
S. domuncula	^{21}QGDCGA**S**YAF^{28}	^{159}SSTSLN**H**AMV^{168}
H. perleve	^{21}QGDCGA**S**YAF^{28}	^{158}SSSSLN**H**AMV^{167}
L. baicalensi	^{17}QGQCGA**S**YAF^{24}	^{154}SSTKLN**H**AML^{163}

Fig. 1 (a) The chemical reaction catalysed by the organic filament of the sponge spicule. (b) Conservation of residues flanking catalytic serine and histidine in silicatein-α from various sponge species. Shown is the amino acid sequence around the catalytic serine/cysteine and histidine in human cathepsin L (UniProtKB/TrEMBL entry P07711) and silicatein α from: *Tethya aurantia* (O76238); *Suberites domuncula* (Q2MEV3); *Hymeniacidon perleve* (Q2HYF6); *Lubomirskia baicalensi* (Q2PC18).

[a] *SIPBS, University of Strathclyde, 27 Taylor Street, Glasgow, Scotland, UK. E-mail: chris.walle@strath.ac.uk*
[b] *The Centre for Biomolecular Science and The Scottish Structural Proteomics Facility, University of St. Andrews, North Haugh, St. Andrews, Scotland, UK. E-mail: naismith@st-andrews.ac.uk*

† Electronic Supplementary Information (ESI) available: Molecular biology, protein purification and crystallographic details.

‡ The final crystal structure and experimental have been deposited with the Protein Data Bank, entry 2VHS. Links to PDB visualisations are included in the online supplementary data.

§ These authors contributed equally.

Table 1 Mutant constructs made

Cathepsin L construct[a]	Mutations	Match to silicatein α
C2S	C25S	Catalytic serine
AS2	S24A W26Y	Residues flanking catalytic serine
AS2H	M161L, D162N, G164A, V165M	Residues flanking catalytic histidine
LOOP	^{173}ESTESDNN^{180} to ISNNQ	To replicate loop
2SER	E159S, D160S	Match serines
4SER	E153S, P154S	Match serines

[a] The mutations are additive from top to bottom, thus 4SER has all the preceding mutations.

water glass as a substrate (Fig. 1a). Mutation of the catalytic cysteine to serine (the C25S mutant) did not confer significant levels of condensing activity. The C25S mutant would have been expected to be sufficient from the predicted mechanism of silicatein α.[6] In order to obtain significant condensing activity, it was found to be necessary to mutate the residues either side of the catalytic S25 residue (Table 1, Fig. 2) before any significant activity was observed. Additional mutations to cathepsin L to further increase its resemblance to silicatein α did not significantly increase the ability to condense silica; in fact many of these additional mutations tended to decrease condensation activity. It is likely, therefore, that these residues play other functions in the sponge spicule such as in the association of silicatein α with silicatein β and galectin,[15,16] or in the templating of silica into specific morphologies, rather than underlie catalysis *per se*. While the various active mutant proteins were found to readily precipitate silica from the natural sodium silicate substrate, we could not detect any precipitate from tetraethoxysilane substrate, which can be utilised by silicatein-α.[4,6]

In an effort to understand the catalytic basis of silica condensation we have solved the X-ray crystal structure of the 4SER mutant (Fig. 3a). Crystals of the 4SER mutant were

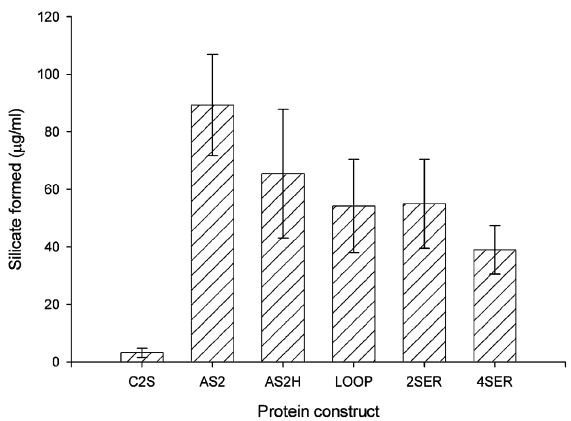


Fig. 2 Silica condensing activity of mutant proteins. To 0.1 mg mL^{-1} of protein solution in 0.1 M sodium phosphate buffer pH 7 was added 4.5 mM sodium silicate. Samples were incubated at room temperature for 24 h; precipitated silica was collected by centrifugation and then assayed using the Merck spectroquant silicon assay kit as described.[17] Shown is the average value of five separate incubations with the error as shown.

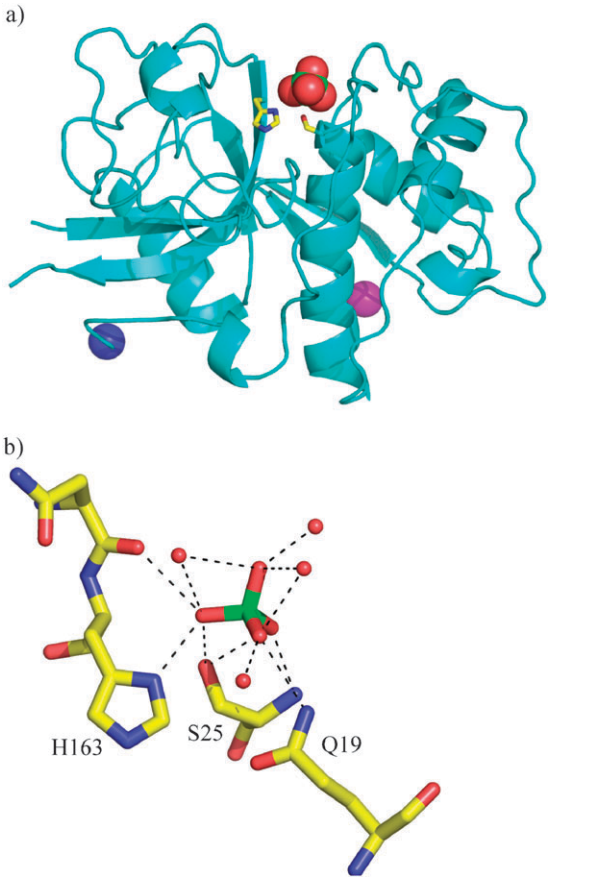


Fig. 3 (a) The structure of the 4SER chimera shown in cartoon form. The sulfate at the active site is shown as space filling spheres. The two 'catalytic' residues H163 and S25 are shown as sticks. Carbon atoms are coloured yellow, oxygen red, sulfur green and nitrogen blue. (b) Close up view of the active site of the 4SER chimera showing the close contacts (<3.5 Å) as dashed lines. The sulfate molecule is shown as sticks; the colour scheme is the same as in part (a). The extensive interactions mediated by water would allow extensive proton shuttling. Figures produced with PYMOL.[18]

obtained by hanging drop vapour diffusions. The structure was determined to 1.5 Å using cathepsin L (PDB 1MHW) as a model for molecular replacement.[19,20] Unsurprisingly, the 4SER chimera is identical to cathepsin L with a RMSD of only 0.5 Å between the two structures. Full experimental details are given in the ESI.†

Fig. 3b shows the active site architecture of the 4SER mutant with that of the C25S pro-cathepsin L crystal structure[21] (PDB 1CJL). The C25S mutant of cathepsin is itself inactive as a protease, indicating the nucleophilicity of the OG atom of S25 is low. In 4SER the distance between the OG atom on S25 and ND1 on H163 has increased to 3.6 Å (beyond hydrogen bonding distance), further decreasing the nucleophilicity of S25. Mutation of the flanking residues perturbs a cluster of residues that sit behind S25 (Fig. 3). The mutations remove a hydrogen bond (S24A) and decrease van der Waal interaction (W26Y). The mutations are both to residues smaller than the native enzyme. The net result would be to allow the S25 to move away from the H163 and to enlarge the volume at the active site.

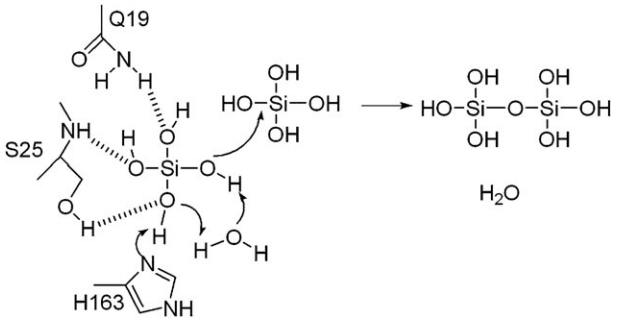


Fig. 4 The chemical mechanism proposed to operate for the polymerisation of silicic acid. We propose H163 deprotonates $Si(OH)_4$; the extensive water network could permit proton shuttling. We have illustrated one possibility.

An understanding of the mechanism of silica condensation by the chimera is aided by the presence of a sulfate ion at the active site. (A second sulfate is at a crystal contact 17 Å distant from S25). The crystal structure reveals it would be difficult to fit the bulky $Si(OEt)_4$ at the active site without fairly significant conformational adjustment. This would seem to explain the lack of activity of the chimera against $Si(OEt)_4$ and would suggest that silicatein α must have a larger active site than 4SER which would allow it to accommodate the bulky $Si(OEt)_4$. We believe the tetrahedral sulfate ion at the active site is a good mimic of the tetrahedral silicic acid. The sulfate is hydrogen bonded to S25. This is consistent with S25 making a nucleophilic attack on $Si(OH)_4$ to generate a covalent enzyme intermediate. However, such a high energy intermediate would seem chemically unlikely and prone to immediate hydrolysis back to $Si(OH)_4$.

The sulfate is hydrogen bonded to H163, the residue that activates the protein nucleophile in proteases. We favour a mechanism in which H163 binds and stabilises the deprotonated form of $Si(OH)_4$ at the active site. The extensive network of water molecules and hydrogen bonds (Fig. 3b) would permit significant proton shuttling, such that the negative charge may reside on a different oxygen. This deprotonated enzyme bound species is a sufficient nucleophile that it will attack $Si(OH)_4$ solution species initiating polymerisation (Fig. 4). Loss of a $Si(OH)_4$ hydroxyl group carried out by protonation and elimination of water could also promote the reaction but we have no evidence for an acid. It is also unlikely that at pH 7 $Si(OH)_4$ will be a sufficient base to first deprotonate H163 (whose pK_a is likely to be lower than 7 due to the presence of D187) and then for the hydroxyl group of the second $Si(OH)_4$ to perform a nucleophilic attack.

In our proposed mechanism, the roles of C25S and flanking mutations are simply to create a sufficiently sized pocket that will allow recognition of the tetrahedral $Si(OH)_4$ molecule in such a way that H163 can deprotonate it. The deprotonated $Si(OH)_4$ protein complex can be thought of as a template for condensation reaction. In this proposal there is no need to involve a high energy covalent intermediate. The presence of a specific $Si(OH)_4$ transporter in sponge[22] suggests the true substrate *in vivo* is indeed silicic acid, not high energy silicon alkoxides. This being so, the simple acid base activation mechanism we propose seems a good model for the biological process.

This work was supported by the Leverhulme Trust, grant F/00 273/H. The Scottish Structural Proteomics Facility is funded by the Scottish Funding Council and the Biotechnology and Biological Sciences Research Council.

Notes and references

1 S. Giri, B. G. Trewyn and V. S. Lin, *Nanomedicine*, 2007, **2**, 99–111.
2 W. E. Muller, S. I. Belikov, W. Tremel, C. C. Perry, W. W. Gieskes, A. Boreiko and H. C. Schroder, *Micron*, 2006, **37**, 107–120.
3 H. C. Schroder, O. Boreiko, A. Krasko, A. Reiber, H. Schwertner and W. E. Muller, *J. Biomed. Mater. Res., Part B*, 2005, **75**, 387–392.
4 K. Shimizu, J. Cha, G. D. Stucky and D. E. Morse, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 6234–6238.
5 J. C. Weaver and D. E. Morse, *Microsc. Res. Tech.*, 2003, **62**, 356–367.
6 J. N. Cha, K. Shimizu, Y. Zhou, S. C. Christiansen, B. F. Chmelka, G. D. Stucky and D. E. Morse, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 361–365.
7 M. N. Tahir, P. Theato, W. E. Muller, H. C. Schroder, A. Janshoff, J. Zhang, J. Huth and W. Tremel, *Chem. Commun.*, 2004, 2848–2849.
8 P. Curnow, D. Kisailus and D. E. Morse, *Angew. Chem., Int. Ed.*, 2006, **45**, 613–616.
9 M. N. Tahir, P. Theato, W. E. Muller, H. C. Schroder, A. Borejko, S. Faiss, A. Janshoff, J. Huth and W. Tremel, *Chem. Commun.*, 2005, 5533–5535.
10 S. M. Smith and M. M. Gottesman, *J. Biol. Chem.*, 1989, **264**, 20487–20495.
11 R. W. Mason, G. D. Green and A. J. Barrett, *Biochem. J.*, 1985, **226**, 233–241.
12 Y. Zhou, K. Shimizu, J. N. Cha, G. D. Stucky and D. E. Morse, *Angew. Chem., Int. Ed.*, 1999, **38**, 780–782.
13 M. Fairhead and C. F. van der Walle, *Protein Pept. Lett.*, 2008, **15**, 47–53.
14 R. Menard, E. Carmona, S. Takebe, E. Dufour, C. Plouffe, P. Mason and J. S. Mort, *J. Biol. Chem.*, 1998, **273**, 4478–4484.
15 P. Curnow, P. H. Bessette, D. Kisailus, M. M. Murr, P. S. Daugherty and D. E. Morse, *J. Am. Chem. Soc.*, 2005, **127**, 15749–15755.
16 H. C. Schroder, A. Boreiko, M. Korzhev, M. N. Tahir, W. Tremel, C. Eckert, H. Ushijima, I. M. Muller and W. E. Muller, *J. Biol. Chem.*, 2006, **281**, 12001–12009.
17 W. E. Muller, A. Krasko, G. Le Pennec and H. C. Schroder, *Microsc. Res. Tech.*, 2003, **62**, 368–377.
18 W. L. DeLano, DeLano Scientific, San Carlos, CA, USA, 2002.
19 A. J. McCoy, R. W. Grosse-Kunstleve, L. C. Storoni and R. J. Read, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2005, **61**, 458–464.
20 L. C. Storoni, A. J. McCoy and R. J. Read, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2004, **60**, 432–438.
21 R. Coulombe, P. Grochulski, J. Sivaraman, R. Menard, J. S. Mort and M. Cygler, *EMBO J.*, 1996, **15**, 5492–5503.
22 H. C. Schroder, S. Perovic-Ottstadt, M. Rothenberger, M. Wiens, H. Schwertner, R. Batel, M. Korzhev, I. M. Muller and W. E. G. Muller, *Biochem. J.*, 2004, **381**, 665–673.