Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Structure of the newly found green turtle egg-white ribonuclease

Somporn Katekaew,[a] Buabarn Kuaprasert,[b] Takao Torikata,[c] Yoshimitsu Kakuta,[d] Makoto Kimura,[d] Kazunari Yoneda[c] and Tomohiro Araki[c]*

[a]Department of Biochemistry, Faculty of Science, Khon Kaen University, Khon Kaen 42000, Thailand, [b]Research and Academic Divisions, Synchrotron Light Research Institute, Nakhon Ratchasima 30000, Thailand, [c]Department of Bioscience, School of Agriculture, Tokai University, Aso, Kumamoto 869-1404, Japan, and [d]Graduate School of Bioresource and Bioenvironmental Science, Kyushu University, Hakozaki 6-10-1, Higashi-ku, Fukuoka 812-8581, Japan

Correspondence e-mail: araki@agri.u-tokai.ac.jp

Received 2 March 2010
Accepted 7 May 2010

PDB Reference: green turtle egg-white ribonuclease, 2zpo.

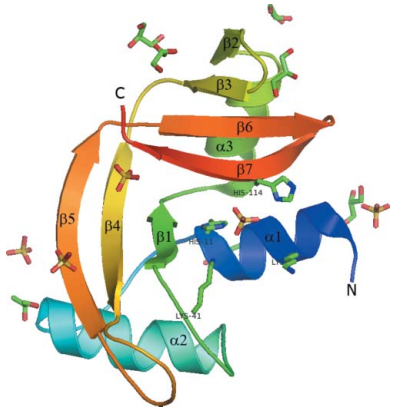
Marine green turtle (*Chelonia mydas*) egg-white ribonuclease (GTRNase) was crystallized from 1.1 *M* ammonium sulfate pH 5.5 and 30% glycerol using the sitting-drop vapour-diffusion method. The structure of GTRNase has been solved at 1.60 Å resolution by the molecular-replacement technique using a model based on the structure of RNase 5 (murine angiogenin) from *Mus musculus* (46% identity). The crystal belonged to the monoclinic space group *C*2, with unit-cell parameters $a = 86.271$, $b = 34.174$, $c = 39.738$ Å, $\alpha = 90$, $\beta = 102$, $\gamma = 90^\circ$. GTRNase consists of three helices and seven β-strands and displays the $\alpha+\beta$ folding topology typical of a member of the RNase A superfamily. Superposition of the C^α coordinates of GTRNase and RNase A superfamily members indicates that the overall structure is highly similar to that of angiogenin or RNase 5 from *M. musculus* (PDB code 2bwl) and RNase A from *Bos taurus* (PDB code 2blz), with root-mean-square deviations of 3.9 and 2.0 Å, respectively. The catalytic residues are conserved with respect to the RNase A superfamily. The three disulfide bridges observed in the reptilian enzymes are conserved in GTRNase, while one further disulfide bond is required for the structural stability of mammalian RNases. GTRNase is expressed in egg white and the fact that its sequence has the highest similarity to that of snapping turtle pancreatic RNase suggests that the GTRNase secreted from oviduct cells to form egg white is probably the product of the same gene as activated in pancreatic cells.

1. Introduction

Bovine pancreatic ribonuclease A (RNase A; EC 3.1.27.5) is one of the most extensively investigated proteins. It was the subject of many pioneering studies in protein chemistry and enzymology involving chemical modification, refolding, amino-acid sequence determination, limited proteolysis, X-ray crystallography and peptide synthesis, among others. It continues to be the focus of considerable attention and a vast quantity of kinetic, mechanistic, thermodynamic and structural information has now been reported (for reviews, see Richards & Wyckoff, 1971; Blackburn & Moore, 1982; Eftink & Biltonen, 1987; Raines, 1998).

The RNase A superfamily is a group of homologous proteins that are present in considerable quantities in the pancreas of a number of mammalian taxa and a few reptiles (Barnard, 1969; Beintema *et al.*, 1973). An extensive study has been conducted on the molecular evolution of pancreatic RNases in many mammalian species (Beintema & Lenstra, 1982), but did not include those of reptiles. However, reasonable quantities of reptile RNases have only been found in turtle (Barnard, 1969) and iguana pancreas (Beintema *et al.*, 1988).

In reptiles, RNase is present in pancreatic tissue and also in eggs. There is a high accumulation of an RNase with a distinct catalytic activity in the egg white of the marine green turtle (*Chelonia mydas*). This green turtle egg-white RNase (GTRNase) was first purified and sequenced by our group (Katekaew *et al.*, 2006). Two isoforms of the enzyme, both composed of 119 residues, have been analyzed and have almost identical molecular weights (13 kDa) and catalytic activities. There is only one amino-acid difference between the two isoforms



© 2010 International Union of Crystallography
All rights reserved

(Ser37 and Leu37). Multiple sequence alignment of GTRNase and related RNase enzymes indicates that GTRNase is very similar to pancreatic RNase from snapping turtle (*Chelydra serpentine*). The amino-acid residues involved in the active site are conserved, including His11, Lys41 and His114 (Fig. 1), suggesting that GTRNase belongs to the RNase A superfamily (Katekaew *et al.*, 2006).

The GTRNase is specific for poly-(C) and, with a lower rate, degraded poly-(U). The effects of pH and temperature on RNase activity against poly-(C) are similar to those on RNase A. However, GTRNase is unlike RNase A in its optimum pH, broad range of optimum temperature and pH stability. The differences found in catalytic activity might be explained by changes in the structure (Katekaew *et al.*, 2007). The structure of GTRNase has not yet been characterized, although a model of GTRNase (MD-GTRNase) has been created based on RNase 5 from *Mus musculus* to evaluate the enzymatic properties of GTRNase (Katekaew *et al.*, 2007).

To obtain structural information on GTRNase, we crystallized GTRNase and collected X-ray data. The molecular-replacement method was used to solve the structure, with MD-GTRNase as a search model. In this study, the crystal structure of GTRNase, a newly found RNase from egg white, is described and structural comparisons based on the active sites of GTRNase and RNases are discussed.

2. Materials and methods

A sample of GTRNase was prepared using methods that have been described previously (Katekaew *et al.*, 2006). All reagents and chemicals were of analytical grade and commercially available.

2.1. Crystallization, data collection and processing

Crystallization trials of GTRNase were performed using the sitting-drop vapour-diffusion method at 293 K with the Crystal Screen I (Hampton Research) crystallization screening kit. Individual crystallization drops containing equal volumes (3 µl) of reservoir solution and 20 mg ml^{-1} GTRNase solution were suspended over 0.5 ml reservoir solution. GTRNase crystals that were suitable for X-ray diffraction were grown in 1.1 *M* ammonium sulfate pH 5.5 and 30% glycerol within one week. The concentration of GTRNase was estimated using the extinction coefficient of RNase A at A_{280} ($E^{1\%}$ = 7 ml mg^{-1} cm^{-1}; Pace *et al.*, 1995). The crystal was cryopro-

Table 1
Data-collection details and crystallographic refinement statistics.

Values in parentheses are for the highest resolution shell.

Data-collection details	
X-ray source	SPring-8 (BL38B1)
Wavelength (Å)	1.0
Resolution (Å)	50.00–1.60 (1.66–1.60)
Space group	*C*2
Unit-cell parameters	
a (Å)	86.271
b (Å)	34.174
c (Å)	39.738
α (°)	90
β (°)	102
γ (°)	90
Observed reflections	14850
Unique reflections	14099
$I/\sigma(I)$	4.4 (10.7)
Completeness (%)	97.9 (86.1)
$R_{\rm merge}$† (%)	8.5 (37.0)
Refinement statistics	
$R_{\rm cryst}$‡ (%)	18.3
$R_{\rm free}$§ (%)	24.7
No. of residues	119
No. of sulfate ions	5
No. of glycerol molecules	7
No. of water molecules	152
R.m.s.d.	
Bond lengths (Å)	0.013
Bond angles (°)	1.513
Average *B* factor (Å^2)	
Protein (all atoms)	22.38
Validation (% of all residues)	
Favoured	97.4
Allowed	100.0
Disallowed	0.0

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$. ‡ $R_{\rm cryst} = \sum_{hkl} ||F_{\rm obs}| - |F_{\rm calc}|| / \sum_{hkl} |F_{\rm obs}|$, where $F_{\rm obs}$ and $F_{\rm calc}$ are observed and calculated structure-factor amplitudes, respectively. § The $R_{\rm free}$ value was calculated from 5% of all data that were not used in refinement.

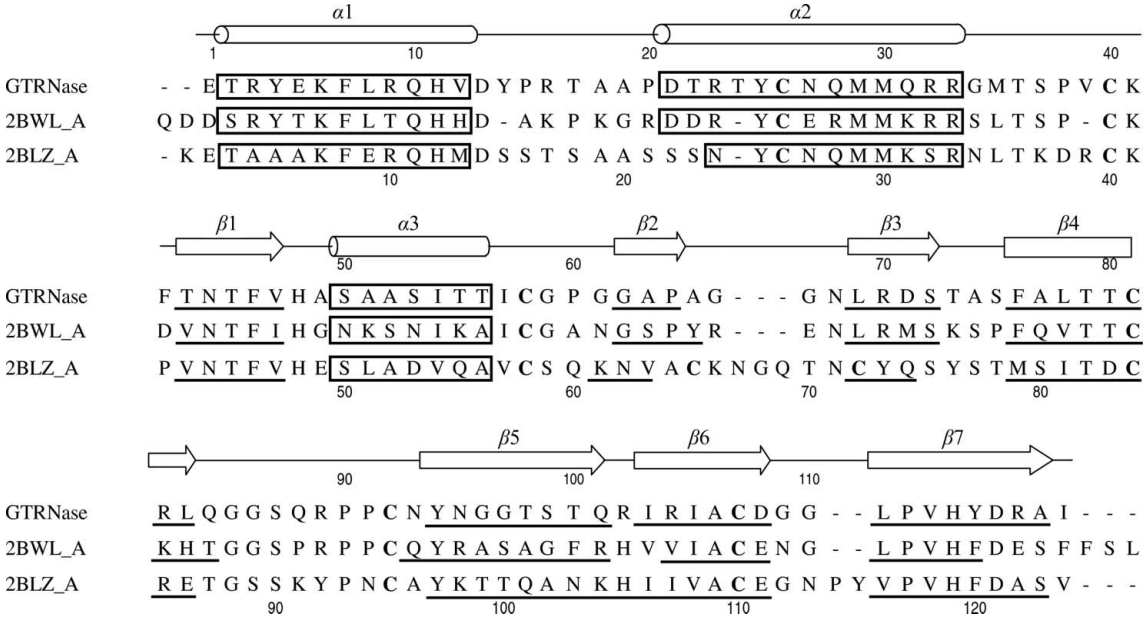


Figure 1
Alignment of the amino-acid sequences of GTRNase, RNase 5 (angiogenin; 2bwl, chain *A*) and RNase A (2blz, chain *A*). The *α*-helices and *β*-strands of GTRNase are displayed as rods and arrows, respectively. The topologically equivalent secondary-structure elements are boxed (*α*-helices) and underlined (*β*-strands). The upper numbers correspond to GTRNase and the lower numbers correspond to RNase A.

tected using crystallization solution and flash-frozen. X-ray data were collected in a nitrogen stream using a MAR CCD detector on BL38B1 at SPring-8. The crystal diffracted X-rays to beyond 1.60 Å resolution. Analysis of the diffraction data using the *HKL*-2000 package (Otwinowski & Minor, 1997) indicated that the crystal belonged to the monoclinic space group *C*2, with unit-cell parameters $a = 86.271$, $b = 34.174$, $c = 39.738$ Å, $\alpha = 90$, $\beta = 102$, $\gamma = 90^\circ$. A total of 14 099 unique reflections were processed with 97.9% completeness. The relative molecular volume, V_M, was 2.21 Å^3 Da^{-1} and was consistent with the presence of one molecule of protein in the asymmetric unit and a solvent content of 44%.

2.2. Structure determination and refinement

The crystal structure of GTRNase was solved by the molecular-replacement method. RNase 5 from *Mus musculus* (PDB code 2bwl; Holloway *et al.*, 2005) was used as a model to provide an initial solution using *MOLREP* (Vagin & Teplyakov, 1997). Crystallographic refinement and structure building were performed using *REFMAC*5 (Murshudov *et al.*, 1997) and *Coot* (Emsley & Cowtan, 2004). *PROCHECK* (Laskowski *et al.*, 1993) was used to evaluate the stereochemistry of the crystal structure. The X-ray diffraction data-collection and refinement statistics are summarized in Table 1. The atomic coordinates of the GTRNase from *Chelonia mydas* as determined in this study have been deposited in the RSCB Protein Data Bank as entry 2zpo. The superposition of GTRNase (2zpo) with RNase 5 from *M. musculus* (PDB code 2bwl) and RNase A from *Bos taurus* (PDB code 2blz; Nanao *et al.*, 2005) was created by *SUPERPOSE* from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994). Structural figures were produced using *PyMOL* (DeLano, 2002). Multiple sequence alignment was performed using *ClustalW* (Larkin *et al.*, 2007).

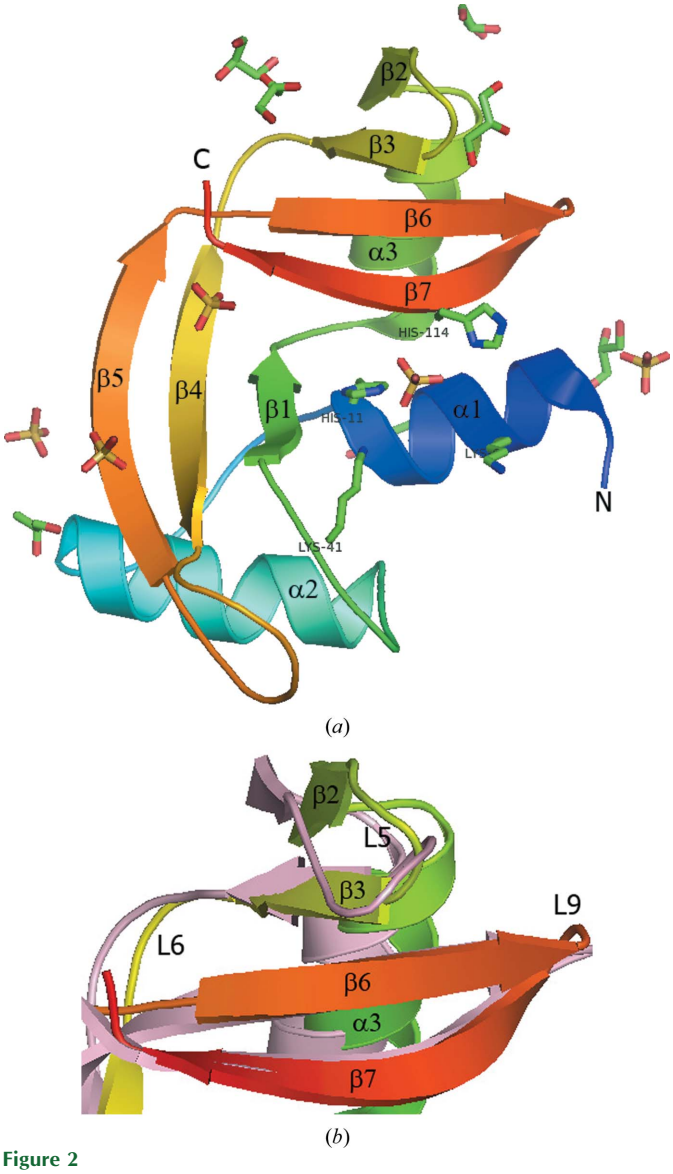


Figure 2
(*a*) Overview of the GTRNase structure complexed with five sulfate ions and seven glycerol molecules. One of the sulfate ions is located in the active site. α, α-helix, β, β-strand, L, loop. Highlighted residues are involved in the conserved catalytic binding site. (*b*) Superimposition of GRTNase (2zpo; rainbow) and RNase A (2blz; light pink) indicating the shorter loops in 2zpo (L5 and L9), in which the amino-acid deletion sites are located, compared with 2blz.

3. Results and discussion

3.1. Overall structure

The GTRNase crystal structure in space group *C*2 was solved at 1.60 Å resolution. After refinement, the R and R_{free} values were 18.3% and 24.7%, respectively. The structure displays an $\alpha+\beta$-type polypeptide chain that folds into a V-shape with the active-site cleft in the middle. The catalytic residues that are involved in the active site, including His11, Lys41 and His114, are conserved as highlighted in Fig. 2(*a*). Five sulfate ions and seven glycerol molecules were found in the structure; one of the sulfate ions is bound in the active site. Three disulfide bridges have been identified (Cys26–Cys81, Cys40–Cys92 and Cys58–Cys107).

The overall structure encompasses three α-helices (α1–α3) in the N-terminal part, seven β-strands (β1–β7) and nine linking loops (L1–L9). The first β-strand is packed between helices α2 and α3 and the others are after helix α3. The V-shape of the structure is formed by a sharp turn in the orientation of two groups of β-strands. One side of the V-shape is composed of β1, β4 and β5, where β4 is located antiparallel to the other strands. The other side consists of the anti-

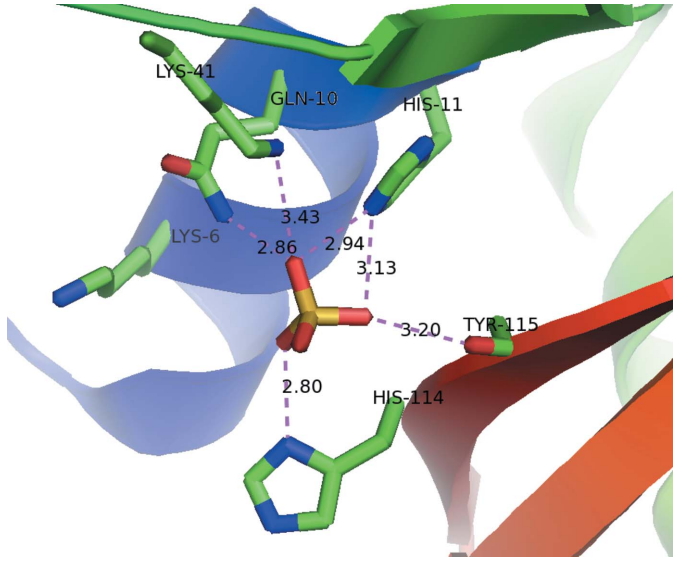


Figure 3
Coordination of a sulfate ion at the active site of GTRNase A. Tyr115 forms a hydrogen bond to the sulfate using the main-chain carbonyl group. Magenta dotted lines represent donor–acceptor interactions; bond distances are indicated.

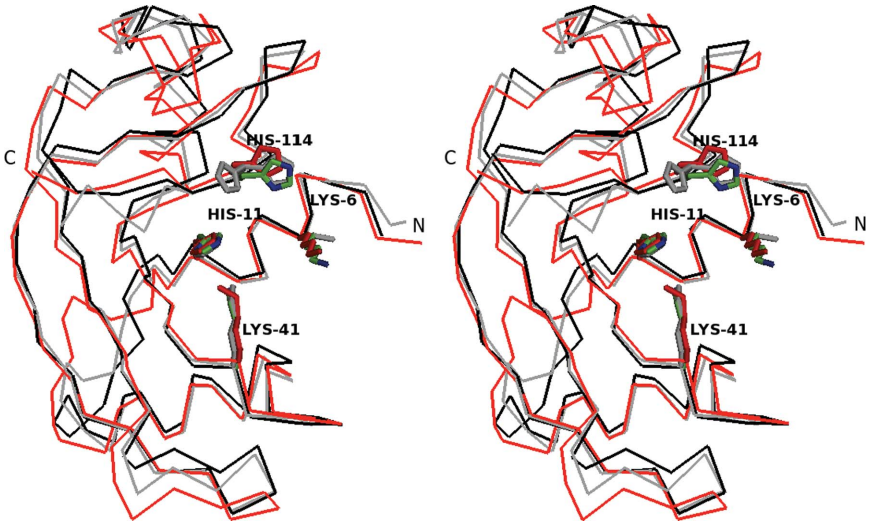


Figure 4
Cross-eyed stereoview of a C^{α} superposition of GTRNase (2zpo; black), RNase A from *B. taurus* (2blz, chain *A*; red) and murine angiogenin from *M. musculus* (2bwl, chain *A*; grey). Highlighted residue numbers are related to the GTRNase sequence. His113 of angiogenin adopts two conformations, one of which is located in the same position as His114 in GTRNase.

parallel strands $\beta2/\beta3$ and $\beta6/\beta7$. The side chains of Lys6, His11, Lys41 and His114 are conserved in the active site compared with the known structures of angiogenin and RNase A. Fig. 3 shows the coordination of a sulfate ion in the active site of GTRNase. The sulfate participates in hydrogen bonds to the N atoms of Lys6 and His11 from $\alpha1$, Lys41 from L2 and His114 from $\beta7$ (Fig. 2*a*). The average *B* factor of the bound sulfate ion of 31.8 Å^2 is higher than the mean *B* value of the overall structure of 22.4 Å^2, suggesting that the sulfate is not a tight-binding ligand. However, four further sulfate ions and seven glycerol molecules are bound on the molecular surface, suggesting that they may play a crucial role in molecular packing during the protein crystallization process.

3.2. Structural comparison of GTRNase with angiogenin and RNase A

The GTRNase structure displays the fold common to all RNases, including angiogenin or RNase 5 from *M. musculus* (PDB code 2bwl; Holloway *et al.*, 2005) and RNase A from *B. taurus* (PDB code 2blz; Nanao *et al.*, 2005). Superposition of the C^{α} atoms of GTRNase on those of the other two structures results in root-mean-square (r.m.s.) deviations of 3.9 and 2.0 Å, respectively. The catalytic site architectures of the three enzymes are highly similar. Although His113 of the angiogenin structure has two conformations, one of them is placed in the same orientation as His114 of the GTRNase.

There are two amino-acid deletion sites in GTRNase (in L5 and L9) and angiogenin compared with RNase A. The superposition of GTRNase and RNase A shown in Fig. 2(*b*) shows the locations of the two shorter loops (L5 and L9). L5 is a linking loop between $\beta2$ and $\beta3$. $\beta7$ contains the catalytic residue His114 and connects to $\beta6$ using a shorter loop L9. However, superposition of the C^{α} positions of GTRNase, angiogenin and RNase A shows a small difference between His114 of GTRNase and His119 of RNase A compared with one of the conformations of His113 in angiogenin (Fig. 4).

The structure of GTRNase is stabilized by three disulfide bonds: Cys26–Cys81 bridges helix $\alpha2$ and strand $\beta4$, Cys40–Cys92 connects loops L2 and L7 and Cys58–Cys107 links loop L4 and strand $\beta6$. A disulfide bridge (Cys65–Cys72) that is conserved in mammalian RNases is absent in GTRNase. In GTRNase this site is replaced by Ala65 and Leu72 on L5 and is continued by $\beta3$ (Fig. 2*b*). These residues are not close to the active site; the nearest catalytic residue is His114 in $\beta7$. However, in mammalian RNase A the additional disulfide bond is likely to contribute to the proper alignment of residues (such as Lys41 and Lys66) that are necessary for the catalytic efficiency of the RNA cleavage (Klink *et al.*, 2000). In addition, the average *B* factor of the catalytic residue His114 of GTRNase is 18.6 Å^2, compared with the relative average *B* factor of the structure of 22.4 Å^2, suggesting that His114 is fixed in this position and that the lack of a disulfide bond does not affect the GTRNase active-site architecture. The minor differences in activity between GTRNase and RNase A may be caused by changes in the micro-environment owing to the deletion and substitution of residues (Katekaew *et al.*, 2007). These properties might be expected to be similar in all reptile and related RNases which display high similarity in sequence and by implication in structure to GTRNase.

The GTRNase structure has been observed to be homologous to those of enzymes of the RNase superfamily and its sequence has the highest identity to snapping turtle pancreatic RNase (84%; Katekaew *et al.*, 2006). The structural features that GTRNase, snapping turtle RNase and angiogenin have in common are deletions in the external loops near residues 69 and 115 and the absence of the disulfide bond linking residues 65 and 72. The differences between angiogenin, RNase A and turtle RNases indicate that a gene duplication leading to separate genes for RNase and angiogenin may have occurred around 300 million years ago when reptiles and mammals diverged. Turtle RNase has possibly diverged from the ancestral sequence somewhat less than other members of the RNase A superfamily (Beintema *et al.*, 1986).

GTRNase is expressed in egg white; however, the sequence was found to have the highest idenitity to snapping turtle pancreatic RNase, suggesting that the RNase secreted from oviduct cells to form egg white is probably the product of the same gene as activated in pancreatic cells.

References

Barnard, E. A. (1969). *Annu. Rev. Biochem.* **38**, 677–732.
Beintema, J. J., Fitch, M. W. & Carsana, A. (1986). *Mol. Biol. Evol.* **3**, 262–275.

Beintema, J. J. & Lenstra, J. A. (1982). *Macromolecular Sequences in Systemic and Evolutionary Biology*, edited by M. Goodman, pp. 43–73. New York: Plenum.
Beintema, J. J., Scheffer, A. J., van Dijk, H., Welling, G. W. & Zwiers, H. (1973). *Nature New. Biol.* **241**, 76–78.
Beintema, J. J., Schüller, C., Irie, M. & Caesana, A. (1988). *Prog. Biophys. Mol. Biol.* **51**, 165–192.
Blackburn, P. & Moore, S. (1982). *The Enzymes*, edited by P. D. Boyer, Vol. 15, pp. 317–433. New York: Academic Press.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
DeLano, W. L. (2002). *PyMOL Molecular Viewer.* http://www.pymol.org.
Eftink, M. E. & Biltonen, R. L. (1987). *Hydrolytic Enzymes*, edited by A. Neuberger & K. Brocklehurst, pp. 333–376. Amsterdam: Elsevier.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Holloway, D. E., Chavali, G. B., Hares, M. C., Subramanian, V. & Acharya, K. R. (2005). *Acta Cryst.* D**61**, 1568–1578.
Katekaew, S., Torikata, T. & Araki, T. (2006). *Protein J.* **25**, 316–327.
Katekaew, S., Torikata, T., Hirakawa, H., Kuhara, S. & Araki, T. (2007). *Protein J.* **26**, 75–87.
Klink, T. A., Woycechowsky, K. J., Taylor, K. M. & Raines, R. T. (2000). *Eur. J. Biochem.* **267**, 566–572.
Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J. & Higgins, D. G. (2007). *Bioinformatics*, **23**, 2947–2948.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Nanao, M. H., Sheldrick, G. M. & Ravelli, R. B. G. (2005). *Acta Cryst.* D**61**, 1227–1237.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pace, C. N., Vajdos, F., Fee, L., Grimsley, G. & Gray, T. (1995). *Protein Sci.* **4**, 2411–2423.
Raines, R. T. (1998). *Chem. Rev.* **98**, 1045–1065.
Richards, F. M. & Wyckoff, H. W. (1971). *The Enzymes*, edited by P. D. Boyer, Vol. 4, pp. 647–806. New York: Academic Press.
Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.