Eur Biophys J (1990) 18:225–233

European Biophysics Journal

Tertiary structure of RNase Pch1 predicted from the model structure of RNase Ms and the crystal structure of RNase T1

Comparison among the model structures – testing the limits of modelling by homology

R. Floegel, P. Zielenkiewicz*, and W. Saenger

Institut für Kristallographie, Freie Universität Berlin, Takustrasse 6, D-1000 Berlin 33, Germany

Received April 19, 1989/Accepted in revised form March 19, 1990

Abstract. In this paper we predict the structure of RNase Pch1 as modelled from the previously predicted structure of RNase Ms and the crystal structure of RNase T1 in the complex with 2′GMP. The predicted structures and their initial energy minimized structural RNase T1 template are compared. The predicted structures of RNase Pch1 show, independent of their prediction form RNase Ms or T1, a higher structural similarity to RNase T1 than to RNase Ms, in agreement with higher sequence similarity and specificity – RNases T1 and Pch1 are specific for guanine whereas RNase Ms is base-unspecific with preference for guanine.

Key words: Ribonuclease T1 (Aspergillus oryzae) – Ribonuclease Ms (Aspergillus saitoi) – Ribonuclease Pch1 (Penicillium chrysogenum) – Structure prediction – Computer graphics – Energy minimization

I. Introduction

Since the primary structure of Ribonuclease (RNase) T1 from *Aspergillus oryzae* was first determined (Takahashi 1985), various other fungal RNases have been isolated and sequenced. Comparisons of their amino acid sequences established that they are similar to RNase T1. A more distant relationship of these eukaryotic RNases with a number of prokaryotic RNases was pointed out by Hartley and Barker (1972), making the RNase T1 family an interesting model for research into molecular evolution as well.

Among the RNases produced by fungi, the complete amino acid sequences of the following guanine-specific enzymes are known: RNase T1 (Takahashi, 1985), N1 (Takahashi 1988), C2 (Bezborodova et al. 1983), Pb1 (Shlyapnikov et al. 1984), F1 (Hirabayashi and Yoshida 1983), Th1 (Polyakov et al. 1987), Pch1 (Shlyapnikov et al. 1986), U1 (Takahashi and Hashimoto 1988) (see Table 1). Fungal RNases with a different specificity are RNase Ms (Watanabe et al. 1982) which is a guanine-preferential, base non-specific RNase and purine-specific RNase U2 (Kanaya and Uchida 1986). Bacterial RNases found to have a relationship to the eukaryotic RNases of the T1-family and of which the complete sequence is known are RNase St (Yoshida et al. 1976) and Sa (Shlyapnikov et al. 1987), both G-specific, and RNase Ba (Hartley and Barker 1972) and Bi (Aphanasyenko et al. 1979), both purine specific. New results, however, indicate that Ba could be G-specific as well (Bielefeld et al. 1988).

Although for various RNases of the T1-family there are reports on successful crystallization attempts dating back to the late sixties and some preliminary X-ray data is available, the first structure determined to good resolution was RNase T1 in a complex with the specific inhibitor 2′GMP (Heinemann and Saenger 1982). It has been further refined to high resolution of 1.9 Å (Arni et al. 1988).

Recent reports on successful crystallization (Nakamura, K., personal communication) of RNase Ms encouraged us to attempt the prediction of the structure of this enzyme based on the known structure of RNase T1 from the T1*2′ GMP crystal complex, using interactive computer graphics and energy minimization techniques (Floegel et al. 1988). RNase T1 provides an excellent template for predicting the sturcture of RNase Ms through modelling by homology, since the structure of T1 is very well defined and the percentage sequence identity to Ms is ~60%. Agreements with experimental results suggested that the predicted structure of RNase Ms is basically correct. The crystal structure of RNase Ms is currently being solved by molecular replacement with the RNase T1 structure (Nakamura, K., personal communication).

In this paper we present the predicted structure of RNase Pch1 – another enzyme of the T1 family from which successful crystallization has been reported (Bezborodova et al. 1988) – and a structural comparison

* On leave from the Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Warsaw, Poland
Offprint requests to: W. Saenger

Table 1. Sequences of Ribonucleases belonging to the RNase T1 family. The line separates the fungal RNases from the RNases of the T1-family produced by bacteria. The alignment is based on those of Hartley (1980)

	10 * 20 * 30 *
T1	A C D Y T C G S N C Y S S - S D V S T - - A Q A A G Y K L H E D G
N1	A C M Y I C G S V C Y S S - S A I S A - - A L N K G Y S Y Y E D G
C2	D C D Y T C G S H C Y S A - S A V S D - - A Q S A G Y Q L E S A G
Ms	E S C E Y T C G S T C Y W S - S D V S A - - A K A K G Y S L Y E S G
F1	E S A T T C G S T N Y S A - S Q V R A - - A A N A A C Q Y Y Q N D
Pb1	A C A A T C G T V C Y T S - S A I S S - - A G A A G Y N L Y S T N
Pch1	A C A A T C G S V C Y T S - S A I S A - - A Q E A G Y D L Y S A N
Th1	D T A T C G K V F Y S A - S A V S A - - A S N A A C N Y V R A G
U1	e G G V S V N C G G T Y Y S S - T Q V N R - - A I N N A K - - - - S G
U2	C D I P Q S T N C G G N V Y S N - D D I N T - - - - A I Q G A L D D V A
St	Q A P C G D T S G F E Q V R L A D L P P E A T D T Y E L I E K G
Sa	D V S G T V C L S A L P P E A T D T L N L I A S D
Ba	A Q V I N T F D G V A D Y L Q T Y H K L P N D Y I T K S E A Q A L G W V A S K G N L A D V A
Bi	A V I N T F D G V A D Y L I R Y K R L P N D Y I T K S Q A S A L G W V A S K G D L A E V A

	40 * 50 * 60 * 70 *
T1	E T V G S N S Y P H K Y N N Y E G F D F S V S S - - - P Y Y - E W P I L S S G D V Y S G
N1	A T A G S S S Y P H R Y N N Y E G F D F P T A K - - - P W Y - E F P I L S S G R V Y T G
C2	Q S V G R S R Y P H Q Y R N Y E G F N F P V S G - - - N Y Y - E W P I L S S G S T Y N G
Ms	D T I - - D D Y P H E Y H D Y E G F D F P V G T - - - S Y Y - E Y P I M S D Y D V Y T G
F1	D S A G S T T Y P H T Y N N Y E G F D F P V D G - - - P Y Q - E F P I K S G G - V Y T G
Pb1	D D V - - S N Y P H E Y H N Y E G F D F P V F G - - - T Y Y - E F P I L K S C K V Y T G
Pch1	D D V - - S N Y P H E Y R N Y E G F D F P V S G - - - T Y Y - E F P I L R S G A V Y S G
Th1	S T A G G S T Y P H V Y N N Y E G F R F K G L S - - - K P F Y E F P I L S S G K T Y T G
U1	Q Y - S S T G Y P H T Y N N Y E G F D F S D Y C - D G P Y K - E Y P L K T S S S G Y T G
U2	N G D R P D N Y P H Q Y Y D - E A S D Q I T L C C G S G P W S E F P L V Y N G P Y Y S S
St	G P Y P Y P E D G T V F E N R E G I L P D C A E - - - G Y Y H E Y T V - - - - K T P S G
Sa	G P F P Y S Q D G V V F Q N R E S V L P T Q S Y - - - G Y Y H E Y T V - - - - I T P G A
Ba	- P G - K S I G G D I F S N R E G K L P G K S G - - - R T W R E A D I - - - - N Y T S G
Bi	- P G - K S I G G D V F S N R E G R L P S A G S - - - R T W R E A D I - - - - N Y V S G

	80 * 90 * 100 *
T1	- - - - G S P G A D R V V F N E N N - Q L A G V I T H T G A - S G N N F V E C T
N1	- - - - G S P G A D R V I F D S H G - N L D M L I T H N G A - S G N N F V A C N
C2	- - - - G G P G A D R V V F N D N D - E L A G L I T H T G A - S G D G F V A C Y
Ms	- - - - G S P G A D R V I F D G D D - E L A G V I T H T G A A G G D D F V A C S S S
F1	- - - - G S P G A D R V V I N T N C - E T A G A I T H T G A - S G N N F V G C S N T G
Pb1	- - - - S S P G A D R V I F N D D D - E L A G V I T H T G A - S G N N F V A C T
Pch1	- - - - N S P G A D R V V F N G N D - Q L A G V I T H T G A - S G N N F V A C D
Th1	- - - - G S P G A D R V V I N G Q C - S I A G I I T H T G A - S G D A F V A C G G T
U1	- - - - G S P G A D R V V Y D S N D G T F C G A I T H T G A - S G N N F V Q C S Y
U2	R D N Y V S P G P D R V I Y Q T N T G E F C A T V T H T G A A S Y D G F T Q C S
St	- - - - D D R G A R R F V V G - D G G E Y F Y T E D H Y E - - - - - S F R L T I V N
Sa	- - - - R T R G T R R I I C G E A T Q E D Y Y T G D H Y A - - - - - T F S L I D Q T C
Ba	- - - F - - R N S D R I L Y S S D W - L I Y K T T D H Y Q - - - - - T F T K I R
Bi	- - - F - - R N A D R L V Y S S D W - L I Y K T T D H Y A - - - - - T F T R I R

of the predicted structures of RNase Pch1 and Ms, with the structure of RNase T1.

Although the percentage sequence identity of RNase Pch1 is higher to T1 (~70%) than to Ms (~60%) we decided to model the structure of RNase Pch1 on both, the crystal structure of RNase T1 and the predicted structure of RNase Ms as templates, in the hope that future comparison with crystal structures of these proteins will give information about the limitations and accumulation of errors in the procedure of modelling proteins by homology.

II. Materials and methods

The RNase Pch1 structure was modelled using an Evans & Sutherland vector graphics system. Energy minimization was performed with the Biosym program package on a Cray X-MP/24. For secondary structure analysis of the models obtained the DSSP program of Kabsch and Sander (1983) was used. The solvent accessible surface was calculated by the method of Lavery et al. (1981). Graphic representations of the predicted structures were prepared with the plotting program SCHAKAL88.

The RNase Pch1 model structure was predicted based on the conformation of RNase T1 in the complex with T1*2′GMP, and the predicted structure of RNase Ms. Modelling criteria were the same as reported previously for RNase Ms (Floegel et al. 1988).

Although the loop motif 92–100 of RNase Pch1 has an identical sequence to that in RNase T1, it had to be modelled de novo upon predicting RNase Pch1 from the model structure of RNase Ms. A more straightforward approach would have been to take the loop motif from the RNase T1 crystal structure and tansfer it onto the RNase Ms model structure. The topology of the RNase Ms template did not allow this approach, however, since a simple transfer would have caused steric clashes between the loop 92–100 and the RNase Ms template. For modelling the loop motif 30–38, in the prediction of RNase Pch1 from RNase T1, we used the main chain conformation of this motif from the initial model of RNase Ms before energy minimization. Since in RNase Pch1 two amino acids are deleted within this loop motif as in RNase Ms and the starting template RNase T1 was the same this approach was sufficient. Replacing the sidechains to match the sequence of RNase Pch1 did not cause steric problems in the RNase T1 template on which RNase Pch1 was modelled.

For the motif 45–60 – when modelling RNase Pch1 from T1 – the same strategy was employed as described previously (Floegel et al 1988).

After bonding the loops onto the RNase Ms template and the RNase T1 crystal structure template, amino acid substitutions were performed by replacing sidechains whilst keeping the same χ_n sidechain torsion angles wherever possible (Summers et al. 1987). For residues where this was not possible (eg. Ala to Asp), torsion angles had to be introduced which fitted the sidechains in the model structure of Pch1. The two predicted models of RNase Pch1 were then examined residue by residue in order to make sure that no steric clashes occurred. Changes were only found to be necessary for the sidechains of Asp 25, Asn 30, Arg 43, Arg 63, Asn 71 in the model structure of RNase Pch1 which was modelled from Ms. No changes were necessary for RNase Pch1 from T1 reflecting the higher percentage sequence identity between the latter two enzymes. The modelling was completed in the case of the RNase Ms template by deleting the one additional N-terminal and the two additional C-terminal residues which are not present in RNase Pch1.

The modelled structures were then energy minimized with the aid of the Consistent Valene Force Field (Hagler et al. 1979). The parameters used in energy minimization were as described previously (Floegel et al. 1988).

III. Results

The predicted, energy minimized structures of RNase Pch1 and RNase Ms and the energy minimized structure of RNase T1 in the crystalline complex with 2′GMP are shown in Fig. 1 a–d. The total solvent accessible surface of the structure was calculated by the method of Lavery et al. (1981) to be 5214 Å^2 in T1, 5706 Å^2 in Ms, 5277 Å^2 in RNase Pch1 predicted from the RNase Ms model structure and 5616 Å^2 in RNase Pch1 modelled from RNase T1. Table 2 contains the secondary structure elements of RNase T1, Ms and Pch1 as defined by the DSSP program.

The distribution of β-strands is very similar in all four structures with only subtle differences. Moving along the main chain from the N-terminus, β-strands 1 and 2 form the first antiparallel β-sheet in RNases T1, Pch1 and Ms. β-sheet 2 is built up from the remaining β-strands.

Since DSSP was designed for analysing crystallographically determined structures, its input are structures without hydrogen atoms. For calculating hydrogen bonds, hydrogen atoms are added in ideal geometry and with a given bond length, which does not necessarily reproduce the positions they have in the energy minimized structures. Consequently some residues might not meet the constraints of the DSSP program for defining them as part of a specific secondary structure element and thus suggest structural differences. Therefore a detailed geometrical analysis of the hydrogen bond pattern of the two β-sheets was performed, Tables 3 and 4.

Table 3 indicates a less well defined hydrogen bond pattern in the two RNase Pch1 structures if compared to RNase Ms and T1. The β-sheet hydrogen bonding is almost ideal in RNase Ms, whereas hydrogen bonds between Cys 6 and Asn 9 are widened in RNase T1. For both of the RNase Pch1 structures there is at least one donor-acceptor distance which is too long for formation of a hydrogen bond.

β-sheet 2 constitutes the central core of the protein in all four structures. A close analysis reveals that in certain features both of the RNase Pch1 structures are more similar to RNase T1, than RNase Ms is to T1 or to either of the two RNase Pch1 conformations, although the mere comparison of hydrogen bond patterns of β-sheet 2 might suggest a somewhat mixed character among the four con-

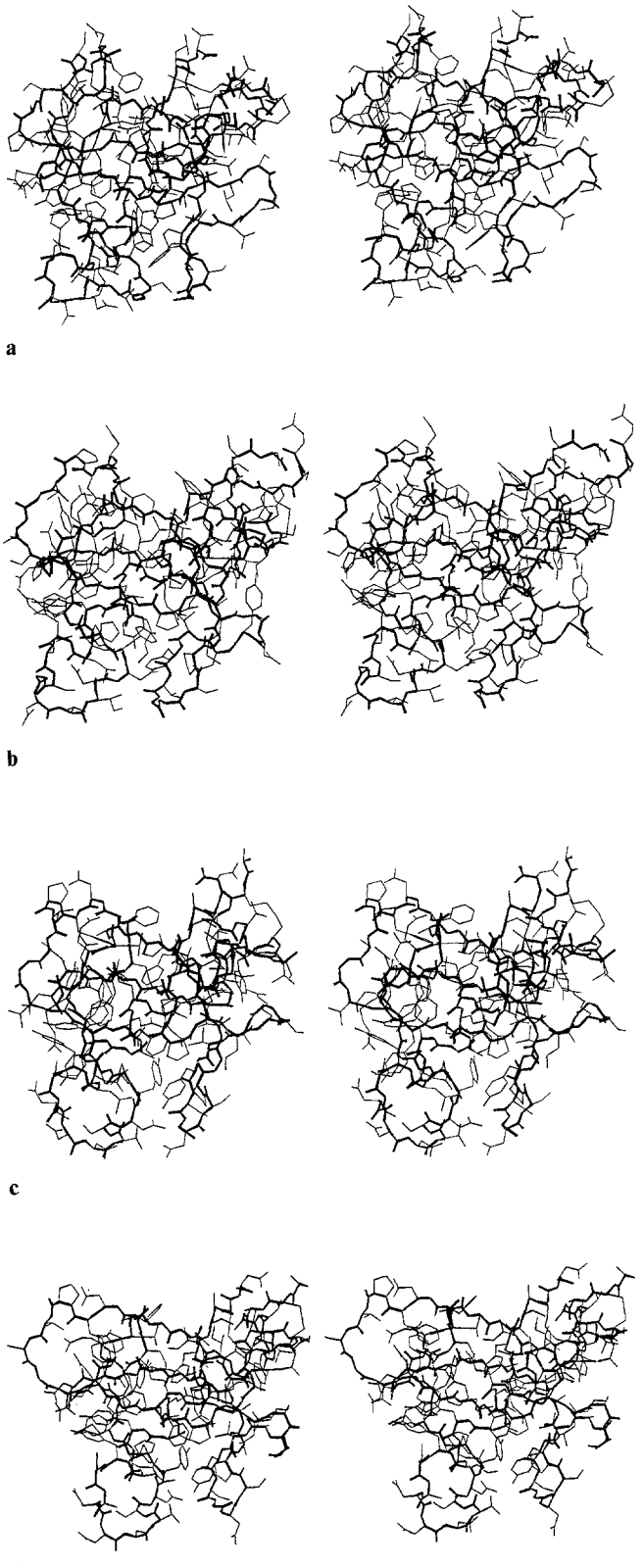


Fig. 1 a–d. Stereo plots of RNase T1 (**a**), RNase Ms (**b**) and RNases Pch1 (**c, d**). Bonds connecting main chain N, Cα and O atoms are indicated by thicker lines. Sidechains are drawn in thinner lines. Hydrogens atoms are omitted. The α-helix lies in the background ranging with its N-terminus from the upper right corner to the lower left with its C-terminus. In order to obtain similar orientations in all plots the α-helical N, Cα and C main chain atoms of RNases T1 and Pch1 were superimposed onto the corresponding atoms of RNase MS

formations. The relative geometries of hydrogen bond donating and accepting HN and O atoms of residues which were defined by DSSP as part of a β-strand in at least one of the structures are compared in Table 4.

The first ladder of the antiparallel β-sheet 2 is formed between strands 3 and 4. In RNase Ms the His 40 N–H..O Glu 58 geometry does not allow formation of hydrogen bonds. Instead the HN hydrogen of His 40 of RNase Ms is oriented to Asp 37 O, with a H..O distance of 2.83 Å and an N–H..O angle of 140°. This kind of hydrogen bonding pattern is not possible in either RNase T1 or the two RNase Pch1 conformations. The corresponding values for the # 40 and # 37 donor acceptor distance and angles are 4.24 Å and 52° in RNase T1, 4.90 Å and 68° in RNase Pch1 predicted from Ms, and 3.49 Å in RNase Pch1 predicted from T1 with an N–H..O angle of 79° indicating that a pattern like in Ms is not possible. Among residues # 42 and # 56 the hydrogen bond geometries between RNase T1 and the two predicted RNase Pch1 models are also more similar compared to RNase Ms.

The next antiparallel ladder of β-sheet 2 being composed of residues # 57, # 59 and # 61 on the one strand and # 76, # 78 and # 80 on the corresponding β-strand shows a somewhat mixed behaviour among the four compared conformations i.e. there is no clear tendency for a closer conformational relationship among both Pch1 structures and T1 compared to Ms.

The distances and angles in Table 4 show that the short β-strand 5 which is connected to strand 4 via residue # 66 to # 62 in RNase Pch1 predicted from Ms is also realised in RNase T1 and, at least with the # 62 HN – # 66 O hydrogen bond also in RNase Pch1 predicted from T1, though it escaped the constraints of the DSSP program for reporting it. The same structural element cannot be realised in RNase Ms with donor acceptor distances between the corresponding residues of 3.30 Å and 4.89 Å.

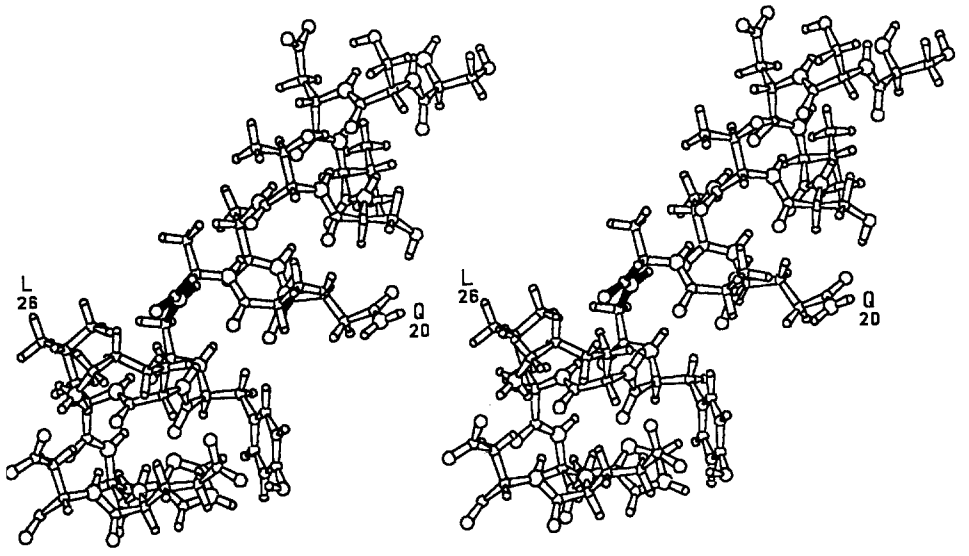
The following ladder of β-sheet 2 is built up from residues # 77, # 79 and # 81 on the one strand and # 85, # 87, # 88 and # 90 on the other. An element common to all four structures is that the ladder is kinked between residues # 79 and # 88. It allows formation of bifurcated hydrogen bonds from the C=O peptide bond oxygen of # 79 not only to the HN of # 88 – which is regular bonding within the β-sheet geometry – but also to the HN of # 87. Then the ladder continues with a hydrogen bond between # 81 HN and # 85 O in RNase T1 and Pch1 from Ms, but not in RNase Ms and RNase Pch1 predicted from T1.

The last ladder of β-sheet 2 is formed between residues # 91 and # 101. When comparing hydrogen bond distances and angles this ladder is almost equally favourable in either of the four conformations.

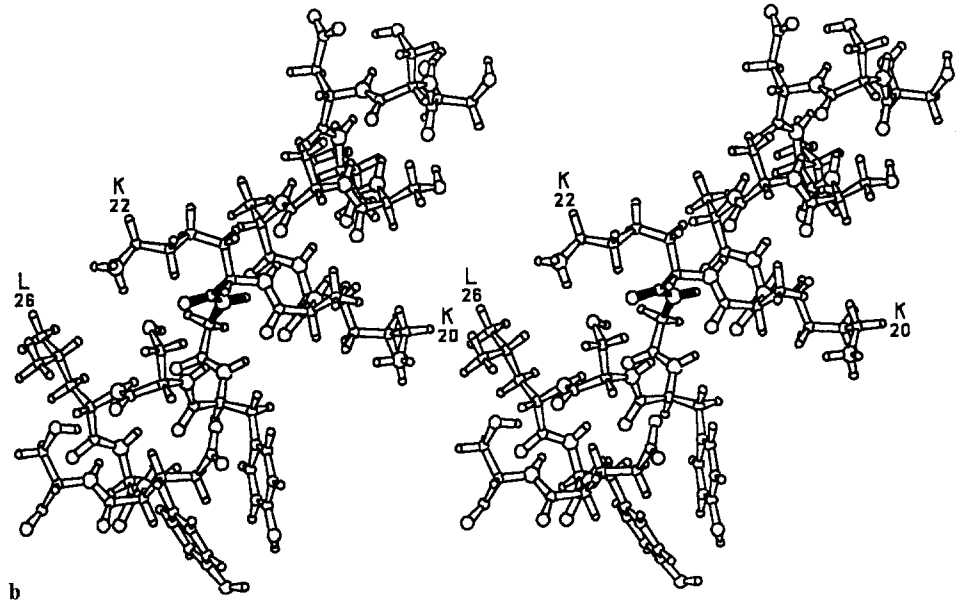
The secondary structure element most conserved in RNases T1, Ms and Pch1 is the α-helix. It comprises the same number of residues in the molecules and is located in the same place, from residues # 13 to # 29. The main chain torsion angles Φ, Ψ are in the normal range, $-60°/-60°$, characteristic for α-helices, except for the one peptide group in *both* RNase Pch1 conformations between Ala 22 and Gly 33.

Table 2. Secondary structure analysis by the DSSP program of the energy minimized structure of RNase T1, and the predicted structures of RNase Ms and RNase Pch1

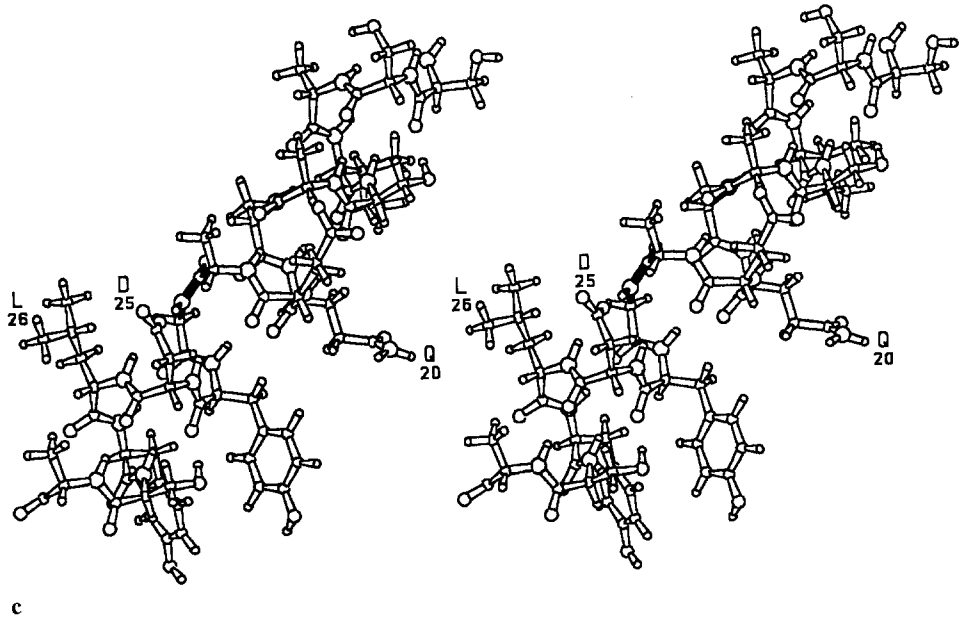
```
                        10        20        30        40        50        60        70        80        90       100
                      .   *   .   *   .   *   .   *   .   *   .   *   .   *   .   *   .   *   .   *   .   *
  RNase T1      ACDYTCGSNCYSSSDVSTAQAAGYKLHEDGETVGSNSYPHKYNNYEGFDFSVSSPYYEWPILSSGDVYSGGSPGADRVVFNENNQLAGVITHTGA-SGNNFVECT
  RNase Ms      ESCEYTCGSTCYWSSDVSAAKAKGYSLYESGDTI--DDYPHEYHDYEGFDFPVGTSYYEYPIMSDYDVYTGGSPGADRVIFDGDDELAGVITHTGAAGGDDFVACSSS
  RNase Pch1    ACAATCGSVCYTSSAISAAQEAGYDLYSANDDV--SNYPHEYRNYEGFDFPVSGTYYEFPILRSGAVYSGNSPGADRVVFNGNDQLAGVITHTGA-SGNNFVACD
                      .   *   .   *   .   *   .   *   .   *   .   *   .   *   .   *   .   *   .   *   .   *
A-HELIX T1                  1111111111111111
        Ms                  1111111111111111
        Pch1                1111111111111111
                      .   *   .   *   .   *   .   *   .   *   .   *   .   *   .   *   .   *   .   *   .   *
B-SHEET T1          11    22                        33            44444             55555    66666
        Ms          111   222                        33           44444              5555      6666          77
        Pch1        111   222                       333           4444444    55      66666    7777777       88
                      .   *   .   *   .   *   .   *   .   *   .   *   .   *   .   *   .   *   .   *   .   *
3-TURN  T1                                      1111                      2222              3333
        Ms      11112222  3333      44-55         6666                    7777
        Pch1      1111    2222 33334444       5--555
                      .   *   .   *   .   *   .   *   .   *   .   *   .   *   .   *   .   *   .   *   .   *
4-TURN  T1
        Ms
        Pch1                                                              11111
                      .   *   .   *   .   *   .   *   .   *   .   *   .   *   .   *   .   *   .   *   .   *
5-TURN  T1                          111111 222222                                                 333333
        Ms                  111111
        Pch1
```



a



b



c

Fig. 2a–d. Stereo plots of the α-helix of RNases Pch1, Ms and T1. The view is from the exterior of the protein onto the helix. Bonds connecting residues #22 and #23 are drawn solid to indicate the conformational differences in the four structures

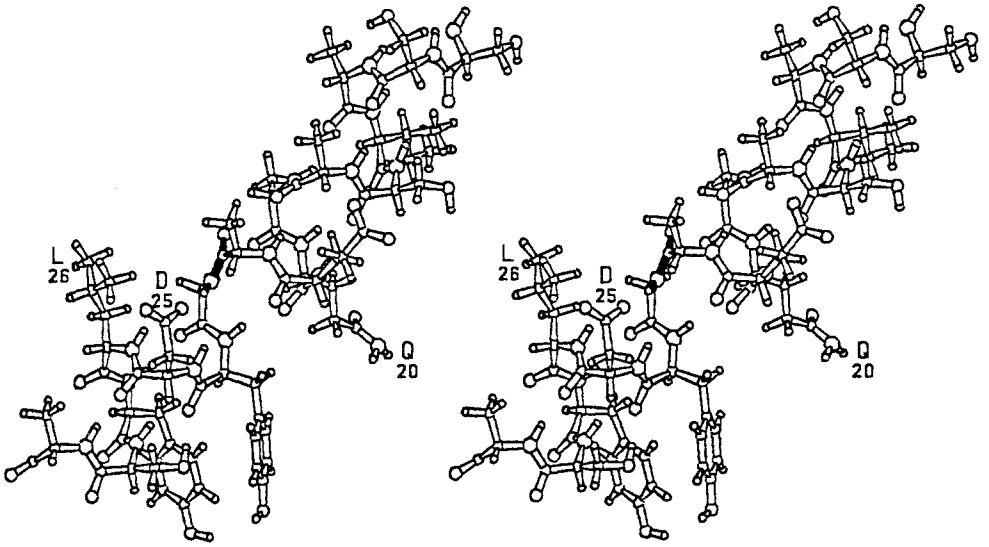


Fig. 2d

Table 3. Distances/angles for possible hydrogen bonds in β-sheet 1. Distances in Angstrom, the angle is measured between N–H..O

	RNase T1	RNase Ms	RNase Pch1
# 4 HN – # 11 O	1.93/171.5	1.95/160.0	2.75/148.9
# 4 O – # 11 HN	2.31/144.4	2.34/141.5	2.01/151.3
# 6 HN – # 9 O	2.35/132.8	2.00/166.7	2.93/112.9
# 6 O – # 9 HN	2.47/132.2	1.98/169.5	3.24/146.2

Table 4. Distances/angles for possible hydrogen bonds in β-sheet 2. Distances in Angstrom, the angle is measured between N–H..O

	RNase T1	RNase Ms	RNase Pch1
# 40 HN – # 58 O	3.00/162.3	4.27/117.0	2.70/157.0
# 40 O – # 58 HN	2.57/139.6	3.64/140.1	2.19/159.7
# 42 HN – # 56 O	1.86/171.9	2.41/122.4	2.35/143.1
# 42 O – # 56 HN	3.86/106.5	2.81/103.7	3.84/136.4
# 57 HN – # 80 O	1.90/145.7	1.81/176.3	3.86/143.8
# 57 O – # 80 HN	2.08/172.9	2.11/173.3	3.96/144.1
# 59 HN – # 78 O	2.08/171.7	2.10/152.0	2.88/148.3
# 59 O – # 78 HN	1.99/163.7	1.86/172.3	1.94/165.7
# 61 HN – # 76 O	1.87/173.2	2.42/150.8	2.01/168.5
# 61 O – # 76 HN	3.87/ 97.5	4.54/107.4	3.56/129.4
# 62 HN – # 66 O	2.07/156.5	3.30/127.6	1.95/159.6
# 62 O – # 66 HN	3.11/112.5	4.89/126.0	2.27/148.5
# 77 HN – # 90 O	2.25/142.2	2.17/145.2	2.16/155.2
# 77 O – # 90 HN	2.04/163.2	1.95/163.7	2.00/163.6
# 79 HN – # 88 O	1.84/170.8	2.01/162.7	2.15/148.6
# 79 O – # 88 HN	2.13/164.8	2.09/163.2	2.25/166.3
# 79 O – # 87 HN	2.09/136.1	2.54/119.1	2.09/143.7
# 81 HN – # 85 O	2.01/163.8	3.29/140.6	1.97/167.0
# 81 O – # 85 HN	3.58/177.0	– –	5.31/152.2
# 91 HN – # 101 O	2.35/147.8	1.85/166.5	2.41/144.0
# 91 O – # 101 HN	1.90/167.2	2.07/164.5	2.04/163.9

This peptide group is rotated by ~180°, disrupting the regular i to $i+4$ hydrogen bonding scheme, and its orientation is defined in the case of RNase Pch1 predicted from Ms by Ala 22 $\Phi = -83°$, $\Psi = 119°$/Gly 23 $\Phi = 129°$, $\Psi = -64°$. The values for RNase Pch1 predicted from T1 are Ala 22 $\Phi = -98°$, $\Psi = 106°$/Gly 23 $\Phi = 153°$, $\Psi = -48°$. Compare Fig. 2a–d.

Gly 23 forms a hydrogen bond with one of the oxygens of the negatively charged carboxy group of Asp 25, which is also in contact with Leu 26 HN. The values for RNase Pch1 predicted from Ms are (Asp 25 OD1–Gly 23 HN 1.79 Å, angle N–H..O 172°/Asp 25 OD1–Leu 26 HN 2.03 Å, 157°). The corresponding numbers for RNase Pch1 predicted from T1 are (Asp 25 OD1–Gly 23 HN 1.82 Å, 164°/Asp 25 OD1–Leu 26 HN 2.54 Å, 148°).

Residues thought to be involved in catalytic action of RNases of the T1 family and which are conserved in pro- and eukaryotic enzymes are Glu 58, Arg 77 and His 92 (Heinemann and Saenger 1983). Arg 77 is also of structural interest, beyond its role in catalytic action which probably is to stabilize the transition state. It is an internal residue with a solvent accessibility of 3 Å^2 in RNase T1, 4 Å^2 in RNase Ms and 8 Å^2 in RNase Pch1 predicted from Ms and 2 Å^2 for the RNase Pch1 conformation predicted from T1. The corresponding values for Glu 58 are 13 Å^2 in T1, 8 Å^2 in Ms, 14 Å^2 in Pch1 form Ms and 12 Å^2 in Pch1 from T1.

Arg 77 and Glu 58, are located in the interior of the catalytic center of the protein, the charges being stabilized by the formation of a salt bridge. The distance between the carboxyl oxygen of Glu 58 and the HE of Arg 77 is 1.83 Å in T1, 2.22 Å in Ms, 2.35 Å in Pch1 from Ms and 1.94 Å in RNase Pch1 from T1. The hydrogen bonds involving the guanidinium group of Arg 77 are almost ideal in RNase Ms, T1 and Pch1 (compare values in Table 5).

The only difference is with the Arg 77 HH12. In RNase T1 and both Pch1 models it bonds to Pro 73 O, in RNase Ms its hydrogen bond partner is Gly 74 O. In

Table 5. Hydrogen bond pattern of Arg 77 in RNase T1, Ms and Pch1 Distances in Angstrom, angles between N–H..O in degrees

	RNase T1	RNase Ms	RNase Pch1
Arg 77 HE	Glu 58 OE2 1.83/152	Glu 58 OE2 2.22/134	Glu 58 OE2 2.35/132
Arg 77 HH11	Asp 76 O 1.84/176	Asp 76 O 2.15/165	Asp 76 O 2.12/170
Arg 77 HH12	Pro 73 O 1.93/122	Gly 74 O 2.01/156	Pro 73 O 1.89/156
Arg 77 HH21	Gly 74 O 2.27/137	Gly 74 O 2.05/150	Gly 74 O 2.04/142
Arg 77 HH22	Tyr 38 OH 2.15/149	Tyr 38 OH 2.26/145	Tyr 38 OH 1.93/162

RNase Ms a bifurcated hydrogen bond between Gly 74 O and Arg 77 HH12 and HH21 occurs, which is not observed in either RNase T1 or in the two predicted Pch1 conformations.

The internal location of Arg 77 in the catalytic site raises the question whether additional interactions are conserved. In all fungal RNases and in two prokaryotic RNases of the T1 family the residue next to Arg 77 is Asp 76, which also is barely accessible to solvent. For RNase T1 the static solvent accessibility of Asp 76 is 0.4 Å^2, for Ms it is 6.8 Å^2, for Pch1 from Ms it is 1.0 Å^2 and for Pch1 from T1 the solvent accessible surface was calculated to be zero. Coulombic interaction adds to the mutual stabilization of Asp 76 and Arg 77.

The situation is slightly different for both RNase Pch1 conformations where Arg 63, which lies on the surface of the protein, can form a salt bridge between its guanidinium group and the carboxy group of Asp 76.

A residue which also interacts with Arg 77 is Tyr 38. It is conserved in all RNases of the T1 family and has a low solvent accessibility in the structures of RNase T1, Ms and Pch1 (T1 7 Å^2/Ms 22 Å^2/Pch1 from Ms 26 Å^2/Pch1 from T1 38 Å^2). Tyr 38 is possibly conserved in the fungal RNases because the hydrogen bond between Tyr 38 OH and Arg 77 HH22 is important. No other favourable hydrogen bonds can be formed by these atoms in all predicted conformations.

IV. Conclusions

The amino acid sequence similarity of RNase Pch1 to RNase T1 is ~70%, compared to ~60% of RNase Ms to both RNase T1 and RNase Pch1. For reasons described in the outset, RNase Pch1 was also predicted based on the template structure of RNase Ms – which itself had been predicted from RNase T1.

Comparisons to future X-ray crystal structures of the two predicted enzymes will elucidate how far one can extend prediction techniques.

It is interesting to not that the discussed structural features of the predicted conformations indicate a closer relationship between the G-specific RNase T1 and Pch1, compared to the G-preferential, base non-specific RNase Ms.

Important in this context may be the change of the hydrogen bond pattern of the His 40 peptide bond in RNase Ms, which does not participate in a double β-bridge to the peptide bond Glu 58. A reason for it could be the interaction of the positively charged His 40 imidazole group with Asp 36 and Asp 37 in RNase Ms, which are not present in RNases T1 and Pch1. Although no direct interaction of the imidazole group and the carboxy groups can be visualised, long-range electrostatic interactions attract the sidechain of His 40 and thus also dislocate the main chain of His 40. The His 40 peptide bond HN in RNase Ms then interacts with Asp 37 O and not with Glu 58 O.

In RNase Pch1 modelled from Ms, where no negatively charged residues are in position # 36 and # 37, the Force Field method reconstituted the β-bridge between His 40 and Glu 58.

It is remarkable that the flip of the peptide group between Ala 22 and Gly 23 in the α-helix occurred in both RNase Pch1 conformations, independent of the starting template used for predictions. It is possibly due to the vicinity of the negatively charged residue Asp 25. As far as the specific conformation of Asp 25 in the predicted structures are concerned, it has to be noted that it is unlikely that conformations of solvent accessible and charged sidechains in the model sets represent the situation which is actually found in solution, where they can extend to the exterior and interact with solvent. Water had to be excluded from the minimization process in order to reduce the required calculation time to a justifiable level. It would be interesting to see whether such as disruption of a regular helical pattern can be detected in the actual crystal structure of RNase Pch1 which is still undetermined. This would require high resolution crystallographic data to better than 3 Å resolution in order to be 'visible'.

The lack of water in the minimization process may also account for the difference in total solvent accessibility of RNase Pch1 predicted from Ms and predicted from T1. When proteins are minimized in the absence of water a shrinking effect in total volume is observed. Since RNase Pch1 predicted from Ms represents a template which had already gone through a minimization process the subsequent minimization could have further reduced the solvent accessible surface of RNase Pch1 predicted from Ms as compared to RNase Pch1 predicted from the RNase T1 crystal structure.

Since the structure of RNase Ms has been determined by molecular replacement based on the RNase T1 crystal structure, the main chain RMS differences between crystal Ms and crystal T1 are available; they are 0.83 Å (K. Nakamura, personal communication). The RMS difference between the predicted structure of RNase Ms and its crystal structure for main chain atoms is 1.89 Å which is in the usual range if prediction techniques combined with force field methods are compared with X-ray data. For the energy minimized structure of RNase T1 the RMS difference for main chain atoms is 1.79 Å when compared to the crystal structure. Furthermore the RMS differences be-

tween *either* of the four energy minimized structures are around ~2 Å for main chain atoms, which is mainly due to high discrepancies in loop regions for which there are not yet any conclusive rules for modelling. Therefore predicting structures based on homologous structures should be interpreted as a means of elucidating common structural features instead of pinpointing at atomic positions.

The present study discloses that the predominant features of the predicted structures of RNase Pch1 are similar, independent of the templates, the model structure of RNase Ms and the crystal structure of RNase T1. A number of structural features indicate a closer relationship between either of the RNase Pch1 conformations to RNase T1. This is in agreement with the observed higher sequence homology of RNases Pch1 and T1, and the identical specificities of the two enzymes. These results show that the procedure of modelling proteins by homology, when combined with energy minimization techniques, can probably be reliably used for step by step predictions of related structures. However, only by comparison with future crystallographic structures we will be able to assess the bias introduced on the predicted structures by the selection of templates.

Acknowledgements. We thank Dr. K. Nakamura for providing comparison of our RNase Ms predicted structure and his unpublished crystal structure of the RNase Ms*3'GMP complex. This work was supported by a fellowship of the Alexander von Humboldt Foundation to P.Z., by Fonds der Chemischen Industrie and the Deutsche Forschungsgemeinschaft (Sonderforschungsbereich 9 and Leibniz-Program).

References

Aphanasenko GA, Dudkin SM, Kaminir LB, Leshchinskaya IB, Severin ES (1979) Primary structure of ribonuclease from bacillus intermedius 7P. FEBS Lett 97:77–80

Arni R, Heinemann U, Tokuoka R, Saenger W (1988) Three-dimensional structure of the ribonuclease T1*2'GMP complex at 1.9-resolution. J Biol Chem 263:15358–15368

Bezborodova SI, Khodova OM, Stepanov VM (1983) The complete amino acid sequence of ribonuclease C2 from Aspergillus clavatus. FEBS Lett 159:375–382

Bezborodova SI, Beletskaya OP, Tishchenko GN (1988) Crystallization, the growth of single crystals and preliminary crystallographic data of RNases C2 and Pch1. In: Sevcik J, Zelinka J (eds) Ribosomes and nucleic acid metabolism, vol 3. Slovak Acad Sci, Bratislava, pp 177–188

Bielefeld M, Lauwereys M, van Vliet A, Stanssens P (1988) The barnase-barstar system: a model to study protein -protein interactions. In: GBF Workshop-Advances in Protein Design. GBF, Brunswick, FRG, p 50

Floegel R, Zielenkiewicz P, Saenger W (1988) The tertiary structure of Aspergillus saitoi minor ribonuclease (Ms) predicted from the structure of RNase T1. FEBS Lett 240:29–32

Hagler AT, Dauber P, Lifson S (1979) Consistent force field studies of intermolecular forces in hydrogen-bonded crystals. 3. The C=O..H–O hydrogen bond and the analysis of the energetics and packing of carboxylic acids. J Am Chem Soc 101:5131–5141

Hartley RW (1980) Homology between prokaryotic and eukaryotic ribonucleases. J Mol Evol 15:355–358

Hartley RW, Barker EA (1972) Amino-acid sequence of extracellular ribonuclease (barnase) of bacillus amyloliquefaciens. Nature: 15–16

Heinemann U, Saenger W (1982) Specific protein-nucleic acid recognition in ribonuclease T1-2'-guanylic acid complex: an X-ray study. Nature 299:27–31

Heinemann U, Saenger W (1983) Crystallographic study of mechanism of ribonuclease T1 catalysed specific RNA hydrolysis. J Biomol Struct Dyn 1:523–538

Hirabayashi J, Yoshida H (1983) The primary structure of ribonuclease F1 from Fusarium moniliforme. Biochem Int 7:255–262

Kabsch W, Sander Ch (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers 22:2577–2637

Kanaya S, Uchida T (1986) Comparison of the primary structures of ribonuclease U2 isoforms. Biochem J 240:163–170

Lavery R, Pullman A, Pullman B (1981) Steric accessibility of reactive centers in B-DNA. Int J Quant Chem 20:49–62

Polyakov KM, Strokopytov BV, Vagin AA, Bezborodova SI, Shlyapnikov SV (1987) In: Crystallization and preliminary X-ray structural studies of RNase Th1 from Trichoderma harzianum. In: Sevcik J, Zelinka J. (eds) Metabolism and enzymology of nucleic acids, vol 6. Slovak Acad Sci, Bratislava, pp 331–334

SCHAKAL88 by E. Keller, Kristallographisches Institut der Albert-Ludwigs-Universität Freiburg, FRG

Shlyapnikov SV, Kulikov VA, Yakovlev GI (1984) Amino acid sequence and S-S bonds of Penicillium brevicom- pactum guanyl-specific ribonuclease. FEBS Lett 177:256–248

Shlyapnikov SV, Bezborodova SI, Kulikov VA, Yakovlev GI (1986) Express analysis of protein amino acid sequences: primary structure of penicillium chrysogenum 152A guanyl-specific ribonuclease. FEBS Lett 196:29–33

Shlyapnikov SV, Both V, Kulikov VA, Dementiev AA, Sevcik J, Zelinka J (1987) Amino acid sequence determination of guanyl-specific ribonuclease Sa from Streptomyces aureofaciens. FEBS Lett 209:335–339

Summers NL, Carlson WD, Karplus M (1987) Analysis of side-chain orientations in homologous proteins. J Mol Biol 196:175–198

Takahashi K (1985) A revision and confirmation of the amino acid sequence of ribonuclease T1. J Biochem 98:815–817

Takahashi K (1988) The amino acid sequence of ribonuclease N1, a guanine-specific ribonuclease from the fungus neurospora crassa. J Biochem 104:375–382

Takahashi K, Hashimoto J (1988) The amino acid sequence of ribonuclease U1, a guanine-specific ribonuclease from the fungus ustilago sphaeroghena. J Biochem 103:313–320

Watanabe H, Ohgi K, Irie M (1982) Primary structure of a minor ribonuclease from Aspergillus saitoi. J Biochem 91:1495–1509

Yoshida N, Sasaki A, Rashid MA, Otsuka H (1986) The amino acid sequence of ribonuclease St. FEBS Lett 64:122–125