Protein Engineering vol. 16 no. 6 pp. 391–396, 2003
DOI: 10.1093/protein/gzg056

Substrate binding and catalysis in trichosanthin occur in different sites as revealed by the complex structures of several E85 mutants

Qing Guo[1,3], Weihong Zhou[1], Hiu-Mei Too[2], Jianhui Li[3], Yiwei Liu[1], Mark Bartlam[1], Yicheng Dong[1,3], Kam-Bo Wong[2], Pang-Chui Shaw[2,4] and Zihe Rao[1,3,4]

[1]Laboratory of Structural Biology and the MOE Laboratory of Protein Science, School of Life Science and Engineering, Tsinghua University, Beijing 100084, [2]Department of Biochemistry, Chinese University of Hong Kong, Hong Kong and [3]National Laboratory of Macromolecules, Institute of Biophysics, Chinese Academy of Science, Beijing, 100101, China

[4]To whom correspondence should be addressed.
E-mail: raozh@xtal.tsinghua.edu.cn; pcshaw@cuhk.edu.hk

Trichosanthin (TCS) is a type I ribosome-inactivating protein (RIP) which possesses rRNA *N*-glycosidase activity. In recent years, its immunomodulatory, anti-tumor and anti-HIV properties have been revealed. Here we report the crystal structures of several E85 mutant TCS complexes with adenosine-5′-monophosphate (AMP) and adenine. In E85Q TCS/AMP and E85A TCS/AMP, near the active site of the molecule and parallel to the aromatic ring of Tyr70, an AMP molecule is bound to the mutant without being hydrolyzed. In the E85R TCS/adenine complex, the hydrolyzed product adenine is located in the active pocket where it occupies a position similar to that in the TCS/NADPH complex. Significantly, AMP is bound in a position different to that of adenine. In comparison with these structures, we suggest that there are at least two subsites in the active site of TCS, one for initial substrate recognition as revealed by the AMP site and another for catalysis as represented by the NADPH site. Based on these complex structures, the function of residue 85 and the mechanism of catalysis are proposed.

Keywords: complex structure/mutants/*N*-glycosidase mechanism/ribosome-inactivating protein/trichosanthin

Introduction

Ribosome-inactivating proteins (RIPs) are a group of cytotoxins that possess a unique rRNA *N*-glycosidase activity. They can hydrolyze a single *N*-glycosidic bond between adenine and ribose at A4324 in the 28S rRNA of rat liver ribosome. This reaction damages the ribosome's elongation factor-binding site, hence the irreversible consequent arrest of protein synthesis (Endo and Tsurugi, 1987; Endo *et al.*, 1987; Stirpe *et al.*, 1988). Two types of RIP are found, type I and type II. Type II RIP consists of a catalytically active A chain linked to a carbohydrate-binding B chain. The B chain possesses lectin properties, which facilities entry of the A chain into the cytoplasm of the cell. Type I RIPs consist of a single polypeptide chain, which is homologous with the A chain of type II RIP.

Trichosanthin (TCS) is a type I RIP isolated from the root tuber of *Trichosanthes kirilowii* Maximowicz (Zhang and Wang, 1986). It is a single-chain polypeptide with 247 residues and consists of two domains: an N-terminal domain (residues 1–184) and a small C-terminal domain (residues 185–247). The active pocket responsible for *N*-glycosidase activity is located in the cleft between the two domains (Gao *et al.*, 1993). TCS has various pharmacological properties, including immunomodulatory, anti-tumor and anti-viral activities (Shaw *et al.*, 1994). In particular, TCS and other related proteins, such as α-momorcharin (αMMC), have been found to act selectively against HIV-infected cells *in vitro* by inactivating HIV replication in infected T cells and macrophages. Clinical trials have been performed and the results showed an increase of $CD4^+$ and $CD8^+$ T cells in patients (Byers *et al.*, 1994; Kahn *et al.*, 1994).

In order to explore the *N*-glycosidase mechanism of RIPs, the structures of TCS, αMMC and their complexes with substrate analogues such as FMP and NADPH have been reported (Ren *et al.*, 1994; Xiong *et al.*, 1994; Yang *et al.*, 1994; Huang *et al.*, 1995. These studies have provided the structural basis for understanding the mechanism of the *N*-glycosidase activity. TCS removes an adenine residue at position A4324 of mammalian 28S ribosome. Therefore, AMP has been used as a substrate analogue in crystallographic studies. However, previously reported attempts to prepare a TCS complex with AMP have been unsuccessful since TCS is able to interact with and hydrolyze AMP. A number of TCS mutants, including E160A, E160D, R163H, E189A and Y70A, have been used previously to prepare a complex with AMP (Li *et al.*, 1999a,b; Yan *et al.*, 1999), but in each case the AMP was hydrolyzed by TCS and the result was a complex with adenine.

In order to examine the effects of the charge and size of the side chain of residue 85 on the *N*-glycosidase activity, three mutants (E85Q, E85A, E85R) of trichosanthin were used to prepare a complex with AMP. Here we report the crystal structures of the three complexes and propose the mechanism of catalysis of *N*-glycosidase suggested by these complex structures.

Materials and methods

Construction of expression vector and protein purification

Three site-directed mutants at position 85 of TCS were generated by polymerase chain reaction (PCR) mutagenesis using Pfu DNA polymerase (Strategene) on pET8C carrying the wild-type TCS sequence. The PCR product was cleaved by *Nco*I and *Bam*HI and cloned to pET8C for expression. The mutation was confirmed by DNA sequencing. A single colony was inoculated into 2 l of LB medium containing 50 μg/ml ampicillin and 25 μg/ml chloramphenicol and the protein was induced by 0.4 mM IPTG at 37°C for 3–4 h. The protein was purified using an 8 × 2.5 cm CM-Sepharose CL6B column (Amersham Pharmacia Biotech) and 20 mM phosphate buffer with a 0–1 M NaCl gradient.

Table I. Crystal parameters, data collection and refinement statistics

	E85A TCS/AMP	E85Q TCS/AMP	E85R TCS/adenine
Crystal parameters and data collection statistics			
Resolution (Å)	1.85	1.7	1.8
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit cell constants (Å)			
a	38.3	38.0	38.0
b	76.7	75.9	75.3
c	79.2	78.4	78.4
Total no. of reflections	20174	24510	20470
R_{merge} (%)	4.9 (19.9[a])	6.1 (23.6[a])	7.7 (51.7[a])
Completeness (%)	97.8 (73.1[a])	95.5 (95.6[a])	99.2 (99.2[a])
Refinement statistics			
Resolution range for refinement (Å)	30.0–1.86	30.0–1.7	30.0–1.8
R-factor	0.173	0.193	0.210
R_{free}[b]	0.226	0.234	0.247
R.m.s. bond length deviation (Å)	0.013	0.012	0.019
R.m.s. bond angle deviation (°)	1.67	1.58	1.29

[a]Numbers in parentheses correspond to the highest resolution shell: 1.89–1.85 Å for E85A TCS/AMP, 1.76–1.70 Å for E85Q TCS/AMP and 1.86–1.80 Å for E85R TCS/adenine.
[b]5% of the total reflections were used for the R_{free} calculation.

Crystallization and preparation of AMP complexes

Crystallization of three mutants was carried out by the hanging drop vapor diffusion method. Solution A contained 40 mg/ml protein and solution B contained 100 mM NaOAc–HOAc buffer (pH 5.7) containing 20% KCl and 100 mM $CaCl_2$. A 2 µl volume of solution A and an equal amount of solution B were mixed in droplets; the reservoir solution used was 0.2 ml of solution B. Crystals were obtained after 24 h at 291 K. The three complex crystals were prepared by the soaking them in a reservoir solution containing 10 mg/ml AMP at 291 K for 60 h.

Data collection and processing

Diffraction data were collected using an MAR345 image plate with a Rigaku RU2000 Cu Kα rotating anode X-ray generator operating at 48 kV and 98 mA (λ = 1.5418 Å). The three complex crystals all belong to the space group $P2_12_12_1$. The data were processed with the programs DENZO and SCALEPACK (Otwinowski and Minor, 1997). The crystal parameters and data collection statistics are listed in Table I.

Structure solution and refinement of the three complexes

The three complex crystals are isomorphous with wild-type TCS, hence the starting model used for molecular replacement was the refined structure of TCS (Gao *et al.*, 1993) with E85 replaced by G85. The three complex structures were all solved by the difference Fourier method. The structure refinements were carried out using the CNS software package (Brünger *et al.*, 1998) and all manual rebuilding of coordinates was performed using the program O (Jones *et al.*, 1991).

In the electron density maps of both E85Q TCS and E85A TCS complexes, clear electron densities were observed in the active pocket (Figure 1). The AMP molecule can be fitted into the electron density smoothly. In the electron density maps of the E85R TCS complex, however, the shape and size of the electron density in the active pocket were sufficient only to accommodate an adenine molecule. The adenine molecule was built into the electron density.

Simulated annealing refinement was carried out by decreasing the temperature in 25 K steps from an initial value of 5000 to 300 K. The crystallographic *R*-factor was reduced from 0.290 to 0.231 for E85Q TCS/AMP. A further 80 steps of energy-restrained least-squares refinement and 20 steps of restrained individual thermal factor refinement further decreased the *R*-factor from 0.231 to 0.211. $2F_o - F_c$ and $F_o - F_c$ maps were calculated and inspected using the program O. Water molecules were assigned in the model if they could form reasonable hydrogen bonds with significant electron density in both the $2F_o - F_c$ map (contoured at 1.0σ) and $F_o - F_c$ map (contoured at 2.5σ). The water molecules with temperature factors >50 $Å^2$ were removed. The same refinement protocol was used for the E85A TCS/AMP and E85R TCS/adenine complexes. The final R_{work} and R_{free} values were 0.193 and 0.234 for E85Q TCS/AMP, 0.173 and 0.226 for E85A TCS/AMP and 0.210 and 0.247 for E85R TCS/adenine, respectively. Refinement statistics for the three complex structures are summarized in Table I.

Results

Overall structure of the three complexes similar to TCS

After refinement, the *R*-factors and r.m.s deviations of bond length and bond angles of the three complexes are reasonable. The distribution of main-chain torsion angles was calculated using the program PROCHECK (Laskowski *et al.*, 1993). All main-chain torsion angles lie within the allowed regions of the Ramachandran plot. Residue 85 fits the electron density map well in each of the three mutant structures. Equally, AMP and adenine show a good fit with the electron density in their respective complex structures. The fitting of AMP in the electron density map for E85Q TCS/AMP is shown in Figure 1a.

The TCS in the three complex structures is very similar to uncomplexed wild-type TCS. Superposition of Cα backbones shows r.m.s deviations of 0.41 Å between E85Q TCS/AMP and TCS, 0.31 Å between E85A TCS/AMP and TCS and 0.33 Å between E85R TCS/adenine and TCS. The main differences in conformation occur in the last three residues of the C-terminus.

Interactions between AMP or adenine and TCS in the active pocket

In the E85Q TCS/AMP complex, a network of hydrogen bonds between TCS and AMP helps to stabilize the substrate in its

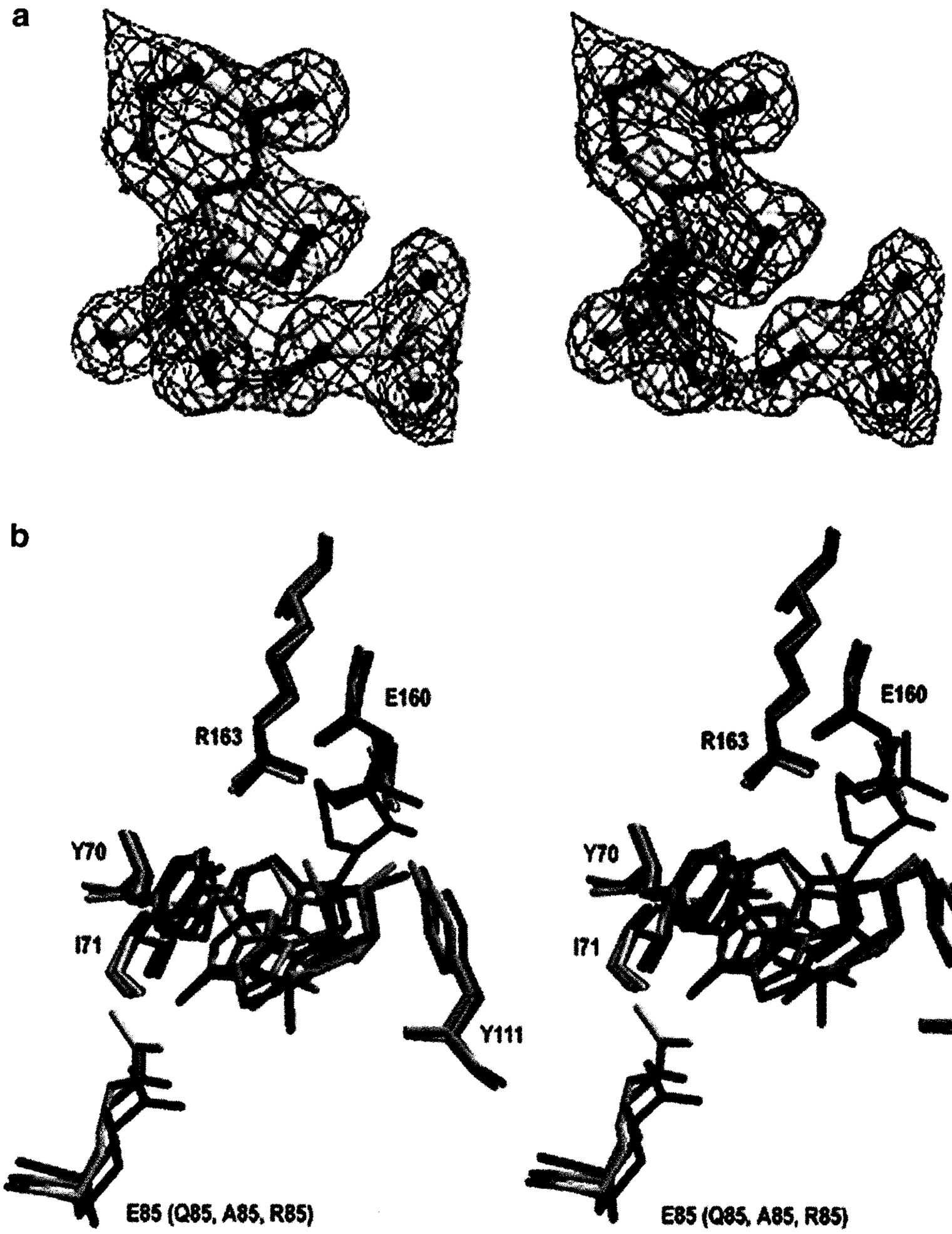

Fig. 1. (**a**) Stereo view of AMP in the $2F_o - F_c$ electron-density map. The map is calculated to 1.7 Å resolution and contoured at the 1σ level. (**b**) Superposition of the active site structures of E85Q TCS/AMP (red), E85A TCS/AMP (green), E85R TCS/adenine (yellow) and TCS/NADPH (blue). Red AMP and green AMP are close, yellow adenine and blue adenine of NADPH are close.

location in the binding pocket of TCS (Figure 2a). The N3 nitrogen of AMP interacts with OE1 of Glu160 and NH1 of Arg163 indirectly through a water molecule. The O atom of Glu189, OE2 of Glu160 and O3′ of AMP interact with each other through a water molecule. The N6 nitrogen of AMP makes two hydrogen bonds with OE1 of Gln85 and O of Ile71. The OH of Tyr70 makes two hydrogen bonds with O3P and O5′ of AMP. There is another hydrogen bond between O4 of AMP and N of Tyr111.

Whereas the position of AMP in the E85A TCS/AMP complex is similar to that in the E85Q TCS/AMP complex, the interactions between AMP and TCS are slightly different in these two complexes. In the E85A TCS/AMP complex, there are no hydrogen bonds formed by the side chain of the Ala85 residue. The hydrogen bonds formed between AMP and residue 85, which are observed in the E85Q TCS/AMP complex, are eliminated. Since the alanine side chain is less bulky than that of glutamine, more water molecules can be accommodated in the active pocket. The net effect of these additional water molecules is to stabilize the AMP molecule in the binding pocket of TCS, as more hydrogen bonds between AMP and TCS are formed indirectly through these water molecules (Figure 2b).

In the E85R TCS/adenine complex, the adenine molecule is located in the bottom of the active pocket. The interactions between adenine and TCS are clearly different from those observed in the E85Q TCS/AMP and E85A TCS/AMP complex structures. In the E85R TCS/adenine complex, the N3 atom of adenine forms a direct hydrogen bond with NH1 of Arg163. This interaction is similar to that between NADPH and TCS in the TCS/NADPH complex (Xiong *et al.*, 1994). The orientation of adenine in the E85R TCS/adenine complex is

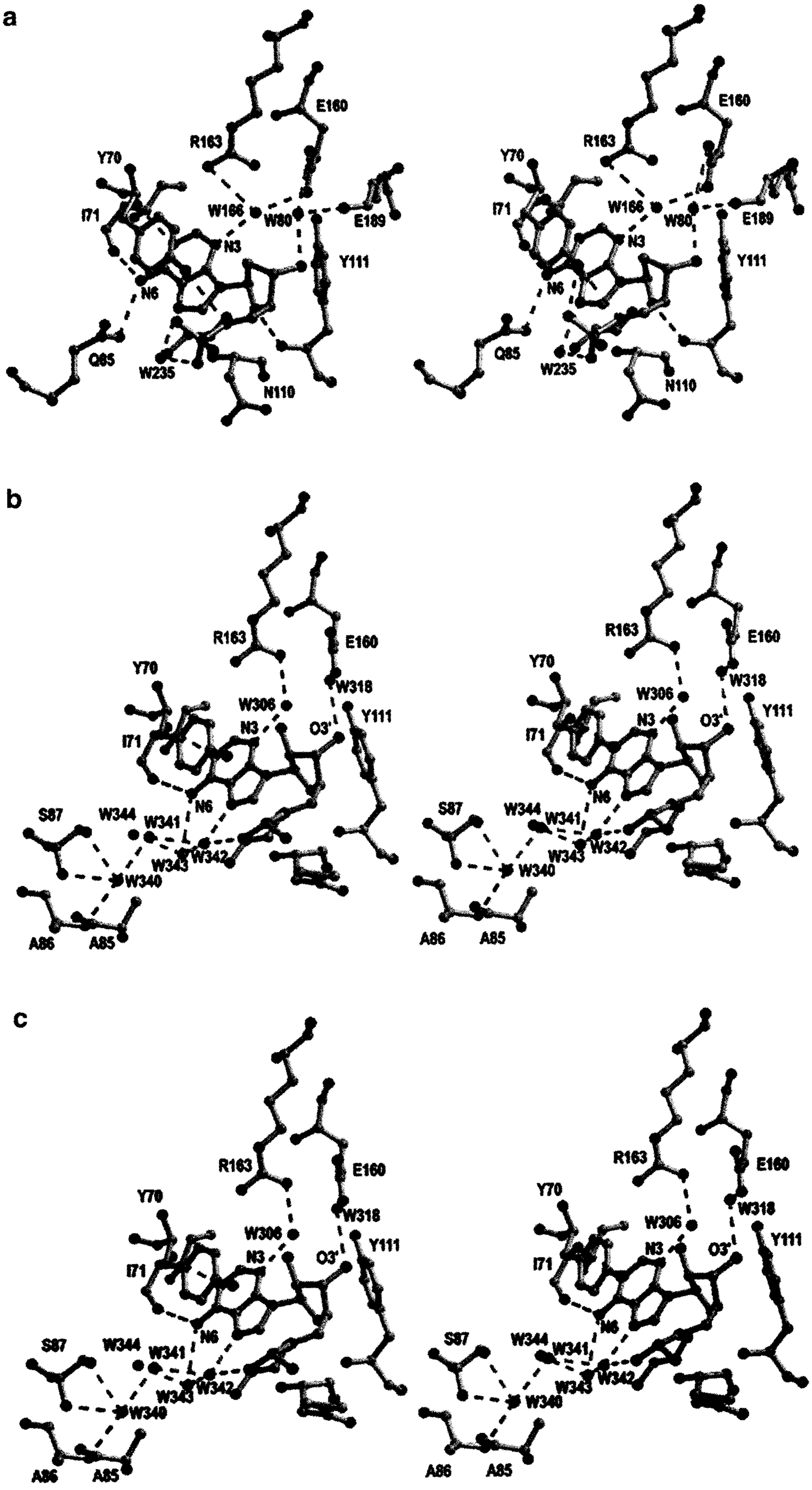

Fig. 2. Stereo drawings of the active site structure in (**a**) E85Q TCS with AMP bound, (**b**) E85A TCS with AMP bound and (**c**) E85R TCS with adenine bound. Hydrogen bonds (<3.4 Å) are indicated using red dashed lines.

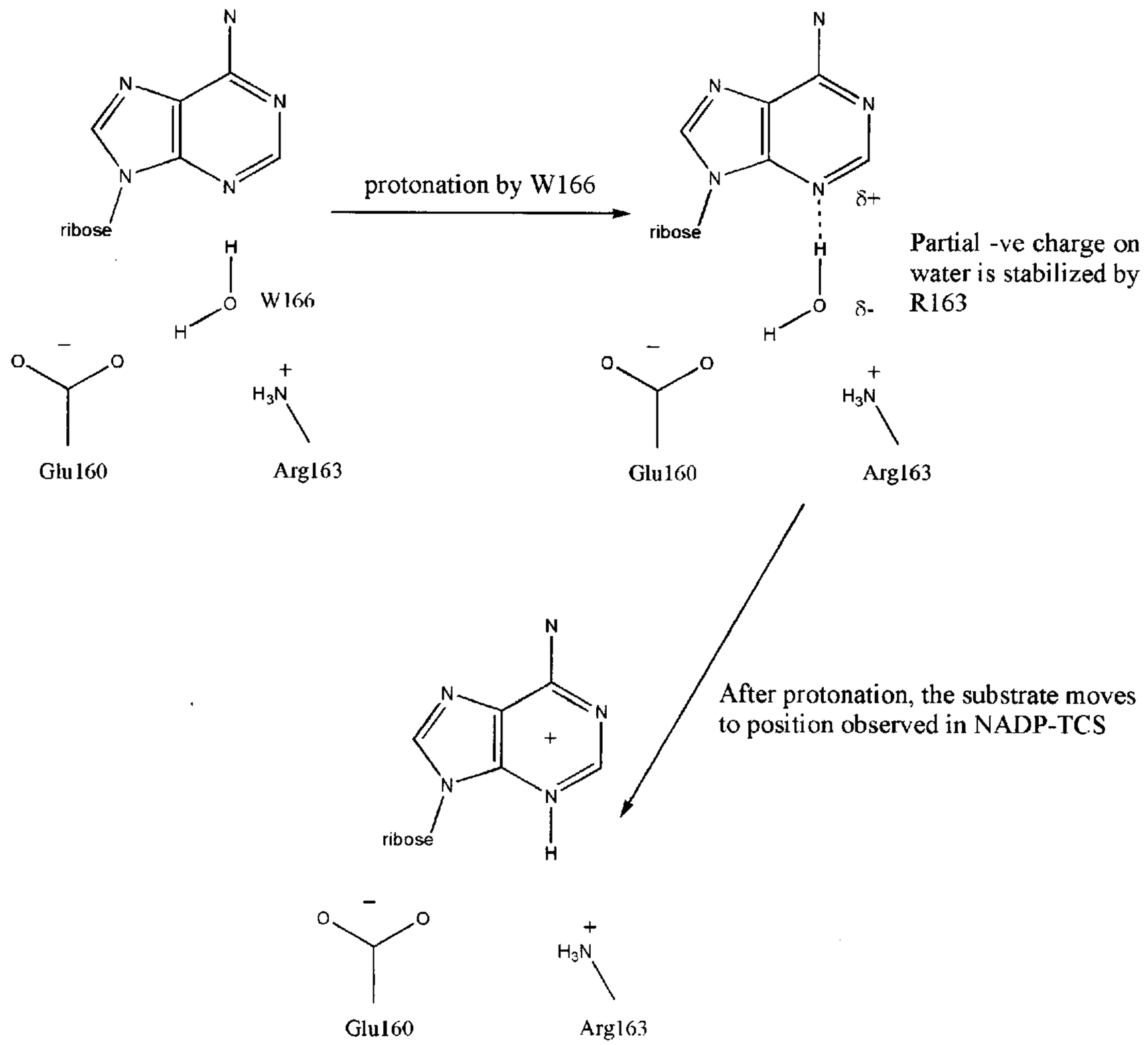

Fig. 3. Protonation process at N3 of adenine.

Table II. The main hydrogen bonds between TCS and substrates (AMP and adenine) in the active pocket

E85A TCS/AMP	E85Q TCS/AMP	E85R TCS/ adenine
N6 of AMP–W343	N6 of AMP–OE1 of Q85	N1 of adenine–N of I71
N6 of AMP–O of I71	N6 of AMP–O of I71	N6 of adenine–O of I71
O3′ of AMP–W318–O of E189	O3′ of AMP–W80–O of E189	
O3′ of AMP–W318–OE1 of E160	O3′ of AMP–W80–OE2 of E160	
N3 of AMP–W306–NH1 of R163	N3 of AMP–W166–NH1 of R163	N3 of adenine–NH2 of R163
		N3 of adenine–NH1 of R163

slightly different to that in the TCS/NADPH complex, resulting in the formation of more hydrogen bonds between adenine and TCS (Figure 2c).

The main hydrogen bonds between substrate analogues and TCS in four known complex structures are listed in Table II. The superimposed map of related residues and substrate analogues in the four structures is shown in Figure 1. The AMP molecule adopts very similar positions in the binding pockets of both E85Q and E85A TCS mutants. The adenine molecules of both the E85R TCS/adenine and TCS/NADPH complexes are similarly oriented in the binding pocket. However, the position of AMP in the binding pocket is clearly different from the adenine position, which suggests the presence of two binding sites in the active pocket.

Effect of residue 85 on the substrate bound in the active pocket

Residue E85 is not highly conserved among the RIP family members. According to the structure determination of three E85 mutant complexes, their overall structures are very similar.

However, two different types of complex were obtained. The first type of complex, obtained for E85Q and E85A mutants, has a complete AMP molecule bound in the active pocket. The second type of complex, obtained for the E85R mutant, has an adenine molecule bound in the active pocket. This adenine molecule is the product of hydrolysis of AMP by TCS, as reported for the complex structures of other TCS mutants (Li *et al.*, 1999a,b). On preparing the complex crystal by the soaking method, the TCS/adenine complex was obtained when residue 85 was charged and the TCS/AMP complex was obtained when residue 85 was uncharged. This suggests that residue 85 could have important implications for *N*-glycosidase activity. The size of the side chain of residue 85 appears to have no effect on the molecule bound in the active pocket.

Protonation occurs at N3 of adenine

Two different hypotheses about protonation of the adenine substrate have been reported to date. First, Huang *et al.* proposed that atom N7 of adenine is protonated by Glu85 of TCS (Huang *et al.*, 1995). However, in the study of the interactions between TCS and tubercidin, Wu *et al.* observed that the *N*-glycoside bond in tubercidin is cleaved by TCS (Wu *et al.*, 1998). In tubercidin, the equivalent to atom N7 in adenine is a saturated carbon that cannot be protonated (Wu *et al.*, 1998). Second, Ren *et al.* proposed that atom N3 of adenine is the most likely protonation site on the basis of the hydrogen bond between N3 and guanidinum of Arg163 (Ren *et al.*, 1994). However, this arginine group is unlikely to protonate adenine owing to its high pK_a value (Ren *et al.*, 1994).

From the E85Q TCS/AMP and E85Q TCS/AMP complex structures, the N3 atom of adenine as a protonation site is possible. A stable water molecule is located at the top of the binding pocket, where it forms hydrogen bonds with N3 of AMP and guanidinum of Arg163. This water molecule is ideally placed to act as a proton donor when the AMP molecule enters at the top of the active pocket. Following protonation of the N3 nitrogen of AMP, Arg163 would be able to stabilize the partial negative charge on the water molecule. The substrate would then move to the bottom of the binding pocket, into the position occupied by NADPH in the TCS/NADPH complex structure. This proposed mechanism of protonation is shown in Figure 3.

Conclusions

From the above structural comparison between TCS/AMP, TCS/adenine and TCS/NADPH complexes, it is reasonable to suggest that the *N*-glycosidase activity can be divided into two steps, recognition and hydrolysis. First the substrate is recognized and bound to the enzyme at the top of the active pocket, in a similar position to AMP in the E85Q TCS/AMP and E85A TCS/AMP complex structures. The position of adenine in the E85R TCS/AMP complex is the position for catalysis, which is similar to NADPH in the TCS/NADPH complex structure. Residue 85 with or without charge has an important effect on *N*-glycosidase activity. However, the size of the side chain of residue 85 has no effect on substrate binding.

Acknowledgements

Z.R. was supported by grants from Project 863 No. 2001AA233011, Project 973 No. G1999075602 and the Ministry of Science and Technology No. 2002BA711A12. The work in Hong Kong was supported by a grant (CUHK 4145/01M) from the Research Grants Council of Hong Kong.

References

Brünger,A.T. *et al.* (1998) *Acta Crystallogr., Sect. D*, **54**, 905–921.
Byers,Y.S., Levin,A.S., Malvino,A., Waites,L., Robins,R.A. and Baldwin,R.W. (1994) *AIDS Res. Hum. Retrovir.*, **10**, 413–420.
Endo,Y. and Tsurugi,K. (1987) *J. Biol. Chem.*, **262**, 8128–8130.
Endo,Y., Mitsui,K., Mutizuki,M. and Tsurugi,K. (1987) *J. Biol. Chem.*, **262**, 5908–5912.
Gao,B., Ma,X.Q., Wang,Y.P., Chen,S.Z., Wu,S. and Dong,Y.C. (1993) *Sci. Chin., Ser. B*, **37**, 59–73.
Huang,Q., Liu,S., Tang,Y., Jin,S. and Wang,Y. (1995) *Biochem. J.*, **309**, 285–298.
Jones,T.A., Zou,J.-Y., Cowan,S.W. and Kjeldgaard,M. (1991) *Acta Crystallogr., Sect. A*, **47**, 110–119.
Kahn,J.O., Gorelick,K.J., Gatti,G., Arri,C.J., Lifson,J.D., Gambertoglio,J.G., Bostrom,A. and Williams,R. (1994) *Antimicrob. Agents Chemother.*, **38**, 260–267.
Laskowski,R.A., MacArthur,M.W., Moos,D.S. and Thornton,J.M. (1993) *J. Appl. Crystallogr.*, **26**, 283–291.
Li,H.G., Xu,S.Z., Wu,S., Yan,L., Li,J.H., Wong,R.N.S., Shi,Q.L. and Dong,Y.C. (1999a) *Protein Eng.*, **12**, 999–1004.
Li,J.H., Wu,S., Yao,H.P., Shaw,P.C. and Dong,Y.C. (1999b) *Chin. J. Biochem. Mol. Biol.*, **15**, 943–947.
Otwinowski,Z.and Minor,W. (1997) *Methods Enzymol.*, **276**, 307–326.
Ren,J.S., Wang,Y.P., Dong,Y.C., Stuart,D. (1994) *Structure*, **2**, 7–16.
Shaw,P.C., Chan,W.L., Yeung,H.W. and Ng,T.B. (1994) *Life Sci.*, **55**, 253–262.
Stirpe,F., Bailey,S., Miller,S.P. and Bodley,J.W. (1988) *Nucleic Acids Res.*, **16**, 1349–1357.
Wu,S., Lu,X.H., Zhu,Y.R., Yang,J. and Dong,Y.C. (1998) *Sci. Chin., Ser.C*, **41**, 174–180.
Xiong,J.P., Xia,Z.X. and Wang,Y. (1994) *Nat. Struct. Biol.*, **1**, 694–700.
Yan,L., Wu,S., Li,H.G., Li,J.H., Wong,R.N.S., Shi,Q.L. and Dong,Y.C. (1999) *Toxicon*, **37**, 961–972.
Yang,J., Wu,S. and Dong,Y.C. (1994) *Acta Biophys. Sin.*, **10**, 537–542.
Zhang,X.J. and Wang,J.H. (1986) *Nature*, **321**, 477–478.

Received December 10, 2002; accepted May 20, 2003