Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and X-ray analysis of the N-terminal core domain of a tumour-associated human DEAD-box RNA helicase, rck/p54

Tsutomu Matsui,[a] Keita Hogetsu,[a] Yukihiro Akao,[b] Masashi Tanaka,[b] Takao Sato,[a] Takashi Kumasaka[a] and Nobuo Tanaka[a]*

[a]Department of Life Science, Tokyo Institute of Technology, 4259 Nagatsuta-cho, Midori-ku, Yokohama 226-8501, Japan, and [b]Department of Genetic Diagnosis, Gifu International Institute of Biotechnology, 1-1 Naka-Fudogaoka, Kakamigahara, Gifu 504-0838, Japan

Correspondence e-mail: ntanaka@bio.titech.ac.jp

Received 11 August 2003
Accepted 22 October 2003

The *RCK* gene was cloned on the basis of the t(11;14)(q23;q32) chromosome translocation observed in human B-cell lymphoma cell line RC-K8. This gene was found to be overexpressed in various kinds of tumours. The gene product, rck/p54, consisting of 472 amino-acid residues with molecular weight 53.2 kDa, belongs to the family of DEAD-box RNA helicases. Its ATP-dependent RNA-unwinding activity toward c-myc RNA molecules *in vitro* has recently been demonstrated. In the present study, limited proteolysis experiments of rck/p54 were used to truncate the N-terminal domain (residues 1–288; 31.8 kDa) of rck/p54, leading to successful crystallization of Nc-rck/p54, *i.e.* the N-terminal core domain (residues 70–288; 24.5 kDa) of rck/p54. Crystals of Nc-rck/p54 were grown to a size suitable for X-ray structure analysis using polyethylene glycol 3350 as the precipitant. The crystal belongs to the orthorhombic space group $P2_12_12_1$, with unit-cell parameters a = 65.5, b = 73.1, c = 84.8 Å, and diffracts X-rays to beyond 2.0 Å resolution.

1. Introduction

Structural changes in RNA play a very important role that is required by all living organisms (reviewed by Tanner & Linder, 2001). DEAD-box RNA-helicase proteins are highly conserved from bacteria and viruses to humans and are involved in all biological processes of transcription, translation, ribosome biogenesis, mRNA splicing, RNA chaperones, RNA maturation and RNA degradation by unwinding and/or rearranging RNA. It is known that DEAD-box RNA helicases possess both ATPase and RNA-helicase activities: an ATP-dependent RNA-helicase activity and an RNA-dependent ATPase activity. Sequence analysis as well as biochemical and genetic experiments have revealed the protein to be composed of two domains with seven to nine conserved motifs (Tanner *et al.*, 2003; Gorbalenya *et al.*, 1989; Pause *et al.*, 1993), some of which play roles in binding NTP and unwinding dsRNA by utilizing the energy of hydrolysis. However, mysteries still abound regarding the reaction and substrate-recognition mechanisms.

Human rck/p54 (472 amino acids; 53.2 kDa) is a member of the DEAD-box RNA-helicase family and is the prototype of the rck/p54 subfamily (Minshall *et al.*, 2001). The *RCK* gene at 11q23, encoding rck/p54, is fused to an immunoglobulin heavy-chain gene upon t(11;14) chromosome translocation, which is observed in about 5% of B-cell lymphomas carrying 14q32 translocations (Akao *et al.*, 1992; Lu & Yunis, 1992). The promoter region and the possible first exon of the *RCK* gene are decapitated from its coding region by the translocation. *RCK* mRNA was found to be ubiquitously expressed. Expression of rck/p54 is very poor in some tissues; however, a significant amount of rck/p54 was found in tumours that originated from these tissues. In some colorectal adenomas, which are premalignant lesions of colon cancer, rck/p54 and c-myc were found to be co-overexpressed (Akao *et al.*, 1995, 1998; Nakagawa *et al.*, 1999; Hashimoto *et al.*, 2001). These findings suggest that rck/p54 may contribute to cell proliferation and/or carcinogenesis at the translation-initiation or RNA-processing step.

eIF4A, eukaryotic initiation factor 4A, has often been studied as a prototype of the DEAD-box RNA-helicase family and functions in the translation-initiation complex. The reaction mechanism of monomer helicases from the inchworm model (Tanner & Linder, 2001) has been advocated for DEAD-box RNA helicase. However, there is almost no structural evidence which can directly demonstrate this model. In the DEAD-box RNA-helicase family, some crystal structures have already been reported (Benz *et al.*, 1999; Caruthers *et al.*, 2000; Johnson & McKay, 1999; Story *et al.*, 2001), but no crystal structures from vertebrates, including humans, have been determined. This paper describes the preliminary results of a crystallographic study that

may elucidate the relationship between the structure and function of human DEAD-box RNA helicase.

2. Experimental procedures and results

2.1. Expression and purification of N-rck/p54

The expression vector was constructed by inserting N-rck/p54, the N-terminal domain of rck/p54 (1–289), into pQE30 (Qiagen). The vector was introduced into *Escherichia coli* strain M15 (Qiagen) and the protein was expressed as a His-tagged fusion protein at the N-terminus. When cultures reached an OD_{600} of 0.5–0.7 at 310 K in LB (Luria–Bertani) broth supplemented with 100 μg ml^{-1} ampicillin and 25 μg ml^{-1} kanamycin, expression of recombinant N-rck/p54 was induced by addition of IPTG (isopropyl-D-thiogalactopyranoside) to a final concentration of 0.1 m*M*; expression took place for 6 h at 303 K. The harvested *E. coli* was lysed by treatment with lysozyme and sonication in a lysis buffer containing 20 m*M* Tris–HCl pH 8.0, 500 m*M* KCl and a protease-inhibitor tablet (Complete EDTA-free Protease Inhibitor Cocktail Tablets, Roche). The resulting cell homogenate was centrifuged at 20 000*g* for 30 min. The supernatant was applied to a Chelating Sepharose Fast Flow column (Amersham Biosciences) immobilized with Co^{2+} and was washed in buffer containing 20 m*M* Tris–HCl pH 6.5, 500 m*M* NaCl, 20 m*M* imidazole. The column was eluted with elution buffer (20 m*M* Tris–HCl pH 8.0, 500 m*M* NaCl and 400 m*M* imidazole). To remove minor impurities, fractions containing N-rck/p54 were combined and dialyzed against 20 m*M* Tris–HCl pH 8.0, 100 m*M* KCl and applied to a TSK-gel SuperQ-Toyoparl (Tohso) column equilibrated with the same buffer. The flow-through containing N-rck/p54 was collected.

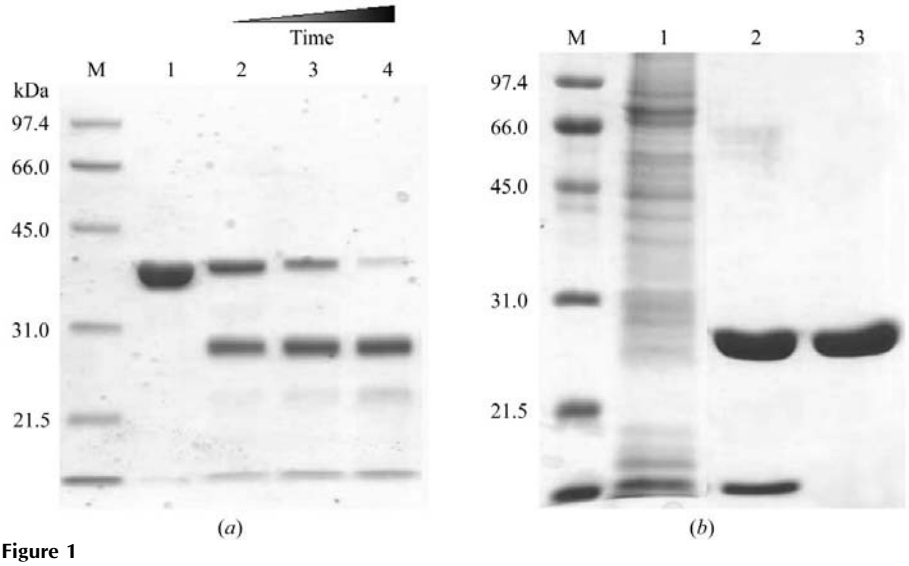

Figure 1
Coomassie-stained SDS–PAGE analysis showing the pattern of trypsin hydrolysis of N-rck/p54 (*a*) and the purification of Nc-rck/p54 (*b*). (*a*) Lane M, molecular-weight markers; lane 1, native N-rck/p54; lanes 2, 3 and 4, time courses of cleavage of 40 μ*M* N-rck/p54 by 380 n*M* trypsin: 1.5, 2 and 4 min, respectively. The major band of about 25 kDa was analyzed by MALDI–TOF mass spectrometry. (*b*) Lane M, molecular-weight markers; lane 1, supernatant after cell lysis; lane 2, protein obtained following affinity chromatography; lane 3, final purified Nc-rck/p54 obtained following gel-filtration chromatography.

```
N-rck/p54   MGLSSQNGQLRGPVKPTGGPGGGGTQTQQQMNQLKNTNTINNGTQQQAQSMTTTIKPGDD 60
A-eIF4A     ------------------------------------------------------SEG--

                        *
N-rck/p54   WKKTLKLPPKDLRIKTSDVTSTKGNEFEDYCLKRELLMGIFEMGWEKPSPIQEESIPIAL 120
A-eIF4A     --ITDIEESQIQTNYDKVVY-----KFDDMELDENLLRGVFGYGFEEPSAIQQRAIMPII

N-rck/p54   SGRDILARAKNGTGKSGAYLIPLLERLDLKKDNIQAMVIVPTRELALQVSQICIQVSKHM 180
A-eIF4A      EGHDVLAQAQSGTGKTGTFSIAALQRIDTSVKAPQALMLAPTRELALQIQKVVMALAFHM

N-rck/p54   GGAKVMATTGGTNLRDDIMRLDDTVHVVIATPGRILDLIKKGVAKVDHVQMIVLDEADKL 240
A-eIF4A     D-IKVHACIGGTSFVEDAEGLRD-AQIVVGTPGRVFDNIQRRRFRTDKIKMFILDEADEM

N-rck/p54   LSQDFVQIMEDIILTLPKNRQILLYSATFPLSVQKFMNSHLQKPYEIN---------- 289
A-eIF4A     LSSGFKEQIYQIFTLLPPTTQVVLLSATMPNDVLEVTTKFMRNPVRILVKKDELTLE
```

Figure 2
Sequence alignment of the N-terminal domain of rck/p54 (N-rck/p54) and the ATPase-domain region of yeast eIF4A (A-eIF4A). The cleavage site by trypsin is shown with an asterisk. The shaded residues are identical in Nc-rck/p54. The underlined residues show the disordered regions in the crystal structure of the ATPase-domain region of eIF4A (PDB code 1qde; Benz *et al.*, 1999). Sequences were aligned using the program *CLUSTAL_W* (Thomson *et al.*, 1994).

2.2. Limited proteolysis experiment

To determine the core-domain structure of N-rck/p56, a limited proteolysis experiment was performed by trypsin treatment and mass spectrometry (Lorsch & Herschlag, 1998). N-rck/p54 (40 μ*M*) was digested with 380 n*M* bovine pancreatic trypsin (Sigma) in a reaction buffer consisting of 20 m*M* Tris–HCl pH 7.8 and 100 m*M* KCl at 310 K. The proteolysis reactions were stopped at various times by the addition of SDS (sodium dodecyl sulfate). The sample solution was applied to SDS–PAGE (10% polyacrylamide gel). After Coomassie blue staining (Fig. 1*a*), the major fragment of approximately 25 kDa was collected and its molecular mass was measured by MALDI–TOF mass spectrometry. One of the cleavage sites was also determined to be at Lys70–Asp71 (Fig. 2).

2.3. Cloning, expression and purification of Nc-rck/p54

The N-terminal core domain of rck/p54 (70–288), Nc-rck/p54, was also synthesized from the vector pQE30/*RCK* containing full-length rck/p54. Both the recognition site for the restriction endonuclease *Sap*I and the insertion of ATG after the *Sap*I site were incorporated into the forward primer (45 bp; 5′-GGTGGTT**GCTCTTC**CAACATGAA-GGATCTAAGAATCAAAACTTCG-3′). The downstream reverse primer contained the recognition site for the restriction endonuclease *Pst*I (45 bp; 5′-CCG**CTG-CAG**TTAGTTAATCTCATAGGGTTTC-TGC-3′). The PCR (polymerase chain reaction) product was ligated into the *Sap*I/*Pst*I sites of the vector pTYB11 of the IMPACT T7 system (New England BioLabs) as the

N-terminal segment fused to the intein and chitin-binding domain unit. The sequence of Nc-rck/p54 was confirmed. The recombinant vector pTYB11-Nc-rck/p54 was introduced into *E. coli* strain ER2566 (New England BioLabs) and cultures were grown to an OD_{600} of 0.5–0.7 at 310 K in LB broth supplemented with 100 µg ml^{-1} ampicillin. The expression of recombinant Nc-rck/p54 was then induced by the addition of IPTG to a final concentration of 0.5 m*M* at 298 K overnight. The harvested *E. coli* was lysed by treatments with both lysozyme and sonication in a lysis buffer containing 20 m*M* HEPES–NaOH pH 8.0, 500 m*M* NaCl, 5 m*M* EDTA and a protease-inhibitor tablet. The resulting cell homogenate was centrifuged at 20 000*g* for 30 min.

The IMPACT (intein-mediated purification with an affinity chitin-binding tag; New England BioLabs) system was employed for protein purification. The supernatant was applied onto a chitin bead column (New England BioLabs) equilibrated with lysis buffer without any protease inhibitors and was washed with the same buffer. On-column self-cleavage was carried out for 40 h using a cleavage buffer consisting of 20 m*M* HEPES–NaOH pH 8.0, 500 m*M* NaCl, 5 m*M* EDTA and 50 m*M* DTT. The fraction containing Nc-rck/p54 was collected and concentrated. Minor impurities were removed by gel-filtration chromatography on a Superdex-16/60 200pg column (Amersham Biosciences) equilibrated in 5 m*M* Tris–HCl pH 7.5, 250 m*M* KCl and 250 m*M* NaCl. The fractions containing Nc-rck/p54 were combined and concentrated to 5 mg ml^{-1} (Fig. 1*b*).

2.4. Crystallization

Crystallization was performed at 298 K using the hanging-drop vapour-diffusion method. The crystallization conditions were determined by a sparse-matrix screen using Crystal Screen and PEG/Ion Screen kits (Hampton Research). Each drop was prepared by mixing 2 µl of protein solution, prepared as described above, with 2 µl precipitant solution. The reservoir precipitant volume was 0.5 ml. The best crystals were obtained from condition No. 36 of PEG/Ion Screen [20%(*w*/*v*) polyethylene glycol 3350, 0.2 *M* sodium tartrate pH 7.2]. Crystals grew to approximately 0.5 × 0.1 × 0.1 mm in 2 d (Fig. 3).

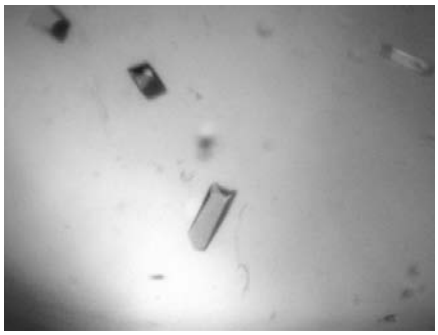

Figure 3
Nc-rck/p54 crystals. Nc-rck/p54 crystals belong to space group $P2_12_12_1$ and were prepared by the hanging-drop vapour-diffusion method in 2 d.

Table 1
Crystallographic data for Nc-rck/p54.

Values in parentheses are for the highest resolution shell.

V_M (Å^3 Da^{-1})	2.1 (2 molecules per AU)
Solvent content (%)	40
Resolution limit (Å)	2.00 (2.08–2.00)
No. reflections	101467 (9845)
No. unique reflections	40891 (3982)
Completeness (%)	99.3 (99.3)
$\langle I/\sigma(I)\rangle$	7.4 (2.1)
R_{merge}†	0.073 (0.263)
Mosaicity (°)	0.39

† $R_{merge} = \sum[|I(h) - I(h)_i/I(h)|]$, where $I(h)$ is the mean intensity after rejections.

2.5. Data collection and molecular replacement

Diffraction data were collected at cryogenic temperature (100 K) using the flash-freezing method with an R-AXIS IV imaging-plate system and a Rigaku RU-300B rotating-anode generator operating at 50 kV and 80 mA. The cryoprotectant solution was prepared by adding 16.7%(*v*/*v*) glycerol to the crystallization reservoir. Diffraction intensities were integrated, merged and scaled using *d*TREK* (Pflugrath, 1999) from *CrystalClear* (Rigaku/MSC). The data-collection statistics are given in Table 1. The crystals belong to the orthorhombic space group $P2_12_12_1$, with unit-cell parameters a = 65.5, b = 73.1, c = 84.8 Å and diffract X-rays beyond 2.0 Å resolution (Table 1).

The structure was determined by molecular replacement using the *CCP*4 program *MOLREP* (Collaborative Computational Project, Number 4, 1994; Vagin & Teplyakov, 1998) using the ATPase domain of translation-initiation factor 4A (eIF4A) from *Saccharomyces cerevisiae* (PDB code 1qde; Benz *et al.*, 1999) as the search model. A solution was found with a correlation coefficient of 0.373 (R factor = 0.518) and two molecules were contained in the asymmetric unit. Details of the structure will be discussed in a subsequent paper which is now being prepared.

3. Discussion

Because rck/p54 and N-rck/p54 did not crystallize to give crystals of a suitable size for X-ray analysis, limited proteolysis experiments of N-rck/p54 were examined in order to design an optimal truncation of rck/p54 for crystallization. Electrophoretic analysis of His-tagged rck/p54 after several weeks of storage showed a bundle of bands corresponding to molecular weights of between approximately 54 and 44 kDa (data not shown). One band was from His-tagged rck/p54 and the others were from degraded His-tagged rck/p54, which did not bind to the cobalt column. This observation implies that the N-terminal region of rck/p54 is easy to degrade.

Limited proteolysis coupled with MALDI–TOF mass spectroscopy were carried out in order to elucidate the cleavage site of N-rck/p54. Cleavage between Lys70 and Asp71 was observed in the major proteolysis product (Fig. 1*a*, lanes 2–4). The sequence alignment in Fig. 2 shows an extra N-terminal region in comparison with eIF4A, which is known to be the smallest DEAD-box protein shown to have unwinding activity *in vitro* (Rogers *et al.*, 2001). Previous electron-microscopy data (Akao *et al.*, 2003), together with our limited proteolysis data, support the occurrence of an additional N-terminal structural domain.

In the present study, Nc-rck/p54, which possesses all the conserved sequence motifs of the N-terminal ATPase domain, was designed by removing the N-terminal residues to Pro69 from N-rck/p54 for crystallization. The well shaped crystals, which are the first crystal samples of human DEAD-box RNA helicase, were grown to a size suitable for X-ray analysis. Further knowledge of the involvement of rck/p54 in the carcinogenesis and/or human translation-initiation step may be obtained from the crystal structure of Nc-rck/p54.

We thank K. Kakinuma and E. Nango (Tokyo Institute of Technology) for their help with the MALDI–TOF mass-spectrometry measurements. This study was supported in part by a grant from the National Project on Protein Structural and Functional Analyses, the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

Akao, Y., Marukawa, O., Morikawa, H., Nakao, K., Kamei, M., Hachiya, T. & Tsujimoto, Y. (1995). *Cancer Res.* **55**, 3444–3449.

Akao, Y., Mizoguchi, H., Ohishi, N. & Yagi, K. (1998). *FEBS Lett.* **16**, 429279–429283.
Akao, Y., Seto, M., Yamamoto, K., Iida, S., Nakazawa, S., Inazawa, J., Abe, T., Takahashi, T. & Ueda, R. (1992). *Cancer Res.* **52**, 6083–6087.
Akao, Y., Yoshida, H., Matsumoto, K., Matsui, T., Hogetsu, K., Tanaka, N. & Usukura, J. (2003). *Genes Cells*, **8**, 671–676.
Benz, J., Trachsel, H. & Baumann, U. (1999). *Structure*, **7**, 671–679.
Caruthers, J. M., Johnson, E. R. & McKay, D. B. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 13080–13085.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Gorbalenya, A. E., Koonin, E. V., Donchenko, A. P. & Blinov, V. M. (1989). *Nucleic Acids Res.* **26**, 4713–4730.
Hashimoto, K., Nakagawa, Y., Morikawa, H., Niki, M., Egashira, Y., Hirata, I., Katsu, K. & Akao, Y. (2001). *Carcinogenesis*, **22**, 1965–70.
Johnson, E. R. & McKay, D. B. (1999). *RNA*, **5**, 1526–1534.
Lorsch, J. R. & Herschlag, D. (1998). *Biochemistry*, **24**, 2194–2206.
Lu, D. & Yunis, J. J. (1992). *Nucleic Acids Res.* **20**, 1967–1972.
Minshall, N., Thom, G. & Standart, N. (2001). *RNA*, **7**, 1728–1742.
Nakagawa, Y., Morikawa H., Hirata, I., Shiozaki, M., Matsumoto, A., Maemura, K., Nishikawa, T., Niki, M., Tanigawa, N., Ikegami, M., Katsu, K. & Akao, Y. (1999). *Br. J. Cancer*, **80**, 914–917.
Pause, A., Methot, N. & Sonenberg, N. (1993). *Mol. Cell Biol.* **13**, 6789–6798.
Pflugrath, J. W. (1999). *Acta Cryst.* D**55**, 1718–1725.
Rogers, G. W. J., Lima, W. F. & Merrick, W. C. (2001). *J. Biol. Chem.* **276**, 12598–12608.
Story, R. M., Li, H. & Abelson, J. N. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 1465–1470.
Tanner, N. K., Cordin, O., Banroques, J., Doere, M. & Linder, P. (2003). *Mol. Cell*, **11**, 127–138.
Tanner, N. K. & Linder, P. (2001). *Mol. Cell*, **8**, 251–262.
Thomson, J. D., Higgins, D. G. & Gibson, T. J. (1994). *Nucleic Acids Res.* **22**, 4673–4680.
Vagin, A. & Teplyakov, A. (1998). *Acta Cryst.* D**54**, 400–402.