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

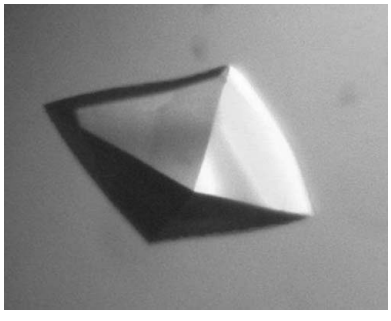
Crystallization and preliminary X-ray analysis of the RAD protein from *Antirrhinum majus*

Clare E. M. Stevenson,[a] Nicolas Burton,[a] Manuela Costa,[b] Utpal Nath,[b] Ray A. Dixon,[c] Enrico S. Coen[b] and David M. Lawson[a]*

[a]Department of Biological Chemistry, John Innes Centre, Norwich NR4 7UH, England, [b]Department of Cell and Developmental Biology, John Innes Centre, Norwich NR4 7UH, England, and [c]Department of Molecular Microbiology, John Innes Centre, Norwich NR4 7UH, England

Correspondence e-mail: david.lawson@bbsrc.ac.uk

Received 30 June 2005
Accepted 26 August 2005
Online 13 September 2005



© 2005 International Union of Crystallography
All rights reserved

Crystals of the RADIALIS protein from *Antirrhinum majus* were grown by vapour diffusion after limited proteolysis. Mass spectrometry indicated that an 8 kDa fragment had been crystallized corresponding to the predicted MYB DNA-binding domain. X-ray data collected at room temperature were consistent with tetragonal symmetry, whereas data collected at 100 K using crystals cryoprotected by supplementing the mother liquor with ethylene glycol conformed to orthorhombic symmetry. It was subsequently shown that crystals soaked in cryoprotectants that were 'osmolality-matched' to the mother liquor retained tetragonal symmetry. Using these crystals, X-ray data were collected in-house to a maximum resolution of 2 Å.

1. Introduction

Wild-type flowers of *Antirrhinum majus* (garden snapdragon) are markedly asymmetric along their dorsoventral axis and this asymmetry is particularly pronounced in the petals (Coen, 1996). A number of mutants have been identified that give rise to abnormal radially symmetric flowers. Amongst these are mutations in the *RADIALIS* (*RAD*) gene that encodes a 93-residue protein (RAD) with a calculated molecular weight of 10 774 Da. Sequence analysis suggests that RAD belongs to the MYB family of transcription factors, as it contains one copy of a putative MYB DNA-binding domain spanning roughly the N-terminal two-thirds of the protein. No other functional domains or motifs can be recognized in the remaining sequence (Corley *et al.*, 2005). MYB transcription factors are widespread in eukaryotes, particularly in plants (over 125 have been identified in *Arabidopsis* alone; Stracke *et al.*, 2001), and they have roles in a diverse range of signalling processes. It has been shown that the *RAD* gene is required for establishing the asymmetry of the flower and is expressed specifically in the dorsal region of the early floral meristem (Corley *et al.*, 2005).

Despite their ubiquity, there are currently very few structures in the current Protein Data Bank (PDB; Bernstein *et al.*, 1977) that are annotated as MYB proteins: a full text search of the entire database using the keyword 'MYB' finds a representative set of ten MYB structures (variants of the same structure were excluded using a 90% sequence-identity cutoff) and only two of these are X-ray structures (PDB codes 1gvd and 1h6p). Fold recognition by the *FUGUE* server (http://www-cryst.bioc.cam.ac.uk/~fugue/prfsearch.html; Shi *et al.*, 2001) using the native RAD sequence finds two further X-ray structures (PDB codes 1ign and 1xc5 with Z scores of 6.45 and 5.94, respectively). However, none of these four structures share greater than 15% sequence identity with RAD and thus do not represent suitable templates for molecular replacement.

We report here the crystallization and preliminary X-ray analysis of the RAD protein as a step towards elucidating the molecular details of its biological activity.

2. Materials and methods

2.1. Protein expression and purification

The *A. majus RADIALIS* gene (GenBank accession No. AY954971) was amplified by PCR from previously isolated genomic

DNA (Coen *et al.*, 1986) and cloned into the pRSET A vector (Invitrogen) to give a plasmid encoding a polypeptide with an N-terminal hexahistidine tag. This added a further 36 residues to the native protein (with sequence MRGSHHHHHHGMASMTGGQQ-MGRDLYDDDDKDRWGS), giving a total deduced molecular weight of 14 904 Da. This plasmid was transformed into *Escherichia coli* strain BL21 (DE3) plys E SBET (Studier & Moffatt, 1986; Schenk *et al.*, 1995) and the cells were grown in 1 l Luria–Bertani medium containing 200 µg ml^{-1} ampicillin, 30 µg ml^{-1} chloramphenicol and 30 µg ml^{-1} kanamycin at 310 K for approximately 5 h. Protein overproduction was induced at an OD_{600} of 0.5 by the addition of 1 m*M* isopropyl β-D-thiogalactopyranoside and was continued overnight. Cells were harvested by centrifugation and stored at 253 K until required. The cell pellets were subsequently resuspended in 50 m*M* Tris–HCl pH 8.0 and 500 m*M* NaCl (buffer *A*) containing 20 µg deoxyribonuclease I (Sigma) per millilitre of resuspended cells. The cells were then lysed by passage through a French press at 6.9 MPa. The cell lysate obtained by centrifugation at 30 600*g* for 10 min was loaded onto an Ni^{2+}-charged Hi-trap metal-chelation column (Amersham Biosciences) and unbound proteins were removed by washing with buffer *A*. His-tagged RAD bound to the column was eluted with an imidazole gradient in buffer *A*. The RAD protein eluted at approximately 0.2 *M* imidazole and the fractions containing the His-tagged RAD, as judged by SDS–PAGE, were pooled and dialysed into buffer *A*. All subsequent protein-concentration steps were performed using Ultrafree 10 kDa cutoff centrifugal concentrators (Millipore).

2.2. Tag removal and further proteolytic cleavage

In order to generate stable fragments of RAD for crystallization, the purified protein was subjected to limited proteolytic digestion. The N-terminal His tag of RAD was removed enzymatically using dipeptidyl aminopeptidase I (DAPase, Qiagen). This is an exopeptidase that cleaves dipeptides sequentially from the N-terminus of a protein until a stop point is reached. The first potential stop point in the His-tagged RAD is located 22 residues into the tag sequence before an Arg residue and would leave 14 residues of linker attached to the native RAD sequence, giving a fragment with a deduced molecular weight of 12 512 Da. Approximately 1 ml of purified RAD protein (5–10 mg ml^{-1}) was dialysed into 20 m*M* MOPS buffer pH 6.9 containing 100 m*M* NaCl (buffer *B*). For each 5 mg of His-tagged RAD to be cleaved, 30 µl DAPase (10 U ml^{-1}) was activated by the addition of 270 µl buffer *B*, 60 µl cysteamine–HCl and 540 µl water.

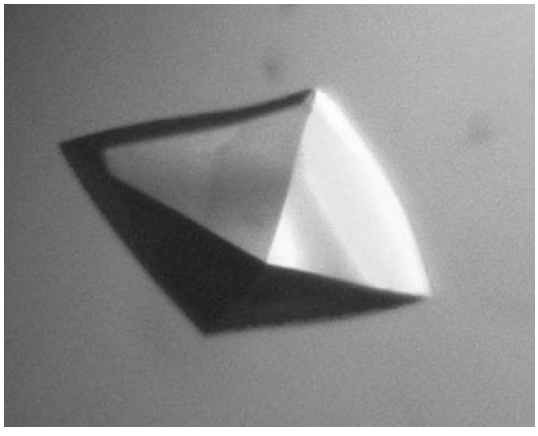


Figure 1
Crystal of the *A. majus* RAD protein with approximate dimensions of 250 × 150 × 150 µm.

After 5 min at 293 K, the DAPase solution was added to the RAD sample and incubated at 293 K for approximately 2 h. The DAPase enzyme contains a C-terminal His tag that cannot be self-cleaved. Thus, tag-free protein could be separated from the DAPase and any remaining uncleaved protein by running a further nickel-affinity step. In this case, the required sample eluted in the flowthrough.

Subsequently, the DAPase-treated protein was subjected to further proteolysis using limited digestion with trypsin (Promega). The trypsin was dissolved in trypsin resuspension buffer (Promega) to a concentration of 0.2 µg µl^{-1}. For each 50 µl of RAD protein at approximately 10 mg ml^{-1} in buffer *A*, 1 µl of trypsin (0.2 µg) was added. This was left for 5 min at 293 K before adding 1 µl (0.5 µg) of the trypsin inhibitor (Sigma). The sample was then used directly for crystallization without further purification.

2.3. Crystallization and X-ray diffraction analysis

Crystallization trials were performed on the His-tagged protein, the DAPase-treated sample and the sample after both DAPase and trypsin treatment. Trials were performed by vapour diffusion in sitting drops using CrystalClear strips (Hampton Research) at constant temperatures of 291 and 277 K. Drops consisted of 1 µl protein and 1 µl well solution with a well volume of 100 µl. In all cases the protein was at a concentration of approximately 10 mg ml^{-1} in buffer *A*. Initial crystallization conditions were sought using a variety of commercially available and in-house screens. Conditions were subsequently optimized and adapted to hanging-drop format using 24-well VDX plates (Hampton Research). In this case the well volume was 1 ml and the protein:precipitant ratio was either 2 µl:1 µl, 1 µl:1 µl or 1 µl:2 µl.

Data collection at room temperature (293 K) was used to establish the diffraction quality of untreated crystals. For this, a crystal was mounted in a 0.7 mm diameter glass capillary (Muller Glas). X-ray data were collected in-house using a MAR345 image-plate detector (X-ray Research) mounted on a Rigaku RU-H3RHB rotating-anode X-ray generator (operated at 50 kV and 100 mA) fitted with Osmic confocal optics and a copper target (Cu $K\alpha$; λ = 1.542 Å). The data were processed using the *HKL* software package (Otwinowski & Minor, 1997); all other downstream data processing and statistical analysis was carried out using programs from the *CCP*4 software suite (Collaborative Computational Project, Number 4, 1994).

For cryogenic data collection, crystals were initially cryoprotected by a short soak (less than 1 min) in mother liquor containing ethylene glycol. In subsequent experiments, the osmolality of the cryoprotectant was matched to that of the mother liquor, as recommended by Garman (1999), in order to minimize the osmotic shock upon transfer of the crystal from one solution to the other. The osmolality of ammonium sulfate and ethylene glycol in the cryoprotectants were estimated from standard tables in Weast (1988–1989). The crystals were then mounted in cryoloops (Hampton Research) and flash-cooled to 100 K in a stream of gaseous nitrogen produced by an X-Stream cryocooler (Rigaku-MSC). Diffraction data were recorded and processed as for the capillary-mounted crystal.

3. Results and discussion

His-tagged RAD was overexpressed and purified with an approximate yield of 10 mg of protein from 1 l of culture and was judged to be greater than 90% pure from SDS–PAGE analysis. Extensive crystallization trials with the His-tagged RAD failed to yield any promising results and the same was true for initial experiments with the DAPase-treated material. However, some 3–4 weeks after setting

Table 1
MALDI–TOF analysis of crystallized RAD fragment.

Weight from MALDI–TOF (Da)	Predicted C-terminus	Weight calculated from sequence (Da)
7659.5	...PN	7661.6
7823.7	...PNY	7824.8
7979.6†	...PNYR	7981.0
8083.9	...PNYRT	8082.0
8183.9	...PNYRTT	8183.2

† Major peak.

up crystallization trials with the DAPase-treated protein, small crystals began to appear under several conditions containing ammonium sulfate at high pH. These could be optimized from samples stored for extended periods at 277 K in the absence of protease inhibitors, whereupon improved crystals were grown from 2.8 *M* ammonium sulfate in 100 m*M* CHES pH 9.5 using drops comprised of 2 µl protein solution and 1 µl precipitant solution. These took up to 7 d to reach maximum dimensions of 250 × 150 × 150 µm (see Fig. 1). Analysis of dissolved crystals using SDS–PAGE indicated that the DAPase-treated material had further proteolysed to a fragment of approximate molecular weight 8 kDa. In order to identify the crystallized fragment, crystals were first washed in the crystallization well solution and then dissolved in Milli-Q water for proteomic analysis. N-terminal sequencing using a Procise model 491 protein sequencer (Applied Biosystems) unambiguously gave the sequence GSGRP that was consistent with residues 6–10 of the native sequence. Mass-spectrometric analysis using a Reflex III MALDI-ToF mass spectrometer (Bruker Daltonics Ltd) gave a value of 7979.6 Da as the major peak, corresponding closely to the calculated weight of 7981.0 Da for residues 6–74 inclusive of the native sequence (hereafter referred to as the 8 kDa fragment). Several minor peaks were also present and these corresponded closely to the predicted weights of the same fragment truncated or extended by one or two amino acids at the C-terminus (see Table 1). Thus, the resultant fragment had been cleaved at both termini and contained no residual residues from the His tag. According to the PFAM database (http://www.sanger.ac.uk/Software/Pfam; Bateman *et al.*, 2002) the Myb-like DNA-binding domain of RAD (PF00249) comprises residues 8–57 and therefore remains intact in the crystallized 8 kDa fragment.

It was subsequently established that the 8 kDa fragment could be generated from freshly prepared DAPase-treated samples by further digestion with trypsin. This was not unexpected, as both of the cleavage sites were predicted to be preferentially recognized by trypsin, being Arg-Gly and Arg-Thr for the N- and C-terminal sites, respectively. SDS–PAGE analysis indicated that all the starting material had been cleaved and that an additional major band was present with a molecular weight roughly intermediate between the starting material (with predicted molecular weight 12.5 kDa) and the desired 8 kDa fragment. The presence of this additional band did not appear to inhibit crystallization, since crystals were readily obtained under the same conditions as before. Several attempts were made to generate the 8 kDa fragment directly from the full-length His-tagged protein by trypsin treatment alone, but unfortunately this yielded very little of the required fragment. Moreover, since the DAPase/trypsin protocol outlined above was reproducible, no attempt was made to directly clone the 8 kDa fragment for crystallization.

The ambient temperature data were recorded from a single crystal to a maximum resolution of 2.9 Å. The symmetry was established as primitive tetragonal, with approximate unit-cell parameters $a = b = 45$, $c = 72$ Å, and the data could be processed satisfactorily in space group *P*422 with an overall $R_{\rm merge}$ of 0.088 (see Table 2). From the inspection of systematic absences in pseudo-precession plots of the reciprocal lattice (using the program *HKLVIEW*), the space group was subsequently assigned as either $P4_12_12$ or $P4_32_12$. Analysis of the contents of the asymmetric unit based on a single copy of the 8 kDa fragment gave a crystal-packing parameter ($V_{\rm M}$) of 2.3 Å^3 Da^{-1}, with a corresponding solvent content of 46% (Matthews, 1968).

Table 2
Summary of X-ray data for RAD.

Values in parentheses indicate the figures for the outer resolution shell.

	Room temperature	Cryoprotected (non-matched)	Cryoprotected (osmolality-matched)
Resolution range (Å)	30–2.9	40–2.35	40–2.02
Space group	$P4_12_12/P4_32_12$	$P2_12_12_1$	$P4_12_12/P4_32_12$
Unit-cell parameters (Å)	$a = b = 45.07$, $c = 72.14$	$a = 44.33$, $b = 44.90$, $c = 70.59$	$a = b = 44.63$, $c = 71.90$
Estimated mosaicity (°)	0.65	1.48	1.02
Unique reflections	1871	6139	5182
Redundancy	6.2	3.5	12.5
Completeness	99.7 (99.5)	97.7 (94.1)	99.9 (99.8)
Wilson *B* value (Å^2)	54.0	32.5	27.5
$R_{\rm merge}$	0.088 (0.221)	0.084 (0.210)	0.058 (0.236)
$\langle I/\sigma(I)\rangle$	20.0 (5.8)	13.5 (4.1)	43.1 (7.9)

For cryogenic data collection, a crystal was cryoprotected by adding 20%(*v*/*v*) ethylene glycol to the mother liquor in place of an equivalent volume of buffer. X-ray data were then recorded from a single crystal to a maximum resolution of 2.35 Å. These data could be indexed in a tetragonal lattice, but the merging statistics were poor (overall $R_{\rm merge}$ = 0.163 when processed as *P*4; other statistics not shown). Reprocessing the data in the orthorhombic space group *P*222 gave a much improved overall $R_{\rm merge}$ of 0.084 (see Table 2). Moreover, systematic absences were indicative of space group $P2_12_12_1$: in particular, reflections satisfying the condition $l = 2n$ were clearly present along the *l* axis, being consistent with a 2_1 but not a 4_1 screw axis. Independent refinement of the unit-cell parameters gave *a* and *b* cell edges that differed by less than 1 Å.

In the mother liquor, 2.8 *M* ammonium sulfate contributes 5.3 osmol kg^{-1} to the overall osmolality. However, 20%(*v*/*v*) ethylene glycol is required for cryoprotection, having an osmolality of 4.3 osmol kg^{-1}. In order to match the osmolality of the two solutions, the ammonium sulfate contribution needs to be reduced to 1.0 osmol kg^{-1} in the cryoprotectant, corresponding to a concentration of 0.52 *M*. Thus, the final composition of the osmolality-matched cryoprotectant was 0.52 *M* ammonium sulfate, 20%(*v*/*v*) ethylene glycol in 100 m*M* CHES pH 9.5. After a short soak (less than 1 min) in this solution, data were collected from a single crystal to a maximum resolution of 2 Å at 100 K. Although the diffraction pattern was contaminated by diffuse rings characteristic of hexagonal ice (at resolutions of 3.90, 3.67 and 2.25 Å), this was not a serious problem as the data were strong and clearly superior to anything collected thus far: data reduction gave an overall $R_{\rm merge}$ of 0.058 (see Table 2).

Although data collection at room temperature was not repeated, several data sets were subsequently recorded at 100 K using both osmolality-matched and non-matched cryoprotectants with reproducible results: the osmolality-matched crystals yielded data that were consistent with tetragonal symmetry, whereas the non-matched crystals gave data that conformed to orthorhombic symmetry (data not shown). It was also noted that whilst the mosaicity tended to increase as the result of cryocooling, which is a common observation (Garman, 1999), it was significantly higher for non-matched crystals.

In this paper we demonstrate that limited proteolysis remains a valuable technique in the crystallographers toolkit, enabling the generation of truncated versions of the target protein that are

presumably more rigid and therefore more amenable to crystallization (reviewed recently by Derewenda, 2004). Nevertheless, one should always consider that the resulting fragments may not be representative of the intact molecule. Indeed, in the absence of a suitable assay, we have been unable to establish whether the truncated RAD protein retains wild-type properties. In addition, we show that the composition of the cryoprotectant can have profound effects on the resultant X-ray data. Specifically, we conclude that crystals of the 8 kDa RAD fragment grow with tetragonal symmetry and this is retained when crystals are soaked in osmolality-matched cryoprotectants. By contrast, soaking in non-matched cryoprotectants disrupts the crystal lattice and reduces the symmetry to orthorhombic. In the absence of suitable search models for molecular replacement, we will need to solve the RAD structure by isomorphous replacement methods.

This work was funded by the BBSRC. We would like to thank I. Usón for assistance with data analysis, and M. Naldrett and A. Bottrill for MALDI–TOF and N-terminal sequence analysis.

References

Bateman, A., Birney, E., Cerruti, L., Durbin, R., Etwiller, L., Eddy, S. R., Griffiths-Jones, S., Howe, K. L., Marshall, M. & Sonnhammer, E. L. L. (2002). *Nucleic Acids Res.* **30**, 276–280.
Bernstein, F. C., Koetzle, T. F., Williams, G. J., Meyer, E. F. Jr, Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977). *J. Mol. Biol.* **112**, 535–542.
Coen, E. S. (1996). *EMBO J.* **15**, 6777–6788.
Coen, E. S., Carpenter, R. & Martin, C. (1986). *Cell*, **47**, 285–296.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Corley, S. B., Carpenter, R., Copsey, L. & Coen, E. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 5068–5073.
Derewenda, Z. S. (2004). *Methods*, **34**, 354–363.
Garman, E. (1999). *Acta Cryst.* D**55**, 1641–1653.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Schenk, P. M., Baumann, S., Mattes, R. & Steinbiss, H. H. (1995). *Biotechniques*, **19**, 196–200.
Shi, J., Blundell, T. L. & Mizuguchi, K. (2001). *J. Mol. Biol.* **310**, 243–257.
Stracke, R., Werber, M. & Weisshaar, B. (2001). *Curr. Opin. Plant Biol.* **4**, 447–456.
Studier, F. W. & Moffatt, B. A. (1986). *J. Mol. Biol.* **189**, 113–130.
Weast, R. C. (1988–1989). Editor. *Handbook of Chemistry & Physics*, 69th ed. Boca Raton, Florida: CRC Press.