Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structural features of the single-stranded DNA-binding protein MoSub1 from *Magnaporthe oryzae*

Jinguang Huang,[a,b,c] Yanxiang Zhao,[b] Dan Huang,[b] Huaian Liu,[b] Neil Justin,[d] Wensheng Zhao,[a,b] Junfeng Liu[b]* and Youliang Peng[a,b]*

[a]State Key Laboratory of Agrobiotechnology, China Agricultural University, No. 2 Yunamingyuanxilu, Beijing 100193, People's Republic of China, [b]MOA Key Laboratory of Plant Pathology, China Agricultural University, No. 2 Yuanmingyuanxilu, Beijing 100193, People's Republic of China, [c]College of Agronomy and Plant Protection, Qingdao Agricultural University, Qingdao, Shandong 266109, People's Republic of China, and [d]Molecular Structure, MRC–NIMR, The Ridgeway, London NW7 1AA, England

Correspondence e-mail: jliu@cau.edu.cn, pengyl@cau.edu.cn

Received 21 March 2012
Accepted 3 May 2012

PDB Reference: MoSub1, 4agh.

The well studied general transcription cofactor Sub1/PC4 has multiple functions in transcription. It plays both a negative and a positive role in transcription initiation and is involved in elongation and downstream transcription processes and as a transcription reinitiation factor. MoSub1, a Sub1/PC4 orthologue from rice blast fungus, binds the single-stranded DNA dT_{12} tightly with an affinity of 186 n*M*. The crystal structure of MoSub1 has been solved to 1.79 Å resolution. The structure of the protein shows high similiarity to the structure of PC4 and it has a similar dimer interface and DNA-binding region to PC4, indicating that MoSub1 could bind DNA using the same motif as other proteins of the Sub1/PC4 family. There are two novel features in the MoSub1 structure: a region N-terminal to the DNA-binding domain and a C-terminal extension. The region N-terminal to the DNA-binding domain of MoSub1 turns back towards the DNA-binding site and may interact directly with DNA or the DNA-binding site. The C-terminal extension region, which is absent in PC4, may not be capable of interacting with DNA and is one possible reason for the differences between Sub1 and PC4.

1. Introduction

The yeast transcription cofactor Sub1 (suppressor of TFIIB) or TSP1 (transcriptional stimulatory protein) was originally isolated as a suppressor of a cold-sensitive TFIIB mutant and has multiple functions in transcription: it plays both a negative and a positive role in transcription initiation and is involved in elongation and downstream transcription processes and as a transcription-reinitiation factor (Henry *et al.*, 1996; Knaus *et al.*, 1996; Rosonina *et al.*, 2009; Tavenet *et al.*, 2009; Conesa & Acker, 2010). Unlike other transcription cofactors, Sub1 and its human orthologue PC4 contain neither a histone-modification domain nor a chromatin-remodelling domain. Instead, they contain a conserved DNA-binding domain that binds ssDNA or unpaired dsDNA (Henry *et al.*, 1996; Knaus *et al.*, 1996; Werten, Langen *et al.*, 1998; Werten, Stelzer *et al.*, 1998).

The DNA-binding domain, which is the key domain for transcription, has been characterized using biochemical and structural methods (Brandsen *et al.*, 1997; Werten, Langen *et al.*, 1998; Werten, Stelzer *et al.*, 1998; Werten & Moras, 2006). The structure of the C-terminal amino acids 63–127 of PC4 (PC4-CTD), which possess DNA-binding activity, has been solved alone and in complex with ssDNA (Brandsen *et al.*, 1997; Werten & Moras, 2006). PC4-CTD monomers, which consist of a curved five-stranded antiparallel β-sheet followed by a 45° kinked α-helix, form dimers using the kinked α-helix,

which packs against the β-sheet of their respective partners (Brandsen *et al.*, 1997; Werten & Moras, 2006). The single-stranded oligonucleotide is tightly bound to the concave β-surfaces of the PC4 homodimer as a hairpin-like structure and the ssDNA hairpin is wrapped around the ridge (Werten & Moras, 2006). Several positively charged side chains (such as Arg70, Arg100 and Lys101) directly interact with the DNA backbone by forming hydrogen bonds. Another positively charged residue, Arg86, which is one of the few absolutely conserved residues in PC4, hydrogen bonds to one of the ssDNA phosphate groups through two waters. In addition, the residues Phe77 and Trp89 are engaged in stacking interactions with several DNA residues *via* their aromatic side chains. The aromatic nature of each of these residues is highly conserved among PC4 orthologues, and the mutation of either residue to alanine abolishes ssDNA binding (Werten, Stelzer *et al.*, 1998).

Sub1, which shares very high homology with PC4 in the DNA-binding region, differs from PC4 in at least two regards. Firstly, Sub1/PC4 plays both a negative and a positive role in transcription initiation, however, the stimulatory effect of PC4 requires an additional activator (Malik *et al.*, 1998), whereas Sub1 was found to directly activate Pol II transcription *in vitro* (Henry *et al.*, 1996). Secondly, Sub1 interacts with TFIIB but not with TFIIA, which instead interacts with PC4 (Ge & Roeder, 1994; Knaus *et al.*, 1996). The reasons for these differences are still unknown and one possible reason may be related to the difference in domain structures between them. The DNA-binding domain of Sub1 is in the N-terminus (residues 32–105), while that of the human counterpart PC4 is in the C-terminus (residues 62–127). The N-terminus of PC4 is an unstructured serine-rich or lysine-rich region, while Sub1 has fewer amino-terminal serine residues (five serines in 17 residues of the yeast protein, compared with nine serines in 16 residues and eight serines in 13 residues in the two SEAC domains of the human protein). Sub1 also contains a carboxyl-terminal extension which is absent in PC4. The structure of Sub1 has not yet been solved.

Rice blast fungus (*Magnaporthe oryzae*) is the pathogen that causes rice blast disease, one of the most devastating diseases to affect rice production (Wilson & Talbot, 2009; Valent & Khang, 2010). It is widely recognized as a model organism for the study of the molecular mechanisms of fungal pathogenesis (Ebbole, 2007). Elucidating the mechanism of development and pathogenicity of the fungus will help us to find better ways of controlling the disease. Several transcription factors have been found to be involved in fungal pathogenicity and development (Kim *et al.*, 2009; Nishimura *et al.*, 2009; Guo *et al.*, 2011; Li *et al.*, 2011; Yan *et al.*, 2011; Zhou *et al.*, 2011). *Mosub1* is the *Sub1* homology gene in the fungal genome and is likely to encode one of the general factors that control its development. MoSub1 has a DNA-binding domain in its N-terminus, like Sub1, and could be used as a homology model of Sub1 to study the differences between Sub1 and PC4.

Here, we purified recombinant MoSub1 and analyzed its DNA-binding activity using surface plasmon resonance (SPR). We solved the crystal structure in order to discover the structural basis of its function.

2. Materials and methods

2.1. Overexpression and purification

Recombinant N-terminally His_6-tagged MoSub1 was cloned by PCR-based subcloning using primers complementary to the 5′ and 3′ ends of full-length MoSub1. The amplicons were inserted between the *Nco*I and *Xho*I restriction sites of pHAT2 vector (kindly supplied by Dr Arie Geerlof, EMBO). Chemically competent *Escherichia coli* BL21 (DE3) cells were transformed with this recombinant vector. Cells were grown at 310 K in 2×YT medium containing 100 µg ml^{-1} ampicillin until the OD_{600} reached 0.5. Expression was induced by adding 0.1 m*M* IPTG and took place at 289 K overnight. The cell pellet was collected by centrifugation and the cells were lysed *via* sonication. The recombinant His_6-MoSub1 was purified using Chelating Sepharose Fast Flow Agarose, Resource Q anion-exchange and Superdex 200 10/300 GL gel-filtration chromatography (GE Healthcare). The protein was concentrated to a final concentration of 10 mg ml^{-1} with an Amicon Ultra-15 centrifugal filter with 10 kDa molecular-weight cutoff (Millipore).

2.2. DNA-binding affinity measurements

The specific binding of MoSub1 and DNA molecules was monitored by the surface plasmon resonance (SPR) biosensor technique using a Biacore 3000 instrument (GE Healthcare). Single-stranded oligo(5′-TTTTTTTTTTTT-3′) (dT_{12}) with a 5′-biotin label was captured on an SA sensor chip (30 response units). The running buffer consisted of 20 m*M* Tris pH 8.0, 150 m*M* NaCl, 0.005%(*v*/*v*) Tween 20. A blank flow cell was used as a reference to correct for instrumental and concentration effects. The protein at different concentrations (each concentration in duplicate) was prepared by step dilution in the running buffer and injected onto the DNA surface and blank flow cell for 90 s at a flow rate of 30 µl min^{-1}. The protein was dissociated for 150 s and regenerated for 60 s with 0.05%(*v*/*v*) SDS at the same rate. All assays were performed at 298 K and the data were analyzed with the Biacore 3000 evaluation software. Equilibrium constants were calculated by fitting to a steady-state affinity model.

2.3. Crystallization

Initial MoSub1 crystals were obtained by the sitting-drop vapour-diffusion method using the PEG/Ion crystallization screen (Hampton Research) with an Oryx4 robot (Douglas Instruments Ltd). 0.15 µl protein solution was mixed with 0.15 µl well solution and the mixture was equilibrated against 73 µl reservoir solution. Crystals appeared after about one month using 25% PEG 3350, 0.2 *M* calcium acetate. An optimized condition was obtained by adding trypsin as a seed (Huang *et al.*, 2012) and the final condition consisted of mixing 0.30 µl protein solution (8–10 mg ml^{-1}) in 20 m*M* Tris–HCl pH 8.0, 50 m*M* NaCl with 0.25 µl reservoir solution consisting of 24–26% PEG 3350, 0.01 *M* calcium acetate and 0.02–0.04 µl trypsin (0.1 mg ml^{-1}) and equilibrating at 293 K. The crystals were flash-cooled by adding 20% PEG 400 as a cryoprotectant

Table 1
Data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

Data collection	
Space group	$P6_122$
Unit-cell parameters (Å, °)	$a = b = 60.1$, $c = 114.4$, $\alpha = \beta = 90.0$, $\gamma = 120.0$
Resolution (Å)	57.2–1.79 (1.83–1.79)
R_{merge} (%)	10.7 (68.1)
$\langle I/\sigma(I)\rangle$	16.4 (3.1)
Completeness (%)	99.8 (98.7)
Multiplicity	17.6 (7.8)
Refinement	
Resolution (Å)	47.4–1.79
No. of reflections	12162
R_{work}/R_{free} (%)	18.1/20.8
No. of atoms	
Protein	649
Water	110
B factors (Å^2)	
Protein	32.5
Water	42.2
R.m.s. deviations	
Bond lengths (Å)	0.003
Bond angles (°)	0.802
Ramachandran plot (*MolProbity*)	
Residues in favoured regions (%)	98.7
Residues in allowed regions (%)	1.3
Outliers (%)	0.0

and plunging them into liquid nitrogen; they were then stored in liquid nitrogen until use in X-ray diffraction experiments.

2.4. Data collection, structure determination and structure analysis

Diffraction data were collected at 100 K on beamline BL-17U at Shanghai Synchrotron Research Facility (SSRF) and station 3W1A at Beijing Synchrotron Research Facility (BSRF). The data were indexed and scaled with *xia*2 in *CCP*4 (Winter, 2010; Winn *et al.*, 2011). The data-collection statistics are summarized in Table 1. The crystals belonged to space group $P6_122$, with unit-cell parameters $a = b = 60.1$, $c = 144.4$ Å. The structure was solved by molecular replacement with *Phaser* (McCoy *et al.*, 2007), using the *A* chain of the apo PC4 structure (PDB entry 1pcf), to which it has a sequence identity of 36.9% for 66 residues, stripped of waters and ligands as the search model. The asymmetric unit is estimated to contain one monomer, with a solvent content of 26%. The model was subsequently improved with *Coot* (Emsley *et al.*, 2010) and was further refined using *PHENIX* with TLS restraints (Adams *et al.*, 2010). The atomic coordinates and structure factors of MoSub1 have been deposited in the PDB as entry 4agh. Crystal parameters and data-collection statistics are also given in Table 1. Stereochemical validation of the model was performed using *MolProbity* (Chen *et al.*, 2010).

Dimer-interface analysis was performed with *PISA* (http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html; Krissinel & Henrick, 2007) and sequence alignment was performed with *ClustalW* (Larkin *et al.*, 2007). *BoxShade* (http://www.ch.embnet.org/software/BOX_form.html) was used to prepare Fig. 3. The structural figures were generated with *PyMOL* (v.1.3; Schrödinger LLC). Possible phosphorylation and acetylation sites were predicted by *NetPhos* (http://www.cbs.dtu.dk/services/NetPhos/; Blom *et al.*, 1999) and *PAIL* (http://bdmpail.biocuckoo.org/prediction.php).

3. Results and discussion

3.1. Overall structure

The structure of MoSub1 was determined by molecular replacement and was refined using data to 1.79 Å resolution. The data-collection and model-refinement statistics are given in Table 1. There is one molecule in the asymmetric unit, with a solvent content of 26%. The final crystallographic model consisted of one molecule of MoSub1 containing residues 38–117, which includes 66 residues from the model and 14 further residues that were manually built by *Coot*. The other N- and C-terminal residues of the molecule could not be located in the electron-density map and were thus not included in the final model.

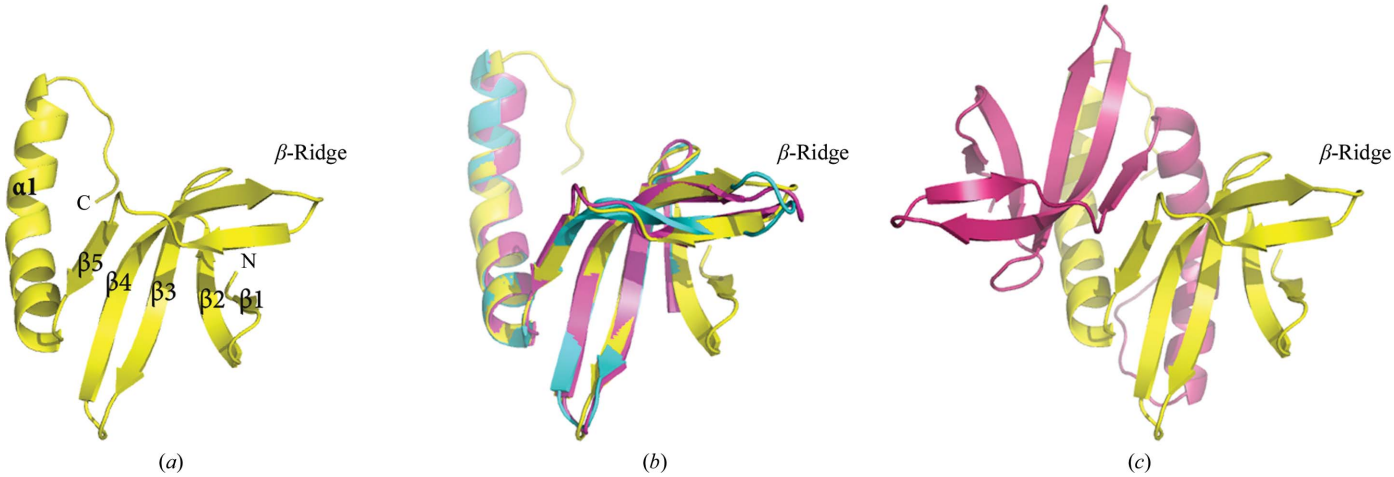

Figure 1
The overall structure and DNA-binding region of MoSub1. (*a*) The overall structure of MoSub1 (residues 38–117). (*b*) The DNA-binding region of MoSub1 (yellow) overlaid onto apo PC4 (purple) and the PC4–DNA complex (blue). (*c*) The MoSub1 dimer (yellow and red).

The molecule consists of a curved five-stranded antiparallel β-sheet followed by a 45° kinked α-helix (Fig. 1*a*). Superposition of the C^{α} atoms of MoSub1 with those of apo and DNA-complexed human PC4 indicates a high level of structural homology between them. The r.m.s.d. between the apo MoSub1 structure and PC4 is about 1.0 Å for apo PC4 and 1.2 Å for the DNA complex for residues 62–127 of PC4. The only difference between the models is in the β-ridge region between β4 and β5 (Fig. 1*b*).

3.2. The dimer interface

In the structure, a dimer can be formed by one monomer and another symmetry-related monomer, although there is only one molecule in the asymmetric unit (Fig. 1*c*). The dimer-interface area of MoSub1 is about 1646 Å^2; this is more than 25% of its total accessible surface area and is a little larger than those of the apo and DNA-complexed PC4 structures (Table 2). There are ten hydrogen bonds and eight salt bridges in the dimer interface and most of them are formed by residues that are conserved compared with the PC4 model (Table 2, Fig. 2). These results indicate that the dimerization interface is conserved among the three transcription cofactors PC4, Sub1 and MoSub1.

Table 2
The dimer-formation interfaces of the three structures MoSub1, apo PC4 (PDB entry 1pcf) and the PC4–DNA complex (PDB entry 2c62).

Model (chain)	Area (Å^2)	Residues	Atoms	Hydrogen bonds	Salt bridges
MoSub1	1646.0	48	166	10	8
1pcf (*F*)	1447.9	36	144	10	4
2c62 (*A*)	1404.9	36	137	12	4

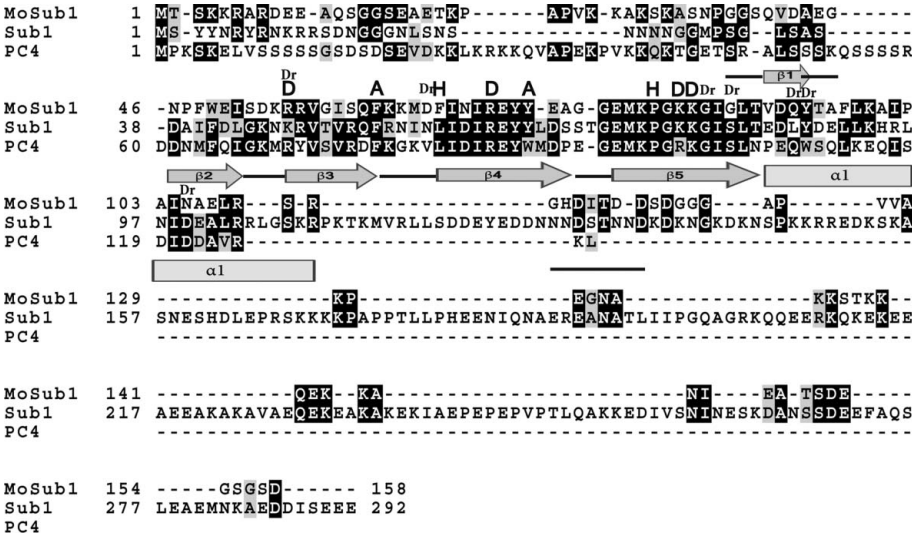

Figure 2
Sequence alignment and secondary-structure elements derived from the crystal structure of MoSub1. Identical residues are coloured black and similar residues are coloured grey; the key residues in DNA binding are labelled D (residues that form hydrogen bonds to DNA), A (aromatic residues that interact with DNA) or H (other hydrophobic residues that interact with DNA); those involved in dimer formation are labelled Dr.

3.3. The DNA-binding site

DNA-binding activity is an essential factor in Sub1/PC4 function. The DNA-binding affinity of MoSub1 for single-stranded DNA was analyzed by SPR. MoSub1 binds the single-stranded DNA dT_{12} tightly, with a dissociation constant of 186.0 n*M* (Fig. 3). The binding affinity of oligo-dTs of different lengths (dT_{10}, dT_{16} and dT_{20}) for the C-terminal DNA-binding domain of PC4 (PC4-CTD) was determined by EMSA (electrophoretic mobility shift assay; Werten, Langen *et al.*, 1998). The longer DNAs such as dT_{16} or dT_{20} have a very high affinity for the protein (K_d is 0.8 n*M* for dT_{16} and

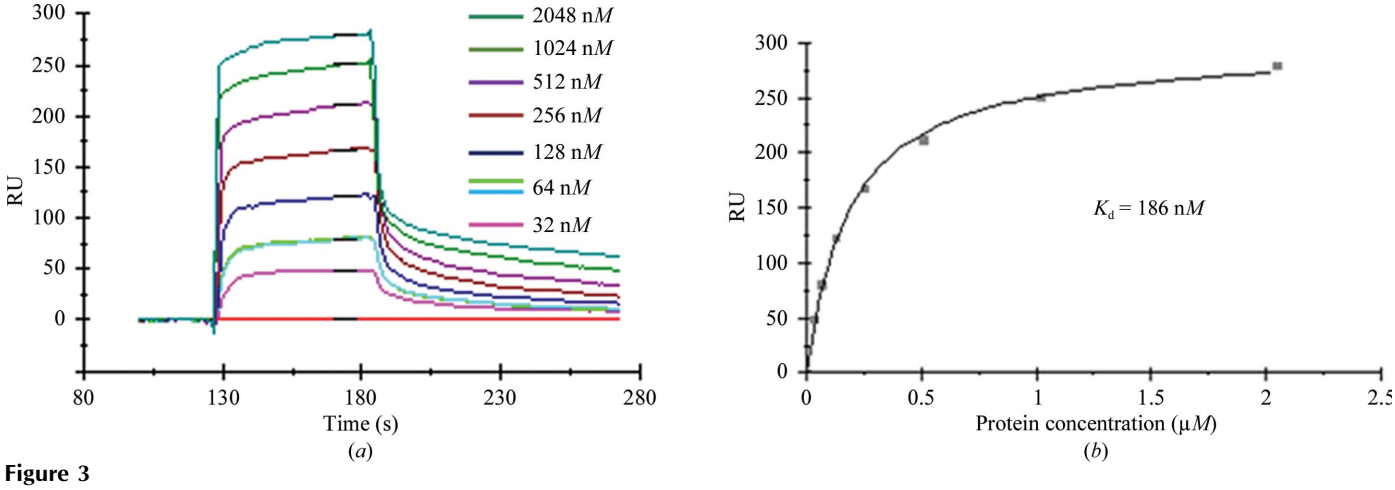

Figure 3
Biacore binding assays of MoSub1 to immobilized single-stranded oligo-(dT_{12}). (*a*) Six concentrations of the protein (from 32 to 2048 n*M*; coloured lines) were measured. RU, response units. (*b*) Equilibrium analysis of MoSub1 binding to immobilized single-stranded oligo-(dT_{12}). Responses at equilibrium were plotted against protein concentration and fitted to a steady-state affinity model.

0.2 n*M* for dT_{20}), while hardly any interaction between dT_{10} and the protein could be detected. MoSub1, which is a homologue of PC4, could bind longer oligo-DNAs (dT_{16} or dT_{20}) more strongly than the short oligo-DNA dT_{12}.

The DNA-binding site of MoSub1 was predicted by superposition of the MoSub1 and PC4–DNA complex structures (Werten & Moras, 2006). The key residues for binding DNA, such as Arg70, Phe77, Arg86, Pro98 and Lys101, are highly conserved between PC4 and MoSub1 (PC4 numbering; Figs. 2 and 4). A change from Arg to Lys at position 100 or from Trp to Tyr at position 89 retained the interaction with DNA (Fig. 4). At the same time, three waters which form the bridge between the main chain of the β-sheet and DNA were found in similar positions in the MoSub1 structure and could be involved in the interaction with DNA (Fig. 4). These results, together with the high sequence homology, indicated that the proteins of the Sub1/PC4 family bind DNA using similar motifs.

3.4. Structural features of the N- and C-terminal regions

There are two extra regions in the MoSub1 model that were absent in the PC4 structure: the region N-terminal to the DNA-binding domain and the region C-terminal to the DNA-binding domain (Fig. 5). The N-terminal region (38–46, MoSub1 numbering) is bent and turns back towards the DNA-binding surface, indicating that the flexible N-terminus may be able to interact directly with DNA or the DNA-binding region. It has been shown that post-translational modifications in this region of PC4 or Sub1 can change the DNA-binding activity (Ge *et al.*, 1994; Jonker *et al.*, 2006). The reason for this could be that the modifications may change the conformation and affect DNA binding or interaction with DNA directly. The lysine-rich and SEAC (serine/acidic residue-rich) regions in the N-terminal tail of PC4 are flexible and could be phosphorylated or acetylated to regulate its function (Jonker *et al.*, 2006). Several residues, such as the seven Ser/Thr and six Lys residues in the N-terminal 40 residues of MoSub1 (Tables 3 and 4), are suitable candidates for modification. Most of these residues are likely to be flexible as no electron density could be observed, with the exception of Ser39 which could be built into the N-terminus of the model (38–117). The function of MoSub1 could be regulated by post-translational modification of these residues in the N-terminus, as has been reported for PC4 (Ge *et al.*, 1994; Jonker *et al.*, 2006). The conformation of the N-terminal domain in the PC4 model is unknown and may be similar to that of MoSub1, although the sequence homology in this region is quite low.

The C-terminal nonconserved residues (112–117, MoSub1 numbering), which are only present in Sub1 or MoSub1, are not found near the potential DNA-binding interface and interact with the two helices that form the dimerization interface (Fig. 5). Instead of binding DNA, they may be able to bind other novel partners that play a role in transcription, such as TFIIB. The C-terminal region extension in Sub1 or MoSub1 is one possible explanation for the differences between Sub1 and PC4.

In conclusion, we have provided a structural insight into MoSub1 from rice blast fungus, which is an ssDNA-binding protein, a potential transcription cofactor and a homology model of yeast Sub1. In the future, it will be of interest to

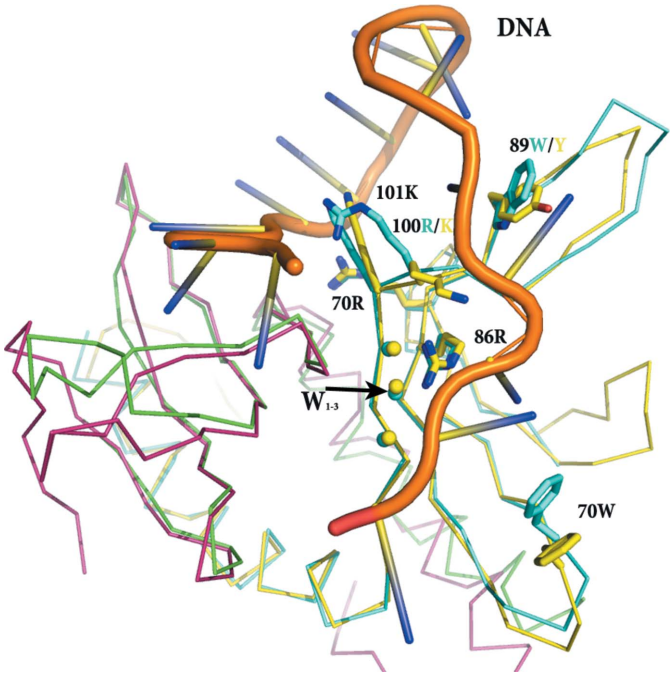

Figure 4
The binding surface of one monomer, including three waters that interact with DNA. The side chains of the key residues are shown as sticks, MoSub1 (yellow and red) is overlaid onto the PC4–DNA complex structure (blue and green) and the DNA is coloured orange.

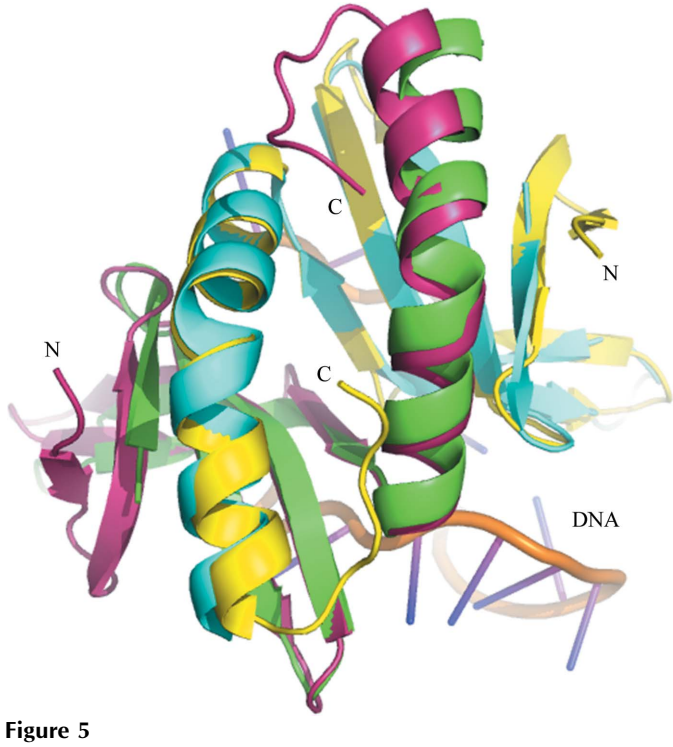

Figure 5
The regions N- and C-terminal to the DNA-binding domain of MoSub1. MoSub1 (yellow and red) is superposed onto the PC4–DNA complex structure (blue and green) and the DNA is coloured orange.

Table 3
Phosphorylation sites predicted by *NetPhos* (http://www.cbs.dtu.dk/services/NetPhos/).

Position	Context	Score	Prediction
3	**MT**S**KKRA	0.763	*S*
14	EEAQ**S**GGSE	0.908	*S*
17	QSGG**S**EAET	0.994	*S*
31	KKAK**S**KASN	0.950	*S*
34	KSKA**S**NPGG	0.687	*S*
39	NPGG**S**QVDA	0.995	*S*
2	***M**T**S**KKR	0.900	*T*

Table 4
Acetylation sites predicted by *PAIL* (http://bdmpail.biocuckoo.org/prediction.php).

Peptide	Position	Score	Threshold
***MTS**K**KRARDE	4	12.08	0.5
MTSKK**RARDEE	5	7.58	0.5
TKPAPV**K**KAKSKA	27	3.37	0.5
KPAPVK**K**AKSKAS	28	2.72	0.5
APVKKA**K**SKASNP	30	2.40	0.5
VKKAKS**K**ASNPGG	32	2.84	0.5

determine the DNA or protein partners that MoSub1 can interact with and determine the mechanism of action.

We thank the staff of beamline BL-17U at Shanghai Synchrotron Research Facility (SSRF) and station 3W1A at Beijing Synchrotron Research Facility (BSRF) and Hongbing Li from the Institute of Microbiology, Chinese Academy of Sciences for help with crystal screening and data collection, and Yuanyuan Chen for technical help with the Biacore 3000 at the Institute of Biophysics, Chinese Academy of Sciences. This work was supported by National Natural Science Foundation of China (NSFC) grant No. 31171800 and 973 Project grant No. 2012CB114000 from the Ministry of Science and Technology of China (MOST), as well as funds from the Chinese Universities Scientific Fund (Project Nos. 2011JS071 and 2012RC008).

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Blom, N., Gammeltoft, S. & Brunak, S. (1999). *J. Mol. Biol.* **294**, 1351–1362.
Brandsen, J., Werten, S., van der Vliet, P. C., Meisterernst, M., Kroon, J. & Gros, P. (1997). *Nature Struct. Biol.* **4**, 900–903.
Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.
Conesa, C. & Acker, J. (2010). *RNA Biol.* **7**, 287–290.
Ebbole, D. J. (2007). *Annu. Rev. Phytopathol.* **45**, 437–456.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Ge, H. & Roeder, R. G. (1994). *Cell*, **78**, 513–523.
Ge, H., Zhao, Y., Chait, B. T. & Roeder, R. G. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 12691–12695.
Guo, M., Chen, Y., Du, Y., Dong, Y., Guo, W., Zhai, S., Zhang, H., Dong, S., Zhang, Z., Wang, Y., Wang, P. & Zheng, X. (2011). *PLoS Pathog.* **7**, e1001302.
Henry, N. L., Bushnell, D. A. & Kornberg, R. D. (1996). *J. Biol. Chem.* **271**, 21842–21847.
Huang, J., Gong, Y., Huang, D., Haire, L., Liu, J. & Peng, Y. (2012). *Acta Cryst.* F**68**, 606–609.
Jonker, H. R., Wechselberger, R. W., Boelens, R., Kaptein, R. & Folkers, G. E. (2006). *Biochemistry*, **45**, 5067–5081.
Kim, S., Park, S.-Y., Kim, K. S., Rho, H.-S., Chi, M.-H., Choi, J., Park, J., Kong, S., Park, J., Goh, J. & Lee, Y.-H. (2009). *PLoS Genet.* **5**, e1000757.
Knaus, R., Pollock, R. & Guarente, L. (1996). *EMBO J.* **15**, 1933–1940.
Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J. & Higgins, D. G. (2007). *Bioinformatics*, **23**, 2947–2948.
Li, G., Zhou, X., Kong, L., Wang, Y., Zhang, H., Zhu, H., Mitchell, T. K., Dean, R. A. & Xu, J.-R. (2011). *PLoS One*, **6**, e19951.
Malik, S., Guermah, M. & Roeder, R. G. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 2192–2197.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Nishimura, M., Fukada, J., Moriwaki, A., Fujikawa, T., Ohashi, M., Hibi, T. & Hayashi, N. (2009). *Biosci. Biotechnol. Biochem.* **73**, 1779–1786.
Rosonina, E., Willis, I. M. & Manley, J. L. (2009). *Mol. Cell. Biol.* **29**, 2308–2321.
Tavenet, A., Suleau, A., Dubreuil, G., Ferrari, R., Ducrot, C., Michaut, M., Aude, J.-C., Dieci, G., Lefebvre, O., Conesa, C. & Acker, J. (2009). *Proc. Natl Acad. Sci. USA*, **106**, 14265–14270.
Valent, B. & Khang, C. H. (2010). *Curr. Opin. Plant Biol.* **13**, 434–441.
Werten, S., Langen, F. W., van Schaik, R., Timmers, H. T., Meisterernst, M. & van der Vliet, P. C. (1998). *J. Mol. Biol.* **276**, 367–377.
Werten, S. & Moras, D. (2006). *Nature Struct. Mol. Biol.* **13**, 181–182.
Werten, S., Stelzer, G., Goppelt, A., Langen, F. M., Gros, P., Timmers, H. T., van der Vliet, P. C. & Meisterernst, M. (1998). *EMBO J.* **17**, 5103–5111.
Wilson, R. A. & Talbot, N. J. (2009). *Nature Rev. Microbiol.* **7**, 185–195.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Winter, G. (2010). *J. Appl. Cryst.* **43**, 186–190.
Yan, X., Li, Y., Yue, X., Wang, C., Que, Y., Kong, D., Ma, Z., Talbot, N. J. & Wang, Z. (2011). *PLoS Pathog.* **7**, e1002385.
Zhou, X., Liu, W., Wang, C., Xu, Q., Wang, Y., Ding, S. & Xu, J.-R. (2011). *Mol. Microbiol.* **80**, 33–53.