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

Expression, crystallization and preliminary X-ray data analysis of NT-Als9-2, a fungal adhesin from *Candida albicans*

Paula S. Salgado, Robert Yan,‡ Fiona Rowan§ and Ernesto Cota*

Division of Molecular Biosciences, Faculty of Natural Sciences, Imperial College London, South Kensington Campus, London SW7 2AZ, England

‡ Current address: National Institute for Medical Research, Medical Research Council, The Ridgeway, Mill Hill, London NW7 1AA, England.
§ Current address: Mitotic Regulation and Cancer, Section of Structural Biology, Institute of Cancer Research, Chester Beatty Laboratories, 237 Fulham Road, London SW3 6JB, England.

Correspondence e-mail: e.cota@imperial.ac.uk

Received 15 October 2010
Accepted 26 January 2011

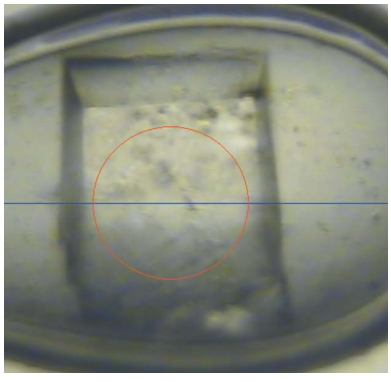
Candida albicans is a common human fungal commensal that can also cause a range of infections from skin/mucosal 'thrush' to severe systemic candidiasis. Adherence to host cells is one of the key determinants of *Candida* pathogenesis. The Als family of surface proteins has been implicated in adhesion of *C. albicans*, yet limited information has been published on the structure and mechanism of these fungal adhesins. The N-terminal region of these proteins has been shown to possess adhesive properties, making it a possible target for new therapeutic strategies. Recombinant NT-Als9-2 from *C. albicans* (residues 18–329) was overexpressed in *Escherichia coli*, purified and crystallized. Diffraction data were collected to 2.0 Å resolution. The crystals belonged to space group $P2_12_12_1$, with unit-cell parameters $a = 34.73$, $b = 68.71$, $c = 120.03$ Å, $\alpha = \beta = \gamma = 90^\circ$ and one molecule in the asymmetric unit. Platinum-derivatized crystals belonged to the same space group, with similar unit-cell parameters, although they were not completely isomorphous.

1. Introduction

Candida species are ubiquitous fungal commensals that are present in ∼80% of the human population. However, impairment of host immunity or the normal host microbiota can lead to *Candida* overgrowth and disease. As a result, *Candida* spp. are also the most frequent cause of opportunistic mycoses worldwide. Skin and oral *Candida* infections are common in HIV-positive individuals (Gaitán-Cepeda *et al.*, 2005) and vaginal candidiasis is frequent in normally healthy women (Sobel *et al.*, 1998). Risk factors for severe systemic candidiasis include immunosuppression, either through disease or immunotherapy (such as that associated with organ transplantation), or the use of medical implants. *C. albicans* and other *Candida* spp. have emerged as the sixth most frequently isolated pathogen in hospitals (Jarvis, 1995; Pfaller & Diekema, 2007) and the fourth most frequent cause of nosocomial bloodstream infections, with a mortality rate of up to 50% when these develop into severe systemic forms (Gudlaugsson *et al.*, 2003).

Adherence to host cells is one of the major determinants of *Candida* pathogenesis and a key step in the infection process. The agglutinin-like sequence (Als) family of surface proteins (Als1–7 and Als9) has been implicated in the adhesion of *C. albicans* to host cells and in binding to a broad range of proteins in cell surfaces and the extracellular matrix (Gaur & Klotz, 1997; Phan *et al.*, 2007). Moreover, Als1 and Als5 have been shown to bind a large variety of sequences from a library of random heptameric peptides (Klotz *et al.*, 2004). Structural studies are essential for the understanding of the mechanism that allows the association of Als proteins with such a diverse variety of ligands and its implication in the infection process.

The Als adhesins are composed of four distinct protein regions: a poorly glycosylated N-terminal region of approximately 320 amino



acids, with 41–84% similarity among Als members, and a nonrepeat Thr-rich region of 103 amino acids followed by two highly glycosylated regions: a central Thr-rich domain with a variable number of 36-amino-acid repeats and a Ser/Thr/Asn-rich membrane-anchoring C-terminal domain of variable length (Hoyer, 2001). The N-terminal region, in which the observed cell-adhesion properties have been shown to reside (Sheppard *et al.*, 2004; Loza *et al.*, 2004), is predicted to form a predominant β-structure.

In order to describe in atomic detail the mode of peptide binding and the conformational changes that occur upon the transition from free protein to protein–ligand complex, we used protein X-ray crystallography and nuclear magnetic resonance (NMR). Our preliminary NMR data from the N-terminal domain of Als revealed an IgG-superfamily secondary-structure topology (Yan *et al.*, 2010). Here, we report the expression, crystallization and preliminary X-ray data analysis of the N-terminal region of one member of the Als family: NT-Als9-2.

2. Materials and methods

2.1. NT-Als9-2 cloning

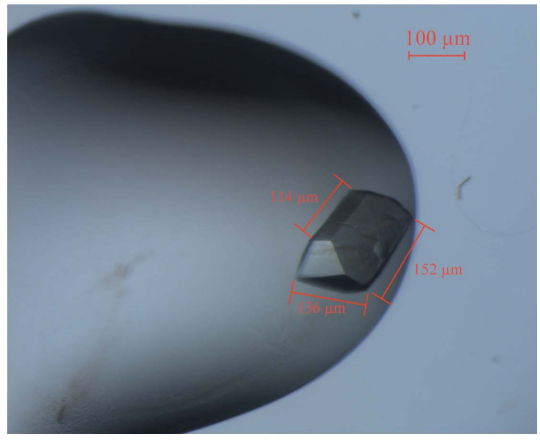
A DNA fragment of *C. albicans* Als9-2 (ExPASy Q5A8T1) encoding residues 18–329 was kindly provided by Professor Lois Hoyer (University of Illinois at Urbana-Champaign). The *Escherichia coli* expression plasmid pET32-LIC/Xa (Novagen) contains an N-terminal thioredoxin-6×His tag. In order to introduce the NT-Als9-2 coding sequence, oligonucleotides 5′-GGTATT-GAGGGTCGCAAGACTATTACTGGTGTTTTCAATAGTTTT-3′ and 5′-AGAGGAGAGTTAGACCCTTAAACTACGATCACATC-CCCGTCAGAGTC-3′ were used as forward and reverse primers, respectively. To produce NT-Als9-2, plasmid pET32-LIC/Xa was transformed into *E. coli* strain Origami B (Novagen), which provides a reducing environment for disulfide-bond formation in the cytoplasm.

2.2. Expression and purification

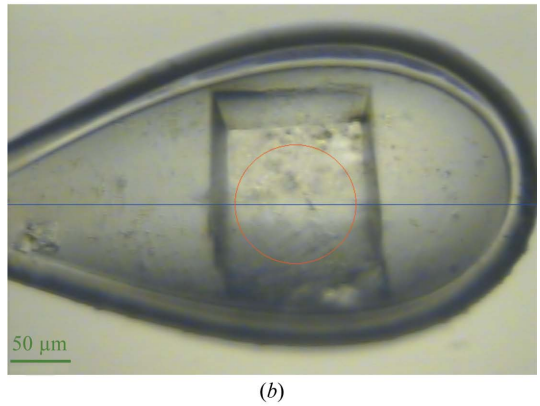
Cells were grown at 310 K in LB medium in a 15 l fermenter with aeration for 6 h, at which point the temperature was lowered to 295 K for overnight growth. The medium was then complemented with 20% sterilized glucose and the cells were grown at 310 K to an A_{600} of ∼2. The cells were then induced with 0.5 m*M* IPTG and expression was allowed to continue overnight. The cells were collected by centrifugation for 15 min at 1500 rev min^{-1} (at room temperature) and resuspended in 100 ml buffer *A* (50 m*M* Tris–HCl, 300 m*M* NaCl pH 8.0). After lysing the cells using a cell disruptor, the supernatant obtained by centrifugation at 15 000 rev min^{-1} for 1 h was loaded onto an Ni–NTA affinity column pre-equilibrated with buffer *A*. Several washes with the same buffer were carried out before eluting the bound protein with buffer *A* containing 200 m*M* imidazole. In order to prepare the eluted fractions for cleavage of the fusion protein, dialysis against 50 m*M* Tris, 100 m*M* NaCl pH 8.0 was carried out overnight at 277 K. At this point, the protein was concentrated using a 3000 molecular-weight cutoff Vivaspin 20 ml concentrator (Sartorius) before adding factor Xa (Novagen) at a concentration of 1 U per 100 µg of protein. The protein concentration was determined by measuring the absorption at A_{280} using a Nanodrop spectrophotometer. Cleavage buffer (10×; 1 *M* NaCl, 500 m*M* Tris–HCl, 50 m*M* $CaCl_2$ pH 8.0) was also added, as well as Milli-Q water, to give a final reaction volume of 6 ml. The cleavage reaction was left at room temperature for at least 16 h with gentle shaking. Cleavage-reaction mixtures were injected onto a Superdex 200 HiLoad 16/60 column (GE Healthcare) equilibrated with buffer *A* and run at room temperature. Fractions were analysed by 15% SDS–PAGE and the protein concentration was determined by measuring the absorption at A_{280} using a Nanodrop spectrophotometer.

2.3. Crystallization

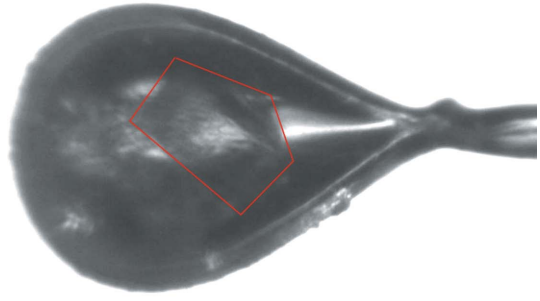
The purified protein was concentrated to 10 mg ml^{-1} by centrifugation in a Vivaspin 20 ml concentrator (Sartorius) in a final buffer consisting of 25 m*M* Tris–HCl, 50 m*M* NaCl pH 8.0. Crystallization condition screening was carried out using a range of commercially available screens with 200 and 400 nl drop sizes (1:1 and 2:1 protein:precipitant ratio, respectively) dispensed by a Mosquito robot. Initial hits were found in more than 50 conditions over all



(*a*)



(*b*)



(*c*)

Figure 1
Crystals of NT-Als9-2. (*a*) Native crystals grown at 295 K in 100 m*M* sodium acetate, 200 m*M* $MgCl_2$ pH 5.0, 20% PEG 6000. (*b*) A native crystal mounted in a cryoloop prior to data collection. (*c*) A $PtCl_4$-soaked crystal (outlined with a red box) mounted in a cryoloop prior to data collection. The approximate dimensions of the crystals are 100 × 100 × 50 µm. Paraffin oil was used as cryoprotectant.

Table 1
Data-collection statistics.

Values in parentheses are for the last resolution shell.

	Native	Pt derivative
Crystal dimensions (µm)	~100 × 100 × 50	~100 × 100 × 50
Source	BM14, ESRF	PXIII, SLS
Wavelength (Å)	0.979	1.068
No. of images	360	720
Oscillation per image (°)	1	0.5
Unit-cell parameters (Å)	a = 34.73, b = 68.71, c = 120.03	a = 35.07, b = 68.99, c = 120.77
Space group	$P2_12_12_1$	$P2_12_12_1$
Resolution (Å)	50.00–2.00 (2.03–2.00)	68.9–1.83 (1.93–1.83)
Reflections (total/unique)	172506/20201	366357/26698
Completeness (total/anomalous) (%)	99.7	99.7/98.2
Multiplicity	8.5 (7.4)	13.7 (12.5)
R_{merge}†	0.12 (0.45)	0.08 (0.45)
$\langle I/\sigma(I)\rangle$	19.8 (3.4)	24.9 (5.6)
Mosaicity (°)	0.8	0.47
No. of molecules in asymmetric unit	1	1
Solvent content (%)	48.0	49
Substructure solution: *SHELX*		
No. of sites	—	2
CC_{all}/CC_{weak}	—	0.48
PATFOM	—	98.67
Phasing: *SHARP*		
FOM	—	0.38
Phasing power (iso/ano)	—	1.04/1.7
Model building: *ARP*/*wARP*		
No. of residues built	—	282 [of 310]
No. docked in sequence	—	279 [98%]
Score	—	2.26
R_{work} (%)	—	20.9
R_{free} (%)	—	27.3

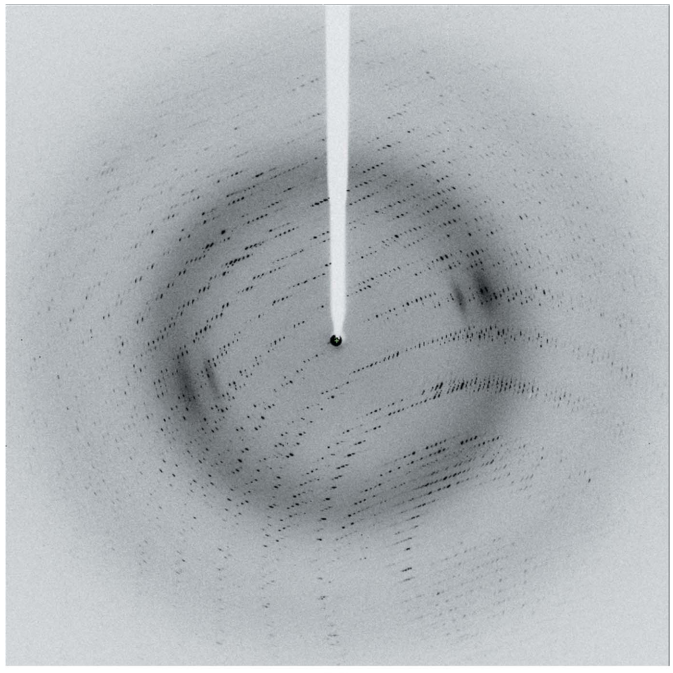
† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the ith observation of reflection hkl and $\langle I(hkl)\rangle$ is the weighted average intensity for all observations i of reflection hkl.

screens tested. However, crystal quality was poor, with several lattices often being present in the crystals even when they appeared to be unique when observed under the microscope (Fig. 1*a*). Optimization was therefore carried out around conditions that apparently yielded fewer multiple crystals. All crystallization trials were carried out at 293 K.

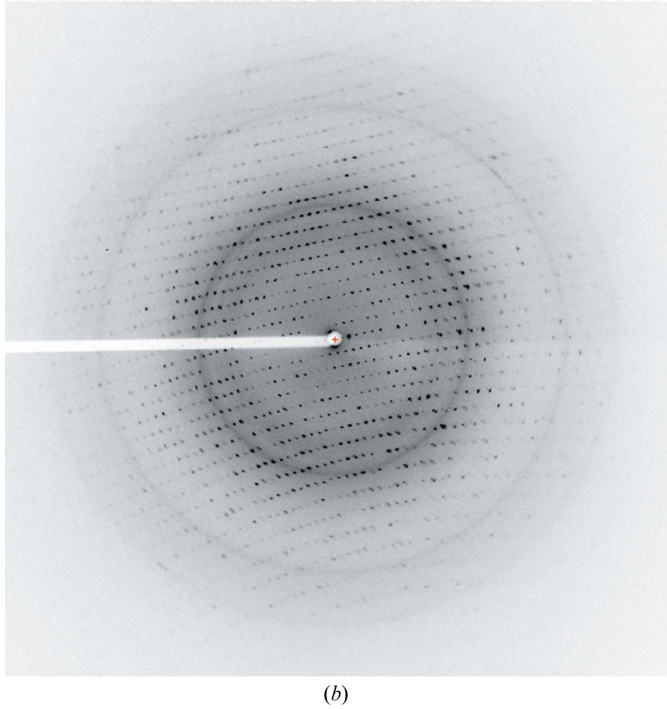
For cryo data collection, a range of common cryoprotectants was tested and finally all data were collected after briefly washing the crystals in a solution of paraffin or silicone oil. Owing to the presence of multiple lattices, together with the poor diffraction power of most of the crystals obtained, extensive crystal screening had to be performed at synchrotron sources before a good processable data set was obtained from a crystal grown using 100 m*M* sodium acetate, 200 m*M* $MgCl_2$ pH 5.0, 20% PEG 6000 (Fig. 1*b*).

Molecular-replacement (MR) methods to solve the phase problem proved unsuccessful, probably owing to the low sequence similarity of NT-Als9-2 to the bacterial adhesins ClfA from *Staphylococcus aureus* (Ponnuraj *et al.*, 2003) and SdrG from *S. epidermidis* (Deivanayagam *et al.*, 2002), which are the closest structural homologues as predicted by *PHYRE* (Kelley & Sternberg, 2009). Inability to obtain a good MR solution, as observed for other proteins with an IgG fold, could also be related to the high β-sheet content of the search models. Selenomethionine derivatization was also unsuccessful, as no protein was expressed in auxotrophic Met^- *E. coli* strains or in Origami B strain grown in minimal medium supplemented with SeMet (data not shown).

Obtaining heavy-atom derivatives was therefore necessary, but was not trivial. Extensive cocrystallization and soaking conditions were screened, initially based on the conditions that yielded the best native data. However, even when good overall quality data were obtained after extensive crystal screening only a very weak anomalous signal was detected. Therefore, re-optimization of crystallization conditions to allow better heavy-atom incorporation was carried out. Crystals grown in 30 m*M* Tris, 100 m*M* $MgCl_2$ pH 8.3, 25% PEG 4000 (Fig. 1*c*) and soaked with an excess of $PtCl_4$ finally provided a good data set with a detectable anomalous signal. Particularly important for successful phasing was a high concentration of the heavy atom to ensure complete saturation of the crystals, combined with data collection at the precisely determined peak energy and corresponding wavelength. This is owing to the presence of only one heavy atom per



(*a*)



(*b*)

Figure 2
Typical X-ray diffraction patterns for (*a*) a native crystal (1° oscillation) and (*b*) a $PtCl_4$-soaked crystal (0.5° oscillation).

molecule, which means that only optimal conditions of anomalous signal detection could provide good starting phasing power.

2.4. Data collection and processing

A native data set was collected at 100 K to 2.00 Å resolution on station BM14 at the ESRF, Grenoble at a wavelength of 0.979 Å using a MAR CCD detector and 1° oscillation per image. Data were processed using *HKL*-2000 (Otwinowski & Minor, 1997; details are given in Table 1; Fig. 2*a*). NT-Als9-2 crystals belonged to space group $P2_12_12_1$, with unit-cell parameters a = 34.73, b = 68.71, c = 120.03 Å and one molecule per asymmetric unit.

A crystal soaked for 1 h in 100 m*M* $PtCl_4$ was used to collect the Pt-derivative data on beamline PXIII at Swiss Light Source, Switzerland. The presence of anomalous signal was confirmed by performing a fluorescence scan around the Pt edge, followed by analysis with *CHOOCH* (Evans & Pettifer, 2001). Data were then collected at the detected peak wavelength of 1.068 Å with 0.5° oscillation per image (Fig. 2*b*). The crystal belonged to space group $P2_12_12_1$, with unit-cell parameters a = 35.07, b = 68.99, c = 120.77 Å. *XDS* (Kabsch, 2010) was used to process the data and analysis of the anomalous signal was performed with the *SHELX* suite of programs (Sheldrick, 2008). Preliminary analysis of the platinum substructure by *SHELX* and *autoSHARP* (Vonrhein *et al.*, 2007) confirmed that one Pt atom was present in the crystal and the phasing power statistics indicate that a solution is obtainable. Preliminary model building with *ARP*/*wARP* (Perrakis *et al.*, 1999) has allowed the building of an initial model with 98% of residues docked into the sequence. Model building and refinement is ongoing.

These results pave the way for the determination of the atomic structure of NT-Als9-2 and demonstrate that even if more than one lattice is present in a crystal a good data set with strong anomalous signal can be obtained if data are collected carefully and at optimal energy levels.

We thank the staff at the UK MAD station, BM14, ESRF, Grenoble and the staff at the PXIII station, Swiss Light Source, Switzerland, in particular M. Walsh and R. Bingel-Erlenmeyer, respectively, for their assistance with data collection. We also acknowledge the Centre for Structural Biology at Imperial College London and J. Moore for data collection and analysis. The work was supported by BBSRC Grant BB/F007566/1.

References

Deivanayagam, C. C., Wann, E. R., Chen, W., Carson, M., Rajashankar, K. R., Höök, M. & Narayana, S. V. (2002). *EMBO J.* **21**, 6660–6672.
Evans, G. & Pettifer, R. F. (2001). *J. Appl. Cryst.* **34**, 82–86.
Gaitán-Cepeda, L. A., Martínez-González, M. & Ceballos-Salobreña, A. (2005). *AIDS Patient Care STDS*, **19**, 70–77.
Gaur, N. K. & Klotz, S. A. (1997). *Infect. Immun.* **65**, 5289–5294.
Gudlaugsson, O., Gillespie, S., Lee, K., Vande Berg, J., Hu, J., Messer, S., Herwaldt, L., Pfaller, M. & Diekema, D. (2003). *Clin. Infect. Dis.* **37**, 1172–1177.
Hoyer, L. L. (2001). *Trends Microbiol.* **9**, 176–180.
Jarvis, W. R. (1995). *Clin. Infect. Dis.* **20**, 1526–1530.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 125–132.
Kelley, L. A. & Sternberg, M. J. (2009). *Nature Protoc.* **4**, 363–371.
Klotz, S. A., Gaur, N. K., Lake, D. F., Chan, V., Rauceo, J. & Lipke, P. N. (2004). *Infect. Immun.* **72**, 2029–2034.
Loza, L., Fu, Y., Ibrahim, A. S., Sheppard, D. C., Filler, S. G. & Edwards, J. E. (2004). *Yeast*, **21**, 473–482.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Perrakis, A., Morris, R. & Lamzin, V. S. (1999). *Nature Struct. Biol.* **6**, 458–463.
Pfaller, M. A. & Diekema, D. J. (2007). *Clin. Microbiol. Rev.* **20**, 133–163.
Phan, Q. T., Myers, C. L., Fu, Y., Sheppard, D. C., Yeaman, M. R., Welch, W. H., Ibrahim, A. S., Edwards, J. E. & Filler, S. G. (2007). *PLoS Biol.* **5**, e64.
Ponnuraj, K., Bowden, M. G., Davis, S., Gurusiddappa, S., Moore, D., Choe, D., Xu, Y., Hook, M. & Narayana, S. V. L. (2003). *Cell*, **115**, 217–228.
Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.
Sheppard, D. C., Yeaman, M. R., Welch, W. H., Phan, Q. T., Fu, Y., Ibrahim, A. S., Filler, S. G., Zhang, M., Waring, A. J. & Edwards, J. E. (2004). *J. Biol. Chem.* **279**, 30480–30489.
Sobel, J. D., Faro, S., Force, R. W., Foxman, B., Ledger, W. J., Nyirjesy, P. R., Reed, B. D. & Summers, P. R. (1998). *Am. J. Obstet. Gynecol.* **178**, 203–211.
Vonrhein, C., Blanc, E., Roversi, P. & Bricogne, G. (2007). *Methods Mol. Biol.* **364**, 215–230.
Yan, R., Simpson, P. J., Matthews, S. J. & Cota, E. (2010). *Biomol. NMR Assign.* **4**, 187–190.