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

Hannes Simader and Dietrich Suck*

Structural and Computational Research Program, European Molecular Biology Laboratory (EMBL), Meyerhofstrasse 1, D-69117 Heidelberg, Germany

Correspondence e-mail: suck@embl.de

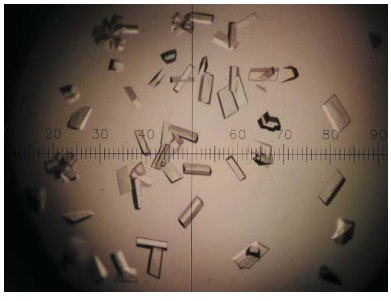
Received 26 January 2006
Accepted 16 February 2006

Expression, purification, crystallization and preliminary phasing of the heteromerization domain of the tRNA-export and aminoacylation cofactor Arc1p from yeast

Eukaryotic aminoacyl-tRNA synthetases (aaRSs) must be integrated into an efficient tRNA-export and shuttling machinery. This is reflected by the presence of additional protein–protein interaction domains and a correspondingly higher degree of complex formation in eukaryotic aaRSs. However, the structural basis of interaction between eukaryotic aaRSs and associated protein cofactors has remained elusive. The N-terminal heteromerization domain of the tRNA aminoacylation and export cofactor Arc1p has been cloned from yeast, expressed and purified. Crystals have been obtained belonging to space group $C2$, with unit-cell parameters a = 222.32, b = 89.46, c = 126.79 Å, β = 99.39°. Calculated Matthews coefficients are compatible with the presence of 10–25 monomers in the asymmetric unit. A complete multiple-wavelength anomalous dispersion data set has been collected from a selenomethionine-substituted crystal at 2.8 Å resolution. Preliminary phasing reveals the presence of 20 monomers organized in five tetramers per asymmetric unit.

1. Introduction

In prokaryotes, aminoacyl-tRNA synthetases (aaRSs) function as monomers or dimers. Owing to the high degree of compartmentalization of eukaryotic cells, eukaryotic aaRSs must be integrated into an efficient cellular tRNA-channelling machinery (Negrutskii & Deutscher, 1991, 1992; Stapulionis & Deutscher, 1995). This functional requirement is frequently met by the presence of additional protein–protein interaction domains and a correspondingly higher degree of complex formation in eukaryotic aaRSs (Martinis *et al.*, 1999). While there is a wealth of structural information available on prokaryotic synthetases, structural insight into complex formation between eukaryotic aaRSs and protein cofactors is lacking. In higher eukaryotes, nine aminoacyl-tRNA synthetases, including methionyl-tRNA synthetase (MetRS), are associated together to form a supramolecular multienzyme complex (Kellermann *et al.*, 1982; Lee *et al.*, 2004). In the yeast *Saccharomyces cerevisiae*, an evolutionary intermediate is represented by a complex formed by MetRS and glutamyl-tRNA synthetase (GluRS) with Arc1p (Simos *et al.*, 1996, 1998; Galani *et al.*, 2001, 2005). The N-terminal 130 amino acids of Arc1p are known to be necessary and sufficient for recruiting both GluRS and MetRS into a heteromeric complex of apparent 1:1:1 stoichiometry (Deinert *et al.*, 2001). To elucidate the structural basis of heteromerization between a eukaryotic aaRS and its cofactor, we have chosen the yeast Arc1p complex as a well characterized model system.



2. Expression, purification and interaction analysis

Residues 1–110, 1–116, 1–122 and 1–131 of *S. cerevisiae* Arc1p were PCR-amplified from plasmid pET-HIS6/pET8c-Arc1p (Simos *et al.*, 1996) and cloned into a modified pETm (Novagen) vector allowing the removal of the N-terminal His_6 tag through TEV protease cleavage, leaving a Gly-His dipeptide N-terminal to the natural starting methionine. The sequence of the expression constructs was confirmed by sequencing with T7 primer. The constructs were transformed into *Escherichia coli* strain BL21(DE3) Star. The cells were grown at 310 K in 6 l TB medium containing 100 µg ml^{-1}

Table 1
Data statistics.

Values in parentheses refer to the highest resolution shell.

	Peak	Edge	Remote
Space group	*C*2		
Unit-cell parameters (Å, °)	a = 222.3, b = 89.5, c = 126.8, β = 99.4		
Wavelength (Å)	0.97925	0.97945	0.95375
Resolution range (Å)	50–2.8 (2.9–2.8)		
Observed reflections	449519	338093	339646
Unique reflections	116298	116682	116970
Redundancy	3.9 (3.8)	2.9 (2.8)	2.9 (2.9)
Completeness (%)	98.9 (90.2)	99.2 (93.9)	99.4 (96.0)
R_{sym}† (%)	7.8 (27.9)	7.5 (27.9)	7.8 (30.1)
Mosaicity (°)	0.168	0.178	0.178
Average $I/\sigma(I)$	10.86 (4.54)	9.51 (3.63)	9.17 (3.43)
Anomalous signal‡	1.15	1.04	1.06

† $R_{sym} = \sum |I(h,i) - I(h)| / \sum I(h,i)$. ‡ Anomalous signal-to-noise ratio as the mean $\sigma(I)$ of acentric reflections assuming Friedel's law to be true divided by the mean $\sigma(I)$ of accentric reflections assuming Friedel's law to be false.

carbenicillin (Sigma) to an OD_{600} of 0.6. Upon induction with 0.1 m*M* IPTG, the incubation temperature was lowered to 291 K and cells were grown for a further 12 h. Residues 1–110 and 1–116 could not be expressed in soluble form, but residues 1–122 (Arc1p_1–122) and 1–131 (Arc1p_1–131) were well expressed and soluble. Cells were harvested by centrifugation, resuspended in lysis buffer (50 m*M* Tris, 150 m*M* NaCl, 20 m*M* imidazole, 5 m*M* $MgCl_2$, 10% glycerol, 10 m*M* β-mercaptoethanol pH 7.5) at 10 ml buffer per gram of wet cell paste and lysed with an Emulsiflex-C5 (Avestin, Canada). The lysate was cleared by centrifugation at 30 000*g* for 20 min and the supernatant was loaded onto nickel–nitrilotriacetic acid Superflow resin (Qiagen). The resin was washed with lysis buffer containing 1 *M* NaCl, followed by elution of the proteins with a linear gradient of 20–400 m*M* imidazole in lysis buffer. Protein peak fractions (typically >95% pure) were pooled, the buffer was exchanged to TEV protease buffer (50 m*M* Tris, 150 m*M* NaCl, 5 m*M* $MgCl_2$, 10% glycerol, 5 m*M* β-mercaptoethanol pH 8.0) using a Sephadex G-25 column and the sample was incubated with TEV protease in a 50:1 molar ratio at room temperature overnight to remove the His_6 tag. TEV protease, impurities and any remaining uncleaved His_6-tagged protein was removed by re-adsorption to nickel–nitrilotriacetic acid resin. The flowthrough was concentrated to 20 mg ml^{-1} using centrifugal filter devices (Centricon) and further purified on a 16/60 Superdex-75 column (Pharmacia) equilibrated and run in gel-filtration buffer (20 m*M* HEPES, 150 m*M* NaCl, 5 m*M* $MgCl_2$, 1 m*M* DTT pH 7.2). Pure peak fractions from the gel filtration were pooled, concentrated to 10 mg ml^{-1} and stored at 193 K for crystallization. The molecular weight of the protein was confirmed by Q-TOF mass spectroscopy. Selenomethionine-substituted protein was expressed in *E. coli* B834 (DE3) and purified following the same protocol. Arc1p_1–131 is monomeric in solution under physiological conditions (Deinert *et al.*, 2001). The monomeric state of Arc1p_1–122 (MW = 14.3 kDa) was confirmed by dynamic light scattering and gel filtration. The ability of Arc1p_1–122 and Arc1p_1–131 to form a stable complex with GluRS and MetRS was confirmed by copurifying both proteins with MetRS and GluRS in gel filtration on a 16/60 Superdex-200 column (results not shown). MetRS and GluRS were purified for this interaction analysis as described in Deinert *et al.* (2001).

3. Crystallization

Crystallization trials were set up at 293 K in 96-well format using a Mosquito nanolitre pipetting robot (TTP Labtech). Sitting drops of 600 nl (1:1 mixture of 10 mg ml^{-1} protein in gel-filtration buffer and reservoir solution) were equilibrated by vapour diffusion against reservoirs containing 75 µl screen solution in Crystal Quick plates (Greiner Bio-One). 15 screens of 96 conditions each were tested with both Arc1p_1–122 and Arc1p_1–131. While Arc1p_1–131 did not yield crystalline material in any of the 1440 conditions tested, Arc1p_1–122 crystallized under two conditions from custom grid screens. Only one of these conditions could be optimized to yield crystals suitable for X-ray analysis (Fig. 1). These crystals were optimized using microseeding in 24-well plates by vapour diffusion from hanging drops composed of equal volumes (2 + 2 µl) of protein and reservoir solution against 1 ml reservoir solution (35% PEG 3350, 100 m*M* $LiSO_4$, 50 m*M* Tris acetate pH 8.0).

4. X-ray data collection and analysis

Crystals were cryoprotected by sequential transfer to mother liquor with stepwise increased additions of glycerol up to 15% prior to freezing in liquid nitrogen. Preliminary crystal characterization was performed on an in-house X-ray source (Enraf–Nonius FR 751 rotating-anode generator, total reflection double-focusing mirrors, MAR 345 IP detector). Indexing was performed with *XDS* (Kabsch, 1993) and *MOSFLM* (Steller *et al.*, 1997) and gave unambigous and consistent results with different crystals (Fig. 2). The space group is *C*2, with unit-cell parameters a = 222.32, b = 89.46, c = 126.79 Å, β = 99.39°. Given the rather low molecular weight of the protein of

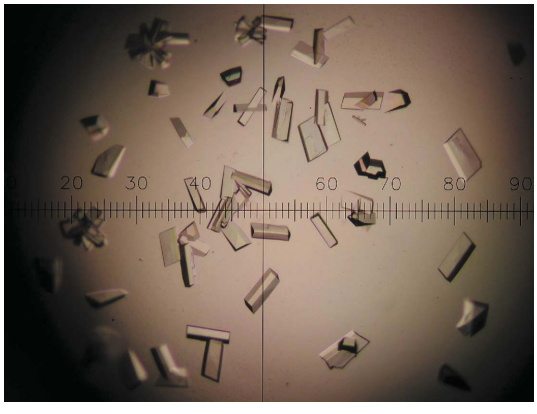


Figure 1
Photograph of typical Arc1p_1–122 crystals (maximum 200 µm in length).

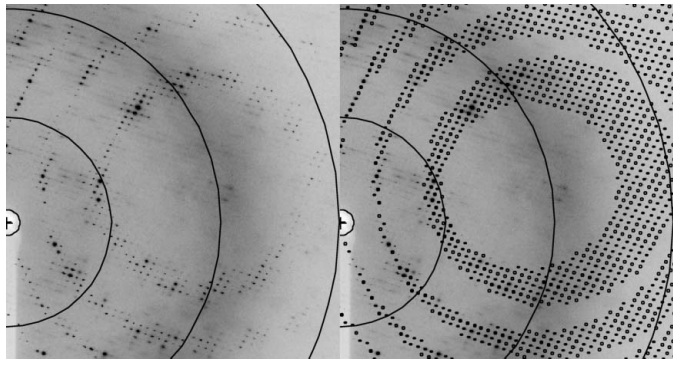


Figure 2
Left, representative area of a diffraction image. Right, the same area with spots predicted upon indexing with *MOSFLM*. The outermost resolution ring corresponds to 2.74 Å; the crystal rotation was 0.5°.

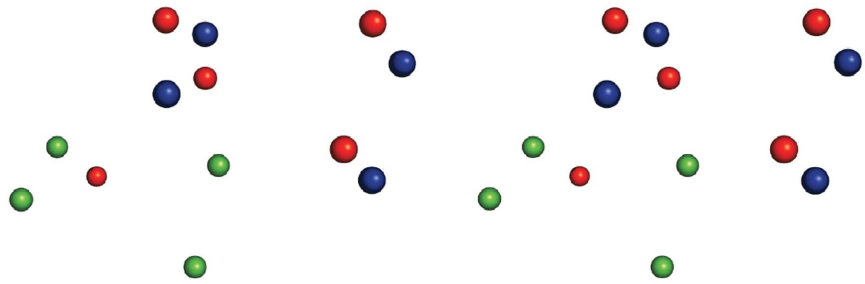


Figure 3
Recurring pattern regularities observed in different *SHELXD* selenium-site solutions suggest a tetrameric arrangement of Arc1p_1–122 monomers. A stereo representation of a partial selenium-site solution obtained from *SHELXD* searching for 16 sites on data in the resolution range 50–4 Å is shown. Two equivalent groups of four sites are shown in green and blue (accepted sites); sites not conforming to the observed regularities are shown in red (discarded sites).

14.3 kDa and the fact that it behaves as a monomer in solution, this is a surprisingly large unit cell. Calculated Matthews coefficients are compatible with the presence of between ten (V_M = 4.38 Å^3 Da^{-1}, V_S = 71.9% solvent) and 25 (V_M = 1.75 Å^3 Da^{-1}, V_S = 29.8% solvent) monomers in the asymmetric unit, suggesting a very high degree of non-crystallographic symmetry. However, the self-rotation function and the native Patterson yielded very complex patterns which did not prove helpful in elucidating the non-crystallographic symmetry within the asymmetric unit (results not shown). Complete multiple-wavelength anomalous dispersion data were collected at 100 K from a selenomethionine-substituted crystal at the peak and inflection-point wavelength of selenium and at a high remote wavelength (Table 1) at European Synchrotron Radiation Facility beamline ID-23. Data reduction and scaling were performed with *XDS*. Data-collection statistics are given in Table 1.

5. Preliminary phasing

In order to solve the phase problem for this structure, we decided to attempt experimental phasing *via* heavy-atom soaking and selenomethionine substitution in parallel; molecular replacement is clearly not an option owing to the lack of any structures significantly related to Arc1p_1–122 by sequence homology. The Arc1p_1–122 protein contains two methionines; however, the starting methionine is unlikely to be ordered and selenomethionine-substituted crystals are therefore expected to yield anomalous signal corresponding to one ordered methionine per 14.3 kDa of protein. Although this is a quite unfavourable ratio, useful phase information may be obtained if the high degree of non-crystallographic symmetry expected to be present in this crystal form could be exploited for NCS averaging successfully. *SHELXC* (Sheldrick, 1990) was used to analyse and prepare the MAD data set for a heavy-atom site search using *SHELXD*. Examination of the anomalous signal-to-noise ratio and of the correlation coefficients between signed anomalous differences of the peak, inflection-point and remote data sets as a function of resolution (not shown) suggested the presence of a useful anomalous signal for a heavy-atom site search to approximately 3.5 Å. As the number of monomers and selenium sites per asymmetric unit were initially unknown, we set up many parallel processes of *SHELXD* searching for variable numbers of selenium sites ranging from eight to 40 and applying different resolution limits to the searches. It was anticipated that the selenium-site occupancy statistics would hint at the actual number of selenium sites present in the asymmetric unit. Unfortunately, this strategy did not reveal a consistent and significant occupancy drop after any given number of sites and did not result in a recognisable solution. However, visual inspection of the output selenium sites corresponding to the top solutions from many different *SHELXD* processes revealed a recurring pattern among some of the sites superseded by noise. The observed pattern regularities suggested an arrangement of selenium sites in groups of four, which is likely to correspond to a tetrameric arrangement of individual Arc1p_1–122 monomers (Fig. 2). Upon closer examination of several *SHELXD* top solutions, we also observed regularities in the spacings between neighbouring putative tetrameric groups of selenium sites, which were approximately 43 Å along the *x*, 44.5 Å along the *y* and 64 Å along the *z* axes, suggesting the presence of 20 tetramers in the

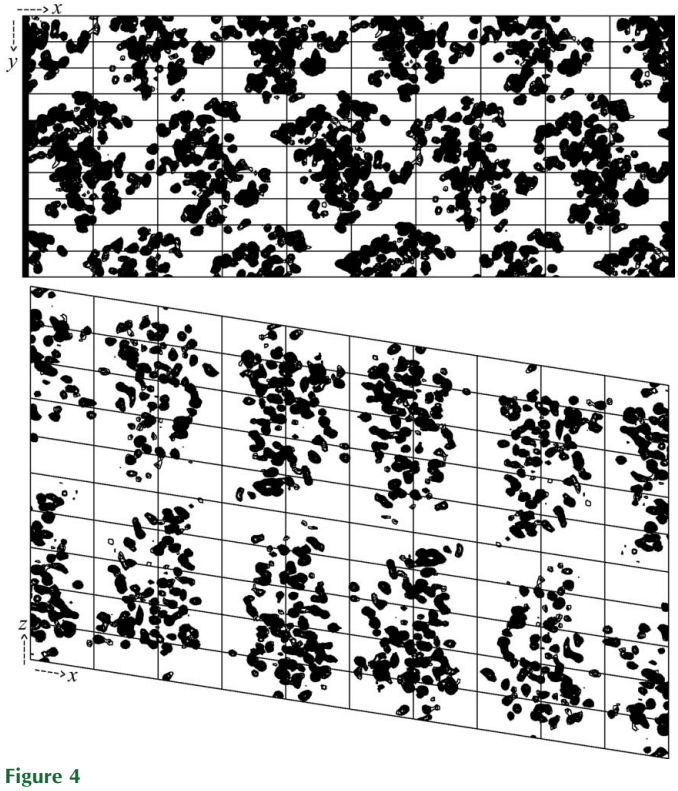


Figure 4
Preliminary phases reveal the presence of 20 tetramers in the unit cell. Data in the resolution range 50–3.5 Å were used to calculate phases and to perform ten cycles of density modification by *SHELXE* on the basis of 20 manually selected heavy atoms from *SHELXD* solutions. Plots of electron density show sections of the unit cell from 42 to 55 Å along the *z* axis and from 20 to 29 Å along the *y* axis. Density is contoured from 1.5 to 100σ in intervals of 0.5σ. Plots were prepared with *CCP*4 (Collaborative Computational Project, Number 4, 1994).

unit cell (five tetramers along *a* and two each along both *b* and *c*) which would correspond to a Matthews coefficient of 2.19 Å^3 Da^{-1} and a solvent content of 43.9%. Based on the observed pattern regularities, 20 selenium sites were manually selected or discarded from *SHELXD* output files (Fig. 3) and used as input to *SHELXE*. Preliminary phase information obtained after ten cycles of density modification in *SHELXE* assuming 43.9% solvent content yielded electron density with good solvent contrast, clearly confirming the presence of 20 regularly spaced tetramers in the unit cell (Fig. 4). While this initial map was not interpretable for model building, fivefold NCS averaging over the five tetramers contained in the asymmetric unit as implemented in *RESOLVE* (Terwilliger, 2000) has subsequently produced an interpretable map and model building is currently in progress.

The authors are grateful to J. Basquin and A. Scholz (EMBL Heidelberg) for crystallization screen design and setup and would like to thank the ID-23 beamline staff at the ESRF (Grenoble, France) for their assistance during data collection. This work was suported by the EU grant 3D-Repertoire contract No. LSHG-CT-2005-512028.

References

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Deinert, K., Fasiolo, F., Hurt, E. C. & Simos, G. (2001). *J. Biol. Chem.* **276**, 6000–6008.
Galani, K., Grosshans, H., Deinert, K., Hurt, E. C. & Simos, G. (2001). *EMBO J.* **20**, 6889–6898.
Galani, K., Hurt, E. C. & Simos, G. (2005). *FEBS Lett.* **579**, 969–975.
Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.
Kellermann, O., Tonetti, H., Brevet, A., Mirande, M., Pailliez, J. P. & Waller, J. P. (1982). *J. Biol. Chem.* **257**, 11041–11048.
Lee, S. W., Cho, B. H., Park, S. G. & Kim, S. J. (2004). *Cell Sci.* **117**, 3725–3734.
Martinis, S. A., Plateau, P., Cavarelli, J. & Florentz, C. (1999). *EMBO J.* **18**, 4591–4596.
Negrutskii, B. S. & Deutscher, M. P. (1991). *Proc. Natl Acad. Sci. USA*, **88**, 4991–4995.
Negrutskii, B. S. & Deutscher, M. P. (1992). *Proc. Natl Acad. Sci. USA*, **89**, 3601–3604.
Sheldrick, G. M. (1990). *Acta Cryst.* A**46**, 467–473.
Simos, G., Sauer, A., Fasiolo, F. & Hurt, E. C. (1998). *Mol. Cell*, **1**, 235–242.
Simos, G., Segref, A., Fasiolo, F., Hellmuth, K., Shevchenko, A., Mann, M. & Hurt, E. C. (1996). *EMBO J.* **15**, 5437–5448.
Stapulionis, R. & Deutscher, M. P. (1995). *Proc. Natl Acad. Sci. USA*, **92**, 7158–7161.
Steller, I., Bolotovsky, R. & Rossmann, M. (1997). *J. Appl. Cryst.* **30**, 1036–1040.
Terwilliger, T. C. (2000). *Acta Cryst.* D**56**, 965–972.