

Structural basis of yeast aminoacyl-tRNA synthetase complex formation revealed by crystal structures of two binary sub-complexes

Hannes Simader, Michael Hothorn, Christine Köhler, Jerome Basquin,
George Simos¹ and Dietrich Suck*

Structural and Computational Biology Unit, European Molecular Biology Laboratory (EMBL), Meyerhofstrasse 1, D-69117 Heidelberg, Germany and ¹Department of Medicine, University of Thessaly, 22 Papakiriazzi Street, Larissa, 41222, Greece

Received May 6, 2006; Revised July 18, 2006; Accepted July 19, 2006

ABSTRACT

The yeast aminoacyl-tRNA synthetase (aaRS) complex is formed by the methionyl- and glutamyl-tRNA synthetases (MetRS and GluRS, respectively) and the tRNA aminoacylation cofactor Arc1p. It is considered an evolutionary intermediate between prokaryotic aaRS and the multi- aaRS complex found in higher eukaryotes. While a wealth of structural information is available on the enzymatic domains of single aaRS, insight into complex formation between eukaryotic aaRS and associated protein cofactors is missing. Here we report crystal structures of the binary complexes between the interacting domains of Arc1p and MetRS as well as those of Arc1p and GluRS at resolutions of 2.2 and 2.05 Å, respectively. The data provide a complete structural model for ternary complex formation between the interacting domains of MetRS, GluRS and Arc1p. The structures reveal that all three domains adopt a glutathione *S*-transferase (GST)-like fold and that simultaneous interaction of Arc1p with GluRS and MetRS is mediated by the use of a novel interface in addition to a classical GST dimerization interaction. The results demonstrate a novel role for this fold as a heteromerization domain specific to eukaryotic aaRS, associated proteins and protein translation elongation factors.

INTRODUCTION

Aminoacyl-tRNA synthetases (aaRS) catalyze charging of their cognate tRNAs with the corresponding amino acid in a highly specific manner and may thus be considered the real translators of the genetic code. Due to the compartmentalization and

larger size of eukaryotic cells, their aaRS must be integrated into an efficient tRNA nuclear export- and shuttling machinery and interconnected closely with the translational apparatus. This functional requirement is reflected by the presence of additional protein- and RNA-binding domains and a correspondingly higher degree of complex formation among eukaryotic aaRS (1–3). The basic enzymatic function of pro- and eukaryotic aaRS is structurally very well understood (4,5). With public databases containing over 200 structural models, many including substrates, a wealth of structural information is available on these enzymes. However, in spite of many reported interactions between eukaryotic aaRS and protein cofactors (6–9), a structural model of a protein–protein interaction involving a eukaryotic aaRS has not been reported.

In prokaryotes, aaRS typically function as monomers or homo-dimers and do hardly ever engage in heteromeric protein interactions whereas heteromeric interactions are a hallmark of many eukaryotic aaRS. In higher eukaryotes, nine aaRS, including MetRS and GluProRS, associate to form a supra-molecular, multi-enzyme complex (10–12). Apart from the enzymes, this complex contains three accessory proteins: p18, p38 and p43 (13–15). In the yeast *Saccharomyces cerevisiae*, an evolutionary intermediate aaRS complex exists, formed by the two class I aaRS MetRS and GluRS and the tRNA aminoacylation cofactor Arc1p, which shares sequence homology with all three accessory components of higher eukaryotes (16). Biochemically, the yeast aaRS complex has been extensively characterized. It has been shown that both synthetases as well as both tRNA and tRNA can be efficiently co-purified with protein A tagged Arc1p from cellular yeast lysates (16). The N-terminal appended domains of yeast MetRS and GluRS as well as the N-terminal domain of Arc1p are necessary and sufficient for formation of the MetRS–Arc1p–GluRS ternary complex which appears to be stoichiometric (9,17,18). The N-terminal domain of Arc1p contains two binding sites for MetRS and GluRS which enable it to bind both synthetases

*To whom correspondence should be addressed. Tel: 0049 6221 387307; Fax: 0049 6221 387306; Email: suck@embl.de

simultaneously and independently (18). Furthermore, association of the synthetases with Arc1p strongly increases the affinity for the cognate tRNAs due to the presence of a tRNA binding domain (TRBD) in the middle and C-terminal part of Arc1p (17). In case of the synthetases, the N-terminal appended domains mediating complex formation are non-catalytic and absent in the prokaryotic counterparts. MetRS cannot directly interact with GluRS in the absence of Arc1p, demonstrating that the association is mediated by Arc1p (17,18). Both the MetRS–Arc1p and the GluRS–Arc1p sub-complexes are stable and appear to be fully functional in terms of aminoacylation efficiency (9,17). A high molecular weight complex containing all three protein components can be formed and maintained in a stable state from recombinant proteins. This complex selectively recruits its cognate tRNAs from total yeast tRNA, demonstrating that no other cellular factors are required for complex formation and stabilization (9). In addition to its well established role in promoting aminoacylation of tRNA^{Met} and tRNA^{Glu} (16,17), Arc1p has been suggested to play a role in nuclear export of tRNAs based on synthetic lethality of *arc1* mutations with mutations of the tRNA exportin Los1p (16), the requirement of aminoacylation for efficient tRNA nuclear export (19) and XpoI—dependent nuclear exclusion of Arc1p (20).

We have chosen the yeast aaRS complex as a well characterized model system in order to investigate at the structural level how two eukaryotic aaRS interact with a protein cofactor. In order to understand how the N-terminal domain of Arc1p can interact simultaneously and independently with both synthetases, we have solved high resolution crystal structures of complexes between the Arc1p and GluRS as well as Arc1p and MetRS interacting domains.

MATERIALS AND METHODS

Cloning, expression and purification

The N-terminal domains of *S.cerevisiae* Arc1p (residues 1–122, referred to as Arc1p-N), GluRS (residues 1–207, referred to as GluRS-N) and MetRS (residues 1–160, referred to as MetRS-N) were PCR amplified from plasmids pET-HIS6/pET8c-Arc1p (16), pEMBLyex4-His8-GluRS and pEMBLyex4-His8-MetRS (9), respectively, and cloned into a modified pETm (Novagen) vector allowing the removal of the 6His-tag through TEV protease cleavage. Proteins were expressed and purified following the procedure described for Arc1p-N (21). For complex preparation, Arc1p-N and GluRS-N or Arc1p-N and MetRS-N were mixed using a slight molar excess of Arc1p-N prior to gel-filtration, concentrated to 20 mg/ml using centrifugal filter devices (Centricon) and the complex was purified from excess Arc1p-N on a 16/60 Superdex-75 column equilibrated in 20 mM HEPES, 150 mM NaCl, 5 mM MgCl₂, 1 mM DTT (pH 7.2 w.) NaOH. Protein samples were concentrated to 15 mg/ml (MetRS-N–Arc1p-N complex) and 20 mg/ml (GluRS-N–Arc1p-N complex) prior to crystallization.

Crystallization and data collection

Crystallization of Arc1p-N has been described elsewhere (21). The GluRS-N–Arc1p-N complex and the MetRS-N–Arc1p-N complex were crystallized following the same basic

procedure. Briefly, crystals were grown by vapor diffusion from hanging drops composed of equal volumes (typically 2 + 2 µl) of protein and reservoir solution (30–35% PEG 3350, 300–500 mM NaSCN for GluRS-N–Arc1p-N crystals; 8–14% PEG 20.000, 1–3% dioxane, 100 mM BICINE (pH 9.0) for MetRS-N–Arc1p-N crystals). MetRS-N–Arc1p-N crystals were cryo-protected by serial transfer through solutions of stepwise increasing concentrations of ethylene glycol up to 30% and flash frozen in liquid nitrogen. Crystals of GluRS-N–Arc1p-N were mounted in sealed capillaries and data collection on an in-house rotating copper anode X-ray source yielded a dataset at 3.0 Å resolutions. A thin plate frozen without addition of cryo-protectant later yielded a dataset to 2.05 Å on the same X-ray source but with significantly altered unit cell parameters (Table 1). MetRS-N–Arc1p-N native data to 2.5 Å used for structure solution were collected on an in-house rotating copper anode X-ray source and data to 2.2 Å used for refinement at the ESRF beam-line ID14-4. Data processing and scaling was performed with XDS (22). Data collection statistics are given in Table 1.

Structure determination and refinement

Both structures were solved by molecular replacement with the programs MOLREP (23) (GluRS-N–Arc1p-N) or Phaser (24) (MetRS-N–Arc1p-N) from the CCP4 suite (25,26) using either chain B of the GluRS-N structure (GluRS-N–Arc1p-N) (H. Simader, M. Hothorn and D. Suck, manuscript submitted) or chain C or the Arc1p-N structure (MetRS-N–Arc1p-N) [(21) and H. Simader, M. Hothorn and D. Suck, manuscript submitted] as a search model. In case of the MetRS-N–Arc1p-N structure, prime and switch minimal bias phasing, density modification and phase extension was carried out as implemented in RESOLVE (27). The resulting electron density map displayed sufficient features for the MetRS-N part of the structure to start model building. Both structures were built in COOT (28) and refined in REFMAC5 (29).

Mutagenesis and interaction analysis

In addition to the plasmids for expression of Arc1p-N, MetRS-N and GluRS-N described above, full-length Arc1p, GluRS and MetRS were cloned into expression vectors as described above. Single amino acid substitution mutants of Arc1p, MetRS and GluRS were created with the Quick Change II Site-Directed Mutagenesis Kit (Stratagen) according to manufacturers information. Mutant and wild-type proteins were expressed and purified as described above except that full-length mutant and wild-type GluRS and MetRS were expressed in BL21 DE3 pRARE cells. Interaction analysis of mutated and wild-type Arc1p, GluRS and MetRS was performed by a pull-down assay using as bait protein 6 His-tagged wild-type or mutant GluRS or MetRS bound to Ni-NTA beads with untagged wild-type or mutant Arc1p as prey proteins. Ni-NTA super-flow resin (QIAGEN) was washed and equilibrated in binding buffer [50 mM Tris, 150 mM NaCl, 20 mM imidazole, 5 mM MgCl₂, 10% glycerol, 10 mM β-mercaptoethanol (pH 7.5 w.) HCl]. A total of 10 µl Ni-NTA beads were mixed with 50 µl of bait protein at 10 µM concentration in binding buffer and incubated 5 min at room temperature for binding bait protein to the beads. Beads were washed twice with 1 ml binding buffer, 100 µl

Table 1. Data collection and refinement statistics

Crystal	Arc1p-N–MetRS-N complex	Arc1p-N–GluRS-N complex Room temperature	Cryo (100K)
Data collection			
X-ray source/wavelength (Å)	ESRF ID14-4 0.9395	EMBL in house 1.5418	EMBL in-house 1.5418
Space group	P4(3)2(1)2	P 2(1)	P 2(1)
Cell dimensions			
α, β, γ (Å)	94.5, 94.5, 88.5	41.4, 87.5, 52.8	40.3, 86.3, 47.1
a, b, c (°)	90, 90, 90	90, 92.6, 90	90, 99.4, 90
Asymmetric unit (chains/residues)	2/286	2/333	2/333
V _M /% solvent content	2.96/58.1	2.94/57.8	2.21/43.8
Resolution (Å)	2.2 (2.3–2.2)	3.0 (3.18–3.0)*	2.05 (2.10–2.05)*
R _{sym}	0.056 (0.263)	0.097 (0.311)	0.038 (0.283)
I/σI	21.5 (4.96)	10.6 (3.92)	27.5 (5.75)
No. observations	228 072	22 888	81 645
Completeness (%)	99.3 (100.0)	98.1 (79.6)	99.9 (100.0)
Redundancy	10.9 (11.3)	3.1 (2.8)	4.1 (4.0)
Refinement			
Resolution (Å)	2.2 (2.26–2.20)	3.0 (3.08–3.00)	2.05 (2.10–2.05)
No. used reflections	19 731	7 075	18 982
R _{work} /R _{free}	0.225/0.279	0.195/0.251	0.191/0.256
No. atoms			
Protein	2203	2322	2314
Ligand/ion	—	0	1
Water	181	0	305
Mean B-factors			
Protein	57.3	47.4	34.5
Ligand/ion	—	—	36.8
Water	48.3	—	48.3
R.m.s.d.			
Bond lengths (Å)	0.01	0.01	0.01
Bond angles (°)	1.27	1.32	1.02
Ramachandran statistics			
% Residues in core	90.1	79.6	92.4
Allowed	7.6	16.8	6.9
Generously allowed	1.1	1.8	0
Disallowed region	1.2	1.8	0.7

*Highest resolution shell is shown in parenthesis.

of prey protein at 10 μM concentration in binding buffer were added to the beads and incubated 5 min at room temperature. Excess and unspecifically bound prey protein was removed by washing once with 1 ml binding buffer, once with 1 ml binding buffer +1 M NaCl and once with 1 ml lysis buffer +40 mM imidazol. Proteins were eluted by addition of 50 μl binding buffer +400 mM imidazol and protein content of the eluate was analyzed on SDS–PAGE. Independent evidence confirming the pull-down results was obtained by analysis of equimolar mixtures of bait and prey proteins using gel-filtration chromatography on 16/60 Superdex-75 and Superdex-200 columns and subsequent analysis of eluting fractions on 8–20% gradient SDS–PAGE.

Miscellaneous

Complex interfaces were analyzed using the Protein interfaces, surfaces and assemblies service PISA at the European Bio-informatics Institute (30). Molecular graphics were prepared with Pymol (31) (<http://www.pymol.org>), multiple structural alignments were computed using MAMMOTH-mult (32) and multiple sequence alignments were prepared with Clustal (33,34). The structural model of the full ternary complex was derived by superimposition of the Arc1p-N moieties of both GluRS-N–Arc1p-N and MetRS-N–Arc1p-N complex crystal structures (Figure 8). This approach

is justified as both complex structures show that the conformation of the Arc1p-N-terminus is not affected by binding of either complex partner (Figure 8). The RMSD for superposition of the Arc1p-N moiety of both complexes is 0.67 Å and thus within the range normally observed for identical structures solved in different crystal forms due to differences in crystal packing contacts. The putative non domain-swapped monomer of MetRS-N was modeled as indicated in Figure 1 by exchanging the swapped domain, i.e. residues 1–56, of the crystal structure with the corresponding residues from the domain-swapped symmetry mate and by subsequently correcting the stereo-chemistry of the connecting residues by regularization over residues 54–58 in COOT (28).

RESULTS

The MetRS-N–Arc1p-N complex reveals domain swapping between symmetry related molecules of MetRS-N

The crystal structure of the MetRS-N–Arc1p-N complex (Table 1) reveals domain swapping between symmetry related molecules of MetRS-N, effectively generating tetramers of 2:2 compositions (Figure 1). The N-terminal 55 amino acids, comprising most of the β-sheet and the

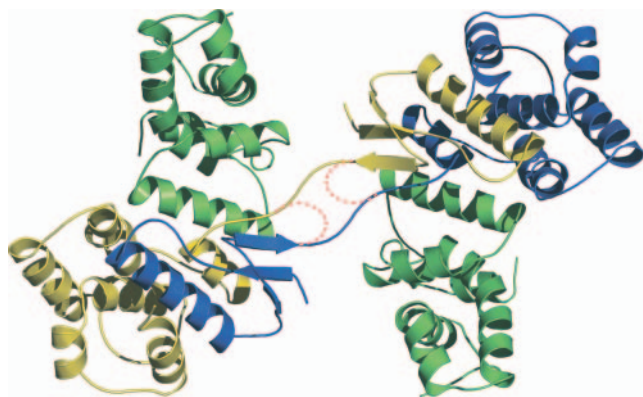


Figure 1. The crystal structure of the MetRS-N–Arc1p-N complex reveals domain swapping between symmetry related molecules of MetRS-N. The N-terminal 55 amino acids, comprising most of the β -sheet and the first α -helix, are exchanged between neighboring molecules related by a true crystallographic 2-fold axis. Monomers of MetRS-N are shown in yellow and blue, Arc1p-N is shown in green. The red dotted lines indicate how the putative monomeric MetRS-N was modeled. See text for explanation and discussion.

first α -helix, are exchanged between neighboring molecules related by a true crystallographic 2-fold axis.

This domain swapping, if it was biologically relevant, would imply the structural architecture of the full yeast aaRS complex as a hexameric complex of 2:2:2 stoichiometry. However, it is unlikely to occur *in vivo* as full-length MetRS purified from yeast behaves as a monomer in gel-filtration experiments under near-physiological conditions (9) and using full-length recombinant protein expressed in yeast, the composition of the yeast aaRS complex is very well established as a ternary complex of 1:1:1 stoichiometry (9). Yeast MetRS-N purified from *Escherichia coli* was found to exist as a mixture of two conformers corresponding in MW to monomers and dimers as judged on the basis of gel-filtration and dynamic light scattering experiments under near-physiological conditions. Both conformers can be efficiently separated by gel-filtration (data not shown) and appear to remain stable, i.e. no shift of small to big conformer or vice versa is observed during a 24 h incubation period (data not shown). The mass distribution between the two conformeric states was found to be mainly determined by the temperature during protein expression. Since the small conformer could not be crystallized, we have used the conformer corresponding in size to a dimer of MetRS-N for formation and crystallization of the MetRS-N–Arc1p-N complex, suggesting that this conformer corresponds to the domain-swapped dimer of MetRS-N revealed by our crystal structure. It is therefore likely that the observed domain swapping is an artifact resulting from either heterologous protein expression in *E. coli*, C-terminal protein truncation or N-terminal modification and processing (expression with affinity tag and subsequent cleavage by TEV protease) or over-expression (molecular crowding) in cells, or a combination thereof. It is unlikely to be an artifact of protein crystallization since presence of MetRS-N conformers corresponding in size to dimers was observed during protein purification prior to crystallization. The observed domain swapping of MetRS-N does not appear to affect specific binding of MetRS-N to Arc1p in solution under near-physiological conditions as gel-filtration fractions

of both MetRS-N conformers equally co-purified with Arc1p at apparent 1:1 (or 2:2, respectively) stoichiometry in gel-filtrations whereas unspecific copurification with other proteins was not observed during protein purification from *E. coli* lysates (data not shown).

Domain swapping is being described in an increasing number of cases and has been analyzed and discussed elsewhere (35–38). The most relevant finding that has emerged from these studies in respect to our case is that native protein topology is the major conformational determinant in domain swapping, implying that the native monomeric conformation of a protein is still truly reflected by its domain-swapped conformer (35–38). In conclusion (i) the observed domain swapping is unlikely to be of biological relevance but most likely represents an artifact of heterologous recombinant protein expression, not crystallization; (ii) it does not appear to affect specific binding to Arc1p under near-physiological conditions in solution; (iii) research from a number of labs suggests that the native conformation of the MetRS-N domain is truly reflected by our domain-swapped structure.

Since domain-swapping dimerization is known not to occur in MetRS expressed endogenously in yeast cells (9), we have chosen to model the putative non-swapped monomer of MetRS-N. All subsequent structural analysis and figures were prepared based on the domain-swap corrected model of the MetRS-N domain.

All three interacting domains adopt a GST-like fold

The Arc1p-N-terminus was originally described as a domain with little or no homology to known proteins (16,17). For the GluRS and Arc1p interacting domains, a relationship to glutathione *S*-transferase (GST)-like proteins has later been predicted based on the conserved sequence motifs (18). In fact, the crystal structures reveal that all three domains adopt a GST-like fold (Figure 3) and searching the Protein Data Bank for proteins structurally related to Arc1p-N, GluRS-N and MetRS-N with DALI (39,40) reveals matches to GSTs from diverse species as bacteria, yeast, plants and humans. A multiple structure based sequence alignment (Figure 2 and Supplementary Table 1) of Arc1p-N, GluRS-N and MetRS-N with their closest structural homologs reported by DALI shows superimposed sequence conservations of 8–19% (average 12%) and reveals RMSDs of 2.2–4.1 Å (average 3.1 Å; Supplementary Table 1). Given that the sequence homology is below the detection level in many cases, the conservation on the structural level is remarkably high (41). The Arc1p-N-terminal domain corresponds to only the C-terminal α -helical part of the classical GST fold [GST_C in PFAM (42)], it lacks the N-terminal α/β motif, which contributes most of the substrate binding residues in classical GSTs. In the GluRS and MetRS-N-termini in contrast, both domains of the classical GST fold are conserved.

The MetRS-N–Arc1p-N hetero-dimer resembles a classical GST homo-dimer

Classical GSTs form homo-dimers and interact in a characteristic way that relates both monomers by 2-fold rotational symmetry, shown exemplarily for *E. coli* GST in Figure 3A. In the MetRS-N–Arc1p-N complex, the orientation of the

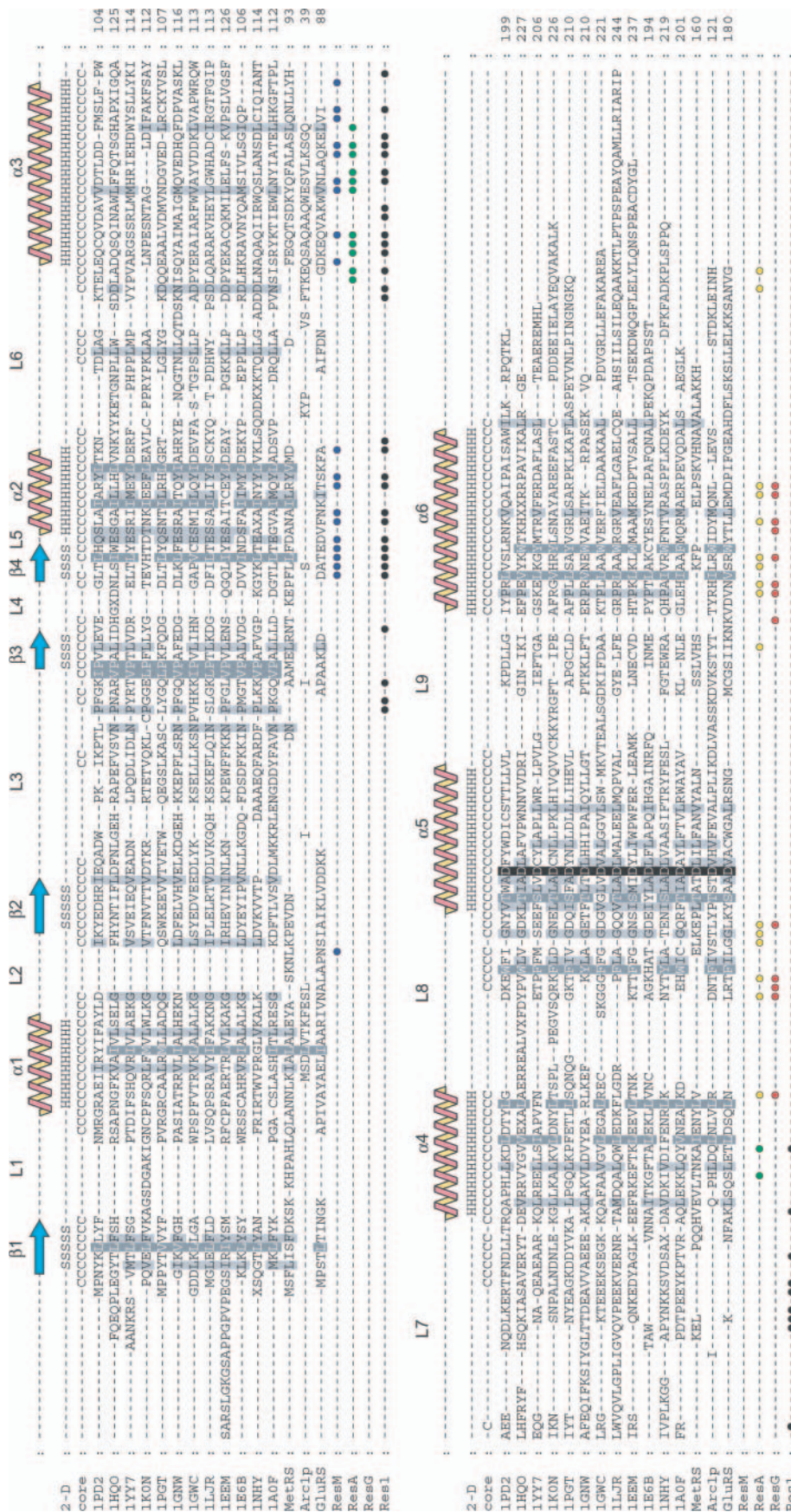


Figure 2. Multiple structure based sequence alignment of Arc1p-N, MetRS-N and GluRS-N with their most similar structures reported by DALI. Shown are the top 12 DALI hits (labeled by their pdb accessions, supplementary Table 1) structurally aligned with Arc1p-N, MetRS-N and GluRS-N shaded by conservation at three levels. Residues labeled (*) in the line termed 'core' are part of the structural alignment core. Secondary structure elements of the classical GST fold are indicated via S for β-sheets and H for α-helices in the line termed '2-D' and represented graphically above the alignment. Residues of MetRS-N interacting with Arc1p-N are indicated by blue dots (in line ResM), residues of Arc1p-N interacting with MetRS-N by green dots (in line ResA), residues of Arc1p-N interacting with GluRS-N by yellow dots (in line ResA), residues of GluRS-N interacting with Arc1p-N by red dots (in line ResG) and contacting residues of the *E. coli* GST homo-dimer (Pdb accession 1A0F) are indicated by black dots (in line Res1).

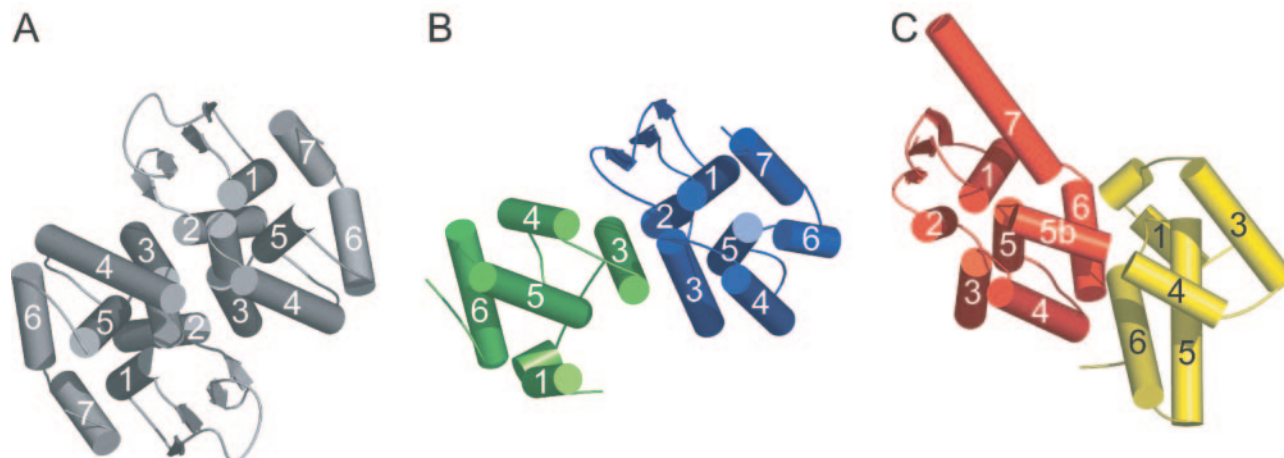


Figure 3. The architecture of the Arc1p-N–MetRS-N complex corresponds to the classical GST homo-dimer while Arc1p-N and GluRS-N interact in a novel way. (A) *E. coli* GST homo-dimer (gray, pdb accession 1A0F) viewed along the 2-fold rotational symmetry axis. (B) The MetRS-N–Arc1p-N complex viewed in an orientation corresponding to that shown for *E. coli* GST in (A). Arc1p-N is shown in green, MetRS-N in blue. (C) The GluRS-N–Arc1p-N complex viewed in an orientation corresponding to GluRS-N superposition over the top chain of *E. coli* GST in (A) and over MetRS-N in (B). Arc1p-N is shown in yellow and GluRS-N in red. Corresponding α -helices of the GST domain are labeled 1–7.

interacting domains towards each other (Figure 3B) and the structural elements most prominently involved in the interaction (Figure 2) correspond to that seen in the classical GST homo-dimerization. While the overall architecture of the MetRS-N–Arc1p-N complex is rather similar to the classical GST homo-dimer, the size of the interface and the number and identity of the residues involved are different. Due to the lack of the N-terminal α/β motif in Arc1p and due to a dramatic reduction of loop 7 in both Arc1p and MetRS compared to classical GSTs (Figure 2), the interface between MetRS-N and Arc1p-N is reduced to roughly half the size of the *E. coli* GST interface (<1600 Å² buried surface area as opposed to over 3200 Å²). The reduced interface surface area is matched by a corresponding reduction in the number of contacting residues. A total of 20 residues from MetRS and 13 residues from Arc1p are mediating the formation of the Arc1p-N–MetRS-N complex as compared to 39 residues from each monomer in the *E. coli* GST homo-dimer (Figures 2 and 4). While the contacting residues are mostly found in structurally corresponding positions, none of them are significantly conserved between Arc1p, MetRS and classical GSTs (Figures 2 and 4). The pseudo-homomeric nature of the complex at the level of the overall architecture is thus not reflected by homology at the level of the contacting residues. Analysis of the interface charge distribution suggests that the complex formation is probably not mediated by electrostatic forces but rather by Van-der-Waals forces. This is supported by the finding that non-polar atoms account for 70% of the interface surface. The polar components of the interface contribute five hydrogen bonds but no salt bridges, consistent with the observation that the electrostatic surface potential of Arc1p-N and MetRS-N does not reveal any strong complementary electrostatic patches for the interface surface or any obvious differences between the interface surface and the remaining solvent accessible surface (data not shown). Multiple protein sequence alignments of Arc1p-N and MetRS-N with orthologous sequences from other yeasts suggests that the contacting residues are on average no more conserved than other residues not involved

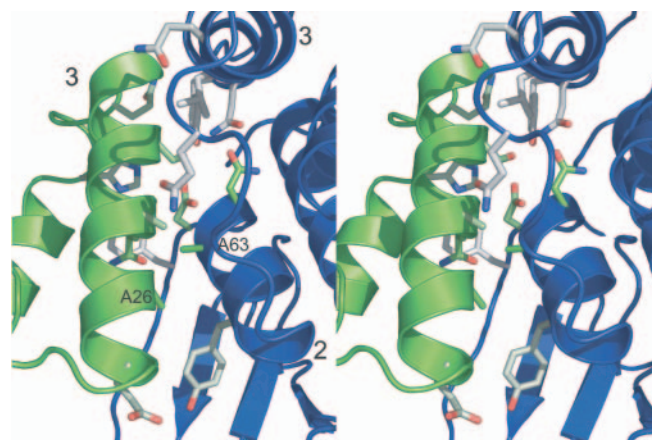


Figure 4. The MetRS-N–Arc1p-N interface. Stereo-view of the MetRS-N (blue)–Arc1p-N (green) complex interface with helix labels as in Figure 3B. Selected contacting residues are shown in stick mode and colored in a linear gradient from white (10%) to green (90%) by sequence conservation among orthologous sequences (Figure 5). The two conserved alanines are labeled A26 (Arc1p) and A63 (MetRS).

in the interaction (Figure 5A and B). Mapping the conservation scores derived from these alignments onto the surface of Arc1p-N and MetRS-N demonstrates that the interface surface is no more conserved than the average solvent accessible surface (data not shown). This unexpected finding suggests a high plasticity of this interface.

However, there are some interface residues that are remarkably conserved and interface analysis suggests why this may be the case. Highly conserved alanines are found at residue number 26 of Arc1p and 63 of MetRS (Figure 5A and B). In our structural model, both of these alanines in central positions of the interface were sterical constraints are incompatible with the presence of much larger side-chains. This interpretation is confirmed by the finding that mutagenesis of either Arc1p alanine 26 to arginine or MetRS alanine 63 to histidine abolishes formation of a stable

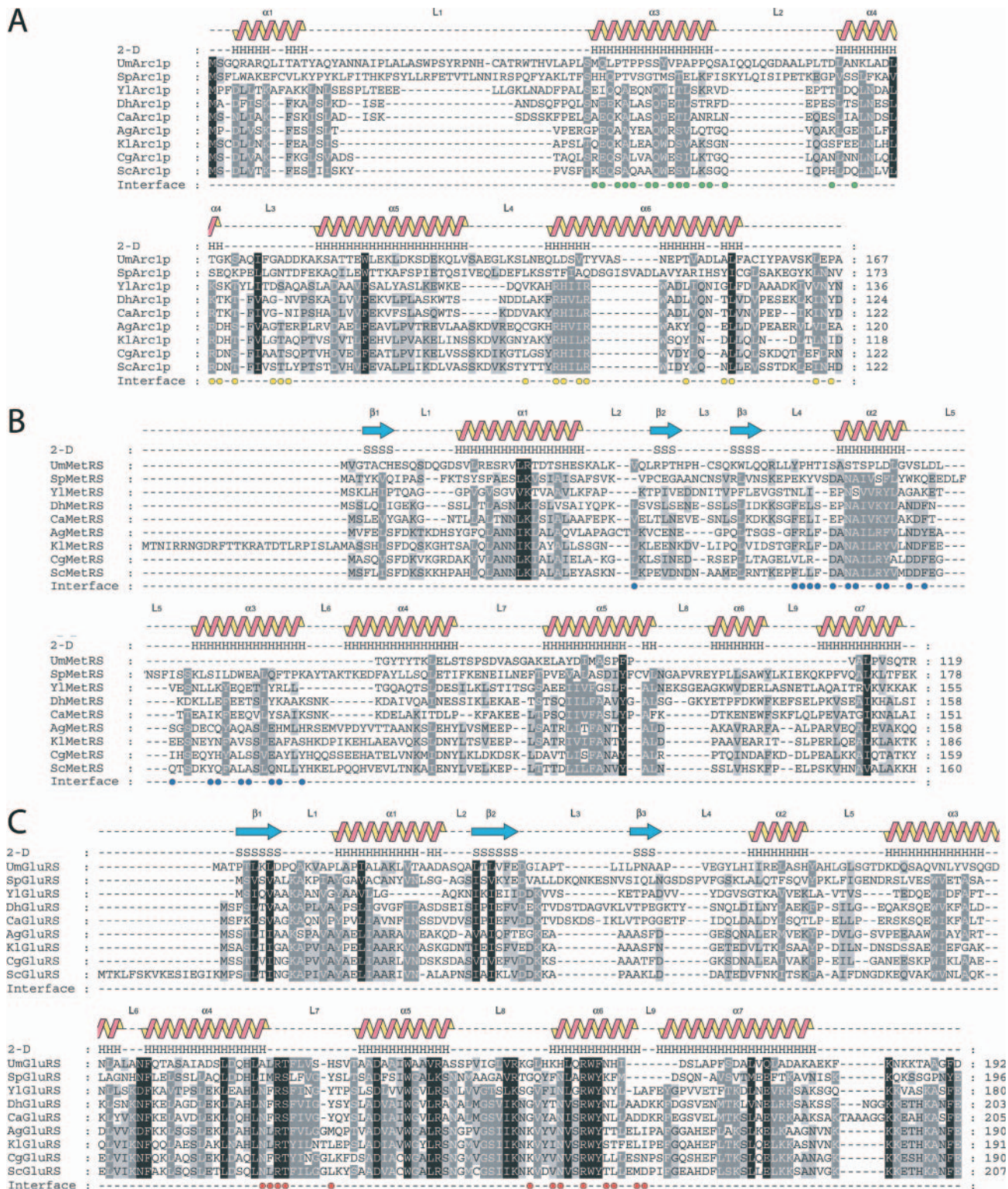


Figure 5. Multiple sequence alignment of *S.cerevisiae* Arc1p-N (A), MetRS-N (B) and GluRS-N (C) with orthologs from other yeasts. Shown are the sequences from *Ustilago maydis* (Um), *Schizosaccharomyces pombe* (Sp), *Yarrowia lipolytica* (Yl), *Debaryomyces hansenii* (Dh), *Candida albicans* (Ca), *Kluyveromyces lactis* (Kl), *Ashbya gossypii* (Ag) and *S.cerevisiae* (Sc) shaded by conservation at three levels. Secondary structure elements of Arc1p-N, MetRS-N and GluRS-N are indicated via S for β -sheets and H for α -helices in the line termed "2-D" and represented graphically above the alignment. Contacting residues are color coded as in Figure 2. Sequences correspond to the following GeneBank accessions: *U.maydis* Arc1p: XP_400509.1, GluRS: XP_402397.1, MetRS: XP_401146.1; *S.pombe* Arc1p: NP_594656.1, GluRS: NP_593483.1, MetRS: NP_595586.1; *Y.lipolytica* Arc1p: XP_503499.1, GluRS: XP_504508.1, MetRS: XP_506024.1; *D.hansenii* Arc1p: XP_456881.1, GluRS: XP_461343.1, MetRS: XP_462423.1; *C.albicans* Arc1p: XP_713255.1, GluRS: XP_720349.1, MetRS: XP_721864.1; *K.lactis* Arc1p: XP_455553.1, GluRS: XP_451028.1, MetRS: XP_451421.1; *A.gossypii* Arc1p: NP_985516.1, GluRS: NP_985811.1, MetRS: NP_986998.1; *S.cerevisiae* Arc1p: X95481, GluRS: P46655, MetRS: P00958.

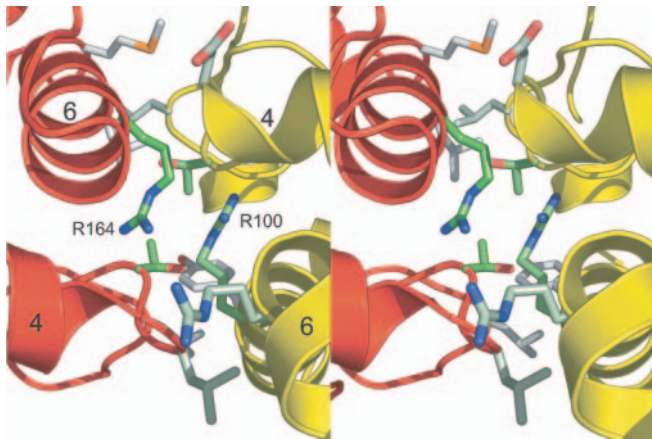


Figure 6. The GluRS-N–Arc1p-N interface. Stereo-view of the GluRS-N (red)–Arc1p-N (yellow) complex interface with helix labels as in Figure 3C. Selected contacting residues are shown in stick mode and colored in a linear gradient from white (10%) to green (90%) by sequence conservation among orthologous sequences (Figure 4). The two conserved arginines engaged in a stacking interaction at the center of the interface are labeled R164 (GluRS) and R100 (Arc1p).

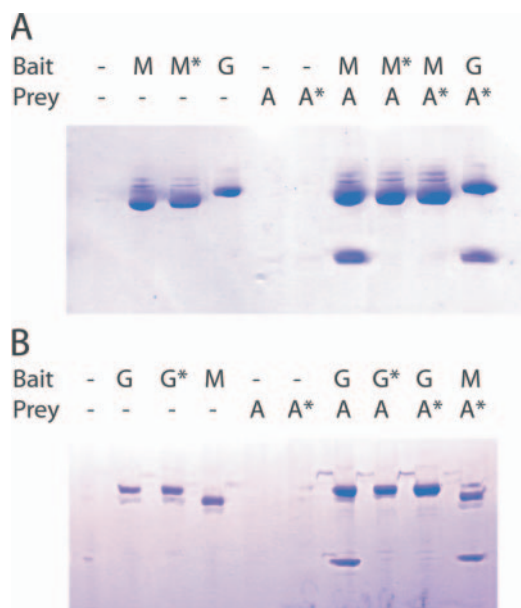


Figure 7. Several single amino acid point mutations rationally designed based on the complex crystal structures abolish formation of a stable complex between full-length Arc1p and MetRS or GluRS in solution under near-physiological conditions while they do not interfere with stable association of Arc1p with the other enzyme. (A) Mutants Arc1p(A26R) and MetRS(A63H) interfere with Arc1p–MetRS association (B) Mutants Arc1p(R100A) and GluRS(R164A) interfere with Arc1p–GluRS association. Shown are coomassie stained SDS–PAGE analysis of pull-down assays using as bait 6His-tagged GluRS (G), GluRS(R164A) (G*), MetRS (M) or MetRS(A63H) (M*) bound on Ni–NTA beads with untagged Arc1p (A), Arc1p(A26R) (A* in A) or Arc1p(R100A) (A* in B) as prey proteins.

complex between Arc1p and MetRS but does not interfere with stable association of Arc1p with GluRS (Figure 7A).

Our structural model and our interpretation of the interaction between MetRS and Arc1p is largely independent of the domain–swapping phenomenon observed between MetRS-N

monomers as the interface between Arc1p-N and MetRS-N is formed to more than 97.5% by the non-swapped core of the MetRS-N domain. This finding explains why both conformer fractions of MetRS-N were found to be equally active in Arc1p binding.

The GluRS-N–Arc1p-N complex represents a novel mode of interaction

Since the interface corresponding to the classical GST dimerization interface is used for the interaction between Arc1p and MetRS, but Arc1p is known to be capable of simultaneous and independent interaction with both MetRS and GluRS, a novel interface must mediate the interaction between Arc1p and GluRS. Indeed, the complex between GluRS-N and Arc1p-N shows both proteins interacting in a novel mode entirely different from that seen in classical GST homo-dimers (compare Figure 3A and C). The complex structure reveals that, although this is a heteromeric interaction, the same structurally corresponding elements from each complex partner (residues of helix 6 and loop 8 of the GST domain, Figures 2 and 3C) are involved in the interaction, relating the GST fold of both components by a pseudo 2-fold rotation axis. In accordance with this pseudo-homomeric nature, as in the case of the Arc1p/MetRS interface, the contacting residues are mostly found in structurally corresponding positions (Figure 2). While the GluRS/Arc1p interface involves different structural elements, it is similar to the Arc1p/MetRS interface in several other respects. The size ($\sim 1600 \text{ \AA}^2$ total buried surface area) is almost identical to that of the Arc1p/MetRS interface and the number of contacting residues (17 from Arc1p and 13 from GluRS) are comparable as well (Figures 2 and 6). As for the Arc1p–MetRS complex, analysis of the physico–chemical properties of the interface surface suggests that the complex formation is mediated predominantly by van-der-Waals forces rather than by electrostatic forces. Non-polar atoms account for over 60% of the interface surface, while the polar components of the interface contribute just one hydrogen bond and one salt bridge to the interface. Multiple protein sequence alignments of Arc1p-N and GluRS-N with orthologous sequences from other yeasts suggests that the contacting residues are on average not significantly more conserved than other residues not involved in the interaction (Figure 5A and C). As for the Arc1p–MetRS interaction, this finding suggests a high plasticity of the interface. However, some of the interacting residues are conserved in the GluRS-N–Arc1p-N interaction. A tetrad of structurally corresponding arginines and threonines appears to play a special role at the center of the complex interface (Figure 6). Interestingly, both residues are conserved not only among the orthologous groups of Arc1p and GluRS (Figure 5A and C), but also between the two (Figure 2). The guanidinium groups of Arc1p R100 and GluRS R164 are $\sim 4.1 \text{ \AA}$ apart and thus properly positioned to engage in a stacking interaction (43). The importance of this arginine stacking interaction for Arc1p–GluRS complex formation is reflected by the finding that mutagenesis of either arginine to alanine abolishes formation of a stable complex between full-length Arc1p and GluRS under near-physiological conditions in solution while it does not interfere with stable association of Arc1p with MetRS (Figure 7B).

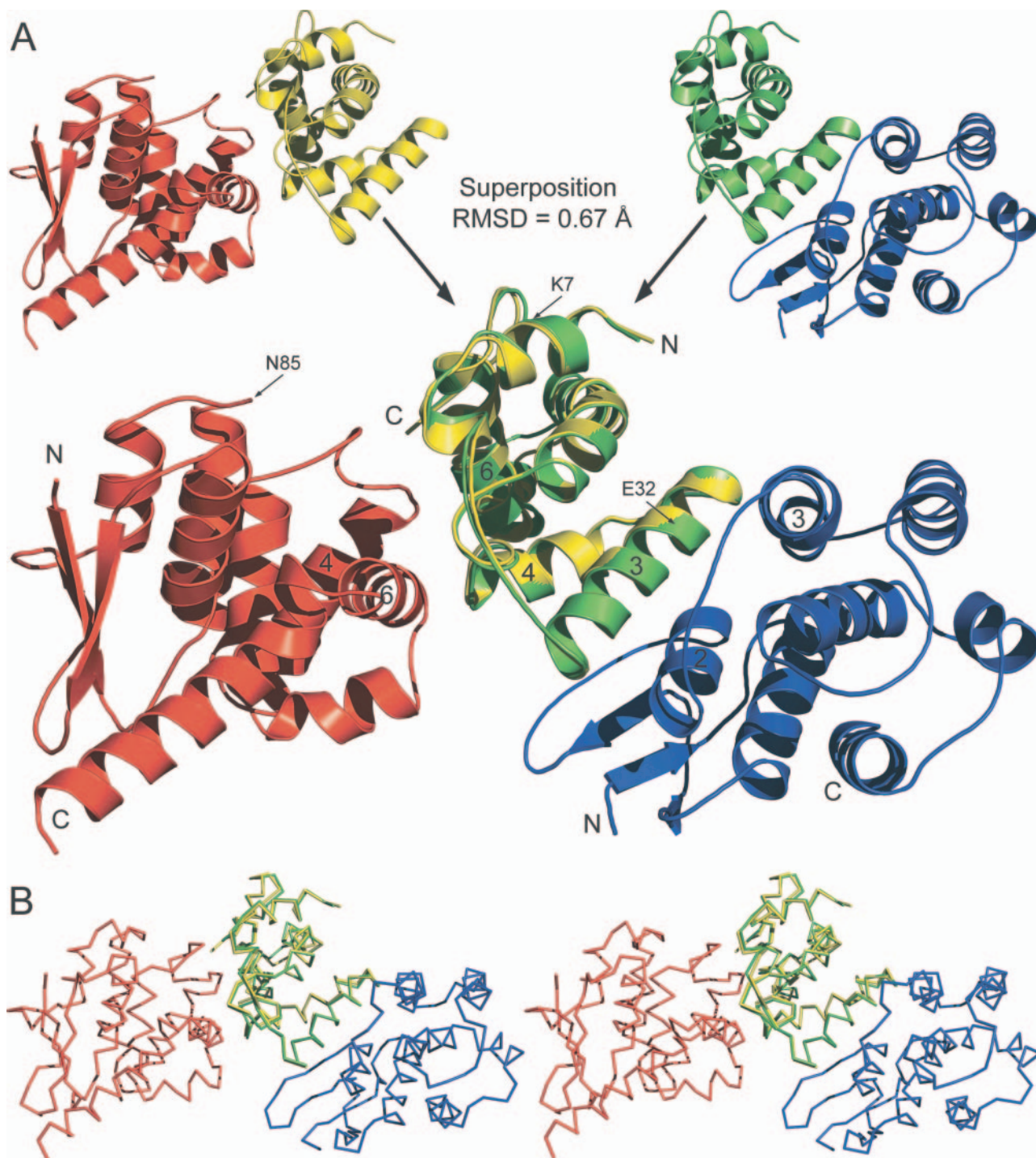


Figure 8. Model of the GluRS-N-Arc1p-N-MetRS-N ternary complex. Superimposing the Arc1p-N parts of the two binary sub-complexes reveals the structural organization of the ternary complex. (A) The Arc1p-N parts from the MetRS-N-Arc1p-N and GluRS-N-Arc1p-N complexes superpose with an RMSD of 0.67 Å, revealing the structural organization of the ternary complex. Helices involved in the interactions are labeled as in Figure 3 and N-termini (N) and C-termini (C) of each protein are indicated. Residue labels (Arc1p K7, E32 and GluRS-N85) indicate boundaries of the deletion constructs used in. (B) Stereo-view of the GluRS-N (red)-Arc1p-N (yellow) and Arc1p-N (green)-MetRS-N (blue) complexes with their Arc1p-N moieties superposed.

The contribution of the threonines to complex formation is less clear except for the constraints requiring small side-chains in these positions: mutagenesis of either threonine to arginine also abolishes formation of a stable complex between Arc1p and GluRS but does not interfere with stable association of Arc1p with MetRS (data not shown).

Structure of the ternary yeast aaRS complex

The structural model of the ternary complex, as derived by superimposing the Arc1p-N domains of the two binary complexes (Figure 8), explains many previous biochemical findings. It explains how Arc1p-N can interact simultaneously

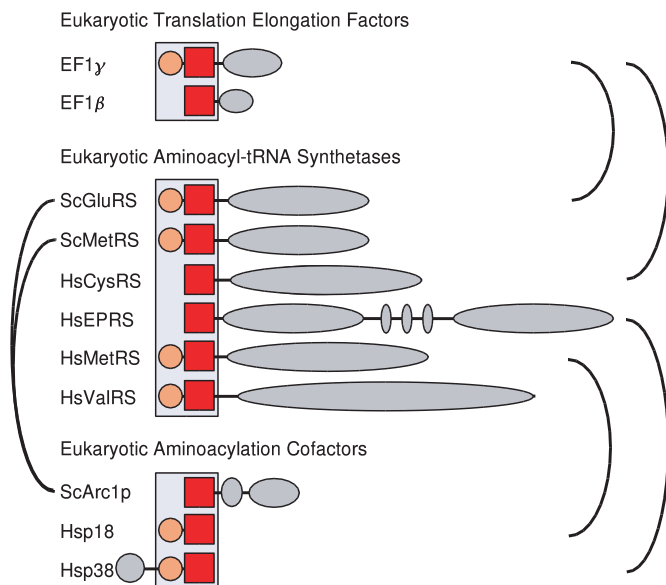


Figure 9. Domains with predicted GST fold are conserved among various components of the eukaryotic translation and aminoacylation machineries and are thought to mediate protein–protein interactions in several cases. GST_N (orange circles) and GST_C (red squares) domains of eukaryotic translation elongation factors, aaRS and aminoacylation cofactors are highlighted, other domains are shown in gray and known interactions are indicated by brackets. Shown are proteins from *S.cerevisiae* (Sc) and *Homo sapiens* (Hs).

and independently with both aaRS via two adjacent, but non-overlapping interfaces (Figure 8): a classical interface for the interaction with MetRS (Figures 3A and 4) and a novel interface for the interaction with GluRS (Figures 3C and 6). It thereby confirms that ternary complex formation is exclusively mediated by Arc1p without direct contacts between MetRS and GluRS (Figure 8), which were previously found not to interact in the absence of Arc1p (18). Based on a deletion analysis of the GluRS-N-terminal domain, it was shown that GluRS residues 9–85 are not required for interaction with Arc1p, but residues 85–131 are (18). Our structure of the complex between GluRS and Arc1p-N-termini shows that GluRS residues 122–131 and 157–173 are involved in the interaction with Arc1p, thus explaining the previous findings. In the same study, an N-terminal deletion mutant of Arc1p lacking amino acids 7–32 was found capable of interacting with GluRS while interaction with MetRS is abolished (18). Our GluRS-N–Arc1p-N complex shows that residues 7–32 of Arc1p are distant from the GluRS binding site (Figure 2 and residues up to the C-terminus of Helix 3 in Figure 3) and none of these residues is involved in the GluRS–Arc1p interface, explaining why deletion of these residues does not interfere with binding of Arc1p to GluRS. In our structure of the complex between MetRS and Arc1p-N-termini, these residues account for 56% of the interface surface (Figures 2 and 3B), explaining why deletion of these residues abolished interaction between MetRS and Arc1p. On the basis of a more detailed deletion mapping analysis of the Arc1p-N-terminus (18), it was previously concluded that Arc1p harbors distinct but overlapping binding sites for MetRS and GluRS. This conclusion was based on the finding that deletion in the C-terminus of Arc1p-N disrupted the interaction with both synthetases

while deletion of the N-terminus, as mentioned above, disrupted only the interaction with MetRS (18). Our structures suggest that this is not strictly true. Although the binding sites for MetRS and GluRS on Arc1p are immediately adjacent to each other, they are not overlapping, neither in terms of the quaternary structure of the ternary complex (Figure 8) nor in terms of the primary structure of Arc1p (Figure 2): while binding to MetRS is mediated exclusively by Arc1p residues 22–46, interaction with GluRS relies exclusively on residues 52–61 and 92–121. Our structures of the Arc1p-N-terminus suggest that some of the truncated Arc1p constructs used in the previous study may be prone to misfolding, possibly explaining the results of the deletion mapping analysis (18).

It has previously been shown that formation of the functional yeast aaRS complex from purified recombinant components occurs spontaneously under physiological conditions, demonstrating that no other cellular factors are required to form or stabilize this complex (9). For the Arc1p and GluRS-N-termini, the crystal structures of the isolated components [(21) and H. Simader, M. Hothorn and D. Suck, manuscript submitted] compared to the structure of the complex between the two demonstrate that both domains do not undergo any significant structural change upon complex formation (data not shown). Our structures thus suggest that both represent rigid modules, ready for a simple ‘plugging’ interaction where the rate and extent of complex formation is only determined by the diffusion rate and dissociation constant. This explains why the complex forms so rapidly and independently of other cellular factors and appears to suggest that complex assembly is probably not subject to regulation. On the other hand, the size of the interfaces in the yeast aaRS complex ($\sim 1600 \text{ \AA}^2$ buried surface area, respectively) is typical for optional protein complexes whereas the interfaces of permanent protein complexes tend to be much bigger ($>3400 \text{ \AA}^2$ on average) (44). This observation is in agreement with the finding that all three components can be stably and functionally maintained in a monomeric state both *in vivo* and *in vitro* and is compatible with a proposed role of complex assembly for regulating the sub-cellular distribution of the complex components (18).

DISCUSSION

In this work we present structural models for the interaction between the three protein components of the yeast aaRS complex. Our results and interpretations are based on the crystal structures of two binary complexes comprising the protein interacting domains of Arc1p and MetRS or Arc1p and GluRS, respectively. Previous research has clearly established that the N-terminal domains of Arc1p, GluRS and MetRS are necessary and sufficient for formation of the yeast aaRS complex both *in vitro* and *in vivo* (17,18). There is no evidence suggesting that other protein or tRNA domains apart from the N-terminal protein domains play a role in the formation of this complex. The importance of the protein–protein interactions mediated by the N-terminal domains of Arc1p, MetRS and GluRS for the biological activity of MetRS (9,16) and GluRS (9,17) as well as for the sub-cellular localization of MetRS and GluRS (18,20) is

very well established. Taken together, all available evidence suggests that our approach of studying the protein–protein interactions within the yeast aaRS complex using the isolated N-terminal domains yields biologically meaningful results. We have assembled and purified specifically forming, kinetically stable and structurally homogeneous binary complexes under physiological conditions in solution prior to crystallization, suggesting that our crystal structures represent biologically relevant interactions. Our finding that we can specifically abolish stable association of Arc1p with one enzyme while preserving specific interaction with the other in the context of the full-length proteins under physiological conditions in solution by either one of several single amino acid point mutations, rationally designed on the basis of our crystal structures (Figure 7) strongly suggests that the structural models of the two binary interactions presented here, which are based on the crystal structure of the interacting domains, truly reflect the interaction of the three full-length proteins in solution.

GST-like proteins have so far been structurally characterized in monomeric or homo-dimeric form. The finding that Arc1p-N, GluRS-N and MetRS-N adopt a GST-like fold is intriguing as they are known to engage exclusively in heteromeric interactions (18) and the structural data available so far did not explain how this fold could mediate heteromeric interactions. The results presented in this work explain how the N-terminal domain of Arc1p can interact simultaneously and independently with both MetRS and GluRS via one classical and one novel interface, respectively. Our finding that the basic architecture of the classical GST homo-dimer interaction has remained conserved in the yeast Arc1p–MetRS complex appears to suggest that this heteromeric interaction may have gradually evolved from a preexisting, originally homomeric interaction. The pseudo-homomeric nature of both interactions raises the question why the three complex components do not engage in homomeric interactions (9). Analysis of the residues involved in contacts (Figures 2 and 5) shows that most of them are not conserved between the interacting partners, generating interfaces complementary to each other, but sufficiently different to allow only heterodimeric contacts, consistent with the finding that neither domain forms homo-dimers in solution (9).

The complex structures presented here may represent ‘prototype’ models of interactions at the core of the mammalian multi-aaRS complex, which is known to contain five components orthologous to MetRS, GluRS and Arc1p: MetRS, the bi-functional GluProRS and the three accessory proteins p18, p38 and p43 (10,13). Four of them contain N-terminal GST-like folds and experimental evidence indicates direct interactions between MetRS and p18 as well as between GluProRS and p38 (7), suggesting that the interactions observed in the yeast aaRS complex may be conserved in the mammalian multi-aaRS complex.

tRNAs are known to be channeled in mammalian cells, i.e. directly transferred from the nuclear export machinery to aaRS and further via eEF-1A to the ribosome and back to the aaRS during many cycles of translation without diffusing freely into the cytoplasm (45–49). Yet our knowledge about the molecular mechanisms underlying the tRNA channeling cycle is still very limited. In this context, several examples of direct interactions between individual

components of the eukaryotic aminoacylation and the translational machinery have been reported in recent years (6,8,50) (Figure 9). However, they did not reveal a general concept on how both processes might be interconnected at the molecular level. Sequence analysis shows that domains with a predicted GST fold are present both in eukaryotic translation elongation factors and proteins of the aminoacylation machinery. In many cases, these domains have been reported to mediate physical interactions (6–8, 18) (Figure 9), suggesting that the modes of heteromerization shown here for the yeast aaRS complex may be conserved in other interactions networking the eukaryotic aminoacylation and translation machineries. The finding that MetRS-N adopts a GST-like fold although sequence homology to any GST-like protein is below the detection level (Figure 2) suggest that GST-like domains may be much more common among this class of proteins than predicted by mere sequence analysis and therefore their frequency and their importance may have been underestimated. The interaction modes observed in our complex structures suggests how this domain could function as a general protein–protein interaction module capable of multiple interactions, bridging and coordinating aminoacylation and translation in the context of the tRNA channeling cycle.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors would like to thank the synchrotron beam-line staff at ESRF ID14-4 for technical support during data collection. M.H. gratefully acknowledges financial support by the Peter and Traudl Engelhorn Foundation, Penzberg, Germany. This work is supported by the EU-grant: 3D repertoire, contract nr. LSHG-CT-2005-512028. Funding to pay the Open Access publication charges for this article was provided by EMBL.

Conflict of interest statement. None declared.

REFERENCES

- Mirande, M. (1991) Aminoacyl-tRNA synthetases and DNA replication. Molecular mimicry between RNAI and tRNA(Lys). *FEBS Lett.*, **283**, 1–3.
- Kisselev, L.L. and Wolfson, A.D. (1994) Aminoacyl-tRNA synthetases from higher eukaryotes. *Prog. Nucleic Acid Res. Mol. Biol.*, **48**, 83–142.
- Yang, D.C. (1996) Mammalian aminoacyl-tRNA synthetases. *Curr. Top Cell. Regul.*, **34**, 101–136.
- Martinis, S.A., Plateau, P., Cavarelli, J. and Florentz, C. (1999) Aminoacyl-tRNA synthetases: a family of expanding functions. Mittelwahr, France, October 10–15, 1999. *EMBO J.*, **18**, 4591–4596.
- Woese, C.R., Olsen, G.J., Ibba, M. and Söll, D. (2000) Aminoacyl-tRNA synthetases, the genetic code, and the evolutionary process. *Microbiol. Mol. Biol. Rev.*, **64**, 202–236.
- Negrutskii, B.S., Shalak, V.F., Kerjan, P., El'skaya, A.V. and Mirande, M. (1999) Functional interaction of mammalian valyl-tRNA synthetase with elongation factor EF-1alpha in the complex with EF-1H. *J. Biol. Chem.*, **274**, 4545–4550.
- Quevillon, S., Robinson, J.C., Berthonneau, E., Siatecka, M. and Mirande, M. Macromolecular assemblage of aminoacyl-tRNA

- synthetases: identification of protein–protein interactions and characterization of a core protein. *J. Mol. Biol.*, **285**, 183–195.
8. Kim, J.E., Kim, K.H., Lee, S.W., Seol, W., Shiba, K. and Kim, S. (2000) An elongation factor-associating domain is inserted into human cysteinyl-tRNA synthetase by alternative splicing. *Nucleic Acids Res.*, **28**, 2866–2872.
 9. Deinert, K., Fasiolo, F., Hurt, E.C. and Simos, G. (2001) Arc1p organizes the yeast aminoacyl-tRNA synthetase complex and stabilizes its interaction with the cognate tRNAs. *J. Biol. Chem.*, **276**, 6000–6008.
 10. Mirande, M., Kellermann, O. and Waller, J.P. (1982) Macromolecular complexes from sheep and rabbit containing seven aminoacyl-tRNA synthetases. II. Structural characterization of the polypeptide components and immunological identification of the methionyl-tRNA synthetase subunit. *J. Biol. Chem.*, **257**, 11049–11055.
 11. Kerjan, P., Cerini, C., Sémériva, M. and Mirande, M. (1994) The ubitienzyme complex containing nine aminoacyl-tRNA synthetases is ubiquitous from *Drosophila* to mammals. *Biochim. Biophys. Acta.*, **1199**, 293–297.
 12. Lee, S.W., Cho, B.H., Park, S.G. and Kim, S. (2004) Aminoacyl-tRNA synthetase complexes: beyond translation. *J. Cell. Sci.*, **117**, 3725–3734.
 13. Quevillon, S. and Mirande, M. (1996) The p18 component of the multisynthetase complex shares a protein motif with the beta and gamma subunits of eukaryotic elongation factor 1. *FEBS Lett.*, **395**, 63–67.
 14. Quevillon, S., Agou, F., Robinson, J.C. and Mirande, M. (1997) The p43 component of the mammalian multi-synthetase complex is likely to be the precursor of the endothelial monocyte-activating polypeptide II cytokine. *J. Biol. Chem.*, **272**, 32573–32579.
 15. Kim, J.Y., Kang, Y.-S., Lee, J.-W., Kim, H.J., Ahn, Y.H., Park, H., Ko, Y.-G. and Kim, S. (2002) p38 is essential for the assembly and stability of macromolecular tRNA synthetase complex: implications for its physiological significance. *Proc. Natl Acad. Sci. USA*, **99**, 7912–7916.
 16. Simos, G., Segref, A., Fasiolo, F., Hellmuth, K., Shevchenko, A., Mann, M. and Hurt, E.C. (1996) The yeast protein Arc1p binds to tRNA and functions as a cofactor for the methionyl- and glutamyl-tRNA synthetases. *EMBO J.*, **15**, 5437–5448.
 17. Simos, G., Sauer, A., Fasiolo, F. and Hurt, E.C. (1998) A conserved domain within Arc1p delivers tRNA to aminoacyl-tRNA synthetases. *Mol. Cell.*, **1**, 235–242.
 18. Galani, K., Grosshans, H., Deinert, K., Hurt, E.C. and Simos, G. (2001) The intracellular location of two aminoacyl-tRNA synthetases depends on complex formation with Arc1p. *EMBO J.*, **20**, 6889–6898.
 19. Simos, G., Grosshans, H. and Hurt, E. (2002) Nuclear export of tRNA. *Results Probl. Cell. Differ.*, **35**, 115–131.
 20. Galani, K., Hurt, E. and Simos, G. (2005) The tRNA aminoacylation co-factor Arc1p is excluded from the nucleus by an Xpo1p-dependent mechanism. *FEBS Lett.*, **579**, 969–975.
 21. Simader, H. and Suck, D. (2006) Expression, purification, crystallization and preliminary phasing of the heteromerization domain of the tRNA-export and aminoacylation cofactor Arc1p from yeast. *Acta Crystallogr. Sect. F. Struct. Biol. Cryst. Commun.*, **62**, 346–349.
 22. Kabsch, W. (1993) Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Crystallogr.*, **26**, 795–800.
 23. Vagin, A. and Teplyakov, A. (1997) Molrep: an automated program for molecular replacement. *J. Appl. Crystallogr.*, **30**, 1022–1025.
 24. McCoy, A.J., Storoni, L.C. and Read, R.J. (2004) Simple algorithm for a maximum-likelihood SAD function. *Acta Crystallogr. D. Biol. Crystallogr.*, **60**, 1220–1228.
 25. Collaborative Computational Project, Number 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D. Biol. Crystallogr.*, **50**, 760–763.
 26. Winn, M.D. (2003) An overview of the CCP4 project in protein crystallography: an example of a collaborative project. *J. Synchrotron Radiat.*, **10**, 23–25.
 27. Terwilliger, T.C. (2000) Maximum-likelihood density modification. *Acta Crystallogr. D. Biol. Crystallogr.*, **56**, 965–972.
 28. Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D. Biol. Crystallogr.*, **60**, 2126–2132.
 29. Murshudov, G.N., Vagin, A.A. and Dodson, E.J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D. Biol. Crystallogr.*, **53**, 240–255.
 30. Krissinel, E. and Henrick, K. (2005) Detection of protein assemblies in crystals. *CompLife 2005, LNBI*. Springer-Verlag, Berlin, Heidelberg, Vol. 3695, pp. 163–174.
 31. DeLano, W.L. The PyMOL Molecular Graphics System. DeLano Scientific, San Carlos, CA, USA.
 32. Lupyán, D., Leo-Macias, A. and Ortiz, A.R. (2005) A new progressive-iterative algorithm for multiple structure alignment. *Bioinformatics*, **21**, 3255–3263.
 33. Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, **22**, 4673–4680.
 34. Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T.J., Higgins, D.G. and Thompson, J.D. (2003) Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res.*, **31**, 3497–3500.
 35. Rousseau, F., Schymkowitz, J.W.H. and Itzhaki, L.S. (2003) The unfolding story of three-dimensional domain swapping. *Structure*, **11**, 243–251.
 36. Bennett, M.J. and Eisenberg, D. (2004) The evolving role of 3D domain swapping in proteins. *Structure*, **12**, 1339–1341.
 37. Yang, S., Cho, S.S., Levy, Y., Cheung, M.S., Levine, H., Wolynes, P.G. and Onuchic, J.N. (2004) Domain swapping is a consequence of minimal frustration. *Proc. Natl Acad. Sci. USA*, **101**, 13786–13791.
 38. Ding, F., Prutzman, K.C., Campbell, S.L. and Dokholyan, N.V. (2006) Topological determinants of protein domain swapping. *Structure*, **14**, 5–14.
 39. Holm, L. and Sander, C. (1993) Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.*, **233**, 123–138.
 40. Holm, L. and Sander, C. (1998) Touring protein fold space with Dali/FSSP. *Nucleic Acids Res.*, **26**, 316–319.
 41. Chothia, C. and Lesk, A.M. (1986) The relation between the divergence of sequence and structure in proteins. *EMBO J.*, **5**, 823–826.
 42. Bateman, A., Coin, L., Durbin, R., Finn, R.D., Hollich, V., Griffiths-Jones, S., Khanna, A., Marshall, M., Moxon, S., Sonnhammer, E.L.L. et al. (2004) The Pfam protein families database. *Nucleic Acids Res.*, **32**, D138–D141.
 43. Singh, J. and Thornton, J.M. (1992) *Atlas of Protein Side-Chain Interactions*. IRL press, Oxford, UK, Vol. 1, pp. 70–71.
 44. Jones, S. and Thornton, J.M. (1996) Principles of protein–protein interactions. *Proc. Natl Acad. Sci. USA*, **93**, 13–20.
 45. Negrutskii, B.S. and Deutscher, M.P. (1991) Channeling of aminoacyl-tRNA for protein synthesis *in vivo*. *Proc. Natl Acad. Sci. USA*, **88**, 4991–4995.
 46. Negrutskii, B.S. and Deutscher, M.P. (1992) A sequestered pool of aminoacyl-tRNA in mammalian cells. *Proc. Natl Acad. Sci. USA*, **89**, 3601–3604.
 47. Stapulionis, R. and Deutscher, M.P. (1995) A channeled tRNA cycle during mammalian protein synthesis. *Proc. Natl Acad. Sci. USA*, **92**, 7158–7161.
 48. Grosshans, H., Hurt, E. and Simos, G. (2000) An aminoacylation-dependent nuclear tRNA export pathway in yeast. *Genes Dev.*, **14**, 830–840.
 49. Grosshans, H., Simos, G. and Hurt, E. (2000) Review: transport of tRNA out of the nucleus—direct channeling to the ribosome? *J. Struct. Biol.*, **129**, 288–294.
 50. Gavin, A.-C., Bösch, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., Schultz, J., Rick, J.M., Michon, A.-M., Cruciat, C.-M. et al. (2002) Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature*, **415**, 141–147.