The presence of an isoform of glyceraldehyde-3-phosphate dehydrogenase (kmGAPDH1p) associated with the cell wall of a flocculent strain of Kluyveromyces marxianus was the first report of a non-cytosolic localization of a glycolytic enzyme, but the mechanism by which the protein is transported to the cell surface is not known. In order to identify structural features that could account for the multiple localizations of the protein, the three-dimensional structure of kmGAPDH1p was determined by X-ray crystallography and small angle X-ray scattering.

The X-ray crystallographic structure of kmGAPDH1p revealed a dimer, while all GAPDH homologs studied so far have a tetrameric structure with 222 symmetry. Interestingly, the structure of kmGAPDH1p in solution revealed a tetramer with a 70° tilt angle between the dimers. Moreover, the separation between the centres of the dimers composing the kmGAPDH1p tetramer diminishes from 34 Å to 30 Å upon NAD⁺ binding, this latter value being similar to the observed in the crystallographic models of GAPDH homologs. The less compact structure of apo-kmGAPDH1p could already be the first snapshot of the transition intermediate between the tetramer observed in solution and the dimeric form found in the crystal structure, which we postulate to exist in vivo due to the protein’s multiple subcellular localizations in this yeast species.

The reversible oxidation of D-glyceraldehyde 3-phosphate to 1,3-bisphosphoglycerate with the concomitant reduction of NAD⁺ to NADH is a key step in glycolysis and gluconeogenesis catalysed by the NAD⁺-dependent D-glyceraldehyde-3-phosphate dehydrogenase (GAPDH; EC 1.1.2.12). Multifunctional roles have been recently described for many proteins, and glycolytic enzymes such as glyceraldehyde-3-phosphate dehydrogenase and enolase exhibit other functions in addition to their catalytic role in glycolysis, both in prokaryotic and eukaryotic cells, including the yeast Saccharomyces cerevisiae (1; 2).

The three-dimensional structures of GAPDHs from a number of different organisms, ranging from Escherichia coli to Homo sapiens, have been determined by X-ray crystallography (3-18). The overall architecture of this enzyme is well known: four identical subunits (O, P, Q and R) are arranged with a molecular 222 symmetry, the two-fold axes being labelled P, Q and R (Figure 1). Each ~330 amino acid subunit is composed of two domains as shown on Figure 2. The NAD⁺-binding domain, comprising amino acids 1 to 148, displays the characteristic
Rossmann fold, with a central parallel β-sheet covered on both sides by α-helices. The catalytic domain (residues 149 to 329) folds into an eight-stranded antiparallel β-sheet and four α-helices. The catalytic domain (residues 149 to 329) folds into an eight-stranded antiparallel β-sheet and four α-helices.

Over the past ten years mounting evidence revealed that this protein has also non-glycolytic functions (1), such as a role in membrane fusion (19), microtubule–organanelle interactions (20), phosphotransferase activity (21), endoplasmic reticulum (ER)-to-Golgi vesicular transport (22;23) and regulation of transcription (24). These additional functions of GAPDH are often attributed to its tetrameric form. Nevertheless, there are reports that indicate that monomers and dimers are the forms involved in DNA (25) and RNA interaction (26). Given its multiple subcellular localizations (e.g. cell wall) the protein must exist, at least transiently, in non-tetrameric forms in order to be translocated across specific membranes (e.g. the ER membrane) (27).

Presence of glycolytic enzymes on the cell surface of microorganisms has been previously reported. Indeed enolase was found to be associated to glucan in the inner layers of the cell wall in *Candida albicans* (28), and to be secreted together with GAPDH when the *S. cerevisiae* cell walls are regenerating (29). In *C. albicans*, phosphoglycerate kinase, alcohol dehydrogenase and GAPDH have also been found associated to the cell wall and glycolytic enzymes appear to be important during its pathogenesis, since they act as major inducers of host immune responses (30). During *Streptococcus pyogenes* infection GAPDH acts as an adhesion molecule and helps escape detection by neutrophils (31). Alcohol dehydrogenase was also detected on the surface of *Entamoeba histolytica* (32) and phosphoglycerate kinase, triose-phosphate isomerase and GAPDH have been found on the surface of *Schistosoma mansoni* (33-35).

*Kluveromyces marxianus* encompasses a multigenic GAPDH family. Of the three genes coding for GAPDH-like polypeptides, only *GAP1* and *GAP2* are transcribed (36). GAPDH isoform 1 from *K. marxianus* (kmGAPDH1p), formerly referred to as "p37" (37), was first identified at the cell surface of flocculent *K. marxianus* cells, being the first report of a non-cytosolic localization of a glycolytic enzyme. Subsequently the same observation was described for *S. cerevisiae* (38). Although kmGAPDH1p is detected at the cell surface and is N-glycosylated (37;39), it does not contain the N-terminal signal peptide (36). Sequence comparison shows that this protein is 81.6 % identical to GAPDH isoform 1 from *S. cerevisiae*.

Although all these GAPDH isoforms are secreted, it has not been possible to identify a classic ER targeting sequence in any of them. It was reported that the GAPDH isoform 3 from *C. albicans* is able to direct incorporation of polypeptides into the *S. cerevisiae* cell wall, and that any motifs responsible for this targeting should be within the N-terminal half of the GAPDH amino acid sequence (40).

The molecular mechanisms underlying the multifunctionality of GAPDH, the regulation of its intracellular localization and oligomeric structure remain to be disclosed. Aiming at finding molecular features that could account for the multi-localization of the GAPDH isoform 1 from *K. marxianus*, we determined its three-dimensional structure. This is the first description of the three-dimensional structure of a GAPDH protein from a yeast species (kmGAPDH1p) obtained by both X-ray crystallography and small angle scattering experiments.

**EXPERIMENTAL PROCEDURES**

Protein expression and purification - kmGAPDH1 ORF was amplified by PCR and cloned into pQE32 (Quiagen®). Recombinant kmGAPDH1p was overexpressed in *E. coli* M15 and purified as follows. Cells were grown at 37 ºC to OD<sub>600 nm</sub> = 0.6 and induced with 0.1 mM IPTG for 10 h at 20 ºC. Cells were harvested by centrifugation (20 min, 11000 g) and the pellet stored at -70 ºC. All purification steps were performed at 4 ºC and all buffers were supplemented with a cocktail of protease inhibitors (MiniComplete®, Roche) according to the manufacturer’s instructions. After thawing the cell pellet was resuspended in buffer A (300 mM NaCl, 10 mM Imidazole, 1 mM β-ME, 1 mg/ml lysozyme, 10 mM Tris-HCl, pH 7.5). After 30 min the suspension was sonicated and incubated with 10 µg/ml RNase A and 5 µg/ml DNase I for 15 min. Cell debris was removed by centrifugation (30 min, 82500 g) and the supernatant was loaded onto a 10x75 mm Ni-NTA agarose (Qiagen®) column. A washing step was carried out with buffer B (1 M NaCl, 20 mM
Imidazole, 1 mM β-ME, 10 mM Tris-HCl, pH 7.5) until the OD\textsubscript{280 nm} value remained invariant (10 column volumes). Protein was eluted with buffer B + 250 mM Imidazole. kmGAPDH1p-containing fractions were pooled and dialyzed against buffer C (150 mM NaCl, 1 mM β-ME, 1 mM EDTA, 10 mM Tris-HCl, pH 7.5), concentrated to 13 mg/ml, frozen in liquid nitrogen and stored at -70 °C. The purity of the sample was verified by SDS–PAGE and western blot, followed by immunodetection with anti-His tag and anti-kmGAPDH1p sera (2).

**Crystallization** - Crystals of apo-kmGAPDH1p were grown using the sitting drop vapour diffusion method from 7 % PEG6000, 4 % MPD, 0.1 M Hepes pH 7.5 at 14 °C using 4 μl drops containing equal volume of the protein (13 mg/ml) and reservoir solutions. Crystals reached maximum size after one week and were then transferred to a solution containing the reservoir solution supplemented with 30 % glucose and cryo-cooled in liquid nitrogen for data collection.

**X-ray diffraction data collection and data reduction** - X-ray diffraction data were collected at 100 K from a single crystal on an ADSC Q4R CCD detector using synchrotron radiation at the ID14-EH1 beamline of the European Synchrotron Radiation Facility (ESRF, Grenoble). Data were evaluated and processed with MOSFLM (41), and reduced with SCALA and TRUNCATE (42). The kmGAPDH1p crystal used in the experiments diffracted to a maximum resolution of 2.3 Å and belonged to space group P4\textsubscript{1}2\textsubscript{1}2\textsubscript{1}. The unit-cell parameters were a = b = 70.9 Å and c = 358.4 Å. Assuming a molecular weight of 36.8 kDa and two monomers per asymmetric unit a Matthews coefficient of 3.06 Å\textsuperscript{3} Da\textsuperscript{-1} is obtained.

**Structure determination and refinement** - The structure was solved by molecular replacement using the program AMoRe (43) and the coordinates of *E. coli* GAPDH (PDB code 1GAD; 1) as search model. Calculations were performed for data in the 15-3.5 Å resolution range, and the rotation and translation searches provided the correct solution that was further improved with a final rigid body refinement, yielding an R-factor of 0.45 and a correlation coefficient of 46.5 for the best solution (R-factor of 0.52 and correlation coefficient of 26.3 for the second best solution). Positional and individual temperature factor refinement was carried out using CNS (44). Automated refinement was alternated with manual fitting of the model to the electron density on a SGI Graphic Workstation using the program Turbo-FRODO (45). Water molecules were added manually at the positions of positive peaks (> 3σ) on the difference Fourier maps, when good hydrogen bond geometry was possible. Structure validation was performed with the programs PROCHECK (46), WHAT_CHECK (47) and SF_CHECK (48). Data processing and refinement statistics are outlined in Table 1. Figures were produced with PyMOL (49) and ESPript (50). Multiple sequence alignment created with ClustalW (51).

**Size exclusion chromatography** - Size exclusion chromatography experiments were performed with a pre-packed Superose 12 10/300 GL column from Amersham Biosciences\textsuperscript{®}. The column was equilibrated with buffer C and a 0.5 ml/min flow rate was used throughout the experiments. Calibration was done with protein standards (MW, Stokes radius (R\textsc{s})): Blue Dextran 2000 (2 MDa), Catalase (232 kDa, 5.22 nm), Aldolase (158 kDa, 4.81 nm), BSA (67 kDa, 3.55 nm), Ovalbumin (43 kDa, 3.05 nm) and Chymotrypsinogen (25 kDa, 2.09 nm). Two hundred μl of purified protein (13 mg/ml) were used in each run. Protein elution was monitored by measuring the absorbance at 280 nm. The K\textsc{a} parameter was determined according to the equation K\textsc{a} = (V\textsubscript{e} - V\textsubscript{0})/(V\textsubscript{t} - V\textsubscript{0}), where V\textsubscript{e} represents the elution volume, V\textsubscript{0} the void volume of the column, and V\textsubscript{t} the total bed volume. Stokes radius (R\textsc{s}) for the experimental data was calculated using: (-logK\textsc{a})\textsuperscript{1/2} = f(R\textsc{s}) (52).

**Small angle X-ray scattering (SAXS) experiments and data analysis** - The synchrotron radiation small angle X-ray scattering data were collected on beamline X33 (53;54) of the EMBL-Hamburg on storage ring DORIS III of the Deutsches Elektronen Synchrotron (DESY), using a MAR 345 Image Plate detector. The scattering patterns from solutions of kmGAPDH1p at protein concentrations of 2.1, 4.2, 6.3 and 13 mg/ml were measured in buffer C with and without a pre-incubation with 1 mM NAD\textsuperscript{+}. At a sample-detector distance 2.7 m and wavelength λ = 1.5 Å the momentum transfer range 0.015 < s < 0.45 Å\textsuperscript{-1} was covered (s=4πsinθ/λ, where 2θ is the scattering angle). To monitor radiation damage, two successive two-minute exposures were compared and no significant change was observed. The data were normalized to the intensity of the
transmitted beam, radially averaged, and the scattering of the buffer was subtracted. The difference curves were scaled for protein concentration and extrapolated to yield the final composite scattering curves. The data processing steps were performed using the program package PRIMUS (55). The forward scattering \(I(0)\), radius of gyration \(R_g\) and the maximum particle dimension \(D_{\text{max}}\) were evaluated using the program GNOM (56). The molecular masses (MM) of the solutes were evaluated by comparison of the \(I(0)\) values with that from a reference solution of bovine serum albumin. Rigid body modeling of the quaternary structure of kmGAPDH1p in solution was performed using the crystallographic coordinates of dimeric kmGAPDH1p. The missing loop (AAs 180-199) was added as in the model of ecGAPDH (PDB code 1GAD). The scattering from the tentative tetrameric assembly of kmGAPDH1p as observed in the crystals of ecGAPDH was calculated using the program CRYSOl (57). This program evaluates the scattering from the protein in solution by computing the scattering from the atomic structure in vacuum, from the excluded volume and from the solvation shell. The values of the excluded volume and of the shell contrast are adjusted to fit the experimental data \(I_{\text{exp}}(s)\) by minimizing the discrepancy:

\[
\chi^2 = \frac{1}{N-1} \sum_j \left[ \frac{I_{\text{exp}}(s_j) - cI_{\text{calc}}(s_j)}{\sigma(s_j)} \right]^2
\]

where \(N\) is the number of experimental points, \(c\) is a scaling factor and \(I_{\text{calc}}(s_j)\) and \(\sigma(s_j)\) are the calculated intensity and the experimental error at the momentum transfer \(s_j\), respectively. The quaternary structure was further modeled by the exhaustive search program GLOBSYMM (58) in terms of two kmGAPDH1p dimers. The program screened the tetramers produced by rigid body movements and rotations of the crystallographic kmGAPDH1p dimer, whereas the second dimer was generated using a 2-fold symmetry axis (Z). The first dimer was oriented to have a two-fold symmetry along X axis and only movements along X and rotations around X were allowed such that the resulting tetramer kept the 222 symmetry. GLOBSYMM performed an exhaustive search of interconnected tetramers without steric clashes with a 2-degree angular step and 0.5 Å spatial increments. For each model, discrepancy (1) between the experimental scattering data and the curve calculated from the entire tetramer was calculated and the best-fit solution was selected. The experimental scattering curves recorded in the absence and presence of NAD\(^+\) were fitted independently to yield the rigid body models for the two cases.

RESULTS

Structure determination by X-ray crystallography

Crystals of kmGAPDH1p belong to the tetragonal space group P4\(_{2}121\), with cell dimensions \(a = b = 70.9\) Å and \(c = 358.4\) Å, two monomers per asymmetric unit and a solvent content of 59.8%. Diffraction data to 2.3 Å resolution were collected using synchrotron radiation. Data processing and refinement statistics are given in Table 1.

The final model has a crystallographic R value of 21.6 % and consists of a dimer with 4662 protein atoms, 230 water molecules, 2 β-mercaptoethanol molecules and 2 α-D-glucose molecules. All protein residues lie in the allowed regions of the Ramachandran plot.

Of the 329 amino acids that compose each kmGAPDH1p monomer, residues 180 to 199 are not well defined in the electron density maps and could not be modeled (Figure 1B). This stretch of residues is part of a highly conserved and well defined loop (S-loop) in most of the GAPDH structures described so far (Figure 2), but kmGAPDH1p is the only GAPDH of known structure where there is a glycine residue at position 201 of the polypeptide chain. The increase in flexibility originated by this substitution may account for the disorder of the S-loop in kmGAPDH1p.

The secondary and tertiary structures are very similar to those of other GAPDH structures. A comparison between the kmGAPDH1p and the E. coli GAPDH (ecGAPDH; PDB code 1GAD) monomers reveals great similarity between the two structures. The overall r.m.s deviation (r.m.s.d.) of the C\(_\alpha\) atomic positions after superimposing both structures is 1.16 Å for 298 aligned residues. Ser48 and Thr49, located in the loop between helix α\(_B\) and strand β\(_C\) (Figure 2), are the two residues that are most far apart in the two models, with C\(_\alpha\) r.m.s.d. of 4.0 and 4.3 Å, respectively.
The crystallographic packing of kmGAPDH1p can be described as an arrangement of OP dimers (Figure 3). Each unit cell is composed of eight dimers arranged in P4_{2,2,2} symmetry. Each dimer has dimensions of approximately 98 Å x 57 Å x 48 Å. In contrast to the other GAPDH structures described so far there is no association of neighboring dimers to form a tetramer. The intersubunit interactions across the R interface of the tetrameric GAPDHs involve amino acids that are conserved in most of the S-loops, including the one from kmGAPDH1p (Figure 2B). The only exception which is visible in the electron density maps is glycine 201. Interestingly, in the N-terminal end of the kmGAPDH1p S-loop, serine 177 makes a novel hydrogen bond with the carbonyl oxygen of threonine 235, which is not observed for other GAPDHs with a serine residue at position 177. The side chain of the following residue, isoleucine 178, rotates 180 degrees relatively to the equivalent residue in tetrameric GAPDHs, thus occupying the position of the side chain of histidine 176, which must move away thereby disrupting the interaction with the other catalytic residue, cysteine 149. The association of these two amino acids is essential for the catalysis, namely for the formation of an ion pair in the acylation step of the chemical mechanism. The dimeric GAPDH from K. marxianus must thus be enzymatically inactive. The absence of NAD⁺ can contribute to the observed disorder of the S-loop in the kmGAPDH1p model, since in NAD⁺-containing structures (3; 4) there is a hydrogen bond network involving NAD⁺, ordered water molecules and the S-loop from the R-related subunit.

Size exclusion chromatography - In order to assess the stability of the quaternary structure of kmGAPDH1p, the protein was subjected to size-exclusion chromatography. The experimental equivalent hydrodynamic radius (Rₐ) determined for kmGAPDH1p by this technique is 43.5 Å, significantly higher than the 35.6 Å calculated using the crystallographic coordinates of the dimer and the program HYDROPRO (59).

The quaternary structure of kmGAPDH1p in solution - To determine the oligomeric arrangement of kmGAPDH1p in solution, the protein preparations were examined by synchrotron SAXS. The scattering patterns from purified kmGAPDH1p without and with pre-incubation with 1 mM NAD⁺ are presented in Figure 4. The MM of the solutes determined from the extrapolated forward scattering was 130±15 kDa indicating that the protein is tetrameric in both cases. The experimental radius of gyration and the maximum size of the free enzyme are 37±1 Å and 110±10 Å. After incubation with NAD⁺ these values are noticeably smaller (34±1 Å and 100±10 Å, respectively), suggesting that NAD⁺ binding leads to a compaction of the structure.

It was natural to assume that the quaternary structure of tetrameric kmGAPDH1p in solution should be similar to the symmetric tetramers observed in the crystal structures of other GAPDHs. To verify this, a tentative kmGAPDH1p tetramer was constructed with the arrangement of the dimers as in the ecGAPDH structure (PDB code 1GAD). The radius of gyration computed from this model (Rᵥ=32.7 Å) was too small in comparison with the SAXS-derived value for the free enzyme, and the computed theoretical scattering pattern failed to fit the experimental data from free kmGAPDH1p (Figure 4, top curve, discrepancy χ²=3.7). The Rᵥ of the tentative model was more compatible with the experimental value of NAD⁺ incubated kmGAPDH1p, but the theoretical pattern did not fit the experimental scattering for this sample either (Figure 4, bottom curve, discrepancy χ²=2.7). The scattering curves computed from the crystallographic tetramer display a pronounced minimum at s=0.1 Å⁻¹ absent in the experimental data, indicating that the model is too compact and isometric compared to the tetramer in solution.

The quaternary structure was further modeled by the exhaustive search program GLOBSYMM (58) in terms of two kmGAPDH1p dimers keeping the 222 symmetry of the tetramer, as described in Materials and Methods. The quaternary structures of apo kmGAPDH1p and of the enzyme incubated with NAD⁺ were modeled independently using the corresponding scattering patterns. The resulting models, displayed in Figure 5, have radii of gyration of 34.6 Å (apo enzyme) and 33.5 Å (NAD⁺-bound enzyme). These values are still somewhat lower than the experimental ones, which may in part be attributed to the influence of the N-terminal His-tags not present in the high resolution model. Overall, the two models...
provide good fits to the experimental data displayed in Figure 4 (with discrepancy $\chi^2=1.56$ and 1.39 with and without NAD$^+$, respectively). The tilt angle between the dimers in apo kmGAPDH1p is 70° and the separation between their centers is 34 Å (instead of the angle of 90° and the separation of 29.9 Å for the ecGAPDH model). Remarkably, the independently-modeled structure for the enzyme incubated with NAD$^+$ yields exactly the same tilt angle between the dimers and only differs from the apo kmGAPDH1p model by a shorter distance between the dimers (separation between the centers of 30 Å). In both cases, with and without NAD$^+$, the solutions provided by GLOBSYMM were practically unique, i.e. only one major minimum in the discrepancy was present among the interconnected tetramers without steric clashes showing a 222 symmetry. Computations without restricting the modeling to the 222 symmetry and also without addition of the missing loop (AAs 180-199) yielded very similar results.

**DISCUSSION**

The unexpected presence of a GAPDH isoform at the cell wall of *K. marxianus* led us to investigate the three-dimensional structure of the protein.

Although kmGAPDH1p is detected at the cell surface of flocculent *K. marxianus* cells, it does not contain an N-terminal ER targeting peptide (36). The evidence accumulated during the past decade, showing that several proteins lacking an N-terminal signal peptide reach the yeast cell wall, has led to the suggestion that an alternative ER-independent mechanism is responsible for the secretion of these proteins (60). However, this seems not to be the case for the secretion of kmGAPDH1p since the protein purified from *K. marxianus* cell walls comprises an N-linked carbohydrate chain of 2 kDa (36), which constitutes strong evidence for the passage of the protein through the ER. Moreover, the analysis of the secretion of engineered kmGAPDH1p polypeptides expressed in *S. cerevisiae* suggests that it takes place via the ER. In these experiments it was observed that the addition of an ER retention motif to kmGAPDH1p blocks the secretion of the protein (data not shown). It was previously shown, using *GAP1 K. marxianus* mutant cells that kmGAPDH1p contributes to the GAPDH cytosolic activity (61). This suggests that the protein is translocated post-translationally across the ER membrane and for this the kmGAPDH1p tetramer has to dissociate into monomers before being unfolded to cross to the ER lumen. The tendency of kmGAPDH1p to dissociate into dimers in the crystallization experiments suggests that the disassembly of the tetramer occurs *via* the formation of OP dimers.

Yeast cell walls have a net negative charge, mainly due to the presence of carboxylic and phosphate groups. The presence of an electrostatic potential at the surface of the cells accounts for cell dispersion, which is maintained unless specific interactions are established (62). Proteins involved in yeast flocculation might facilitate the adhesion of cells by neutralizing the overall negative charge that draws cells apart in the culture medium. However, inspection of the surface electrostatic potential of kmGAPDH1p reveals no significant positive surface potential that could account for the electrostatic neutralization of the yeast cell wall. Therefore, we conclude that the role of kmGAPDH1p in flocculation is most probably a diverse and more specific one.

All GAPDH structures described so far are tetrameric while our results from X-ray crystallography revealed a dimeric structure. Therefore it was imperative to determine the quaternary structure of the protein in solution. The SAXS experiments revealed that the protein is tetrameric in solution, but the arrangement of the dimers is different from the one found in the crystallographic structures of other GAPDH proteins. In solution, there is a rotation around the molecular P axis of one dimer in relation to the other, which does not disrupt the molecular 222 symmetry but makes the tetramer essentially less isometric (Figure 5a). Moreover, the binding of NAD$^+$ to the protein leads to a more compact structure by a 4 Å shift of the dimers towards each other as revealed in Figure 5b.

Interestingly, the $R_s$ calculated with HYDROPRO from the SAXS-derived tetrameric apo form is 44.0 Å, in good agreement with the experimental value of 43.5 Å obtained by size exclusion chromatography.

The SAXS-derived model of the apo-kmGAPDH1p tetramer shows a substantial
deviation from the isometric tetramers observed in the crystallographic structures of other GAPDHs. Given that kmGAPDH1p does not display these tetramers in the crystal either, this result may reflect differences between species. Such differences were observed e.g. between the quaternary structures of tetrameric pyruvate decarboxylase (EC 4.1.1.1) from Zymomonas mobilis and brewers' yeast. The former species displayed a compact isometric arrangement of dimers both in the crystal and in solution. The latter species revealed a rather loose tetramer in the crystal, and rigid body modeling against SAXS data yielded a similar but somewhat more compact dimer arrangement, with r.m.s.d. of 5.8 Å to the loose crystal structure.

We compared the Q and R interfaces of the crystallographic ecGAPDH model with the equivalent regions of kmGAPDH1p. Although the three-dimensional structure of the polypeptide chain at the ecGAPDH Q interface is similar to the homologous sections of kmGAPDH1p, the same is not true for the R interface, which is the largest one. The R interface of ecGAPDH involves residues 39 to 49 of one subunit and the S-loop residues 179 to 202 of the R-related subunit. All of the 24 hydrogen bonds in this interface involve residues from the S-loop, which could not be modeled for the kmGAPDH1p structure, thus impairing direct comparison of these interactions. The high degree of similarity at sequence level in this region (Figure 2B) probably indicates that the observed disorder is not the cause but rather a consequence of the absence of the R-related dimer (Figure 1).

The only residue of the S-loop distinctly different in kmGAPDH1p is glycine 201. The increased flexibility introduced by this residue probably explains the disorder of most of the loop, as well as allows the movements necessary for the formation of the additional interactions in region 177-178, with the concomitant disruption of the catalytic site, and most likely also allows for the disruption of the interactions with residues 39-49 of the R-related subunit.

It is noteworthy that Ser48 and Thr49, two residues in the loop between helix $\alpha_B$ and strand $\beta_C$ involved in more than half of the hydrogen bonds across the R interface, are so distant in the two structures (C$_\alpha$ r.m.s.d higher than 4 Å) that these interactions would not be possible in an isometric kmGAPDH1p tetramer. Interestingly, Ser48 was described as important for tetramer stability in the GAPDH from Bacillus stearothermophilus. In the same study it was shown that the Y46G/S48G dimeric mutant is of the OP type.

Our results reveal a GAPDH tetrameric structure in solution which is different from that observed in the crystallographic structures of other GAPDHs. While keeping the 222 symmetry, the kmGAPDH1p tetramer is less isometric than the crystallographic tetramers of other GAPDHs, revealing an unstable structure that dissociates into OP dimers in the experimental conditions used to obtain the protein crystals. We believe that the higher propensity of kmGAPDH1p to dissociation portrayed in this study may favor the protein translocation across the ER membrane, although additional studies are needed to determine how this occurs in vivo.

REFERENCES


**FOOTNOTES**

* F. F.-S. would like to thank FCT, Portugal, for financial support through fellowship Praxis XXI / BD / 19777 / 99.

The atomic coordinates and structure factors (PDB code 2I5P) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

1The abbreviations used are: GAPDH, glyceraldehyde-3-phosphate dehydrogenase; kmGAPDH1p, GAPDH isoform 1 from *K. marxianus*; r.m.s.d., root mean square deviation; AAs, amino acids; ER, endoplasmic reticulum; Rs, Stokes radius; MM, molecular mass. Rg, radius of gyration

9
ACKNOWLEDGMENTS

We acknowledge the European Synchrotron Radiation Facility for provision of synchrotron radiation facilities and we would like to thank the ESRF staff for assistance in using beamline ID14-EH1. We also thank EMBL-Hamburg for the use of synchrotron radiation facilities on beamline X33.

FIGURE LEGENDS

Fig. 1. Comparison between the *E. coli* GAPDH tetramer and the *K. marxianus* GAPDH1p dimer. The tetrameric ecGAPDH (A) as viewed along the molecular R axis and the kmGAPDH1p dimer (B), as revealed by their crystallographic structures. The S-loop is highlighted in black. The S loops represented with dashed lines in (B) are from the superposition of the monomer from ecGAPDH and are absent from the model of kmGAPDH1p.

Fig. 2. (A) kmGAPDH1p monomer. The β strands A to F and the α helices A to D are part of the NAD-binding domain. The catalytic domain is composed of β strands 1 to 8 and the α helices 1 to 3. The S loop (orange) and the NAD⁺ molecule represented are from the superposed monomer of ecGAPDH and are absent from the kmGAPDH1p model. (B) kmGAPDH1p sequence is shown aligned against the sequence of GAPDHs from different organisms (common name, PDB code): *Escherichia coli* (1GAD); *Palinurus versicolor* (Lobster, 1CRW); *Homo sapiens* (3GPD); *Bacillus stearothermophilus* (1NQO) and *Oryctolagus cuniculus* (Rabbit, 1J0X). The elements of secondary structure of kmGAPDH1p are illustrated. The positions with residues involved in hydrogen bonds across the R interface of the GAPDH from *E. coli* are indicated by black triangles.

Fig. 3. Crystal packing of kmGAPDH1p. The eight GAPDH dimers present in the unit cell are shown as Cα traces. The unit cell is shown as viewed along the b axis (A) and the a axis (B).

Fig. 4. Experimental and calculated scattering patterns of kmGAPDH1p. (1): experimental data (dots with error bars); (2) and (3) scattering from the isometric crystallographic tetramer as observed for ecGADPH and from the rigid body models, respectively. Upper curve, Apo form of kmGAPDH1p; lower curves, NAD-bound kmGAPDH1p (displaced by one logarithmic unit down for clarity).

Fig. 5. Quaternary structure of kmGAPDH1p in solution. (A) Superposition between the apo-kmGAPDH1p tetramer (black and blue dimers) obtained by SAXS and the ecGAPDH tetramer (black and red dimers). (B) Superposition between the SAXS-derived tetramer of NAD-kmGAPDH1p (black and red dimers) and the apo-kmGAPDH1p tetramer (black and blue dimers).
### Table 1

**Data collection and structure refinement statistics**

<table>
<thead>
<tr>
<th>Data Collection</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Resolution range (Å)</strong></td>
<td>65.9-2.30</td>
</tr>
<tr>
<td><strong>Space group</strong></td>
<td>P4₁2₁2</td>
</tr>
<tr>
<td><strong>Unit-cell parameters (Å)</strong></td>
<td>a = 70.9, b = 70.9, c = 358.4</td>
</tr>
<tr>
<td><strong>Nº of monomers per asymmetric unit</strong></td>
<td>2</td>
</tr>
<tr>
<td><strong>Reflections (measured / unique)</strong></td>
<td>339229 / 42054</td>
</tr>
<tr>
<td>**Multiplicity **</td>
<td>8.1(8.2)</td>
</tr>
<tr>
<td>**Completeness (%) **</td>
<td>99.9(100.0)</td>
</tr>
<tr>
<td><strong>Rmerge a</strong></td>
<td>0.073 (0.380)</td>
</tr>
<tr>
<td><strong>I / σ(I)</strong></td>
<td>6.3 (2.0)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Structure Refinement</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Resolution range (Å)</strong></td>
<td>65.9-2.30</td>
</tr>
<tr>
<td><strong>Rcryst b</strong></td>
<td>0.216</td>
</tr>
<tr>
<td><strong>Rfree c</strong></td>
<td>0.245</td>
</tr>
<tr>
<td><strong>r.m.s.d., bond lengths (Å)</strong></td>
<td>0.007</td>
</tr>
<tr>
<td><strong>r.m.s.d., bond angles (º)</strong></td>
<td>1.3</td>
</tr>
<tr>
<td><strong>Nº of protein atoms</strong></td>
<td>4662</td>
</tr>
<tr>
<td><strong>Nº of solvent atoms</strong></td>
<td>230</td>
</tr>
<tr>
<td><strong>Average B factors (Å²)</strong></td>
<td></td>
</tr>
<tr>
<td>Main-chain atoms</td>
<td>45.25</td>
</tr>
<tr>
<td>Side-chain atoms</td>
<td>47.09</td>
</tr>
<tr>
<td>α-D-Glucose</td>
<td>70.95</td>
</tr>
<tr>
<td>β-Mercaptoethanol</td>
<td>68.29</td>
</tr>
<tr>
<td>Water molecules</td>
<td>51.60</td>
</tr>
<tr>
<td><strong>r.m.s.d. of bonded Bs for main chain residues.</strong></td>
<td>0.95</td>
</tr>
<tr>
<td><strong>r.m.s.d. of bonded Bs for side chain residues</strong></td>
<td>1.54</td>
</tr>
<tr>
<td><strong>Ramachandran plot statistics</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Nº of non-glycine and non-proline residues</strong></td>
<td>528</td>
</tr>
<tr>
<td><strong>residues in most favoured regions</strong></td>
<td>467 (88.4 %)</td>
</tr>
<tr>
<td><strong>residues in additional allowed regions</strong></td>
<td>60 (11.4 %)</td>
</tr>
<tr>
<td><strong>residues in generously allowed regions</strong></td>
<td>1 (0.2 %)</td>
</tr>
<tr>
<td><strong>residues in disallowed regions</strong></td>
<td>0 (0 %)</td>
</tr>
</tbody>
</table>

*a Values in parentheses refer to the highest-resolution shell

*a_Rmerge = \sum |I(h)_i - \langle I(h) \rangle| / \sum I(h)_i\,\text{, where } I(h)_i\text{ is the observed intensity of the }i\text{th source and }\langle I(h) \rangle\text{ is the mean intensity of reflection }h\text{ over all measurements of }I(h)\text{.}

*b_Rcryst = \sum \mid F_{\text{obs}} \mid - \mid F_{\text{calc}} \mid / \sum \mid F_{\text{obs}} \mid , \text{ where } \mid F_{\text{obs}} \mid \text{ and } \mid F_{\text{calc}} \mid \text{ are observed and calculated structure factor amplitudes, respectively.}

*c_Rfree is equivalent to Rcryst except that it concerns 5 % of the total reflections which were omitted in the refinement process.
Figure 3
The crystal and solution structures of glyceraldehyde-3-phosphate dehydrogenase reveal different quaternary structures
Frederico Ferreira-da-Silva, Pedro J.B. Pereira, Luís Gales, Manfred Roessle, Dmitri I. Svergun, Pedro Moradas Ferreira and Ana M. Damas

J. Biol. Chem. published online September 7, 2006

Access the most updated version of this article at doi: 10.1074/jbc.M605267200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 0 references, 0 of which can be accessed free at
http://www.jbc.org/content/early/2006/09/07/jbc.M605267200.citation.full.html#ref-list-1