BIOLOGICAL CRYSTALLOGRAPHY
ISSN 1399-0047

Unravelling the shape and structural assembly of the photosynthetic GAPDH–CP12–PRK complex from *Arabidopsis thaliana* by small-angle X-ray scattering analysis

Alessandra Del Giudice,[a] Nicolae Viorel Pavel,[a]‡ Luciano Galantini,[a] Giuseppe Falini,[b] Paolo Trost,[c] Simona Fermani[b]* and Francesca Sparla[c]*

[a]Department of Chemistry, University of Rome 'Sapienza', Rome, Italy, [b]Department of Chemistry 'G. Ciamician', University of Bologna, Bologna, Italy, and [c]Department of Pharmacy and Biotechnology – FaBiT, University of Bologna, Bologna, Italy. *Correspondence e-mail: simona.fermani@unibo.it, francesca.sparla@unibo.it

Received 11 March 2015
Accepted 3 October 2015

Edited by R. C. Garratt, University of Sao Paulo, Brazil

‡ Co-senior author.

Keywords: Calvin–Benson cycle; SAXS; *Arabidopsis thaliana*; multiprotein complex.

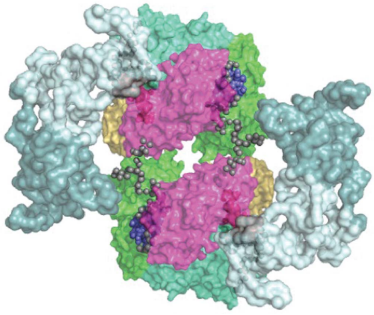
Supporting information: this article has supporting information at journals.iucr.org/d

Oxygenic photosynthetic organisms produce sugars through the Calvin–Benson cycle, a metabolism that is tightly linked to the light reactions of photosynthesis and is regulated by different mechanisms, including the formation of protein complexes. Two enzymes of the cycle, glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and phosphoribulokinase (PRK), form a supramolecular complex with the regulatory protein CP12 with the formula $(GAPDH–CP12_2–PRK)_2$, in which both enzyme activities are transiently inhibited during the night. Small-angle X-ray scattering analysis performed on both the GAPDH–CP12–PRK complex and its components, GAPDH–CP12 and PRK, from *Arabidopsis thaliana* showed that (i) PRK has an elongated, bent and screwed shape, (ii) the oxidized N-terminal region of CP12 that is not embedded in the GAPDH–CP12 complex prefers a compact conformation and (iii) the interaction of PRK with the N-terminal region of CP12 favours the approach of two GAPDH tetramers. The interaction between the GAPDH tetramers may contribute to the overall stabilization of the GAPDH–CP12–PRK complex, the structure of which is presented here for the first time.

1. Introduction

Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and phosphoribulokinase (PRK) are two enzymes involved in the photosynthetic carbon fixation of oxygenic phototrophs (land plants, algae and cyanobacteria), a metabolism also known as the Calvin–Benson cycle. GAPDH is responsible for the single reducing step of the cycle, while PRK allows the regeneration of ribulose-1,5-bisphosphate, the acceptor of CO_2. Together, GAPDH and PRK consume most of the NADPH and ATP produced by the light phase of photosynthesis to fuel the Calvin–Benson cycle.

Besides converting light into ATP and NADPH, the photosynthetic electron-transport chain of oxygen phototrophs keeps reduced a pool of thioredoxins that regulate several enzymes of the Calvin–Benson cycle through dithiol/disulfide exchange reactions, including GAPDH and PRK, and other metabolisms (Michelet *et al.*, 2013). By connecting the photosynthetic electron flow to the Calvin–Benson cycle, dithiol/disulfide redox cascades are believed to represent an important adaptation of oxygen phototrophs to fluctuating light conditions (Balsera *et al.*, 2014). In addition, other monothiol-based redox modifications (*e.g.* glutathionylation and nitrosylation) that are induced by reactive oxygen and nitrogen species may also contribute to the dynamic



regulation of photosynthetic metabolism under stress (Michelet *et al.*, 2013; Zaffagnini *et al.*, 2012).

Even though GAPDH and PRK catalyze two non-consecutive reactions of the Calvin–Benson cycle, they can form a supramolecular complex in which two GAPDH tetramers and two PRK dimers interact together by means of four CP12 molecules (Marri *et al.*, 2008). CP12 is an intrinsically disordered protein (IDP) of about 80 amino acids, including four cysteines that can form two disulfide bridges under thioredoxin control (Graciet *et al.*, 2003; Marri *et al.*, 2008). When both disulfide bridges are formed, CP12 may assemble the ternary complex with GAPDH and PRK, with both enzymes consequently being inactivated (Wedel & Soll, 1998; Marri *et al.*, 2005). Through modulation of its redox state, CP12 controls the assembly and disassembly of the ternary complex and consequently the regulation of the GAPDH and PRK activities (Marri *et al.*, 2009). CP12 is universally distributed in oxygenic phototrophs from cyanobacteria to higher plants (Marri *et al.*, 2010; Stanley *et al.*, 2013; Gontero & Maberly, 2012; López-Calcagno *et al.*, 2014), where CP12-assembled complexes tend to accumulate in the dark (Scheibe *et al.*, 2002; Tamoi *et al.*, 2005; Howard *et al.*, 2008; Howard, Lloyd *et al.*, 2011). In tobacco, CP12 was shown to be essential for normal growth and development (Howard, Fryer *et al.*, 2011), and in the cyanobacterium *Synecococcus* PCC7942 genetic disruption of CP12 gene caused delayed growth in a normal light/dark cycle (Tamoi *et al.*, 2005).

Beside a homotetrameric GAPDH isoform (A_4-GAPDH), higher plants possess a second isoform made up of A and B subunits (AB-GAPDH) (Scagliarini *et al.*, 1998; Scheibe *et al.*, 2002; Trost *et al.*, 2006; Howard, Lloyd *et al.*, 2011). The main difference between the A and B subunits is the presence in subunit B of a C-terminal extension (CTE) of 30 amino acids that is homologous to the C-terminal end of CP12 (Baalmann *et al.*, 1996; Sparla *et al.*, 2002; Trost *et al.*, 2006; Thieulin-Pardo *et al.*, 2015). Like CP12, the CTE is unstructured and contains a pair of redox-regulated cysteines (Fermani *et al.*, 2007). The CTE is essential for the autonomous (CP12-independent) regulation of the AB-GAPDH isoform, including the capability of A_2B_2-GAPDH to associate into inactive oligomers (A_8B_8) in the dark (Baalmann *et al.*, 1996; Sparla *et al.*, 2005).

Oxygen phototrophs contain a unique form of dimeric PRK made of 40 kDa subunits. This form is only distantly related to the octameric PRK with 32 kDa subunits from aerobic photosynthetic bacteria such as *Rhodobacter sphaeroides*, the crystal structure of which has been solved (Harrison *et al.*, 1998). PRK of green algae and land plants is directly regulated by thioredoxins affecting the redox state of two cysteines located in the active site of the enzyme (Brandes *et al.*, 1996), and it may be further regulated through the formation of the ternary complex with GAPDH and CP12 (Wedel & Soll, 1998; Scheibe *et al.*, 2002; Graciet *et al.*, 2003; Marri *et al.*, 2005; Tamoi *et al.*, 2005; Howard *et al.*, 2008; Howard, Lloyd *et al.*, 2011).

In *Arabidopsis thaliana*, the pathway of complex assembly has been elucidated. Firstly, oxidized CP12 and A_4-GAPDH interact together, forming a binary complex (Graciet *et al.*, 2003). This interaction is only possible if GAPDH binds NAD(H) in its coenzyme-binding sites (Marri *et al.*, 2005; Matsumura *et al.*, 2011). Although CP12 populates different conformations in solution, only a subset of CP12 conformers may initially bind to GAPDH and successively fold their C-terminal regions within the active site of the enzyme (Fermani *et al.*, 2012; Reichmann & Jakob, 2013; Uversky, 2013). After the formation of the binary complex, in which two CP12 molecules bind to one GAPDH tetramer, oxidized PRK promotes the assembly of the ternary complex in which two binary complexes are held together by two PRK dimers (Graciet *et al.*, 2003; Marri *et al.*, 2005, 2008).

In the last decade, small-angle X-ray scattering (SAXS) has become a powerful technique to structurally characterize biological macromolecules in solution (Koch *et al.*, 2003; Lipfert & Doniach, 2007; Rambo & Tainer, 2013*a*). In this manuscript, we exploited the potential of SAXS to describe a low-resolution model of the GAPDH–CP12–PRK complex from *A. thaliana* and of its components the GAPDH–CP12 complex and PRK alone. We demonstrate that PRK has an elongated shape that is suitable to connect two GAPDH–CP12 complexes and form a compact ternary complex in which the GAPDH tetramers are much closer than previously suspected, suggesting a role of GAPDH–GAPDH interactions in ternary-complex stabilization. Besides depicting the shape and organization of the GAPDH–CP12–PRK complex for the first time, the results confirm the renowned capability of photosynthetic GAPDH tetramers to associate into supramolecular complexes with regulatory functions (GAPDH–CP12–PRK; A_8B_8-GAPDH).

2. Materials and methods

2.1. Protein expression and purification

Recombinant and mature forms of GAPDH (EC 1.2.1.13), PRK (EC 2.7.1.19) and CP12 (isoform 2) of *A. thaliana* were expressed and purified as described in Marri *et al.* (2008). Different aliquots of pure PRK were treated with 5 m*M* reduced or oxidized DTT to ensure the presence of a unique redox form. After overnight incubation at 4°C, samples were desalted in 25 m*M* potassium phosphate buffer solution pH 7.5 and concentrated to 10 mg ml^{-1} using a Centricon (Millipore) device with a 10 kDa cutoff.

The GAPDH–CP12 complex was obtained after overnight incubation at 4°C of pure GAPDH and CP12 in a 1:1 ratio (on a subunit basis) in the presence of 5 m*M* NAD and 5 m*M* oxidized DTT. To remove excess CP12, the sample was concentrated to 200 µl using a Centricon (Millipore) device with a 3 kDa cutoff and loaded onto a gel-filtration column (Superdex 200 10/300 GL; GE Healthcare) pre-equilibrated with 25 m*M* potassium phosphate buffer solution pH 7.5. The eluted GAPDH–CP12 binary complex was collected, 1 m*M* NAD was immediately added and it was brought to a final concentration of 10 mg ml^{-1}.

The GAPDH–CP12–PRK complex was obtained after overnight incubation at 4°C of the purified protein

Table 1
SAXS data-acquisition parameters and analysis.

n.p., not performed.

	PRK			
	Oxidized	Reduced	GAPDH–CP12	GAPDH–CP12–PRK
Data-collection parameters				
Detector	PILATUS 1M (67 × 420 mm)			
Beam geometry	0.7 × 0.7 mm			
Wavelength (Å)	0.99			
Sample-to-detector distance (mm)	2430			
Capillary diameter (mm)	1.8			
q range† (Å^{-1})	0.005–0.45			
Exposure time (s)	10 × 1			
Temperature (K)	278.15			
Concentration range (mg ml^{-1})	2.0–5.5	1.1–3.9	0.5–5.0	1.0–5.0
Structural parameters‡				
$I(0)$ BSA	66.5	66.8	66.8	68.2
q interval for Fourier inversion (Å^{-1})	0.029–0.35	0.03–0.35	0.01–0.35	0.008–0.35
$I(0)$ [from $p(r)$]	74.9 ± 0.5	81.4 ± 0.5	146.0 ± 0.1	281.6 ± 0.1
R_g [from $p(r)$] (Å)	38.7 ± 0.1	40.1 ± 0.5	34.0 ± 0.1	57.2 ± 0.1
q interval for Guinier linear fit (Å^{-1})	0.020–0.033	0.025–0.032	0.013–0.037	0.010–0.023
$I(0)$ (from Guinier approximation)	77.2 ± 0.5	84.5 ± 0.5	146.0 ± 0.1	280.0 ± 0.5
R_g (from Guinier approximation) (Å)	39.4 ± 0.5	40.9 ± 0.5	34.0 ± 0.1	56.7 ± 0.1
D_{max} (Å)	126 ± 5	130 ± 5	115 ± 2	175 ± 5
Porod volume estimate (nm^3)	162 ± 5	165 ± 5	207 ± 10	732 ± 10
DAMMIN excluded volume (nm^3)	199 ± 2	201 ± 5	240 ± 2	675 ± 2
Dry volume calculated from sequence (nm^3) (v = 0.733 cm^3 g^{-1})	95	95	197	584
Molecular mass (kDa)				
From Porod volume (×0.625)	81	103	129	458
From excluded volume (×0.5)	99	100	120	338
From $I(0)$	77.2	84.5	146	280
From volume of correlation	96.4	95.6	120	355
From sequence	78.4	78.4	161.6	480
Software employed				
Primary data reduction	*PIPELINE*	*PIPELINE*	*PIPELINE*	*PIPELINE*
Data processing	*PRIMUS*	*PRIMUS*	*PRIMUS*	*PRIMUS*
Ab initio modelling	*GASBOR*	n.p.	*DAMMIN*	*DAMMIN*
Validation and averaging	*DAMAVER*/*DAMCLUST*	n.p.	*DAMAVER*/*DAMCLUST*	*DAMAVER*/*DAMCLUST*
Rigid-body modelling	n.p.	n.p.	*CORAL*, *EOM*	*SASREFMX*
Computation of model intensities	n.p.	n.p.	*CRYSOL*	*CRYSOL*
Three-dimensional representations	*PyMOL*	n.p.	*PyMOL*	*PyMOL*

† $q = 4\pi\sin(\theta)/\lambda$, where 2θ is the scattering angle and λ = 0.99 Å is the wavelength. ‡ Calculated merging data for concentrations of 2.0 and 5.53 mg ml^{-1} for oxidized PRK and 0.50 and 5.0 mg ml^{-1} for GAPDH–CP12. Calculated averaging data for concentrations of 2.0 and 2.86 mg ml^{-1} for GAPDH–CP12–PRK.

components in the presence of 5 m*M* NAD and 5 m*M* oxidized DTT. GAPDH, PRK and CP12 were present in a 2:1:2 ratio (on a subunit basis) and excess CP12 was removed by gel filtration as described above for the binary complex. 1 m*M* NAD was immediately added to the eluted GAPDH–CP12–PRK ternary complex and it was brought to a final concentration of 10 mg ml^{-1}.

The purity of all samples was verified by SDS–PAGE followed by Coomassie staining.

2.2. SAXS data acquisition and analysis

SAXS data were collected on the BioSAXS beamline BM29 at ESRF, Grenoble, France (Pernot *et al.*, 2013). The data-collection parameters are reported in Table 1. Flushed volumes of 60 µl were used for each protein sample. In order to optimize data acquisition, various protein concentrations below 10 mg ml^{-1} were measured (Table 1). A set of ten consecutive 1 s exposures were made on each sample and were compared to assess radiation damage.

The scattering contribution of the capillary filled with buffer was subtracted and the data were placed on a relative scale using bovine serum albumin (BSA) as a standard (molecular weight 66.5 kDa).

The $I(0)$ and radius of gyration (R_g) were calculated using the Guinier approximation $I(q) = I(0)\exp[-(qR_g)^2/3]$ (Guinier, 1939) as implemented in *PRIMUS* (Konarev *et al.*, 2003; Petoukhov *et al.*, 2012). The $p(r)$ function, from which $I(0)$, the maximum particle dimension (D_{max}) and R_g were estimated, and the reciprocal-space fit (model-independent) of the experimental data were computed using *GNOM* (Svergun, 1992). This fit was used as a good reference in the selected angular range for the model-dependent fitting attempts.

The molecular weight (MW; Table 1) was estimated from (i) the Porod invariant (Porod, 1982) as 0.625 times the Porod volume (V_p) for roughly globular particles (Petoukhov *et al.*, 2012), (ii) the excluded volume of averaged hydrated particles (V_d) computed using *DAMMIN* (Svergun, 1999) as 0.5 × V_d, (iii) comparison of the $I(0)$ with that of native BSA at pH 7.4

(Mylonas & Svergun, 2007) and (iv) the SAS invariant volume-of-correlation length (V_c) through a power-law relationship between V_c, R_g and MW that has been parametrized for proteins, mixed nucleic acid–protein complexes and RNA structures (Rambo & Tainer, 2013*b*) using the program *ScÅtter*.

Among the various methods of mass estimation, the MW obtained from V_p and V_d are affected by an error of 20%, while the estimated error is 10% when the forward scattering $I(0)$ is used (Graewert *et al.*, 2015). For native BSA, largely used for calibrating the intensities, $I(0)$ can be affected by the presence of oligomers, with a deviation of 17% from the expected value (Mylonas & Svergun, 2007; Petoukhov *et al.*, 2013).

For a MW determined using a native monomeric protein as a molecular-weight standard (Kozak, 2005) or the absolute calibrated $I(0)$ expressed in cm^{-1}, the major source of errors is generally owing to the calculation of the electron-density contrast, which depends on the number of electrons per mass of dry protein, the electron density of the solvent and the partial specific volume of the protein (monomeric or heteromeric).

2.3. *Ab initio* analysis and rigid-body modelling

The calculation of model intensities from atomic structures was performed by *CRYSOL* (Svergun *et al.*, 1995) using default parameters.

For the GAPDH–CP12 and the GAPDH–CP12–PRK complexes, the *ab initio* low-resolution envelopes were reconstructed by *DAMMIN* (Svergun, 1999); 12 and ten structures were generated, respectively. The *ab initio* model of the ternary complex was computed using a starting model obtained applying *DAMAVER* (Volkov & Svergun, 2003) to a number of similar rigid-body structures with a cutoff volume equal to 500 nm^3.

For PRK, the *ab initio* program *GASBOR* (Svergun *et al.*, 2001) was used. This software uses the average scattering factor of amino acids and allows the particle to be modelled as a protein chain-like assembly of dummy residues, the number of which is constrained according to the known sequence of the protein. The reliability of the model calculated from *GASBOR* is considered to be good up to a higher resolution than the *DAMMIN* dummy-atom method.

The sequence and the homodimeric state of PRK (349 amino acids × two monomers = 698 amino acids) were given as inputs in *GASBOR*. The calculations were performed with both *P*1 and *P*2 symmetry. A series of 20 models was generated. We complemented the *ab initio* analysis of PRK using the dummy-atom method implemented in *DAMMIF* (Franke & Svergun, 2009) for fitting the data collected from both the oxidized and the reduced form. 20 calculations with both *P*1 and *P*2 symmetry were performed.

Rigid-body modelling of the GAPDH–CP12 complex was performed by the program *CORAL* (Petoukhov *et al.*, 2012). For the GAPDH–CP12–PRK complex, the software *SASREF* (Petoukhov & Svergun, 2005) and the implemented version *SASREFMX* (Petoukhov *et al.*, 2013) that takes into account partial complex dissociation were used.

For the binary complex, a twofold symmetry was applied to a pair of GAPDH subunits, one binding CP12 and one free. In order to build the complete complex, the missing CP12 amino acids were generated as dummy residues. Crystallographic contacts highlighted by the *Contact Map Analysis* software (Sobolev *et al.*, 2005) were used as constraints in the calculation.

Since the system of the ternary complex presented numerous degrees of freedom, some constraints were imposed in the calculations. The stoichiometry suggested a twofold symmetry of the system and mild contacts were imposed in order to keep every CP12 bound to GAPDH not further than 10 Å from a PRK model moiety. This condition is based on the knowledge that PRK has no significant affinity for GAPDH alone, but only for the GAPDH–CP12 complex (Marri *et al.*, 2005, 2008).

In all cases, repeated runs of three-dimensional reconstruction algorithms were performed. The similarity of the obtained structures was checked by *DAMAVER* (Volkov & Svergun, 2003) and *DAMCLUST* (Petoukhov *et al.*, 2012), in which the superposition was performed by the *SUPCOMB* code (Kozin & Svergun, 2001). A normalized spatial discrepancy (NSD) value was calculated and used as a parameter to determine the difference between two three-dimensional models.

EOM (Bernadó *et al.*, 2007) was used to assess the flexibility of the CP12 N-terminal region (residues 1–57) and to model the binary complex. A pool of conformers for the missing CP12 amino acids were randomly generated as a chain of dummy residues. The structures with unacceptable superimpositions between dummy atoms and the rigid-body structure were discarded. From a starting random pool consisting of about 10 000 conformers, an ensemble was selected in order to best fit the experimental SAXS curve.

All programs used for SAXS data analysis and reconstruction listed in Table 1 belong to the *ATSAS* package v.2.5 (Konarev *et al.*, 2006; Petoukhov *et al.*, 2012). The graphical representations of the three-dimensional models were generated with *PyMOL* (v.1.4.1; Schrödinger; DeLano, 2002).

2.4. Hydrodynamic calculations

The program *HYDROPRO* (Garcìa de la Torre *et al.*, 2000) was used to estimate the hydrodynamic radius (R_h) of the proposed models.

The calculation was performed on the *ab initio* model of oxidized PRK and on hybrid all-atom/dummy-residue models for the binary and ternary complexes. The values of the radii were optimized as suggested by the authors (Ortega *et al.*, 2011), and radii of 4.5, 3.3 and 3.7 Å for PRK and the binary and ternary complexes, respectively, were chosen. Room temperature and viscosity of water were considered in the calculations.

3. Results

PRK, CP12 and GAPDH (isoform A_4) from *A. thaliana* were all produced as recombinant proteins in *E. coli* and analyzed in different conditions and combinations. The PRK samples were either fully reduced or oxidized. The binary complex between GAPDH and oxidized CP12 (GAPDH–$CP12_2$; Fermani *et al.*, 2012) and the ternary complex also involving oxidized PRK (Marri *et al.*, 2008) were both stabilized by NAD. Information on SAXS data collection and analysis is summarized in Table 1. The final scattering curves are reported as normalized Kratky plots (Fig. 1) that allow a comparison between proteins with different dimensions and masses (Durand *et al.*, 2010). The symmetrical plots of the binary and ternary complexes suggest that they behave as globular proteins, while the nonzero (positive) skewness of the PRK plot (either reduced or oxidized) indicates a more elongated structure (Durand *et al.*, 2010).

3.1. Oxidized and reduced phosphoribulokinase

The low q-range behaviour of the experimental scattering pattern of oxidized PRK (Fig. 2*a*) showed a slight increase in the slope of the Guinier plot on increasing the protein concentration from 2.0 to 5.5 mg ml^{-1} (Supplementary Fig. S1*a*), suggesting partial protein aggregation. In order to minimize this effect, only samples at lower concentrations were further analyzed and experimental points before the linear Guinier region were excluded. The data provided a reliable Guinier linear fit with little or no dependence of the slope on the first angular points when considered in the q interval 0.02–0.03 Å^{-1} (Supplementary Fig. S1*b*), from which an R_g of 39.4 Å was estimated. From the forward scattering [$I(0)$] a molecular weight (MW) of 77.2 kDa was estimated using BSA as a standard (Table 1). This mass is compatible with the presence of dimeric PRK in solution (the MW calculated from the sequence is 78.4 kDa; Table 1). The curve derived by merging data obtained at the lowest ($q < 0.08$ Å^{-1}) and the highest concentration was used for further analysis of oxidized PRK.

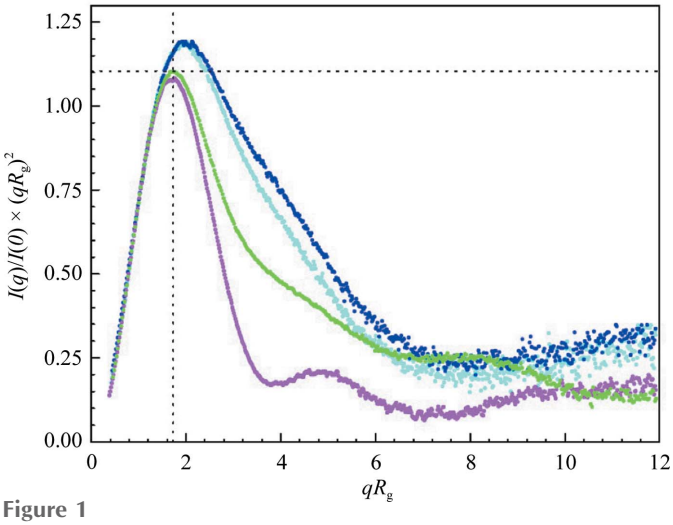


Figure 1
Dimensionless Kratky plots of the ternary complex and its single components. Dimensionless Kratky plots calculated from the experimental data for oxidized PRK (cyan; R_g = 39.4 Å), reduced PRK (blue; R_g = 40.9 Å), the GAPDH–CP12 complex (magenta; R_g = 34.0 Å) and the GAPDH–CP12–PRK complex (green; R_g = 56.7 Å).

The SAXS data from reduced PRK samples were less reliable since no correlation was found between the low-q behaviour (expressed as R_g) of the scattering pattern and the sample concentration (Supplementary Fig. S1*a*). In addition, the linear region of the Guinier plot was slightly shifted towards higher q values (Fig. 2*b*, inset). In order to detect possible differences between reduced and oxidized PRK, the corresponding experimental curves (from q = 0.03 Å^{-1}) were superimposed and found to be virtually identical (Supplementary Fig. S2). For this comparison, the experimental curves were chosen among the concentration series that least deviated from ideal behaviour at low angles.

The pair-distance distributions [$p(r)$] computed from oxidized and reduced PRK data (Fig. 2*b* and Supplementary Fig. S3) showed an asymmetric bell-shaped profile, suggesting dimers with an elongated shape. The maximum particle dimension (D_{max}) and the R_g calculated from the $p(r)$ for the two PRK forms are reported in Table 1. By comparing the two $p(r)$ functions and the normalized Kratky plots (Figs. 1 and 2*b* and Supplementary Fig. S3), a slightly more compact conformation for oxidized PRK compared with the reduced form could be envisaged. However, an accurate assessment of the conformational changes caused by disulfide-bridge formation/reduction could not be achieved because of the partial aggregation of the reduced PRK sample that probably affected the scattering curve.

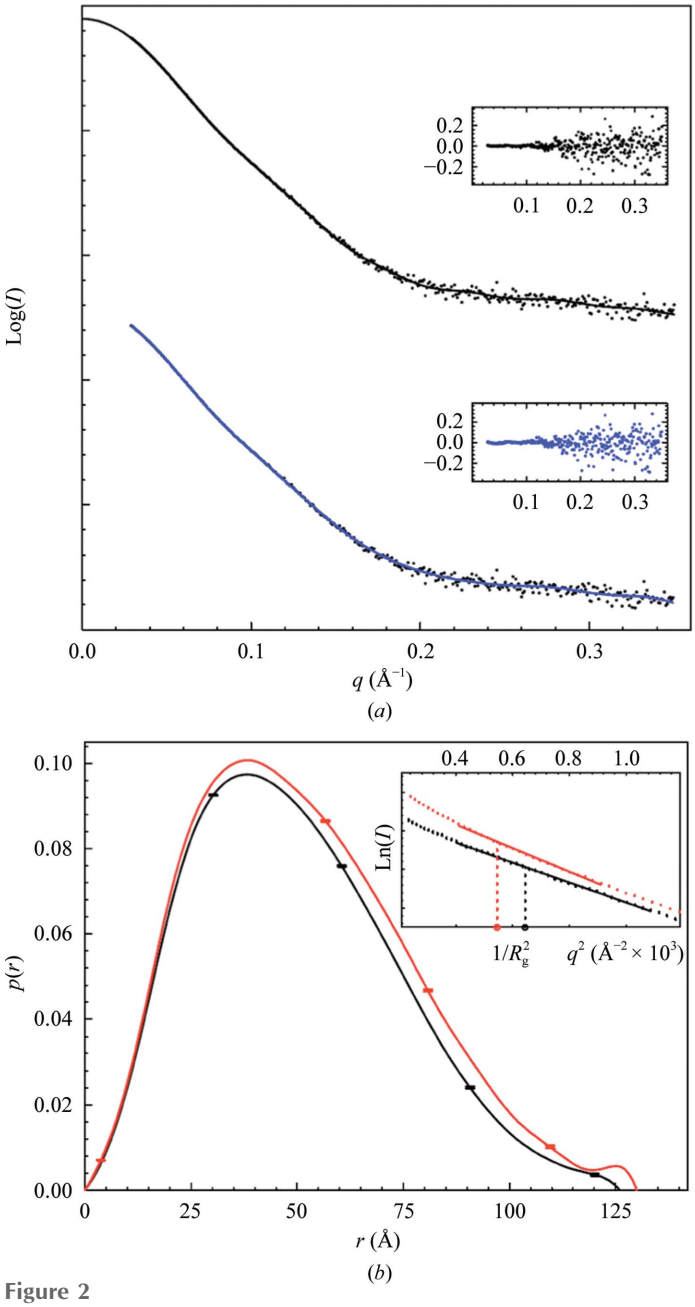
Since no high-resolution structure of *Arabidopsis* PRK is available, and its sequence identity with *Rhodobacter* PRK (Harrison *et al.*, 1998) is less than 19%, together with the consideration that *Arabidopsis* PRK is about 60 residues longer than the bacterial enzyme (349 *versus* 290 amino acids), only an *ab initio* modelling approach could be attempted. The modelling using *GASBOR* was performed only for oxidized PRK since the quality of the experimental data of reduced PRK was too low for a similar approach. Repeated calculations converged to a satisfactory agreement between the intensity computed from the model and the experimental data. The goodness of fit provided by the most representative low-resolution structure was comparable to that of the model-independent analysis (Fig. 2*a*) reported as a reference for a good fit. The residuals were similar in both cases (Fig. 2*a*, insets). The most representative model of oxidized PRK obtained applying *P*2 symmetry is shown in Fig. 2(*c*), while the *P*1 model is reported in Supplementary Fig. S4 together with the *P*2 model for better comparison. Substantial agreement between the *P*1 and *P*2 models can be observed. The normalized spatial discrepancy (NSD) values for the *P*1 and *P*2 models are reported in Supplementary Table S1. PRK, with dimensions of about 50 Å in width and 125 Å in length, is bent and screwed in shape, with two well defined lobes possibly corresponding to the two monomers. This overall shape proposed for PRK is confirmed by a complementary *ab initio* analysis of PRK using a dummy-atom method (*DAMMIF*)

applied to both redox forms (Supplementary Figs. S5 and S6). As expected from the close similarity between the scattering data of PRK in the two redox forms, the *ab initio* calculations gave comparable results. The NSD values for the models are reported in Supplementary Table S2. Even if a more pronounced bent shape for reduced PRK compared with the oxidized form could be envisaged from inspection of Supplementary Figs. S5 and S6, the NSD values found by superimposing the most representative models of oxidized and reduced PRK (0.660 and 0.703 for *P*1 and *P*2, respectively) suggested they are not statistically different.

Since PRK embedded in the ternary complex is oxidized (Marri *et al.*, 2005, 2009), the *ab initio GASBOR* model of oxidized PRK was employed for further modelling of the GAPDH–CP12–PRK complex.

3.2. GAPDH–CP12 complex

The experimental scattering data of the binary complex (Fig. 3*a*) did not show significant concentration effects, except for a very slight downward deviation observed in the $q <$ 0.02 Å range for the most concentrated sample (5 mg ml^{-1}), possibly owing to weak repulsive interactions. Indeed, the net



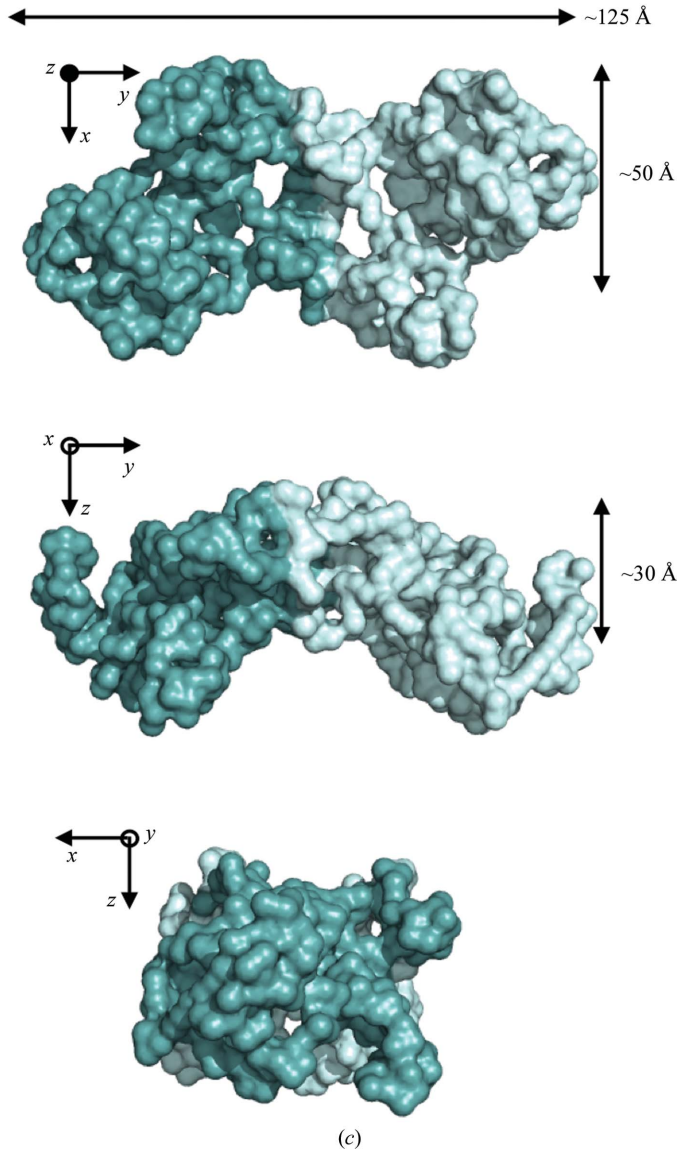


Figure 2
SAXS analysis and proposed *ab initio* model of oxidized PRK. (*a*) Oxidized PRK experimental scattering curve (black dots) and fits by model-independent *GNOM* (Svergun, 1992) analysis (as a reference for a good fit; black line) and *ab initio* modelling with *GASBOR* (Svergun *et al.*, 2001; blue line). Normalized residuals are shown in insets. (*b*) Model-independent $p(r)$ function for oxidized PRK (black line) and reduced PRK (red line). Only a few error bars are shown for the sake of clarity. In the inset the Guinier plot (dots) and the linear fit are shown for oxidized PRK (black line) and reduced PRK (red line). (*c*) The *ab initio* model of oxidized PRK computed by applying *P*2 symmetry is represented by three orthogonal orientations. Dummy waters bound to the surface of the molecule and introduced in the calculations are omitted. One half of the model is coloured darker for clarity. Lengths are shown for reference.

charge of the GAPDH–CP12 complex at pH 7.5, as calculated from the sequence, is highly negative and largely dependent on the N-terminal regions of the two CP12 proteins.

A scattering pattern obtained by merging low-concentration (0.5 mg ml^{-1}) and high-concentration (5 mg ml^{-1}) data was used for all further analyses. The R_g values obtained from either the Guinier fit or the $p(r)$ values were in good agreement (Table 1). The mass estimated from $I(0)$ was slightly lower than the value calculated from the protein sequences (Table 1).

The GAPDH–CP12 $p(r)$ function derived from the experimental data looked similar to that of a spherical particle with a radius of about 43 Å, with the addition of extended portions that caused the D_{max} to increase to 115 Å and made the curve slightly asymmetric (Fig. 3*b*).

In the crystal structure of the GAPDH–CP12 complex (PDB entry 3qv1; Fermani *et al.*, 2012) two CP12 fragments were bound to the A_4-GAPDH tetramer. Each CP12 fragment included 21 C-terminal amino acids engaged in interactions with GAPDH, whereas the remaining 56 N-terminal residues were disordered and were exposed to the solvent. The GAPDH sites that bind the CP12 fragments were either on the same side with respect to the *R* symmetry axis of the tetramer or on opposite sides (Fermani *et al.*, 2012). These alternative conformations of the binary complex are here named 'same-side' and 'opposite', respectively (Supplementary Fig. S7). The computed $p(r)$ functions of the two conformations are indistinguishable (Fig. 3*b*).

The intensity calculated from the high-resolution models, either 'same-side' or 'opposite' (not shown), was roughly in agreement with the experimental data, even though a systematic deviation was observed around $q = 0.1$ Å^{-1} (Fig. 3*a*, red line) similarly to that previously reported for cytosolic GAPDH alone (Torres-Bugeau *et al.*, 2012). This disagreement was probably owing to a different arrangement of the GAPDH subunits in solution. Indeed, the theoretical intensity computed for the optimized model of the GAPDH–CP12 complex obtained by a rigid-body procedure confirmed the above hypothesis, indicating that a subunit rearrangement (Supplementary Fig. S8) was sufficient to improve the fit (Fig. 3*a*, pink line). Repeated runs showed almost identical subunit arrangements, as indicated by the computed

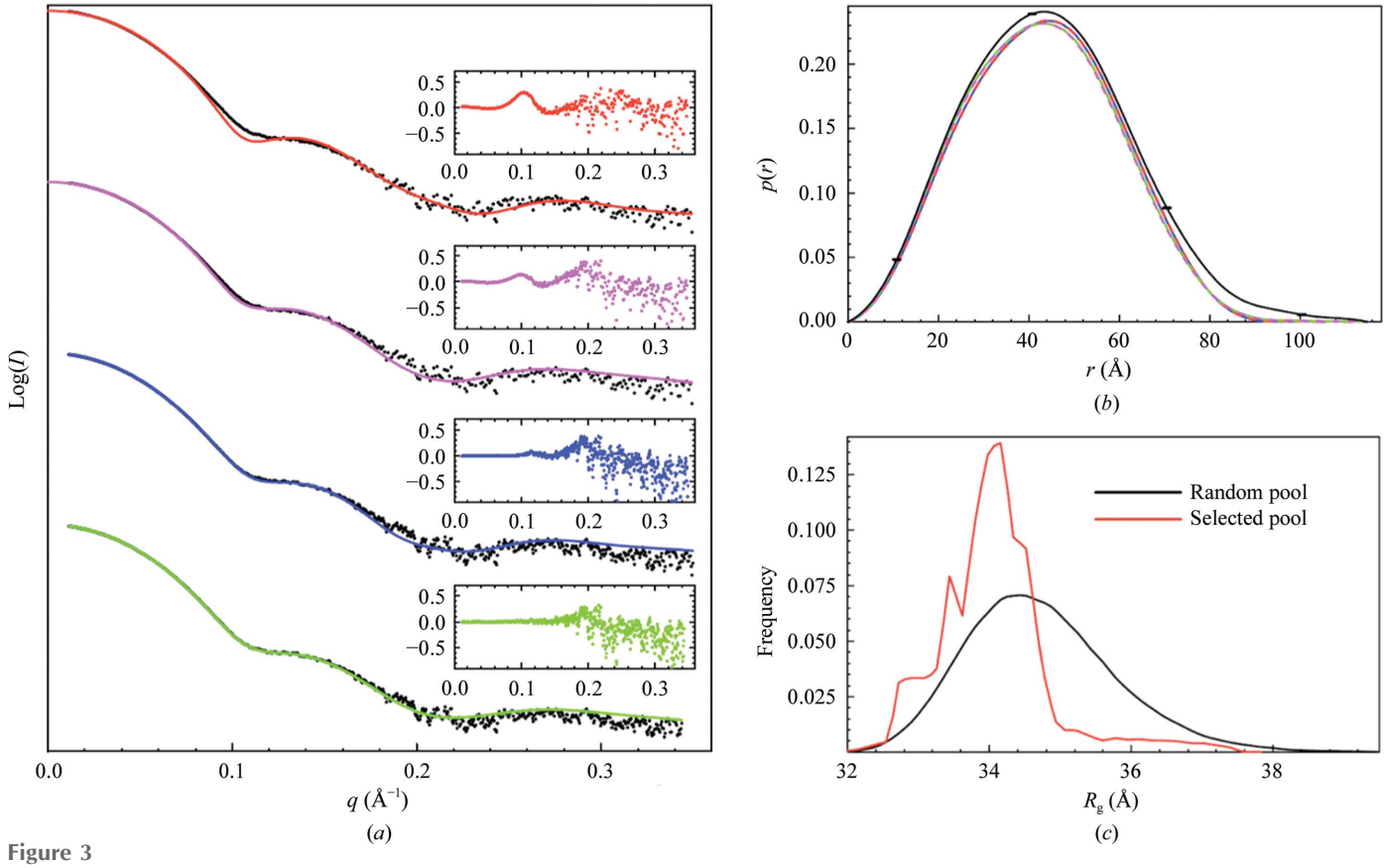


Figure 3
SAXS analysis of the GAPDH–CP12 complex. (*a*) Fits between the experimental (black dots) and the calculated intensities from the high-resolution model of the binary complex 'same-side' conformation (red line; PDB entry 3qv1), the *CORAL* (Petoukhov *et al.*, 2012) optimized rigid-body model without (magenta line) and including (blue line) dummy residues representing missing CP12 amino acids and the *EOM* (Bernadó *et al.*, 2007) optimized ensemble of coexisting different conformations of CP12 residues (green line) are shown. In the insets normalized residues are reported. (*b*) $p(r)$ functions calculated from the experimental data (black solid line), crystal structures (PDB entry 3qv1) of the binary complex 'same-side' (blue dashed line) and 'opposite' (red dashed line) conformations, and rigid-body optimized models without dummy residues of the binary complex 'same-side' (green dashed line) and 'opposite' (pink dashed line) conformations. Only a few error bars are shown for the sake of clarity. (*c*) The R_g distribution for models selected after the genetic algorithm procedure (red line) by *EOM* (Bernadó *et al.*, 2007) are biased towards lower values compared with the randomly generated pool (black line).

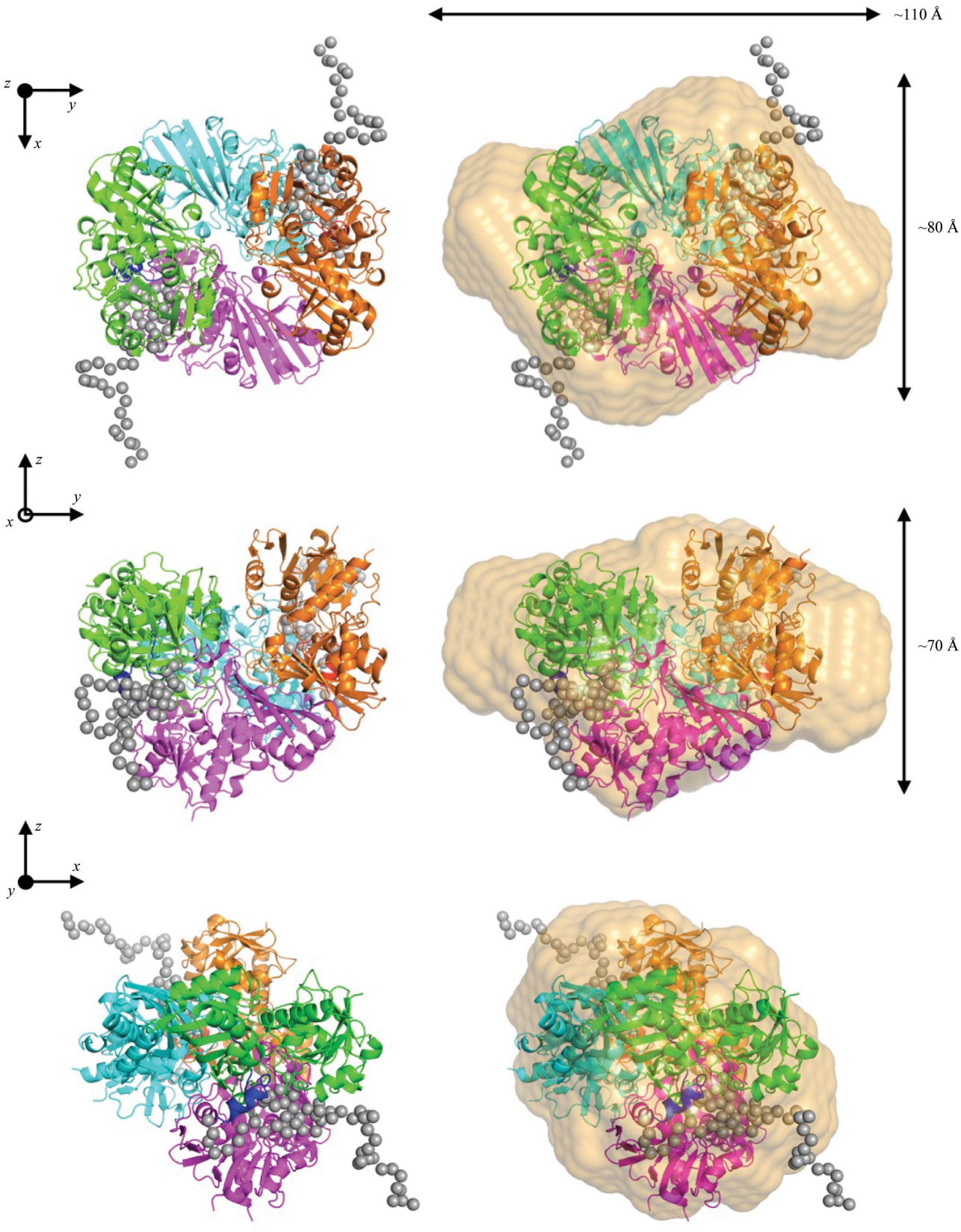


Figure 4
Proposed model of the GAPDH–CP12 complex in solution. The binary-complex model obtained with a combination of rigid-body, dummy-residue modelling and the ensemble-optimization method, and its superimposition with the *ab initio* model (orange envelope), are represented in the left and right columns, respectively. Three orthogonal orientations are shown. The GAPDH subunits and the C-terminal portion of CP12 are represented as cartoons and depicted in different colours (GAPDH chain *A* in cyan, chain *B* in magenta, chain *C* in green and chain *D* in orange; CP12 chain *G* in red and chain *H* in blue), whereas the N-terminal dummy residues are shown as grey spheres. Lengths are shown for reference.

normalized spatial discrepancy index ($\langle NSD \rangle = 0.611 \pm 0.054$). Moreover, the use of an *ab initio* approach in order to model the 112 residues of CP12 that the crystallographic model did not include (more than 10% of the overall mass) resulted in a further improvement of the overall fit ($\chi = 1.36$; Fig. 3*a*, blue line). The rearranged models of the two complex conformations showed identical $p(r)$ functions (Fig. 3*b*). On this basis, further analyses were only performed on the 'same-side' conformation.

Since the N-terminal region of CP12 bound to GAPDH is disordered, the best description of the protein complex in solution should involve the possibility for this flexible portion to explore different conformations. The pool selected after the genetic algorithm procedure (Bernadó *et al.*, 2007) showed an R_g distribution biased towards lower values compared with the randomly generated pool (Fig. 3*c*), indicating that more compact conformations of CP12 bound to GAPDH were preferentially populated in the complex. Since the complex was formed under oxidizing conditions, compact conformations are possibly favoured by the N-terminal disulfide bridge of oxidized CP12. The selected ensemble, the calculated intensity of which fits the experimental data with $\chi = 1.10$ (Fig. 3*a*, green line), was composed of about 70% of conformers with $R_g < 35$ Å and $D_{max} \simeq 130$ Å and about 30% of more extended structures with $R_g > 35$ Å and $D_{max} > 150$ Å. A particularly recurring conformer (in repeated runs of the genetic algorithm) with $R_g = 33.25$ Å and $D_{max} = 132.76$ Å (Fig. 4) was chosen as the best model for the binary complex to be used in the modelling of the ternary complex.

The presence of an intrinsically disordered protein such as CP12 in the binary complex makes this system partially unsuitable for *ab initio* modelling. However, the method was applied and the most representative *ab initio* envelope, when superimposed with the model obtained with a rigid-body/*ab initio* approach, showed a reasonable agreement (Fig. 4). The lobes that are not occupied by the GAPDH tetramer may represent the conformational space sampled by the flexible N-terminal portions of the CP12 chains.

3.3. GAPDH–CP12–PRK complex

The scattering data from the ternary complex did not display any significant concentration effect in the low-q range (Fig. 5*a*), but radiation damage and problems in background subtraction affected the data at the lowest concentrations. A curve obtained by averaging data at concentrations of 2.0 and 2.86 mg ml^{-1} was used for further analysis.

The $p(r)$ computed from the experimental data (Fig. 5*b*) gave an R_g of 57 Å and a D_{max} of 175 Å, which are in good agreement with the values obtained from the Guinier approximation (Table 1). The mass estimated from the $I(0)$ was lower than the value calculated from the amino-acid sequences and the experimentally determined stoichiometry of the complex (Table 1; Marri *et al.*, 2008). However, the values estimated using other methods (Table 1) were closer to the expected molecular mass.

Rigid-body modelling of the ternary complex was performed considering two GAPDH–CP12 complexes and two oxidized PRK dimers as structural components (Marri *et al.*, 2008). These models have been described in previous sections. The introduction of a small volume fraction of free components (GAPDH–CP12 complex and PRK) in the rigid-body modelling of the ternary complex improved the fit of the

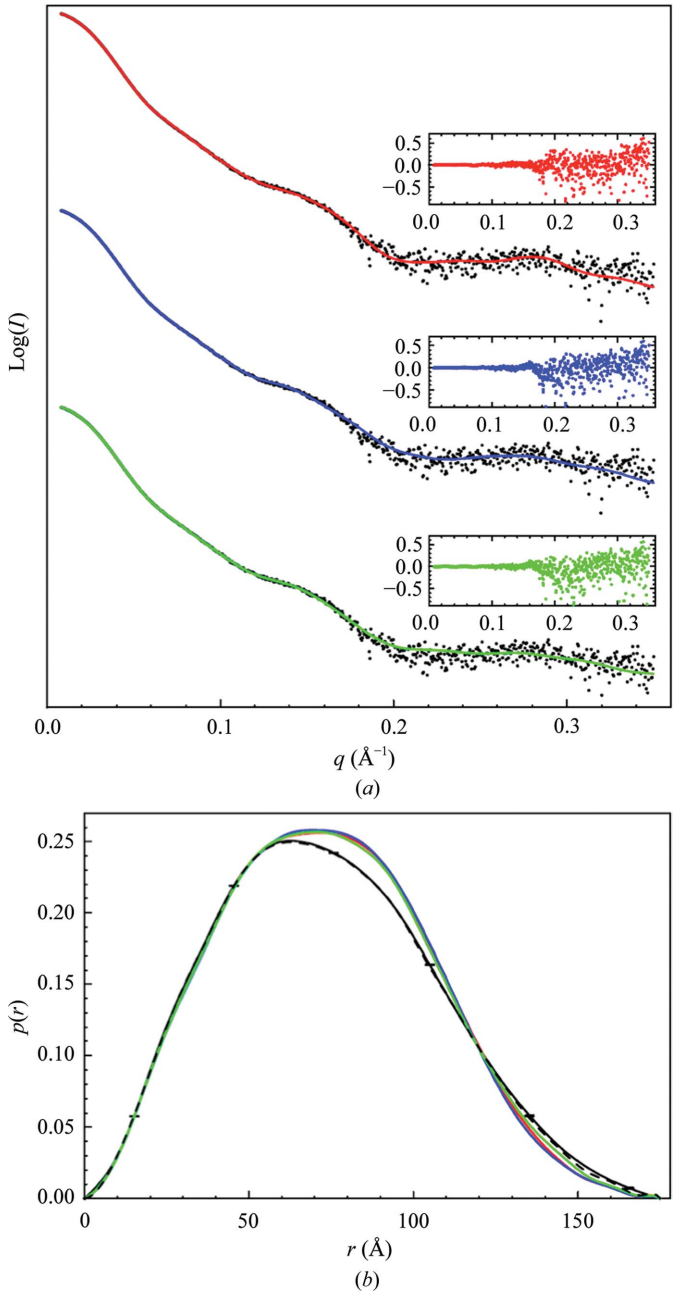


Figure 5
SAXS analysis of the GAPDH–CP12–PRK complex. (*a*) Fits between the experimental (black dots) and calculated (solid lines) intensities computed from the optimized rigid-body models 1 (red), 2 (blue) and 3 (green). The normalized residuals are shown in the insets. (*b*) $p(r)$ functions calculated from experimental data (black line), proposed models 1 (red), 2 (blue) and 3 (green) and a volume-fraction-weighted mixture composed of model 1, the GAPDH–CP12 complex model and the PRK model (black dashed line). Only a few error bars are shown for the sake of clarity.

experimental data. Given the numerous degrees of freedom of the system, the minimization procedures were repeated several times, obtaining an average agreement index of $\langle\chi\rangle = 1.28 \pm 0.02$ and a volume fraction of the associated complex of

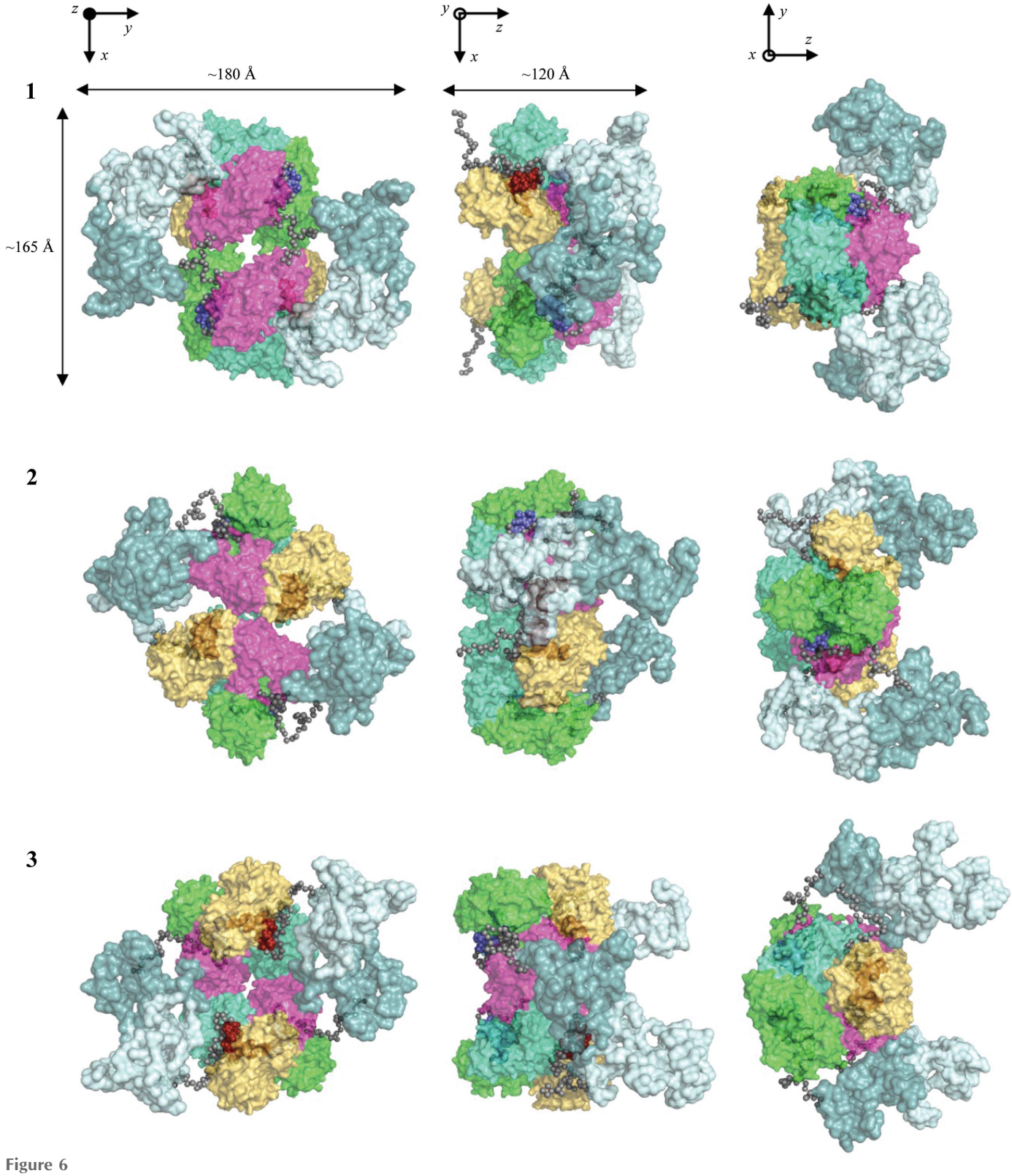


Figure 6
Proposed model of the GAPDH–CP12–PRK complex in solution. Three model assemblies indicated as 1, 2 and 3 were obtained by rigid-body modelling. The most representative structure for each assembly is shown. GAPDH tetramers (chain *A* in cyan, chain *B* in magenta, chain *C* in green and chain *D* in orange) and PRK dimers (monomers in cyan and deep cyan) are represented as surfaces, CP12s as spheres (chain *G* in red, chain *H* in blue and the N-terminal dummy residues in grey). Three orthogonal orientations are visualized. Lengths are shown for reference.

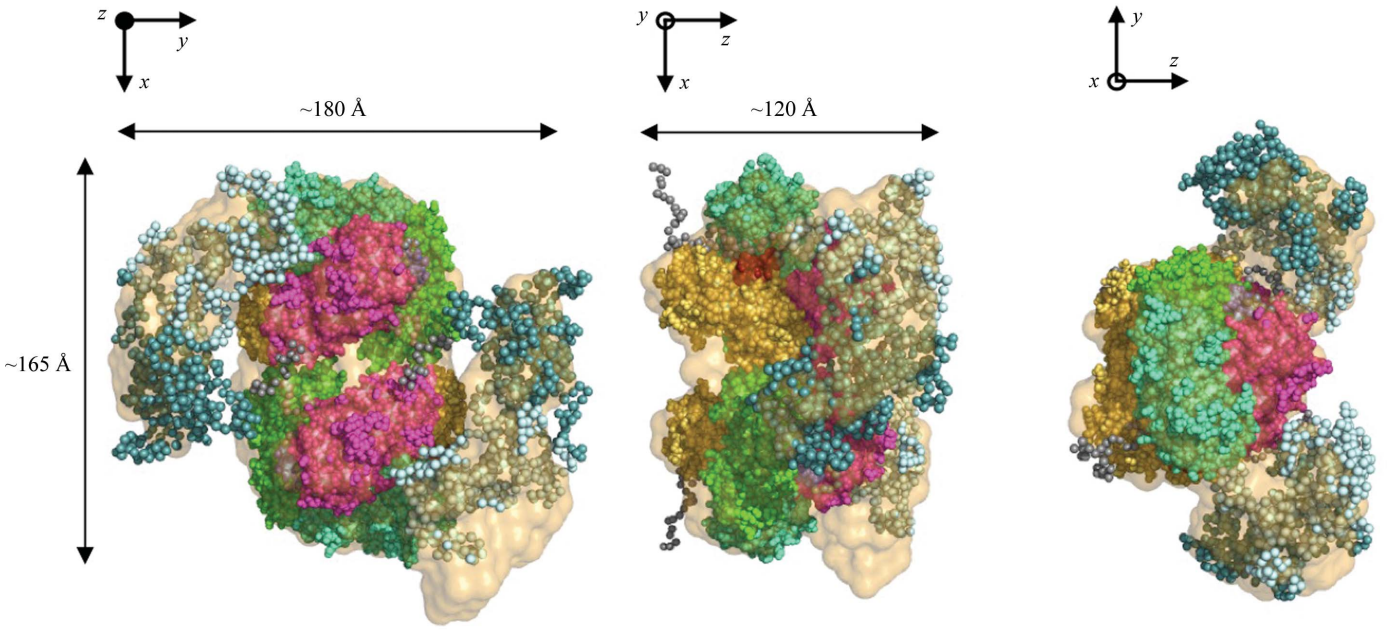


Figure 7
Superimposition between the rigid-body and *ab initio* models of the GAPDH–CP12–PRK complex. The reference structure of assembly 1 is superimposed on the *ab initio* model (orange envelope) of the ternary complex. GAPDH tetramers (chain *A* in cyan, chain *B* in magenta, chain *C* in green and chain *D* in orange), PRK dimers (monomers in cyan and deep cyan) and CP12s (chain *G* in red, chain *H* in blue and the N-terminal dummy residues in grey) are represented as spheres. Three orthogonal orientations are visualized. Lengths are shown for reference.

0.66 ± 0.01. The large majority of the optimized structures were compatible with a general assembly in which two PRK dimers tied together two binary complexes by interacting with the CP12s according to a previously proposed model based on biochemical grounds (Marri *et al.*, 2008). Notably, the two GAPDH tetramers were very close in all optimized structures and possibly interacted with each other. Three types of assembly were recognized and are indicated in descending order of occurrence frequency as models 1, 2 and 3. The most representative structure of each type of assembly is shown in Fig. 6. The superimposition and the classification of the computed ternary-complex models within each assembly type highlighted a general reproducibility of the GAPDH tetramer assembly, which probably gave the main contribution to the computed intensity, while the orientation of the PRK dimers was more variable (Supplementary Fig. S9).

The dimensions of the ternary complex were about 165 and 180 Å along the *x* and *y* axes, respectively (whichever type of assembly was considered), while along the *z* axis model 1 appeared more compact with respect to models 2 and 3. The three models diverged substantially in the relative orientation of the GAPDH tetramers, which hampered the reliable identification of a region of interaction. The orientation of the anisometric PRK dimers was also different between models, as the concave side of PRK faced the centre in models 1 and 2 while it was exposed outwards in model 3 (Fig. 6).

The calculated intensity was almost the same in all cases and showed a very good agreement with the experimental data (Fig. 5*a* and Table 2). Assembly 1 showed a good overall agreement index ($\langle\chi\rangle$) and a lower discrepancy among models. Moreover, the computed $p(r)$ functions of the three representative models are almost superimposed and the $p(r)$ function calculated from model 1 (considering complex dissociation) was in optimal agreement with the experimental $p(r)$ (Fig. 5*b*).

An *ab initio* model was calculated with the aim of validating the ternary-complex assembly obtained by the rigid-body approach. The imposition of a starting model in the *ab initio* calculations (see §2.3) provided as good a fitting as the default calculations (Supplementary Fig. S10). The superimposition between the *ab initio* and the rigid-body models was highly satisfying (Fig. 7), thereby confirming the overall validity of our procedure.

3.4. Hydrodynamic calculations from recovered models

The low-resolution models obtained for oxidized PRK and for the GAPDH–CP12 and GAPDH–CP12–PRK complexes were validated by computing the corresponding hydrodynamic radii (R_h) and comparing the results with the experimental values reported by Marri *et al.* (2008).

Although the computed R_h value (4.2 nm) of PRK was almost 30% higher than the previously published value (3.3 nm; Table 3), the agreement for both the binary and the ternary complex was very satisfying (Table 3), further confirming the validity of our models of complex structures.

4. Discussion

The pathway of CO_2 fixation in oxygenic photosynthetic organisms, the Calvin–Benson cycle, is the major limiting factor to photosynthetic efficiency under saturating light

Table 2
Average parameters for the three types of assembly (1, 2 and 3) obtained in the rigid-body modelling of the ternary complex.

The normalized spatial discrepancy (NSD) index calculated among the members of the same group, the χ value of the fit and the associated volume fraction f_{assoc} of the complex are reported. The NSD calculated among representative solutions of the different assemblies (1–2–3) is also reported.

Models	⟨NSD⟩	⟨χ⟩	⟨f_{assoc}⟩
1	1.10 ± 0.09	1.28 ± 0.02	0.654 ± 0.005
2	1.45 ± 0.10	1.28 ± 0.01	0.658 ± 0.014
3	0.88 ± 0.24	1.30 ± 0.01	0.661 ± 0.003
1–2–3	1.80 ± 0.04		

Table 3
Comparison between the hydrodynamic radii estimated by translational diffusion coefficients computed with *HYDROPRO* (Garcia de la Torre *et al.*, 2000) for the proposed models of oxidized PRK and the binary and ternary complexes and those previously estimated by SEC-MALS-QELS in Marri *et al.* (2008).

	R_h (nm)	
Models	*HYDROPRO*	SEC-MALS-QELS
Oxidized PRK	4.2 ± 0.1	3.3 ± 0.3
GAPDH–CP12	4.6 ± 0.1	4.3 ± 0.5
GAPDH–CP12–PRK assemblies		
1	7.5 ± 0.1	
2	7.5 ± 0.1	7.0 ± 0.1
3	7.7 ± 0.1	

conditions. A deep knowledge of the Calvin–Benson cycle is essential for understanding, and possibly improving, the photosynthetic conversion of light into chemical energy at the base of the production of food, biofuels and biomass by plants, algae and cyanobacteria.

The GAPDH–CP12–PRK complex is composed of two enzymes of the Calvin–Benson cycle (PRK and GAPDH) and the small intrinsically disordered scaffold protein CP12. The complex accumulates in chloroplasts during the night, when the metabolic cycle needs to slow down (Scheibe *et al.*, 2002; Tamoi *et al.*, 2005; Howard *et al.*, 2008; Howard, Lloyd *et al.*, 2011). This complex purified from *Chlamydomonas reinhardtii* has been studied at low resolution by cryoelectron microscopy (cryoEM; Mouche *et al.*, 2002). The reported dimensions of the observed particles (200 × 100 Å) were roughly compatible with the dimensions of the *Arabidopsis* complex determined in this work by SAXS (about 180 × 165 × 120 Å). Attempts to localize GAPDH and PRK within the three-dimensional cryoEM volume of the *Chlamydomonas* complex failed for the second enzyme. Fitting GAPDH within the complex was less problematic and resulted in two well separated GAPDH tetramers with no direct interactions betwen each other. The presence of CP12 in the *Chlamydomonas* complex was still under debate at the time and its positioning within the complex was not even performed (Mouche *et al.*, 2002).

All attempts to crystallize the GAPDH–CP12–PRK complex have failed so far, preventing high-resolution structural studies. However, the stoichiometry of the *Arabidopsis* complex was ascertained and a speculative model was proposed based on biochemical data, in which two GAPDH tetramers were connected to two PRK dimers by four CP12 molecules. In the model, CP12 was envisaged as a flexible linker with one extreme (C-terminus) interacting with GAPDH and the other (N-terminus) with PRK (Marri *et al.*, 2008). Here, we provide for the first time an experimental model of the GAPDH–CP12–PRK ternary complex that, while confirming the basic features of the original speculative model (Marri *et al.*, 2008), reveals structural insights that no other technique has provided so far. In order to shed light on the assembly process, the single components of the complex, *i.e.* PRK (both in the reduced and oxidized form), and the binary complex GAPDH–CP12 were also structurally characterized by SAXS.

Owing to the level of automation at advanced synchrotron beamlines and the progress in computational methods in the last decade, SAXS has become especially useful for protein systems that are difficult to crystallize or when the crystallographic structure alone cannot explain the biological functions in solution. By using *ab initio* and rigid-body modelling (Svergun, 1999; Chacón *et al.*, 2000; Heller *et al.*, 2003), the shapes of monomeric forms and protein complexes can be assessed (Putnam *et al.*, 2007; Mertens & Svergun, 2010; Nogales *et al.*, 2010; Pons *et al.*, 2010).

Analyses of the SAXS data suggested that PRK has an elongated shape in either redox condition, with a slightly more compact conformation in the case of the oxidized enzyme (Figs. 1, 2*b* and Supplementary Fig. S3). However, an accurate assessment of the conformational changes derived from the formation/reduction of the disulfide bond could not be achieved. The only PRK for which a crystal structure is available is that from *R. sphaeroides* (PDB entry 1a7j; Harrison *et al.*, 1998). Unfortunately, *Rhodobacter* PRK is less than 19% identical to *Arabidopsis* PRK, which is also about 60 amino acids longer. For this reason, only an *ab initio* model could be computed from the SAXS data for *Arabidopsis* PRK. The elongated, bent and screwed shape of oxidized PRK (Fig. 2*c*, Supplementary Figs. S4 and S6) appeared to be well adapted to function as a bridge between two GAPDH–CP12 complexes. As a major limitation of *ab initio* models, single amino acids could not be localized, preventing any speculation on which residues may be specifically involved in protein–protein interactions.

The $p(r)$ function, the Kratky plot and the *ab initio* model computed from the scattering data indicate that the GAPDH–CP12 complex (Figs. 1, 3*b* and 4) is constituted by a spherical main moiety with the addition of two opposite unfolded tails corresponding to the N-terminal regions of the two bound CP12 molecules. The rigid-body modelling approach using the high-resolution structure of the binary complex (in which only the C-terminal ends of CP12 were visible; PDB entry 3qv1; Fermani *et al.*, 2012) highlighted a clear rearrangement of the GAPDH subunits in solution (Supplementary Fig. S8). The CP12 amino acids missing from the high-resolution structure of the binary complex have been modelled as dummy residues in the SAXS model. Consistent with the flexible and unfolded state of unbound CP12 moieties (Matsumura *et al.*, 2011;

Fermani *et al.*, 2012), an ensemble of possible structures were generated. *EOM* analysis (Bernadó *et al.*, 2007) indicated that CP12 preferentially adopted more compact conformations (R_g < 35 Å) than extended conformations randomly oriented into the solvent (Fig. 3*c*). This preference is probably determined by the N-terminal disulfide of oxidized CP12, but nonspecific interactions with GAPDH could not be excluded. Complex formation between GAPDH and CP12 is known to depend on the C-terminal disulfide of CP12, which stabilizes a circular structural motif recognized by GAPDH (Fermani *et al.*, 2012). Similarly, a compact conformation of the N-terminal half of CP12 bound to GAPDH may be essential for the interaction with PRK during ternary-complex assembly.

The fit of the experimental data obtained from ternary-complex samples was greatly improved by assuming a fraction of dissociated complex in the calculations. This dissociated fraction was estimated at around 35% (Table 2), in agreement with the reversible nature of this regulatory complex, but was possibly also caused by the dilution required for the SAXS analysis. Consistently, it has previously been shown that the ternary complex from *Chlamydomonas* can partially dissociate upon dilution (Mouche *et al.*, 2002).

The partial dissociation cannot completely explain the lower mass values estimated from the $I(0)$ and excluded volume (Table 1) with respect to the theoretical molecular weight calculated from stoichiometry and sequence (Table 1). The expected $I(0)$ for the ternary complex, partially dissociated into PRK dimers and GAPDH–CP12 complexes, was estimated using the formula $I(0) = \sum_k v_k I_k(0)$ (Petoukhov *et al.*, 2013), where v_k is the volume fraction and $I_k(0)$ is the forward-scattering intensity contribution derived from the sequence of each of k species in solution. Assuming a volume fraction for the complex equal to 0.65, $I(0)$ turned out to be about 360 kDa, which is higher than the observed $I(0)$ value but still lower than the MW calculated from the sequence. Since $I_k(0)$ is a function of the particle electron-density contrast, the observed discrepancy may depend on a lower mean electron density of the multiprotein complex compared with a globular protein such as BSA, which was used as a standard. On the other hand, the consistency of the SAXS fit results that was achieved by assuming a $(\text{GAPDH–CP12}_2\text{–PRK})_2$ stoichiometry in the calculations was not achieved by using an alternative heteromeric composition with a lower mass (data not shown), thus strengthening the conclusion that the previously reported stoichiometry is the only one compatible with the scattering data.

Three model assemblies of the ternary complex were obtained (Fig. 6), all characterized by a good fit to the experimental data (Table 2). Assembly 1, showing a higher frequency and a lower discrepancy among models, could be considered as the most representative. It is possible that three complex populations with conserved stoichiometry but different shapes could coexist in solution, implying an intrinsic flexibility of the ternary complex. However, it should be taken into account that the low resolution of the *ab initio* model of PRK might impede a clear discrimination between the proposed ternary-complex models.

In all model assemblies each PRK dimer interacts with two GAPDH–CP12 complexes through their CP12s, although PRK adopts different orientations in the different assemblies (Fig. 6). The reciprocal orientation of the two GAPDH tetramers is relatively conserved within each assembly and the most surprising result was the interaction between GAPDH tetramers. Such a compact structure for the ternary complex was not predicted by the speculative model of Marri *et al.* (2008) nor was it suggested by cryoEM studies on the algal complex (Mouche *et al.*, 2002).

The propensity of plant GAPDH to aggregate has long been known. The main plant GAPDH isoform (A_2B_2) forms stable and inactive hexadecamers (A_8B_8) when it binds NAD(H) under oxidizing conditions. The effect is strictly dependent on both the redox state of the CTE of the B subunits and on the coenzyme bound to GAPDH (Baalmann *et al.*, 1996; Sparla *et al.*, 2002, 2005). The CTE is homologous to, and is evolutionarily derived from, the C-terminal end of CP12 (Petersen *et al.*, 2006). The crystal structure of oxidized A_2B_2-GAPDH bound to NADP shows that the CTE is located within a cleft close to the coenzyme-binding site and is delimited by two GAPDH subunits (Fermani *et al.*, 2007). The oligomerization of oxidized A_2B_2-GAPDH into its hexadecameric form requires the replacement of NADP(H) by NAD(H). Unfortunately, the crystal structure of A_8B_8-GAPDH is unknown and the molecular basis of the GAPDH auto-assembly process remains elusive. Interestingly, the novel structural insights provided by the GAPDH–CP12–PRK complex suggest that the capability to build supramolecular complexes may be an intrinsic property of any type of photosynthetic GAPDH, either the A_4 or A_2B_2 isoforms. In this hypothesis, the interaction between CP12 (or its evolutionarily derived CTE) and GAPDH bound to NAD(H) is supposedly responsible for a conformational effect on GAPDH that favours interaction between tetramers. Since GAPDH–CP12 complexes do not further associate in the absence of PRK, we speculate that the interaction between binary complexes is sterically hindered and electrostatically unfavoured by the flexible and negatively charged N-terminal ends of CP12. In this view, the role of PRK in the formation of the ternary complex would consist of dislocating the CP12 N-terminal ends, clearing the interaction surfaces of GAPDH tetramers. In contrast, the assembly of A_2B_2-GAPDH would be independent of PRK because no dislocation of cumbersome polypeptides (like the N-terminal region of CP12) is required.

In conclusion, a polypeptide of about 20 amino acids found either at the C-terminus of CP12 or as a C-terminal extension in the B subunits is the reason why GAPDH isoforms of oxygen phototrophs are so different from their counterparts in non-photosynthetic organisms. Not only does this small polypeptide confer sensitivity to thioredoxins (Marri *et al.*, 2009) but it also confers auto-assembly properties controlled by pyridine nucleotides [NAD(H)/NADP(H)], both functions being involved in the light/dark regulation of the Calvin–Benson cycle.

Acknowledgements

We thank the ESRF for synchrotron beam-time allocation (BAG proposal MX1640) and the staff of beamline BM29 for technical support. This work was supported by the University of Bologna, project FARB2012. SF and GF thank the CIRCSMB.

References

Baalmann, E., Scheibe, R., Cerff, R. & Martin, W. (1996). *Plant Mol. Biol.* **32**, 505–513.
Balsera, M., Uberegui, E., Schürmann, P. & Buchanan, B. B. (2014). *Antioxid. Redox Signal.* **21**, 1327–1355.
Bernadó, P., Mylonas, E., Petoukhov, M. V., Blackledge, M. & Svergun, D. I. (2007). *J. Am. Chem. Soc.* **129**, 5656–5664.
Brandes, H. K., Hartman, F. C., Lu, T.-Y. S. & Larimer, F. W. (1996). *J. Biol. Chem.* **271**, 6490–6496.
Chacón, P., Díaz, J. F., Morán, F. & Andreu, J. M. (2000). *J. Mol. Biol.* **299**, 1289–1302.
DeLano, W. L. (2002). *PyMOL*. http://www.pymol.org.
Durand, D., Vivès, C., Cannella, D., Pérez, J., Pebay-Peyroula, E., Vachette, P. & Fieschi, F. (2010). *J. Struct. Biol.* **169**, 45–53.
Fermani, S., Sparla, F., Falini, G., Martelli, P. L., Casadio, R., Pupillo, P., Ripamonti, A. & Trost, P. (2007). *Proc. Natl Acad. Sci. USA*, **104**, 11109–11114.
Fermani, S., Trivelli, X., Sparla, F., Thumiger, A., Calvaresi, M., Marri, L., Falini, G., Zerbetto, F. & Trost, P. (2012). *J. Biol. Chem.* **287**, 21372–21383.
Franke, D. & Svergun, D. I. (2009). *J. Appl. Cryst.* **42**, 342–346.
García de la Torre, J., Huertas, M. L. & Carrasco, B. (2000). *Biophys. J.* **78**, 719–730.
Gontero, B. & Maberly, S. C. (2012). *Biochem. Soc. Trans.* **40**, 995–999.
Graciet, E., Gans, P., Wedel, N., Lebreton, S., Camadro, J. M. & Gontero, B. (2003). *Biochemistry*, **42**, 8163–8170.
Graewert, M. A., Franke, D., Ruskule, D., Kuhle, K., Flieger, A., Schäfer, B., Tartsch, B. & Svergun, D. I. (2015). *Sci. Rep.* **5**, 10734.
Guinier, A. (1939). *Ann. Phys.* **12**, 161–237.
Harrison, D. H., Runquist, J. A., Holub, A. & Miziorko, H. M. (1998). *Biochemistry*, **37**, 5074–5085.
Heller, W. T., Krueger, J. K. & Trewhella, J. (2003). *Biochemistry*, **42**, 10579–10588.
Howard, T. P., Fryer, M. J., Singh, P., Metodiev, M., Lytovchenko, A., Obata, T., Fernie, A. R., Kruger, N. J., Quick, W. P., Lloyd, J. C. & Raines, C. A. (2011). *Plant Physiol.* **157**, 620–631.
Howard, T. P., Lloyd, J. C. & Raines, C. A. (2011). *J. Exp. Bot.* **62**, 3799–3805.
Howard, T. P., Metodiev, M., Lloyd, J. C. & Raines, C. A. (2008). *Proc. Natl Acad. Sci. USA*, **105**, 4056–4061.
Koch, M. H. J., Vachette, P. & Svergun, D. I. (2003). *Q. Rev. Biophys.* **36**, 147–227.
Konarev, P. V., Petoukhov, M. V., Volkov, V. V. & Svergun, D. I. (2006). *J. Appl. Cryst.* **39**, 277–286.
Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J. & Svergun, D. I. (2003). *J. Appl. Cryst.* **36**, 1277–1282.
Kozak, M. (2005). *J. Appl. Cryst.* **38**, 555–558.
Kozin, M. B. & Svergun, D. I. (2001). *J. Appl. Cryst.* **34**, 33–41.
Lipfert, J. & Doniach, S. (2007). *Annu. Rev. Biophys. Biomol. Struct.* **36**, 307–327.
López-Calcagno, P. E., Howard, T. P. & Raines, C. A. (2014). *Front. Plant Sci.* **5**, 9.
Marri, L., Pesaresi, A., Valerio, C., Lamba, D., Pupillo, P., Trost, P. & Sparla, F. (2010). *J. Plant Physiol.* **167**, 939–950.
Marri, L., Trost, P., Pupillo, P. & Sparla, F. (2005). *Plant Physiol.* **139**, 1433–1443.
Marri, L., Trost, P., Trivelli, X., Gonnelli, L., Pupillo, P. & Sparla, F. (2008). *J. Biol. Chem.* **283**, 1831–1838.
Marri, L., Zaffagnini, M., Collin, V., Issakidis-Bourguet, E., Lemaire, S. D., Pupillo, P., Sparla, F., Miginiac-Maslow, M. & Trost, P. (2009). *Mol. Plant.* **2**, 259–269.
Matsumura, H., Kai, A., Maeda, T., Tamoi, M., Satoh, A., Tamura, H., Hirose, M., Ogawa, T., Kizu, N., Wadano, A., Inoue, T. & Shigeoka, S. (2011). *Structure*, **19**, 1846–1854.
Mertens, H. D. T. & Svergun, D. I. (2010). *J. Struct. Biol.* **172**, 128–141.
Michelet, L., Zaffagnini, M., Morisse, S., Sparla, F., Pérez-Pérez, M. E., Francia, F., Danon, A., Marchand, C. H., Fermani, S., Trost, P. & Lemaire, S. D. (2013). *Front. Plant Sci.* **4**, 470.
Mouche, F., Gontero, B., Callebaut, I., Mornon, J.-P. & Boisset, N. (2002). *J. Biol. Chem.* **277**, 6743–6749.
Mylonas, E. & Svergun, D. I. (2007). *J. Appl. Cryst.* **40**, s245–s249.
Nogales, A., García, C., Pérez, J., Callow, P., Ezquerra, T. A. & González-Rodríguez, J. (2010). *J. Biol. Chem.* **285**, 1023–1031.
Ortega, A., Amorós, D. & García de la Torre, J. (2011). *Biophys. J.* **101**, 892–898.
Pernot, P. *et al.* (2013). *J. Synchrotron Rad.* **20**, 660–664.
Petersen, J., Teich, R., Becker, B., Cerff, R. & Brinkmann, H. (2006). *Mol. Biol. Evol.* **23**, 1109–1118.
Petoukhov, M. V., Billas, I. M., Takacs, M., Graewert, M. A., Moras, D. & Svergun, D. I. (2013). *Biochemistry*, **52**, 6844–6855.
Petoukhov, M. V., Franke, D., Shkumatov, A. V., Tria, G., Kikhney, A. G., Gajda, M., Gorba, C., Mertens, H. D. T., Konarev, P. V. & Svergun, D. I. (2012). *J. Appl. Cryst.* **45**, 342–350.
Petoukhov, M. V. & Svergun, D. I. (2005). *Biophys. J.* **89**, 1237–1250.
Pons, C., D'Abramo, M., Svergun, D. I., Orozco, M., Bernadó, P. & Fernández-Recio, J. (2010). *J. Mol. Biol.* **403**, 217–230.
Porod, G. (1982). *Small Angle X-ray Scattering*, edited by O. Glatter & O. Kratky. London: Academic Press.
Putnam, C. D., Hammel, M., Hura, G. L. & Tainer, J. A. (2007). *Q. Rev. Biophys.* **40**, 191–285.
Rambo, R. P. & Tainer, J. A. (2013*a*). *Annu. Rev. Biophys.* **42**, 415–441.
Rambo, R. P. & Tainer, J. A. (2013*b*). *Nature (London)*, **496**, 477–481.
Reichmann, D. & Jakob, U. (2013). *Curr. Opin. Struct. Biol.* **23**, 436–442.
Scagliarini, S., Trost, P. & Pupillo, P. (1998). *J. Exp. Bot.* **49**, 1307–1315.
Scheibe, R., Wedel, N., Vetter, S., Emmerlich, V. & Sauermann, S. M. (2002). *Eur. J. Biochem.* **269**, 5617–5624.
Sobolev, V., Eyal, E., Gerzon, S., Potapov, V., Babor, M., Prilusky, J. & Edelman, M. (2005). *Nucleic Acids Res.* **33**, W39–W43.
Sparla, F., Pupillo, P. & Trost, P. (2002). *J. Biol. Chem.* **277**, 44946–44952.
Sparla, F., Zaffagnini, M., Wedel, N., Scheibe, R., Pupillo, P. & Trost, P. (2005). *Plant Physiol.* **138**, 2210–2219.
Stanley, D. N., Raines, C. A. & Kerfeld, C. A. (2013). *Plant Physiol.* **161**, 824–835.
Svergun, D. I. (1992). *J. Appl. Cryst.* **25**, 495–503.
Svergun, D. I. (1999). *Biophys. J.* **76**, 2879–2886.
Svergun, D., Barberato, C. & Koch, M. H. J. (1995). *J. Appl. Cryst.* **28**, 768–773.
Svergun, D. I., Petoukhov, M. V. & Koch, M. H. J. (2001). *Biophys. J.* **80**, 2946–2953.
Tamoi, M., Miyazaki, T., Fukamizo, T. & Shigeoka, S. (2005). *Plant J.* **42**, 504–513.
Thieulin-Pardo, G., Avilan, L., Kojadinovic, M. & Gontero, B. (2015). *Front. Mol. Biosci.* **2**, 23.
Torres-Bugeau, C. M., Ávila, C. L., Raisman-Vozari, R., Papy-Garcia, D., Itri, R., Barbosa, L. R., Cortez, L. M., Sim, V. L. & Chehín, R. N. (2012). *J. Biol. Chem.* **287**, 2398–2409.
Trost, P., Fermani, S., Marri, L., Zaffagnini, M., Falini, G., Scagliarini, S., Pupillo, P. & Sparla, F. (2006). *Photosynth. Res.* **89**, 263–275.
Uversky, V. N. (2013). *Biopolymers*, **99**, 870–887.
Volkov, V. V. & Svergun, D. I. (2003). *J. Appl. Cryst.* **36**, 860–864.
Wedel, N. & Soll, J. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 9699–9704.
Zaffagnini, M., Bedhomme, M., Lemaire, S. D. & Trost, P. (2012). *Plant Sci.* **185–186**, 86–96.