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

Structure of *Pisum sativum* Rubisco with bound ribulose 1,5-bisphosphate

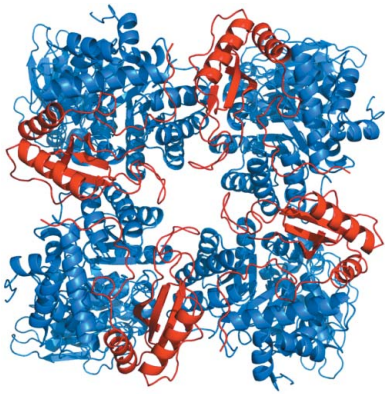
Peter C. Loewen,[a] Allison L. Didychuk,[b] Jacek Switala,[a] Rosa Perez-Luque,[c] Ignacio Fita[c] and Michele C. Loewen[b,d]*

[a]Department of Microbiology, University of Manitoba, 418 Buller Building, Winnipeg, MB R3T 2N2, Canada, [b]National Research Council of Canada, 110 Gymnasium Place, Saskatoon, SK S7N 0W9, Canada, [c]Institut de Biologia Molecular de Barcelona (IBMB–CSIC), Parc Científic, Baldiri Reixac 10, 08028 Barcelona, Spain, and [d]Department of Biochemistry, University of Saskatchewan, 107 Wiggins Road, Saskatoon, SK S7N 5E5, Canada

Correspondence e-mail: michele.loewen@nrc.ca

Received 19 October 2012
Accepted 19 November 2012

PDB Reference: Rubisco, 4hhh



The first structure of a ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) from a pulse crop is reported. Rubisco was purified from *Pisum sativum* (garden pea) and diffraction-quality crystals were obtained by hanging-drop vapour diffusion in the presence of the substrate ribulose 1,5-bisphosphate. X-ray diffraction data were recorded to 2.20 Å resolution from a single crystal at the Canadian Light Source. The overall quaternary structure of non-activated *P. sativum* Rubisco highlights the conservation of the form I Rubisco hexadecameric complex. The electron density places the substrate in the active site at the interface of the large-subunit dimers. Lys201 in the active site is not carbamylated as expected for this non-activated structure. Some heterogeneity in the small-subunit sequence is noted, as well as possible variations in the conformation and contacts of ribulose 1,5-bisphosphate in the large-subunit active sites. Overall, the active-site conformation most closely correlates with the 'closed' conformation observed in other substrate/inhibitor-bound Rubisco structures.

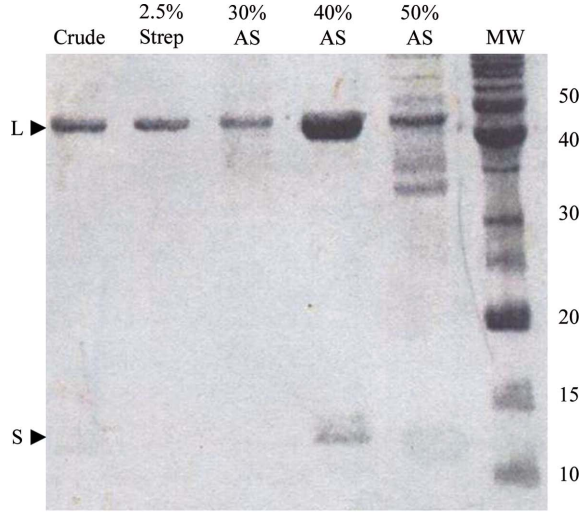
1. Introduction

The photosynthetic assimilation of atmospheric CO_2 into carbohydrates in plants occurs *via* the Calvin cycle and utilizes ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) for the primary CO_2-fixation step (Ellis, 1979). A molecule of CO_2 is first added to the five-carbon substrate ribulose 1,5-bisphosphate (RuBP), and the resulting six-carbon intermediate is cleaved to two molecules of 3-phosphoglycerate (3-PGA). The effectiveness of Rubisco in CO_2 fixation is compromised by its competing catalysis of RuBP oxygenation, which yields one molecule of 3-PGA and one molecule of phosphoglycolate. The first Rubisco structures revealed complexes comprised of large or of large and small subunits (Knight *et al.*, 1989; Andersson *et al.*, 1989; Chapman *et al.*, 1987; Suh *et al.*, 1987; Schneider *et al.*, 1986). In the case of higher plants, the complex was found to consistently have a hexadecameric structure with an L_8S_8 (L, large subunit, 475 residues; S, small subunit, 123 residues) arrangement referred to as form I Rubisco (Andersson, 2008). Catalysis begins with an initial activation step involving the conversion of an active-site lysine (Lys201 in the spinach enzyme) in the L subunit to its carbamylated derivative by the addition of CO_2, which also facilitates the binding of a magnesium ion required to complete the active site. RuBP can bind to both activated and non-activated Rubisco, but the latter forms a nonproductive complex that leads to a gradual loss of enzymatic activity *in vitro* (Zhu & Jensen, 1991). A comparison of structures reveals the substrate-bound or inhibitor-bound forms to have a more constrained or isolated RuBP-binding site, which is a result of localized changes in flexible loop regions (Taylor & Andersson, 1997*b*). Despite decades of study and the availability of Rubisco structures from spinach, tobacco, rice, algae, cyanobacteria and phototropic bacteria, many aspects of the regulation of Rubisco activity, such as the partitioning between carboxylation and oxygenation reactions, remain enigmatic (Andersson & Backlund, 2008). Recently, in the context of climate change and food security, Rubisco has become a major target of international crop-improvement consortia (Parry *et al.*, 2011). The focus of these efforts is on the modulation of Rubisco to increase its expression level,

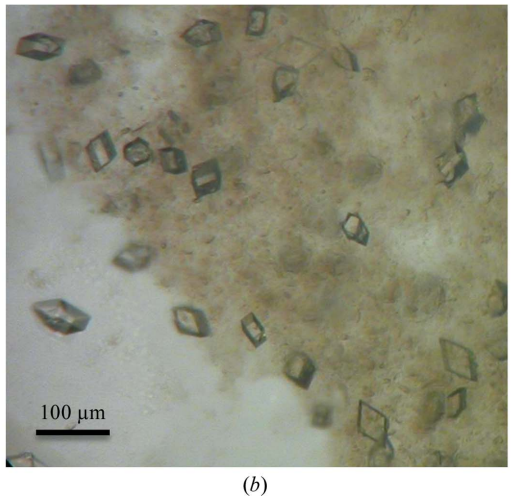
Table 1
Data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

Data collection	
Space group	$P2_122_1$
Unit-cell parameters (Å, °)	a = 109.81, b = 109.96, c = 201.270, $\alpha = \beta = \gamma = 90$
Resolution (Å)	48.25–2.20 (2.32–2.20)
No. of unique reflections	116075 (17084)
Completeness (%)	94.1 (95.5)
R_{merge}	0.222 (0.387)
$\langle I/\sigma(I)\rangle$	3.8 (2.0)
Multiplicity	2.7 (2.6)
Refinement statistics	
No. of reflections	110030
R factor (%)	20.9
R_{free} (%)	25.5
No. of non-H atoms	19681
No. of waters	789
R.m.s.d., bonds (Å)	0.009
R.m.s.d., angles (°)	1.44
Average B factor (Å^2)	9.3
Average B factor, waters (Å^2)	11.4



(*a*)



(*b*)

Figure 1
Preparation of diffracting crystals of *Ps*Rubisco. (*a*) Purification of *Ps*Rubisco. SDS–PAGE showing ammonium sulfate (AS) fractions. The 40% fraction contained the highest proportion of Rubisco and was used directly for crystallization after dialysis to remove AS. L, large subunit; S, small subunit. (*b*) Crystals of *Ps*Rubisco. The approximate dimensions of the *Ps*Rubisco crystals were 50 × 50 × 20 µm.

catalytic rate, thermal stability and availability of substrates, with the aim of developing higher yielding crops that sequester larger amounts of carbon dioxide from the environment. In this light, a recent structure of rice Rubisco has been published that yielded novel insights into mechanistic aspects of its regulation by specific effector molecules (Matsumura *et al.*, 2012). Here, we report the structure of non-activated *Pisum sativum* (garden pea) Rubisco (*Ps*Rubisco) with bound RuBP, which is the first representative structure from a pulse crop.

2. Materials and methods

2.1. Protein purification and crystallization

Rubisco was purified from *P. sativum* plants. Plant homogenates in 50 m*M* Tris–HCl pH 7.2 were strained through gauze, centrifuged at 15 000*g* for 20 min and subjected to 2.5% streptomycin and 30–50% ammonium sulfate (AS) precipitations. The 40% AS fraction was resuspended in 50 m*M* Tris–HCl pH 7.2 and dialyzed overnight into the same buffer. This simple protocol yielded a final dialyzed Rubisco concentration of 25–33 mg ml^{-1} in repeated preparations, with >95% purity as estimated by SDS–PAGE visualization (Fig. 1*a*). This fraction was co-incubated with 20 m*M* RuBP for 10 min on ice before setting up crystallization trials with 2 µl protein–RuBP solution mixed with an equal volume of reservoir solution consisting of 10–12% polyethylene glycol 6000 in 100 m*M* HEPES pH 7.0. The mixed solution was then equilibrated against 750 µl reservoir solution at room temperature. Diffraction-quality crystals were obtained by the hanging-drop vapour-diffusion method. Crystals of approximately 50 × 40 × 20 µm formed within 24–48 h and were harvested and stored in liquid nitrogen with no additional cryopreservative (Fig. 1*b*).

2.2. Crystal diffraction and structure elucidation

Diffraction data were obtained from cryopreserved crystals on the 08ID-1 beamline at the Canadian Light Source, Canadian Macromolecular Crystallography Facility. The data were processed using *MOSFLM* and *SCALA* within the *CCP*4 software package (Battye *et al.*, 2011; Leslie & Powell, 2007; Winn *et al.*, 2011; Potterton *et al.*, 2003). The space group was $P2_122_1$; the unit-cell parameters and processing statistics are given in Table 1. The structure of *P. sativum* Rubisco was solved by molecular replacement using a single LS unit complex (*A* and *I* subunits) from the structure of spinach Rubisco (PDB entry 8ruc; Andersson, 1996) using *Phaser* through *AutoMR* in the *PHENIX* software package (Adams *et al.*, 2010; McCoy *et al.*, 2007) to yield an asymmetric unit containing one L_4S_4 unit. The protein structure was refined using a combination of *REFMAC* (Murshudov *et al.*, 2011) and *BUSTER* (Bricogne *et al.*, 2009) and manual modelling using the molecular-graphics program *Coot* (Emsley *et al.*, 2010). Assessment of data quality with *phenix.xtriage* (Adams *et al.*, 2010) indicated the possibility of pseudo-merohedral twinning. This twinning, combined with the coincidence of the *a* and *b* unit-cell parameters being virtually identical (Table 1), prompted the use of the less common axis choice in the space group. Twinning was further addressed during refinement using the amplitude-based twinning function in *REFMAC* in the final rounds, which reported twin fractions of 0.57 and 0.43 (twin operator k, h, $-l$). Noncrystallographic symmetry (NCS) restraints were implemented during refinement in *REFMAC* using the 'automatically generated local NCS restraints' with 'medium' restraint option. Water molecules were added automatically with *Coot* and picked manually. Subunits refined in the asymmetric unit were named *A*, *B*, *C* and *D* for the L subunits and *S*, *T*, *U* and *V* for the S subunits based on a consensus evaluation

of available Rubisco structure nomenclature (Supplementary Table S1). Refinement statistics are shown in Table 1. Figures were generated using *PyMOL* (Schrödinger LLC). The atomic coordinates and structure factors have been deposited in the PDB as entry 4hhh.

3. Results and discussion

Refinement of the four LS subunit pairs comprising the asymmetric unit shows they are identical in structure and are related by a local fourfold axis (Fig. 2*a*). The overall quaternary structure of biologically active *Ps*Rubisco is shown in Fig. 2(*b*), highlighting the conservation of the form I Rubisco hexadecameric complex. Closer examination of an L_2S_2 unit shows that the L subunit maintains the expected discrete N- and C-terminal domains (Fig. 3). The C-terminal domain is comprised of an eight-stranded parallel α/β-barrel and includes the active site formed by loops connecting the C-terminal end of β-strands to the subsequent α-helix. The active sites are located at the interface of each L_2 dimer. The N-terminal domain of the L subunit, comprised of a five-stranded mixed β-sheet with helices on one side of the sheet, contributes two loops to the active site. Electron density for the C-terminal and N-terminal regions of the L subunits was weak and thus residues 1–11 and 470–475 were not included in the model. The S subunits are comprised of a four-stranded antiparallel β-sheet flanked by two helices on one side. Overall, the structure is very similar to the spinach, rice and tobacco Rubisco structures (Andersson & Backlund, 2008).

Analysis of the density for the side chains for the L subunit yielded a sequence that is in agreement with the available *Ps*Rubisco L-subunit sequence (NCBI accession No. NP_054944). In the case of the S subunits, several published residues (see accession Nos. P00868, P00869, P07689 and AAA33686) were found to be inconsistent with the electron-density maps, including (expected) residues 47 (Glu or Leu), 91 (Val), 92 (Ala), 93 (Ala) and 96 (Glu or Gln) in all four chains. The side chains that best fit the density in both $F_o - F_c$ OMIT and $2F_o - F_c$ maps in all four determined S-subunit chains are Lys47, Lys91, Lys92, Glu93 and Arg96, yielding an additional unique pea S-subunit sequence. Heterogeneity in the nuclear-encoded S-subunit gene family has been noted previously and may be light-dependent, organ-specific or chloroplast-coordinated (Andersson, 1996; Shibata *et al.*, 1996).

Comparisons of the RuBP-binding conformations in the four *Ps*Rubisco L subunits in the asymmetric unit revealed some degree of variation. Generally, the conformation of RuBP in all chains is similar to that observed in the non-activated RuBP-bound spinach Rubisco structure (PDB entry 1rcx; Taylor & Andersson, 1997*b*), showing contacts between O4 and Glu204 and between the P2 phosphate group and His327 and showing P1–P2 distances ranging from 8.6 and 8.7 Å in chains *B* and *C* to 9.0 and 9.1 Å in chains *A* and *D*, respectively (Fig. 4*a*). P1–P2 distances of 9.1 Å or less (as observed here and in PDB entry 1rcx) have been correlated to a 'closed' active-site conformation (Duff *et al.*, 2000; discussed below). However, the RuBP O3—Lys201 distances of 4.7, 5.7, 3.7 and 4.4 Å in chains *A*, *B*, *C* and *D* of *Ps*Rubisco, respectively, all differ from the 3.23 Å distance observed in every L-subunit chain of PDB entry 1rcx. Similarly, the adoption of variable binding conformations by RuBP in *Ps*Rubisco is highlighted by shifts in the distance between O4 and Glu204 ranging from 2.8 Å in chain *B* to 3.2 Å in chain *A*, 3.4 Å in chain *C* and 4.8 Å in chain *D*, compared with precisely 2.96 Å in every L subunit of the 1rcx structure. The observed variations in the RuBP-binding site of *Ps*Rubisco compared with the remarkable consistency of binding in PDB entry 1rcx are likely to be related to differences in the refinement methods. The simultaneous refinement of RuBP in the model with the protein and the reliance upon automatic selection of NCS restraints during refinement may have led to the observed variability. Speculatively, these subtle variations may be representative of ligand flexibility in the non-activated RuBP-binding site arising from a lack

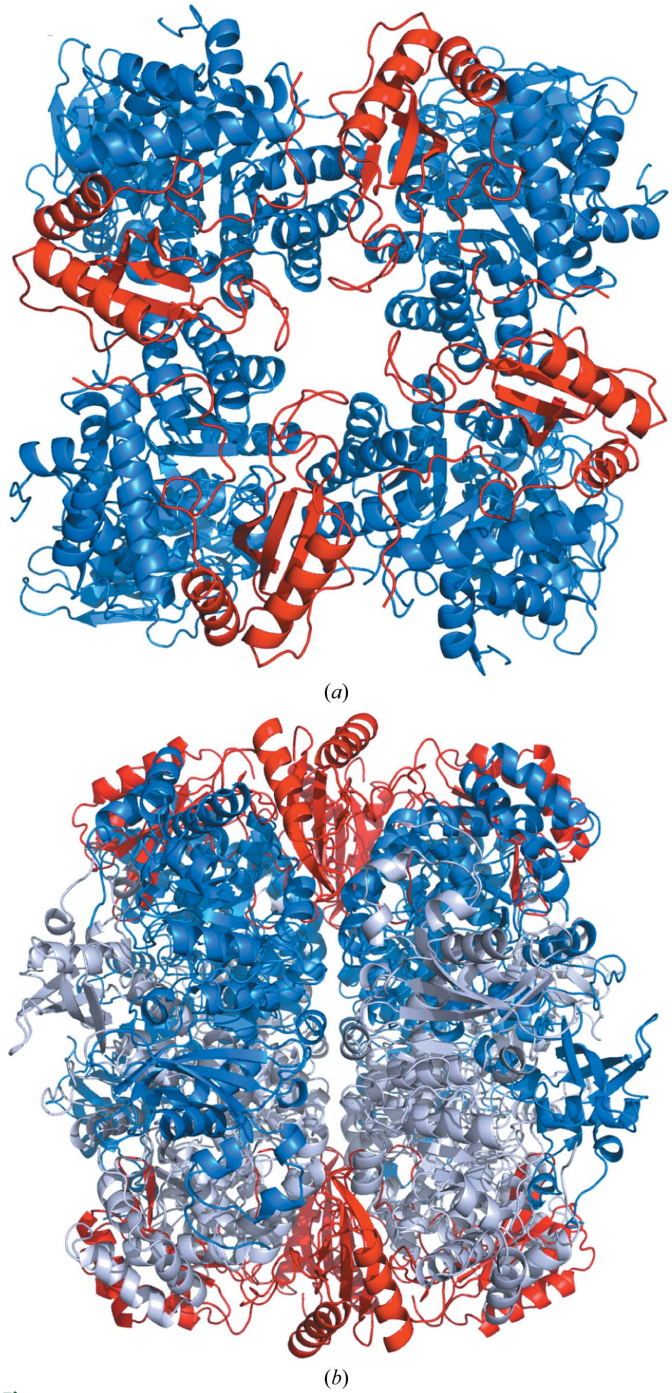


Figure 2
*Ps*Rubisco maintains the most common hexadecameric Rubisco form. (*a*) The refined subunits comprising the L_4S_4 asymmetric unit viewed along the fourfold symmetry axis. L subunits are shown in blue and S subunits are shown in red. (*b*) The L_8S_8 biologically relevant hexadecameric complex viewed along the twofold symmetry axis. The L-subunit symmetry mates are shown in grey, with the asymmetric unit refined L subunits shown in blue and both refined and symmetry-related S subunits shown in red.

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference: TT5035).

of constraint in the absence of a carbamylated Lys201 and the Mg^{2+}. However, the possibility that the variations might arise from crystal-packing defects or errors in estimation of the twin fraction cannot be strictly eliminated.

As alluded to above, spinach Rubisco has been shown to adopt two different active-site conformational states. In the unliganded form or in the presence of the product 3-PGA (*e.g.* PDB entries 1aus and 1aa1; Taylor & Andersson, 1996, 1997*a*) the active site is in an 'open' conformation. In these structures, the C-terminal region (residues 466–475) as well as loop 6 (residues 333–338) are disordered and are not actually represented in the available models. In the presence of ligand, be it RuBP in a non-activated enzyme (PDB entry 1rcx; Taylor & Andersson, 1997*b*) or inhibitor in the activated form (PDB entry 8ruc; Andersson, 1996), the C-terminal strand is stabilized against loop 6, causing loop 6 to fold over the ligand-access channel yielding a stable 'closed' conformation. A third, possibly intermediate, conformation was isolated upon activation of spinach Rubisco with Ca^{2+} instead of Mg^{2+}, yielding an inactive activated enzyme which could be cocrystallized with RuBP (PDB entry 1rxo; Taylor *et al.*, 1996). The

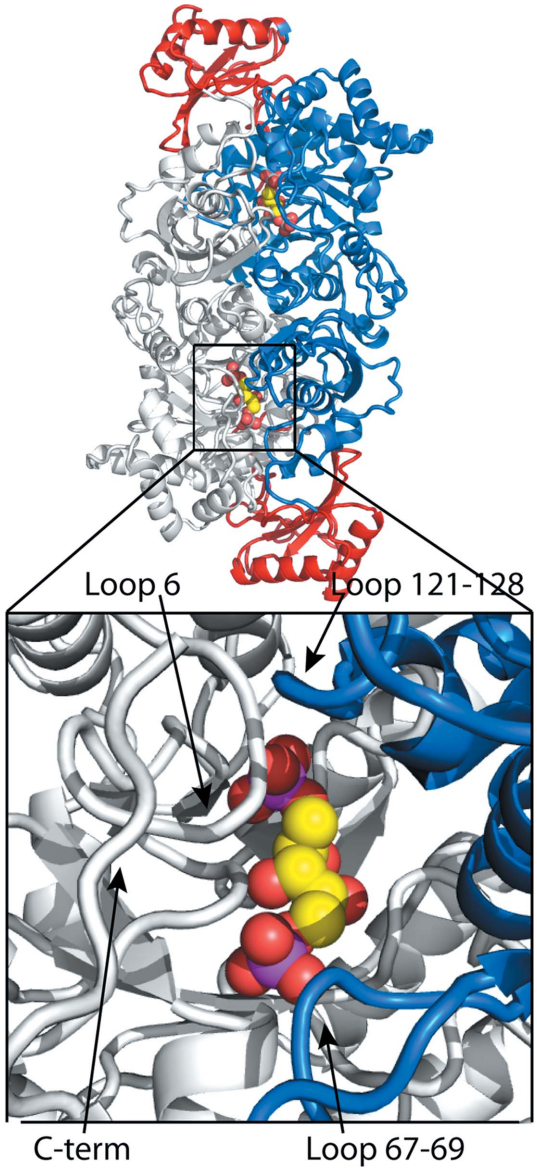


Figure 3
The catalytic site of *Ps*Rubisco is located at the interface of interacting L subunits. An L_2S_2 unit is viewed, showing two active sites with bound RuBP substrate at the interface of the two twofold symmetry-related L subunits. The expanded image shows detail of the loops from both L-subunit symmetry mates comprising the catalytic site. Loops that undergo conformational variation in the 'open' *versus* 'closed' forms of the enzyme are labelled. The L-subunit symmetry mate is shown in grey, with the asymmetric unit refined L subunit shown in blue and both refined and symmetry-related S subunits shown in red. The substrate RuBP is shown with pink phosphate, red O and yellow C atoms.

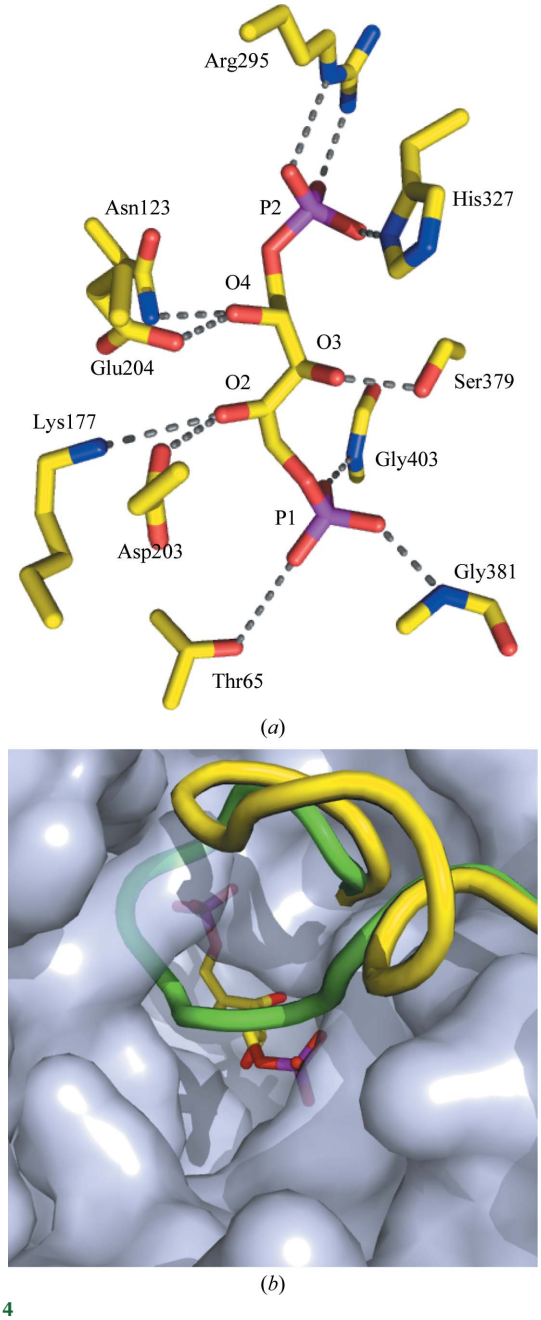


Figure 4
Details of the catalytic site. (*a*) A representation of RuBP in the active site of non-activated *Ps*Rubisco. (*b*) Comparison of non-activated RuBP-bound *Ps*Rubisco with Ca^{2+}-activated inactive RuBP-bound spinach Rubisco. The substrate-binding channel in the *Ps*Rubisco structure is shown as a grey surface. Loop 6 of the *Ps*Rubisco L subunit (green loop) is folded over and pinned down by the overlapping Lys128 residue (grey surface), yielding a 'closed' active-site conformation. This is in contrast to the 'open' conformation observed for spinach Rubisco when activated in the presence of Ca^{2+} with RuBP bound (PDB entry 1rxo), which shows loop 6 (yellow loop) folded back and away from the channel. In both panels, discrete atoms are shown with pink phosphate, red O, yellow C and blue N atoms.

resulting structure showed a stabilized loop 6 but in a relatively 'open' conformation, with Lys128 from an N-terminal loop taking on a compact conformation and loop 6 extending away from the active site. Superposition of the RuBP-bound non-activated *Ps*Rubisco structure and the Ca^{2+}-activated RuBP-bound spinach Rubisco structure (PDB entry 1rxo) shows that the *Ps*Rubisco structure maintains a relatively 'closed' active-site conformation (Fig. 4*b*). Indeed, the *Ps*Rubisco fold is virtually identical to those observed for the RuBP-bound non-activated spinach enzyme (PDB entry 1rcx) and the inhibitor-bound activated spinach enzyme (PDB entry 8ruc). In particular, loop 6 is folded over the active-site channel, with Lys334 extending into the active site and making contacts with the P1 phosphate group of RuBP and with Thr65 and Glu60 in the N-terminus of the adjacent subunit, effectively holding loop 6 down. Lys128 (from the associated dimeric L-subunit partner) takes on an extended conformation, packing over loop 6 and making good contacts with the main-chain carboxyl groups of residues 331 and 333 as well as residue 467 in the C-terminal arm. This latter contact (Lys128–Phe467) stabilizes the C-terminal arm sufficiently for good density to be observed for residues 466–469 of the C-terminal arm. The remainder of the chain (residues 470–475) was too disordered to be evident in the electron-density maps.

In conclusion, the first structure of a pulse-crop Rubisco is presented. The garden pea Rubisco structure is similar in most respects to those of other higher plant Rubiscos, maintaining the common form I quarternary structure and catalytic active site. This non-activated RuBP-bound *Ps*Rubisco structure has a 'closed' active-site conformation consistent with other similar structures. Overall, this work will enable ongoing research in the context of new international crop-improvement consortia.

We are grateful to Dr Stan Moore (University of Saskatchewan) for discussions and Drs Sean Hemmingsen and Jon Page (National Research Council of Canada) for reviewing the manuscript. Research described in this paper was performed using beamline 08ID-1 at the Canadian Light Source, which is supported by the Natural Sciences and Engineering Research Council of Canada, the National Research Council Canada, the Canadian Institutes of Health Research, the Province of Saskatchewan, Western Economic Diversification Canada and the University of Saskatchewan. This work was supported by the National Research Council of Canada (to MCL), the Canada Research Chair Program (to PCL) and the Natural Sciences and Engineering Research Council Discovery Grants Program (to both MCL and PCL). This manuscript represents NRC Communication No. 54669.

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Andersson, I. (1996). *J. Mol. Biol.* **259**, 160–174.
Andersson, I. (2008). *J. Exp. Bot.* **59**, 1555–1568.
Andersson, I. & Backlund, A. (2008). *Plant Physiol. Biochem.* **46**, 275–291.
Andersson, I., Knight, S., Schneider, G., Lindqvist, Y., Lundqvist, T., Brändén, C.-I. & Lorimer, G. H. (1989). *Nature (London)*, **337**, 231–234.
Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* D**67**, 271–281.
Bricogne, G., Blanc, E., Brandl, M., Flensburg, C., Keller, P., Paciorek, W., Roversi, P., Smart, O. S., Vonrhein, C. & Womack, T. O. (2009). *BUSTER* v.2.8.0. Cambridge: Global Phasing Ltd.
Chapman, M. S., Suh, S. W., Cascio, D., Smith, W. W. & Eisenberg, D. (1987). *Nature (London)*, **329**, 354–356.
Duff, A. P., Andrews, T. J. & Curmi, P. M. (2000). *J. Mol. Biol.* **298**, 903–916.
Ellis, R. J. (1979). *Trends Biochem. Sci.* **4**, 241–244.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Knight, S., Andersson, I. & Brändén, C.-I. (1989). *Science*, **244**, 702–705.
Leslie, A. G. W. & Powell, H. R. (2007). *Evolving Methods for Macromolecular Crystallography*, edited by R. J. Read & J. L. Sussman, pp. 41–51. Dordrecht: Springer.
Matsumura, H., Mizohata, E., Ishida, H., Kogami, A., Ueno, T., Makino, A., Inoue, T., Yokota, A., Mae, T. & Kai, Y. (2012). *J. Mol. Biol.* **422**, 75–86.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.
Parry, M. A., Reynolds, M., Salvucci, M. E., Raines, C., Andralojc, P. J., Zhu, X.-G., Price, G. D., Condon, A. G. & Furbank, R. T. (2011). *J. Exp. Bot.* **62**, 453–467.
Potterton, E., Briggs, P., Turkenburg, M. & Dodson, E. (2003). *Acta Cryst.* D**59**, 1131–1137.
Schneider, G., Lindqvist, Y., Brändén, C.-I. & Lorimer, G. (1986). *EMBO J.* **5**, 3409–3415.
Shibata, N., Inoue, T., Fukuhara, K., Nagara, Y., Kitagawa, R., Harada, S., Kasai, N., Uemura, K., Kato, K., Yokota, A. & Kai, Y. (1996). *J. Biol. Chem.* **271**, 26449–26452.
Suh, S. W., Cascio, D., Chapman, M. S. & Eisenberg, D. (1987). *J. Mol. Biol.* **197**, 363–365.
Taylor, T. C. & Andersson, I. (1996). *Nature Struct. Biol.* **3**, 95–101.
Taylor, T. C. & Andersson, I. (1997*a*). *Biochemistry*, **36**, 4041–4046.
Taylor, T. C. & Andersson, I. (1997*b*). *J. Mol. Biol.* **265**, 432–444.
Taylor, T. C., Fothergill, M. D. & Andersson, I. (1996). *J. Biol. Chem.* **271**, 32894–32899.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Zhu, G. & Jensen, R. G. (1991). *Plant Physiol.* **97**, 1354–1358.