The crystal structure of *Escherichia coli* maltodextrin phosphorylase provides an explanation for the activity without control in this basic archetype of a phosphorylase

K.A.Watson^{1,2}, R.Schinzel³, D.Palm³ and L.N.Johnson¹

¹Laboratory of Molecular Biophysics, Rex Richards Building, South Parks Road, Oxford OX1 3QU, UK and ³Theodor-Boveri-Institut für Biowissenschaften (Biozentrum), Physiologische Chemie I, Universität Würzburg, Am Hubland, 97074 Würzburg, Germany

²Corresponding author

In animals, glycogen phosphorylase (GP) exists in an inactive (T state) and an active (R state) equilibrium that can be altered by allosteric effectors or covalent modification. In *Escherichia coli*, the activity of maltodextrin phosphorylase (MalP) is controlled by induction at the level of gene expression, and the enzyme exhibits no regulatory properties. We report the crystal structure of *E.coli* maltodextrin phosphorylase refined to 2.4 Å resolution. The molecule consists of a dimer with 796 amino acids per monomer, with 46% sequence identity to the mammalian enzyme. The overall structure of MalP shows a similar fold to GP and the catalytic sites are highly conserved. However, the relative orientation of the two subunits in *E.coli* MalP is different from both the T and R state GP structures, and there are significant changes at the subunit–subunit interfaces. The sequence changes result in loss of each of the control sites present in rabbit muscle GP. As a result of the changes at the subunit interface, the 280s loop, which in T state GP acts as a gate to control access to the catalytic site, is held in an open conformation in MalP. The open access to the conserved catalytic site provides an explanation for the activity without control in this basic archetype of a phosphorylase.

Keywords: bacteria/maltodextrin/phosphorylase/regulation/X-ray structure

Introduction

Priority in the utilization of carbon sources is finely controlled in microorganisms. In *Escherichia coli*, glucose is the preferred source but, in its absence, other sources can be used, such as maltose and maltodextrins. Utilization of these sources requires relief of catabolite repression of the *mal* genes, which is mediated by the catabolite activator protein (CAP) in response to cAMP. The *E.coli* maltodextrin phosphorylase (MalP) is part of this maltose and maltodextrin transport and utilization system. The maltose regulon contains several genetic regions which encode proteins involved in the transport and metabolism of maltose and maltodextrin (Schwartz, 1987, and references therein). Expression of the *mal* operons requires activation and expression of a transcriptional activator (*malT*) which is controlled directly by the cAMP–CAP system (Chapon, 1982). Recognition of the *malT* product

by maltose or maltodextrin induces expression of the other genes (Raibaud and Schwartz, 1984). *Escherichia coli* also contains a genuine glycogen phosphorylase encoded by a distinct gene, *glgP* (Yu *et al.*, 1988). The glycogen phosphorylase appears to be constitutively expressed by bacteria and is present in low activity, participating in the slow degradation of glycogen during extended periods of substrate deprivation (Preiss *et al.*, 1983). The *E.coli* maltodextrin and glycogen phosphorylases are readily separated during purification and have different substrate specificity.

Crystallographic studies on phosphorylated and unphosphorylated forms of rabbit muscle glycogen phosphorylase, and forms complexed with substrates and effectors, have provided a description of the structural changes undergone by a protein in response to phosphorylation and allosteric effectors (Barford and Johnson, 1989; Barford *et al.*, 1991; Sprang *et al.*, 1991; Johnson, 1992). Activation by phosphorylation is based on an allosteric response whereby covalent attachment of a phosphate group at one site brings about conformational changes at remote sites. In phosphorylase, the interactions are mediated through subunit–subunit interactions that lead to an opening of the catalytic site in the active form and rearrangement of certain amino acids to create a high affinity substrate recognition site. In contrast to the more widely studied mammalian glycogen phosphorylase (GP), *E.coli* MalP exhibits no regulatory properties. It is active in the absence of AMP and is not controlled by phosphorylation/dephosphorylation. The enzyme is regulated at the level of gene transcription in response to the availability of maltodextrins and, once synthesized, requires no post-translational modification or allosteric effectors for activity.

The bacterial and mammalian enzymes also exhibit differences in substrate affinity that reflect their biological roles. The *E.coli* MalP has a high affinity for linear oligosaccharides and <1% activity against glycogen (Schwartz and Hofnung, 1967), while the rabbit muscle enzyme has a poor affinity for linear oligosaccharides and a high affinity for glycogen (Hu and Gold, 1975). Nevertheless, the bacterial and mammalian enzymes share similar enzymatic properties. Both catalyse the phosphorylation of the α -1–4 glucosyl link between glucose residues at the non-reducing end of the glucosyl chain and utilize the 5'-phosphate group of the cofactor pyridoxal phosphate (PLP) in catalysis (Palm *et al.*, 1990). The two enzymes exhibit similar pH dependence and rapid equilibrium bi-bi kinetics (Chao *et al.*, 1969) and both require the dimeric state of oligomeric assembly for active enzyme (Palm *et al.*, 1975).

The *E.coli* MalP (796 amino acids) is smaller than the mammalian GP (842 amino acids), and lacks 17 residues at the N-terminus and 13 residues at the C-terminus (Palm

Fig. 1. Structural alignment of the sequences of MalP and GP obtained with STAMP (Russell and Barton, 1992) and ALSCRIPT (Barton, 1993). Boxed, upper case regions correspond to regions of conserved secondary structure as defined by Russell and Barton (1992). The numbering system is for rabbit muscle GP. Secondary structural elements, as defined by DSSP (Kabsch and Sander, 1983), are shown as black cylinders for α helices, grey cylinders for 3_{10} helices and arrows for β strands. The secondary structural elements are numbered according to Acharya *et al.* (1991).

et al., 1985, 1987). There is 46% identity over all aligned residues. The majority of sequence differences occur in those regions that are involved in control (Newgard *et al.*, 1989; Hudson *et al.*, 1993). In particular, the first 80 residues, which in GP contain the phosphorylatable Ser14, exhibit minimal similarity in amino acid sequence. This region corresponds to the first exon in the human muscle phosphorylase (Burke *et al.*, 1987), which has led to the suggestion that the selective control by phosphorylation could have been conferred on an ancestral phosphorylase by splicing of a phosphorylatable peptide (Palm *et al.*, 1985).

In order to elaborate on the differences between a regulatory and non-regulatory phosphorylase, we have determined the crystal structure of *E. coli* MalP, and compared its structure with the active and inactive structures of rabbit muscle GP. Structural studies have shown that to a first approximation the activation of GP either by phosphorylation or by allosteric activation by AMP can be understood in terms of the Monod–Wyman–Changeux theory in which the enzyme exists in two different quaternary and tertiary structural states (Barford and Johnson, 1989; Barford *et al.*, 1991). The less active T state is characteristic of the non-phosphorylated form of the enzyme, GPb, in the absence of activatory ligands, while the more active R state conformation is characterized either by the phosphorylated GPa form of the enzyme in the absence of inhibitors or by GPb activated by AMP. Studies on *E. coli* MalP were initiated in the 1960s with the expectation that they would provide a system for genetic exploration of protein allosteric mechanisms. The studies uncovered instead the elaborate regulation at the

level of the gene (Schwartz, 1987). The non-regulatory *E. coli* phosphorylase has provided a valuable tool with which to study catalysis without the complications of control, and this has led to many important advances (Palm *et al.*, 1990).

Results

Description of overall structure

The structural alignment of sequences of *E. coli* MalP and rabbit muscle GP is similar to that predicted on the basis of sequence alone (Palm *et al.*, 1985), except that the deletion of four residues which was placed after residue 54 now follows residue 66, and the deletion of eight residues which was placed after residue 311 now follows residue 316 (Figure 1). The numbering system for the rabbit muscle GP is used throughout this manuscript. Overall, the secondary structural features of MalP are almost identical to those of GP (Acharya *et al.*, 1991), particularly throughout the internal core. As in the rabbit muscle enzyme, the bacterial enzyme is an α/β protein that exhibits a two domain fold, the N-terminal domain (residues 19–482) and the C-terminal domain (residues 483–829) separated by a catalytic site cleft (Figure 2). There are four deletions (3–8 residues in length) and one insertion (two residues) in MalP compared with GP, and each of these occurs on the surface loops, as predicted (Hudson *et al.*, 1993). The result is a more compact structure for the MalP enzyme. The structure of the individual subunits shows greater similarity to the active R state GP than to the less active T state GP, and there is a correlation between sequence identity and structural

Fig. 2. Schematic views of MalP (**A** and **C**) and GPα (**B** and **D**) viewed down the 2-fold axis of symmetry of the dimer with the N-terminal region towards the viewer (**A** and **B**) and viewed normal to the 2-fold axis (**C** and **D**). The subunits of MalP are coloured cyan and orange; the subunits of GPα are coloured red and blue. The green regions of MalP represent those regions from Figure 1 where no structural similarity was found between MalP and GPα.

Table I. R.m.s. deviations in Cα positions for *E.coli* MalP and T and R state GP^a

	MalP versus R state GPα (Å)	MalP versus T state GPβ (Å)	R state GPα versus T state GPβ (Å)
N-terminal domain residues 19–482	1.37 399 atoms	1.42 388 atoms	0.89 440 atoms
C-terminal domain residues 483–829	0.88 341 atoms	0.96 342 atoms	0.42 355 atoms
Whole molecule residues 19–829	1.19 740 atoms	1.32 731 atoms	0.79 795 atoms

^aCalculations performed using program ‘O’ (Jones *et al.*, 1991).

superposition, as expected (Table I). The more conserved C-terminal domain (54% sequence identity) shows greater similarity than the more variable N-terminal domain (36% sequence identity). In the catalytic site and the vicinity of the essential cofactor, PLP, the root mean square (r.m.s.)

difference is 0.4 Å and there is almost 100% conservation of residues. On the outside of the conserved structural framework, there are substantial differences in conformation between MalP and GP, which mainly affect subunit–subunit contacts (Figure 2).

In MalP, the two subunits related by a non-crystallographic 2-fold rotation axis are identical (r.m.s. difference in C α positions 0.14 Å). Strict non-crystallographic restraints were used in the refinement and the calculations gave no indication that these restraints should be relaxed. Superficially, the dimeric structure of *E.coli* MalP is intermediate between the R and T state GP structures. With respect to the T and R state structures, there is approximately a 4° and 6° rotation, respectively, of one subunit with respect to the other subunit about an axis normal to the 2-fold axis of the dimer which intersects the 2-fold axis at the centre of the dimer interface, compared with a rotation of 10° which is required for the transformation between T and R states of GP. In addition, there is an even greater rotation normal to both this axis and the 2-fold axis which involves approximately a 14° and 8° rotation between the subunits of MalP and R state GPa and T state GPb, respectively, compared with a 6° rotation between the T and R states of GP. The essential features of the subunit–subunit interactions, which are affected by these transformations, are shown in Figure 3A–F. Figure 3A–C represents an enlargement of the cap'– α 2 interface and tower–tower' helix interface viewed normal to the 2-fold axis as shown in Figure 2C and D. Figures 3D–F illustrates these subunit–subunit features viewed normal to Figure 3A–C with the rotation around the 2-fold crystallographic axis. In GP, these regions form the focus for the transmission of allosteric effects. The details of the changes are summarized below.

N-terminal/C-terminal regions

In rabbit muscle GP, the conversion of inactive GPb to active GPa on phosphorylation of Ser14 results in the ordering and repositioning of the N-terminal tail and the displacement and disordering of five C-terminal residues (Barford *et al.*, 1991). These changes affect the subunit–subunit interface in the region of the allosteric effector interface and are communicated by further tertiary and quaternary structural changes to the catalytic site which is >30 Å away. In GPa, the Ser14 phosphate contacts two arginine residues, one Arg69 from the α 2 helix and the other Arg43' from the cap' region of the other subunit (where the prime denotes a residue or region from the other subunit). The hydrophobic and positively charged residues of the N-terminal tail dock against the subunit interface, which results in the tightening at the α 2–cap' interface compared with T state GP (Figure 3B and C).

Sequence changes in MalP greatly alter this region. The first 17 residues are missing in *E.coli* MalP and neither Arg43 or Arg69 are conserved. There are no corresponding interactions that mimic the Ser14 phosphate and N-terminal tail interactions of GP. Residues 18–80 share only seven identities in sequence between mammalian and bacterial phosphorylases, a level of identity which is below the threshold where sequence similarity alone would indicate structural similarity (Sander and Schneider, 1991). Nevertheless, these residues adopt a conformation that has some remarkable similarities yet some major differences to GP (Figure 3A and B). The first five N-terminal residues of *E.coli* MalP occupy a position that is intermediate between those adopted by the N-terminal tails of GPb and GPa. The chain then enters the α 1 helix (Figure 3A), but the superposition of this helix with that of GPa shows

only a short region of similarity (Figure 1). At the end of the α 1 helix there follows the cap region (residues 36–47), and here MalP and GPa take different paths with some differences as great as 3.3 Å for equivalent residues. The position of the cap in MalP would be blocked in GP by the N-terminal tail of the other subunit (Figure 3A and B). There are several complementary changes in sequence and structure in this region that provide different but almost equally close-packed structures. For example, Trp35 in MalP structurally takes the place of His36 in GPa, Tyr38 the place of Leu39 and Met46 the place of Leu35.

After the differences in the cap region, the two chains are structurally similar at the beginning of the α 2 helix and continue in step until residue 65. In GPb, the helix is unwound at this point and there is a structural discontinuity where the helix bends by 30°. In GPa, the helix contracts at this point to a more regular helix, although there is still a bend (Figure 3B and C). These changes are correlated with the shifts encountered on phosphorylation of the N-terminal tail, and the α 2 helix is an important determinant in the inactive to active conversion. In *E.coli* MalP, there is a dramatic change in structure of the α 2 helix in this region. After the deletion of four residues following residue 66, the chain adopts an extended conformation, made necessary by proline residues in positions 71 and 75. At residue 81, the chain enters the first strand of the β sheet and at this point and for the next 100 residues the MalP and GP structures align almost perfectly.

The details of the subunit–subunit interactions for MalP are shown in Table II. At the cap'– α 2– β 7 interface, there are hydrogen bonds between the main chain carbonyl oxygen of cap' residue Arg37' and the main chain nitrogen of His192, between the Arg37' side chain and Glu61, and between Gln50' and Asn193 (Figure 4A). In GPa, there are hydrogen bonds between the main chain carbonyl oxygens of cap residues 39' and 40' and Arg193, between His36' main chain oxygen and the side chain of Arg60 and between Lys41' and Glu195 (Figure 4B). Thus, although there are sequence changes in the mammalian and the bacterial enzymes, both proteins contrive a similar pattern of hydrogen bonds.

In GP, there is a further interaction at the interface between Pro194 and Tyr185' (Figure 4B), an interaction which appears significant for regulation of GPb by the potent inhibitor glucose-6-P and which allows a non-covalent connection between the cap'– α 2– β 7 interface and the tower helices (Johnson *et al.*, 1993; Rath *et al.*, 1996). Residues 183'–186', in the loop region between strands β 6' and β 7' (and, of course, their symmetry-related equivalents), adopt a very different conformation in MalP and GP in a central part of the structure that is otherwise reasonably conserved in sequence and very well conserved in structure (Figures 1 and 4A and B). In GP, residue Gly186' adopts a conformation only accessible to glycine, and the change from a glycine to an asparagine in MalP necessitates a significant change in ϕ, ψ values. In GP, Arg184' is exposed but in MalP it is buried and in van der Waals contact with residues which include Asn160 (not shown in Figure 4A). In GP, residue 160 is an arginine, and the presence of this arginine would presumably deter the burial of Arg184 in GP. The result of these changes is a larger hole on the 2-fold axis between the subunits in MalP compared with GP (Figure 2C and

Fig. 3. Two views normal to the 2-fold axis of symmetry of *E.coli* MalP (A and D), R state GPa (B and E) and T state GPb (C and F) showing the relative orientations of the subunits and the structural elements that comprise the subunit-subunit interface. To the right is the cap'- α 2- β 7 interface with residues 18 (MalP), 10 (GPa) and 19 (GPb) to residue 85 (for all three structures). To the left is the tower-tower' helix interface showing residues 241-314. There is a break in the chain between residues 281 and 287 following the tower helix in GPa. The pyridoxal 5'-phosphate (PLP) is shown for reference. (A), (B) and (C) are similar to the views of Figure 2C and D, whereas (D), (E) and (F) are normal to this.

Table II. van der Waal contacts (<4.0 Å) and hydrogen bond interactions (<3.4 Å) for residues involved in subunit–subunit contacts in MalP^a

Residue	vdW contact	H bond contact
Arg33 Gln36 Arg37	Arg33', Glu61' Arg64', Ala65', Arg191' Arg64', Glu61', Arg191'	NH2 Arg37...OE1 Glu61' O Arg37...N His192'
Thr47 Arg49 Gln50 Glu61 Phe252 Leu254 Thr255 Phe257 Asn258	Glu194' Glu194' Asn193', Glu194' Arg33', Arg37', Glu61' Tyr163', Pro179', Lys277', Val278' Phe166', Glu177', Ala178' Wat294' Lys274', Lys277', Val278', Pro281' Pro281', Asn282', His285'	NE2 Gln50...OD1 Asn193' OD1 Asn258...N Asn282' O Asn258...ND2 Asn282' O Asn258...ND1 His285' O Asn258...NE2 His285'
Asp259 Gly260 Phe262 Leu263 Glu266	His285' His285', Ala287' Lys274', Ala287', Leu291' Leu263', Arg264', Gln267', Ile270' Ile270', Lys277', Wat117', Wat182'	 OE2 Glu266...NZ Lys277'

^aWater molecules shown are involved in hydrogen bond contact with residues of their own subunit.

D) and an uncoupling of subunit–subunit interactions that connect the allosteric interface to the tower interface.

The changes in the N-terminal region are related to concomitant changes occurring at the C-terminus. Leading up to the C-terminus, the two structures correspond closely up to residue 826, after which the final three residues of *E.coli* MalP adopt a unique conformation. Rabbit muscle GP has an additional 13 residues at the C-terminus. The space created by the absence of these residues in the *E.coli* structure is partially replaced by its N-terminal residues and the unwound $\alpha 2$ helix (illustrated in Figure 5).

Allosteric site and phosphate recognition

Rabbit muscle GP is activated by AMP which binds at the allosteric site and promotes changes in the subunit–subunit contacts similar to those observed for activation by phosphorylation. The recognition site involves residues from the cap' and the $\alpha 2$ and $\alpha 8$ helices. Comparison of R and T state structures of GPb show that the high affinity AMP binding site in R state GPb is created by the tertiary and quaternary structural changes that involve the cap' and $\alpha 2$ helix, resulting in the interior of the allosteric site becoming more tightly packed. The *E.coli* MalP does not require activation by AMP. Examination of the structure shows that there is no AMP binding site. As a result of the shift in the cap and the unwinding of the $\alpha 2$ helix, the equivalent region to the GP AMP site is open and accessible to solvent (Figure 5A). In the rabbit muscle R state GPb–AMP complex (Barford *et al.*, 1991; Sprang *et al.*, 1991), the adenine is sandwiched between the side chains of Tyr75 from the $\alpha 2$ helix and Asn44' from the cap' region of the other subunit, the ribose makes a hydrogen bond to Asp42' and is in van der Waals contact with Val45' from the cap' and Gln71 from the $\alpha 2$ helix, and the phosphate makes specific interactions with two adjacent arginines from the $\alpha 8$ helix, Arg309 and Arg310 (Figure 5B). The important AMP binding residues of GP are not conserved in MalP. Asp42 is a serine, Asn44 an alanine, Val45 a glutamate and Tyr75 a lysine. Both Glu45

and Lys75 in MalP are exposed to the solvent and their spatial location is very different to the equivalent residues in GP.

Contacts to the phosphate group of AMP in GP provide a strong determinant of specificity. In *E.coli* MalP, both Arg309 and Arg310 are conserved but their side chains point in opposite directions so that there is no phosphate recognition site. However, in MalP, the side chains of Arg310 and Arg316 come close together (Figure 5A). The electron density map indicates either a very strongly bound water molecule or a phosphate dianion bound between the guanidinium groups of these residues, since the crystals were grown in 0.1 M sodium phosphate buffer. This site is not present in GP because the C-terminal end of the $\alpha 2$ helix blocks the site (Figure 5B) and because residue 316 is a phenylalanine. Following residue 316, there is an eight residue deletion in MalP and the chain takes a short surface turn of four residues and then goes into the $\alpha 8b$ helix. The longer loop between helices $\alpha 8$ and $\alpha 8b$ in GP is not well ordered in either the T state or the R state native enzymes. However, in crystals of GPb grown in the presence of a modified cofactor that promotes the R state, this loop is ordered, which allows it to make important contacts to AMP bound at the allosteric site (Sprang *et al.*, 1991). Because of the deletion (residues 317–324), these contacts are not possible in MalP.

Subunit–subunit interface at the tower helices leading to the 280s active site loop

The other major subunit–subunit contact region involves the tower ($\alpha 7$)–tower' ($\alpha 7'$) helices on the opposite side of the structure to the cap'– $\alpha 2$ helix interactions (Figures 2 and 3). In GP, this region undergoes a major conformational change on the T to R transition and provides the means of communication between the subunit–subunit interface and the catalytic site (Barford and Johnson, 1989). In GP, the tower helices themselves (residues 266–277) make sparse contacts within their own subunit but make strong subunit–subunit interactions. The link region (residues

Fig. 4. The subunit-subunit contacts for (A) MalP and (B) R state GPa at the allosteric interface of GP. The view is close to that down the 2-fold axis (as in Figure 2A and B) but rotated so as to show the $\alpha 2$ helix and the $\beta 7$ strand from one subunit and the cap' region and the loop preceding the $\beta 7'$ strand from the other subunit. Only selected side chains are shown for clarity. For further details see text.

252–258) going up to the tower is poorly ordered. The C-terminal end of the tower helix is anchored by main chain contacts. The chain then enters the 280s loop (residues 280–287) followed by the $\alpha 8$ helix, the helix that contains at its C-terminal end the residues forming the phosphate recognition site for AMP. In T state GPb,

Fig. 5. The AMP allosteric effector site in R state GPb and the equivalent region in MalP. (A) MalP and (B) R state GPb complexed with AMP. There is no conservation between MalP and GPb of residues important for binding the adenosine moiety of AMP in GPb. The two arginines which recognize the phosphate group are conserved but their orientations with respect to each other are different in MalP and GPb. The diagram also shows the relative orientations of the N- and C-terminal parts of the polypeptide chains. In MalP, the first five residues at the N-terminus and the unwound $\alpha 2$ helix partly occupy the region taken by the 13 residue extension in GPb.

the 280s loop acts as a gate which blocks access to the catalytic site. On activation either by phosphorylation or by AMP, the 280s loop becomes disordered and this allows access to the catalytic site.

The details of the tower-tower' helix interactions viewed down the 2-fold axis for MalP, GPa and GPb are shown in Figure 6A–C. Despite 40% sequence similarity between GP and MalP over residues 246–289, the *E.coli* MalP chain adopts a very different conformation with

Fig. 6. The tower–tower' helix interactions for (A) MalP, (B) R state GPa and (C) T state GPb viewed down the 2-fold axis. Residues 245–314 are shown for each structure and their symmetry-related counterparts.

Fig. 7. Details of subunit–subunit contacts in MalP. (A) Interactions across the tower helices. There are hydrophobic interactions between Leu263', Ile270 and Leu263 and a hydrogen bond between Glu266' and Lys277. The view is the same as that shown in Figure 6. (B) Interactions between Asn282 from the 280s loop and Asn258' from the $\alpha 6b'$ helix which help locate the 280s loop in a position removed from the catalytic site. This view is rotated 180° in the vertical plane from that shown in Figure 6.

very different subunit contacts. There are shifts as great as 17 Å for equivalent residues of GPa and MalP at the start of the tower. In MalP, residues 254–260 form a short helix ($\alpha 6b$) (Figure 6A). They are well ordered and make substantial contacts both within their own subunit and to the other subunit (Table II). Residues 261–265 form an irregular helix and also make substantial intersubunit contacts that include Leu263 packing to its symmetry counterpart Leu263' (Figure 7A), a contact which partially mimics a contact observed in R state GP where Val266 contacts Val266' but the conformations of GP and MalP are different at this stage. The chain enters the tower helix (residues 267–276) which contacts the other tower helix by an ion pair between Glu266 and Lys277' and by non-polar interactions such as Leu263 contact to Ile270' (Figure 7A). Ironically, superposition of MalP with GPb shows greater similarity in the positions of the tower helices in one subunit than between MalP and GPa

(Figures 3 and 6). However, the details of the interactions are different and arise from quaternary structural changes and sequence changes.

The sequences at the start of the 280s loop from 278 to 284 are identical in MalP and GP (Figure 1). From residue 278 to residue 280 the main chain positions for *E.coli* MalP and both R and T states of rabbit muscle GP are also similar. After residue 280, the 280s loop adopts a very different conformation in all three structures (Figure 6). In a striking difference from GP, the 280s loop in MalP is held well away from the entrance to the catalytic site and makes contacts to the $\alpha 6b$ helix at the top of the tower of the other subunit. There is a ring of hydrogen bonds formed by the side chain ND2 of Asn282 to the main chain oxygen of Asn258' (top of the tower) and the side chain OD1 of Asn258' to the main chain nitrogen of Asn282 (Figure 7B). From residue 286, the chain enters the $\alpha 8$ helix early with an extra turn of helix at the

Fig. 8. Stereo diagram showing details of the interactions at the pyridoxal phosphate site for MalP. Lys680, to which the cofactor is linked via a Schiff base, and the inorganic phosphate ion are visible. Water molecules are shown as spheres.

N-terminus compared with GP (Figure 6). This additional turn of helix in MalP can be rationalized from amino acid changes. For example, in GP, Asn282 hydrogen-bonds to Glu287 across the 280s loop. In MalP, residue 287 is an alanine, the hydrogen bond is not possible and Asn282 hydrogen-bonds instead to the top of the tower. In T state GP, Phe285 is partially shielded by a van der Waals interaction with Tyr613, and the packing of these two aromatic groups gives rise to the nucleoside inhibitor site, a site that binds caffeine and a number of other fused ring compounds. The adjacent phenylalanine, Phe286, stacks against Trp387. In MalP, residue 285 is a histidine and residue 286 is a threonine. The non-conservation of these two surface phenylalanine residues (285 and 286) appears to direct changes that allow the chain to adopt a different conformation and to enter the helix early in MalP. After the $\alpha 8$ helix, the MalP and GP structures superimpose closely, with the exception of the glycogen storage site, until residue 826 at the C-terminus.

The tight localization of the 280s loop in MalP appears the most significant feature that explains why MalP is constitutively active without the need for covalent or allosteric activation. In T state GPb, the loop is localized in a position which results in electrostatic repulsion for the substrate phosphate recognition site and steric blocking of access to the catalytic site for oligosaccharide substrates (but not for monosaccharide substrates). In R state GPa, the loop is not well ordered and its position has not been established definitively. The disordered conformation allows access to the catalytic site and also removes an acidic group (Asp283) and allows its replacement by a basic residue (Arg569) to create the phosphate recognition site for substrate. In MalP, the 280s loop is ordered and the loop is held away from the catalytic site, and there is ready access for oligosaccharide substrate binding. Arg569 is in its active conformation and contributes to the substrate phosphate recognition.

Catalytic site and conserved cofactor binding site

Despite their regulatory differences, all known phosphorylases share similar catalytic properties and a highly conserved PLP site (Newgard *et al.*, 1989). Therefore, it is

no surprise that the catalytic site of rabbit muscle GPa and *E.coli* MalP superimpose precisely. Residues surrounding the PLP are almost 100% conserved, with the exceptions of Tyr90 and Phe681 in GP which are both leucines in MalP (Figure 8). The 5'-phosphate group of the PLP makes hydrogen bonds to the side chain of Lys568 and the main chain nitrogens of Thr676 and Gly677 and six water molecules, exactly as in the rabbit muscle GP. A phosphate ion is located nearby (crystals were grown in 0.1 M phosphate). There is a direct hydrogen bond between the cofactor phosphate (3.1 Å) and the inorganic phosphate ion, and the position of the latter is also stabilized by interactions with the main chain nitrogen of Gly135, the side chains of Arg569 and Lys574 and three water molecules, again exactly as in the activated states of GP in the presence of phosphate or sulfate (Figure 8). The separation of phosphorus atoms of the cofactor phosphate and inorganic phosphate is 4.8 Å. The positions of each of the residues identified in rabbit muscle GP for binding of glucosyl substrates, products and transition state analogues (Johnson *et al.*, 1990; Martin *et al.*, 1990; Mitchell *et al.*, 1996) are almost identical. There is a difference in the position of the 380s loop (residues 377–384) (Barford and Johnson, 1992). In T state GPb, this loop helps close the catalytic site through the interaction between Trp387 and Phe286 and an ion pair between Glu382 and Arg770. In R state GPa, these elements move apart and the shifts are correlated with the movement of Arg569 into the catalytic site. In MalP, the regions move even further apart. Shifts in this region in GP may provide a means of communication between the catalytic site and the glycogen storage site.

'The glycogen storage site'

Mammalian GP exhibits a glycogen storage site through which the enzyme may be attached to its large substrate in the form of glycogen particles *in vivo* (Johnson *et al.*, 1989). The site is on the surface of the molecule some 30 Å from the catalytic site and partly linked to it via the 380s loop. In crystallographic binding studies with maltoheptaose and other maltodextrins and T state GP, the presence of five well-localized subsites at the major

binding site and two further subsites at a minor site have been identified (Johnson *et al.*, 1988, 1990; Goldsmith *et al.*, 1989). In GP, the major contacts involve residues from the α 12 helix, the α 13 helix and the loop between the antiparallel β strands, β 15 and β 16 (Figure 1).

Escherichia coli MalP shows no affinity for glycogen. Comparison of sequences (Figure 1) shows that of the five residues (Tyr404, Asn407, Gln408, Glu433, Lys437) in GP that form the major contacts with the most strongly bound sugars of the oligosaccharide in GP, only two of these residues are conserved in the bacterial enzyme (Asn407 and Lys437). Further, the three residue deletion (residues 434–436) between β 15 and β 16, which occurs on a surface loop, places Asp433 and Lys437 in different conformations so that they no longer contribute to the binding site. Leu425, one of the residues that makes important hydrophobic contacts in GP to the sugar, is a tryptophan in *E.coli* MalP and the large side chain partially blocks the site. Hence, at least three essential components are different in MalP at this glycogen binding site compared with GP: absence of hydrophobic glucose ring packing with Tyr404 because of the Tyr→Asn change in *E.coli* MalP; loss of the hydrogen bonding network to Glu433 and Lys437 arising from conformational changes conferred by a three residue deletion; and steric blocking of the site resulting from the Leu425→Trp change.

The minor glycogen binding site in GP is situated above the major site and makes contacts to residues at the top of α 12 and to the top loop of the antiparallel strands β 8 and β 9. The changes in sequence of contact residues are not so dramatic at this site, but binding at this site will no doubt be affected by the deletion at the turn between the two β strands in the top loop of residues 209–211 and the change of Arg358 to a glutamine (Figure 1).

The fact that there is poor conservation of the glycogen binding site in the bacterial enzyme provides a partial explanation for the preference for maltodextrins over large, branched glycogen molecules. It has been postulated that the glycogen storage site in mammalian phosphorylase aids in the positioning of the oligosaccharide into the active site crevice, and it is the absence of this correct position through a non-functional glycogen recognition site which may account for the low activity of glycogen in *E.coli* MalP (Kasvinsky *et al.*, 1978).

Discussion

The crystal structure of *E.coli* MalP has shown a compact phosphorylase fold which exhibits none of the allosteric control sites recognized in the mammalian GP. There are major changes at the subunit–subunit interface of the dimer which, in the region of the cap'– α 2 interface, result in stabilizing interactions of an order of magnitude similar to those observed in GP despite significant changes in sequence in this region. However, the seryl phosphate recognition site and the AMP allosteric effector site of GP do not exist in MalP because of deletions of the first 17 residues, the unwinding of the α 2 helix and sequence changes. Despite very little sequence conservation in the first 80 residues, these residues nevertheless adopt similar folds in GP and MalP, indicating that if this subdomain had been added through gene splicing, as previously suggested, there is a need to occupy this space of the

structure with roughly similar topologies. The larger hole on the 2-fold axis of symmetry between the two subunits, caused by sequence changes, results in loosening of the contacts between the cap'– α 2 interface and the tower subunit–subunit interface.

In contrast to GP, the loop to the tower is ordered in MalP. The tower helices are more similar to the orientation observed in T state GPb than in R state GPa. However, the 280s loop, which in T state GPb forms a gate that blocks access to the catalytic site, is held open in MalP through contacts to residues in the new α 6b' helix in the region going up to the tower of the other subunit. The catalytic site residues and the contacts to the PLP are identical to those observed for the active GPa and activated GPb structures. In particular, Arg569 is in the catalytic site pocket and makes an ionic interaction with a phosphate group bound at the catalytic site.

The more open access of the catalytic site to the 5'-phosphate group of the cofactor is consistent with NMR experiments (Palm *et al.*, 1979) which have shown that the 5'-phosphate in MalP gives rise to a single resonance characteristic of a rapid equilibrium between mono and dianionic forms of the phosphate, in contrast to the 5'-phosphate group in GP which exhibits two distinct peaks. Moreover, in MalP, the 5'-phosphate group can be titrated with pH and has a pK of 5.6 for the substrate-free form of the enzyme. The cofactor phosphate group in rabbit muscle GP shows no titration over the pH range 5.5–7.5, consistent with its restricted access to bulk solvent by the 280s loop. A series of site-directed mutagenesis experiments in *E.coli* MalP (Schinzel and Palm, 1990; Schinzel, 1991; Schinzel and Druckes, 1991; Schinzel *et al.*, 1992) have identified the roles and contributions of some key residues to catalysis and substrate recognition, where the choice of residue for modification was based on the rabbit muscle GP structure. The MalP structure now confirms that from structural superposition these residues do indeed appear to play identical roles in the two enzymes, although the details of their contributions to glucosyl substrate recognition need to be elucidated by X-ray studies with complexes of MalP with substrates. The residues identified include the key cluster of basic residues Lys568, which interacts with the 5'-phosphate of the PLP, and Arg569 and Lys574, which interact with the substrate phosphate.

The bacterial and mammalian phosphorylases represent the extremes of a non-regulatory and a regulatory phosphorylase, respectively. Yeast phosphorylase is an intermediate which is dependent on phosphorylation for activity. The non-phosphorylated form is inhibited by glucose-6-P and shows no activation by AMP. Structural studies of yeast phosphorylase (Rath and Fletterick, 1994; Lin *et al.*, 1996; Rath *et al.*, 1996) have shown that the mechanism for activation by phosphorylation is different from rabbit muscle GP but operates within the same structural framework. The yeast phosphorylase contains an N-terminal 39 amino acid extension compared with GP, with the phosphorylatable threonine at position –10. In the inactive glucose-6-P-inhibited form of the enzyme, the long N-terminal tail wraps around the subunit covering part of the region occupied by the N-terminal tail in GPa and then partially occludes access to the catalytic site. On phosphorylation, the N-terminal tail is removed from the

catalytic site and the phosphothreonine displaces the glucose-6-P and binds with the phospho group occupying the glucose-6-P phospho recognition site that is comprised of arginine residues 309 and 310. Thus, the interactions of the phosphoamino acid are quite different from those in GP_a but they serve a similar role in strengthening the subunit interface by interactions with the cap' residues and the N-terminal tail. As in MalP, the 280s loop gate to the catalytic site is held open but in a different conformation and with different interactions. In evolutionary terms, it is as if control by phosphorylation in mammalian and yeast phosphorylases has evolved twice, resulting in different solutions to the same evolutionary challenge.

Despite very different biological control mechanisms between mammalian and *E. coli* phosphorylases, the initiating events for activation or expression have a common factor. Both mechanisms involve cAMP, which in the mammalian system acts to stimulate cAMP-dependent protein kinase for phosphorylase kinase activation and in the bacterial system to relieve catabolite repression through interaction with CAP. The cAMP regulatory binding subunit of cAMP-dependent protein kinase is homologous in sequence and structure to CAP, indicating a common evolutionary origin in response to cAMP (Weber *et al.*, 1982; Su *et al.*, 1995). In both cases, cAMP is synthesized in response to external stimuli, hormonal stimulation for the mammalian system and nutrient deprivation for the bacteria.

Escherichia coli MalP appears to be a minimal phosphorylase which contains within the large chain all the necessary constellation of atoms to promote phosphorylysis of maltodextrins but which lacks each of the regulatory sites of mammalian GP. The changes that have occurred in the separate evolutionary developments of these enzymes allow rationalization of their different properties in terms of their three-dimensional structure. Some of these differences were possible to predict from comparison of sequences (Palm *et al.*, 1985), but the subunit interface changes that lock open the catalytic site were unexpected.

Materials and methods

Expression and purification of *E. coli* maltodextrin phosphorylase

MalP was purified from *E. coli* Δ malA518 harbouring the expression plasmid pMAP101 as described previously (Schinzel and Palm, 1990; Schinzel *et al.*, 1992). The enzyme was pure to ~98% as judged from SDS-PAGE. Further purification was achieved by a gel filtration step. The concentrated phosphorylase solution was applied on a Superdex 200 column (Pharmacia LKB, Uppsala) equilibrated with 0.1 M NaCl in buffer B. The phosphorylase eluted first from the column. The purified enzyme was stored in an 80% ammonium sulfate solution at 4°C in buffer B. If necessary, chromatography on an FPLC Mono-Q anion exchange column (Pharmacia LKB, Uppsala) was introduced as an additional purification step. The enzyme was eluted from this column equilibrated with buffer B by a linear gradient (0–0.3 M NaCl in buffer B). Typically, 10 g of bacteria (wet weight) yielded ~60 mg of pure enzyme (>98%). Protein concentration was measured by the method of Bradford (1976) or from the absorbance at 280 nm, using $E_{1\text{ cm}}^{0.1\%} = 1.36$.

Crystallization

Crystals of *E. coli* MalP were first obtained from 1.8–1.9 M ammonium sulfate in 0.1 M sodium phosphate buffer pH 6.4 (Buehner and Bender, 1978). Precession photography (to 5 Å resolution) showed the crystals

to be in an orthorhombic cell with space group C222₁ and cell dimensions $a = 104.8$ Å, $b = 191.5$ Å and $c = 297$ Å. The asymmetric unit for this cell contained one dimer (based on a mol. wt of 164 kDa). The crystals obtained by this method were also shown to exhibit disorder beyond ~6 Å in one direction. It was necessary, therefore, to find a more suitable set of crystallization conditions that would yield crystals of higher diffraction quality and which showed no signs of disorder.

The best crystals of *E. coli* MalP were grown in 20–22% 4K PEG, 0.1–0.3 M NaCl, in sodium phosphate buffer at pH 6.4 using a Z/3 plate (Luft *et al.*, 1994) in an effort to control the rate of crystal growth and produce more highly ordered crystals. These crystals are orthorhombic, space group P2₁2₁2₁ with unit cell dimensions $a = 110.2$ Å, $b = 112.3$ Å and $c = 151.0$ Å with one dimer per asymmetric unit. These were the first crystals to diffract beyond 3.0 Å and show no signs of disorder. Due to the length of time (~2 months) to grow crystals in the Z/3 plate, duplicate conditions were set up in the smaller Linbro plates. Coincidentally, a further column (Mono-Q anion exchange as described above) was added to the purification steps during the enzyme preparation since there was evidence of trace impurities still present at the end of the procedure in some of the enzyme samples. The result was reproducible crystals of a smaller size (0.3×0.3×0.8 mm) but exhibiting the same space group and cell dimensions as those grown in the Z/3 plate. These crystals typically appeared within 10 days, and showed diffraction to 2.4 Å resolution using cryo-cooling techniques (25% glycerol as cryoprotectant, T = 100 K) at the SRS Daresbury on station PX9.6. The crystals are also orthorhombic space group P2₁2₁2₁ with unit cell dimensions $a = 109.2$ Å, $b = 110.7$ Å and $c = 150.2$ Å, and one dimer per asymmetric unit.

Data collection and processing

Data to 3.3 Å were collected in-house using the oscillation method, on an 18 cm MAR image plate mounted on a Rigaku RU200 rotating anode source operating at 60 kV and 70 mA, from a single crystal grown in the Z/3 plate. The crystal diffracted to 3 Å at the beginning of the data collection but, due to radiation damage arising from the long exposure times necessary, the final data set used was to 3.3 Å resolution. The data were integrated using the program DENZO (Otwinowski, 1993), and reflections subsequently were scaled and merged using the program SCALEPACK (Otwinowski, 1993), giving a final R_{merge} of 12.4% for 35 207 unique reflections (Table III).

A subsequent data set was collected using cryo-cooling to 100 K at the SRS Daresbury Laboratory station PX9.6 ($\lambda = 0.87$ Å) on a 30 cm Mar Research imaging plate, with a crystal–detector distance of 230 mm giving 2.4 Å at the edge. The crystal was soaked in cryo-protectant (25% glycerol added to the mother liquor) for ~2–3 min prior to mounting using the method of Teng (1990). The images were integrated, and the reflections scaled and merged as before using the programs DENZO and SCALEPACK, giving a final R_{merge} of 8.4% for 69 375 unique reflections (Table III). Structure factors were obtained from intensities using the CCP4 program TRUNCATE (CCP4, 1994).

Phasing

One monomer of the R state rabbit muscle GP_b 2.9 Å refined structure (Barford and Johnson, 1989) was used as the search model in a molecular replacement strategy. Regions of low sequence identity and those regions corresponding to insertions or deletions were excluded from the search model. Using a sequence alignment between rabbit muscle, yeast and *E. coli* enzymes, the following residues were excluded: 206–217 (β-hairpin turn), 309–327 (Type I turn and αN distortion in the region of the AMP binding site), 431–439 (β-hairpin turn in the region of the glycogen storage site) corresponding to areas where deletions occur in the *E. coli* sequence, and residues 766–771 (Type I turn) representing a two residue insertion. Residues 826–842 of the C-terminal tail were also excluded since the *E. coli* sequence only extends as far as residue 829 and the conserved tryptophan at position 825 was conveniently kept. Residues 10–79 which show low sequence identity between the two species were also excluded. In summary, the model used for the molecular replacement search contained the following residues: 80–205, 218–308, 328–430, 440–765 and 772–825.

The molecular replacement search based on one monomer was performed using the program AMoRe (Navaza, 1994), and the 3.3 Å data collected in-house using all the data between 10.0 and 4.0 Å resolution. The subroutine ROTING was used to calculate the cross-rotation function with an integration radius of 25 Å. The first run of the rotation search correctly located a monomer solution with a correlation coefficient Cc of 19.9% and an R-factor of 52.4%. A second search, using the first monomer solution as input, revealed the position of the

Table III. Data collection and refinement statistics

	In-house (273 K)	Daresbury PX9.6 (100 K)
Data collection		
Resolution range (Å)	30.0–3.3	30.0–2.4
No. of crystals used	1	1
No. of observations	123 939	253 500
No. of unique reflections	35 207	69 375
Completeness of all data (%)	91.5	96.2
Completeness of the data in the final resolution shell (3.41–3.30 Å) and (2.51–2.40 Å), respectively	86.5	90.8
R_{merge} (%) ^a	12.4	8.4
Final refinement parameters		
Resolution range (Å)		6.0–2.4
No. of protein atoms		6369
No. of solvent atoms		316
No. of cofactor atoms		20
R_{conv} (%) ^b		23.3
R_{free} (%) ^b		29.2
R.m.s. deviation from ideal geometry for the protein		
Bond lengths (Å)		0.007
Bond angles (°)		1.3
Planar groups (°)		1.3
Average B -factors (Å ²)		
Main chain		27.8
Side chains		31.0
Solvent		39.8

^a $R_{\text{merge}} = \sum_h \sum_i |I_{h,i} - I_h| / \sum_h \sum_i I_{h,i}$, where $I_{h,i}$ and I_h are the i th measurement and mean measurement of reflection h , respectively, and the sum is over all reflections for which more than one measurement is recorded.

^b R_{conv} , $R_{\text{free}} = \sum_h ||F_{\text{obs}}| - |F_{\text{calc}}|| / \sum_h |F_{\text{obs}}|$, where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes respectively. R_{conv} is calculated for 95% of the available data, and R_{free} is calculated for 5% of the data selected at random before beginning structure refinement.

Fig. 9. Stereo diagram of the electron density from the final $2F_o - F_c$ map contoured at 2σ in the vicinity of the pyridoxal phosphate. The view is similar to that shown in Figure 8.

other monomer with a Cc of 36.8% and an R -factor of 47.3%. The difference in the correlation coefficients at this stage was significant. The next highest solution had a Cc of 11.5% and an R -factor of 54.7%. The correctness of the two solutions was confirmed by examining graphically the crystal packing between the two monomers. This revealed that the two monomers formed a dimer in the asymmetric unit.

The dimer solution from the translation search was then used as input to FITING, where 60 cycles of rigid body refinement further brought the overall R -factor to 46.6% and increased the correlation coefficient to 37.8%. At this stage, the electron density map was of sufficient quality to allow unambiguous assignment of many secondary structural features and clear placement of some side chains. The resulting phases

from this procedure were improved using density modification in the program DM (Cowtan, 1994), using the solvent flattening, histogram matching and 2-fold averaging options, in a resolution-based extension scheme from 6.0 to 3.3 Å for the in-house data, and then from 4.5 to 2.4 Å resolution for the data collected at 100 K.

Model building and refinement

Using the in-house 3.3 Å data and real space averaging RAVE software (Kleywegt and Jones, 1994), a 2-fold averaged map was calculated to which 80% of the model could easily be fit. This corresponded to the core of the structure, with the remaining 20% consisting of loops and the N-terminus. Small insertions and deletions were accounted for, but

the density in the regions of large differences were much more difficult to locate with accuracy. The final model building (including all 796 residues and 316 waters) was achieved using the 2.4 Å data collected at 100 K.

The refinement procedure (using only data collected at 100 K) was carried out using alternate cycles of simulated annealing in which strict non-crystallographic symmetry was imposed using X-PLOR (Brünger, 1992), and manual refitting using the program O (Jones *et al.*, 1991) which ultimately gave a complete model (residues 1–796) with an *R*-factor = 27% and *R*_{free} = 32%. At this point in the refinement, water molecules were added to the model using the program ARP (Lamzin and Wilson, 1993). The final model (including 316 waters) was subjected to simulated annealing (with harmonic restraints on the water molecules and maintaining strict non-crystallographic symmetry) and energy minimization, followed by restrained individual *B*-factor refinement using X-PLOR. In the X-PLOR refinement, a 2σ cutoff was applied to the data. The final *R*-factor is 23.3% with *R*_{free} = 29.2% (Table III).

A section of the final 2*F*_o–*F*_c electron density map is presented in Figure 9, showing the region where the essential cofactor PLP is bound. Also seen in this region is the bound phosphate (since the crystals were obtained in the presence of phosphate).

The refined model

The program PROCHECK (Morris *et al.*, 1992) was used to assess the geometry of the final structure. This showed that the main chain dihedral angles for the majority of the residues lie in the most favourable regions of the Ramachandran plot, with only one amino acid (Lys423) found in an energetically unfavourable geometry. The overall 'G-factor' for the final structure was shown to be three standard deviations better than the mean for structures of the same nominal resolution. Comparisons with GP were performed with program 'O' using the coordinate sets from the Protein Data Bank with codes; 1GPA for R state GP, 1GPB for T state GPb and 7GPB for R state GPb with AMP bound.

The coordinates have been submitted to the Brookhaven Protein Data Bank.

Acknowledgements

The staff at the SRS Daresbury Laboratory (station 9.6) provided excellent facilities for data collection. The authors would like to thank K.Harlos (Oxford) and E.Garman (Oxford) for assistance during in-house and synchrotron data collection, respectively, M.Noble for his program XOBJECTS and support producing the figures, and R.Copley and G.Barton for their assistance running STAMP and ALSCRIPT. We also thank members of the Johnson laboratory for helpful discussions throughout the course of this work. This work is supported by the Medical Research Council, the Deutsche Forschungsgemeinschaft Pa92/23-1 and an EC grant B102-CT943025.

References

Acharya,K.R., Stuart,D.I., Varvill,K.M. and Johnson,L.N. (1991) *Glycogen Phosphorylase: Description of the Protein Structure*. World Scientific, London and Singapore.

Barford,D. and Johnson,L.N. (1989) The allosteric transition of glycogen phosphorylase. *Nature*, **340**, 609–614.

Barford,D. and Johnson,L.N. (1992) The molecular mechanism for the tetrameric association of glycogen phosphorylase promoted by protein phosphorylation. *Protein Sci.*, **1**, 472–493.

Barford,D., Hu,S.-H. and Johnson,L.N. (1991) The structural mechanism for glycogen phosphorylase control by phosphorylation and by AMP. *J. Mol. Biol.*, **218**, 233–260.

Barton,G.J. (1993) ALSCRIPT: a tool to format multiple sequence alignments. *Protein Engng.*, **6**, 37–40.

Bradford,M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.*, **72**, 248–254.

Brünger,A.T. (1992) *X-PLOR Manual. Version 3.1*. Yale University, New Haven, CT.

Buehner,M. and Bender,C. (1978) Crystallisation and crystallographic data of *Escherichia coli* maltodextrin phosphorylase. *FEBS Lett.*, **85**, 91–94.

Burke,J., Hwang,P., Anderson,L., Lebo,R., Gorin,G. and Fletterick,R.J. (1987) Intron/exon structure of the human gene for the muscle isozyme of glycogen phosphorylase. *Proteins*, **2**, 177–187.

CCP4 (Collaborative Computational Project, Number 4) (1994) The

CCP4 suite: programs for protein crystallography. *Acta Crystallogr.*, **D50**, 760–763.

Chao,J., Johnson,G.F. and Graves,D.J. (1969) The kinetic mechanism of maltodextrin phosphorylase. *Biochemistry*, **8**, 1459–1466.

Chapon,C. (1982) Role of the catabolite activator protein in the maltose regulon of *Escherichia coli*. *J. Bacteriol.*, **150**, 722–729.

Cowan,K. (1994) 'dm': an automated procedure for phase improvement by density modification. *Joint CCP4 ESF-EACBM Newslett.*, **31**, 34–38.

Goldsmith,E.J., Sprang,S.R. and Fletterick,R.J. (1989) Alternative binding modes for maltopentaose in the activation site of glycogen phosphorylase a. *Trans. ACA*, **25**, 87–104.

Hu,H.-Y. and Gold,A.M. (1975) Kinetics of glycogen phosphorylase a with a series of semisynthetic, branched saccharides: a model for binding of polysaccharide substrates. *Biochemistry*, **14**, 2224–2230.

Hudson,J.W., Golding,G.B. and Crerar,M.M. (1993) Evolution of allosteric control in glycogen phosphorylase. *J. Mol. Biol.*, **234**, 700–721.

Johnson,L.N. (1992) Glycogen phosphorylase: control by phosphorylation and allosteric effectors. *FASEB J.*, **6**, 2274–2282.

Johnson,L.N., Cheetham,J., McLaughlin,P.J., Acharya,K.R., Barford,D. and Phillips,D.C. (1988) Protein-oligosaccharide interactions: lysozyme, phosphorylase, amylases. *Curr. Top. Microbiol. Immunol.*, **139**, 81–134.

Johnson,L.N., Hajdu,J., Acharya,K.R., Stuart,D.I., McLaughlin,P.J., Oikonomakos,N.G. and Barford,D. (1989) Glycogen phosphorylase b. In Herve,G. (ed.), *Allosteric Enzymes*. CRC Press, Boca Raton, FL, pp. 81–127.

Johnson,L.N., Acharya,K.R., Jordan,M.D. and McLaughlin,P.J. (1990) The refined crystal structure of the phosphorylase-heptulose 2-phosphate-oligosaccharide-AMP complex. *J. Mol. Biol.*, **211**, 645–661.

Johnson,L.N., Snape,P., Martin,J.L., Acharya,K.R., Barford,D. and Oikonomakos,N.G. (1993) Crystallographic binding studies on the allosteric inhibitor glucose-6-phosphate to T state glycogen phosphorylase b. *J. Mol. Biol.*, **232**, 253–267.

Jones,T.A., Zou,J.Y., Cowan,S.W. and Kjeldgaard,M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr.*, **A47**, 110–119.

Kabsch,W. and Sander,C. (1983) Dictionary of protein secondary structure. *Biopolymers*, **22**, 2577–2637.

Kasvinsky,P., Madsen,N.B., Fletterick,R.J. and Sygusch,J. (1978) X-ray crystallographic and kinetic studies of oligosaccharide binding to phosphorylase. *J. Biol. Chem.*, **253**, 1290–1296.

Kleywegt,G.J. and Jones,T.A. (1994) Halloween...masks and bones. In Bailey,S., Hubbard,R. and Waller,D. (eds), *Proceedings of the CCP4 Study Weekend*. SERC Daresbury Laboratory, Warrington, UK, pp. 59–66.

Lamzin,V.S. and Wilson,K.S. (1993) Automated refinement of protein models. *Acta Crystallogr.*, **D49**, 129–147.

Lin,K., Rath,V.L., Dai,S.C., Fletterick,R.J. and Hwang,P.K. (1996) Protein phosphorylation control by competition between a phosphorylated side chain and an inhibitory metabolite. *Science*, **273**, 1539–1541.

Luft,J.R., Arakali,S.V., Kirisits,M.J., Kalenik,J., Wawrzak,I., Cody,V., Pangborn,W.A. and DeTitta,G.T. (1994) A macromolecular crystallisation procedure employing diffusion cells of varying depths as reservoirs to tailor the time course of equilibration in hanging- and sitting-drop vapor-diffusion and microdialysis experiments. *J. Appl. Crystallogr.*, **27**, 443–452.

Martin,J.L., Withers,S.G. and Johnson,L.N. (1990) Comparison of the binding of glucose and glucose-1-phosphate derivatives to T state glycogen phosphorylase b. *Biochemistry*, **29**, 10745–10757.

Mitchell,E.P., Withers,S.G., Ermert,P., Vasella,A.T., Garman,E.F., Oikonomakos,N.G. and Johnson,L.N. (1996) Ternary complex crystal structures of glycogen phosphorylase with a transition state analogue nojirimycin tetrazole and phosphate in the T and R states. *Biochemistry*, **35**, 7341–7355.

Morris,A.L., MacArthur,M.W., Hutchinson,E.G. and Thornton,J.M. (1992) Stereochemical quality of protein structure coordinates. *Proteins*, **12**, 354–364.

Navaza,J. (1994) AMoRe: an automated package for molecular replacement. *Acta Crystallogr.*, **A50**, 157–163.

Newgard,C.B., Hwang,P.K. and Fletterick,R.J. (1989) The family of glycogen phosphorylases: structure and function. *CRC Crit. Rev. Biochem. Mol. Biol.*, **24**, 69–99.

- Otwinowski, Z. (1993) Data collection and processing. In Sawyer, L., Isaacs, N. and Bailey, S. (eds), *Proceedings of the CCP4 Study Weekend, DL/SCI/R34*. SERC Daresbury Laboratory, Warrington, UK, pp. 56–62.
- Palm, D., Thanner, F. and Schwenk, L. (1975) Purification and properties of *E. coli* maltodextrin phosphorylase. In Shaltiel, S. (ed.), *Metabolic Interconversion of Enzymes*. Springer-Verlag, New York, pp. 37–43.
- Palm, D., Schachtele, K., Feldman, K. and Helmreich, E.J.M. (1979) Is the active form of pyridoxal phosphate in α -glucan phosphorylases a 5'-phosphate dianion? *FEBS Lett.*, **101**, 403–406.
- Palm, D., Goerl, R. and Burger, K.J. (1985) Evolution of catalytic and regulatory sites in phosphorylase. *Nature*, **313**, 500–502.
- Palm, D., Goerl, R., Weidinger, G., Zeier, R., Fischer, B. and Schinzel, R. (1987) *E. coli* maltodextrin phosphorylase: primary structure and deletion mapping of the C-terminal site. *Z. Naturforsch.*, **42c**, 394–400.
- Palm, D., Klein, H.W., Schinzel, R., Buehner, M. and Helmreich, E.J.M. (1990) The role of pyridoxal-5'-phosphate in glycogen phosphorylase catalysis. *Biochemistry*, **29**, 1099–1107.
- Preiss, J., Yung, S.G. and Baecker, P.A. (1983) Regulation of bacterial glycogen synthesis. *Mol. Cell. Biochem.*, **57**, 61–80.
- Raibaud, O. and Schwartz, M. (1984) Positive control of transcription initiation in bacteria. *Annu. Rev. Genet.*, **18**, 207–231.
- Rath, V. and Fletterick, R.J. (1994) Parallel evolution in two homologues of phosphorylase. *Nature Struct. Biol.*, **1**, 681–690.
- Rath, V., Lin, K., Hwang, P. and Fletterick, R.J. (1996) The evolution of an allosteric site in phosphorylase. *Structure*, **4**, 463–473.
- Russell, R.B. and Barton, G.J. (1992) Multiple protein sequence alignment from tertiary structure comparison: assignment of global and residue confidence levels. *Proteins: Structure, Funct. Genet.*, **14**, 309–323.
- Sander, C. and Schneider, R. (1991) Database of homology-derived protein structures and the structural meaning of sequence alignment. *Proteins*, **9**, 56–68.
- Schinzel, R. (1991) Active site lysine promotes catalytic function of pyridoxal 5'-phosphate in α -glucan phosphorylases. *J. Biol. Chem.*, **266**, 9428–9431.
- Schinzel, R. and Druckes, P. (1991) The phosphate recognition site of *Escherichia coli* maltodextrin phosphorylase. *FEBS Lett.*, **286**, 125–128.
- Schinzel, R. and Palm, D. (1990) *Escherichia coli* maltodextrin phosphorylase: contribution of active site residues Glu637 and Tyr538 to the phosphorolytic cleavage of α -glucans. *Biochemistry*, **29**, 9956–9962.
- Schinzel, R., Palm, D. and Schnackerz, K.D. (1992) Pyridoxal 5'-phosphate as a ^{31}P reporter observing functional changes in the active site of *Escherichia coli* maltodextrin phosphorylase after site directed mutagenesis. *Biochemistry*, **31**, 4128–4133.
- Schwartz, M. (1987) The maltose regulon. In Neidhardt, F.C., Ingraham, J.L., Low, K.B., Magasanik, B., Schaechter, M. and Umberger, H.E. (eds), *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*. American Society for Microbiology, Washington, DC, vol. 2, pp. 1482–1502.
- Schwartz, M. and Hofnung, M. (1967) La maltodextrine phosphorylase d'*Escherichia coli*. *Eur. J. Biochem.*, **2**, 132–145.
- Sprang, S.R., Withers, S.G., Goldsmith, E.J., Fletterick, R.J. and Madsen, N.B. (1991) The structural basis for the association of phosphorylase b with AMP. *Science*, **254**, 1367–1371.
- Su, Y., Dostmann, W.R.G., Herberg, F.W., Durick, K., Xuong, N.-H., Ten Eyck, L., Taylor, S.S. and Varughese, K.I. (1995) Regulatory subunit of protein kinase A: structure of deletion mutant with cAMP binding domains. *Science*, **269**, 807–813.
- Teng, T.-Y. (1990) Mounting of crystals for macromolecular crystallography in a free-standing thin film. *J. Appl. Crystallogr.*, **23**, 387–381.
- Weber, I.T., Takio, K., Titani, K. and Steitz, T.A. (1982) The cAMP-binding domains of the regulatory subunit of cAMP-dependent protein kinase and the catabolite gene activator protein are homologous. *Proc. Natl Acad. Sci. USA*, **79**, 7679–7683.
- Yu, F., Jen, Y., Takeuchi, E., Inouye, M., Nakayama, H., Tagaya, M. and Fukui, T. (1988) α -glucan phosphorylase from *Escherichia coli*: cloning of the gene and characterisation of the protein. *J. Biol. Chem.*, **263**, 13706–13711.

Received on August 13, 1996; revised on September 18, 1996