E. coli Maltodextrin Phosphorylase: Primary Structure and Deletion Mapping of the C-Terminal Site

D. Palm, R. Goerl, G. Weidinger, R. Zeier, B. Fischer, and R. Schinzel

Physiologisch-Chemisches Institut, Universität Würzburg, Koellikerstraße 2, D-8700 Würzburg, Bundesrepublik Deutschland

Z. Naturforsch. 42c, 394–400 (1987); received December 17, 1986

Dedicated to Professor Helmut Simon on the occasion of his 60th birthday

Maltodextrin Phosphorylase, Sequence, Mutagenesis, Evolution, malP Gene

The complete 796 residue amino acid sequence of maltodextrin phosphorylase was deduced from the E. coli malP nucleotide sequence. The calculated molecular weight of 90,500, including pyridoxal phosphate, is significantly larger than experimentally determined values. Enzymatically active and inactive mutants following deletion or exchange of up to 8 codons (7 amino acids) at the 3' end (C-terminus) confirm the size of the mature native enzyme and disclose the essential functional or structural role of the highly conserved C-terminal region of phosphorylases.

Introduction

Phosphorylases (EC 2.4.1.1.) catalyze the release of glucose 1-phosphate from α-glucans, which is the first step in the intracellular degradation of starch or glycogen. All phosphorylases so far investigated have many structural properties in common [1, 2], more over, they share an absolute pyridoxal phosphate requirement which reflects the key role of the cofactor in the intricate catalytic mechanism [3, 4]. On the other hand phosphorylase activity regulation occurs at vastly different levels of structural or cellular complexity in nonrelated organisms [5–8]. From the completely resolved amino acid sequence and X-ray crystallographic analysis of rabbit muscle phosphorylase, it was suggested that an exceedingly large number of amino acid residues are involved in substrate binding and catalysis as well as regulatory responses [2, 9, 10]. The predicted interactions correlate closely with segments of strong homology in E. coli maltodextrin phosphorylase [11] potato α-glucan phosphorylase [12] and yeast glycogen phosphorylase [13]. But partial or complete sequence divergence characterizes those regions of the individual phosphorylases which were attributed to participate in individual responses to polysaccharide binding, allosteric effectors or protein phosphorylation [11–13].

Maltodextrin phosphorylase, which is devoid of allosteric or covalent regulation, has served as a model to assess the essential structural requirements for the catalytic mechanism. 92–100% sequence conservation as compared to rabbit muscle phosphorylase was observed for the cofactor binding site and catalytic site in a stretch of roughly 600 out of 687 amino acids sequenced so far from the bacterial enzyme. This includes all known ligand interactions [10, 11]. When we completed the DNA deduced amino acid sequence of maltodextrin phosphorylase, we noted that more and equally strongly conserved homologous regions were found in the vicinity of the C-terminus where no functional connotation was expected or predicted. The essential role of this subdomain is probed by the sequence-function relationship of a small number of deletion mutants.

Materials and Methods

Escherichia coli pop 750 carries the phage 080dmalA2 [15, 16] and a helper phage. E. coli pop 2158 carries the deletion ΔmalA518 and is of mal− phenotype [14]; this strain was used preferentially for phosphorylase activity assay after transformation with mutagenized plasmids. Plasmid pOM13 (see Fig. 1) was carried in E. coli pop 2152 [14]. All strains were obtained from M. Schwartz.

Phage 080dmalA2 was isolated from pop 750 by lysis with CHCl3 5 h after thermal induction. The DNA was extracted as described in ref. [16]. Constructions of recombinant plasmids and phages, isolation of DNA and use of restriction endonucleases was carried out essentially following published protocols in Maniatis et al. [17], Clover [18], or suggestions of the suppliers. Transformation of E. coli
HB101, pop 2158 and JM103 was performed according to Hanahan [18, 19]. DNA sequencing was carried out by the methods of Sanger et al. [20] and Maxam-Gilbert [21]; short fragments were sequenced by the solid-phase technique [22, 23] using CSpaper (Denta-MED Service, Berlin).

**Maltodextrin phosphorylase assay**

Cells grown in 5 ml LB medium [18] were resuspended in 1 ml 10 mM Tris, 1 mM EDTA, pH 7.0 buffer [25] and sonicated. The assay was performed by previously published procedures [25]. Colony screening for phosphorylase activity was first described by Schwartz [26].

**Immunological techniques**

Homogeneous maltodextrin phosphorylase [25], 1 mg in 0.5 ml saline, was injected intracutaneously into rabbits in complete Freund’s adjuvant. Boosters were given 3–4 weeks later in incomplete Freund’s adjuvant. Competent antisera were pooled and affinity purified on immobilized maltodextrin phosphorylase. Elisa’s were performed according to Engvall [27] and Western blots according to Burnette [28]; anti rabbit IgG-POD from goat was used as a second antibody.

**Results and Discussion**

The N-terminal genomic DNA-fragment from the *E. coli* malP gene encoding maltodextrin phosphorylase was previously isolated [16] and sequenced [11, 29]. In this study we took advantage of the known sequence to search for an overlapping gene fragment in phage 080odmalA2, which carries the complete *malP,Q* genes [15]. These requirements were met by a 2.4 kb *BgIII* fragment which was cloned into pBR322 giving pGRW5 (Fig. 1). From this plasmid a 514 bp *HindIII-BgII* fragment was sequenced by the Maxam-Gilbert method in both directions except for 40 bp at the 3′ end. The 1060 bp *EcoRV-HincII* fragment was subcloned into M13mp9 and sequenced using the dideoxy method of Sanger et al.

A 797-residue amino acid sequence was deduced from the uninterrupted open reading frame (Fig. 2). The Met at position −1 is not found in the mature native enzyme protein [30]. The predicted C-terminal sequence of Lys–Arg agrees with the results of a time-dependent carboxypeptidase b digestion. Cys 775 – Arg 785 accounts for a hitherto unassigned tryptic peptide obtained from [¹⁴C]carboxymethylated phosphorylase [31]. Below we will give further evidence for the existence of the essentially unshortened polypeptide chain in the native enzyme. From the 796 amino acid residues comprising maltodextrin phosphorylase, plus the essential cofactor pyridoxal phosphate, the relative molecular mass (*M*ₐ) 90500 is calculated, which exceeds significantly the experimentally determined value of 81000 [11].

The newly disclosed C-terminal extension of the polypeptide chain came quite unexpected in view of the hitherto known structural and functional properties: considering the location of all known substrate binding sites and the cofactor binding site within the first 700 amino acids of rabbit muscle phosphorylase and potato phosphorylase, the smaller experimental molecular weight of maltodextrin phosphorylase was interpreted in support of evolutionary selection with loss of dispensable termini [11, 12]. This hypothesis is also at variance with the molecular weight of 90000 found for *Klebsiella pneumoniae* 1,4α-glucan phosphorylase [32] which corresponds to that of *E. coli* as determined now.

The cloning of the bacterial phosphorylase gene and its sequence determination let us apply for the first time the potential of site directed mutations to probe for structural and functional correlations. Our first aim was to further characterize the C-terminus of the translated polypeptide chain.

The vector used for these studies was constructed from pOM13, which, in a *Clal* insert, carries the complete *malP,Q* operon (Fig. 1). In pMAP101 the upstream *ClaI* site is deleted by *Bal31* digestion, leaving the *malP,Q* promoter intact (Fig. 1). Transformants express maltodextrin phosphorylase at a slightly higher level than fully induced *E. coli* K12 [25]. To introduce a set of deletion mutations at or close to the predicted 3′ end of the *malP* gene we developed a strategy making use of a single *NsiI* restriction site 130 bp downstream of the stop codon at position 2394 (Fig. 1 and 2). pMAP101 was linearized with *NsiI* and digested with *Bal31*. The downstream *malQ* DNA insert is excised by *ClaI*. Religated recombinants were transformed into *E. coli* HB101 or pop 2158.

Restriction analysis of single colonies showed that 33% of the transformants had lost the *BgII* to *ClaI* restriction site 130 bp downstream of the stop codon.
Fig. 1. Map of the malA operon (16) and malP (enlarged). The position of the restriction sites are given as a distance in base pairs from the translation start of malP; restriction sites in parenthesis refer to the pBR322 sequence. The map further shows the bacterial DNA present in phage 080dmalA2 and plasmids derived from it. The width of the solid black lines symbolizes the origin of the DNA. 080dmalA2 and pOM13 were obtained from M. Schwartz. pGRW5: a BglII fragment was cloned into pBR322. pMAP101 and pMAP101Δn: The arrows indicate the extent of deletions constructed in vitro by restriction and Bal31 (see text).
Fig. 2.
D. Palm et al. · E. coli Maltodextrin Phosphorylase
(position 2434—5300) fragment, leaving intact Sphl, position 2107. These clones and those which in addition have retained the NruI site at position 2357 were further assayed for phosphorylase activity, total protein and immunoreactive protein (Table I). Two fully active and two inactive clones were selected for resequencing the altered site. The EcoRI-NruI restriction fragment was 5’ end labelled and used for solid phase sequencing of small fragments [23, 24] (Fig. 3).

The deletions fall into two classes: I) Deletions located downstream of the presumed malP stop codon and removing the malQ gene. The cell-free supernatant contains equal or slightly higher phosphorylase activity and immunoreactive phosphorylase protein than the parent plasmids pOM13 and pMAP101. This fact indicates a slightly more efficient expression of the functionally intact phosphorylase protein. Enzyme activity was unchanged when the last three triplets of the malP gene were replaced by fusion with pBR322 DNA. We assume that the DNA fusion was translated upstream the first stop codon of the appropriate reading frame, leading to a hexadecapeptide hybride. II) All deletions or vector fusions taking off the C-terminal 7 or more amino acids of phosphorylase showed a mal- phenotype. This was best seen after transformation into pop 2158 which did not change its phenotype despite expressing approximately 3—6 times more immunoreactive phosphorylase protein.

The functional characterization provides conclusive evidence for the correct interpretation of the open reading frame in the malP gene. The messenger must be translated beyond the NruI site which makes the phosphorylase polypeptide chain at least 785 amino acids long. Mutations Δ52 and Δ107 map the minimal size of the fully active protein chain even more closely within 790—794 amino acid residues. In the preceding construction it is assumed that a fused hexadecapeptide chain is not deleterious to the function.

A comparative analysis of the newly sequenced C-terminal chain of maltodextrin phosphorylase from residues 687—796 with the rabbit muscle phosphory-
The observation that a deletion or exchange of only 7 amino acid residue at the C-terminus effects a total loss of activity in a bacterial phosphorylase must be considered a favorable coincidence, since from all cloned phosphorylases available only those of *E. coli* and yeast [13] might have responded to a short deletion. A prediction to the contrary was made by Cerrar [35], despite implicating an important structural or functional role of the C-terminal position of muscle phosphorylases.

Although at present we cannot deduce the steps leading to the loss of activity from the change in the primary structure, it is interesting to speculate about a possible role of the conserved sequence 816*–831*, especially a role of Trp-826*. It appears that these residues close the bottom of the active site assembly as seen from the entrance at the I-site [10]. Actually the entrance encompasses the loop 275*–290* which is exactly as strongly conserved as the site discussed here.

Our present approach of a segment mutation should be variable enough to map this region more precisely. Another way to probe functions, even of inactive mutants under standard conditions, will be making use of newly developed glycosylc substrate

---

*a* Residue numbers labelled with an asterisk refer to the numbering of rabbit muscle phosphorylase [33].
analogs which excel in significantly different primer requirements [4].

Acknowledgements

We thank Maxime Schwartz, Institut Pasteur, Paris, for the original clones and for continuous support of the study. Thanks are due to E. Schiltz, Freiburg, for amino acid analyses, M. Buehner, for sequence comparison and K. J. Burger, for practical and theoretical advice. This work was supported by grant Pa 92/19 from the Deutsche Forschungsgemeinschaft.