Sequence of the Human Erythrocyte Phosphoglycerate Mutase by Microsequencer and Mass Spectrometry*

(Received for publication, April 18, 1988)

Yves Blouquit‡, Marie-Claude Calvin‡, Raymond Rosta§, Marie-Claude Promé§, Françoise Brachet†, Michel Cohen-Solalt†, and Jean Rosta‡

From the 1Institut National de la Santé et de la Recherche Médicale U.91 and Centre National de la Recherche Scientifique UA 607, Hôtel Henri Mondor, 94010 Créteil, France and the 2Centre de Recherche de Biochimie et Génétique Cellulaires du Centre National de la Recherche Scientifique 118, route de Narbonne, 31062 Toulouse Cedex, France

We have previously reported the isolation in pure form of the human erythrocyte phosphoglycerate mutase isozyme B. We now report the sequence of the whole protein and the identification of its N-terminal blocking group. The protein tryptic peptides of phosphoglycerate mutase isozyme B were isolated by high performance liquid chromatography and their sequence determined by microsequencing. The sequence and the nature of the blocking group of the N-terminal tryptic peptide was shown to be N-acetyl-Ala-Ala-Tyr-Lys by mass spectrometry. Overlaps of the tryptic peptides were obtained by studying the V8 Staphylococcus aureus protease peptides of the aminoethylated phosphoglycerate mutase isozyme B either by microsequencing or by mass spectrometry. The procedure used allowed us to obtain the sequence in a very small amount of material and in a short period of time. Our data agree well with those derived from the cDNA nucleotide sequence described by Sakoda et al. (Sakoda, S., Shanske, S., DiMauro, S., and Schon, E. A. (1988) J. Biol. Chem. 263, 16899–16905). In addition, our data directly indicate that the initiation codon does not introduce a methionine as N-terminal amino acid and allowed the identification of the acetyl N-terminal group.

Two isozymes have been described for the glycolytic enzyme phosphoglycerate mutase (EC 2.7.5.3) in mammals (1). The M isozyme is found in muscle, and the B isozyme in brain, erythrocyte, liver, and kidney, although both isozymes coexist in heart tissue (2–4).

The amino acid sequence of the human muscle M isozyme has been published recently by derivation from the nucleotide sequence of the corresponding cDNA (5); in this paper, we describe the amino acid sequence of the B-type isozyme determined from the purified human erythrocyte enzyme and compared with those obtained from the cDNA sequence published in the preceding paper. In addition, the exact nature of the blocking N-terminal residue was obtained by mass spectrometry studies.

*This work was supported by grants from the Institut National pour la Santé et la Recherche Médicale (Unité INSERM U.91), the Centre National de la Recherche Scientifique (UA CNRS 607), the University of Paris XII, and the Fondation pour la Recherche Médicale. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡To whom correspondence should be addressed: Unité INSERM U.91, Hôtel Henri Mondor, 94010 Créteil, France.

MATERIALS AND METHODS

Purification of the Human Erythrocyte Phosphoglycerate Mutase—The human phosphoglycerate mutase isozyme B protein was purified as described previously (6) from 1.250 liters of outdated normal blood from a blood bank (Centre Départemental de Transfusion Sanguine, Créteil) representing the pool of five donors.

Enzyme assays were performed as described previously (7). The protein used in this study was pure according to the following criteria: (i) a single band by sodium dodecyl sulfate gel electrophoresis, (ii) immunoelectrophoresis showing no cross-reaction with 2,5-diphosphoglycerate mutase, and (iii) a specific activity of 652 units/mg.

Enzyme assays were performed as described previously (7). The protein used in this study was pure according to the following criteria: (i) a single band by sodium dodecyl sulfate gel electrophoresis, (ii) immunoelectrophoresis showing no cross-reaction with 2,5-diphosphoglycerate mutase, and (iii) a specific activity of 652 units/mg.

Sequence Determination of Peptides—Amino acid microsequences of all tryptic peptides and of several of the V8 peptides were determined using an Applied Biosystems 470A microsequencer (12). Position of the tryptic peptides was obtained by overlapping sequences of the V8 peptides or by consideration of the mass of several V8 peptides obtained by mass spectrometry.

Mass spectrometry was performed on a HS-ZAB 2F mass spectrometer (VG-analytical, Manchester, United Kingdom) interfaced to a VG-2035 Data system, via a PDP-8A computer (13). The FAB spectra were generated by a neutral xenon atom beam of 8 keV. The calibration curve was made by cesium iodide FAB spectra. The mass scan range and resolution settings of the instrument were adjusted for sampling a cesium iodide peak at least three times. The mass spectrometry spectra were acquired with a CAD-MIKE spectrum of the N-terminal tryptic peptide.

VG-MIKE software was used for multichannel analysis. In this software, the time data is converted into mass data when a linear relationship between time and mass is provided by the spectrometer scan law. The calibration curve was made by cesium iodide FAB specta. The mass scan range and resolution settings of the instrument were adjusted for sampling a cesium iodide peak at least three times. The CAD-MIKE experiments were done using the multichannel recording system at a speed of 800 eV/s.

1The abbreviations used are: HPLC, high performance liquid chromatography; FAB, fast atom bombardment; CAD-MIKE, collision activation dissociation-mass analyzed ion kinetic energy.
FIG. 1. HPLC profile of the tryptic peptides from the human phosphoglycerate mutase B isozyme. The separation was carried out on a Macherey Nagel C18 Nucleosil column (5 μm, inner diameter 4.6 mm, 130 mm) using a linear gradient from 0 to 100% in 120 min made from solvent A (0.1% trifluoroacetic acid in water) and solvent B (0.1% trifluoroacetic acid in 60% acetonitrile). Asterisks represent peptides in their disulfide forms. Sample 100 μg was dissolved in 120 μl of solvent A.

FIG. 2. HPLC profile of the V8 S. aureus protease peptides from the human aminoethylated phosphoglycerate mutase B isozyme. The separation was obtained as indicated in legend of Fig. 1. Sample 250 μg was injected in 200 μl of hydrolysis buffer (0.2 M NRHCO3, 2 M urea).

RESULTS

As described previously (6), the phosphoglycerate mutase isozyme B protein used in this study was purified to homogeneity, as shown by enzymatic staining after nondenaturing electrophoresis and by immunoreaction showing the absence of any contamination with 2,3-diphosphoglycerate mutase, and by polyacrylamide-SDS gel electrophoresis.

Fig. 1 shows the elution pattern of the tryptic digestion of 100 μg of phosphoglycerate mutase isozyme B protein isolated by reverse phase HPLC. Forty-five major peaks were obtained, each containing one or a few peptides which were repurified when necessary. Each pure peptide was sequenced by automatic microsequencing. In some instances the results of the microsequence determination indicated the presence of a homogeneous disulfide peptide (i.e. peptide T7-T7 or T21-T21) or of the mixed disulfide (T7-T21). In these cases the disulfide peptides were treated with 4-vinylpyridine and their sequences determined. Peak number T1 did not yield any amino acid sequence, suggesting that the N-terminal residue of the human phosphoglycerate mutase isozyme B was blocked.

Fig. 2 shows the elution pattern of the V8 digestion of 250 μg of aminoethylated phosphoglycerate mutase isozyme B protein isolated by reverse phase HPLC. The mass of most of the major peptide peaks was determined by mass spectrometry. Aminoethylation was performed on this sample in order to prevent the occurrence of large disulfide peptides whose mass would have been too large to be detected. In several instances the V8 peptides were sequenced by microsequencing. For V8 peptides ending in EE pairs, such as 35-36, 168-169, 201-202, and 235-236, it could not be determined with certainty during microsequencing whether the last E residue identified by HPLC was due to a real EE sequence or to a contamination carried over from the last degradation step during the sequence determination. A similar problem is present with trypsin for the double K or the K,R sequences (residues 38-39, 60-61, 115-116, 139-140, 239-240, and 252-253). For both of these cases, mass spectrometry was the most reliable method to verify the presence of the pairs.

The phosphoglycerate mutase isozyme B N-terminal sequence was determined by mass spectrometry analysis of the T1 tryptic peptide and of the V1 Staphylococcus peptide which were found by microsequencing to be N-terminal-blocked. The mass spectrometry analysis of peptide T1 indicated a mass of 494 (Fig. 3) which was compatible with that of the tryptic peptide T1 Ala-Ala-Tyr-Lys and an acetyl group. The definitive N-terminal sequence of the protein was obtained by CAD-MIKE spectrum analysis of this tryptic peptide T1 of mass 494. Results gave the sequence N-acetyl-Ala-Ala-Tyr-Lys (Figs. 4 and 5). In addition the mass spectrometry analysis of S. aureus peptide V1 indicated a mass of 1410 which is compatible with that of the tryptic peptide T2 (Leu-Val-Leu-Leu-Ile-Arg) plus the N-acetylated T1 peptide and part of peptide

"Human Phosphoglycerate Mutase Isozyme B Protein Sequence"
**Human Phosphoglycerate Mutase Isozyme B Protein Sequence**

**FIG. 3.** Partial FAB spectrum (400 < m/z < 650) of the phosphoglycerate mutase isozyme B tryptic peptide HPLC fraction containing the blocked N-terminal peptide T1.

**FIG. 4.** Collision-activation FAB spectrum of the protonated molecular ion m/z 494 (CAD-MIKE analysis). The fragment ions at m/z 361, 310, and 147 correspond to ammonium-sequence ions bearing the carboxyl end group (Y“, fragments according to the Roepstorff nomenclature) (21). The ions at m/z 348, 185, 114, and 43 are acylium-sequence ions possessing the amino end group (B-fragments). This last series was associated with the [B-CO]+ series (A-fragments) at m/z 320, 157 (weak), and 86. The Z-series, corresponding formerly to [Y“-NH2] fragments, was observed at m/z 435, 364, 293, and 130. The intense ion at m/z 136 could be attributed to a doubled cleavage of the peptide backbone and might correspond to the immonium ion of tyrosine.

**FIG. 5.** Amino acid sequence deduced from the CAD-MIKE spectrum of Fig. 4. The symbols correspond to the nomenclature of Roepstorff (20).

The total sequence of the human erythrocyte phosphoglycerate mutase isozyme B shown in Fig. 6 extends over 253 residues which are distributed over 35 tryptic peptides. Alignment of these tryptic peptides was obtained by the use of several *S. aureus* overlapping peptides whose sequence had been either determined by microsequencing or deduced from the mass measured by FAB mass spectrometry. Sequences of the double tryptic peptides T16-17 and T19-20 allowed the localization of peptides T17 and T20 which consist of a single basic residue. Sequences 38-85 and 117-138 could not be obtained as described above. For this part of the molecule, data were obtained for only tryptic peptides and the sequence was obtained by alignment with that of the muscle isozyme recently published (5), which shows a high degree of homology. The same procedure was used to localize residues 39 and 61 which were not present in any of the overlapping peptides. Sequences 38-85 and 117-138 are located in four V8 *S. aureus* protease peptides: V5, V6, V8, and V9. The failure to detect these four peptides by FAB mass spectrometry could be explained for all cases except V5. The mass of peptide V6 (residues 51-88) was too large to be analyzed with the FAB mass spectrometer used in this study. The bound E126-P127...
The mass obtained...complementary information necessary to confirm the cDNA sequence data (14–18). In the case of the B isozyme of phosphoglycerate mutase, the nucleotide sequence results were corrected by the protein data during the course of this work. Conversely, positions 106–110 and 108–110 were assigned mainly from overlaps from data obtained at the nucleotide level of both the B and the M isozymes of the phosphoglycerate mutase (5–19).

In the present study, our amino acid sequence is in perfect agreement with that derived from the cDNA clone of the liver isozyme described by Sakoda et al. (19) in the preceding paper. Such a result indicates that the phosphoglycerate mutase isozone observed in erythrocytes and in the liver (and probably the brain and the kidney) corresponds to the same protein.

The N-terminal residue of the phosphoglycerate mutase isozone B is blocked and cannot be obtained by conventional amino acid sequencing. The precise N terminus residue cannot be determined precisely from nucleotide sequence data, i.e. it could be either Met or Ala, and in any case the nature of the blocking group could not be determined. This study indicates the usefulness of CAD-MIKE mass spectrometry to solve the exact nature of the N-terminal sequence of blocked proteins.

A comparison was made at the protein level of the amino acid sequence of the human erythrocyte (or brain) isozone of phosphoglycerate mutase with that of the muscle isozone (5). The size of the M isozone is 282 amino acids, compared to 253 for the B isozone, due to the missing last residue. Homology of the protein sequences is very high (78%). The N-terminal residue of the muscle isozone has not been exactly assigned yet. From the nucleotide sequence, it was shown that the N-terminal sequence is Met-Ala-Thr with a blocked N-terminal residue. Studies are in progress in our laboratory to determine the exact nature of this blocking group, using the same strategy followed in the present work, e.g. by mass spectrometry of the corresponding peptide.

In a previous work we have described the sequence of the human erythrocyte 2,3-diphosphoglycerate mutase both by the determination of the cDNA sequence and by analysis of the amino acid composition of the tryptic peptides (20). The strategy used in the present work combining microsequencing and mass spectrometry study of tryptic and V8 peptides was very efficient. The sequence could be established on 0.5 mg of sample and in a short period of time. We think that this kind of strategy could be very useful for the full assessment of a protein structure derived from the cDNA sequence determination.

Acknowledgments—We thank Dr. E. A. Schon for the communication of the human phosphoglycerate mutase isozone B cDNA nucleotide sequence prior to publication and for helpful discussion. Dr. D. M. Ojcius for reviewing the manuscript, and A. M. Dulac for the preparation of the manuscript and the figures.

REFERENCES

Human Phosphoglycerate Mutase Isozyme B Protein Sequence