THE NH_2-TERMINAL SEQUENCES OF GALACTOKINASE FROM *ESCHERICHIA COLI* AND *SACCHAROMYCES CEREVISIAE*

David H. SCHLESINGER, Mark A. SCHELL and David B. WILSON

Endocrine Unit, Massachusetts General Hospital, Boston, Massachusetts 02114 and *Department of Biochemistry, Cornell University, Ithaca, New York 14853, USA*

Received 8 July 1977

1. Introduction

Galactose is metabolized in almost all organisms by a set of three reactions called the Leloir pathway [1]. The first step in this pathway, catalyzed by the enzyme galactokinase, is the transfer of the γ phosphate of ATP to galactose to form α-galactose-1-phosphate. Galactokinase has been purified from a number of organisms [2,3] including *Escherichia coli* [4] and *Saccharomyces cerevisiae* [5].

In both these organisms, the structural genes of the three galactose enzymes are linked [6–8], and the enzymes show coordinate induction [9,10]. In both organisms, the galactokinase structural gene maps at one end of the operon. As part of a biochemical study of the regulation of the galactose enzymes in the bacterium *E. coli* and in yeast, we determined the NH_2-terminal amino acid sequence of galactokinase from both organisms.

2. Materials and methods

2.1. *Strains*

E. coli strain SA1310, which contains a defective λ prophage brought close to the galactose operon by deleting the intervening genetic material, was used as the source of *E. coli* galactokinase and was obtained from Dr Sankar Adhya. *S. cerevisiae* strain X108 was used as the source of yeast galactokinase and was obtained from Dr Bruce Adams [11].

2.2. *Isolation of* E. coli *galactokinase*

Strain SA1310 was grown at 30°C in 100 liters of tryptone broth medium in a New Brunswick Fermentation Fermentor model FM-150 to a $A_{600\,nm}$ 0.60. The temperature was rapidly shifted to 42°C, and the cells were allowed to grow for 2 more hours. The cells were then harvested in a refrigerated Sharples Continuous-Flow Centrifuge (cell yield, 1.5 g/liter). Galactokinase was purified by the procedure of Wilson and Hogness [4].

2.3. *Isolation of yeast galactokinase*

Yeast galactokinase was purified by the procedure of Schell and Wilson [5].

2.4. *Sequence determination*

Galactokinase from *S. cerevisiae* and from *E. coli* were each subjected to automated sequential Edman degradations on a Beckman model 890 peptide protein sequencer using a modified double-coupling Quadrol program of Edman and Begg [12]. Samples of the enzymes were dissolved in 1 M NH_4OH and dried under high vacuum in the spinning cup. Coupling with phenylisothiocyanate was performed twice before initiating automated sequencing. The resulting thiazolinone amino acids were converted to their phenylthiohydantoin (PTH) amino-acid isomers by heating in 1 M HCl at 80°C for 10 min. The PTH amino acids were then identified by thin-layer [13,14] and gas chromatography [15].

3. Results

The specific activity of the *E. coli* galactokinase was 3000 units/mg, and when tested by electrophoresis

Table 1
NH_2-terminal sequences of *E. coli* and *S. cerevisiae* galactokinases

E. coli galactokinase	1 5 10 NH_2 – Ser – Leu – Lys – Glu – Lys – Thr – Gln – Ser – Leu – Phe – Ala – Asn – (42.8) (39.5) (31.2) (28.4) (23.0) 15 – Ala – Phe – Gly – Tyr – Pro – Ala – Thr – (22.6)(19.8)(12.9) (12.0) (16.3)
S. cerevisiae galactokinase	1 5 10 NH_2 – Thr – Lys – Ser – His – Arg – Glu – Arg/Glu – Val – Ile – Val – Pro – (12.2)(10.0)(9.4) (7.8)
E. coli epimerase [17]	1 5 10 NH_2 – Met – Arg – Val – Leu – Val – Thr – Gly – Gly – Ser – Gly – Tyr – Ile – Gly – 15 20 – Ser – His – Thr – Cys – Val – Gln – Leu –

on a polyacrylamide gel in sodium dodecyl sulfate (SDS) by the procedure of Studier [16], 75% of the bound Coomassie Blue was in the galactokinase band, and the rest was present in four bands, each containing about 6% of the bound dye. SDS–polyacrylamide gel electrophoresis of purified yeast galactokinase by the same method [16] showed a single band containing more than 95% of the bound Coomassie Blue present on the gel.

The native enzymes were subjected to automated Edman degradations. The sequences determined in these analyses are given in table 1 along with the previously reported NH_2-terminal sequence of *E. coli* UDP galactose-4-epimerase [17]. The numbers in parentheses indicate the recovery of the PTH amino acids at selected steps in the degradation.

The automated degradation on yeast galactokinase proceeded with a lower repetitive yield (average yield of PTH amino acid per degradation cycle) than did that on the *E. coli* enzyme because of the poor solubility at the pH of the coupling buffer (9.0). Therefore, marked overlap developed early in the degradation, which was terminated after 12 steps because of a lack of confidence in the identification of residues beyond that point in the NH_2-terminal sequence. Of note is the observation of microheterogeneity at position 7, in which both glutamic acid and argenine were found.

On the other hand, the automated degradation on *E. coli* galactokinase, which revealed a very different NH_2-terminal sequence, proceeded with a repetitive yield of about 95% through 18 cycles.

4. Discussion

Serine was reported to be the NH_2-terminal residue of *E. coli* galactokinase in a previous study [18] and that result is confirmed by the sequence determined here. The finding of a single sequence for the first 19 residues of the *E. coli* galactokinase confirms the finding that this protein consists of a single polypeptide chain, as reported previously [18]. Similarly, the sequence data for *S. cerevisiae* galactokinase also indicate that this enzyme consists of a single polypeptide chain, with possible microheterogeneity as proposed by Schell and Wilson [5].

The NH_2-terminal sequence of the *S. cerevisiae* and the *E. coli* galactokinases do not show any sign of homology. This finding does not rule out an evolutionary relationship between the two enzymes, because homology could exist between the NH_2-terminal sequence of the *E. coli* enzyme and an internal sequence in the *S. cerevisiae* enzyme since the *S. cerevisiae* enzyme is 50% larger than the *E. coli* enzyme.

There is also little homology between the NH_2-terminal sequence of *E. coli* galactokinase and the NH_2-terminal sequence of *E. coli* UDP galactose-4-epimerase [17] (see table 1). In this case the monomeric

proteins are the same size, and so this finding, coupled with the fact that the amino acid compositions of these enzymes are very different [18], indicates that if these two enzymes did arise by gene duplication, there has been very extensive divergence.

The NH_2-terminal sequences determined here are extensive enough to allow an unambiguous determination of the translational initiation sequence for each structural gene when the appropriate regions of DNA are sequenced.

Acknowledgements

We are pleased to acknowledge the skilled technical assistance of Diana Irwin in the purification of the *E. coli* galactokinase. This work was supported by Grant AM11923 and NO1-CB-64079 from the National Institutes of Health.

References

[1] Cardini, C. E. and Leloir, L. F. (1961) in: Biochemist's Handbook (King, E. and Sperry, W. eds) pp. 404–405, 439–440, D. van Nostrand Company Inc., New York.
[2] Ballard, F. J. (1966) Biochem. J. 98, 347–352.
[3] Blume, K. G. and Beutler, E. (1971) J. Biol. Chem. 246, 6507–6510.
[4] Wilson, D. B. and Hogness, D. S. (1966) Meth. Enzymol. 8, 229–235.
[5] Schell, M. A. and Wilson, D. B. (1977) J. Biol. Chem. 252, 1162–1166.
[6] Douglas, H. C. and Hawthorne, D. C. (1964) Genetics 49, 837–844.
[7] Bassel, J. and Mortimer, R. (1971) J. Bacteriol. 108, 179–183.
[8] Adler, J. and Kaiser, A. D. (1963) Virology 18, 117–126.
[9] Buttin, G. (1963) J. Mol. Biol. 7, 164–182.
[10] Winge, O. and Roberts, C. (1948) CR Trav. Lab. Carlsberg. Ser. Physiol. 24, 135–263.
[11] Adams, B. G. (1972) J. Bacteriol. 111, 308–315.
[12] Edman, P. and Begg, G. (1967) Eur. J. Biochem. 1, 80–91.
[13] Edman, P. (1960) Ann. NY Acad. Sci. 88, 602–610.
[14] Schlesinger, D. H., Niall, H. D. and Goldstein, G. (1975) Biochemistry 14, 2214–2218.
[15] Pisano, J. J. and Bronzert, T. J. (1969) J. Biol. Chem. 244, 5597–5607.
[16] Studier, W. (1973) J. Mol. Biol. 79, 237–248.
[17] Schlesinger, D. H., Niall, H. D. and Wilson, D. B. (1974) Biochem. Biophys. Res. Commun. 61, 282–289.
[18] Wilson, D. B. and Hogness, D. S. (1969) J. Biol. Chem. 244, 2137–2142.