Characterization of *Neurospora crassa* catabolic dehydroquinase purified from *N. crassa* and *Escherichia coli*

Alastair R. HAWKINS,*† William R. REINERT*‡ and Norman H. GILES*

*Department of Molecular and Population Genetics, University of Georgia, Athens, GA 30602, U.S.A., and

†Department of Biochemistry, University of Leicester, Leicester LE1 7RH, U.K., and ‡Campbell Institute of Research and Technology, 2611 Branch Pike, Cinnaaminsin, NJ 08077, U.S.A.

(Received 27 January 1982/Accepted 24 March 1982)

1. *Neurospora crassa* catabolic dehydroquinase has been purified from *N. crassa* and *Escherichia coli*. 2. Protein-sequence and gel-electrophoretic data show that apparently pure, homogeneous native dehydroquinase is a mixture of intact and proteinase-cleaved enzyme monomers. 3. Protein-sequence data and steady-state kinetics show that the catabolic dehydroquinase gene of *N. crassa* is expressed with fidelity in *E. coli*.

Degradation by endogenous proteinases in crude extracts of *Neurospora crassa* mycelium has been observed by several workers purifying different enzymes (Lampkin *et al.*, 1976; Lumsden & Coggins, 1977; Makoff & Radford, 1977; Yu *et al.*, 1973). The nature of these proteinases has been investigated, and at least five activities have been identified, several having specific inhibitors (Yu & Kula, 1975). The extent of enzyme degradation varies, but is often of a limited nature, causing loss of enzyme activity and/or changes in molecular weight (Rigby & Radford, 1981).

The multimeric *N. crassa* enzyme catabolic dehydroquinase catalyses the second of the first three reactions in catabolism of quinic acid. The *qa2⁺* gene encoding catabolic dehydroquinase, which is part of the inducible, tightly linked *qa* gene cluster, has been cloned and expressed in *Escherichia coli*, and the nucleotide sequence of the first 136 amino acids determined (Alton *et al.*, 1978, 1982; Barea & Giles, 1978; Chaleff, 1974*a,b*; Case & Giles, 1975, 1976; Case *et al.*, 1977*a,b*; Hautala *et al.*, 1975; Stroman *et al.*, 1978; Vapnek *et al.*, 1977). Protein-sequencing studies on *N. crassa* catabolic dehydroquinase purified from *N. crassa* and *E. coli* were undertaken to complement the elucidation of the DNA sequence of the *qa2⁺* gene and to determine if this eukaryotic gene is being expressed with fidelity in *E. coli*. We report here that the *N*-terminal amino acid sequences of catabolic dehydroquinase purified from *N. crassa* and *E. coli* show no homology, and that the enzyme monomers undergo proteolytic cleavage during purification to yield a population of truncated molecules that vary in length. These truncated molecules co-purify to

yield an apparently homogeneous native enzyme. Amino acid compositions, steady-state kinetics, and protein sequencing of the enzyme purified from *E. coli* support the view that the *qa2⁺* gene is expressed with fidelity in *E. coli*.

Materials and methods

Strains

N. crassa strain M18, lacking quinate dehydrogenase, was used in all extractions of catabolic dehydroquinase from *N. crassa* (Hautala *et al.*, 1975). *E. coli* strain SK2884 (*pnp7*, *thi*, *endA*, *sbcB15*) containing either plasmid pV_k57 or pV_k88 was used in all extractions of catabolic dehydroquinase from *E. coli*. This strain produces elevated levels of catabolic dehydroquinase (Hautala *et al.*, 1979).

Reagents

Amino-polystyrene, 1-ethyl-3-(3-diethylamino-propyl)carbodi-imide hydrochloride, acetonitrile, trifluoroacetic acid, phenyl isothiocyanate, phenylene di-isothiocyanate, *N*-methylmorpholine, aminopropyl-glass (CPG 88A; 400 mesh), SDS (Sequal grade), pyridine, dansyl (5-dimethylaminonaphthalene-1-sulphonyl) chloride, dansyl-amino acid standards, acetone, polyamide thin-layer sheets (manufactured by Schleicher and Scheull) and amino acid phenylthiohydantoin derivatives were obtained from Pierce Chemical Co.; Sephadex G-200 and Bio-Gel P4 were from Sigma; DEAE-cellulose (DE-52) was from Whatman; AG 50X4 and AG 1X2 resins were from Bio-Rad; all other chemicals were of reagent grade.

Growth of *N. crassa*

N. crassa was grown in aqueous cultures as previously described (Hautala *et al.*, 1975).

Abbreviation used: SDS, sodium dodecyl sulphate.

Growth of *E. coli*

E. coli strain SK2284 was grown to late-exponential phase in minimal medium (Wilets *et al.*, 1969), supplemented with 2% glucose, 1% cas-amino acids, 10 μ M-thiamin and ampicillin (20 μ g·ml⁻¹) in a Lab Line S.M.S. Hi Density Fermentor at 37°C in a P2 (level-two physical-containment) facility.

Enzyme assays

Catabolic dehydroquinase activity was measured spectrophotometrically (Beckman Acta III) as previously described (Vapnek *et al.*, 1977).

Purification of catabolic dehydroquinase

(a) Catabolic dehydroquinase was purified from *N. crassa* as previously described (Hautala *et al.*, 1975).

(b) Catabolic dehydroquinase was purified from *E. coli* by using the following schedule.

Step 1. Washed *E. coli* cell pastes were frozen to -30°C in an Aminco French pressure cell and pressed at 138 MPa (20000 lb/in²), and the lysate diluted to 0.5 g/ml with 100 mM-potassium phosphate (pH 7.2)/1 mM-EDTA/1 mM-phenylmethanesulphonyl fluoride/1 mM-dithiothreitol (buffer A).

Step 2. After a 40-min incubation at 37°C with deoxyribonuclease and ribonuclease (10 μ g/ml) the diluted lysates were cleared by centrifugation at 10000 g for 20 min. The remainder of the purification was as previously described for *N. crassa* (Hautala *et al.*, 1975) with two modifications: (1) (NH₄)₂SO₄ to 50% saturation was required to precipitate the enzyme; (2) an additional final step of chromatography on a column (0.5 cm × 110 cm) of Bio-Gel P4, with 4% formic acid as eluent, was used. Table 1 shows the enzyme-purification data.

Polyacrylamide-gel electrophoresis

Analytical polyacrylamide-disc-gel electrophoresis and SDS/gradient slab-gel electrophoresis utilizing a linear 7–25% polyacrylamide gradient were as previously described (Davis, 1964; O'Farrell, 1975).

Manual protein sequencing

Enzymic digestions of catabolic dehydroquinase, subsequent peptide purification, manual sequencing, amide assignments and amino acid analysis were as previously described (Matsubara & Sasaki, 1969; Wootton *et al.*, 1975a,b; Holder *et al.*, 1975).

Automated protein sequencing

Coupling of catabolic dehydroquinase to aminopropyl-glass (CPG 88A; 200–400 mesh) by using phenylene di-isothiocyanate and 1-ethyl-3-(3-dimethylaminopropyl)carbodi-imide, sequencing (Sequemat Mini 15) and identification of amino acid phenylthiohydantoin derivatives (Perkin-Elmer series 3 high-pressure liquid chromatograph and Zorbax ODS reverse-phase column) were as previously described (Laursen, 1977; Machleidt & Wachter, 1977; Zimmerman *et al.*, 1977). Liquid-phase protein sequencing of catabolic dehydroquinase was as previously described (Hunkapillar & Hood, 1978).

Results and discussion

Characterization of catabolic dehydroquinase purified from *N. crassa*

Catabolic dehydroquinase was purified from *N. crassa* as previously described (Hautala *et al.*, 1975). The native enzyme electrophoresed as a single species in non-denaturing polyacrylamide gels, and after heating to 90°C in 10% (w/v) SDS for 1–10 h showed bands corresponding to approximate molecular masses 10, 20, 30 and 40 kDa. The 10 kDa species increased with a concomitant decrease in the other higher-molecular-mass species with increasing incubation times. Electrophoresis of the heated samples in 7–25%-polyacrylamide gradient denaturing gels fractionated the 10 kDa protein band into multiple bands of molecular mass varying from 5 to 10 kDa. *N*-Terminal determinations with dansyl chloride showed fluorescence corresponding to the dansyl derivatives of ϵ -lysine, *O*-tyrosine, proline, or leucine, or both. Samples of protein were incubated for 15 h at 15°C with

Table 1. Purification of catabolic dehydroquinase from 172 g of *E. coli* cell paste

Purification	Total protein (mg)	Total activity (units)	Specific activity (units·mg ⁻¹)	Recovery (%)
Crude supernatant	33 000	924	0.028	100
71°C supernatant	2560	640	0.25	69
(NH ₄) ₂ SO ₄ (0–50%-satd.)	237	260	1.1	28
Sephadex G-200	21	243	11.6	26
DEAE-cellulose	0.92	127	138	14
Bio-Gel P4	0.90	0.0*	0.0	13.7

* Enzyme activity was irreversibly lost after solubilizing in 4% (v/v) formic acid.

anhydrous 1M-HCl in methanol to denature the enzyme monomers. Dansylation and hydrolysis of treated samples showed fluorescence of dansyl derivatives of proline, leucine, serine, threonine, alanine, glycine, histidine, aspartate and glutamate in addition to ϵ -lysine and *O*-tyrosine. Dansylation of dried portions of the acid/methanol showed no dansyl derivatives corresponding to amino acids. Despite reports that the acid/methanol treatment does not cause significant peptide-bond hydrolysis (Sheehan & Yang, 1958), this possibility cannot be entirely ruled out. However, in this case we feel that this is unlikely, as identical treatment of catabolic dehydroquinase purified from *E. coli* did not show the dramatic increase in the number of amino groups able to react with dansyl chloride after incubation with acid/methanol. These observations lead to the conclusion that the catabolic dehydroquinase monomers are a heterogeneous population of intact and truncated molecules, the majority of which are unable to react with dansyl chloride until treated with anhydrous-acid/methanol.

Protein-sequence analysis of catabolic dehydroquinase purified from *N. crassa*

Purified catabolic dehydroquinase that had not been treated with anhydrous-acid/methanol was subjected to *N*-terminal sequence analysis in both solid- and liquid-phase automated sequencers and to manual sequence analysis of peptides purified after digestion with staphylococcal V8 proteinase and trypsin. The results are summarized in Fig. 1, where it can be seen that the automated sequencing detected the presence of at least four related populations of molecules from five different enzyme purifications, differing in the position of the *N*-terminus. Eight peptides were purified from the proteinase digests and were sequenced manually; these peptides both overlap the automated sequences and define new sequences. These sequences confirm the conclusion that the purified native enzyme consists of a heterogeneous population of intact and truncated molecules. The intact and truncated molecules co-purify and complement one another to retain enzyme activity. We postulate that the *N*-termini of the monomer polypeptide fragments are either buried within the tertiary structure and/or stabilized by electrostatic interactions with the *C*-termini of truncated monomers, therefore failing to react with dansyl chloride. Slow or non-reactive buried amino groups that are stabilized by electrostatic interactions have been described in glyceraldehyde 3-phosphate dehydrogenase and aldolase (Lambert & Perham, 1977; Lambert *et al.*, 1977). The buried amino groups of the truncated dehydroquinase monomers are only able to react with dansyl chloride after denaturation with anhydrous-acid/methanol. The *N*-terminal sequences detected

Fig. 1. Protein sequences derived from catabolic dehydroquinase purified from (a) *N. crassa* and (b) *E. coli* (a) *N. crassa* sequences marked (i)–(iii) were derived from solid- and liquid-phase *N*-terminal sequencing; peptides marked S1–S7 and T1 were sequenced manually after purification from protein digested with staphylococcal V8 proteinase and trypsin respectively. Peptides were purified by sequential chromatography on AG 50X4 resin, Bio-Gel P4 and AG 1X2 resin. Sequences marked ? indicate ambiguously determined residues, and rightward-pointing arrows delineate the extent of sequence determination. Protein was bound to the glass beads by phenylene di-isothiocyanate or 1-ethyl-3-(3-diethylaminopropyl)carbodi-imide via amino or carboxy groups. (b) The sequence was determined by solid-phase automated protein sequencing. The protein was linked to the glass beads by using 1-ethyl-3-(3-diethylaminopropyl)carbodi-imide via carboxy groups. The one-letter notation for amino acid sequences is defined in *Biochem. J.* (1969) 113, 1–4.

by automated sequencing of catabolic dehydroquinase monomers do not correspond to the true *N*-terminus of the enzyme, but to residues 88–93, showing that these sequences derive from a truncated molecule (see Fig. 2). The automated sequencing runs consistently showed a background 'noise' of 'scrambled' sequence that we believe is due to sequencing from the *N*-terminal of proteinase-cleaved monomers that are present in a much lower yield.

Protein-sequence analysis of catabolic dehydroquinase purified from *E. coli*

N. crassa catabolic dehydroquinase was purified from *E. coli* by the protocol detailed in the Materials and methods section. Protein was purified from *E. coli* containing either plasmid pVk57 or pVk88, which carry the *qa2⁺* gene on independently isolated different-size DNA fragments. The protein was subjected to automated solid-phase sequencing

and the sequence obtained, which was the same with strains carrying either plasmid, is shown in Fig. 1(b). Dansylation of the purified protein showed dansylalanine as the *N*-terminus, but dansyl derivatives corresponding to alanine, proline and threonine were observed after dansylation of protein treated with anhydrous-acid/methanol. These results show that the native enzyme purified from *E. coli* has some monomers that are truncated, but that the extent of degradation is much lower. The protein sequence deduced clearly corresponds to the true *N*-terminal catabolic dehydroquinase sequence (see Fig. 2), although it shows no homology with the *N*-terminal sequence of the enzyme purified from *N. crassa*.

Fig. 2. summarizes the protein-sequence data and aligns the sequenced peptides with the theoretical protein sequence deduced from the nucleotide sequence of the *qa2⁺* gene. The amino acid sequence of the protein purified from *E. coli* represents the true *N*-terminus, and the sequences derived from protein purified from *N. crassa* derive from molecules truncated predominantly around position 88. It is apparent that some of the staphylococcal V8 proteinase-digested peptides have *N*-termini distal to residues other than glutamate or aspartate, and these are not simply interpreted as the *N*-terminal peptides of previously truncated molecules. The presence of peptides S1, S2 and T1, which precede the

apparent *N*-terminus of the protein purified from *N. crassa*, demonstrate the presence of the true *N*-terminal half of the protein in the native enzyme. The finding that the true *N*-terminal sequence was not detected by automated protein sequencing in enzyme purified from *N. crassa* may be due to chemical blockage such as formylation or acetylation. The reason that different populations of molecules of catabolic dehydroquinase were sequenced in different preparations from *N. crassa* is probably due to essentially random differences in the extent of proteolytic degradation in each preparation.

The amino acid compositions of catabolic dehydroquinase purified from *N. crassa* and *E. coli* showed minor differences between each other and the theoretical composition deduced from the DNA sequence. Steady-state kinetics of the enzyme purified from both sources showed that K_m determinations were indistinguishable. Taken as a whole, the protein sequence of the enzyme purified from *E. coli*, amino acid compositions and steady-state kinetics support the view that the *N. crassa qa2⁺* gene is expressed with fidelity in *E. coli*. Protein-sequence and gel-electrophoretic data of the protein purified from *N. crassa* show that an apparently homogeneous active native enzyme consists of proteinase-cleaved and intact monomers. Previous estimates of the molecular mass of the catabolic dehydroquinase monomer were 9–10 kDa; however, the value deduced from the DNA sequence of the *qa2⁺* gene is 18.2 kDa. The explanation for the original underestimate of the molecular mass of the enzyme monomers seems likely to be related to the extreme heat treatment used to solubilize and denature the enzyme in SDS. Before the *N*-terminal sequences of the enzyme purified from *N. crassa* is a proline-rich region from residues 72 to 79. We suggest that this region is unordered, heat-labile and exposed to the solvent, factors favouring degradation of the molecule around this region. The heat treatment of 95°C for 10 h used to denature polypeptides for the determination of subunit molecular masses is postulated to break effectively the already partially proteinase-cleaved molecules around this proline-rich region, giving an anomalous low value for the subunit molecular mass.

Fig. 2. Sequence of the first 139 amino acids of catabolic dehydroquinase (Alton *et al.*, 1982) as deduced from the DNA sequence shown in single-letter code [Biochem. J. (1969) 113, 1–4].

Amino acids 2–17 and 88–106, singly and doubly underlined and labelled 'NT' were determined from *E. coli* and *N. crassa* protein respectively. Underlined sequences labelled S1–S7 and T1 were determined by manual sequencing of peptides generated by staphylococcal V8 proteinase and tryptic digests respectively. The glutamate residues at positions 22 and 136 were inferred from the specificity of staphylococcal V8 proteinase.

This work was supported by an E.M.B.O. (European Molecular Biology Organization) Long-Term Fellowship to A. R. H. and by U.S. National Institutes of Health (NIH) grant GM 21454 to N. H. G. We thank Dr. S. Kushner for use of the Lab Line S.M.S. Hi Density Fermentor and access to a P2 facility, and to Dr. Leroy Hood for use of the automated liquid phase sequencer.

References

- Alton, N. K., Hautala, J. A., Giles, N. H. & Kushner, S. R. (1978) *Gene* 4, 241–259

- Alton, N. K., Buxton, F., Patel, V. B., Giles, N. H. & Vapnek, D. (1982) *Proc. Natl. Acad. Sci. U.S.A.* in the press
- Barea, J. L. & Giles, N. H. (1978) *Biochim. Biophys. Acta* **524**, 1–14
- Case, M. E. & Giles, N. H. (1975) *Proc. Natl. Acad. Sci. U.S.A.* **72**, 553–557
- Case, M. E. & Giles, N. H. (1976) *Mol. Gen. Genet.* **147**, 83–89
- Case, M. E., Hautala, J. A. & Giles, N. H. (1977a) *J. Bacteriol.* **129**, 166–172
- Case, M. E., Pueyo, C., Barea, J. L. & Giles, N. H. (1977b) *Genetics* **86**, 68 (abstr.)
- Chaleff, R. S. (1974a) *J. Gen. Microbiol.* **81**, 337–355
- Chaleff, R. S. (1974b) *J. Gen. Microbiol.* **81**, 357–372
- Davis, B. J. (1964) *Ann. N.Y. Acad. Sci.* **121**, 404–427
- Hautala, J. A., Jacobson, J. W., Case, M. E. & Giles, N. H. (1975) *J. Biol. Chem.* **250**, 6008–6014
- Hautala, J. A., Bassett, C. L., Giles, N. H. & Kushner, S. R. (1979) *Proc. Natl. Acad. Sci. U.S.A.* **76**, 5774, 5778
- Holder, A. A., Woolton, J. C., Baron, A. J., Chambers, G. K. & Fincham, J. R. S. (1975) *Biochem. J.* **149**, 757–773
- Hunkapillar, M. W. & Hood, L. E. (1978) *Biochemistry* **17**, 2124–2133
- Lambert, J. M. & Perham, R. N. (1977) *Biochem. J.* **161**, 49–62
- Lambert, J. M., Perham, R. N. & Coggins, J. R. (1977) *Biochem. J.* **161**, 63–71
- Lampkin, S. L., Cole, K. W., Vitto, A. & Gaertner, F. W. (1976) *Arch. Biochem. Biophys.* **177**, 561–565
- Laursen, R. A. (1977) *Methods Enzymol.* **47**, 277–288
- Lumsden, J. & Coggins, J. R. (1977) *Biochem. J.* **161**, 599–607
- Machleidt, W. & Wachter, E. (1977) *Methods Enzymol.* **47**, 263–277
- Makoff, A. J. & Radford, A. (1977) *Biochim. Biophys. Acta* **485**, 314–329
- Matsubara, H. & Sasaki, R. M. (1969) *Biochem. Biophys. Res. Commun.* **35**, 175–181
- O'Farrell, P. H. (1975) *J. Biol. Chem.* **250**, 2007–2021
- Rigby, D. J. & Radford, A. (1981) *Biochim. Biophys. Acta* **661**, 315–322
- Sheehan, J. C. & Yang, D. D. H. (1958) *J. Am. Chem. Soc.* **80**, 1154–1159
- Stroman, P., Reinert, W. R. & Giles, N. H. (1978) *J. Biol. Chem.* **253**, 4593–4598
- Vapnek, D., Hautala, J. A., Jacobson, J. W., Giles, N. H. & Kushner, S. R. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 3508–3512
- Willems, N. S., Clark, A. J. & Low, K. B. (1969) *J. Bacteriol.* **97**, 244–249
- Wootton, J. C., Taylor, J. G., Jackson, A. A., Chambers, G. K. & Fincham, J. R. S. (1975s) *Biochem. J.* **149**, 739–748
- Wootton, J. C., Baron, A. J. & Fincham, J. R. S. (1975b) *Biochem. J.* **149**, 749–755
- Yu, P. H. & Kula, M. R. (1975) *J. Biochem. (Tokyo)* **56**, 271–287
- Yu, P. H., Kula, M. R. & Tsai, H. (1973) *Eur. J. Biochem.* **32**, 129–135
- Zimmerman, C. L., Apella, E. & Pisano, J. S. (1977) *Anal. Biochem.* **77**, 569–573