Identification of the active-site lysine residues of two biosynthetic 3-dehydroquinases

Subhendu CHAUDHURI,* Kenneth DUNCAN,† Lloyd D. GRAHAM and John R. COGGINS

Department of Biochemistry, University of Glasgow, Glasgow G12 8QQ, Scotland, U.K.

The lysine residues involved in Schiff-base formation at the active sites of both the 3-dehydroquinase component of the pentafunctional *arom* enzyme of *Neurospora crassa* and of the monofunctional 3-dehydroquinase of *Escherichia coli* were labelled by treatment with 3-dehydroquinone in the presence of NaB^3H_4 . Radioactive peptides were isolated by h.p.l.c. following digestion with CNBr (and in one case after further digestion with trypsin). The sequence established for the *N. crassa* peptide was ALQHGDVVKLVGVAR, and that for the *E. coli* peptide was QSFADADIPKIA. An amended nucleotide sequence for the *E. coli* gene (*aroD*) that encodes 3-dehydroquinase is also presented, along with a revised alignment of the deduced amino acid sequences for the biosynthetic enzymes.

INTRODUCTION

In micro-organisms and plants, the third step of the early common pathway (the shikimate pathway) for the biosynthesis of aromatic amino acids is catalysed by the enzyme 3-dehydroquinase (EC 4.2.1.10). This enzyme introduces the first double bond in the aromatic ring, converting 3-dehydroquinone into 3-dehydroshikimate by the *syn* elimination of water. In the reaction, the N^ϵ -amino group of a lysine side chain forms a Schiff base with the keto group of 3-dehydroquinone (Butler *et al.*, 1974), and a basic group, recently shown to be the imidazole side chain of a histidine residue, facilitates proton abstraction (Walsh, 1979; Chaudhuri *et al.*, 1986).

Four classes of 3-dehydroquinase have been identified. Members of three classes catalyse the biosynthetic reaction described above, and illustrate the diversity found in the organization of the shikimate pathway enzymes and genes in various organisms. Monofunctional 3-dehydroquinases have been identified in bacteria (Berlyn & Giles, 1969); the *Escherichia coli* enzyme has been purified to homogeneity and the complete sequence is known (Chaudhuri *et al.*, 1986; Duncan *et al.*, 1986). The plant enzyme occurs on a bifunctional polypeptide chain in association with shikimate dehydrogenase (Polley, 1978; Koshiba, 1978; Mousdale *et al.*, 1987), which catalyses the fourth step on the shikimate pathway. In the third class, exemplified by the fungi, 3-dehydroquinase is one of the five activities carried on the *arom* multifunctional enzyme, the others being the activities of steps, 2, 4, 5 and 6 of the same pathway (Lumsden & Coggins, 1977, 1978; Gaertner & Cole, 1977; Smith & Coggins, 1983; Lambert *et al.*, 1985; Coggins *et al.*, 1985). The *arom* multifunctional enzyme has been isolated from *Neurospora crassa* and studied in detail (Lumsden & Coggins, 1977, 1978; Gaertner & Cole, 1977; Smith & Coggins, 1983; Lambert *et al.*, 1985). Limited-proteolysis experiments have shown that it is possible to isolate an *arom* fragment that carries both the 3-dehydroquinase and shikimate dehydrogenase activities, and is similar in size and structure to the bifunctional plant enzyme (Smith & Coggins, 1983; Coggins *et al.*, 1985). The complete sequences of two fungal biosynthetic 3-dehydroquinases, the enzymes from *Aspergillus nidulans* (Charles *et al.*, 1985, 1986) and *Saccharomyces cerevisiae* (Duncan *et al.*, 1987), are known. A comparison of the *E. coli* and fungal sequences has revealed considerable similarity at the primary sequence level (Duncan

et al., 1987). The fourth class of 3-dehydroquinases contains additional monofunctional enzymes found only in fungi, such as *N. crassa* and *A. nidulans*. In *N. crassa*, a monofunctional 3-dehydroquinase is induced in response to the presence of quinic acid in the growth medium, along with two other enzymes and two regulatory proteins that together make up the quinate degradative pathway (for a review see Giles *et al.*, 1985). The inducible 3-dehydroquinase of *N. crassa* shows little overall similarity to the biosynthetic 3-dehydroquinases at the primary sequence level (Duncan *et al.*, 1987).

As a first step in the more detailed structural analysis of 3-dehydroquinases, we report here the determination of the sequence around the active-site lysine residue of two of the biosynthetic forms of the enzyme, the monofunctional *E. coli* 3-dehydroquinase and the 3-dehydroquinase component of the *N. crassa arom* multifunctional enzyme.

MATERIALS AND METHODS

Materials

NaB^3H_4 was purchased from Amersham International (Amersham, Bucks., U.K.); CNBr was from Sigma Chemical Co. (Poole, Dorset, U.K.); bovine trypsin (Tos-Phe-CH₂Cl-treated) was from Worthington Biochemical Corp. (Freehold, NJ, U.S.A.). Sephadex G-75 (superfine grade) was from Pharmacia (Milton Keynes, Bucks., U.K.). All other materials were purchased from the sources listed in the appropriate references or from BDH Chemicals (Poole, Dorset, U.K.).

Enzyme purification

The *arom* multifunctional enzyme was purified from *N. crassa* by the method of Smith & Coggins (1983). *E. coli* 3-dehydroquinase was purified as described in Duncan *et al.* (1986).

Active-site labelling

***N. crassa* enzyme.** A solution of the *arom* multifunctional enzyme (90 nmol) in 50 mM-potassium phosphate buffer, pH 7.5, containing 1 mM-3-dehydroquinone and 0.4 mM-dithiothreitol (total volume 7.5 ml), was treated with a series of small portions of NaB^3H_4 (14 Ci/mmol) in 50 mM-NaOH. Approx. 10 μ l portions were added at 5 min intervals, following which the solution was assayed for 3-dehydroquinase activity. After ad-

* Present address: Department of Dairy and Animal Science, Pennsylvania State University, University Park, PA 16802, U.S.A.

† Present address: Glaxo Group Research, Greenford, Middx. UB6 0HE, U.K.

dition of a total volume of 100 μ l (100 mCi) of radioactive borohydride, the residual enzyme activity was 30 % of the initial value. Additions of unlabelled NaBH_4 solution were then made until the final enzyme activity had fallen to 1 % of the initial value; this required 15 μ l of NaBH_4 (50 mg/ml) in 50 mM-NaOH. The labelled protein was exhaustively dialysed against 100 mM-*N*-ethylmorpholine acetate, pH 8.5, and then freeze-dried.

***E. coli* enzyme.** A solution of *E. coli* 3-dehydroquinase (186 nmol) in 50 mM-potassium phosphate buffer, pH 7.5, containing 2 mM-3-dehydroquinone and 0.4 mM-dithiothreitol (total volume 8.0 ml), was treated with three sequential additions of 20 μ l (10.7 mCi) of NaB^3H_4 (175 mCi/mmol) in 50 mM-NaOH, following which the residual activity was 2.6 % of the initial value. The labelled protein was exhaustively dialysed against 0.5 % (w/v) NH_4HCO_3 , and freeze-dried.

Reduction and carboxymethylation

This was carried out on all of the material from the above sections by using iodoacetic acid in 5 M-guanidinium chloride, as described previously (Lumsden & Coggins, 1978). Excess reagents were removed by dialysis against 0.5 % (w/v) NH_4HCO_3 , and the protein solution was freeze-dried.

Digestion with CNBr

As the oxidation of methionine residues prevents the reaction with CNBr from taking place, 2-mercaptoethanol or dithiothreitol was used to minimize methionine sulfoxide formation during purification and storage of the enzymes (Jori *et al.*, 1968).

***N. crassa* enzyme.** All of the reduced and carboxymethylated *arom* protein (see the previous section) was dissolved in 70 % (v/v) formic acid. A 500-fold molar excess of CNBr over methionine residues (of which there are 32 per subunit) was added, and the reaction was allowed to continue in the dark at room temperature for 24 h. The reaction mixture was diluted 10-fold with water and freeze-dried several times.

***E. coli* enzyme.** All of the reduced and carboxymethylated enzyme (see the previous section) was dissolved in 1.9 ml of 0.5 % (w/v) NH_4HCO_3 . A 100 μ l portion of 2-mercaptoethanol was added and the mixture was incubated at 20 °C for 20 h, under N_2 . The sample was then dialysed against 0.5 % (w/v) NH_4HCO_3 and then freeze-dried. For CNBr cleavage, the enzyme was dissolved in 70 % (v/v) formic acid. A 50-fold molar excess of CNBr over methionine residues (of which there are 11 per subunit) was added, and the mixture (2 ml final volume) was incubated at room temperature in the dark for 24 h. The reaction mixture was diluted 10-fold with water and freeze-dried several times.

Purification of active-site peptides

***N. crassa* peptides.** (i) Separation of CNBr-cleavage peptides. This was carried out by gel filtration on a Sephadex G-75 (superfine grade) column (90 cm $\times$ 2.1 cm) with 30 % (v/v) acetic acid as eluent; the flow rate was 1 ml/min, and 2 ml fractions were collected. Eluted peptides were detected fluorimetrically with fluorescamine after total digestion of 100 μ l portions with NaOH (Liu *et al.*, 1979). Fractions were pooled as in Fig. 1, and freeze-dried.

(ii) Digestion with trypsin. The peptide pools CNI, CNII and CNIII from the gel-filtration column were dissolved in 0.5 % (w/v) NH_4HCO_3 and digested with trypsin for a total of 8 h at 37 °C. Two additions of trypsin (15 μ g) were made at zero and

4 h. The reactions were terminated by boiling, and the mixtures of peptides were freeze-dried.

(iii) Separation of radioactive tryptic-digest peptides. The tryptic-digest peptides were separated first by h.p.l.c. on a TSK G2000SW column (0.75 cm $\times$ 60 cm) with 0.1 % (v/v) trifluoroacetic acid as eluent and a flow rate of 0.2 ml/min. Each radioactive peak was pooled from several runs, giving pools CNI/T, CNIII/T and CNIV/T. Each pool was freeze-dried and then separated by reverse-phase h.p.l.c. on a Waters C_{18} μ Bondapak column (0.39 cm $\times$ 30 cm). Peptides were eluted with a linear gradient (0–60 %) of propan-2-ol in 0.1 % (v/v) trifluoroacetic acid; the flow rate was 1 ml/min, and the gradient was developed over 60 min. Individual u.v.-absorbing peaks were collected manually and checked for radioactivity (Fig. 2). The most radioactive fractions from each run were retained; material from several runs was pooled and freeze-dried before amino acid analysis and sequencing.

***E. coli* peptides.** (i) Separation of CNBr-cleavage peptides. The *E. coli* CNBr-cleavage peptides were separated by h.p.l.c. gel-filtration chromatography on a TSK G2000SW column (0.75 cm $\times$ 60 cm) with 0.1 % (v/v) trifluoroacetic acid as eluent; the flow rate was 0.5 ml/min, and 0.5 ml fractions were collected. Fractions from several identical runs were collected, and two radioactive 215 nm-absorbing peaks (Fig. 3; EC1 and EC2) were accumulated and freeze-dried.

(ii) Separation of radioactive peptides. Pool EC2 was dissolved in 1 ml of 0.1 % (v/v) trifluoroacetic acid and separated by

Fig. 1. Separation of peptides obtained by CNBr digestion of substrate-labelled *N. crassa arom* enzyme complex by gel filtration on Sephadex G-75 (superfine grade)

Chromatography conditions are described in the Materials and methods section. The fractions from the column were pooled as indicated. The upper trace (○) shows relative fluorescence and the lower trace (●) shows the radioactivity determined by liquid-scintillation counting (c.p.m./5 μ l of fraction).

reverse-phase h.p.l.c. (in several runs) on a Waters C₁₈ μ Bondapak column (0.39 cm $\times$ 30 cm). Peptides were eluted with a linear gradient (0–50%) of acetonitrile in 0.1% (v/v) trifluoroacetic acid; the flow rate was 1 ml/min, 0.5 ml fractions were collected, and the gradient was developed over 60 min. Portions (10 μ l) from each fraction were liquid-scintillation-counted for radioactivity. Three radioactive 215 nm-absorbing peaks were pooled (Fig. 4; EC3, EC4 and EC5). Each pool was assessed for radioactivity and then freeze-dried.

Sequence analysis of purified peptides

***N. crassa* peptide.** The amino acid composition of the peptide CNIV/T was determined, and this peptide was also sequenced on a modified Beckman model 890 liquid-phase sequencer as described by Runswick & Walker (1983). The radioactivity released at each cycle was determined by liquid-scintillation counting.

***E. coli* peptide.** The peptide EC4 was sequenced by automated gas-phase sequencing on a Beckman model 890 liquid-phase sequencer as described previously (Smith *et al.*, 1982). The radioactivity released at each cycle was determined by liquid-scintillation counting.

RESULTS

Isolation of active-site peptides

From the known amino acid composition of the *N. crassa*

arom protein (Lumsden & Coggins, 1978), it was predicted that complete digestion with CNBr would yield 33 peptides, only one of which would be radiolabelled. The elution profile (Fig. 1) shows that several radioactive peaks separable by gel-filtration chromatography on Sephadex G-75 were obtained. Three of the radioactive peaks were digested with trypsin, and then separated by h.p.l.c. gel-filtration chromatography followed by reverse-phase h.p.l.c. (Fig. 2). It is clear from the traces in Fig. 2 that the radiolabelled tryptic fragment obtained by digestion of each CNBr-cleavage peptide pool is eluted from the reverse-phase column at the same concentration of propan-2-ol. It is therefore likely that CNI/T, CNIII/T and CNIV/T are identical. This supposition was confirmed by determination of the amino acid composition of each peptide (results not shown). The explanation for the three original CNBr-cleavage peptide pools must be that the CNBr cleavage of the intact *arom* enzyme was incomplete,

Fig. 2. Final purification of the substrate-labelled peptides by reverse-phase chromatography on a Waters C₁₈ μ Bondapak column

The tryptic peptide pools obtained from *N. crassa arom* were purified on a reverse-phase column as described in the Materials and methods section. Peptide elution was monitored at 215 nm. Fractions were collected as shown, and radioactivity was detected by liquid-scintillation counting of samples from each fraction (sample volumes shown in parentheses). (a) Peptide CNI/T (25 μ l); (b) peptide CNIII/T (20 μ l); (c) peptide CNIV/T (10 μ l).

Fig. 3. Separation of peptides obtained by CNBr digestion of substrate-labelled *E. coli* 3-dehydroquinase by gel-filtration chromatography on a TSK G2000SW column

Chromatography was performed as described in the Materials and methods section. Peptide elution was monitored at 215 nm, and 10 μ l samples of fractions were liquid-scintillation-counted for radioactivity. Fractions from the column were pooled as indicated.

Fig. 4. Final purification of *E. coli* substrate-labelled peptides by reverse-phase h.p.l.c. on a Waters C₁₈ μ Bondapak column

The peptide pool EC2 (described in Fig. 3) was purified using a reverse-phase column as described in the Materials and methods section. The pooled fractions are indicated.

resulting in the production of partially digested CNBr-cleavage fragments. This may be due either to the poor accessibility to CNBr of some of the methionine residues in the very long *arom* polypeptide (subunit M_r 165000) or to oxidation of some methionine residues in the protein during its isolation.

The corresponding monofunctional *E. coli* 3-dehydroquinase was also radiolabelled with NaB^3H_4 in the presence of 3-dehydroquinase, as described in the Materials and methods section. Only CNBr digestion of this enzyme was performed. As the complete amino acid sequence of *E. coli* 3-dehydroquinase is available from the DNA sequence of the *aroD* gene, which encodes this enzyme (Duncan *et al.*, 1986; also the present paper), we were not only able to predict the exact number of CNBr-cleavage fragments that would be produced, but also their precise size and composition. We expected to find 11 peptides, ranging in size from two to 47 amino acid residues. The mixture

Table 1. Amino acid composition of the active-site peptide CNIV/T isolated from *N. crassa*

	Amino acid composition (mol of residue/mol)		
	Relative composition*	Nearest integer	Composition from sequence
Aspartic acid	1.09	1	1
Threonine	0.15	0	0
Serine	0.35	0	0
Glutamic acid	1.14	1	1
Proline	0.10	0	0
Glycine	2.20	2	2
Alanine	2.00	2	2
Valine	2.21	2	4
Isoleucine	0.05	0	0
Leucine	1.78	2	2
Histidine	1.05	1	1
Lysine	0.16	0	†
Arginine	0.97	1	†

*Normalized with respect to alanine = 2 residues.

†Not detected in the sequencer.

Table 2. Amino acid sequence of the active-site peptide CNIV/T isolated from *N. crassa*

Residue no.	Amino acid phenylthiohydantoin derivative identified	Radioactivity (c.p.m.)*
1	Ala	179
2	Leu	171
3	Gln	265
4	His	269
5	Gly	527
6	Asp	2103
7	Val	3235
8	Val	5009
9	†	43347
10	Leu	9098
11	Val	3791
12	Gly	2172
13	Val	1755
14	Ala	1527
15	†	1232

*Radioactivities of 10 μl samples were counted from the 140 μl fractions collected from the sequencer.

†No amino acid phenylthiohydantoin derivative could be identified.

Table 3. Amino acid sequence of the active-site peptide EC4 isolated from *E. coli*

Residue no.	Amino acid phenylthiohydantoin derivative identified	Yield (nmol)	Radioactivity (c.p.m.)
1	Gln	2.0	220
2	Ser	0.5	248
3	Phe	1.9	186
4	Asp	1.3	334
5	Ala	1.5	422
6	Asp	1.2	896
7	Ile	1.1	715
8	Pro	0.46	665
9	†	—	2364
10	Ile	0.005	896
11	Ala	0.002	495

*Samples of the fractions collected from the sequencer were liquid-scintillation-counted, in a manner similar to Table 2.

†No amino acid phenylthiohydantoin derivative could be identified.

(a)

```

1  ATGAAAACCG TAAGTGTAAA AGATCTCGTC ATTGGTACGG GCGCACCTAA
51  AATCATCGTC TCGCTGATGG CGAAGATAT CGCCAGCGTG AAATCCGAAG
101 CTCTCGCCTA TCGTGAAGCG GACTTTGATA TTCTGGAATG GCGTGTGGAC
151 CACTATGCCG ACCTCTCCAA TGTGGAGTCT GTCATGGCGG CAGCAAAAAT
201 TCTCCGTGAG ACCATGCCAG AAAAACCGCT GCTGTTTACC TTCCGCAGTG
251 CCAAAGAAGG CGGCGAGCAG GCGATTTCCA CCGAGGCTTA TATTGCACTC
301 AATCGTGCAG CCATCGACAG CGGCCTGGTT GATATGATCG ATCTGGAGTT
351 ATTTACCGGT GATGATCAGG TTAAGAAAC CGTCGCCTAC GCCCAGCGCG
401 ATGATGTGAA AGTAGTCATG TCCAACCATG ACTTCCATAA AACGCCGGAA
451 GCCGAAGAAA TCATTGCCCG TCTGCGCAAA ATGCAATCCT TCGACGCCGA
501 TATTCTTAAG ATTGCGCTGA TGCCGCAAA TACCAGCGAT GTGCTGACGT
551 TGCTTGCCCG GACCTGGAG ATGAGGAGC AGTATGCCGA TCGTCCAATT
601 ATCAGCATGT CGATGGCAAA AACTGGCGTA ATTTCTCGTC TGCTGGTGA
651 AGTATTTGGC TCGGCGGCAA CTTTGTGTGC GGTAAGAAAA GCGTCTGCGC
701 CAGGGCAAA CTGCGTAAAT GATTGCGCA CGGTATTAAC TATTTACAC
751 CAGGCATAA

```

(b)

```

1  MKTVTVKDLV IGTGAPKIV SLMAKDIAV KSEALAYREA DFDILEWRVD
51  HYADLSNVES VMAAKILRE TMPEKPLIFT FRSAKEGGEQ AISTEAYIAL
101 NRAAIDSLV DMIDLELFTG DDQVKETVAY AHADVKKVVM SNHDFHKTP
151 AEEIIARLRK MQSFDADIPK IALMPQSTSD VLTLLAATLE MQEQYADRPI
201 ITMSMAKTGV ISRLAGEVFG SAATFGAVKK ASAPGGISVN DLRTVLTLH
251 QA

```

Fig. 5. (a) Revised nucleotide sequence of the coding region of the *E. coli* K12 *aroD* gene and (b) revised protein sequence of *E. coli* 3-dehydroquinase

(a) The DNA was sequenced in both strands by using the dideoxy chain-termination method. (b) Residues different from or additional to those in the original sequence (Duncan *et al.*, 1986) are underlined. The revised M_r of the protein is 27466.

	1		48
<i>E. coli</i>	MKTVT.VKDL .VIGTGAPKI IVSLMAKDIA SVKSEALAYR EADFDILEWR		
<i>S. cerevisiae</i>	IATITGVREI EI .PSGRSA FVCLTFDDLT E.QTENLTPI CYGCEAVEVR		
<i>A. nidulans</i>	.AT.GQIDSL SIIKEKEHSF FASLTLPDLR EA.GDILEEV CVGSDAVALR		
	49		92
<i>E. coli</i>	VDHY..... ADLSNVESVM AAKILRETM PEKPLLFTR SAKEGGEQAI		
<i>S. cerevisiae</i>	VDHL..... ANYS.ADFVS KQLSILRKAT DSIPIIFTVR TMKQGGN..F		
<i>A. nidulans</i>	VDLLKDPASN NDIPSVDYVV EQLSFLRSRV T.LPIIFTIR TQSQGR..F		
	93		138
<i>E. coli</i>	STEAYIALNR A....AIDSG LVDMIDLELF TGDDQVKETV AYAHADVVKV		
<i>S. cerevisiae</i>	PDEEFKTL.R ELYDIALKNG .VEFLDLELT LPTDIQYEV I .NKRGNTKI		
<i>A. nidulans</i>	PDNAHDAAL. ELYRLAFRSG .CEFVDLDIA FPEDMLRAVT ..EMKGFSKI		
	139		186
<i>E. coli</i>	VMSNHDFHKT PEAE..IIA RLRKMQSFDA DIPKIALMPQ STSDVLTLLA		
<i>S. cerevisiae</i>	IGSHHDFQGL YSWDDAEWEN RFNQALTLDV DVVKFVGTA V NFEDNL.RLE		
<i>A. nidulans</i>	IASHHDPKGE LSWANMSWIK FYNKALEYG. DIIKLVGVAR NIDDN.TALR		
<i>N. crassa</i>		ALQHG. DVVKLVGVAR	
	187		236
<i>E. coli</i>	ATLEMQEQYA DRPIITMSMA KTGVISRLAG EVFGSAATFG AVKKASAPGQ		
<i>S. cerevisiae</i>	HFR...DTHK NKPLIAVNM T SKGSISRVLN NVLT.PVTS D LLPNSAAPGQ		
<i>A. nidulans</i>	KFKNWAAEAH DVPLIAINMG DQQLSRILN GFMT.PVSH P SLPFKAAPGQ		
	237	252	
<i>E. coli</i>	ISVNDLRTVL TILHQA.		
<i>S. cerevisiae</i>	LTVAQINKMY TSMGGIE		
<i>A. nidulans</i>	LSATEIRKGL SLMGEIK		

Fig. 6. Alignment of the protein sequences of biosynthetic dehydroquinases

The numbering scheme is that for the *E. coli* enzyme. The active-site lysine residue is at position 170 in this enzyme, and its position is marked by an asterisk (*). The active-site peptide obtained from the *N. crassa* arom protein in the present work is also shown. In the *S. cerevisiae* arom protein the 3-dehydroquinase and shikimate dehydrogenase domains are now contiguous, and there is no extra 'linker' region separating them as suggested earlier (Duncan *et al.*, 1987).

of CNBr-cleavage peptides was separated by h.p.l.c. gel-filtration chromatography (Fig. 3), followed by reverse-phase h.p.l.c. (Fig. 4). It is again likely that the three radioactive 215 nm-absorbing peaks represent partially digested CNBr-cleavage fragments.

Peptide sequence analysis

The amino acid composition of the *N. crassa* active-site peptide CNIV/T is given in Table 1. The sequence of this peptide is shown in Table 2, together with the radioactivity released at each sequencing cycle. At position 9 there was a very low yield of the phenylthiohydantoin derivative of unmodified lysine, accompanied by a large release of radioactivity. This led us to conclude that the lysine at this position was the site of the reduced Schiff base. The expected C-terminal arginine was barely detected by the sequencer, but it was clearly seen in the amino acid analysis (Table 1).

Of the three peptides isolated from *E. coli* 3-dehydroquinase, peptide EC4 was the most radioactive. EC4 was therefore chosen for sequence analysis. The result of the sequence determination is detailed in Table 3. Most of the radioactivity was released in cycle 9, and no standard amino acid phenylthiohydantoin derivative was observed for this cycle. Residue 9 must therefore be the position of the modified lysine.

DISCUSSION

NaBH₄ inhibits 3-dehydroquinase by reducing the Schiff-base

intermediate formed when the enzyme and substrate interact. This ability to trap 3-dehydroquinase at the 3-dehydroquinase active site has permitted the labelling of the active-site lysine residue by using NaB³H₄ in the reaction. Subsequent digestion of the labelled protein and isolation and characterization of labelled peptides has allowed us to determine the sequence around the active-site lysine residue in 3-dehydroquinase from two species, *N. crassa* and *E. coli*.

A complete nucleotide sequence for the *E. coli* *aroD* gene has been published (Duncan *et al.*, 1986). Subsequent DNA-sequencing work in our laboratory has revealed that this sequence contained errors at four locations. The corrected nucleotide sequence and the revised predicted protein sequence is given in Fig. 5. The corrected protein sequence differs in 24 amino acid residues from the original sequence, and in addition it contains 12 extra residues at the C-terminus. All the alterations are distant from the active-site lysine residue, which the experiments described in this paper identify as residue 170. The sequences of the biosynthetic 3-dehydroquinase component of *A. nidulans* (Charles *et al.*, 1985, 1986) and *S. cerevisiae* (Duncan *et al.*, 1987) arom polypeptides are also known, and these, together with the sequence of the *N. crassa* peptide CNIV/T determined in the present work, may be aligned with the revised *E. coli* sequence as shown in Fig. 6. Clearly all four sequences are related. One notable feature of the revised *E. coli* 3-dehydroquinase sequence is the absence of cysteine residues; the original sequence had predicted that the protein would contain three cysteine residues

(Duncan *et al.*, 1986). This means that the reduction and alkylation step used in the work-up of the active-site-modified *E. coli* enzyme was not necessary.

The availability of an overproducing strain of *E. coli* has made possible the isolation of 100 mg quantities of 3-dehydroquinase from this organism (Duncan *et al.*, 1986). We have recently improved enzyme yields considerably by constructing a second-generation overproducer, which relies not on copy number but on placement of *aroD* behind the inducible and very powerful *tac* promoter (K. Duncan, unpublished work). *E. coli* 3-dehydroquinase has also now been crystallized (K. Kleanthous, L. Sawyer & J. R. Coggins, unpublished work), and the work reported here on the location of the active-site lysine residue in the primary structure of the enzyme should, in the near future, allow the active site to be located within the tertiary structure.

This work was supported by grants from the Science and Engineering Research Council. It is a pleasure to acknowledge the advice of Dr. John Walker and Dr. Ian Fearnley (M.R.C. Laboratory of Molecular Biology, Cambridge) and Professor John Fothergill (University of Aberdeen), in whose laboratories the protein sequencing was carried out. We are grateful to Fred Northrop and Brian Dunbar for operating the protein sequencers, and to Emma Borthwick for assistance with the DNA sequencing.

REFERENCES

- Berlyn, M. B. & Giles, N. H. (1969) *J. Bacteriol.* **99**, 222–230
- Butler, J. R., Alworth, W. L. & Nugent, M. J. (1974) *J. Am. Chem. Soc.* **96**, 175–184
- Charles, I. J., Keyte, J. W., Brammar, W. J. & Hawkins, A. R. (1985) *Nucleic Acids Res.* **13**, 8119–8128
- Charles, I. J., Keyte, J. W., Brammar, W. J., Smith, M. & Hawkins, A. R. (1986) *Nucleic Acids Res.* **14**, 2201–2213
- Chaudhuri, S., Lambert, J. M., McColl, L. A. & Coggins, J. R. (1986) *Biochem. J.* **239**, 699–704
- Coggins, J. R., Boocock, M. R., Campbell, M. S., Chaudhuri, S., Lambert, J. M., Lewendon, A., Mousdale, D. M. & Smith, D. D. S. (1985) *Biochem. Soc. Trans.* **13**, 299–303
- Duncan, K., Chaudhuri, S., Campbell, M. S. & Coggins, J. R. (1986) *Biochem. J.* **238**, 475–483
- Duncan, K., Edwards, R. M. & Coggins, J. R. (1987) *Biochem. J.* **246**, 375–386
- Gaertner, F. H. & Cole, K. W. (1977) *Biochem. Biophys. Res. Commun.* **75**, 259–264
- Giles, N. H., Case, M. E., Baum, J., Geever, R., Huie, L., Patel, V. & Tyler, B. (1985) *Microbiol. Rev.* **49**, 338–358
- Jori, G., Galiazzo, G., Marzotto, A. & Scoffone, E. (1968) *J. Biol. Chem.* **243**, 4272–4278
- Koshiba, T. (1978) *Biochim. Biophys. Acta* **522**, 10–18
- Lambert, J. M., Boocock, M. R. & Coggins, J. R. (1985) *Biochem. J.* **226**, 817–829
- Liu, Y. S. V., Low, T. L. K. & Putnam, F. M. (1979) *J. Biol. Chem.* **254**, 2859–2864
- Lumsden, J. & Coggins, J. R. (1977) *Biochem. J.* **161**, 599–607
- Lumsden, J. & Coggins, J. R. (1978) *Biochem. J.* **169**, 441–444
- Mousdale, D. M., Campbell, M. S. & Coggins, J. R. (1987) *Phytochemistry* **26**, 2665–2670
- Polley, L. D. (1978) *Biochim. Biophys. Acta* **526**, 259–266
- Runswick, M. J. & Walker, J. E. (1983) *J. Biol. Chem.* **258**, 3081–3089
- Smith, D. D. S. & Coggins, J. R. (1983) *Biochem. J.* **213**, 405–415
- Smith, M. A., Gerrie, L. M., Dunbar, B. & Fothergill, J. E. (1982) *Biochem. J.* **207**, 253–260
- Walsh, C. (1979) *Enzymatic Reaction Mechanisms*, pp. 553–556, W. H. Freeman, San Francisco

Received 7 June 1990/15 August 1990; accepted 20 August 1990