Proc. Natl. Acad. Sci. USA
Vol. 78, No. 9, pp. 5381–5385, September 1981
Biochemistry

Primary structure of porcine heart citrate synthase

(sequence determination/citric acid cycle/mitochondrial enzyme/trimethyllysine)

DAVID P. BLOXHAM[†‡], DAVID C. PARMELEE[§¶], SANTOSH KUMAR[§], ROGER D. WADE[†], LOWELL H. ERICSSON[†], HANS NEURATH[†¶], KENNETH A. WALSH[†], AND KOITI TITANI[†§]

[§]Howard Hughes Medical Institute and [†]Department of Biochemistry, University of Washington, Seattle, Washington 98195

Contributed by Hans Neurath, May 26, 1981

ABSTRACT **The sequence of 437 amino acid residues of porcine heart citrate synthase [citrate oxaloacetate-lyase (*pro*-3*S*-CH_2COO→acetyl-CoA), EC 4. 1. 3. 7] has been determined by the alignment of fragments generated by cleavage with cyanogen bromide and with trypsin. Isolation of the peptides was facilitated by recent developments in the high-performance liquid chromatography of peptide mixtures. The alignment of these peptides was consistent with that previously deduced from fragments derived by restricted cleavage of citrate synthase by limited proteolysis and cleavage of aspartyl-prolyl bonds and asparaginyl-glycyl bonds. The enzyme contains a modified amino acid, trimethyllysine, at residue 368, showing that the enzyme is subjected to post-translational modification.**

Citrate synthase [citrate oxaloacetate-lyase (*pro*-3*S*-CH_2OO→acetyl-CoA), EC 4.1.3.7] is the most common of the enzymes involved in citrate cleavage/condensation reactions and is present in virtually all cells capable of oxidative metabolism. It has been isolated from numerous sources and in its most common form consists of a dimer of molecular weight 90,000–100,000 (1, 2). A tetrameric form with larger subunit molecular weight (55,000–60,000) is found in some prokaryotic organisms (3).

The most widely studied form of citrate synthase is that from porcine heart. The stereochemistry and mechanism of this enzyme have been studied in fine detail and reveal the following principal features.

(*i*) The absolute stereochemistry of citrate synthase involves the *si*-face attack of acetyl-CoA on the carbonyl of oxaloacetate, and the acetyl-CoA carbon atoms are incorporated into the pro-*S* carboxymethyl arm of citrate (4).

(*ii*) Condensation proceeds with the inversion of configuration of the remaining two methylene protons of acetyl-CoA (5).

(*iii*) The enzyme-catalyzed mechanism involves the generation and stabilization of a carbanion at the methyl group of acetyl-CoA and occurs in the presence of the second substrate or an appropriate analog such as *S*-malate (6).

(*iv*) The initial product of condensation is probably *S*-citryl-CoA and completion of the reaction requires enzyme-catalyzed hydrolysis of the ester (7). At present, it is not established whether there are one or two distinct active sites per subunit.

A 3.5-nm-resolution structure of the tetragonal crystal form of citrate synthase has been published (8); however, it was apparent that little further progress would be made without the amino acid sequence. This determination is the subject of the present communication. Conventional fragmentation techniques were used but a key development of the sequence work was the efficient exploitation of high-performance liquid chromatography (HPLC) procedures for peptide purification.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

Pig heart citrate synthase was either isolated by the method of Srere (9) or purchased from Sigma. The enzyme was homogeneous by $NaDodSO_4$ gel electrophoresis and had a specific activity of 140 units/mg of protein at 20°C.

Native enzyme was converted to [*carboxymethyl*-^{14}C] citrate synthase by reaction with iodo[1-^{14}C]acetic acid (10). Cleavages with cyanogen bromide at methionine and with nitrophenylsulfenylbromoskatole (BNPS-skatole) at tryptophan were performed by adaptations of the methods of Gross and Witkop (11) and Omenn *et al.* (12). Prior to tryptic cleavage, lysyl residues were modified with citraconic anhydride (13). Digestions with trypsin (tosyl-L-phenylalanyl chloromethyl ketone-treated, Worthington) were carried out for 1 hr at pH 8.0 and 37°C in the pH-stat. Enzymatic cleavages of glutamyl bonds with staphylococcal protease followed the procedure of Houmard and Drapeau (14).

Peptide mixtures were initially separated by chromatography on Sephadex G-50 SF in either 9% formic acid (cyanogen bromide peptides) or 0.1 M ammonium bicarbonate (tryptic peptides). Citraconyl groups were removed by acidification after lyophilization (15). Peptide mixtures were then dissolved in a small volume of 6 M guanidine·HCl and subjected to HPLC analysis. Waters' μBondapak-CN or $-C_{18}$ columns (300 × 4.9 mm) were equilibrated with 50 mM phosphoric acid titrated to pH 3.0 with 8 M NaOH (system 1) or 8.7 mM trifluoroacetic acid (system 2). Preparative experiments generally used flow rates of 1 ml/min and a maximal loading of 1–2 mg of peptide material. Gradient elution was performed with acetonitrile and a Varian 5000 HPLC unit. Peptide fractions were collected, pooled, and freeze-dried. When system 1 was used, the peptide was dissolved in 9% formic acid and desalted on a Sephadex G-25 column equilibrated with this same acidic solution.

For cleavage reactions that required a blocked amino terminus, peptides were initially succinylated by the method of Yaoi *et al.* (16). Cleavage at aspartyl-prolyl bonds was achieved by reaction in 70% formic acid containing 0.1% 2-mercaptoethanol at 38°C for 36 hr. Hydroxylamine cleavage was performed by the method of Bornstein (17). The conditions for limited proteolysis by subtilisin in the presence of palmitoyl-CoA were as described by Bloxham *et al.* (18).

Amino acid compositions were determined with a Dionex amino acid analyzer (model D500) according to the maker's operating instructions. Sequenator analysis was performed with a Beckman instrument (model 890B) according to the method of Edman and Begg (19) as modified by Brauer *et al.* (20). Phe-

Abbreviations: HPLC, high-performance liquid chromatography.

‡ On leave from the Department of Biochemistry, University of Southampton, Southampton, S09 3TU, England.
¶ Present address: Deutsches Krebsforschungszentrum, 6900 Heidelberg, Federal Republic of Germany.

nylthiohydantoin derivatives of amino acids were identified by HPLC (21).

RESULTS AND DISCUSSION

Outline of Sequence Strategy. The complete amino acid sequence of porcine heart citrate synthase is shown in Fig. 1; the amino acid composition is shown in Table 1. The protein consists of 437 residues and has a calculated subunit molecular weight of 48,987. Most estimates of subunit molecular weight by $NaDodSO_4$ gel electrophoresis have given values between 45,000 and 50,000 (1, 13, 18).

The experimental amino acid composition agrees well with that obtained from the sequence analysis. This composition indicates that cyanogen bromide cleavage or trypsin digestion of the citraconylated enzyme should release a theoretical maximum of 16 or 20 peptides, respectively. This presents a fairly daunting prospect in terms of the problems associated with purifying all these peptides for sequence determination. The recent development of HPLC techniques for peptide purification (22, 23) provides a powerful alternative to conventional purification techniques because of its high resolving power and speed. Our observations suggest that the following strategy for peptide purification may prove optimal. Initially, peptides are separated on the basis of molecular weight by gel filtration. For peptides larger than 30 residues, system 1 should be used; for smaller ones, system 2 is preferable. Alternatively, peptides larger than 60 residues may be purified by ion exchange chromatography in the presence of urea. This general approach has led to the rapid solution of citrate synthase structure.

To illustrate the successful application of the HPLC method, Fig. 2 shows the separation of two cyanogen bromide peptides containing the sequences 176–216 and 229–266. These two peptides could not be resolved by gel filtration or by a number of ion exchange chromatography systems. A feature of the HPLC system 1 was that single cyanogen bromide peptides eluted as multiple peaks. By sampling across the peaks we showed that fractions had identical amino acid compositions.

Table 1. Composition of porcine heart citrate synthase

Amino acid		Composition	Amino acid		Composition
Ala	A	33 (32.7 ± 1.5)	Leu	L	53 (48.1 ± 5)
Arg	R	19 (19.2 ± 0.8)	Lys	K	24 (25.4 ± 1.4)†
Asn	N	18 } (40.3 ± 2)	Met	M	15 (13.2 ± 0.6)
Asp	D	21 }	Phe	F	(12.5 ± 0.55)
Cys	C	4 (4.15 ± 0.32)	Pro	P	22 (21.7 ± 2.1)
Gln	Q	17 } (41.2 ± 0.8)	Ser	S	29 (27.3 ± 1.5)
Glu	E	24 }	Thr	T	(22.8 ± 0.76)
Gly	G	33 (34.9 ± 1.9)	Trp	W	9
His	H	14 (13.3 ± 1.1)	Tyr	Y	19 (17.4 ± 1.9)
Ile	I	19 (18.9 ± 0.32)	Val	V	28 (27.5 ± 1.5)
Me_3Lys	K*	1			

Molecular weight = 48,969; number of residues = 437. Values in parentheses were determined by amino acid analysis and are shown as mean ± SEM. Integral numbers are from the sequence.
† Analysis of whole protein did not resolve trimethyllysine from lysine.

Most of these multiple peaks probably represent interconversion between homoserine and homoserine lactone, partial deamidation, or conformational differences in the peptides.

Features of the Sequence Determination. The amino terminus of citrate synthase is unblocked and its sequence was determined on both the intact protein and appropriate large fragments such as the 37,000-dalton amino-terminal polypeptide liberated by subtilisin (18). The carboxyl-terminal sequence was identified by isolating a cyanogen bromide 12-residue peptide [426–437] which was devoid of homoserine. Determination of its sequence and amino acid composition indicated that lysine formed the carboxyl terminus. This was confirmed by a tryptic subdigestion of the peptide which released Leu-Val-Asp-Ser-Lys. The location of lysine at the carboxyl terminus is somewhat unusual and it is tempting to speculate whether it might have been derived from a tryptic-like proteolysis of the primary translation product of citrate synthase mRNA.

The remainder of the sequence was deduced by isolating 14

```
                10                  20                  30                  40
A S S T N L K D I L A D L I P K E Q A R I K T F R Q Q H G N T V V G Q I T V D M

                50                  60                  70                  80
M Y G G M R G M K G L V Y E T S V L D P D E G I R F R G Y S I P E C Q K M L P K

                90                 100                 110                 120
A K G G E E P L P E G L F W L L V T G Q I P T E E Q V S W L S K E W A K R A A L

               130                 140                 150                 160
P S H V V T M L D N F P T N L H P M S Q L S A A I T A L N S E S N F A R A Y A E

               170                 180                 190                 200
G I H R T K Y W E L I Y E D C M D L I A K L P C V A A K I Y R N L Y R E G S S I

               210                 220                 230                 240
G A I D S K L D W S H N F T N M L G Y T D A Q F T E L M R L Y L T I H S D H E G

               250                 260                 270                 280
G N V S A H T S H L V G S A L S D P Y L S F A A A M N G L A G P L H G L A N Q E

               290                 300                 310                 320
V L V W L T Q L Q K E V G K D V S D E K L R D Y I W N T L N S G R V V P G Y G H

               330                 340                 350                 360
A V L R K T D P R Y T C Q R E F A L K H L P H D P M F K L V A Q L Y K I V P N V

               370*                380                 390                 400
L L E Q G K A K N P W P N V D A H S G V L L Q Y Y G M T E M N Y Y T V L F G V S

               410                 420                 430
R A L G V L A Q L I W S R A L G F P L E R P K S M S T D G L I K L V D S K
```

FIG. 1. Amino acid sequence of porcine heart citrate synthase. The one-letter code is defined in Table 1.

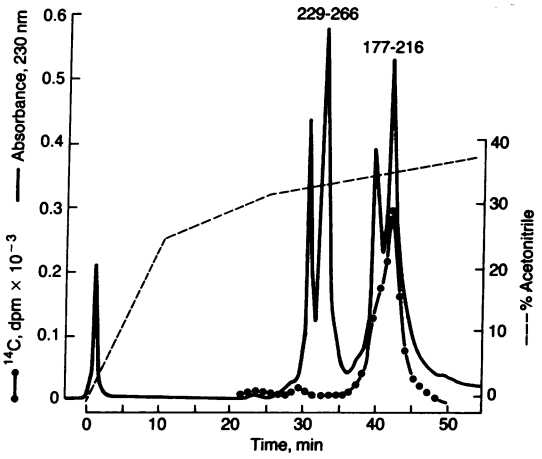

FIG. 2. Chromatography of 680 μg of cyanogen bromide peptide mixture from citrate synthase by HPLC with system 1. The radioactivity present was determined by using 5% of the sample.

cyanogen bromide and 17 tryptic peptides and determining their sequences. The peptides isolated and the actual residues determined by automated sequence analysis are shown in Fig. 3. In general, the majority of amino acids were located by sequence analysis on the primary peptides, and large sequence overlaps were obtained for the two classes of peptides. However, for certain parts of the sequence, additional experiments were required as discussed below.

The sequence in the region 94–117 was resolved by BNPS-skatole degradation (12) at tryptophans in a tryptic peptide containing the sequence 68–117. The peptides isolated corresponded to residues 68–94, 95–109, 110–114, and 115–117. Because of incomplete cleavage by BNPS-skatole, an overlap peptide containing residues 110–117 also was isolated and analyzed. The carboxyl-terminal peptide was identified by the observation of the arginine residue of the original tryptic peptide. To establish the overlap between Arg 117–Ala 118, the cyanogen bromide peptide corresponding to the sequence 78–127 was subjected to digestion by staphylococcal protease at glutamate residues (14). Two subpeptides were isolated corresponding to residues 106–113 and 114–127. The first peptide was used to confirm tryptophan-109 which could not be identified directly in the corresponding BNPS-skatole-cleaved peptide (residues 95–109). The second peptide contained the sequence Trp-Ala-Lys-Arg-Ala-Ala-Leu-Pro and confirmed the expected overlap at residues 117–118.

For determination of the sequence between residues 257 and 267, a tryptic peptide (residues 230–313) containing this sequence was blocked by succinylation and the peptide was cleaved at the single aspartyl-prolyl bond (residues 257–258) by acid cleavage. An overlap through residue 267 was established by sequence analysis of the newly exposed amino terminus.

To locate unambiguously the cyanogen bromide peptide containing residues 347–390 in the sequence, a large tryptic peptide (residues 335–413) was cleaved by cyanogen bromide and the sequence of the total mixture was analyzed. This revealed three major sequences—Glu-Phe-Ala-Leu-etc., Phe-Lys-Leu-Val-etc., and Asn-Tyr-Tyr-Thr-etc.—as well as the minor sequence Thr-Glu-Hse (<10% of total yield). Because the sequence was overlapped through the first and last series of residues in the major peptides, the peptide containing residues 347–390 was placed in the center of the tryptic peptide. The complete sequence 347–390 was determined after tryptic diges-

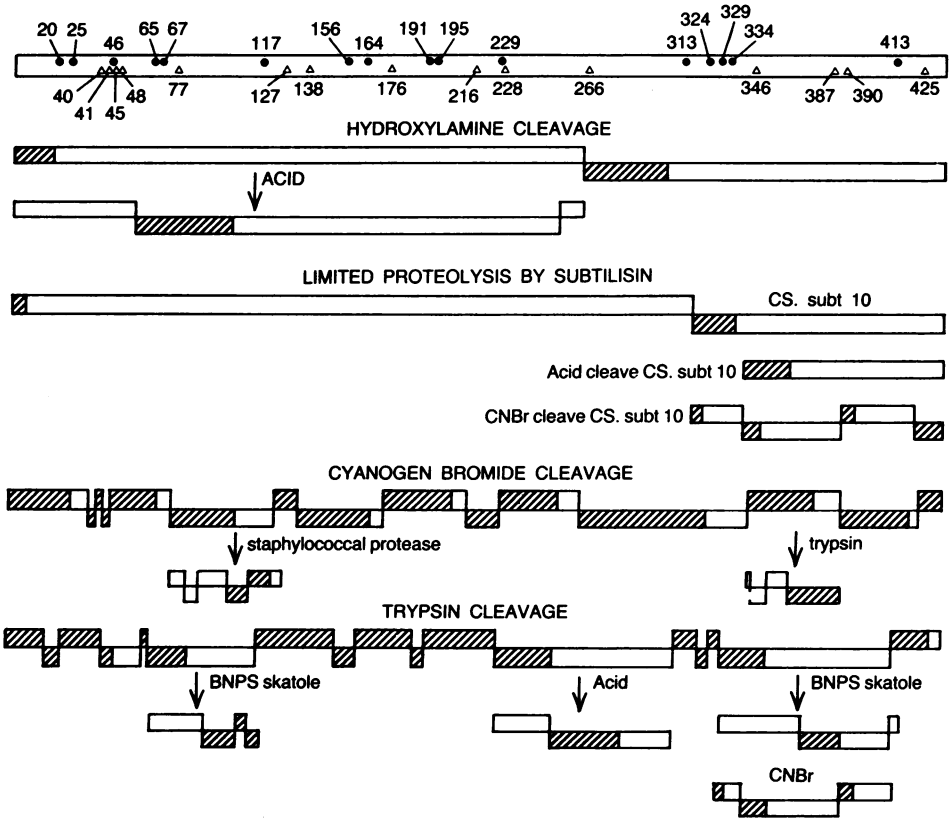

FIG. 3. Summary of fragments generated for the sequence analysis of citrate synthase. The top bar represents the whole protein and indicates the location of cleaved arginine residues (●) and methionine residues (△). The hatched section of each bar represents the portion of the sequence determined by automated Edman degradation.

tion which released four peptides containing the sequences 347–348, 349–355, 356–366, and 367–390. The last peptide contained no lysine but did contain homoserine, confirming that it was the carboxyl-terminal segment of the starting peptide. Determination of this peptide sequence showed two unexpected features. First, the second turn of the Edman degradation [residue 368] showed an unusual phenylthiohydantoin amino acid. Second, the carboxyl-terminal sequence was Gly-Hse-Thr-Glu-Hse, showing the presence of two homoserine residues in the sequence. Presumably the homoserine at residue 387 arises from the cyanogen bromide reaction with a methionine at this position. However, cleavage does not occur in the usual way because of a competing reaction with the hydroxyl group on the adjacent residue threonine-388 (24, 25). This explains the poor yield of a cyanogen bromide peptide with the sequence Thr-Glu-Hse corresponding to 388–390.

Trimethyllysine at Residue 368. Sequence analysis of the tryptic peptide corresponding to residues 367–390 indicated the presence of an unusual basic amino acid at residue 368. Comparison of the retention time on HPLC of the unknown with numerous standards, indicated that the derivative was the phenylthiohydantoin of trimethyllysine. Therefore, amino acid analyses by a system that resolved trimethyllysine from lysine (26) were repeated on two peptides (residues 367–390 and 347–390) that contained residue 368. These analyses established the presence of one residue of trimethyllysine per subunit. The peptide containing trimethyllysine was formed by tryptic cleavage, indicating that the bond containing the modified lysine is resistant to tryptic attack. A similar observation was made by DeLange *et al.* (30) on the tryptic formation of peptide T9 from *Neurospora* cytochrome *c*.

The identification of trimethyllysine has previously been restricted to the calmodulins (27), a ribosomal protein (28), histone III (29), and several species of cytochrome *c* (26, 30). Trimethyllysine has never been observed in an animal cytochrome (30). The formation of this residue must occur as a post translational modification involving methyl transfer from *S*-adenosylmethionine (31). The sequences that are modified in cytochrome *c* contain two adjacent lysine residues (29) and, interestingly, the nascent sequence in citrate synthase would contain two lysine residues in the tripeptide sequence Lys-Ala-Lys at residues 366–368. The presence of a small alanine side chain may render this sequence functionally equivalent to the dilysyl sequence recognized in the modification of cytochrome *c*. Accessibility to the methylating enzyme system is an important factor because the identical sequence, Lys-Ala-Lys at 80–82 is unmodified.

Tryptic Fragmentation of Citraconyl-Citrate Synthase. Three bonds containing arginine, at residues 302, 401, and 421, were resistant to tryptic proteolysis. Failure to cleave the arginyl-aspartyl bond at 302–303 is not unexpected because the presence of an adjacent acidic residue in the scisscile bond is known to reduce the rate of tryptic hydrolysis (32, 33). The failure of trypsin to cleave bond 421–422 is also consistent with the specificity of trypsin because arginyl-prolyl bonds are virtually resistant to proteolysis (33, 34). There is no clear reason why trypsin should fail to cleave at arginine-401, especially when an identical sequence was hydrolyzed just 12 residues away. The only explanation here is that a physical effect may have prevented the bond from becoming exposed to trypsin. Consistent with this suggestion, the tryptic peptide containing this region, residues 335–413, was always isolated as a large aggregate on gel filtration.

Cysteine Residues. The sequence of citrate synthase shows that there are four cysteine residues per subunit, located at residues 74, 175, 184, and 332. This is consistent with the amino acid analysis showing a total of 4.2 carboxymethylcysteine residues per subunit (Table 1). Titration with 5,5′-dithiobis(2-nitrobenzoate) under denaturing conditions revealed 3.65 free thiols per subunit (2), indicating that there are no disulfide bridges in the protein.

Throughout the analytical work, the location of cysteine residues in peptides was monitored by the presence of [^{14}C]carboxymethyl groups. Gel filtration on Sephadex G-50 of the cyanogen bromide digest revealed three radioactive peaks distributed in the ratio 1:1:2. The last fraction was resolved into two distinct peptides by HPLC. These peptides accounted for the total radioactivity and proved that there were four cysteine residues. The pattern of radioactivity distribution in the tryptic peptides of [*carboxymethyl*-^{14}C] citrate synthase confirmed these results. In this case a peptide containing residues 165–191 contained two carboxymethylcysteine residues.

Sequence Overlaps. Fig. 3 shows that the majority of sequence overlaps were unambiguously established, except at residues 45–46 and 228–229 due to the presence of the sequence Met-Arg (Fig. 1) in both places. However, the sequences from 1–45 and from 229–437 were completely overlapped and identified as the amino and carboxyl termini, respectively. Therefore, there is only one solution possible in placing the segments 1–45, 46–228, and 229–437 in sequence. A cyanogen bromide peptide containing the sequence Arg-Gly-Met (residues 46–48) was isolated, establishing the presence of the sequence Met-Arg-Gly-Met (residues 45–48) in the protein. Despite this finding, it is conceivable that a small peptide could be omitted in the regions of minimal overlap.

Limited Proteolysis of Citrate Synthase. In the presence of palmitoyl-CoA, subtilisin cleaves citrate synthase into an amino-terminal 37,000-dalton fragment and a 10,000- to 12,000-dalton (CS.subt 10) carboxyl-terminal fragment (18). The main site of subtilisin cleavage can now be identified as Ala-Val at 321–322. The correct amino-terminal sequence of the small fragment is Val-Leu-Arg-Lys-Thr-Asp-Pro-Arg-Tyr. The presence of arginine in this sequence was omitted in the original report because of difficulties of interpreting the staggered sequence produced by "ragged" subtilisin cleavage. The isolation of this fragment and the determination of its amino-terminal sequence was essential to place the tryptic peptides 325–329, 330–334, and 335–413 in their correct order (Fig. 3). The determination of the site of cleavage by subtilisin means that the extended binding site can be identified as Gly(P_3)-His(P_2)-Ala(P_1)-Val(P_1')-Leu(P_2') [see nomenclature of Schecter and Berger (35)]. The presence of glycine at P_3, which gives a large rate enhancement in synthetic peptides, and valine at P_1' is reasonably consistent with predictions from model peptide studies (36). The presence of histidine at P_2 is unexpected because optimal hydrolysis with synthetic peptides is obtained with alanine in this position and histidine reduces the rate of proteolysis.

Acid Cleavage of Citrate Synthase. There are four aspartyl-prolyl bonds in citrate synthase, at positions 59–60, 257–258, 327–328, and 344–345. Cleavage of the protein at these bonds (37) in acid was incomplete, and analysis of the products by $NaDodSO_4$ gel electrophoresis yielded a complex mixture of polypeptides. The main polypeptides had molecular weights of 41,000 (residues 60–437), 36,500 (1–327), 28,500 (1–257; 60–327), 22,500 (60–257), 10,500 (345–437; 258–344), and 7000 (1–59; 258–327). The two largest polypeptides disappeared as the cleavage reaction time was extended. The molecular weights of all the polypeptides agree well with the sequence shown in Fig. 1.

Hydroxylamine Cleavage of Citrate Synthase. Cleavage of citrate synthase by hydroxylamine at the single asparaginyl-glycyl bond liberated two major polypeptides in equimolar

amounts with molecular weights of 27,000–29,000 (NGα) and 17,000–19,000 (NGβ), respectively. Both polypeptides were purified by chromatography on Sephacryl S-200 in 0.1 M ammonium bicarbonate/6 M guanidine·HCl. The larger fragment contained the native amino-terminal sequence whereas the smaller polypeptide contained the amino terminus, Gly-Leu-Ala-Gly-Pro-etc. This confirms the single asparginyl-glycyl bond at 267–268.

Acid Cleavage of NGα (1–267). The NGα fragment contains two aspartyl-prolyl bonds at 59–60 and 257–258. These bonds were readily cleaved by acid which resulted in the formation of two main polypeptides of molecular weights *ca.* 7000 (residues 1–59) and 20,000–21,000 residues (60–257). A peptide containing the sequence 258–267 is too small to be detected under our conditions of $NaDodSO_4$/polyacrylamide gel electrophoresis. The molecular weights of the cleavage fragments from NGα are in complete accord with the sequence of the amino-terminal 267 residues (Fig. 1). This finding is important because it makes it unlikely that undetected peptides are present in the regions of minimal overlap (residues 45–46 and 228–229).

Correlation of Amino Acid Sequence with Crystallographic Results. The amino acid sequence determined in the present work has been used to refine the crystallographic model of citrate synthase (8). Although it is not possible to identify all 437 residues in the improved crystallographic model, those that can be located (≈410) are completely consistent with the chemical sequence (S. J. Remmington and R. Huber, personal communication). Most important, the crystallographic sequence has confirmed the presence of two stretches of sequence: Met-Met-Tyr-Gly-Gly-Met-Arg-Gly-Met- and Leu-Met-Arg-Leu-Tyr-Leu. Both of these sequences confirm the minimal overlaps in the two Met-Arg regions of the protein and do not require the insertion of any additional sequence.

The authors thank Richard Granberg for performing the amino acid analyses. We are indebted to Prof. R. Huber and Dr. S. J. Remmington for informing us of the crystallographic results prior to publication. This work has been supported by research grants from the National Institutes of Health (GM-15731) and the Wellcome Trust.

1. Srere, P. A. (1975) *Adv. Enzymol. Relat. Areas Mol. Biol.* **47**, 57–101.
2. Singh, M., Brooks, G. C. & Srere, P. A. (1970) *J. Biol. Chem.* **245**, 4636–4640.
3. Weitzman, P. D. J. & Danson, M. J. (1976) *Curr. Top. Cell. Regul.* **10**, 161–204.
4. Hanson, K. R. & Rose, I. A. (1963) *Proc. Natl. Acad. Sci. USA* **50**, 981–988.
5. Lenz, H., Buckel, W., Wunderwald, P., Biedermann, G., Buschmeir, V., Eggerer, H., Cornforth, J. W., Redmond, J. W. & Mallaby, R. (1971) *Eur. J. Biochem.* **24**, 207–215.
6. Eggerer, H. (1965) *Biochem. Z.* **343**, 111–137.
7. Eggerer, H. (1963) *Biochem. Z.* **337**, 202–223.
8. Wiegand, G., Kukla, D., Scholze, H., Jones, T. A. & Huber, R. (1979) *Eur. J. Biochem.* **93**, 41–50.
9. Srere, P. A. (1969) *Methods Enzymol.* **13**, 3–11.
10. Crestfield, A. M., Moore, S. & Stein, W. H. (1963) *J. Biol. Chem.* **238**, 622–627.
11. Gross, E. & Witkop, B. (1962) *J. Biol. Chem.* **237**, 1856–1860.
12. Omenn, G. S., Fontana, A. & Anfinsen, C. B. (1970) *J. Biol. Chem.* **245**, 1895–1902.
13. Dixon, H. B. F. & Perham, R. N. (1968) *Biochem. J.* **109**, 312–314.
14. Houmard, J. & Drapeau, G. R. (1972) *Proc. Natl. Acad. Sci. USA* **69**, 3506–3509.
15. Atassi, M. A. & Habeeb, A. F. S. A. (1972) *Methods Enzymol.* **25**, 546–553.
16. Yaoi, Y., Titani, K. & Narita, K. (1964) *J. Biochem. (Tokyo)* **56**, 222–229.
17. Bornstein, P. (1970) *Biochemistry* **9**, 2408–2421.
18. Bloxham, D. P., Ericsson, L. H., Titani, K., Walsh, K. A. & Neurath H. (1980) *Biochemistry* **19**, 3979–3985.
19. Edman, P. & Begg, G. (1967) *Eur. J. Biochem.* **1**, 80–91.
20. Brauer, A. W., Margolies, M. N. & Haber, E. (1975) *Biochemistry* **14**, 3029–3035.
21. Ericsson, L. H., Wade, R. D., Gagnon, J., McDonald, R. M., Granberg, R. & Walsh, K. A. (1977) in *Solid Phase Methods in Protein Sequence Analysis*, ed. Previero, A. & Coletti-Previero, M. A. (Elsevier, Amsterdam), pp. 137–142.
22. Gerber, G. E., Anderegg, R. J., Herlihy, H. C., Gray, C. P., Biemann, K. & Khorana, H. G. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 217–231.
23. Mahoney, W. C. & Hermodson, M. A. (1980) *J. Biol. Chem.* **255**, 11199–11203.
24. Waxdal, M. J., Konigsberg, W. H., Henley, W. L. & Edelman, G. M. (1968) *Biochemistry* **7**, 1959–1966.
25. Schroeder, W. A., Shelton, J. B. & Shelton, J. R. (1969) *Arch. Biochem. Biophys.* **130**, 551–556.
26. VanEldik, L. J., Grossman, A. R., Iverson, D. B. & Watterson, D. M. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 1912–1916.
27. Vanaman, T. C., Sharief, F. & Watterson, D. M. (1977) in *Calcium Binding Proteins and Calcium Function*, eds. Wasserman, R. H., Corradino, R. A., Carafoli, E., Kretsinger, R. H., MacLennan, D. H. & Siegel, F. L. (North-Holland, New York), pp. 106–116.
28. Dognin, M. J. & Wittman-Liebold, B. (1977) *FEBS Lett.* **84**, 342–346.
29. Vickery, H. B. (1972) *Adv. Protein Chem.* **26**, 81–171.
30. Delange, R. J., Glazer, A. N. & Smith, E. L. (1970) *J. Biol. Chem.* **245**, 3325–3327.
31. Paik, W. K. & Kim, S. (1970) *J. Biol. Chem.* **245**, 6010–6015.
32. Hirs, C. H. W., Moore, S. & Stein, W. H. (1956) *J. Biol. Chem.* **219**, 623–642.
33. Koningsberg, W. H. & Steinman, H. M. (1977) in *The Proteins*, eds. Neurath, H. & Hill, R. L. (Academic, New York), pp. 1–178.
34. Tsugita, A., Gish, D. T., Young, J., Fraenkel-Conrat, H., Knight, C. A. & Stanley, W. M. (1960) *Proc. Natl. Acad. Sci. USA* **46**, 1463–1469.
35. Schechter, I. & Berger, A. (1967) *Biochem. Biophys. Res. Commun.* **27**, 157–162.
36. Morihara, K., Oka, T. & Tsuzuki, H. (1969) *Biochem. Biophys. Res. Commun.* **35**, 210–214.
37. Piszkiewicz, D., Landon, M. & Smith, E. L. (1970) *Biochem. Biophys. Res. Commun.* **40**, 1173–1178.