Eur. J. Biochem. 187, 235 – 239 (1990)
© FEBS 1990

The 2-oxoglutarate dehydrogenase complex from *Azotobacter vinelandii*

2. Molecular cloning and sequence analysis of the gene encoding the succinyltransferase component

Adrie H. WESTPHAL and Arie de KOK

Department of Biochemistry, Agricultural University, Wageningen, The Netherlands

(Received April 13/August 9, 1989) – EJB 89 0467

The nucleotide sequence encoding the succinyltransferase component (E2o) of the 2-oxoglutarate dehydrogenase complex from *Azotobacter vinelandii* has been determined. Previously the cloning in *Escherichia coli* of the gene encoding lipoamide dehydrogenase from *A. vinelandii* was reported [Westphal, A. H. & de Kok, A. (1988) *Eur. J. Biochem. 172*, 299 – 305]. The 3.2-kb fragment used for the sequence determination contained the main part of the gene encoding succinyltransferase. The complete E2o gene, as well as the gene encoding the 2-oxoglutarate dehydrogenase component, resided on a 14.7-kb fragment from which the 3.2-kb fragment was subcloned.

The protein-coding sequence of the gene consists of 1200 bp (400 codons, including the AUG start codon and the UGA stop codon). It is separated from the gene encoding the 2-oxoglutarate dehydrogenase component by 42 bp. No *E. coli*-like promoter sequence was found. A putative ribosome-binding site is located 9 – 15 bp upstream from the start codon. No terminator sequences were found downstream of the stop codon. This makes it likely that the three genes of the oxoglutarate dehydrogenase complex are transcribed as a single mRNA transcript analogous to the pyruvate dehydrogenase complex in *E. coli*. The intact gene was subcloned from the 14.7-kb fragment and brought to high expression under the influence of the vector-encoded *lacZ* promotor.

The similarity with the *E. coli* enzyme is high with 63% identity. Like the enzyme from *E. coli*, it consists of a single lipoyl-binding domain, a putative E1- and E3-binding domain and a catalytic domain. The main difference is found in a 31-residue sequence rich in alanine and proline located between the lipoyl domain and the putative E1- and E3-binding domain. This sequence, usually found in acetyltransferases and there identified as a highly mobile region by ^{1}H-NMR, is replaced by a more polar, charged region in the *E. coli* enzyme.

Succinyl-CoA dihydrolipoate *S*-succinyltransferase, or succinyltransferase (E2o), is the core component of the 2-oxoglutarate dehydrogenase complex [1]. Succinyltransferase and acetyltransferase, the core component of the analogous functioning pyruvate dehydrogenase complex, have many properties required for the function of these multienzyme complexes in common. Both enzymes contain at the N-terminal part one or more lipoyl domains, each of which covalently binds a lipoyl group [2, 3]. These lipoyl domains form flexible arms which transport acyl groups and/or reducing equivalents between the various active sites. Whereas acetyltransferases are reported to contain one, two or three lipoyl domains [4 – 7], succinyltransferases seem to contain only one lipoyl domain [8]. A functional reason for this diversity is not known. It has been demonstrated that removal of two of the three lipoyl domains from the acetyltransferase from *Escherichia coli* has no effect on its function in the intact complex [9].

Both enzymes form oligomers. Acetyltransferases from Gram-negative bacteria contain 24 subunits [10, 11], while those from Gram-positive bacteria or eukaryotic origin contain 60 subunits [12]. Succinyltransferases seem to contain invariably 24 subunits [13, 14]. The information for the quaternary structure, as well as the transacylase active site, is located at the C-terminal part of the peptide chain (called the catalytic domain) [15, 16].

Both enzymes contain binding sites for the two peripheral components, 2-oxoglutarate dehydrogenase (E1o) and pyruvate dehydrogenase (E1p) and the common component lipoamide dehydrogenase (E3). In *E. coli* E2p and E2o, different regions for the binding of the E1 and E3 components are proposed [17]. The E3 component is thought to bind on a folded yet probably mobile region between the lipoyl domain and the catalytic domain, whereas the E1 component is thought to bind on the catalytic domain. In *Azotobacter vinelandii* E2p the region between the lipoyl domain and the catalytic domain is indicated as the binding site for both components [11, 16]. In *B. stearothermophilus* E2p and bovine kidney E2p [17a, 18] the same region is proposed for binding both components.

Correspondence to A. de Kok, Department of Biochemistry, Agricultural University, Dreijenlaan 3, 6703 HA Wageningen, The Netherlands

Enzymes. 2-Oxoglutarate dehydrogenase (E1o), 2-oxoglutarate:lipoate oxidoreductase (EC 1.2.4.2); succinyltransferase (E2o), succinyl-CoA:dihydrolipoate *S*-succinyltransferase (EC 2.3.1.61); lipoamide dehydrogenase (E3), NADH:lipoamide oxidoreductase (EC 1.8.1.4); calf intestinal alkaline phosphatase (EC 3.1.3.1); DNA polymerase I (Klenow fragment) (EC 2.7.7.7); restriction endonucleases: *Sau*3A, *Sma*I, *Cla*I, *Stu*I, *Pst*I, *Dra*I, *Kpn*I (EC 3.1.24.4).

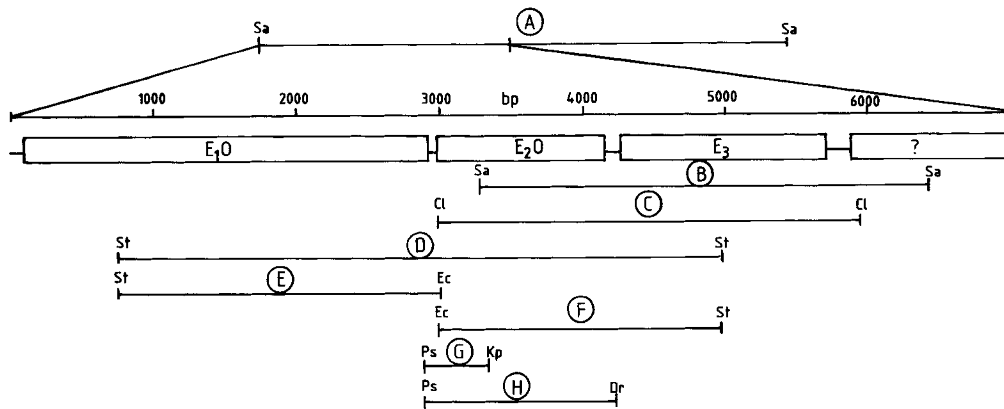

Fig. 1. *Organization of the genes encoding the E1o, E2o and E3 component of the 2-oxoglutarate dehydrogenase complex, and DNA fragments used for cloning and sequencing of the E2o gene.* (A) 14.7-kb fragment. (B) 3.2-kb fragment used for sequencing the lipoamide dehydrogenase gene and part of the succinyltransferase gene. (C) 3-kb *Cla*I fragment. (D) 4.2-kb *Stu*I fragment. (E) 2.23-kb *Stu*I – *Eco*RV fragment. (F) 1.98-kb *Eco*RV – *Stu*I fragment. (G) 438-bp *Pst*I – *Kpn*I fragment. (H) 1.33-kb *Pst*I – *Dra*I fragment used for constructing plasmid pAE2. Abbreviations for restriction sites are: Cl for *Cla*I, Dr for *Dra*I, Ec for *Eco*RV, Kp for *Kpn*I, Ps for *Pst*I, Sa for *Sau*3A and St for *Stu*I

A third class of multienzyme complexes, the branched-chain complexes, contain an acetyltransferase which has many properties in common with succinyltransferase [19, 20].

Thus far, the gene from *E. coli* is the only succinyltransferase gene that has been cloned and sequenced [8]. It forms a cluster with the gene encoding the 2-oxoglutarate dehydrogenase component and it has been shown that both genes are transcribed into a single transcript [21].

The gene for lipoamide dehydrogenase from *A. vinelandii* has been cloned and sequenced [22]. This gene does not form a cluster with the other genes of the pyruvate dehydrogenase complex, as observed for *E. coli* [21], but belongs to the cluster of the 2-oxoglutarate dehydrogenase complex genes. As a result, a part of the succinyltransferase gene was found on a 3.2-kb fragment used for sequence analysis of the E3 gene [22]. In this paper we report the cloning and the nucleotide sequence determination of the complete gene.

MATERIALS AND METHODS

Materials

Materials, bacterial strains and vectors were given in the preceding paper [23]. Succinyl-lipoamide was prepared from reduced lipoamide essentially as described for the synthesis of succinyl-CoA [24].

Sources of DNA

The previously reported sequence of *A. vinelandii* lipoamide dehydrogenase [22] was determined by sequencing a 3.2-kb fragment. This fragment was subcloned from a plasmid containing a 14.7-kb insert (pAW77) which was obtained after partial digestion of *A. vinelandii* chromosomal DNA with *Sau*3A.

The 3.2-kb fragment also contained a major part of the succinyltransferase gene (E2o gene). The original 14.7-kb fragment was then used for the cloning of the E2o gene and the complete determination of the nucleotide sequence.

Southern blotting

Restriction endonuclease digests of the 14.7-kb *A. vinelandii* DNA fragment were run on 0.9% high-gelling-temperature agarose gels in 40 mM Tris/acetate, 2 mM EDTA, pH 8.0 [25] and transferred to nitrocellulose using the blotting procedure of Southern [26]. Universal sequencing primer was annealed to single-stranded M13 DNA containing a part of the E2o gene obtained during sequencing the *A. vinelandii* lipoamide dihydrogenase gene [22]. Next the primer was extended using dGTP, dCTP, dTTP, [α-^{32}P]dATP and Klenow fragment. Excess label was removed by the spun column method [25]. Hybridization was carried out for 16 h at 65 °C using solutions described in [25].

Nucleotide sequence determination and analysis

The major part of the E2o gene situated on the 3.2-kb fragment was fragmentated by sonication, end-repaired and size-fractionated. Fragments of 0.4 – 0.6 kb were blunt-end ligated into the *Sma*I site of M13mp19 as described in [22]. Fragments of the other part of the gene were isolated after digestion of the 14.7-kb fragment with various restriction enzymes and ligated in the multiple cloning site of M13mp18 or M13mp19. After transformation of the recombinant phages and single-stranded DNA isolation, the phages were sequenced using the dideoxy-chain-termination method of Sanger et al. [27] as described in [28]. The sequencing reaction was carried out at 50 °C. In some cases instead of dGTP, 7-deaza-dGTP was used. The data were compiled with the aid of a VAX computer using the programs of Staden [29, 30].

Other techniques

Enzyme assay. Succinyltransferase activity was assayed spectrophotometrically at 340 nm by measuring the amount of NADH formed from NAD^+ (1 mM) in the presence of succinyl-lipoamide (1 mM), CoA (0.16 mM) and *A. vinelandii* lipoamide dehydrogenase (1 unit) according to the following reaction:

$$\text{Succinyl-S-lip-SH} + \text{CoA-SH} + NAD^+ \rightleftharpoons \text{Succinyl-S-CoA} + \text{lipS}_2 + NADH + H^+ .$$

The reaction was carried out in 50 mM potassium phosphate pH 7.0 at 25 °C.

Enzyme isolation. The succinyltransferase component was purified from transformed *E. coli* TG2 (pAE2) cells as follows.

A 20-l culture grown in TY medium [31] containing 50 μg/ml ampicillin and 10 μg/ml isopropyl β-D-thiogalactoside was harvested after 40 h. Cell-free extract was obtained after passage of the cells (100 g), suspended in 250 ml 50 mM potassium phosphate pH 7.0, 3 mM EDTA and 0.1 mM phenylmethylsulfonyl fluoride, through a Manton-Gaulin laboratory homogenizer at 50 MPa followed by centrifugation (20 min, $30000 \times g$). Protamine sulfate solution (2%) was added to a concentration just before precipitation of the succinyltransferase occurred. The solution was centrifuged (20 min, $30000 \times g$) and the enzyme was precipitated from the supernatant by adding poly(ethylene glycol) 6000 to a final concentration of 10%. The pellet was dissolved in 50 ml 10 mM potassium phosphate pH 7.0, 1 mM EDTA, 0.1 mM phenylmethylsulfonyl fluoride and applied to a DEAE-Sephacel column (4×30 cm). The protein was eluted with a potassium chloride gradient and active fractions were pooled. Next the protein was concentrated by poly(ethylene glycol) precipitation (final concentration 10%) and loaded on a Sephacryl S400 HR column (2.5×120 cm). Active fractions were then concentrated on an Amicon YM 100 filter.

RESULTS AND DISCUSSION

Cloning strategy

Southern blots of restriction digests of the 14.7-kb fragment were made using a M13 single-stranded DNA probe containing an insert coding for the C-terminal part of E2o (obtained upon sequencing the gene encoding the lipoamide dehydrogenase component [22]). A strong hybridization was found with a 3-kb *Cla*I fragment (Fig. 1, fragment C). One end of this fragment could be located 245 bp downstream of the stop codon of the E3 gene, the other end could be located 6 bp downstream of the start codon of the E2o gene. Translation of this part of the gene yielded a protein sequence which was 100% identical with amino acid residues 3–10 of the native protein determined by automated Edman degradation. This fragment was then used as a probe and strong hybridization was found with a 4.2-kb *Stu*I fragment (Fig. 1, fragment D). It showed that one end of this fragment was located 404 bp downstream of the start codon of the E3 gene and the other end 651 bp downstream of the start codon of the E1o gene, as judged from the similarity with the *E. coli* E1o gene. This fragment thus contained the complete *A. vinelandii* E2o gene. Next this 4.2-kb fragment was digested with *Eco*RV which cuts once 4 bp downstream of the N-terminal *Cla*I site already found. This yielded two fragments (Fig. 1, fragments E and F) which were both ligated into M13mp19. Sequencing of the ends of these fragments showed part of the gene coding for the C-terminal sequence of E1o, an intracistronic region and part of the gene coding for the N-terminal sequence of E2o. The 4.2-kb *Stu*I fragment was then digested with *Pst*I and *Dra*I yielding a fragment containing the complete E2o gene

```
<---------   Structural gene for 2-oxoglutarate dehydrogenase
 F  L  Q  Y  A  G  R  E  A  S  A  A  P  A  C  G  Y  A  S  M
GTTCCTCCAGTATGCTGGTCGCGAGGCGTCGGCCGCTCCGGCTTGCGGTTACGCGTCGAT
      2830      2840      2850      2860      2870      2880

 H  A  E  Q  Q  E  K  L  L  Q  D  A  F  T  V  *
GCATGCCGAGCAGCAGGAAAAACTGCTGCAGGACGCTTTCACTGTTTAACGCCTTCTCGC
      2890      2900      2910      2920      2930      2940

                 rbs                  A  I  D  I  K  A  P  T  F
AGAAGCAACGAATCAATAGGAACTCTAGATAATGGCTATCGATATCAAAGCCCCGACTTT
      2950      2960      2970      2980      2990      3000

 P  E  S  I  A  D  G  T  V  A  T  W  H  K  K  P  G  E  P  V
CCCGGAATCGATCGCGGACGGCACCGTCGCCACCTGGCACAAAAAGCCCGGCGAACCGGT
      3010      3020      3030      3040      3050      3060
                                          ●
 K  R  D  E  L  I  V  D  I  E  T  D  K  V  V  M  E  V  L  A
CAAGCGTGACGAGCTGATCGTCGACATCGAAACCGACAAGGTCGTGATGGAAGTGCTTGC
      3070      3080      3090      3100      3110      3120

 E  A  D  G  V  I  A  E  I  V  K  N  E  G  D  T  V  L  S  G
TGAAGCGGACGGCGTGATCGCCGAGATCGTCAAGAACGAAGGCGACACCGTGCTCAGCGG
      3130      3140      3150      3160      3170      3180

 E  L  L  G  K  L  T  E  G  G  A  A  T  A  A  P  A  A  A  P
CGAACTGCTCGGCAAGCTGACCGAAGGTGGTGCGGCCACCGCCGCTCCGGCTGCTGCCCC
      3190      3200      3210      3220      3230      3240

 A  P  A  A  A  A  P  A  A  A  E  A  P  I  L  S  P  A  A  R
GGCTCCGGCTGCCGCCGCTCCGGCCGCTGCCGAAGCGCCGATCCTGTCCCCGGCCGCCCG
      3250      3260      3270      3280      3290      3300

 K  I  A  E  E  N  A  I  A  A  D  S  I  T  G  T  G  K  G  G
CAAGATCGCCGAGGAAAACGCCATCGCTGCCGACAGCATCACCGGTACCGGCAAGGGCGG
      3310      3320      3330      3340      3350      3360

 R  V  T  K  E  D  A  V  A  A  A  E  A  K  K  S  A  P  A  G
TCGCGTGACCAAGGAAGACGCGGTCGCCGCCGCGGAAGCCAAGAAGTCCGCTCCTGCCGG
      3370      3380      3390      3400      3410      3420

 Q  P  A  P  A  A  T  A  A  P  L  F  A  A  G  D  R  V  E  K
CCAGCCGGCGCCCGCCGCCACCGCCGCTCCGCTGTTCGCGGCCGGCGATCGCGTCGAGAA
      3430      3440      3450      3460      3470      3480

 R  V  P  M  T  R  L  R  A  K  V  A  E  R  L  V  E  A  Q  S
GCGCGTGCCGATGACCCGCCTGCGCGCCAAGGTCGCCGAGCGCCTGGTCGAGGCCCAGTC
      3490      3500      3510      3520      3530      3540

 S  M  A  M  L  T  T  F  N  E  V  N  M  K  P  V  M  E  L  R
CTCCATGGCCATGCTGACGACCTTCAACGAAGTCAACATGAAGCCGGTCATGGAGCTGCG
      3550      3560      3570      3580      3590      3600

 A  K  Y  K  D  L  F  E  K  T  H  N  G  V  R  L  G  F  M  S
CGCCAAGTACAAGGATCTGTTCGAGAAGACCCACAACGGCGTGCGCCTGGGCTTCATGTC
      3610      3620      3630      3640      3650      3660

 F  F  V  K  A  A  V  E  A  L  K  R  Q  P  G  V  N  A  S  I
CTTCTTCGTCAAGGCCGCTGTCGAAGCCCTGAAGCGCCAGCCGGGCGTCAACGCCTCCAT
      3670      3680      3690      3700      3710      3720

 D  G  N  D  I  V  Y  H  G  Y  Q  D  I  G  V  A  V  S  S  D
CGACGGTAACGACATCGTCTACCACGGCTACCAGGACATCGGCGTGGCCGTTTCCAGCGA
      3730      3740      3750      3760      3770      3780

 R  G  L  V  V  P  V  L  R  N  A  E  F  M  S  L  A  E  I  E
TCGTGGTCTGGTGGTGCCGGTGCTGCGCAACGCCGAGTTCATGAGCCTGGCCGAGATCGA
      3790      3800      3810      3820      3830      3840

 G  G  I  N  E  F  G  K  K  A  K  A  G  K  L  T  I  E  E  M
AGGTGGCATCAACGAGTTCGGCAAGAAGGCCAAGGCCGGCAAGCTGACCATCGAGGAAAT
      3850      3860      3870      3880      3890      3900

 T  G  G  T  F  T  I  S  N  G  G  V  F  G  S  L  L  S  T  P
GACCGGCGGCACCTTCACCATCTCCAACGGTGGCGTGTTCGGTTCCCTGCTCTCGACTCC
      3910      3920      3930      3940      3950      3960

 I  V  N  P  P  Q  T  A  I  L  G  M  H  K  I  Q  E  R  P  M
GATCGTCAACCCGCCGCAGACCGCCATCCTCGGCATGCACAAGATCCAGGAGCGTCCGAT
      3970      3980      3990      4000      4010      4020

 A  V  N  G  Q  V  V  I  L  P  M  M  Y  L  A  L  S  Y  D  H
GGCCGTCAATGGTCAGGTCGTCATCCTGCCGATGATGTATCTTGCCCTGTCCTACGATCA
      4030      4040      4050      4060      4070      4080

 R  L  I  D  G  K  E  A  V  T  F  L  V  T  M  K  D  L  L  E
CCGTTTGATCGATGGCAAGGAGGCCGTTACCTTCCTGGTGACCATGAAGGATCTCCTGGA
      4090      4100      4110      4120      4130      4140

 D  P  A  R  L  L  L  D  V  *
AGACCCGGCTCGCCTGCTGCTGGACGTCTGATCTGAGATCGCAAGCTGCCCGCACAGGCT
      4150      4160      4170      4180      4190      4200


TTTCCAGGAAAGCCCGTGCCGGGCGCTTTACGCCGTTTTAAAGAAGAGGATTCGCTATGA
      4210      4220      4230      4240      4250      4260

 Structural gene for dihydrolipoamide dehydrogenase   ------>
S  Q  K  F  D  V  I  V  I  G  A  G  P  G  G  Y  V  A  A  I
GCCAGAAGTTCGACGTGATTGTGATTGGTGCAGGCCCCGGCGGCTACGTGGCTGCCATCA
      4270      4280      4290      4300      4310      4320
```

Fig. 2. *Nucleotide sequence of the* A. vinelandii *succinyltransferase gene and primary structure of its product.* The nucleotide sequence of the non-coding strand containing the gene encoding E2o is shown in the 5′-3′ direction. The amino acid sequence of 398 residues comprising succinyltransferase is shown directly above the sequence, as are the C-terminus of 2-oxoglutarate dehydrogenase and the N-terminus of lipoamide dehydrogenase. The ribosome binding site (rbs) is upperlined and relevant stop codons are indicated by asterisks. The lysine residue believed to be lipoylated is denoted by a filled circle (●)

Table 1. *Codon usage in the succinyltransferase gene from* A. vinelandii The AUG initiation codon is not included

F UUU	0	S UCU	0	Y UAU	1	C UGU	0
F UUC	12	S UCC	10	Y UAC	4	C UGC	0
L UUA	0	S UCA	0	* UAA	0	* UGA	1
L UUG	1	S UCG	2	* UAG	0	W UGG	1
L CUU	2	P CCU	1	H CAU	0	R CGU	4
L CUC	5	P CCC	2	H CAC	5	R CGC	12
L CUA	0	P CCA	0	Q CAA	0	R CGA	0
L CUG	24	P CCG	21	Q CAG	7	R CGG	0
I AUU	0	T ACU	2	N AAU	1	S AGU	0
I AUC	23	T ACC	20	N AAC	11	S AGC	4
I AUA	0	T ACA	0	K AAA	2	R AGA	0
M AUG	14	T ACG	1	K AAG	25	R AGG	0
V GUU	2	A GCU	14	D GAU	7	G GGU	10
V GUC	20	A GCC	40	D GAC	13	G GGC	23
V GUA	0	A GCA	0	E GAA	17	G GGA	0
V GUG	13	A GCG	8	E GAG	14	G GGG	0

(Fig. 1, fragment H) which was ligated into pUC9 (pAE2). The *Pst*I – *Kpn*I fragment indicated (Fig. 1, fragment G) which overlaps with both the original 3.2-kb fragment and the *Stu*I – *Eco*RV fragment, was sequenced completely from both ends, thus completing the sequence of the E2o gene.

Nucleotide sequence analysis

The DNA sequence of the E2o gene and the derived sequence are presented in Fig. 2. Complete overlap was obtained and 83% of the sequence was derived from both strands (residues 390 – 435, 535 – 624 and 1014 – 1094 were derived from one strand of DNA with a minimum of two completely overlapping fragments). If it is assumed that the initiating formylmethionine is removed post-translationally, alanine will become the first residue. This agrees with the N-terminal sequence determined with automated Edman degradation (10 amino acids: AIDIKAPTFP) of the protein isolated from *A. vinelandii*. The enzyme molecule contains 398 amino acid residues with a total mass of 41 872 Da. The chain mass as determined with SDS-gel electrophoresis (45.2 kDa) is over-estimated by 8%, probably due to the poor SDS binding of the Ala-Pro sequence in the lipoyl domain, similar to that found for acetyltransferases [6, 7].

Regulating elements of the nucleotide sequence and codon usage

An intercistronic region of 42 bp is found between the structural genes of 2-oxoglutarate dehydrogenase and succinyltransferase. No *E. coli*-type promoter sequence [32] is found in this intercistronic region, making it probable that transcription is initiated from a promoter upstream of the E1o gene, as found for the *E. coli* genes. A ribosome binding site is located 9 – 15 bp upstream from the AUG start codon.

Table 2. *Purification of succinyltransferase of* A. vinelandii *from* E. coli *TG2(pAE2)*

Step	Volume	Protein	Specific activity	Total activity	Yield
	ml	mg	U/mg	U	%
Crude extract	335	8442	0.064	539	100
Protamine sulfate	345	3933	0.132	520	96
DEAE-Sephacel	325	426	0.941	402	75
Sephacryl S400 HR	91	76	2.84	217	40

No large inverted repeats are found downstream from the stop codon. This could indicate that the mRNA transcript contains all three genes required for synthesis of the 2-oxoglutarate dehydrogenase complex. The E3 gene contains its own promoter [22] and therefore it is likely that excess E3 can be synthesized for assembly of the pyruvate dehydrogenase complex. A similar situation is found in *Pseudomonas fluorescens* [33]. In contrast, the E3 gene from *E. coli*, being organized as part of the pyruvate dehydrogenase gene cluster, is used for assembly of the 2-oxoglutarate dehydrogenase complex [21]. The reason for this difference in organization is not known. It could be argued that in *E. coli* carbohydrate metabolism, and thus the pyruvate dehydrogenase complex, plays a central role, while the glutamate N-cycle, and thus the 2-oxoglutarate dehydrogenase complex, is more fundamental to soil organisms such as *A. vinelandii* and *P. fluorescens*.

The codon usage (Table 1) is nonrandom and very similar to that of other genes of *A. vinelandii* sequenced [7, 22, 23, 34]. The overall C + G content is 64.7%.

Expression

E. coli TG2 (pAE2) produces high quantities of succinyltransferase under the influence of the vector-encoded *lacZ* promoter. Table 2 shows the result of a purification from a 20-l culture. The final product was pure as judged from SDS gels stained with Coomassie brilliant blue R. The isolated enzyme will be used for crystallization and NMR studies.

Sequence comparison

The sequences of the transsuccinylases from *A. vinelandii* and *E. coli* are compared in Fig. 3. The similarity is high and 63% identity exists between both sequences. The domain structure is identical for both proteins. It is clear that both enzymes contain a single N-terminal lipoyl domain. This is followed by a region thought to be involved in E3 binding, which is not only very similar to *E. coli* transsuccinylase but also to other regions found in the transacetylases from both organisms. The C-terminal part contains the transsuccinylase active site and the information for the quaternary structure; it shows a high degree of similarity to the *E. coli* enzyme. Based on the similarity, Lys-42 is assumed to contain the covalently bound lipoyl group. The major difference between the sequences is observed between the lipoyl domain and the lipoamide-dehydrogenase-binding domain. In the *A. vinelandii* sequence a large Ala-Pro-rich sequence (31 amino acids) is found, which is similar to that found at the end of all three lipoyl domains in the *A. vinelandii* and *E. coli* transacetylase components. In the transsuccinylase from *E. coli* and also in the transacetylase from *B. stearothermophilus* this region contains many charged residues, although the latter example

```
 1          11         21         31         41         51
 AIDIKAPTFP ESIADGTVAT WHKKPGEPVK RDELIVDIET DKVVMEVLAE ADGVIAEIVK
   **  *  * ** ** **** ******  *  ***  * *** **** ** *  ***
SSVDILVPDLP ESVADATVAT WHKKPGDAVV RDEVLVEIET DKVVLEVPAS ADGILDAVLE
1           12         22         32         42         52
61          71         81         91        100         104
 NEGDTVLSGE LLGKLTEGGA ATAAPAAAPA PAAAAPAA-A EAPI------ -LSPAARKIA
  ** ** *    ** * **   *     *     *  **     *           **** *
 DEGTTVTSRQ ILGRLREGNS AGKETSAKSE EKASTPAQRQ QASLEEQNND ALSPAIRRLL
62          72         82         92        102        112
113        123        133        140        150        160
 EENAIAADSI TGTGKGGRVT KEDA---VAA AEAKKSAPAG QPAPAATAAP LFAAGDRVEK
  *    *  *  *** *** *  **     *  * ** ****    * *  * *   **  * **
 AEHNLDASAI KGTGVGGRLT REDVEKHLAK APAKESAPA- --AAAPAAQP ALAA--RSEK
122        132        142        152         161       169
170        180        190        200        210        220
 RVPMTRLRAK VAERLVEAQS SMAMLTTFNE VNMKPVMELR AKYKDLFEKT HNGVRLGFMS
 ********   ***** **   ● ******** ***** * **   *   ***  * * ******
 RVPMTRLRKR VAERLLEAKN STAMLTTFNE VNMKPIMDLR KQYGEAFEKR H-GIRLGFMS
177        187        197        207        217        227
230        240        250        260        270        280
 FFVKAAVEAL KRQPGVNASI DGNDIVYHGY QDIGVAVSSD RGLVVPVLRN AEFMSLAEIE
 * *** **** ** * ***** ** * *** *  *   ***   **** ****        * **
 FYVKAVVEAL KRYPEVNASI DGDDVVYHNY FDVSMAVSTP RGLVTPVLRD VDTLGMADIE
236        246        256        266        276        286
290        300        310        320        330        340
 GGINEFGKKA KAGKLTIEEM TGGTFTISNG GVFGSLLSTP IVNPPQTAIL GMHKIQERPM
   * *   *    **** *   *** *** ** ****** *** * **** *** *** *  ***
 KKIKELAVKG RDGKLTVEDL TGGNFTITNG GVFGSLMSTP IINPPQSAIL GMHAIKDRPM
296        306        316        326        336        346
350        360        370        380        390
 AVNGQVVILP MMYLALSYDH RLIDGKEAVT FLVTMKDLLE DPARLLLDV*
 ****** *** ********** ***** * *  **** * *** ** ******
 AVNGQVEILP MMYLALSYDH RLIDGRESVG FLVTIKELLE DPTRLLLDV*
356        366        376        386        396
```

Fig. 3. *Comparison of the amino acid sequences of succinyltransferase from* A. vinelandii *and* E. coli *[8]*. Residues are given in single-letter amino acid code. Identical residues are indicated by asterisks. Upper sequence, *A. vinelandii*; bottom sequence, *E. coli*

also contains a segment of Ala-Pro-rich sequence. Proton NMR studies [35–39] have clearly shown the internal mobility of the residues in the Ala-Pro-rich regions. These regions are assumed therefore to give flexibility to the lipoyl domain in addition to the movement of the lipoyl group in order to visit the respective active sites. It has not been shown thus far that the charged regions confer a similar mobility to the lipoyl domain. A direct proof of this mobility is given by fluorescence anisotropy measurements of the acetyltransferase from *A. vinelandii* [40]. However, individual mobility of the three lipoyl domains could not be demonstrated but a movement of the three domains as a whole was found. Furthermore high mobility of the individual lipoyl domains (mass about 8 kDa) would result in sharp resonances in the proton NMR spectrum of all the constituting amino acids and not just those of the Ala-Pro-rich regions. With the high amounts of protein now available, this aspect can be studied in more detail.

We wish to thank Dr R. Amons for the determination of the N-terminal sequence with the gas-phase sequenator. This investigation was supported by the Netherlands Foundation for Chemical Research (SON), with the financial aid from the Netherlands Organization for Scientific Research (NWO).

REFERENCES

1. Reed, L. J. & Cox, D. J. (1966) *Annu. Rev. Biochem. 35*, 57–84.
2. Danson, M. J. & Perham, R. N. (1976) *Biochem. J. 159*, 677–682.
3. White, R. H., Bleile, D. M. & Reed, L. J. (1980) *Biochem. Biophys. Res. Commun. 94*, 78–84.
4. Packman, L. C., Borges, A. & Perham, R. N. (1988) *Biochem. J. 252*, 79–86.
5. Thekkumkara, T. J., Ho, L., Wexler, I. D., Pons, G., Liu, T. & Patel, M. S. (1988) *FEBS Lett. 240*, 45–48.
6. Stephens, P. E., Darlison, M. G., Lewis, H. M. & Guest, J. R. (1983) *Eur. J. Biochem. 133*, 481–489.
7. Hanemaaijer, R., Janssen, A., de Kok, A. & Veeger, C. (1988) *Eur. J. Biochem. 174*, 593–599.
8. Spencer, M. E., Darlison, M. G., Stephens, P. E., Duckenfield, I. K. & Guest, J. R. (1984) *Eur. J. Biochem. 141*, 361–374.
9. Guest, J. R., Lewis, H. M., Graham, L. D., Packman, L. C. & Perham, R. N. (1985) *J. Mol. Biol. 185*, 743–754.
10. Reed, L. J. & Oliver, R. M. (1968) *Brookhaven Symp. Biol. 21*, 397–411.
11. Hanemaaijer, R., Westphal, A. H., van der Heiden, T., de Kok, A. & Veeger, C. (1989) *Eur. J. Biochem. 179*, 287–292.
12. Reed, L. J., Damuni, Z. & Merryfield, M. L. (1985) *Curr. Top. Cell. Regul. 27*, 41–49.
13. Pettit, F. H., Hamilton, L., Munk, P., Namihira, G., Eley, M. H., Willms, C. R. & Reed, L. J. (1973) *J. Biol. Chem. 248*, 5282–5290.
14. Koike, M., Hamada, M., Koike, K., Hiraoka, T. & Nakaula, Y. (1976) in *Thiamin* (Gubler, C. J., Fujiwara, M. & Dreyfus, P. M., eds) pp. 5–18, John Wiley, New York.
15. Bleile, D. M., Munk, P., Oliver, R. M. & Reed, L. J. (1979) *Proc. Natl Acad. Sci. USA 76*, 4385.
16. Hanemaaijer, R., de Kok, A., Jollès, J. & Veeger, C. (1987) *Eur. J. Biochem. 169*, 245–252.
17. Packman, L. C. & Perham, R. N. (1986) *FEBS Lett. 206*, 193–198.

17a. Perham, R. N. & Wilkie, A. O. M. (1980) *Biochem. Int. 1*, 470–478.

18. Kresze, G. B., Ronft, H. & Dietl, B. (1980) *Eur. J. Biochem. 105*, 371–379.
19. Hummel, B. K., Litwer, S., Bradford, A. P., Aitken, A., Danner, D. J. & Yeaman, S. J. (1988) *J. Biol. Chem. 263*, 6165–6168.
20. Lau, S. K., Griffin, T. A., Hu, C. W. C. & Chuang, D. T. (1988) *Biochemistry 27*, 1972–1981.
21. Spencer, M. E. & Guest, J. R. (1988) *Mol. Gen. Genet. 200*, 145–154.
22. Westphal, A. H. & de Kok, A. (1988) *Eur. J. Biochem. 172*, 299–305.
23. Schulze, E., Westphal, A. H., Hanemaaijer, R. & de Kok, A. (1989) *Eur. J. Biochem. 187*, 229–234.
24. Simon, E. J. & Shemin, D. (1953) *J. Am. Chem. Soc. 75*, 2520.
25. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) *Molecular cloning: a laboratory manual*, Cold Spring Harbour Laboratory, Cold Spring Harbour, NY.
26. Southern, E. (1975) *J. Mol. Biol. 98*, 503–517.
27. Sanger, F., Nicklen, S. & Coulsen, A. R. (1977) *Proc. Natl Acad. Sci. USA 74*, 5363–5367.
28. Bankier, A. T. & Barrell, B. G. (1983) in *Techniques in nucleic acid biochemistry* (Flavell, R. A., ed.) pp. 1–34, Elsevier Scientific Publishers, Amsterdam.
29. Staden, R. (1982) *Nucleic Acids Res. 10*, 4731–4751.
30. Staden, R. (1984) *Nucleic Acids Res. 12*, 499–503.
31. Miller, J. H. (1977) *Experiments in molecular genetics*, 3rd edn, Cold Spring Harbour Laboratory, Cold Spring Harbour, NY.
32. Harley, C. B. & Reynolds, R. P. (1987) *Nucleic Acids Res. 15*, 2343–2361.
33. Benen, J. A. E., van Berkel, W. J. H., van Dongen, W. M. A. M., Müller, F. & de Kok, A. (1989) *J. Gen. Microbiol. 135*, 1787–1797.
34. Brigle, K. E., Newton, W. E. & Dean, D. R. (1985) *Gene 37*, 37–44.
35. Hanemaaijer, R., Vervoort, J., Westphal, A. H., de Kok, A. & Veeger, C. (1988) *FEBS Lett. 240*, 205–210.
36. Perham, R. N., Duckworth, H. W. & Roberts, G. C. K. (1981) *Nature 292*, 474–477.
37. Roberts, G. C. K., Duckworth, H. W., Packman, L. C. & Perham, R. N. (1983) *Ciba Found. Symp. 93*, 47–62.
38. Duckworth, H. W., Jaenicke, R., Perham, R. N., Wilkie, A. O. M., Finch, J. T. & Roberts, G. C. K. (1982) *Eur. J. Biochem. 124*, 47–62.
39. Packman, L. C., Perham, R. N. & Roberts, G. C. K. (1984) *Biochem. J. 217*, 219–227.
40. Hanemaaijer, R., Masurel, R., Visser, A. J. W. G., de Kok, A. & Veeger, C. (1988) *FEBS Lett. 238*, 285–288.