Isono, K. (1974), *Mol. Gen. Genet. 133,* 77-86.
Kaltschmidt, E., and Wittmann, H. G. (1970), *Proc. Natl. Acad. Sci. U.S.A. 67,* 1276-1282.
Moore, P. B. (1975), *J. Mol. Biol. 98,* 439-444.
Nagano, K. (1977), *J. Mol. Biol. 109,* 251-274.
Robson, B., and Suzuki, E. (1976), *J. Mol. Biol. 107,* 327-356.
Roseau, G., and Pantel, P. (1969), *J. Chromatogr. 44,* 392-395.
Sanger, F., and Thompson, E. O. P. (1953), *Biochem. J. 53,* 366-374.
Schiltz, E., Schnackerz, K. D., and Gracy, W. R. (1977), *Anal. Biochem. 79,* 33-42.
Schwabe, E. (1972), *Hoppe-Seyler's Z. Physiol. Chem. 353,* 1899-1906.
Smithies, O., Gibson, D. M., Fanning, E. M., Goodfliesh, R. M., Gillman, J. G., and Ballantyne, G. L. (1971), *Biochemistry 10,* 4912-4921.
Tschesche, H., and Kupfer, S. (1972), *Eur. J. Biochem. 26,* 33-36.
Vandekerckhove, J., and Van Montagu, M. (1974), *Eur. J. Biochem. 44,* 279-288.
Weber, H. J. (1972), *Mol. Gen. Genet. 119,* 233-248.
Weiner, A. M., Platt, T., and Weber, K. (1972), *J. Biol. Chem. 247,* 3242-3251.
Wittmann-Liebold, B., and Dzionara, M. (1976), *FEBS Lett. 65,* 281-283.
Yaguchi, M., Wittmann, H. G., Cabézon, T., DeWilde, M., Villarroel, R., Herzog, A., and Bollen, A. (1975), *Mol. Gen. Genet. 142,* 35-43.

Primary Structure of Protein L19 from the Large Subunit of *Escherichia coli* Ribosomes†

Jürgen Brosius*,‡ and Ursula Arfsten

ABSTRACT: Protein L19, a component of the *Escherichia coli* 50S ribosomal subunit implicated in 30S-50S subunit interaction, was sequenced by the dansyl-Edman method. L19 consists of a single polypeptide chain of 114 amino acids giving a calculated molecular weight of 13 002. Peptides obtained from various enzymatic cleavages were isolated on thin-layer peptide maps or by gel filtration. Automated Edman degradation using a liquid phase sequenator was carried out on the whole protein as well as on a large 58-residue fragment arising from digestion with *Staphylococcus aureus* protease. Every position in protein L19 was confirmed at least twice. Results of secondary structure estimation and homologies with other *E. coli* ribosomal protein sequences are presented.

Protein L19 is a component of the structure of the 50S ribosomal subunit, and has been implicated in the 30S-50S subunit association (Morrison et al., 1973; Tischendorf et al., 1974). It is accessible to antibody in the free 50S subunit (Morrison et al., 1977) but not in the undissociated 70S particle. Thus, it is one of the 50S proteins that is shielded by the 30S subunit against antibodies (Zeichhardt, 1976). Variable results have been reported for the binding of protein L19 to ribosomal RNA (Garrett et al., 1974). When chloramphenicol bound to 50S subunits is irradiated with ultraviolet (UV) light, the antibiotic is found to become covalently attached to protein L19 (Sonenberg et al., 1976). Protein L19 has been shown to have an antigenic determinant near that of L16 (Tischendorf et al., 1975), a protein which has also been implicated in the binding of chloramphenicol (Nierhaus and Nierhaus, 1973; Pongs et al., 1973). These findings suggest that L19 is located at the 30S-50S interface, and that it may play a role in the structure and function of the ribosomal A site. Knowledge of its primary structure may thus lead to further insights into the mechanism of protein synthesis at the molecular level.

Materials and Methods

If not stated otherwise chemicals were analytical grade purchased from Merck.

Isolation of Protein L19. Protein L19 was isolated according to Hindennach et al. (1971) and provided by Dr. H. G. Wittmann. Its identity and purity were checked by two-dimensional polyacrylamide gel electrophoresis (Kaltschmidt and Wittmann, 1970).

Amino Acid Analysis. Unoxidized and oxidized protein (10 μg each) were subjected to amino acid analysis as described in the preceding paper (Brosius, 1978).

Tryptophan-containing peptides were located by spraying tryptic peptide maps with 1% dimethylbenzaldehyde (Riedel de Haen, Hanover) in 2 N HCl (Spies and Chambers, 1949) or by fluorescence under long-wavelength UV light, as described in the preceding paper (Brosius, 1978). Furthermore, Trp can be detected as its Pth[1] derivative by silica gel thin-layer chromatography (Brosius, 1978).

The N-terminal amino acid was determined by dansylation of 2 μg of the intact protein.

Enzymatic Cleavage. Methods used for enzymatic cleavage of protein L19 are summarized in Table I. Blocking of the lysine residues with citraconic acid anhydride prior to tryptic

† From the Max-Planck-Institut für Molekulare Genetik, Abt. Wittmann, Berlin-Dahlem, West Germany. *Received September 8, 1977.*

‡ Present address: Thimann Laboratories, University of California, Santa Cruz, Calif. 95064.

[1] The same abbreviations and symbols are used as in the preceding paper (Brosius, 1978). In addition, the following were used: Tos-LysCH_2Cl, *N*-α-*p*-tosyl-L-lysylchloromethane; BT, peptides resulting from digestion with trypsin after blocking the lysine residues with citraconic acid anhydride; CH, peptides resulting from digestion with Tos-LysCH_2Cl-chymotrypsin; CH′, peptides resulting from digestion with α-chymotrypsin; SP, peptides resulting from digestion with *Staphylococcus aureus* protease at pH 4.0; SP′, peptides resulting from digestion with *Staphylococcus aureus* protease at pH 7.8; —, solid phase Edman degradation.

0006-2960/78/0417-0508$01.00/0

TABLE I: Enzymatic Digestion of Protein L19.

Expt	Enzyme and Source	Amt of protein (mg)	Ratio enzyme: substrate (w/w)	Buffer used	Time and temp (°C)
1	Trypsin, Tos-PheCH_2Cl treated[a]	0.60	1:50	1 mL of buffer I	4 h, 37
2	*Staph. aureus* protease[b]	3.10	1:40	1.5 mL of 0.05 M ammonium acetate (pH 4.0)	16 h, 37
3	*Staph. aureus* protease[b]	0.07	1:40	0.3 mL of 0.05 M potassium phosphate (pH 7.8)	18 h, 37
4	Chymotrypsin, Tos-LysCH_2Cl treated[a]	0.65	1:60	0.8 mL of buffer I	15 min, 37
5	α-Chymotrypsin[a]	0.13	1:60	0.4 mL of buffer I	15 min, 37
6	Thermolysin[c]	0.45	1:50	0.5 mL of buffer I	90 min, 37
7	Trypsin, Tos-PheCH_2Cl treated[a] after blocking of Lys residues with citraconic acid anhydride	1.35	1:50	3.5 mL of 2 ‰ ammonium bicarbonate	1 h, 37
8	Pepsin[c]	0.80	1:50	1.0 mL of 0.05 N HCl	1 h, 37
9	Carboxypeptidase A[d]	0.02	1:20	0.05 mL of buffer I	30 min, 37
10	Trypsin, Tos-PheCH_2Cl treated[a]	Peptide SP5 0.13 7.17, total	1:40	0.20 mL of buffer I	3 h, 37

[a] Merck, Darmstadt, West Germany. [b] Miles, Slough, U. K. [c] Serva, Heidelberg, West Germany. [d] Böhringer, Mannheim, West Germany. [e] Buffer I: 0.2 M *N*-methylmorpholine acetate (pH 8.0).

cleavage was carried out with 1.35 mg of protein according to Dixon and Perham (1968). After tryptic digestion of the modified protein the lysine residues were unblocked by lowering the pH to 3.3 with 50% acetic acid followed by incubation at room temperature overnight. The solution was lyophilized and peptides separated as described below.

Isolation of Peptides. Most of the peptides were isolated and visualized as described in the preceding paper (Brosius, 1978). One peptide map was usually carried out on 70 μg of protein, digested as summarized in Table I. In addition to ninhydrin and fluorescamine detection one tryptic fingerprint was sprayed with 0.02% phenanthrenequinone (Fluka) in ethanol and then with 10% NaOH–ethanol (1:1) according to Yamada and Itano (1966). Arginine-containing peptides were then located under long-wavelength UV light. Some peptide maps were visualized according to Reindel and Hoppe (1954) with the following modification: the thin-layer plate was first exposed to an atmosphere of chlorine generated from $KMnO_4$ and 10% HCl. After aeration for 1–2 h the plate was sprayed with the following mixture: *o*-toluidine, 80 mg; KI, 0.5 mg; acetic acid, 1.5 mL; and water to 250 mL. Peptides and amino acids generated brownish spots.

In order to separate large peptides from smaller ones, aliquots of the peptides obtained after limited tryptic digestion or cleavage with *S. aureus* protease were separated on a Sephadex G50 (superfine) column (135 × 1 cm). Samples dissolved in 400 μL of 20% acetic acid were applied and eluted with 10% acetic acid. The separation was monitored at 280 nm with a Spectrochrome M (Gilson) and fractions of 0.5 mL were collected.

Sequence Determination. Almost all peptides were sequenced by the dansyl-Edman method (Gray and Hartley, 1963) with minor modifications, as described in detail in the previous paper (Brosius, 1978).

Peptide BT5, together with peptides BT2 and BT8, was isolated from Sephadex G50 from the mixture of blocked tryptic peptides and 30 nmol was sequenced in a solid-phase sequenator (Laursen, 1971; Laursen et al., 1972; Wittmann-Liebold and Lehmann, 1975). The peptides were attached to an aminopolystyrene resin as described previously (Schiltz and Reinbolt, 1975).

TABLE II: Amino Acid Composition of Protein L19.[a]

	Values derived from hydrolysis	Values derived from sequence
Cys		
Asn		4
Asp	7.40	3
Thr	4.04	4
Ser	7.18	8
Gln		6
Glu	15.19	9
Pro	3.33	3
Gly	8.89	8
Ala	7.52	7
Val	14.12	15
Met	1.11	1
Ile	6.55	7
Leu	6.28	6
Tyr	1.95	2
Phe	3.61	4
His	1.98	2
Lys	11.18	11
Arg	13.07	13
Trp	+	1

[a] Amino acid composition as determined after hydrolysis at 110 °C for 20 h (averaged and uncorrected values from four different runs). Trp was determined on peptide maps after spraying with Ehrlich's reagent.

Intact protein (100 nmol) and 100 nmol of the large fragment SP5 were twice subjected to Edman degradation (Edman and Begg, 1967) in an improved Beckman liquid-phase sequenator (Wittmann-Liebold, 1973). Experimental details of the program and procedure are given elsewhere (Wittmann-Liebold et al., 1976; Stadler and Wittmann-Liebold, 1976). Quadrol buffer (0.5 M) was used and two couplings per cycle were performed (Hitz et al., 1977).

Results

Amino Acid Composition. The amino acid composition of protein L19 is presented in Table II. After oxidation, no cysteic acid was detected. One blue fluorescent spot was detected on

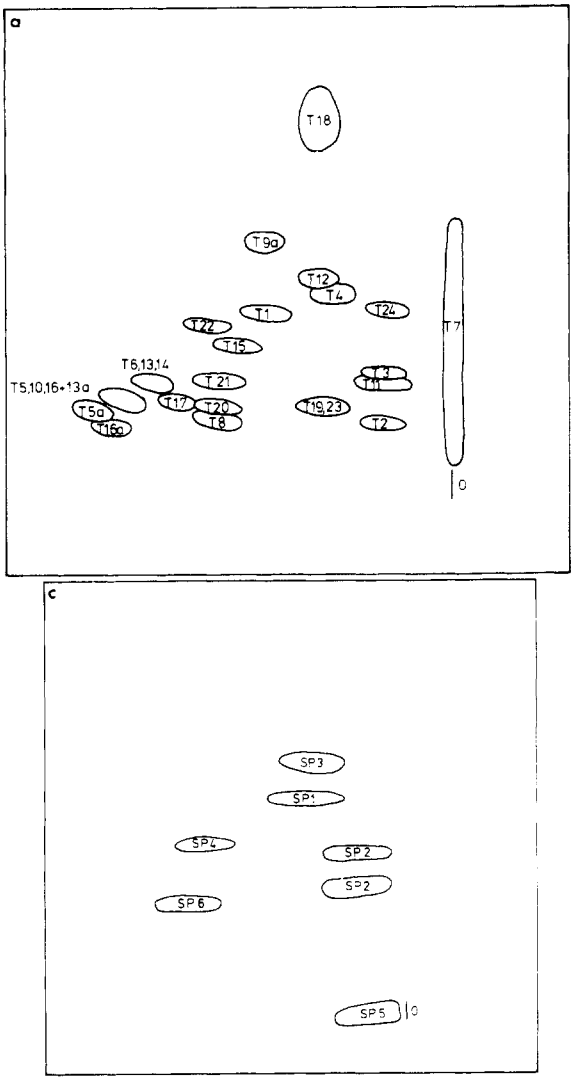


FIGURE 1: Thin-layer peptide maps from enzymatic digests of protein L19; O = origin. In all peptide maps shown the direction of electrophoresis is horizontal (cathode left and anode right) and the direction of chromatography is vertical (bottom to top): (a) tryptic digest; (c) *Staphylococcus aureus* protease digest. Figures 1b, 1d, 1e, and 1f can be found in the supplementary material.

each peptide map, revealing the presence of tryptophan in the respective peptides. The presence of Trp was confirmed by a positive color reaction with Ehrlich's reagent.

Enzymatic Digestion of the Protein; Isolation and Sequence of the Peptides. Peptides are numbered according to their sequence in the protein. The same symbols and abbreviations are used as in the preceding paper (Brosius, 1978).

Tryptic Peptides. All tryptic peptides were isolated from peptide maps (Figure 1a). Their amino acid compositions are summarized in Table IIIa (see Supplementary Material Available paragraph at end of paper). Peptides were not always well separated from each other (e.g., T3/T11 or T4/T12). In this case only the nonoverlapping regions of the spots were eluted from the thin-layer plates. In order to provide a sufficient amount, the material from several peptide maps was utilized.

Ser-Asn-Ile-Ile-Lys
T1

Tryptic peptide T1 was sequenced by the dansyl-Edman procedure. In addition to Dns-Ile, a spot was seen migrating in solvent A somewhat slower than Dns-Phe and in solvent B with Dns-Phe. This spot migrates in solvent C with Dns-Ile and presents the Dns derivative of the dipeptide Ile-Ile. Some peptide bonds, especially between hydrophobic residues, are not cleaved readily under the conditions employed. Asparagine was identified as its Pth derivative on silica gel thin-layer plates. Lys was identified as α-Dns-ϵ-Ptc-Lys on the polyamide thin-layer plate. Positions 3–5 were confirmed by hydrolysis of the ATZ derivatives and identification of the resulting amino acids on the amino acid analyzer. T1 is the only tryptic peptide with N-terminal Ser. Thus, T1 must be the N-terminal tryptic peptide as the N-terminal amino acid of the intact protein was determined to be serine.

Gln-Leu-Glu-Gln-Glu-Gln-Met-Lys
T2

Only one of the two possible spots which arose due to partial oxidation of methionine was isolated and sequenced from the fingerprint. The Gln and Glu residues were distinguished by conversion of the respective ATZ derivatives into Pth-amino acids.

Gln-Asp-Val-Pro-Ser-Phe-Arg-Pro-Gly-Asp-
Thr-Val-Glu-Val-Lys
T3

Peptide T3 was completely sequenced by the dansyl-Edman method. The Pth-aa of position 13 was not identified as the yield was too low. From other peptides, it was determined to be glutamic acid.

Val-Trp-Val-Val-Glu-Gly-Ser-Lys
T4

Peptide T4 was detected as a blue fluorescent spot under UV light at 366 nm, and gave a pink color with Ehrlich's reagent. In position 2 no Dns-aa was detected. The corresponding Pth derivative gave a bright yellow-green spot on the silica gel thin-layer plate after color reaction with ninhydrin, identical with that of a Pth-Trp standard. Position 5 was determined to be glutamic acid. In position 3 Dns-Val as well as Dns-Val-Val, which migrates in solvents B and C slower than Dns-Val, were present.

Lys
T5, T10, T16
Arg
T6, T13, T14
Lys-Arg
T5a
Arg-Arg
T13a

Peptide T13a and free lysine occur in the same spot on the peptide map (Figure 1a). This is in agreement with the finding that an authentic sample of the dipeptide Arg-Arg and the free amino acid lysine migrate identically in both the electrophoresis and chromatography systems. Lys and Arg were found to be the N-terminal amino acids in the spot containing peptide T13a and free lysine. After one degradation step only Dns-Arg was detected. The dipeptide T5a which was also sequenced migrates faster than T13a in the electrophoresis and chromatography systems and occurs as a discrete spot. T5a and T13a arise from incomplete digestion between two basic residues. In addition free Arg was released after tryptic digestion of L19, and was found in the spot (T6, 13, 14). A further degradation step did not reveal another amino acid.

Leu-Gln-Ala-Phe-Glu-Gly-Val-Val-Ile-Ala-Ile-Arg

T7

Peptide T7 could neither be detected with ninhydrin or fluorescamine spray nor with the Reindel reagent. Only spraying with phenanthrenequinone gave a clue to its position on the peptide map. T7 does not migrate during electrophoresis and smears in the chromatography system. Therefore, the peptide spot has an elongated shape and is distributed over a large area. This is one reason for the difficulty in detecting it. It was eluted in relatively low yield and the combined eluate from five peptide maps was used to sequence it. The Glu and Gln residues could not be distinguished, but were established independently from sequencing overlapping peptides (see below). Dns-Val and Dns-Val-Val were observed in position 7. In position 8 Dns-Val was detected and a second spot was seen in the Dns-Phe position after chromatography in solvent B. The second spot migrates to a position somewhat slower than Dns-Ile in solvent C and corresponds to Dns-Val-Ile. Position 9 revealed Dns-Ile as well as Dns-Ile-Ala which runs slower than Dns-Ala in the first dimension and with Dns-Ala in solvents B and C.

Asn-Arg

T8

The sequence was carried out by the dansyl-Edman procedure. Pth-Asn was identified by chromatography.

Gly-Leu-His-Ser-Ala-Phe-Thr-Val-Arg

T9

Gly-Leu-His-Ser-Ala-Phe

T9a

Thr-Val-Arg

T9b

Peptide T9 was usually obtained as two subfragments due to the chymotryptic activity in the trypsin used. The two halves T9a and T9b were sequenced by the dansyl-Edman method. Dns-His was identified in chromatography systems F and G.

Ile-Ser-Asn-Gly-Glu-Gly-Val-Glu-Arg

T11

This nonapeptide (T11) was also sequenced by the dansyl-Edman method. The amides were distinguished from the respective acids by conversion of the ATZ derivatives into Pth-aa and identification on silica gel thin-layer plates.

Val-Phe-Gln-Thr-His-Ser-Pro-Val-Val-Asp-
Ser-Ile-Ser-Val-Lys

T12

T12 was completely sequenced with the exception that no Pth-aa could be detected in position 10 due to its low yield. This position was determined to be Asp by sequencing thermolytic peptides TL16 and TL16a (see below). In the first position Dns-Val and an additional spot corresponding to Dns-Val-Phe migrating somewhat slower in solvent A and somewhat faster in solvents B and C than Dns-Phe were seen. Positions 1 and 2 were confirmed by hydrolysis of the respective ATZ derivatives.

Gly-Ala-Val-Arg

T15

This peptide (T15) gave a yellow color after spraying with ninhydrin, indicating that glycine is the N-terminal amino acid, which was confirmed by its sequence.

Lys-Ala-Lys

T16a

Ala-Lys

T17

T16a arises from incomplete tryptic digestion between peptides T16 and T17.

Leu-Tyr-Tyr-Leu-Arg

T18

Due to its hydrophobic character T18 migrates fast in the chromatography system during peptide mapping. It was sequenced by the dansyl-Edman method. In the first step Dns-Leu and a spot migrating in solvents B and C somewhat faster than bis-Dns-Tyr were detected. The latter spot likely corresponds to Dns-Leu-Tyr or a dansyl derivative of a larger hydrophobic peptide. All positions were confirmed by hydrolysis of the ATZ derivatives and their identification by amino acid analysis.

Glu-Arg

T19, T23

This dipeptide (T19, T23) occurs twice within the sequence of L19. The Glu residue was identified by its Pth derivative.

Thr-Gly-Lys

T20

Ala-Ala-Arg

T21

These two tripeptides (T20 and T21) were sequenced by the dansyl-Edman method.

Ile-Lys

T22

A spot migrating in solvent C somewhat faster than bis-Dns-Lys was detected in addition to Dns-Ile. This spot corresponds to Dns-Ile-(ε-Dns)-Lys. Similar findings were made with sequences like Leu-Lys, Val-Lys, or Met-Lys when lysine is in the second position of a peptide. If these amino acid combinations occur further back in a sequence of a peptide the situation is different since the ε-amino group of lysine has already reacted with phenyl isothiocyanate. In a peptide such as X-Val-Lys, for example, not only is Dns-Val seen in the second position but also a spot migrating in the first dimension faster than Dns-Phe and in solvents B and C like Dns-Phe. This spot corresponds to the Dns-Val-(ε-Ptc)-Lys compound.

Leu-Asn

T24

Asn was identified after conversion of the respective ATZ derivative into Pth-Asn. T24 must be the C-terminal tryptic peptide of L19 as it lacks Arg or Lys. Furthermore, carboxypeptidase A, when incubated with the intact protein, released the amino acids asparagine and leucine.

Peptides Obtained from Limited Tryptic Digestion. Most of the peptides originating from limited digestion with trypsin were isolated from peptide maps (Figure 1b; supplementary material) and sequenced by the dansyl-Edman method. Their amino acid compositions are given in Table IIIb (supplementary material). Two-thirds of the digest was passed over a Sephadex G-50 (superfine) column as described in Materials and Methods. Fractions 71–82 contained BT1; 97–105, BT6 and BT6a; 110–119, BT2, BT5, and BT8; 122–127, BT4,

BT10, and BT11; 129–140, BT3, BT7, BT7a, and BT9 (graph not shown).

BT1 (sequence shown in Figure 2) was isolated in pure form after gel filtration. This peptide was not sequenced as this part of the protein was determined by a liquid-phase sequenator run and confirmed several times by sequencing peptides from different digestions. BT1 comprises tryptic peptides T1–T6.

Leu-Gln-Ala-Phe-Glu-Gly-Val-Val-Ile-Ala-Ile-Arg

BT2

BT2 was isolated from an aliquot of fractions 110–119 by peptide mapping and was sequenced by the dansyl-Edman method.

Asn-Arg

BT3

BT3 is identical with tryptic peptide T8.

Gly-Leu-His-Ser-Ala-Phe-Thr-Val-Arg

BT4

BT4, which is identical with T9, remained intact since after tryptic digestion for only 1 h no detectable chymotryptic activity was observed.

Lys-Ile-Ser-Asn-Gly-Glu-Gly-Val-Glu-Arg

BT5

BT5 contains the tryptic peptide T11 and an N-terminal lysine residue. In order to confirm the identification of the two glutamic acid residues in positions 6 and 9, a mixture of peptides BT5, BT2, and BT8, isolated by gel filtration, was sequenced by the solid phase method. Even though the three peptides were sequenced in parallel it was shown unambiguously that positions 6 and 9 contain glutamic acid and not glutamine.

Lys-Ile-Ser-Asn-Gly-Glu-Gly-Val-Glu-Arg

BT5

Leu-Gln-Ala-Phe-Glu-Gly-Val-Val-Ile-Ala-Ile-Arg

BT2

Lys-Ala-Lys-Leu-Tyr-Tyr-Leu-Arg

BT8

Val-Phe-Gln-Thr-His-Ser-Pro-Val-Val-Asp-Ser-Ile-Ser-Val-Lys-Arg

BT6

BT6a, which was also isolated and sequenced, has an additional arginine at the C-terminal end due to incomplete tryptic digestion between two Arg residues.

Arg-Gly-Ala-Val-Arg

BT7

Gly-Ala-Val-Arg

BT7a

BT7a is identical with tryptic peptide T15. BT7 contains an additional N-terminal arginine.

Lys-Ala-Lys-Leu-Tyr-Tyr-Leu-Arg

T16a T18

BT8

BT8 contains tryptic peptides T16a and T18.

Glu-Arg

BT9

BT9 is identical with T19.

Thr-Gly-Lys-Ala-Ala-Arg

T20 T21

BT10

BT10 links tryptic peptide T20 to T21.

Ile-Lys-Glu-Arg

T22 T23

BT11

BT11 is the bridge between peptides T22 and T23.

Leu-Asn

BT12

Peptide BT12 is identical with T24, the C-terminal tryptic peptide, in agreement with the fact that it does not contain Arg.

Peptides Obtained from Cleavage with Staphylococcus aureus Protease. The peptides obtained after digestion with *S. aureus* protease at pH 4.0 (for amino acid compositions see Table IIIc in the supplementary material) were isolated from peptide maps (Figure 1c) and sequenced by the dansyl-Edman method. An aliquot of the digest was passed over Sephadex G-50 (see Materials and Methods). The first peak contained SP5, the largest peptide, and the only one isolated by gel filtration. Its N-terminal region was sequenced by means of the liquid-phase sequenator.

Ser-Asn-Ile-Ile-Lys-Gln-Leu-Glu

T1 T2

SP1

SP1 provides the overlap between tryptic peptides T1 and T2.

Gln-Glu-Gln-Met-Lys-Gln-Asp-Val-Pro-Ser-

T2 T3

Phe-Arg-Pro-Gly-Asp-Thr-Val-Glu

SP2

The enzyme did not cleave after glutamic acid in the second position. It was confirmed several times (also by the liquid-phase sequenator runs of the intact protein) that the residue is glutamic acid and not glutamine. SP2 combines the C-terminal part of T2 with the N-terminal part of T3.

Val-Lys-Val-Trp-Val-Val-Glu

T3 T4

SP3

This heptapeptide (SP3) gives a fluorescent spot on the peptide map when examined under UV light. There are two tryptic peptides with C-terminal Val-Lys, T3 and T12, but only T3 has a preceding Glu residue. Therefore, SP3 combines the C-terminal portion of T3 with the N-terminal portion of T4.

Gly-Ser-Lys-Lys-Arg-Leu-Gln-Ala-Phe-Glu

T4 T5a T7

SP4

SP4 overlaps tryptic peptides T4–T5a–T7.

The sequence of SP5 is given in Figure 2.

The liquid-phase sequenator run of the fragment gave clear results up to position 30 of SP5. Thus, it provided overlaps from tryptic peptide T7 to T12. In addition an aliquot (0.13 mg) of

peptide SP5 was cleaved with trypsin. The peptides were isolated on peptide maps and their amino acid composition was determined (Table IIIa of the supplementary material). The peptide corresponding to the seven N-terminal residues of SP5 could not be detected and the C-terminal Glu residue was not found.

The remaining peptides were identical with the tryptic peptides covering the region of the large fragment. An additional peptide, SP5/T4 = Lys + T11, was recovered due to the shorter cleavage time.

Arg-Thr-Gly-Lys-Ala-Ala-Arg-Ile-Lys-Glu-Arg-Leu-Asn
T20 T21 T22 T23 T24

SP6

The C-terminal peptide (SP6) from the *Staphylococcus aureus* protease digest overlaps tryptic peptides T20–T24.

Peptides obtained from cleavage of L19 with *Staph. aureus* protease in phosphate buffer at pH 7.8 were separated on a peptide map and their amino acid composition was determined (Table IIIc of the supplementary material). Cleavage occurred not only after all nine glutamic acid residues but also after some aspartic acid residues. The latter peptides (named SP′) fit well in the established sequence (see Figure 2).

Chymotryptic Peptides. The chymotryptic peptides were isolated from peptide maps (Figure 1d of the supplementary material) and were sequenced by the dansyl-Edman method. Their amino acid compositions are shown in Table IIId (supplementary material). One peptide, corresponding to positions 43–58, was not detected on the peptide maps. Most digestion experiments were performed with Tos-LysCH_2Cl-treated chymotrypsin with which usually only one tryptic site (between Lys-86 and Arg-87) was cleaved. Sometimes another site (between Arg-92 and Lys-93) was hydrolyzed. However, when untreated α-chymotrypsin from the same source was used, considerable tryptic activity was observed. The amino acid compositions of the resulting peptides (named CH′) are shown in Table IIId (supplementary material) and reconfirm the sequence of L19.

Ser-Asn-Ile-Ile-Lys-Gln-Leu-Glu-Gln-Glu-Gln-Met
T1 T2

CH1

This peptide (CH1) confirms the overlap between T1 and T2.

Lys-Gln-Asp-Val-Pro-Ser-Phe-Arg-Pro-Gly-Asp-Thr-
SP2

Val-Glu-Val-Lys-Val-Trp
SP3

CH2

Even though Pth-Trp could not be identified (its yield in position 18 is too low) it was concluded that CH2 contains Trp, as the peptide map spot showed a blue fluorescence under long-wavelength UV light. The C-terminal position of Trp is in accord with chymotryptic specificity and with the sequence of peptide SP3. CH2 provides the overlap between SP2 and SP3.

Val-Val-Glu-Gly-Ser-Lys-Lys-Arg-Leu-Gln-Ala-Phe
SP3 SP4

CH3

CH3 overlaps tryptic peptides T4–T7 and SP3–SP4, respectively.

Thr-Val-Arg-Lys-Ile-Ser-Asn-Gly-Glu-Gly-Val-
T9 T10 T11

Glu-Arg-Val-Phe
T12

CH4

CH4 provides overlaps between tryptic peptides T9, T10, T11, and T12.

Gln-Thr-His-Ser-Pro-Val-Val-Asp-Ser-Ile-Ser-Val-Lys

CH5

This peptide (CH5) is identical with the C-terminal portion of tryptic peptide T12.

Arg-Arg-Gly-Ala-Val-Arg-Lys-Ala-Lys-Leu-Tyr
T13a T15 T16a T18

CH6

CH6 overlaps tryptic peptides T13a, T15, T16a, and T18. A fragment of this peptide (Lys-Ala-Lys-Leu-Tyr) was observed occasionally.

Tyr-Leu-Arg-Glu-Arg-Thr-Gly-Lys-Ala-Ala-Arg-
T18 T19 T20 T21

Ile-Lys-Glu-Arg-Leu-Asn
T22 T23 T24

CH7

CH7 is the C-terminal chymotryptic peptide of L19 and contains tryptic peptides T18–T24. CH7a is identical with CH7 except for the lack of a Tyr residue in the first position. CH7a was sequenced up to the sixth residue.

Thermolysin Peptides. Thermolysin peptides were isolated from peptide maps (Figure 1e; see supplementary material). They provided additional information for overlaps and confirmed the previously sequenced regions. Most of the thermolytic peptides were sequenced completely or partially. Their amino acid compositions are given in Table IIIe and can, together with the sequence results, be found in the supplementary material.

A thermolytic peptide which was predicted to overlap positions 85–90 was not isolated in pure form. Therefore, a peptic digest of L19 was performed in order to establish this final overlap.

Peptic Peptides. A peptide map of a peptic digest of protein L19 is shown in Figure 1f (see supplementary material). The amino acid compositions of the peptides (Table IIIf; see supplementary material) are in good agreement with the established sequence of L19. The following peptides were sequenced by the dansyl-Edman method.

In P2 (for the sequence, see Figure 2, positions 13–26), only the five N-terminal residues were sequenced and were in agreement with results obtained independently.

Val-Lys-Val

P3

Ile-Arg-Asn-Arg-Gly-Leu-His-Ser-Ala-Phe
T7 T8 T9

P5

P5 again confirms the linkage of tryptic peptides T7, T8, and T9.

Ile-Ser-Val-Lys-Arg-Arg-Gly-Ala-Val-Arg-
CH5 CH6

Lys-Ala-Lys-Leu

P8

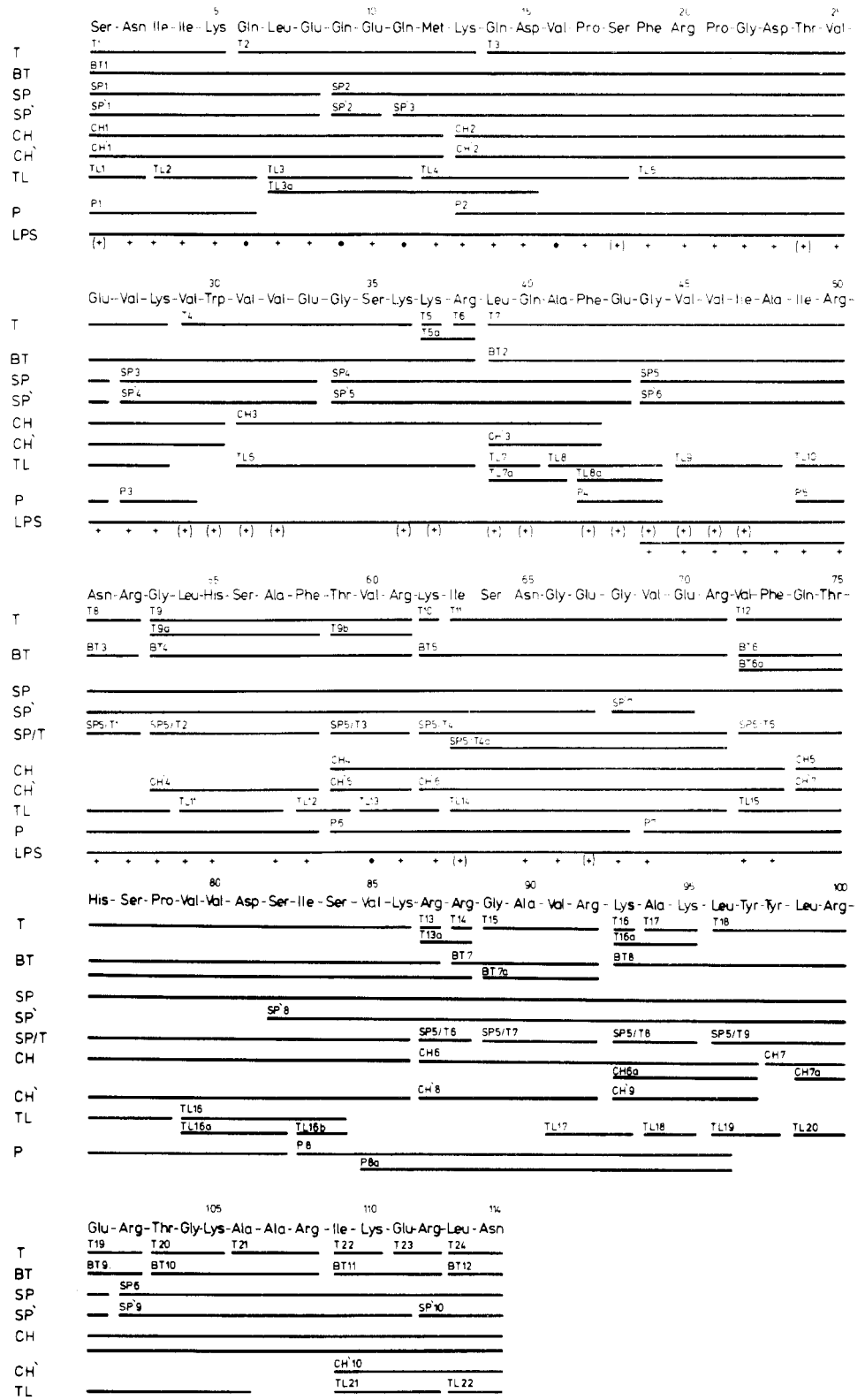
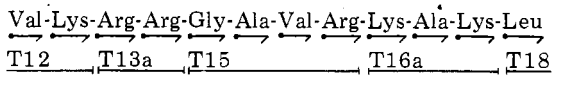


FIGURE 2: Complete amino acid sequence of ribosomal protein L19: T, tryptic peptides; BT, limited tryptic digest; SP, *S. aureus* protease digest at pH 4.0; SP′, *S. aureus* protease digest at pH 7.8; CH, Tos-LysCH_2Cl-chymotrypsin digest; CH′, untreated α-chymotrypsin digest; TL, thermolysin digest; P, peptic digest; LPS, automated Edman degradation in the liquid phase sequenator; +, unambiguous identification of the released Pth-amino acid by thin-layer chromotography; (+), weaker evidence.



P8a

Both of the above peptides (P8 and P8a) contain the region 85–90, the overlaps of which had not yet been established. When sequenced by the dansyl-Edman method considerable amounts of the previous amino acid were observed beginning with step 2. Thus, a clear result was not obtained concerning the missing link between peptides CH5 and CH6. However, by performing both the phenyl isothiocyanate coupling reaction and the trifluoroacetic acid cleavage reaction twice (B. Witt-

mann-Liebold, personal communication), contamination by the previous amino acid could be avoided. Thus, P8a was sequenced completely and the first eight residues of P8 were elucidated. Both peptides reveal the connection between tryptic peptides T12–T18 or CH5–CH6, respectively.

Liquid-Phase Edman Degradation. The N-terminal region of protein L19 was confirmed by liquid-phase Edman degradation on the intact protein giving clear results up to position 28, thus confirming the overlaps of tryptic peptides T1–T3. Furthermore, the less certain results (positions 29–47) were in accord with the established sequence. The results of the liquid-phase sequenator runs on SP5 are shown in the section Peptides Obtained from Cleavage with *S. aureus* Protease and in Figure 2.

The Complete Sequence of Protein L19. All tryptic peptides of protein L19 were completely sequenced. In a few cases when the discrimination between Asp/Asn and Glu/Gln was not possible, this information was obtained from other peptides. The alignment of all tryptic peptides can be summarized as follows: SP1 combines tryptic peptides T1 and T2 while SP2 provides the link between T2 and T3. SP3 is the bridge between T3 and T4. Peptide SP4 gives the order of T4, T5, T6, and T7 while the sequenced part of SP5 overlaps T7 through T11. Peptide CH4 provides an overlap from the C-terminal part of T9 through T10, T11, and T12. Peptic peptides P8 and P8a contain sequences overlapping tryptic peptides T12, T13, T14, T15, T16, and T17 ending at the N-terminal residue of T18. Finally, CH7 overlaps T18, T19, T20, T21, T22, and T23 to the C-terminal tryptic peptide T24. Thus, all tryptic peptides were aligned unequivocally and this was further confirmed by various other peptides (for the complete sequence see Figure 2). Every position in protein L19 was therefore sequenced several times. The total amount of protein used for sequence determination was 9.8 mg ($\simeq$ 750 nmol).

Discussion

The sequence of L19 is in good agreement with the amino acid composition obtained from hydrolysis of the whole protein. The single Trp residue was located initially by the fluorescence of the Trp-containing peptides, and confirmed by conversion of its 2-anilino-5-thiazolinone derivative into Pth-Trp followed by identification on silica gel thin-layer plates. Cysteine is absent in protein L19.

The molecular weight of protein L19 calculated from the sequence is 13 002. Its molecular weight was previously estimated to be 14 900 (Dzionara et al., 1970) or 12 800 (Zimmermann and Stöffler, 1976) by sodium dodecyl sulfate gel electrophoresis. The latter value is in excellent agreement with that derived from sequence analysis.

The net charge of +14 (including two His residues) is compatible with the very high isoelectric point of protein L19 (Kaltschmidt, 1971).

Peptide T7, which was not stained by ninhydrin or fluorescamine, was detected by spraying with phenanthrenequinone (Yamada and Itano, 1966). The amino acid compositions of phenanthrenquinone-stained peptides showed no measurable destruction of amino acid residues. The sole limitation is that only Arg-containing peptides are stained. Another sensitive procedure used in these studies was that of Reindel and Hoppe (1954). Peptides were well visualized and amino acid analysis showed no major destruction of amino acids.

Four glutamic acid containing peptide bonds were not cleaved by *Staphylococcus aureus* protease. This has previously been observed, for example, in protein L16, in which cleavage did not occur after Glu-75 (Brosius, 1976). The identification of Glu in these positions was unambiguously established. In addition, *S. aureus* protease was found to cleave after all Glu residues at pH 7.8 in phosphate buffer and also after some Asp residues, in agreement with the findings of Houmard and Drapeau (1972).

Clustering of hydrophobic amino acids occurs frequently, e.g.: Ile-Ile, Val-Val (three times), Val-Trp, Val-Val-Ile-Ala-Ile, Val-Phe, Leu-Tyr-Tyr-Leu. Basic amino acids are also clustered in doublets and triplets: Lys-Lys-Arg, Arg-Lys (twice), Lys-Arg-Arg. The hydrophobic and basic clusters are distributed throughout the protein chain. The doublet Glu-Gln is repeated in positions 8–11.

Protein L19 shares two identical pentapeptides and four tetrapeptides with other ribosomal proteins whose primary structures have been elucidated (Table IV; see supplementary material). This degree of homology is probably not statistically significant.

Based on the amino acid sequence of L19, secondary structure calculations were performed according to four predictive methods (Burgess et al., 1974; Chou and Fasman, 1974a,b; Chou et al., 1975; Nagano, 1977; Robson and Suzuki, 1976). Details of these results are given elsewhere (Dzionara et al., 1978). At least three of the four methods predict that positions 5–11, 37–43, 90–95, and 106–111 have helical structure, that positions 44–47, 59–62, and 79–80 show β structure, and that turns occur in positions 17–18, 21–23, 50–53, 65–66, and 103–104.

Acknowledgment

We wish to thank Mr. A. Lehmann for solid-phase sequencing, Mrs. E. Marzinzig and Mr. A. Geissler for liquid-phase sequencing, Dr. K. Katze for discussions, Dr. B. Wittmann-Liebold for providing computer data including secondary structure predictions and homologies, and Drs. H. G. Wittmann, Roger Garrett, and Harry Noller for critically reading the manuscript.

Supplementary Material Available

The sequence data of the thermolysin peptides, all tables containing the amino acid compositions of peptides (Tables IIIa–f), Table IV presenting homologies to other ribosomal proteins, and figures of several peptide maps (Figures 1b, 1d–f) as indicated in the text (13 pages). Ordering information is given on any current masthead page.

References

Brosius, J. (1976), Dissertation, Johann Wolfgang Goethe-Universität, Frankfurt am Main.

Brosius, J. (1978), *Biochemistry* (preceding paper in this issue).

Burgess, A. W., Ponnuswamy, P. K., and Scheraga, A. H. (1974), *Isr. J. Chem. 12,* 239–286.

Chou, P. Y., and Fasman, G. D. (1974), *Biochemistry 13,* 211–222.

Chou, P. Y., and Fasman, G. D. (1974), *Biochemistry 13,* 222–245.

Chou, P. Y., Adler, A. J., and Fasman, G. D. (1975), *J. Mol. Biol. 96,* 29–45.

Dixon, H. B. F., and Perham, R. N. (1968), *Biochem. J. 109,* 312–314.

Dzionara, M., Kaltschmidt, E., and Wittmann, H. G. (1970), *Proc. Natl. Acad. Sci. U.S.A. 67,* 1909–1913.

Dzionara, M., Robinson, S. M. L., and Wittmann-Liebold, B. (1978), *J. Supramol. Struct.* (submitted for publication).

Edman, P., and Begg, G. (1967), *Eur. J. Biochem. 1,* 80–91.
Garrett, R. A., Müller, S., Spierer, P., and Zimmermann, R. A. (1974), *J. Mol. Biol. 88,* 553–557.
Gray, W. R., and Hartley, B. S. (1963), *Biochem. J. 89,* 379–380.
Hindennach, I., Kaltschmidt, E., and Wittmann, H. G. (1971), *Eur. J. Biochem. 23,* 12–16.
Hitz, H., Schäfer, D., and Wittmann-Liebold, B. (1977), *Eur. J. Biochem. 75,* 497–512.
Houmard, J., and Drapeau, G. R. (1972), *Proc. Natl. Acad. Sci. U.S.A. 69,* 3506–3509.
Kaltschmidt, E. (1971), *Anal. Biochem. 43,* 25–31.
Kaltschmidt, E., and Wittmann, H. G. (1970), *Proc. Natl. Acad. Sci. U.S.A. 67,* 1276–1282.
Laursen, R. A. (1971), *Eur. J. Biochem. 20,* 89–102.
Laursen, R. A., Horn, M. J., and Bonner, A. G. (1972), *FEBS Lett. 21,* 67–70.
Morrison, C. A., Garrett, R. A., Zeichhardt, H., and Stöffler, G. (1973), *Mol. Gen. Genet. 127,* 359–368.
Morrison, C. A., Tischendorf, G., Stöffler, G., and Garrett, R. A. (1977), *Mol. Gen. Genet. 151,* 245–252.
Nagano, K. (1977), *J. Mol. Biol. 109,* 251–274.
Nierhaus, D., and Nierhaus, K. H. (1973), *Proc. Natl. Acad. Sci. U.S.A. 7,* 2224–2228.
Pongs, O., Bald, R., and Erdmann, V. A. (1973), *Proc. Natl. Acad. Sci. U.S.A. 70,* 2229–2233.
Reindel, F., and Hoppe, W. (1954), *Chem. Ber. 87,* 1103–1107.
Robson, B., and Suzuki, E. (1976), *J. Mol. Biol. 107,* 327–356.
Schiltz, E., and Reinbolt, J. (1975), *Eur. J. Biochem. 56,* 467–481.
Sonenberg, N., Wilchek, M., and Zamir, A. (1976), *Proc. Int. Congr. Biochem., 10th.*
Spies, J. R., and Chambers, D. C. (1949), *Anal. Chem. 21,* 1249–1266.
Stadler, H., and Wittmann-Liebold, B. (1976), *Eur. J. Biochem. 66,* 49–56.
Tischendorf, G. W., Zeichhardt, H., and Stöffler, G. (1974), *Mol. Gen. Genet. 134,* 187–208.
Tischendorf, G., Zeichhardt, H., and Stöffler, G. (1975), *Proc. Natl. Acad. Sci. U.S.A. 72,* 4820–4824.
Wittmann-Liebold, B. (1973), *Hoppe-Seyler's Z. Physiol. Chem. 354,* 1415–1431.
Wittmann-Liebold, B., and Lehmann, A. (1975), in Solid Phase Methods in Protein Sequence Analysis, Laursen, R. A., Ed., Rockford, Ill., Pierce Chemical Co., pp 81–90.
Wittmann-Liebold, B., Graffunder, H., and Kohls, H. (1976), *Anal. Biochem. 75,* 621–633.
Yamada, S., and Itano, H. A. (1966), *Biochim. Biophys. Acta 130,* 538–540.
Zeichhardt, H. (1976), Dissertation, Freie Universität, Berlin.
Zimmermann, R. A., and Stöffler, G. (1976), *Biochemistry 15,* 2007–2017.

Escherichia coli Uridine Diphosphate Galactose 4-Epimerase: Circular Dichroism of the Protein and Protein Bound Dihydronicotinamide Adenine Dinucleotide†

Shan S. Wong, Joseph Y. Cassim, and Perry A. Frey*

ABSTRACT: The circular dichroism spectra of *E. coli* UDP-galactose-4-epimerase in its native (epimerase·NAD^+) and reduced (epimerase·NADH·UMP) forms between 190 and 400 nm are presented. The reduced form exhibits a large positive circular dichroism band at 340 nm attributed to NADH in the complex. Relative to the small negative band exhibited at this wavelength by free NADH itself, the rotational strength of enzyme-bound NADH is some 50 times larger than that of free NADH, while the oscillator strengths and other spectral characteristics are similar. This enhancement reflects dissymetric interactions involving the 340-nm transition and is most consistent with the dihydropyridine ring of NADH being highly immobilized in the reduced complex. In the 200- to 230-nm region both enzyme forms exhibit a negative band at 220 nm and a negative shoulder at 208 nm. The ellipticities of the reduced form are minimally 7% greater at both band positions than those of the native form. The spectra are interpreted to indicate that conversion of the native to the reduced form is accompanied by an increase in α-helix structure at the expense of unordered structure.

The UDP-galactose 4-epimerase catalyzed interconversion of UDP-galactose and UDP-glucose has been shown to proceed by the following pathway

$$\begin{aligned} &\text{E·NAD}^+ \\ &\quad + \quad \rightleftharpoons \text{E·NAD}^+\text{·UDP-Gal} \\ &\text{UDP-Gal} \\ &\qquad \rightleftharpoons \text{E·NADH·UDP-4-ketopyranose} \\ &\qquad\qquad\qquad\qquad\qquad\qquad \text{E·NAD} \\ &\qquad\qquad \rightleftharpoons \text{E·NAD}^+\text{·UDP-Glc} \rightleftharpoons \quad + \\ &\qquad\qquad\qquad\qquad\qquad\qquad \text{UDP-Glc} \end{aligned}$$

† From the Department of Chemistry (S.S.W. and P.A.F.) and the Department of Biophysics (J.Y.C.), The Ohio State University, Columbus, Ohio 43210. *Received July 21, 1977.* This work was supported by Grant No. AM 13502 from the National Institute of Arthritis, Metabolism, and Digestive Diseases.

0006-2960/78/0417-0516$01.00/0