Potter, M., Mushinski, E. B., and Glaudemans, C. P. J. (1972), *J. Immunol. 108*, 295.
Potter, M., Rudikoff, S., Padlan, E. A., and Vrana, M. (1976), Antibodies in Human Diagnosis and Therapy, Haber, E., and Krause, R. M., Ed., New York, N.Y., Raven Press.
Rudikoff, S., Potter, M., Segal, D., Padlan, E. A., and Davies, D. R. (1972), *Proc. Natl. Acad. Sci. U.S.A. 69*, 3689.
Segal, D. M., Padlan, E. A., Cohen, G. H., Rudikoff, S., Potter, M., and Davies, D. R. (1974), *Proc. Natl. Acad. Sci. U.S.A. 71*, 4298.
Stone, M. J., and Metzger, H. (1968), *J. Biol. Chem. 243*, 5977.
Streefkerk, D. G., de Bie, M. J. A., and Vliegenthart, J. F. G. (1974), *Carbohydr. Res. 33*, 249.
Vrana, M., Tomašic, J., and Glaudemans, C. P. J. (1976), *J. Immunol. 116*, 1662.

Amino Acid Sequence of Normal (Microheterogeneous) Porcine Immunoglobulin λ Chains†

Jiří Novotný,* František Franěk, Michael N. Margolies,§ and Edgar Haber

ABSTRACT: The partial amino acid sequence of pooled, microheterogeneous pig immunoglobulin λ chains was determined previously (Franěk, F. (1970), *FEBS Lett. 8*, 269; Novotný, J., and Franěk, F. (1975), *FEBS Lett. 58*, 24). In the present study, citraconylated pig λ chains were digested by trypsin under conditions in which some of the ε-amino groups of lysine residues unmask. The resulting fragments were purified by gel filtration and ion-exchange chromatography at pH 3.0 in buffers containing urea; some of the fragments were found to be of intermediate size (i.e., larger than normal tryptic peptides but smaller than "citraconyl" peptides), thus permitting overlap information and amino acid sequences of all the 14 tryptic peptides to be deduced from amino acid compositions and partial amino acid sequences of selected fragments. In addition to completing the major amino acid sequence of pig immunoglobulin λ chains, the present study demonstrates that it is possible to sequence microheterogeneous proteins with a suitable fragmentation strategy.

Sequencing of microheterogeneous polypeptide chains (such as immunoglobulin chains from pooled nonspecific immunoglobulins) involves the difficulty of isolating and sequencing fragments of non-uniform amino acid sequence. Most of the current approaches to this problem are due to pioneering work of Porter and his group on rabbit immunoglobulin γ chains (Porter, 1967, 1973) and Cebra's group on guinea pig immunoglobulin γ chains (Cebra et al., 1971). Previous work in our laboratory on pooled porcine λ chains led to the tentative amino acid sequence of large segments of these chains, namely, residues 1–46 and 72–93 from the V region[1] (Franěk et al., 1969b; Franěk, 1970), and residues 113–168, 169–206, and 207–214 (Novotný and Franěk, 1975a) from the constant (C) region (i.e., the sequence of the constant region was complete except for overlaps between residues 168–169 and 206–207). The two intrachain disulfide bonds were localized to half-cystine residues 21–89 and 136–195, respectively (Franěk et al., 1968). The purpose of the present publication is twofold: (1) To provide amino acid sequence data sufficient for derivation of the complete λ chain sequence; and (2) to report fragmentation and isolation procedures permitting better quantitation and improved yields of chain fragments. The latter point is of considerable importance in sequence work on microheterogeneous proteins; if a fragment is obtained from a homogeneous protein in low yield (e.g., 15% of the theoretical), this fact has minor impact on the elucidation of the complete primary structure of the protein. If, however, a fragment has been obtained from a microheterogeneous protein in 15% yield, it cannot be excluded that the remaining 85% is represented by variants with properties (amino acid sequence, electric charge, solubility) other than those of the fragment obtained. It would be extremely laborious and often impossible to prove unequivocally that the low yield of the fragment is solely due to technical difficulties such as incomplete cleavage and losses during isolation procedures.

The present work shows that it is possible to derive the complete amino acid sequence of major variants of pooled immunoglobulin λ chains by applying a single step fragmentation procedure, namely, partial tryptic hydrolysis.

Materials and Methods

Materials. Preparation of pig λ chain and its resolution into λ(15), λ(16), and λ(17) subpopulations was described previously in detail (Franěk and Zorina, 1967). DCC-trypsin was purchased from Calbiochem. Pepsin, thermolysin, citraconic anhydride, trinitrobenzenesulfonic acid, ethylenimine, and dithiothreitol were also commercial products. The reagents used for automatic Edman degradation were purchased from Beckman. Those used for manual Edman degradation were pure commercial products repurified as recommended by Edman (1970). Urea and guanidine hydrochloride used for preparation of buffers were prepared from technical products by the Service Laboratory of the Institute of Organic Chemistry and Biochemistry (Prague). Optical absorbance of 6 M solutions of these compounds at 280 nm was less than 0.1. The

† From the Institute of Organic Chemistry and Biochemistry, Czechoslovak Academy of Sciences, 16610 Prague 6, Czechoslovakia (J.N. and F.F.), and the Harvard Medical School and Massachusetts General Hospital, Boston, Massachusetts 02114 (J.N., M.N.M., and E. H.). *Received March 10, 1977.*

§ M.N.M. is an Established Investigator of the American Heart Association.

[1] Abbreviations used: V region, variable region; C region, constant region; DCC-trypsin, diphenylcarbamoyl chloride treated trypsin; Pth-amino acid, phenylthiohydantoin amino acid; SE-Sephadex, sulfoethyl-Sephadex C-25.

specific conductivity of 6 M urea was less than 5×10^{-5} siemens (S)/cm (siemens = ohm^{-1}).

Reaction of λ Chains with Citraconic Anhydride and Digestion of the Modified Chains with Trypsin. Three grams of λ(16) chains were dissolved in 500 mL of buffer (6 M guanidine hydrochloride–0.1 M sodium pyrophosphate, pH adjusted to 9.0 with sodium hydroxide). To this solution 7.2 mL of citraconic (2-methylmaleic) anhydride was added in 1.8-mL portions under constant stirring, the pH of the reaction mixture being maintained at 8.5 by dropwise addition of 4 M sodium hydroxide. After the reaction was completed, the modified chains were transferred into 0.2% ammonium bicarbonate (pH adjusted to 9.0 with ammonia) by passage through a column of Sephadex G-25 (6 × 62 cm, volume 1.8 L) equilibrated with ammonium bicarbonate, pH 9.0. For mild tryptic hydrolysis, 15 mg of DCC-trypsin (substrate-to-enzyme ratio 1:200) was added to 600 mL of the eluate containing the modified chains; after 2 h at 38 °C, the reaction mixture was lyophilized. For more extensive tryptic hydrolysis, 30 mg of DCC-trypsin was added to 600 mL of the eluate containing the modified chains; after 2 h at 55 °C, 15 mg of DCC-trypsin (total substrate-to-enzyme ratio 1:66) was added, and after 2.5 h the reaction mixture was frozen and freeze-dried.

Unmasking of Citraconylated λ Chain Tryptic Digests. Six grams of citraconylated λ chain tryptic hydrolysate was dissolved in 200 mL of buffer (6 M guanidine hydrochloride–0.1 M potassium formate, pH adjusted to 3.0 with formic acid) and incubated 24 h at 38 °C. The solution was directly applied to a column of Sephadex G-75 for gel filtration.

Determination of Free Amino Groups. Two milligrams of either unmodified or citraconylated λ chains was dissolved in 2 mL of 0.1 M sodium pyrophosphate, pH 10. To 1 mL of this solution, 20 μL of 1.8 M trinitrobenzenesulfonic acid was added and, after standing 15 min in the dark at room temperature, the reaction mixture was diluted with 2 mL of freshly dissolved sodium sulfite (1.5 mM) in 0.1 M sodium pyrophosphate, pH 9.0. Immediately after dilution, absorbance of the test solution (A^t) was measured at 420 nm and 280 nm against a blank solution (A^b). Percentage of free amino groups (N%) as compared with the standard blank solution was computed according to the formula $N(\%) = 100(A^b_{280}A^t_{420}/A^b_{420}A^t_{280})$.

Digestion of λ Chains with Pepsin. Five hundred milligrams of λ(16) chains was dissolved in 15 mL of 1 M formic acid and digested with 10 mg of pepsin at 38 °C overnight. The reaction mixture was then lyophilized.

Digestion of Peptide Fragments with Thermolysin. Four milligrams of a peptide was dissolved in 1 mL of 0.2% ammonium bicarbonate, pH 8.6, and treated for 4 h with thermolysin (1:50 w/w) at 37 °C. The hydrolysate was then diluted with 1 mL of water and lyophilized.

Gel Filtration. Tryptic digests of λ chains, dissolved in guanidine/formate buffer, pH 3.0, for unmasking the amino groups, were applied to a column of Sephadex G-75 equilibrated with 0.05 M formic acid, 6 M in urea. For gel filtration and rechromatography of chain fragments, the peptide was dissolved in a minimal volume of 0.05 M formic acid, 6 M in urea, and applied to a column of Sephadex G-75 equilibrated with the same buffer. The columns were eluted at a rate of 2 to 3 mL cm^{-1} h^{-1}. The eluent was monitored for absorbance at 280 nm.

Ion-Exchange Chromatography. Ion-exchange chromatography was performed as described previously (Novotný, 1971; Novotný and Franěk, 1975a).

Separation of Small Peptides. The peptides were purified by high-voltage electrophoresis (4 kV) at pH 1.9 in a buffer containing 50 mL of formic acid and 150 mL of acetic acid in a 1-L solution, or at pH 5.6 in a buffer containing 16 mL of pyridine and 4 mL of acetic acid in a 1-L solution, and by descending chromatography in the system 1-butanol–pyridine–acetic acid–water (15:10:3:12 by volume).

S-Sulfonation of Half-Cystine Residues. The method was adapted from Franěk and Zorina (1967). In a typical experiment, 1 g of chain fragments was dissolved in 30 mL of buffer (6 M guanidine hydrochloride–0.5 M ammonium chloride, pH adjusted to 8.6 with ammonia) and 2.25 mL of 0.1 M cupric sulfate was added. After addition of 2.25 g of solid sodium sulfite and dissolution of the crystals, the reaction mixture was incubated for 3 h at room temperature and then transferred into 0.2% ammonium bicarbonate, pH 8.6, by passage through a column of Sephadex G-25 equilibrated with 0.2% NH_4HCO_3. To prevent precipitation of peptide during desalting, 15 mL of a freshly prepared 6 M urea solution in 0.2% ammonium bicarbonate and 15 mL of guanidine–ammonium chloride buffer were applied to the column before applying the sample. The eluate containing S-sulfonated fragments was lyophilized.

Reduction of Disulfide Bonds and Aminoethylation of Half-Cystine Residues. The method of Raftery and Cole (1966) was used as modified by Franěk and Novotný (1969).

Preparation of Peptide Maps. Two to three milligrams of peptide was S-sulfonated as described above. The solution of the S-sulfo fragment in 0.2% ammonium bicarbonate, pH 8.6 (obtained as eluate from the Sephadex G-25 column after desalting the reaction mixture), was incubated with DCC-trypsin (1:50) for 2 h at 38 °C; the digest was diluted with 2 mL of water, frozen, and freeze-dried. One milligram of the hydrolysate was dissolved in several drops of buffer, pH 1.9 (formic acid–acetic acid–water 1:3:130 by volume), and spotted on Whatman 3 MM paper. Separation by high-voltage horizontal electrophoresis (4 kV) proceeded at pH 1.9 and, in the second direction, by descending chromatography in 1-butanol–pyridine–acetic acid–water (15:10:3:12 by volume). The peptides were detected with 0.2% solution of ninhydrin in acetone; arginine-containing peptides were detected by the Sakaguchi (1950) reaction as described by Jepson and Smith (1953).

Manual and Automated Edman Degradation. Manual Edman degradation was performed as described by Blombäck et al. (1966). The phenylthiohydantoin (Pth) amino acids were identified by thin-layer chromatography on silica gel precoated aluminum plates ("Silufol", Kavalier, Czechoslovakia) using ethyl acetate–acetic acid (Edman, 1970), chloroform–acetone (Grüner, 1970), and dimethylformamide–water solvent system (to detect Pth-arginine and Pth-histidine). Typically, 1 μmol of a fragment was degraded. Some fragments were degraded automatically with the use of the Beckman sequencer 890 C. Fragment T4–T8 (0.4 μmol) was degraded according to a modified Beckman "fast protein" program (1 M Quadrol buffer, benzene and ethyl acetate washings, double cleavage, vacuum drying steps moderately prolonged). Fragments T4–T14 and T6b–T14 (150 nM each) were sequenced using a 0.1 M Quadrol program as described by Brauer et al. (1975). Pth-amino acids obtained from all automated degradations were identified by gas–liquid chromatography (Brauer et al., 1975), and by two-dimensional thin-layer chromatography on polyamide sheets (Summers et al., 1973). Histidine was identified from the aqueous phase by high-pressure liquid chromatography on a Waters (Milford, Mass.) instrument by elution with 10% acetonitrile (Jackson and Burdick, spectral grade) in 0.01 M sodium acetate buffer, pH 4.0.

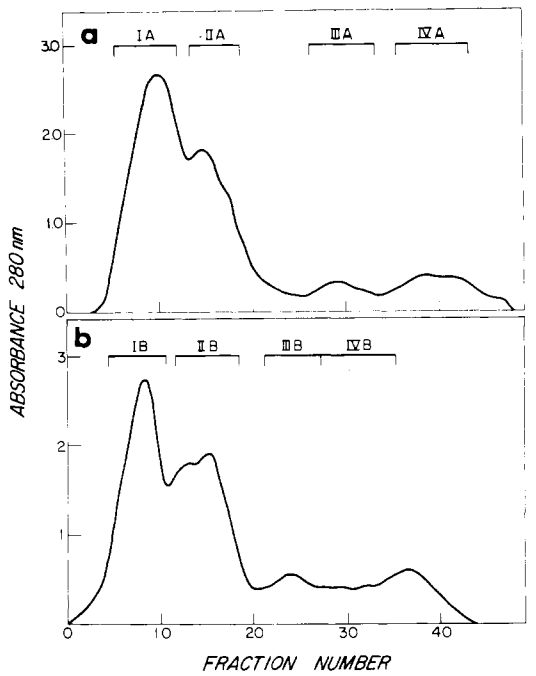

FIGURE 1: Fractionation of unmasked tryptic hydrolysates of citraconylated λ chains. (a) Six grams of a "mild" digest (enzyme-to-substrate ratio 1:200) in 200 mL of guanidine–formate buffer (pH 3.0) was applied to a column of Sephadex G-75 (11 × 130 cm, volume 12.4 L) equilibrated with 0.05 M formic acid, 6 M in urea. Peptide material from indicated fractions was pooled, freed from urea on Sephadex G-25 columns, and freeze-dried. Weight yields of individual pools: IA, 2.2 g; IIA, 1.4 g; IIIA, 0.440 g; IVA, 0.520 g (76% total yield). (b) Six grams of a "profound" digest (enzyme-to-substrate ratio 1:66) chromatographed under conditions identical with those described in a. Weight yields of individual pools: IB, 2.1 g; IIB, 1.9 g; IIIB, 0.330 g; IVB, 0.550 g (81% total yield).

Amino Acid Analysis. This was performed according to Spackman et al. (1958). The values given for threonine and serine represent extrapolations to zero time of hydrolysis; the values for valine and isoleucine are those taken from 70-h hydrolysates. Cysteine was determined either as aminoethylcysteine or as cysteic acid in samples oxidized by performic acid (Hirs, 1956). Tryptophan was determined either spectrophotometrically (Edelhoch, 1967) or after acid hydrolysis with addition of thioglycolic acid according to Matsubara and Sasaki (1969).

Peptide Nomenclature. Tryptic peptides are numbered T1 to T14 starting from the N terminus (Figure 6). The nomenclature of large fragments that consist of several tryptic peptides is based on the numbering system of tryptic peptides (e.g., T4–T8 designates a fragment comprising peptides T4, T5, T6, T7, and T8).

Results

Reaction of λ Chains with Citraconic Anhydride. To study the extent of reaction between citraconic anhydride and λ chain free amino groups, a preliminary experiment in which the chains were reacted with either a 5 M or a 30 M excess of anhydride relative to total free amino groups was performed. The fraction of amino groups not masked by the anhydride was determined by reaction with trinitrobenzenesulfonic acid. While 5 M anhydride excess was insufficient to mask all the amino groups, 30 M excess was found to be sufficient. However, after standing overnight at pH 8.5 or after freeze-drying, the extent of masking of amino groups dropped from 100% to 65–70%. At a higher pH value (pH 9.0), approximately 90% of the amino groups remained blocked after overnight standing at room temperature. Therefore, for tryptic hydrolysis limited to arginine residues, only freshly prepared solutions of citraconylated chains at pH 9.0 were used.

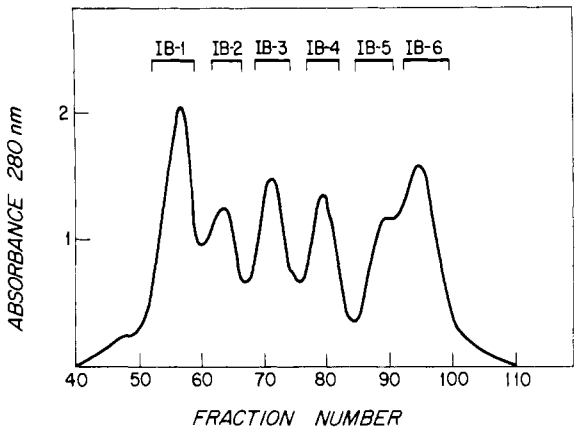

FIGURE 2: Fractionation of S-sulfonated fragments from fraction IB (Figure 1); 1.5 g was dissolved in 30 mL of 0.05 M formic acid, 6 M in urea, and applied to a column of Sephadex G-75 (7.6 × 87 cm, volume 3.8 L), equilibrated with 0.05 M formic acid, 6 M in urea. Peptide material from indicated fractions was pooled, freed from urea on Sephadex G-25 columns, and freeze-dried. Weight yields of individual pools: IB-1, 300 mg; IB-2, 125 mg; IB-3, 205 mg; IB-4, 133 mg; IB-5, 112 mg; IB-6, 145 mg (68% total yield).

Preparation of Large Fragments. To digest the citraconylated λ chains with trypsin, two different procedures were adopted. Mild hydrolysis (2 h at 38 °C, enzyme:substrate ratio 1:200) results in large "citraconyl" fragments; more extensive hydrolysis (4.5 h at 55 °C, enzyme:substrate ratio 1:66) results, in addition to citraconyl fragments, in fragments of intermediate size that are due to relative instability of citraconylamido groups at higher temperature (and, consequently, lower pH). Unmasked hydrolysates of citraconyl λ chains obtained by both mild and more extensive tryptic cleavage were resolved into four fractions by gel filtration on a Sephadex G-75 column (Figures 1a and 1b, respectively).

Fraction IB was S-sulfonated and rechromatographed on a smaller column of Sephadex G-75 (Figure 2). Each of the six fractions obtained (IB1–IB6) was further fractionated by ion-exchange chromatography on SE-Sephadex at pH 3.0. Apart from extremely acidic fragments emerging from the columns with the breakthrough volume and known to originate from the N-terminal variable region of the λ chains (Franěk and Novotný, 1969), four large fragments were isolated by this procedure from IB1, IB2, IB3, and IB6 (Figure 3). They were rechromatographed on Sephadex G-75 and SE-Sephadex either in S-sulfo form (fragments T4–T14, T4–T11, and T12b–T14) or in S-(2-aminoethyl) form (fragment T4–T8; aminoethylation of fragment T4–T8 facilitated further purification on SE-Sephadex at pH 3.0 by raising the net electrical charge of the fragment at this pH). Pure fragments obtained by these purification procedures displayed uniform N-terminal groups and amino acid compositions, which did not change by further rechromatography (Table I). As expected, on the basis of chromatography at pH 3, amino acids half-cystine, lysine, and histidine are present in stoichiometric amounts (S-sulfocysteine, ε-amino group of lysine, and the imidazole ring of histidine are the only side chains charged at this pH). The number of certain amino acids in some of the fragments cannot be expressed by integers; this fact suggests variability in the amino acid sequence of these fragments (Table I).

Fraction IIA (Figure 1a) was resolved into two subfractions by ion-exchange chromatography on SE-Sephadex at pH 3.0 (Figure 3, bottom left). The subfraction T6b–T14 was then S-sulfonated and rechromatographed on Sephadex G-75 and

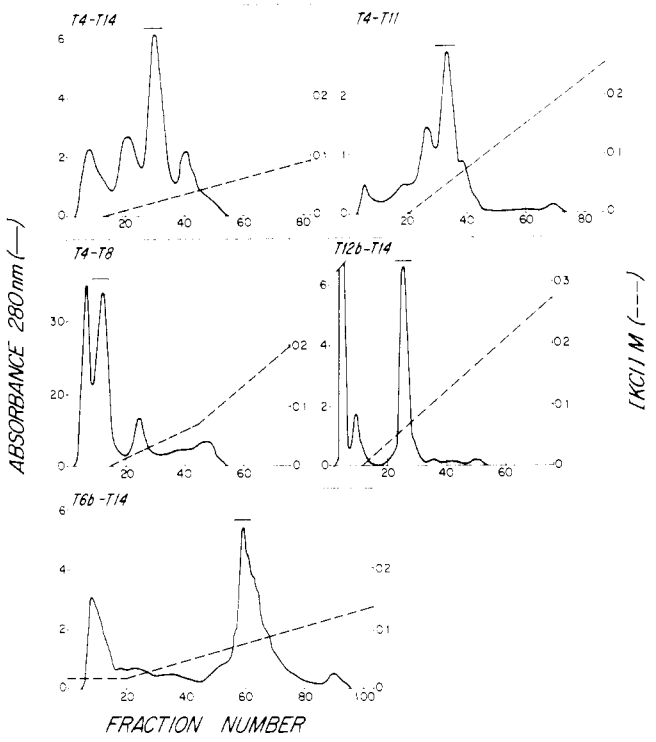

FIGURE 3: Ion-exchange chromatography, on SE-Sephadex at pH 3.0, of peptide fractions obtained by gel filtration (Figure 2). (Top left) Three hundred milligrams of fraction IB-1 was dissolved in 17 mL of the urea-formate buffer (pH 3.0) and applied to a 1.3 × 31 cm column (volume 41 mL). The mixer contained 140 mL of starting buffer, the reservoir 140 mL of 0.20 M KCl solution in the starting buffer. Weight yield of the fraction T4–T14 was 36 mg. (Top right) One hundred and twenty-five milligrams of fraction IB-2 was dissolved in 7 mL of buffer and applied to a 1.3 × 30 cm column. The mixer contained 120 mL of starting buffer. Weight yield of the fraction T4–T11 was 44 mg. (Middle left) Two hundred and five milligrams of fraction IB-3 was dissolved in 10 mL of the urea-formate buffer, pH 3.0, and applied onto a 1.3 × 31 cm column. The mixer contained 130 mL of starting buffer, the reservoir 130 mL of 0.20 M KCl solution in the starting buffer. Weight yield of the fraction T4–T8 was 60 mg. (Middle right) One hundred and forty-five milligrams of fraction IB-6 was dissolved in 5 mL of the urea-formate buffer, pH 3.0, and applied to a 1.3 × 32 cm column. The mixer contained 100 mL of starting buffer, the reservoir 100 mL of 0.25 M KCl solution in the starting buffer. Weight yield of fraction T12b–T14 was 36 mg. (Bottom left) Fraction IIA (1.4 g) were dissolved in 20 mL of the urea-formate buffer, pH 3.0, and applied to a 2.25 × 79 cm column (volume 315 mL). The mixer contained 1 L of starting buffer, the reservoir 1 L of 0.20 M KCl solution in the starting buffer. Weight yield of the fraction T6b–T14 was 314 mg.

SE-Sephadex thereby yielding the pure peptide T6b–T14 (Table I).

Preparation of Small Peptides. Peptides from fractions IVA and IVB (Figures 1a and 1b) were subjected to high-voltage paper electrophoresis. Sakaguchi-positive arginine-containing strips were cut out, sown on Whatman 3 MM paper, and further purified by descending chromatography. Two arginine-containing peptides, Arg-A and Arg-B, were isolated by this procedure. The amino acid compositions of these peptides are given in Table I. A peptic hydrolysate of intact λ(16) chains was resolved in the same way; one arginine-containing peptide, Arg-C, was isolated in a pure form (Table I).

Amino Acid Sequences of Peptides. The amino acid sequence of two small arginine-containing peptides, Arg-A and Arg-B, was determined by manual degradation (Figure 4). The nonapeptide Arg-B was sequenced up to the C terminus without difficulty. The hexadecapeptide Arg-A was sequenced through residue 13 providing an overlap of the Arg-B peptide. At positions 5 to 7 of the Arg-A peptide, no major amino acid residue could be identified because of extensive variability.

From 1 μmol, of fragment T12b–T14, 19 N-proximal amino acid residues were determined by manual degradation (Figure 4). An additional 2 μmol of the fragment T12b–T14 was hydrolyzed with thermolysin and the resulting peptides were resolved by paper techniques. In this way a pentapeptide was isolated in pure form, the amino acid sequence of which was shown by manual Edman degradation to be Ile-Val-Glu-Lys-Thr (Figure 4, residues 203–207).

The fragment T4–T8 was subjected to automated Edman degradation using a 1 M Quadrol program as described in Materials and Methods. The repetitive yield of the degradation, based on the gas chromatographic yields of Pth-alanine residues, was 96%. The N-proximal amino acid sequence of 31 residues thus established was identical with the sequence of the fragment T4–T14 described below (see also Figure 4).

Fragments T4–T14 and T6b–T14 were sequenced automatically using 0.1 M Quadrol procedure; the repetitive yields of the degradations were close to 95%. Of the fragment T4–T14, 37 N-proximal residues were sequenced, 33 of which were positively identified (Figure 4). Three steps (cycles 32–34) displayed a high degree of variability which prevented a single major sequence to be deduced; positive identification of S-sulfocysteine residue (cycle 27) was not attempted. For the fragment T6b–T14, 39 N-proximal residues were sequenced with positive identification of amino acid residues in 38 positions (the S-sulfocysteine residue in position 37 was not identified). In positions 7 and 9–13, two amino acid residues were found in different proportions, permitting the deduction of one major and one minor amino acid sequence. There was apparently a single uniform sequence after position 13 (see Figure 4).

The fragments T4–T14, T6b–T14, T12b–T14, T4–T8, and T4–T11 were further characterized by peptide maps of their tryptic digests (Figure 5). From the peptide map of each of the fragments, the individual peptide spots were eluted and their amino acid composition was determined. They were found to be in accord with the composition expected on the basis of previously published tentative sequence of the constant region of the λ chains (Novotný and Franěk, 1975a,b).

Derivation of the Entire Amino Acid Sequence. Figure 4 summarizes the sequence data from which the major amino acid sequence of the whole λ chains can be derived (Figure 6). In a previous publication (Novotný and Franěk, 1975a), the sequence of the 101 C-proximal amino acid residues was reported with overlaps of residues 168–169 and 206–207 lacking. The 206–207 overlap is now established by the partial sequence of the fragment T12b–T14 (the pentapeptide Ile-Val-Glu-Lys-Thr). The 168–169 overlap is deduced from the compositions and peptide maps of fragments T4–T8 and T4–T11 spanning residues 63–151 and 63–173, respectively. Thus, sequence of residues 113–214 is confirmed and together with the N-proximal sequences of the fragments T4–T14 and T6b–T14 accounts for a continuous sequence of residues 63–214. Since 46 N-terminal amino acid residues were determined previously (Franěk, 1970), it remains to establish the amino acid sequence of 16 residues from the variable part, i.e., residues 47–62 (Figure 4); these constitute the peptide Arg-A (Table II), the amino acid sequence of which was deduced from manual Edman degradation of the peptides Arg-A and Arg-B. The assignment of this peptide to residues 47–62 is further supported by the amino acid composition and the N-terminal residue of the peptide Arg-C. This peptide, containing 1 lysine residue, 2 arginine residues, and an N-terminal tyrosine, spans

TABLE I: Amino Acid Compositions of Large Fragments and Small Peptides.

	T4-T14	T6b-T14	T4-T11	T4-T8	T12b-T14	Arg-A	Arg-B	Arg-C
Lys	10.0	8.0	7.9	4.0	2.0			1.2
His	2.1	2.0	1.2	1.1	1.0			
Arg						2.0	2.0	1.9
Asp	10.2	6.4	9.0	7.7	0.7	1.8		1.7
Thr	19.5	15.4	14.3	10.1	4.7	1.9[a]	1.0	2.4[a]
Ser	15.8	12.4	10.0	6.4	5.5	2.1	1.7	2.5[a]
Glu	13.4	9.6	8.7	6.0	4.0			
Pro	7.3	6.9	6.8	5.9	0.9	2.2	2.0	1.9
Gly	15.0	10.5	11.9	9.8	2.0	1.3	1.1	2.8
Ala	16.0	10.0	11.8	10.0	3.0	1.2		1.1
½-Cystine	4.1	2.9	2.2	2.2	2.0			
Val	10.7	10.0	7.3	5.8	2.8	1.7	1.0	1.0
Ile	4.8	3.0	3.2	3.7	0.9	1.0		1.0
Leu	11.0	8.1	8.2	8.6	1.7			1.1
Tyr	5.0	3.3	4.4	2.9		1.0		0.8
Phe	6.5	4.1	4.8	5.1	1.1			1.1
Trp	2.0	1.2	1.8	0.8	+[b]			
Total:	153.4	113.8	113.5	90.1	32.3	16.2	8.8	20.5
N terminus:	Phe	Gly	Phe	Phe	Leu	Thr	Ser	Tyr
Yield of peptide:	8%	23%	3%	4%	6%			

[a] Values not extrapolated to zero time of hydrolysis. [b] Determined qualitatively.

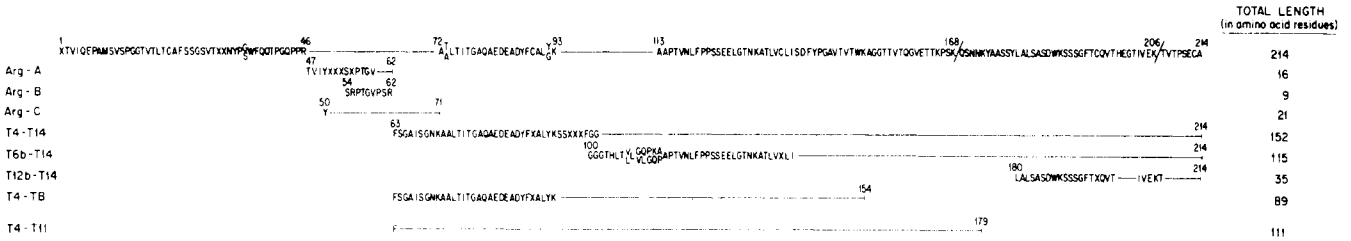

FIGURE 4: Partial amino acid sequences of selected peptides and fragments. At the top is the partial amino acid sequence determined previously on λ(15) chains (Franěk et al., 1969b; Franěk, 1970; Novotný and Franěk, 1975a). The remainder of the peptides, obtained from λ(16) chains, are reported here. The sequences are given in the one letter code. X at the N terminus stands for 2-pyrrolidone carbonyl residue; at other positions, X denotes a hypervariable position or an unidentified amino acid residue.

residues 50–71. On the basis of the derived sequence, its amino acid composition is expected to be Lys_1, Arg_2, Asx_2, Thr_3, Ser_4, Pro_2, Gly_3, Ala_1, Val_1, Ile_1, Tyr_1, Phe_1; the actual amino acid composition found (Table I) is in good accord with this expectation.

Discussion

Fragmentation Procedures and Derivation of the Entire Sequence. Since in working with microheterogeneous proteins the yields (and, implicitly, solubility and ease of handling) of protein fragments are of prime importance, a fragmentation strategy was sought that would provide maximum overlap information with minimal number of fragmentation and separation steps required to isolate pure fragments. In pilot experiments, a certain degree of instability of citraconylamido groups at pH 8.5–8.9 was observed; this property has been exploited to advantage to obtain fragment overlap information by digesting the citraconylated λ chains with trypsin under conditions in which a fraction of lysine ε-amino groups unmask (increased temperature and decreased pH). Some of the fragments originating from this hydrolysate are of intermediate size (they are larger than normal tryptic peptides but smaller than expected "citraconyl" fragments, e.g., T4–T11, T6b–T14, T12–T14, T4–T8), and complete overlap information is provided by comparing the length of individual fragments with their amino acid compositions and peptide maps (Figures 4 and 5). Since incomplete amino acid sequences of porcine λ chains

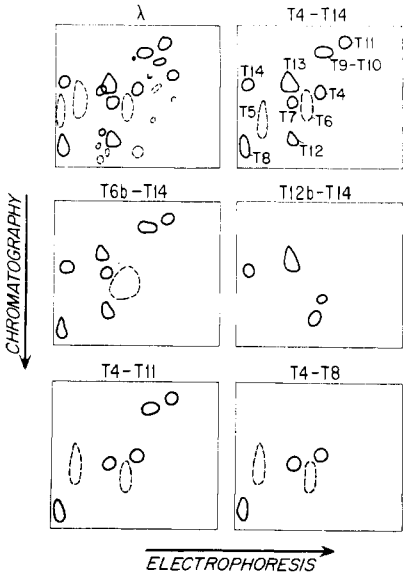

FIGURE 5: Peptide maps of tryptic hydrolysates of intact λ chains and selected peptides. Origin is at top left corner, the anode is at the right.

described in previous papers were also obtained from partial tryptic digest (Franěk, 1970; Novotný and Franěk, 1975a), the present work demonstrates that it is feasible to derive nearly the entire sequence of a microheterogeneous polypeptide chain

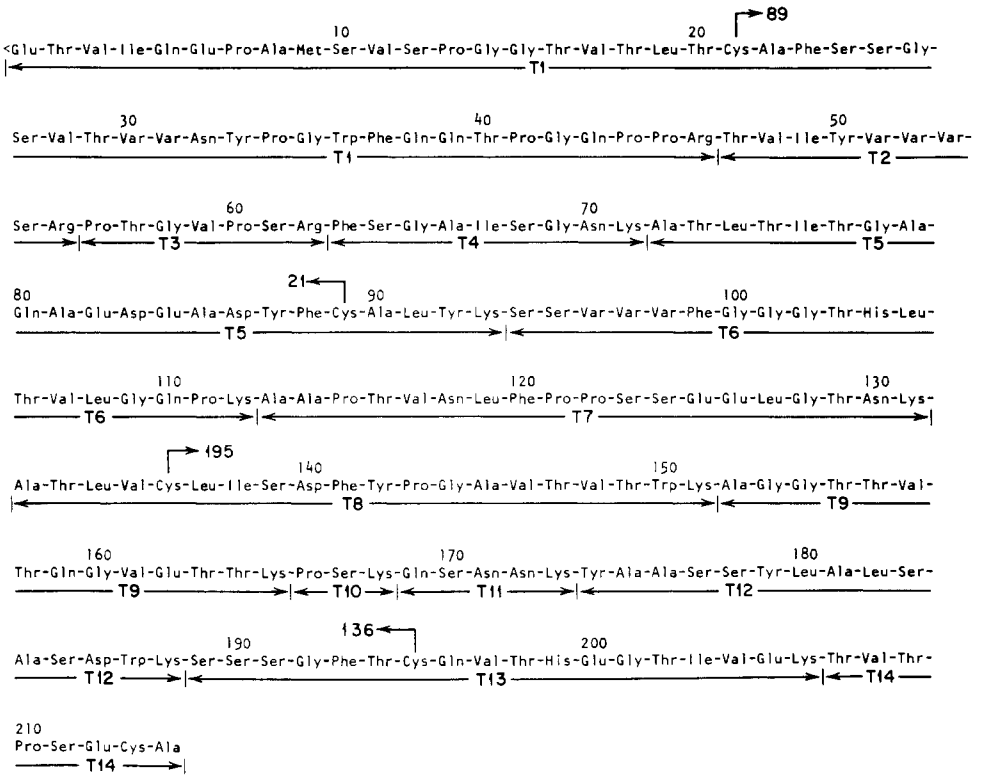

FIGURE 6: The major amino acid sequence of porcine λ chains. Var denotes positions of high variability where no major amino acid residue has been found (see text for details). Right angle arrows denote connections of half-cystine residues in the two intrachain disulfide bridges.

TABLE II: Comparison of Calculated and Found Amino Acid Compositions of the Peptide Arg-A.

	Whole λ(16)[a]	T4-T14	T2a	T2a + (T4-T14)	Arg-A	
					Calcd: whole λ(16)-T2a-(T4-T14)	Found
Lys	9.9	10.0		10.0		
His	2.2	2.1		2.1		
Arg	3.3		1.0	1.0	2.3	2.0
Asp	13.4	10.2	1.4	11.6	1.8	1.8
Thr	28.9	19.5	6.4	25.9	3.0	1.9
Ser	24.7	15.8	6.6	22.4	2.3	2.1
Glu	19.3	13.4	6.1	19.5	(−0.2)	
Pro	15.2	7.3	5.9	13.2	2.0	2.2
Gly	20.8	15.0	4.6	19.6	1.2	1.3
Ala	18.3	16.0	2.3	18.3		1.2
½-Cystine	5.0	4.1	1.0	5.1		
Val	16.0	10.7	4.3	15.0	1.0	1.7
Met	1.0		1.0	1.0		
Ile	6.3	4.8	1.0	5.8	0.5	1.0
Leu	12.6	11.0	1.6	12.6		
Tyr	7.0	5.0	1.1	6.1	0.9	1.0
Phe	7.6	6.5	1.8	8.3	(−0.7)	
Trp	3.4	2.0	1.0	3.0	0.4	
Total:	214.9	153.4	47.1	200.5	15.4	16.2

[a] From Franěk (1967).

with the use of a single enzyme (trypsin) by varying the conditions of hydrolysis or the degree of masking of lysine ε-amino groups or both; only two short peptides originated from peptic and thermolytic hydrolysates, respectively. Although peptides T4-T8 and T4-T11 arose from tryptic cleavage at unmasked Lys residues, peptides T6b-T14 and T12b-T14 were cleaved carboxy terminal to an aromatic amino acid. This suggests that there was chymotryptic activity present in DCC-trypsin preparation.

The only methods used in this work to separate large fragments were gel filtration on Sephadex G-75 (or G-50) and ion-exchange chromatography on SE-Sephadex at pH 3.0, in buffers containing a high concentration of urea. Unlike separations on Dowex-type resins where peptide yields are often poor, the total recovery of peptide material was high (70-80%), permitting accurate quantitation of the whole fragmentation and separation process. In general, 20-25% of the porcine λ chains were split in the switch region between variable and

constant halves, giving rise to the fragment T6b–T14. Yields of the four other fragments described in this work varied from 3 to 8%. Altogether these fragments account for approximately 40% of the material expected to be recovered from the fragmentation and separation process. The remaining 60% is represented by peptides and fragments scattered in minor fractions (e.g., IA, IIB, IIIB, IB-4, and IB-5) that were partially characterized during the sequence work and, revealing no additional sequence information, were not described in the present paper. Though only a single method, namely, absorbance at 280 nm, was used to detect the fragments during separation procedures, it is highly unlikely—in view of the length of the fragments separated and their relatively high content of aromatic amino acid residues—that there are variants that escaped detection. Summing the results of fragmentation and separation procedures, it is apparent that (a) the methods used to purify the fragments meet the requirements imposed for sequencing of microheterogeneous proteins, and that (b) the main difficulties encountered in accounting for the quantitative balance of the whole protein under investigation are due to incomplete fragmentation (partial instability of the citraconyl group, nonspecific cleavages and, possibly, incomplete cleavage of certain bonds with trypsin).

Degree of Variability in the Variable Region. As discussed above, not more than about a half of the existing variants could be studied in sufficient detail; for this reason, the variability reported in the λ chain sequence should be considered as a minimum. Some of the variants escaped detection while others could not be identified with certainty due to technical problems inherent in sequencing methods. However, for reasons discussed above, the overall picture of amino acid variability gained from the present work is representative of the whole porcine λ chain population. The only amino acid replacements that could be identified with certainty in the framework region were the Gly × Ser exchanges in position 35 (Franěk, 1971), Thr × Ala at 73, and Gly × Tyr at 92 (Franěk et al., 1969a,b). On the other hand, variability encountered in complementarity regions (positions 31–32, 51–53, and 96–98) was high and no prevailing amino acid residue could be identified in these positions. The overall pattern of amino acid variability of pooled nonspecific porcine λ chains resembles closely that found with some murine λ chains (Weigert et al., 1970); i.e., they possess almost constant framework regions and highly variable complementarity regions. This is in marked contrast to the variability pattern of pooled equine κ chains (Gibson, 1974) and porcine κ chains (Novotný et al., 1970) in which N-proximal framework regions display significant variability.

Degree of Variability in the Constant Region. Pooled pig immunoglobulin λ chains were shown (Franěk, 1967; Franěk and Zorina, 1967) to be resolved by ion-exchange chromatography into three distinct subpopulations [λ(15), λ(16), λ(17)] according to the total number of basic amino acid residues they contain. Most of the previous sequence work was performed on λ(15) chains, while the results described in this paper were obtained from λ(16) chains. No differences between the amino acid sequences of these two subpopulations have been found, suggesting that the same major sequence exists in the constant region of the whole porcine λ chain population. However, as discussed above, in view of incomplete yields of fragments during fragmentation procedures, one cannot exclude the possibility of single amino acid replacements (analogous to allotypic and isotypic replacements found in rabbit and human immunoglobulin light chains) occurring in pig λ chain constant region.

Interesting interspecies similarities, as well as differences, exist among the sequences of murine, human, and porcine λ chains. It has been shown previously (Novotný and Franěk, 1975b) that, in the constant region of these chains, polypeptide chain segments that are folded inside the domain are highly conservative in phylogeny, whereas segments that are exposed to the surface and only marginally participate in folding architecture are conserved much less. Using the completed amino acid sequence of the porcine λ chains described in this paper, the analysis of interspecies homology can be extended to the variable domains. Results of such an analysis, described in detail in a separate paper (Novotný et al., 1977) reveal that (a) certain V region surface loops, in contrast to C region surface loops, are highly conserved among species, and (b) the highest degree of phylogenetic conservation is always associated with segments comprising the core of the domain; this is particularly striking since the V domain has been shown to be packed inside-out in relation to the C domain, the core-forming segments of the V domain being different from those that form the core of the C domain (Edmundson et al., 1975).

Acknowledgments

We thank Mr. K. Grüner (Prague) and Mr. A. Brauer (Boston) for their expert assistance with automatic Edman degradation.

References

Blombäck, B., Blombäck, M., Edman, P., and Hessel, B. (1966), *Biochim. Biophys. Acta 115,* 371.
Brauer, A. W., Margolies, M. N., and Haber, E. (1975), *Biochemistry 14,* 3029.
Cebra, J. J., Ray, A., Benjamin, D., and Birshtein, B. (1971), *Prog. Immunol. Int. Congr. Immunol. 1st,* 269.
Edelhoch, H. (1967), *Biochemistry 6,* 1948.
Edman, P. (1970), in Protein Sequence Determination, Needleman, S. B., Ed., Berlin, Springer-Verlag, p 211.
Edmundson, A. B., Ely, K. R., Abola, E. E., Schiffer, M., and Panagiotopoulos, N. (1975), *Biochemistry 14,* 3953.
Franěk, F. (1967), *Nobel Symp. 3,* 211.
Franěk, F. (1970), *FEBS Lett. 8,* 269.
Franěk, F. (1971), *Eur. J. Biochem. 19,* 176.
Franěk, F., Keil, B., Novotný, J., and Šorm, F. (1968), *Eur. J. Biochem. 3,* 422.
Franěk, F., Keil, B., and Šorm, F. (1969b), *Eur. J. Biochem. 11,* 170.
Franěk, F., Keil, B., Thomas, D. W., and Lederer, E. (1969a), *FEBS Lett. 2,* 309.
Franěk, F., and Novotný, J. (1969), *Eur. J. Biochem. 11,* 165.
Franěk, F., and Zorina, O. M. (1967), *Collect. Czech. Chem. Commun. 32,* 3229.
Gibson, D. (1974), *Biochemistry 13,* 2776.
Grüner, K. (1970), *Chem. Listy 64,* 1160.
Hirs, C. W. H. (1956), *J. Biol. Chem. 219,* 611.
Jepson, J. B., and Smith, I. (1953), *Nature (London) 172,* 1100.
Matsubara, H., and Sasaki, R. M. (1969), *Biochem. Biophys. Res. Commun. 35,* 175.
Novotný, J. (1971), *FEBS Lett. 14,* 7.
Novotný, J., and Franěk, F. (1975a), *FEBS Lett. 58,* 24.
Novotný, J., and Franěk, F. (1975b), *Nature (London) 258,* 641.
Novotný, J., Franěk, F., and Šorm, F. (1970), *Eur. J. Biochem. 14,* 309.
Novotný, J., Vítek, A., and Franěk, F. (1977), *J. Mol. Biol.* (in press).
Pisano, J. J., and Bronzert, T. J. (1969), *J. Biol. Chem. 224,*

5597.
Porter, R. R. (1967), *Biochem. J. 105*, 417.
Porter, R. R. (1973), *Science 180*, 713.
Raftery, M. A., and Cole, R. D. (1966), *J. Biol. Chem. 241*, 3457.
Sakaguchi, S. (1950), *J. Biochem. (Tokyo) 37*, 231.
Spackman, D. H., Stein, W. H., and Moore, S. (1958), *Anal. Biochem. 30*, 1190.
Summers, M. R., Smythers, G. W., and Oroszlan, S. (1973), *Anal. Biochem. 53*, 624.
Weigert, M. G., Cesari, I. M., Yonkovich, S. J., and Cohn, M. (1970), *Nature (London) 228*, 1045.

Functional *lac* Carrier Proteins in Cytoplasmic Membrane Vesicles Isolated from *Escherichia coli*. 1. Temperature Dependence of Dansyl Galactoside Binding and β-Galactoside Transport†

Hélène Therisod, Lucienne Letellier, Rudolf Weil, and Emanuel Shechter*

ABSTRACT: Dansyl galactoside (6′-(*N*-dansyl)aminohexyl 1-thio-β-D-galactopyranoside) binds in an energy-dependent way to the *lac* carrier proteins of membrane vesicles isolated from *Escherichia coli* cells. The binding is not followed by transport but it is accompanied by a large increase in fluorescence and a shift of the emission maximum to a lower wavelength. These properties make it possible to titrate the amount of *lac* carrier proteins which become accessible upon energizing the membrane. The temperature dependence of the binding of dansyl galactoside has been determined in membrane vesicles of *E. coli* K 1059 and ML 308225. The number of binding sites of dansyl galactoside decreases with decreasing temperature. The range over which this occurs overlaps that at which the conformational order–disorder transition of the membrane lipids takes place. The dissociation constant is temperature independent. The initial rates of binding are graphed according to the Arrhenius equation. This is represented as nmol of dansyl galactoside bound per min either per mg of protein or per nmol of accessible *lac* carrier proteins. In the former case, the best fit to the data requires two lines. The line in the low-temperature range has a steeper slope. However, in the latter case, the data can be fit either by only one line (K 1059 membranes) or by two lines but with a much less pronounced difference in slope (ML 308225 membranes). There is a striking parallel between the temperature dependence of the initial rates of dansyl galactoside binding to a given membrane and the temperature dependence of the initial rates of β-galactoside transport by the same membrane. This parallelism supports the notion that the carriers which participate in the transport (functional carriers) are the ones which become accessible to the external medium upon energizing. It is concluded that the change in slope usually observed in the Arrhenius plot of transport in the temperature range extending over the order disorder transition is mainly a result of a change in the number of functional *lac* carriers.

Isolated cytoplasmic membrane vesicles have become a very useful tool for the study of active transport (Kaback, 1974). Kaback and co-workers (Schuldiner et al., 1976a) have carefully analyzed the mechanism by which the active transport of a variety of metabolites by cytoplasmic membranes of *E. coli* is coupled to the oxidation of D-lactate or to the artificial electron donor, phenazine methosulfate. A recent approach to the study of β-galactoside transport has been the use of reporter groups covalently linked to the sugar (Reeves et al., 1973; Schuldiner et al., 1975a; Rudnick et al., 1975). Among these, dansyl galactosides have been shown to bind specifically and in an energy-dependent way to the *lac* carrier proteins without being transported across the membrane (Schuldiner et al., 1975b,c, 1976b). Studies involving these dansyl galactosides have yielded new information on the mechanism of β-galactoside transport at the molecular level (Schuldiner et al., 1976a).

We have shown previously that, in cytoplasmic membrane vesicles isolated from *E. coli*, the active transport coupled to D-lactate oxidation responds to a change in conformation of the membrane lipids (Shechter et al., 1974, 1975). This change can be induced by temperature variations and involves a transition from a disordered fluid state of the lipids at high temperature to an ordered state of the lipids at low temperature (Ranck et al., 1974). At temperatures at or below those of the conformational transition, Arrhenius plots of active transport display a slope which is in general more pronounced than that at temperatures above the transition (Shechter et al., 1974, 1975; Overath and Träuble, 1973; Overath et al., 1976). The exact reason for this change in slope has been the subject of considerable study.

This report describes the use of dansyl galactoside in the investigation of the mechanism by which the disorder-to-order transition associated with the membrane lipids affects the active transport of β-galactosides. We show that the relationship between lipid conformation and transport is an indirect one and that it involves a change in the number of functional *lac* carrier proteins that participate in the transport process as the conformation of the lipids changes.

Materials and Methods

Growth of Bacteria and Preparation of Membrane Vesicles.

† From the Laboratoire des Biomembranes (LA 272), Département de Biochimie et de Biophysique, Université de Paris-Sud, 91405 Orsay, France (H.T., L.L., and E.S.), and the Sandoz Forschungsinstitut, Wien, A 1235, Austria (R.W.). *Received February 2, 1977.* This work was supported in part by a grant from the Délégation Générale à la Recherche Scientifique et Technique, Comité des Membranes Biologiques.