Journal of Chromatography, 266 (1983) 491–500
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROMSYMP. 043

TWO-DIMENSIONAL PEPTIDE MAPPING BY REVERSED-PHASE COLUMN CHROMATOGRAPHY, APPLIED TO THE SEQUENCE DETERMINATION OF CYTOCHROME *c* FROM THE WILD TYPE AND A MUTANT OF THE BUTTERFLY, *PIERIS BRASSICAE*

WILLIAM H. VENSEL*,*
Department of Biochemistry, Molecular Biology and Cell Biology, Northwestern University, Evanston, IL 60201 (U.S.A.)
VALERIE S. FUJITA
Department of Biochemistry, Molecular Biology and Cell Biology, Northwestern University, Evanston, IL 60201, and Department of Biological Chemistry, University of Michigan, Ann Arbor, MI (U.S.A.)
GEORGE E. TARR
Department of Biological Chemistry, University of Michigan, Ann Arbor, MI (U.S.A.)
E. MARGOLIASH
Department of Biochemistry, Molecular Biology and Cell Biology, Northwestern University, Evanston, IL 60201 (U.S.A.)
and
HARTMUT KAYSER
Department of Biology I, University of Ulm, Ulm (G.F.R.)

SUMMARY

Two-dimensional peptide mapping has been very effective in the characterization of protein digests, particularly for the detection of small structural differences between homologous proteins. The classical thin-layer strategy, which exploits differences in charge and hydrophobicity, has been realized as a method based on reversed-phase high-performance liquid chromatography. An initial fractionation at pH 7.2 with 100 m*M* potassium phosphate, followed by chromatography with 0.1% trifluoroacetic acid, has been applied to chymotryptic digests of cytochromes *c*. The use of UV-transparent and (in the final stage) volatile solvents allows detection and rapid recovery of nanomole amounts of peptides suitable for sequence determination. As an example of the application of this method we report the comparison of two variants of cytochrome *c* from the butterfly, *Pieris brassicae*, one being the wild type and the other a spontaneous mutant isolated from a laboratory colony. The single residue difference was easily detected and identified.

* Present address: Food Proteins Research Unit, U.S. Department of Agriculture, Agricultural Research Service, Western Regional Research Center, Berkeley, CA 94710, U.S.A.

0021-9673/83/$03.00 © 1983 Elsevier Science Publishers B.V.

INTRODUCTION

A protein amino acid sequence represents a wealth of information and is especially valuable as a direct translation of the genome. One strategy for sequence determination is to fragment the protein enzymatically or chemically into small peptides which may be fractionated according to hydrophobicity in combination with charge differences. The ability to separate and recover peptides in nanomole amounts appropriate for sequencing makes reversed-phase liquid chromatography a particularly convenient means of preparing peptides for studies of protein structure. Numerous procedures for the reversed-phase separation of peptides have been published (for a review, see ref. 1). The systematic high-performance liquid chromatographic (HPLC) separation of peptides at two carefully chosen conditions of pH and ionic strength allows construction of two-dimensional peptide maps analogous to those obtained by two-dimensional thin-layer or paper chromatography. We have developed such a method for cytochromes *c*. The method employs a phosphate buffer at neutrality[2] and 0.1 % trifluoroacetic acid (TFA) at pH 2.0 (ref. 3).

In the course of studies of the biosynthesis of cytochrome *c* from *Pieris brassicae*, the white cabbage butterfly, it was found that the butterflies in a laboratory colony contained two varieties of cytochrome *c* which could be separated by cation-exchange chromatography. The cytochrome *c* which represented *ca.* 85 % of the total was chromatographically identical with the homogeneous cytochrome *c* found in a strain of *Pieris* from a different laboratory and was therefore considered to be the

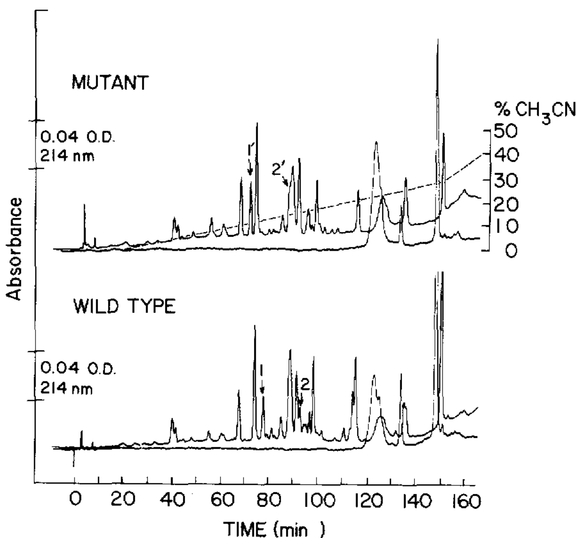

Fig. 1. Analytical map of 2 n*M* each of chymotryptic digests of *Pieris brassicae* mutant and wild-type cytochromes *c* chromatographed at pH 7.2 in a linear gradient between 100 m*M* K_2HPO_4 in water *versus* 45 % acetonitrile in 100 m*M* K_2HPO_4 in water. The upper tracing in each panel is the optical density at 214 nm (0.4 a.u.f.s.); the lower tracing in each panel is the optical density at 410 nm (0.2 a.u.f.s.). The gradient conditions are indicated for the mutant and are the same for both mutant and wild type digests. Flow-rate was 1 ml/min for the first 8 min and was reduced in 2 min to 0.5 ml/min. The acetonitrile was started at 10 min and is plotted with correction for holdup volume of the chromatographic system (broken line). Peak numbers refer to text.

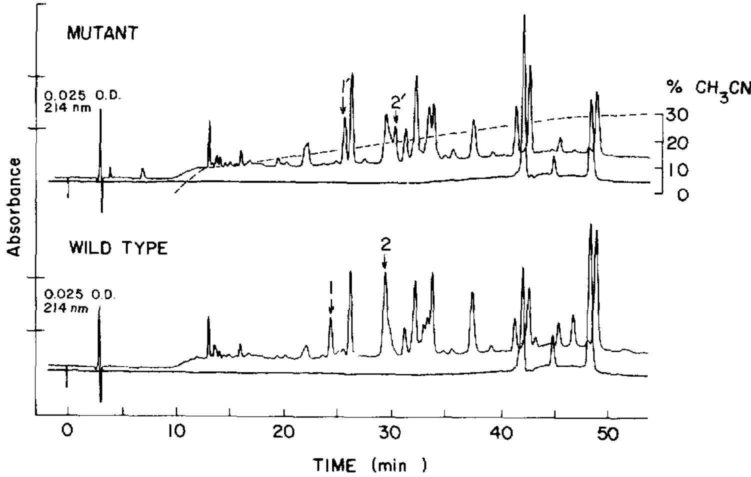

Fig. 2. Analytical map of 1 n*M* each of mutant and wild-type cytochrome *c* chymotryptic digests chromatographed at pH 2.0 in 0.1 % TFA in water *versus* 0.1 % TFA in acetonitrile. The upper tracing in each panel represents the optical density at 214 nm (0.25 a.u.f.s.) and the lower tracing in each panel represents the optical density at 410 nm (0.5 a.u.f.s.). Flow-rate 1 ml/min. The gradient as plotted represents the acetonitrile concentration at the detector and was identical for both chromatograms. Peak numbers refer to text.

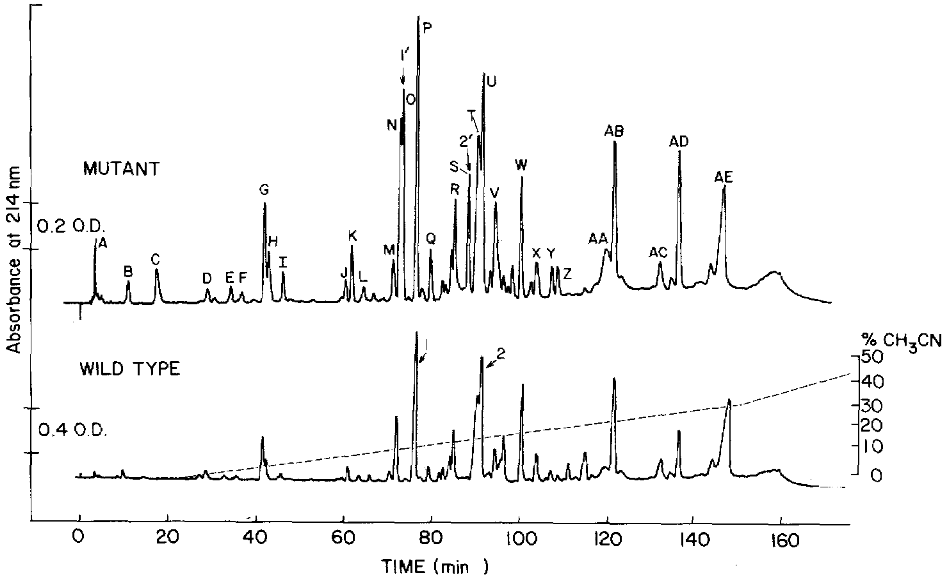

Fig. 3. Comparison of pH 7.2 preparative separation of *Pieris brassicae* mutant and wild-type cytochrome *c* chymotryptic digests. The flow and gradient conditions for the mutant and wild type separations where the same and are as indicated in Fig. 1. The optical density setting was 2.0 a.u.f.s. Recorder attenuation was 20 mV full scale for the wild type digest and 10 mV full scale for the mutant digest. Letters above mutant peptides correspond to labels in Table I.

wild-type protein. Colonies of butterflies were established that were homozygous for the wild-type protein and for the mutant protein. Using two-dimensional peptide mapping by reversed-phase HPLC and amino acid and sequence analysis, we have demonstrated that the mutant cytochrome *c* differs from the wild type at residue 39, the lysine present in the wild type being replaced by a glutamic acid. The complete sequences will be reported elsewhere[4].

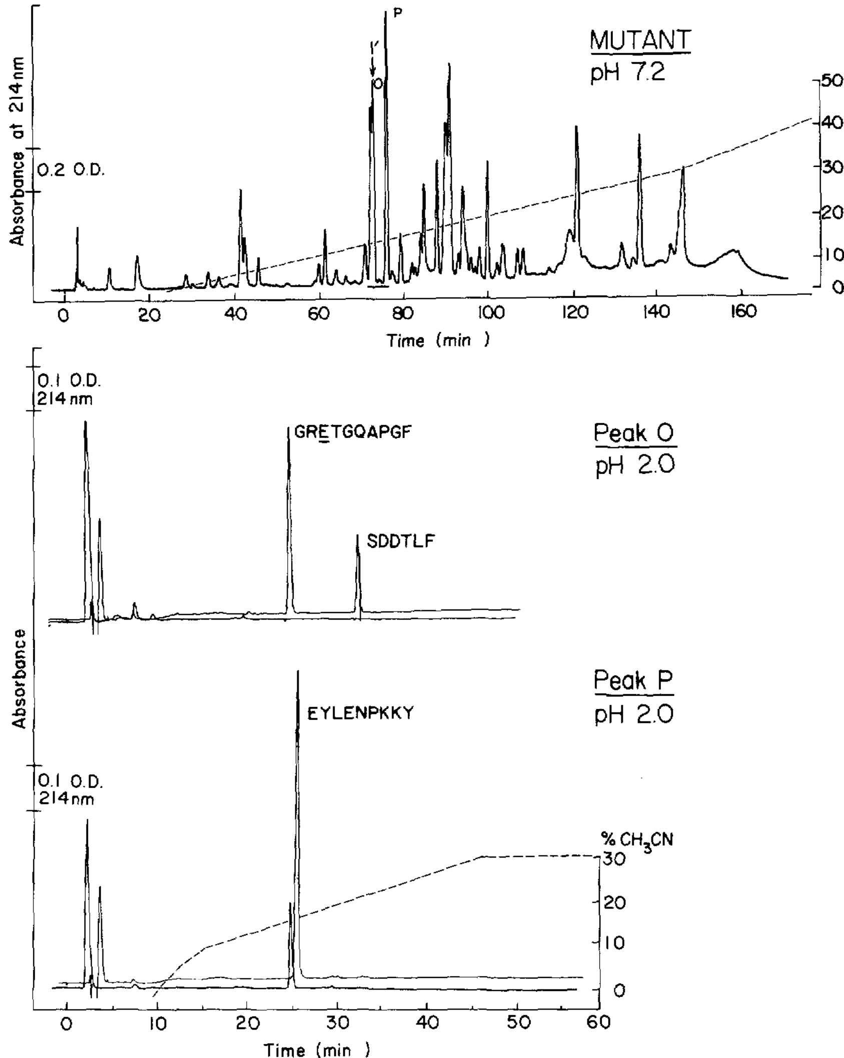

Fig. 4. Preparative separation at pH 7.2 and pH 2.0 preparative rechromatography of peaks O and P for the mutant type *Pieris brassicae* cytochrome *c* chymotryptic digest. Peak O includes peptide 1. The conditions for the pH 7.2 chromatogram are given in Fig. 1 and the flow-rate and gradient conditions for the pH 2.0 rechromatography are given in Fig. 2. Gradient conditions are as shown for rechromatography of P and are identical for chromatography of O and P. The upper tracing in rechromatography of O and P is the optical density at 214 nm (1.0 a.u.f.s.) and the lower tracing in chromatograms of O and P, the optical density at 290 nm (0.1 a.u.f.s.).

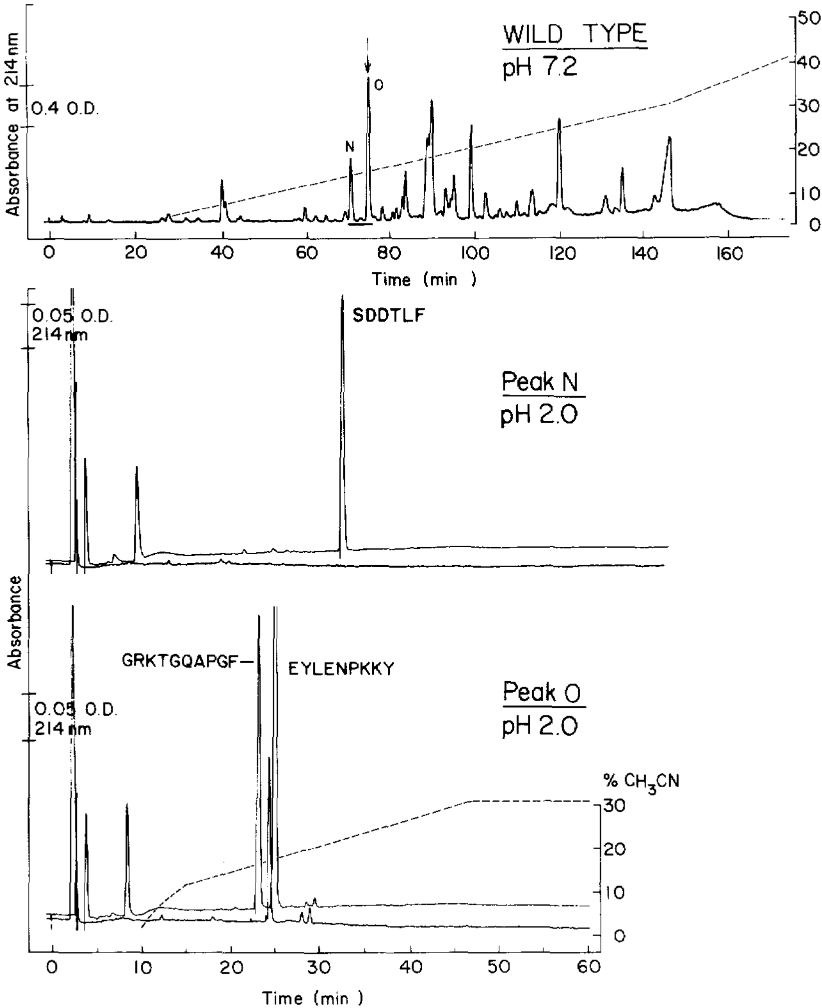

Fig. 5. Preparative separation at pH 7.2 and pH 2.0 preparative rechromatography of peaks N and O for the wild-type *Pieris brassicae* cytochrome *c* chymotryptic digest. Peak O includes peptide 1. Conditions are as described in Fig. 4, except that the upper 214 nm tracing in rechromatography of N and O is 0.5 a.u.f.s. and the lower 290 nm tracing in rechromatography of N and O is 0.5 a.u.f.s.

METHODS

Cytochrome *c* was prepared by a modification of the procedure of Margoliash and Walasek as modified by Brautigan *et al.*[5]. Chymotryptic digestion of the cytochromes *c* was carried out for 7 h at 27°C in 10 m*M* K_2HPO_4 at pH 7.6 with an enzyme/substrate ratio of 1:100 (w/w).

The separations on HPLC were performed with Waters μBondapak C_{18} columns (30 cm × 3.9 mm I.D.). A Beckman Model 342 gradient liquid Chromatograph was used to form gradients, and a Beckman Model 165 variable-wavelength UV detector to detect and scan peaks. Channel 1 of the detector was used routinely at 214 nm, channel 2 was used at 410 nm for the initial separation and at 290 nm for the

TABLE I

RECOVERIES AND ELUTING CONCENTRATIONS OF ACETONITRILE FOR SELECTED *PIERIS BRASSICAE* MUTANT AND WILD-TYPE CYTOCHROME *c* PEPTIDES AT pH 7.2 AND AT pH 2.0

Peptide sequences are given in single-letter IUB code. The underlined glutamic acid residues are in peptides from the mutant containing the single amino acid substitution. Peptides 1 and 2 correspond to peptides 1 and 2 of Figs. 1–5. The pH 2.0 elution concentration of acetonitrile was not determined for peptide AE as no peak eluted. N.D. = Not determined.

Peptide from Fig. 3	*Peptide amino acid sequence*	*Concentration of acetonitrile (%)*		*Recovery (%)*
		pH 7.2	*pH 2.0*	
B	SY	0.0	9.0	12
G	ENPKKY	4.4	14.0	4
I	KKANERDAL	5.3	12.9	9
K	IAY	8.6	14.3	26
M	VF	10.6	14.4	22
N	SDDTLF	10.8	21.7	85
O (1′)	GRETQAPGF	11.1	16.3	52
(1)	GRKTQAPGF	11.8	15.8	56
P	EYLENPKKY	11.8	16.8	69
R	KKANERADLIAY	12.4	20.1	19
S (2′)	GRETQAPGFSY	13.5	19.7	17
T	GVPAGNTENGKKIF	14.3	19.7	18
(2)	GRKTQAPGFSY	14.7	19.3	16
V	SEANKAKGITW	14.7	21.3	31
W	SYSEANKAKGITW	15.8	22.5	13
AA	Heme peptide	20.9	28.2	2
AB	IPGTKMVF	21.4	24.6	37
AD	KVGPNLHG	24	27	22
AE	Heme-containing peptide	29	N.D.	0

reexamination of individual peaks, and channel 3 was used in the scan mode.

Water for HPLC solvents was obtained from a Barnstead Nanopure water purification unit with the ion-exchange cartridge placed after the charcoal cartridge. Phosphate buffers contained NaF (10 mg/l) to retard bacterial growth. In order to remove UV-absorbing material for high-sensitivity work, aqueous components of the phosphate buffers were pumped (before the addition of acetonitrile) through 25 cm × 4.6 mm I.D. columns, containing Waters C_{18} bulk packing. TFA buffers were prepared from 1-ml sealed glass ampoules of TFA (Pierce). In order to reduce the dissolution of silica in the analytical column at neutrality, the effluent from each HPLC pump was passed, before entering the mixing chamber of the chromatograph, though a 25 cm × 4.6 mm I.D. column of Waters C_{18} bulk packing. The initial preparative separation was carried out with a gradient from 0 to 45% acetonitrile, which contained 100 m*M* K_2HPO_4, pH 7.2 throughout. The individual peak fractions were collected, reduced in volume *in vacuo* to remove acetonitrile, and adjusted to *ca.* pH 2 with an amount of phosphoric acid calculated on the basis of peak volume. Fractions were then rechromatographed (and desalted) in a gradient from 0

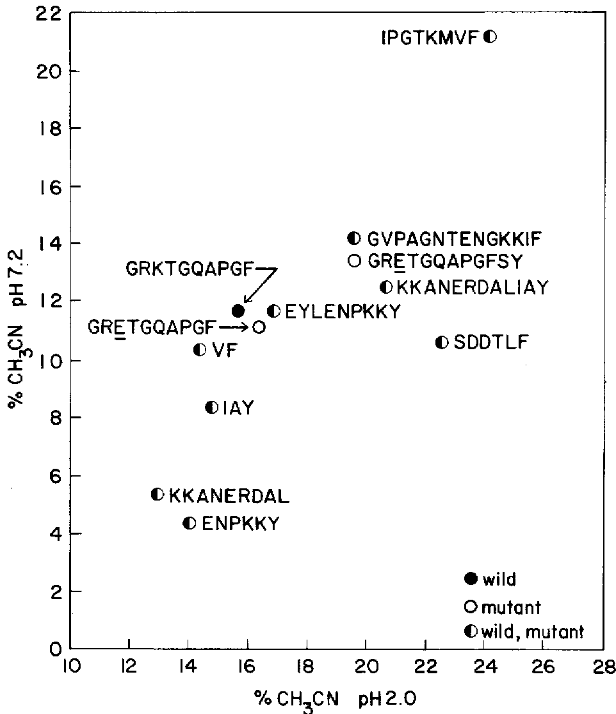

Fig. 6. Plot of eluting concentrations of acetonitrile at pH 7.2 and at pH 2.0 for selected peptides from *Pieris brassicae* mutant and wild-type cytochrome *c* chymotryptic digests.

to 30% acetonitrile which contained 0.1% TFA throughout. Fractions were dried *in vacuo*. Samples for hydrolysis were dried in 50 × 6 mm Pyrex tubes and placed in 40-ml screw-cap vials (up to 12 per vial) with 0.5 ml of 6 *M* hydrochloric acid containing 0.2% phenol. The vial was sealed with and evacuated through a Mininert valve (Pierce). After 24 h at 110°C, the vial was cooled, then opened, and the tubes were removed and vacuum dried[6]. Amino acid analysis was performed with a Durrum D-500 amino acid analyzer. Edman degradation of peptides was carried out by the manual batchwise method of Tarr[7]. Phenylthiohydantoin (PTH)-amino acids were analyzed isocratically[8] with the aid of an automatic sampler[9].

RESULTS

The analytical maps of the *Pieris* mutant and wild-type cytochrome *c* chymotryptic digests at neutral pH are shown in Fig. 1. For purposes of comparison the chromatograms may be divided into two parts. Peptides eluted after 110 min include three containing heme, as indicated by the 410 nm profile, and other large partially digested fragments. Those eluted earlier have the characteristics of a classical "limit" digest for both proteins and may be readily compared. The most striking differences are indicated by the arrows. The peptide from the mutant protein labelled 1′ is eluted more quickly than the corresponding wild-type peptide. A similar relationship holds for the peptides labelled 2.

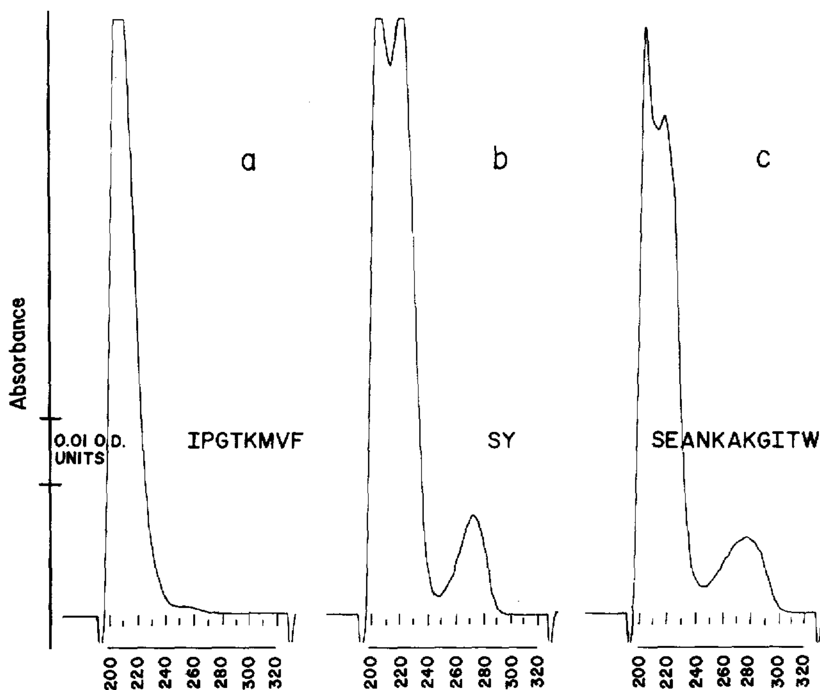

Fig. 7. Spectral scan of three different peptides obtained during rechromatography. The vertical axis represents absorbance. The horizontal axis represents the wavelength in nanometres. Peak a contained 7.3 n*M* of peptide, peaks b and c contained 2.5 n*M* and 6.1 n*M* of peptide, respectively.

The analytical maps in acid for both cytochrome digests are shown in Fig. 2. Again there are three late-eluted heme-containing peptides. Their chromatographic behavior as indicated by sharpness and peak symmetry is better than at pH 7.2. As expected, there are among the early peaks pairs of peptides that vary systematically between mutant and wild type, the wild-type peptides 1 and 2 in this case being eluted earlier.

The preparative separation at neutral pH of 20 nmol each of the mutant and wild-type chymotryptic digests are shown in Fig. 3, the main difference being indicated by arrows as in Fig. 1. Chromatographic conditions for the preparative separation are described in Fig. 1. Fractions ranging in volume from 500 to 1500 μl were collected from the preparative separation of the *Pieris* cytochrome *c* digests. A total of 33 peptides from the mutant and 14 from the wild type were rechromatographed at pH 2.0. As examples of the rechromatography, the purification and mapping of peptide 1 and adjacent peptides from each variant is shown in Figs. 4 and 5. Identifications were made by composition and sequence.

The elution positions for some of the mutant and wild-type peptides are given in Table I and plotted in two dimensions in Fig. 6. Rechromatographing peaks at pH 2.0 resulted in resolution of some peaks which were not separated at pH 7.2. For example, the peptides VF and SDDTLF were barely resolved at neutrality but separated widely at pH 2.0 owing to the marked increase in hydrophobicity in acid of the acidic peptide SDDTLF relative to the peptide VF. In the wild-type digest (Fig. 5) the

peptides GRKTGQAPGF and EYLENPKKY were inseparable at neutral pH in the preparative chromatogram but were resolved at acid pH. Again, the suppression of charge on carboxyl groups in acid causes the more acidic peptide, EYLENPKKY, to elute later. Conversely, the differential increase in charge at neutrality permits resolution of peptides co-eluting in acid, such as mutant peptides GRETGAPGFSY and GVPAGNTENGKKIF.

As shown in Table I, recoveries varied from as low as 0% for a heme-containing peptide to as high as 85% for the peptide SDDTLF. The former may be attributable to poor chromatographic behavior or chemical degradation, but in most cases lower recoveries are due to the occurrence of the sequence in more than one peptide because of incomplete digestion and/or "minor splits". One example of this is the variant region, which shows up in at least one other peptide (peptides 2) beyond the limit digest product (peptides 1). If chymotrypsin were a "perfect" enzyme and only cleaved cytochrome *c* at major cleavage sites, and cleaved those to completion, then the digest would contain only one heme-containing peptide and seventeen other peptides.

During rechromatography of each peak, a spectral scan from 200 to 325 nm (in addition to the optical density of 214 and 410 nm), was automatically obtained when each peak passed through the detector flow cell. Representative spectral scans for three different peptides are shown in Fig. 7. Amino acid analysis and Edman degradation confirmed the inference of phenylalanine, tyrosine and tryptophan from spectra a, b, and c, respectively. The spectrum of each aromatic amino-acid-containing peptide was similar in the range 240–290 nm to spectral data for the corresponding aromatic amino acids[10].

The ability to perform a wavelength scan or to monitor more than one wavelength while separating peptides is an aid in sequence determination. In particular, the detection of the presence of tryptophan which is destroyed by hydrolysis with hydrochloric acid is insured by spectral scans. In this study the presence of tryptophan-containing peptides was also corroborated by monitoring the optical density at 290 nm. The technique is not limited to peptides containing aromatic amino acids but can be used to locate any peptides containing a chromophore.

DISCUSSION

The strategy of enzymatic or chemical fragmentation of proteins followed by reversed-phase HPLC of the fragments is a rapid means of producing peptide maps and pure peptides for the sequence determination of proteins. Using a separation at neutral pH and one at acid pH we have exploited the resultant differential change in peptide hydrophobicity to map and isolate peptides. Significantly this method also allows the experimenter to distinguish between peptides containing acids and amides. The method, while developed for cytochrome *c*, is applicable to other proteins as well.

If this method is to be useful over a long period as a mapping technique, then it is essential that digestion and chromatographic conditions be highly reproducible. We feel these ideals are attainable. In Figs. 1 and 3 the normal and mutant digests clearly differ in ways not accounted for by the single amino acid replacement which we know to be the only difference between the two proteins. In other experiments we have produced digests extremely similar in detail from the cytochrome *c* of several

insect species, so we believe that the unexpected differences seen here are anomalous and attributable to differences in purity and/or denaturational state of these two preparations of protein. The marked difference in amount of "breakthrough" observable in Fig. 3 may be considered support for the former idea. This problem may be eliminated by better standardization of preparative conditions, more stringest criteria of purity, or digestion conditions relatively insensitive to impurities and protein conformation.

The low recovery of the heme-containing peptides is probably not due to their poor chromatographic behavior *per se* but probably due to destruction and modification of the heme during volume reduction. Desalting and rechromatography at pH 2.0 of a single heme-containing peptide (peptide AA, Fig. 3) resulted in recovery of a very small heme-containing peak and in a broad shallow rise in baseline at both 214 and 410 nm. A solution to this problem is to de-heme and alkylate the cytochrome *c* before enzymatic digestion.

The recoveries obtained for some of the individual peptides required that sequencing be carried out at the picomole level. This was readily accomplished using the batchwise methodology of Tarr[7]. The large number of peptides which can be rapidly generated by this procedure make it most useful in conjunction with manual batchwise protein sequencing procedures. The HPLC procedure as described can be speeded up considerably. Gradients for rechromatography, generated at 6–8 %/min produce effective separations and allow two to three repetitions per hour. Thus it is possible to generate fifteen to twenty pure chymotryptic peptides of cytochrome *c* in two days. The potential for the combined methodology of HPLC and manual batchwise sequencing is that less than complete sequencing of a homologous set of proteins combined with HPLC mapping and amino acid analysis might in most cases provide adequate proof of structure in less than a week.

ACKNOWLEDGEMENTS

This work was supported by Grants GM-19121 and HL-11119 from the National Institutes of Health.

REFERENCES

1 M. T. W. Hearn, *J. Liquid Chromatogr.*, 3 (1980) 1255.
2 M. T. W. Hearn, B. Grego and W. S. Hancock, *J. Chromatogr.*, 185 (1979) 429.
3 H. P. J. Bennett, C. A. Browne and S. Solomon, *J. Liquid Chromatogr.*, 3 (1980) 1353.
4 V. S. Fujita, W. H. Vensel, G. E. Tarr, E. Margoliash and H. Kayser, in preparation.
5 D. L. Brautigan, S. Ferguson-Miller and E. Margoliash, *Methods Enzymol.*, 53 (1978) 128.
6 G. E. Tarr, unpublished results.
7 G. E. Tarr, in M. Elzinga (Editor), *Methods in Protein Sequence Analysis*, Humana Press, Clifton, NJ, 1981, p. 223.
8 G. E. Tarr, *Anal. Biochem.*, 111 (1981) 27.
9 W. H. Vensel and G. E. Tarr, in preparation.
10 G. D. Fasman (Editor), *Handbook of Biochemistry and Molecular Biology*, Vol. I, CRC Press, Cleveland, OH, 3rd ed., 1976, p. 187.