Biochem. J. (1986) 236, 871–878 (Printed in Great Britain)

Complete structure of the hydrophilic domain in the porcine NADPH–cytochrome *P*-450 reductase

Frank VOGEL and Ludwig LUMPER
Biochemisches Institut der Justus Liebig-Universität, Fachbereich Humanmedizin, 6300 Gießen, Federal Republic of Germany

The 622-residue amino acid sequence of the hydrophilic domain in the porcine NADPH–cytochrome *P*-450 reductase (EC 1.6.2.4) is reported. The structural data required to complete the sequences published previously [Vogel, Kaiser, Witt & Lumper (1985) Biol. Chem. Hoppe-Seyler **366**, 577–587] and to establish the primary structure of the porcine hydrophilic domain have been obtained by sequencing proteolytic subfragments derived from CNBr fragments and by characterizing the overlapping *S*-[^{14}C]methylmethionine-containing peptides isolated from tryptic digests of the [^{14}C]methyl-labelled hydrophilic domain. The hydrophilic domain displays 91.8% positional identity with that of the corresponding domain in the rat NADPH–cytochrome *P*-450 reductase. The region Val^{528}–Ser^{678} in the NADPH–cytochrome *P*-450 reductase shows a significant homology to the sequence Ile^{165}–Tyr^{314} in the spinach ferredoxin–NADP$^+$ oxidoreductase. A model for the secondary structure of the hydrophilic domain has been derived by computer-assisted analysis of the amino acid sequence. Cys^{472} and Cys^{566} are protected against chemical modification in the NADP$^+$ complex of the NADPH–cytochrome *P*-450 reductase.

INTRODUCTION

The membrane protein NADPH–cytochrome *P*-450 reductase is an essential component of steroid-, haem- and xenobiotic-metabolizing enzyme systems in eukaryotes (Masters & Okita, 1980; Yoshinaga *et al.*, 1982; Gilette *et al.*, 1972). NADPH–cytochrome *P*-450 reductase contains an *N*-terminal hydrophobic membrane segment (hydrophobic domain) and a hydrophilic domain (identical with the trypsin-solubilized NADPH–cytochrome *P*-450 reductase) containing the catalytically active site (Black *et al.*, 1979). Beginning with the studies on the primary structure of the hydrophobic domain (Black & Coon, 1982), several groups reported on the primary structure of the NADPH–cytochrome *P*-450 reductase. The analysis of cloned genomic DNA led to the elucidation of the complete chain sequence of the NADPH–cytochrome *P*-450 reductase from rat liver (Porter & Kasper, 1985). Using protein sequencing techniques, partial structures of the hydrophilic domain in the porcine NADPH–cytochrome *P*-450 reductase have been determined (Vogel & Lumper, 1984; Haniu *et al.*, 1984, 1985; Vogel *et al.*, 1985).

The present paper describes the complete structure of the trypsin-solubilized NADPH–cytochrome *P*-450 reductase from pig liver. Structural features of the protein and the results of secondary structure assignments by predictive methods are discussed.

MATERIALS AND METHODS

Materials

BC-18 [Bakerbond Wide Pore Octadecyl (C_{18}) column material (5 μm, pore width 33 nm)] was purchased from J. T. Baker Chemikalien (Gross-Gerau, Germany) and SC-18 [Shandon ODS-Hypersil (3 μm)] from Shandon Labortechnik (Frankfurt, Germany). ^{14}C-labelled methyl iodide (10 mCi/mmol) was obtained from Du Pont de Nemours (Dreieich, Germany).

Isolation of the trypsin-solubilized NADPH–cytochrome *P*-450 reductase from pig liver

The reductase was isolated from pig liver microsomes as described in Vogel & Lumper (1983) and Vogel *et al.* (1985).

Purification and sequence analysis of peptide fragments from the trypsin-solubilized NADPH–cytochrome *P*-450 reductase from pig liver

Digestion of the enzyme by CNBr and subsequent degradation by proteinases were performed by following the procedures described previously (Vogel & Lumper, 1983, 1985; Vogel *et al.*, 1985). CNBr fragment CB16, derived from the *S*-carboxymethylated hydrophilic domain, was eluted shortly after CB5 during h.p.l.c. of gel permeation chromatography pool IV peptides on a BC-18 column (0.46 cm × 25 cm). Elution was achieved at ambient temperature by a linear gradient over 60 min from acetonitrile/water (1:4, v/v) containing 0.1% trifluoroacetic acid to acetonitrile/water (1:1, v/v) containing 0.1% trifluoroacetic acid (flow rate 0.7 ml/min). To avoid a considerable loss of material, the h.p.l.c. pool (430 μg) containing CB5 and CB16 was split by trypsin under the conditions described previously. The tryptic digest was dried by using a Speed-Vac concentrator, dissolved in 5 μl of 70% formic acid and subsequently diluted with 95 μl of 0.1% trifluoroacetic acid. The peptides were purified by h.p.l.c. on a BC-18 column (0.46 cm × 26 cm) using a 60 min linear gradient from acetonitrile/water (1:99, v/v) containing 0.1% trifluoroacetic acid to acetonitrile/water (1:1, v/v) containing 0.1% trifluoroacetic acid (flow rate 0.7 ml/min; 25 °C).

Abbreviations used: monobromobimane, 4-(bromomethyl)-3,6,7-trimethyl-1,5-diazabicyclo[3.3.0]octa-3,6-diene-2,8-dione; bimanyl group 3,6,7-trimethyl-2,8-dioxo-1,5-diazabicyclo[3.3.0]octa-3,6-dien-4-ylmethyl; NADPH–cytochrome *P*-450 reductase, NADPH–ferrihaemoprotein oxidoreductase (EC 1.6.2.4) (the hydrophilic or catalytic domain of this enzyme is synonymous with trypsin-solubilized NADPH–cytochrome *P*-450 reductase).

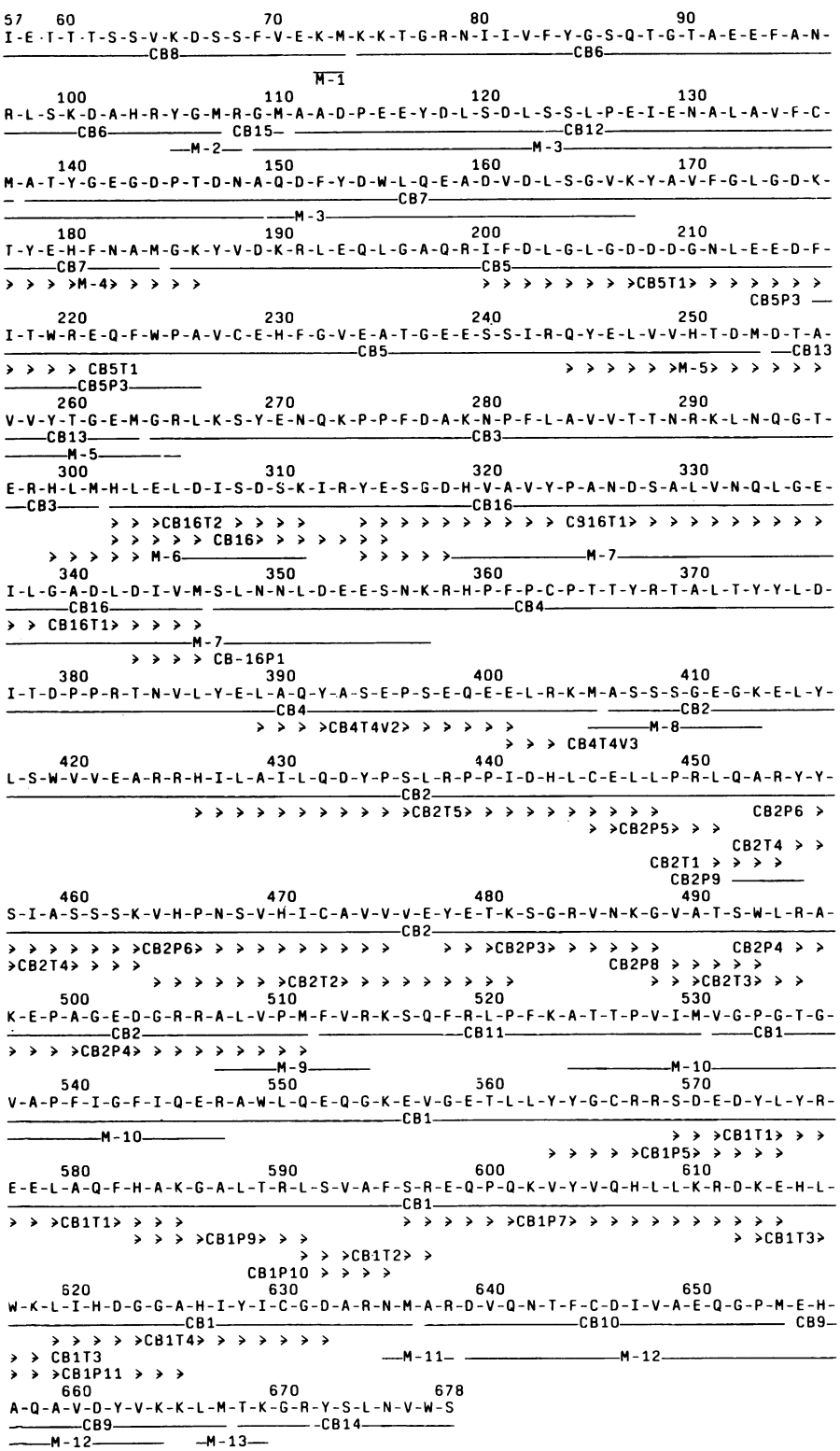

Fig. 1. Summary of proof of the amino acid sequence of the hydrophilic domain of the porcine NADPH–cytochrome *P*-450 reductase Proteolytic subfragments are not shown when described previously (Vogel *et al.*, 1985). Sequence data on individual peptides are indicated as follows: ›, step sequenced by automated Edman degradation, CB. . ., CNBr peptide; CB. . T. . ., tryptic subfragment of CNBr peptide; CB. . P. . ., peptic subfragment of CNBr peptide; CB. . T.V., peptide derived from CB. . T. with V8 proteinase (endoproteinase Glu-C, EC 3.4.21.19). The M_r of the hydrophilic domain in the porcine NADPH–cytochrome *P*-450 reductase calculated on the basis of the sequence data shown is 71098.3.

Table 1. Amino acid compositions of proteolytic peptides obtained from CNBr fragments

Peptides are characterized as follows: (1) by the elution volume (GPC:EV) from a Sephadex G-75 column (1.7 cm × 130 cm); (2) and (3) by the conditions used for isolation on a BC-18 column (0.46 cm × 25 cm) (HPLC BC-18) or an SC-18 column (0.46 cm × 25 cm) (HPLC SC-18). The gradient is described by the initial concentration (%) of acetonitrile in 0.1% trifluoroacetic acid/time of gradient (min)/the final concentration (%) of acetonitrile in 0.1% trifluoroacetic acid, and retention times (RT) are given in min. H.p.l.c. was performed at ambient temperature or at 60 °C (#). Amino acids were determined as *o*-phthaldialdehyde derivatives, cysteine as *S*-carboxymethyl- or as *S*-bimanyl-cysteine (Vogel & Lumper, 1983) (*), methionine as the sum of homoserine and homoserine lactone, tryptophan by peptide fluorescence (excitation 295 nm, emission 354 nm) and proline qualitatively by dansylation of the peptide hydrolysate and t.l.c.

Fragment	CB1T1	CB1T2	CB1T3	CB1T4	CB1P5	CB1P7	CB1P9	CB1P10	CB1P11	CB2T1	CB2T2	CB2T3	CB2T4	CB2T5
1. GPC: EV (ml)	234-284	257-305	257-284	234-257	-	180-245	209-264	264-305	209-264	270-305	230-270	270-305	245-305	202-224
2. HPLC BC-18	1/60/50	1/60/50	1/60/50	1/60/50	1/60/50	1/60/50	1/60/50	1/60/50	1/60/50	1/60/50	1/60/50	1/60/50	1/60/50	1/60/50
RT (min)	41.93	35.80	33.52	40.96	38.3-39.3	39-39.3	32.2-32.8	32.50	32.8-33.3	22.19	44.75	39.20	28.97	54.20
3. HPLC SC-18	15/60/50#	-	-	-	15/60/50	-	-	15/60/50#	20/60/50#	15/60/50#	-	20/60/50#	-	-
RT (min)	31.12	-	-	-	31.20	-	-	16.67	27-29.2	6.31	-	20.72	-	-
N-Terminus	Ser	Leu	Asp	Leu	Tyr	Ser	(His)	Ser	(Trp)	Leu	Val	Gly	Tyr	(His)
Sequenz	569-585	591-597	612-618	619-634	563-574	596-625	583-591	592-595	617-625	451-454	464-481	488-495	455-463	426-450
Cys	-	-	-	0.60 (1)*	0.96 (1)	-	-	-	-	-	1.27 (1)	-	-	1.07 (1)*
Asx	2.00 (2)	-	1.10 (1)	2.24 (2)	1.90 (2)	2.07 (2)	-	-	1.16 (1)	-	1.00 (1)	-	-	2.07 (2)
Thr	-	-	-	-	-	-	0.74 (1)	-	-	-	0.86 (1)	0.72 (1)	-	-
Ser	0.98 (1)	1.72 (2)	-	-	0.79 (1)	1.36 (1)	-	0.85 (1)	-	-	1.00 (1)	0.97 (1)	3.49 (4)	1.49 (1)
Glx	4.31 (4)	-	1.32 (1)	-	1.10 (1)	6.52 (5)	-	-	-	1.31 (1)	2.11 (2)	-	-	2.64 (2)
Gly	-	-	-	2.08 (3)	1.00 (1)	1.28 (2)	1.05 (1)	-	1.67 (2)	-	-	1.01 (1)	-	-
Ala	1.49 (2)	1.00 (1)	-	1.95 (2)	-	0.80 (1)	2.05 (2)	1.00 (1)	1.02 (1)	1.26 (1)	1.00 (1)	1.26 (1)	1.29 (1)	1.24 (1)
Val	-	1.10 (1)	-	-	-	1.98 (2)	-	1.12 (1)	-	-	2.37 (5)	0.87 (1)	-	-
Met	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Ile	-	-	-	1.61 (3)	-	0.42 (1)	-	-	1.34 (1)	-	0.80 (1)	-	1.03 (1)	1.85 (3)
Leu	2.03 (2)	1.10 (1)	1.32 (1)	1.59 (1)	0.98 (1)	3.89 (4)	1.80 (2)	-	1.10 (1)	1.00 (1)	-	1.09 (1)	-	4.51 (6)
Tyr	1.71 (2)	-	-	1.19 (1)	2.67 (3)	1.08 (1)	-	-	-	-	0.86 (1)	-	2.01 (2)	1.50 (1)
Phe	0.86 (1)	0.95 (1)	-	-	-	-	-	1.11 (1)	-	-	-	-	-	-
His	0.84 (1)	-	0.92 (1)	1.22 (2)	-	1.72 (3)	0.67 (1)	-	0.92 (1)	-	1.17 (2)	-	-	1.01 (2)
Lys	1.04 (1)	-	2.00 (2)	-	-	2.70 (4)	0.89 (1)	-	0.82 (1)	-	0.86 (1)	-	0.83 (1)	-
Arg	0.99 (1)	1.01 (1)	-	1.26 (1)	1.90 (2)	2.37 (2)	1.40 (1)	-	-	0.96 (1)	-	1.00 (1)	-	1.99 (2)
Pro	-	-	-	-	-	(1)	-	-	-	-	(1)	-	-	(4)
Trp	-	-	0.66 (1)	-	-	1.05 (1)	-	-	0.98 (1)	-	-	1.40 (1)	-	-
Total residues	17	7	7	16	12	30	9	4	9	4	18	8	9	25
Yield (%)	29	33	47	51	24	13	66	41	7	41	21	2	33	42

Fragment	CB2P3	CB2P4	CB2P5	CB2P6	CB2P8	CB2P9	CB4T4V2	CB4T4V3	CB5T1	CB5P3	CB16T1	CB16T2
1. GPC: EV (ml)	219-253	219-253	-	-	253-305	253-305	-	-	-	-	-	-
2. HPLC BC-18	1/60/50	1/60/50	1/60/50	1/60/50	1/60/50	1/60/50	1/60/50#	1/60/50#	1/60/50	1/60/50	1/60/50	1/60/50
RT (min)	20.49	30.59	44.13	34.39	44.98	23.55	28.39+33.42	24.19	53.79	42.70	58.15	37.36
3. HPLC SC-18	-	-	-	-	-	15/60/50#	-	-	-	-	-	-
RT (min)	-	-	-	-	-	7.11	-	-	-	-	-	-
N-Terminus	Tyr	(Arg)	(Cys)	Tyr	Val	Glu	Leu	Glu	Ile	Phe	Tyr	(His)
Sequenz	478-488	495-511	445-451	456-475	489-493	452-455	389-401	401-403	200-220	216-226	314-346	302-311
Cys	-	-	0.96 (1)*	0.95 (1)	-	-	-	-	-	-	-	-
Asx	0.99 (1)	1.04 (1)	-	2.16 (1)	-	-	-	-	6.04 (6)	-	5.40 (6)	1.71 (2)
Thr	0.82 (1)	-	-	-	0.70 (1)	-	-	-	0.94 (1)	0.77 (1)	-	-
Ser	0.87 (1)	-	-	4.56 (5)	0.78 (1)	-	1.89 (2)	-	-	-	1.79 (2)	1.80 (2)
Glx	1.19 (1)	1.88 (2)	1.11 (1)	-	-	1.03 (1)	6.10 (6)	1.33 (1)	2.60 (2)	2.78 (2)	3.52 (3)	1.46 (1)
Gly	2.04 (2)	1.57 (2)	-	-	-	-	-	-	2.58 (3)	-	2.90 (3)	-
Ala	-	2.95 (3)	-	2.26 (2)	1.00 (1)	1.06 (1)	2.08 (2)	-	-	0.74 (1)	2.39 (4)	-
Val	0.89 (1)	1.05 (1)	-	2.00 (4)	0.72 (1)	-	-	-	-	-	2.92 (4)	-
Met	-	0.62 (1)	-	-	-	-	-	-	-	-	0.82 (1)	-
Ile	-	-	-	2.01 (2)	-	-	-	-	1.48 (2)	0.86 (1)	1.40 (2)	0.82 (1)
Leu	-	1.61 (1)	3.01 (3)	0.65 (-)	-	-	0.93 (1)	1.00 (1)	2.92 (3)	-	3.02 (4)	2.48 (2)
Tyr	0.93 (1)	-	-	1.64 (1)	-	1.01 (1)	0.99 (1)	-	-	-	1.75 (2)	-
Phe	-	-	-	-	-	-	-	-	1.67 (2)	1.87 (2)	-	-
His	-	-	-	2.00 (2)	-	-	-	-	-	-	0.98 (1)	1.17 (1)
Lys	1.47 (2)	0.91 (1)	-	1.50 (1)	-	-	-	-	-	-	-	1.23 (1)
Arg	0.92 (1)	3.85 (3)	0.89 (1)	-	-	0.88 (1)	-	0.97 (1)	0.98 (1)	1.04 (1)	-	-
Pro	-	(2)	(1)	(1)	-	-	(1)	-	-	(1)	(1)	-
Trp	-	-	-	-	0.60 (1)	-	-	-	0.82 (1)	2.14 (2)	-	-
Total residues	11	17	7	20	5	4	13	3	21	11	33	10
Yield (%)	29	17	24	21	27	30 %	58	46	77	2	10	50

Isolation of methionine peptides from the trypsin-solubilized NADPH–cytochrome *P*-450 reductase from pig liver

Methionine-containing peptides were prepared from the reductase modified with [^{14}C]methyl iodide according to Sasagawa *et al.* (1983) by digestion with trypsin.

Sequence analysis

Automated Edman degradation using the solid-phase technique was performed as in Vogel *et al.* (1985).

Secondary structure predictions

Secondary structure predictions were carried out by using the methods described by Chou & Fasman (1974) and Garnier *et al.* (1978). Hydrophilicity and flexibility profiles were calculated according to Hopp & Woods (1981) and Karplus & Schulz (1985).

RESULTS AND DISCUSSION

Primary structure of the hydrophilic domain in the porcine NADPH–cytochrome *P*-450 reductase

Recently the sequences of 15 CNBr fragments representing 93% of the molecular mass determined for the apoprotein of the hydrophilic domain (70 kDa) have been published (Vogel *et al.*, 1985). A further CNBr peptide, CB16, has been isolated from the CNBr digest of the hydrophilic domain by hplc of the peptide pool IV obtained by gel permeation chromatography on Sephadex G-75 (Fig. 1). A summary of proof of the complete sequence in the CNBr fragments CB1 (Val^{531}–Met^{636}), CB2 (Ala^{406}–Met^{511}), CB4 (Ser^{347}–Met^{405}) and CB16 (His^{302}–Met^{346}) is given in Fig. 1. Data concerning the purification and the amino acid composition of proteolytic peptides establishing the structure of these CNBr fragments are summarized in Table 1. During recon-

Table 2. Methionine-containing peptides isolated from the hydrophilic domain of the porcine NADPH–cytochrome *P*-450 reductase by tryptic digestion

The hydrophilic domain (105 nmol) was modified with monobromobimane (enzyme: reagent = 1/100, mol/mol) in the presence of 6 M-guanidinium chloride at pH 8.0 and subsequently treated with [^{14}C]methyl iodide at pH 4.5 and 25 °C for 24 h [enzyme (99 nmol): [^{14}C]methyl iodide = 1/250, mol/mol]; the yield of [^{14}C]methylated reductase was 79 nmol. The tryptic digest was made by incubation of [^{14}C]methylated reductase (79 nmol; 133 μCi/μmol) for 3 h at 37 °C (reductase: trypsin = 50/1, w/w). The digestion was continued for another 3 h after the addition of the same amount of proteinase. The incubation mixture was separated in 10% (v/v) acetic acid on a Sephadex G-75 column (1.7 cm × 130 cm; flow rate 13.1 ml/h); elution volumes (GPC:EV) are given. Peptides were purified by h.p.l.c. on BC-18 and/or on SC-18; conditions for h.p.l.c. are presented as described in the legend to Table 1. The ^{14}C-labelled peptides were detected in the column eluates by liquid-scintillation counting using a Tricarb 4550 (Packard Instruments, Downers Grove, IL, U.S.A.). The vacuum-dried samples were counted in 2.5 ml of Rotiszint 2211 (C. Roth, Karlsruhe, Germany). The experimental conditions for amino acid analysis are described in the legend to Table 1.

Fragment	M-1	M-2	M-3	M-4	M-5	M-6	M-7	M-8	M-9	M-10	M-11	M-12	M-13
1. GPC: EV (ml)	245-300	267-300	178-204	245-267	204-245	204-245	178-204	245-267	245-267	204-245	245-300	204-245	245-267
2. HPLC BC-18	-	-	1/60/50	1/60/50	1/60/50	1/60/50	1/60/50	1/60/50	1/60/50	1/60/50	1/60/50	1/60/50	1/60/50
RT (min)	-	-	62.73	30-32	40.70	44.35	62.73	8-16	38-42	50.04	8-16	45.72	27.45
3. HPLC SC-18	1/60/50#	1/60/50#	15/60/50	1/60/50#	15/60/50	15/60/50	15/60/50	15/60/50	15/60/50#	15/60/50	1/60/50	15/60/50	1/60/50#
RT (min)	11.577	19.633	72.02	29.94	32.18	28.69	67.04	16.13	32.95	52.94	11.39	39.48	22.30
N-Terminus	(Met)	Tyr	Gly	Thr	Glu	(His)	Tyr	(Met)	Ala	Ala	Asp	Asp	Leu
Sequenz	73-74	105-108	109-167	177-186	244-265	299-311	314-357	405-413	507-514	524-547	635-638	639-664	666-669
Cys	-	-	0.62 (1)	-	-	-	-	-	-	-	-	1.20 (1)	-
Asx	-	-	12.22 (11)	1.03 (1)	2.18 (2)	2.88 (2)	12.45 (10)	-	-	-	1.03 (1)	4.20 (4)	-
Thr	-	-	1.63 (2)	0.80 (1)	1.50 (3)	-	0.62 (-)	-	-	2.83 (3)	-	1.40 (1)	0.78 (1)
Ser	-	-	4.15 (4)	-	-	1.80 (2)	2.32 (4)	1.92 (3)	-	-	-	-	-
Glx	-	-	9.85 (8)	1.35 (1)	2.67 (3)	1.64 (1)	6.17 (5)	1.12 (1)	-	2.17 (2)	-	6.06 (5)	-
Gly	-	1.08 (1)	4.50 (4)	1.20 (1)	2.30 (2)	-	4.31 (3)	2.04 (2)	-	3.08 (4)	-	2.26 (1)	-
Ala	-	-	5.95 (7)	1.03 (1)	1.78 (1)	-	2.70 (4)	1.12 (1)	1.17 (1)	1.77 (2)	0.94 (1)	3.74 (3)	-
Val	-	-	2.30 (3)	-	1.40 (4)	-	2.35 (4)	-	2.04 (2)	2.45 (3)	-	4.54 (4)	-
Met	0.81 (1)	0.90 (1)	1.62 (2)	1.00 (1)	1.53 (2)	1.01 (1)	0.97 (1)	1.04 (1)	1.00 (1)	0.89 (1)	0.99 (1)	1.00 (1)	1.00 (1)
Ile	-	-	0.61 (1)	-	-	1.10 (1)	0.83 (2)	-	-	2.24 (3)	-	0.94 (1)	-
Leu	-	-	5.23 (6)	-	1.40 (1)	3.20 (3)	5.12 (6)	-	1.21 (1)	-	-	-	1.23 (1)
Tyr	-	1.09 (1)	2.21 (3)	0.72 (1)	1.65 (2)	-	1.75 (2)	-	-	-	-	1.01 (1)	-
Phe	-	-	2.17 (2)	0.67 (1)	-	-	-	-	0.94 (1)	1.73 (2)	-	1.52 (1)	-
His	-	-	-	0.62 (1)	1.02 (1)	1.50 (2)	1.01 (1)	-	-	-	-	0.92 (1)	-
Lys	1.27 (1)	-	0.97 (1)	0.95 (1)	-	1.14 (1)	0.95 (1)	0.80 (1)	-	-	-	1.04 (1)	0.98 (1)
Arg	-	1.32 (1)	-	-	1.02 (1)	-	-	-	0.90 (1)	0.90 (1)	0.98 (1)	-	-
Pro	-	-	(3)	-	-	-	(1)	-	(1)	(3)	-	(1)	-
Trp	-	-	(1)	-	-	-	-	-	-	-	-	-	-
Total residues	2	4	59	10	22	13	44	9	8	24	4	26	4

sideration of our published sequences, the residue at position 447 has to be changed into Leu and at position 457 into Ser. Furthermore the *C*-terminal sequences in CB2 and CB10 have to be corrected to -Ala-Leu-Val-Pro-Met (residues 507–511) and -Glu-Gln-Gly-Pro-Met (residues 650–654) respectively.

Haniu *et al.* (1984, 1985) sequenced methionine-containing peptides connecting some porcine CNBr fragments. However, the alignment of the peptides CB3, CB4, CB5, CB7, CB13, CB15 and CB16 in the porcine reductase has not yet been shown. The tryptic methionine-containing peptides M4–M8 (Table 2) provided the missing overlaps. The peptides were aligned with the sequences of the porcine CNBr fragments on the basis of amino acid composition, *N*-terminal sequencing and M_r as determined by gel permeation chromatography. The overlapping sequences in question occur beyond the strongly conserved region, when compared with the rat enzyme, and for this reason the complete primary structure of the porcine hydrophilic domain could be established (Figs. 1 and 2).

Comparison between the amino acid sequences of rat and porcine NADPH–cytochrome *P*-450 reductase

The overall homology between the hydrophilic domains (residues 57–678) from porcine and rat NADPH–cytochrome *P*-450 reductase is 91.8%. The high similarity is demonstrated by the feature that all glycine, methionine, leucine, phenylalanine, tryptophan and proline residues present in the pig liver enzyme are conserved in the rat reductase, except for Pro^{600}. The porcine hydrophilic domain shows two additional exchanges (Tyr^{168} → Phe^{168}, His^{230} → Phe^{230}), and the accessible Cys^{645} (Vogel *et al.*, 1985) instead of Tyr^{645} in the rat enzyme (Fig. 2). Local accumulations of amino acid replacements have been observed in the region (Ile^{57}–Ser^{67}) connecting to the membrane segment and in the sequences Ile^{128}–Val^{134}, His^{250}–Val^{258} and Ala^{580}–Lys^{602} (Fig. 2). The number of amino acid replacements, with an average of 3.5 conservative exchanges/100 residues, is least in the region 264–511 (Fig. 2). This chain segment contains sequences exhibiting homology to other nucleotide-binding proteins (Porter & Kasper, 1985).

The cosubstrate NADPH protects the accessible Cys^{472} and Cys^{566} residues against modification with -SH reagents [monobromobimane and 5,5′-dithiobis-(2-nitrobenzoate)] (Vogel *et al.*, 1985). This result locates the NADPH-binding areas in the steric or sequential vicinity of these two cysteine residues. The essential Cys^{566} is in addition positioned within the sequence Gln^{551}–His^{583} which displays not only homology with the NADPH-binding domain of the ferredoxin-NADP$^+$ oxidoreductase (EC 1.18.1.2) (Fig. 3) but also with the sequences His^{233}–Gln^{265} in fumarate reductase (EC 1.3.99.1) of *Escherichia coli* and His^{242}–Glu^{274} in succinate dehydrogenase (EC 1.3.99.1) of *E. coli* (six matches and 20 conservative substitutions), which represent parts of the active sites in these flavoproteins (Wood *et al.*, 1984). Furthermore there is some similarity of the sequence Gly^{558}–Leu^{588} to that of Gly^{85}–Lys^{113} (including two gaps) in *Desulfovibrio vulgaris* flavodoxin [a sequence containing FMN-contacting residues ("close to Tyr^{98}"; Watenpaugh *et al.*, 1973; Dubourdieu & Fox, 1977)] and of Gly^{488}–Val^{509} in the NADPH–cytochrome *P*-450 reductase to Gly^{27}–Val^{48} in the FAD-binding

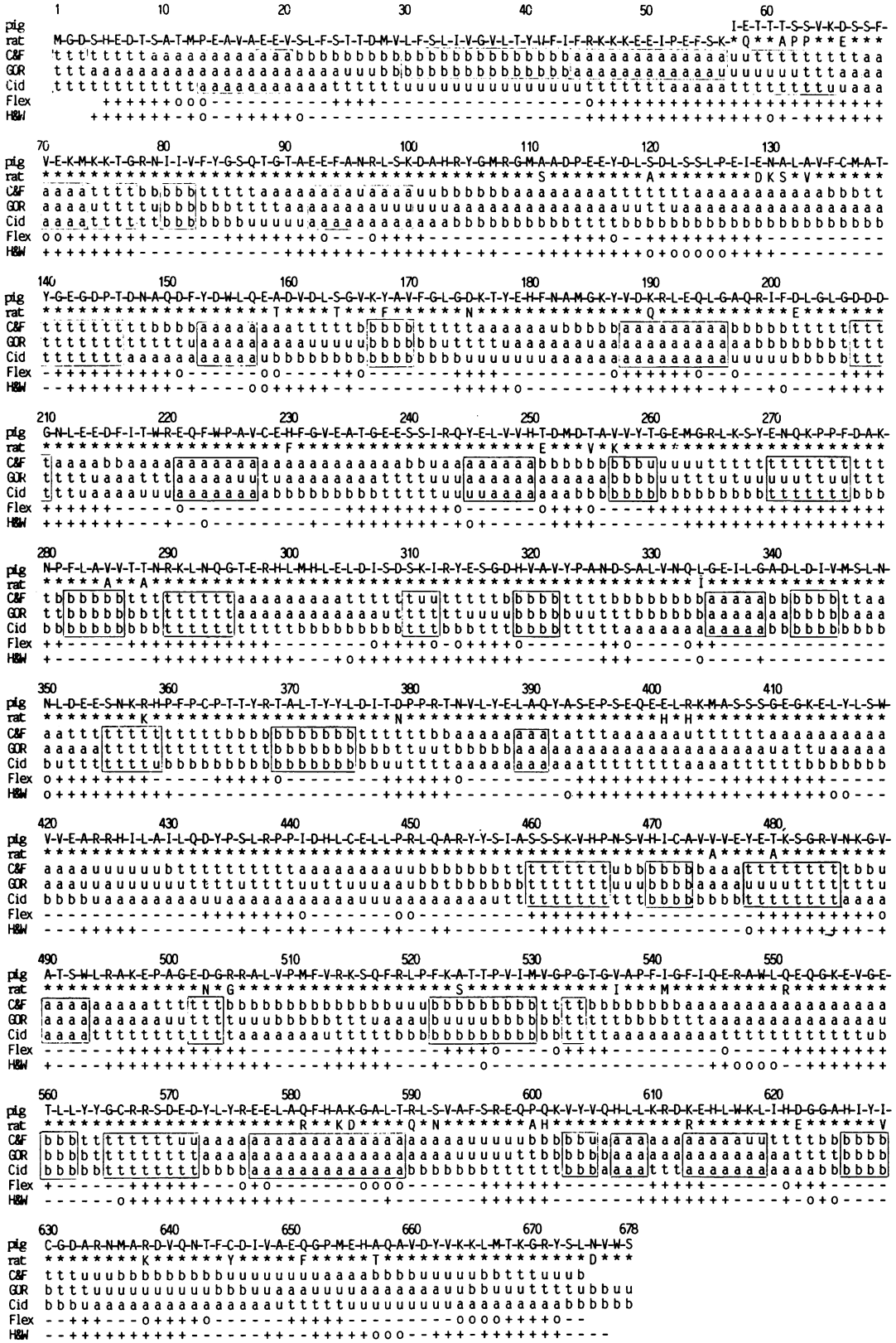

Fig. 2. Comparison of the primary structure of porcine and rat NADPH–cytochrome *P*-450 reductases and predicted secondary structure of the porcine NADPH–cytochrome *P*-450 reductase

Abbreviations: *, identical amino acid residue in the rat reductase. Predictive methods: C&F, Chou & Fasman (1974); GOR, Garnier *et al.* (1978); Cid, Cid *et al.* (1982); residues are assigned a, helical; b, β-strand; t, β-turn/irregular structure; u, not predicted. Hydropathy profile (H&W; Hopp & Woods, 1981): +, hydrophilicity value > 0; −, hydrophilicity < 0. Flex (chain flexibility): +, $B_{norm.}$ value > 0; −, $B_{norm.}$ value < 0.

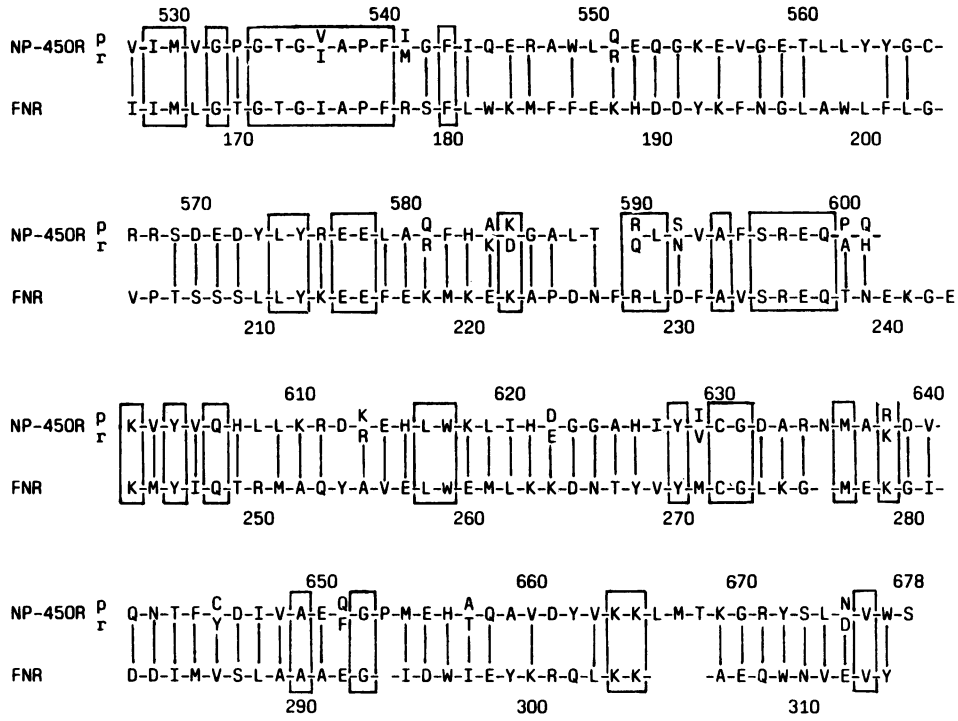

Fig. 3. Alignment of NADPH–cytochrome *P*-450 reductase (NP-450R; p, pig; r, rat) (sequence 528–678) and ferredoxin–NADP$^+$ oxidoreductase (FNR) from spinach (sequence 165–314)

Identical residues are enclosed in boxes and conservatively substituted residues with scores $\geqq 0.1$ in the MDM_{78} matrix (Schwartz & Dayhoff, 1978) are identified by vertical bars.

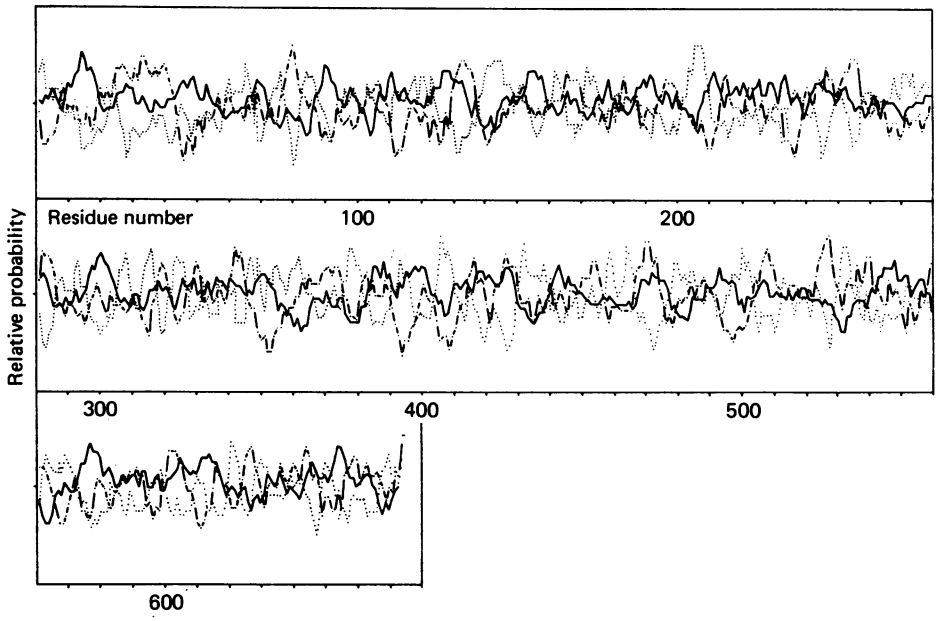

Fig. 4. Chou–Fasman profile of pig liver NADPH–cytochrome *P*-450 reductase

Helix is indicated by ——, β-strand or extended structure by ·–·–· and β-turn by · · · · ·. The average helical-, β-sheet-potential and β-turn probability of five-residue segments are shown.

domain of glutathione reductase (EC 1.6.4.2). Sequences of FAD-binding domains in other flavoproteins [*p*-hydroxybenzoate hydroxylase (EC 1.14.13.2), fumarate reductase and succinate dehydrogenase from *E. coli*] show, according to Porter & Kasper (1985), homology with sequences within the region Asp^{202}–Val^{509} of the NADPH–cytochrome *P*-450 reductase. In summary, the localization of the NADPH-protected cysteine residues of the NADPH–cytochrome *P*-450 reductase in areas exhibiting homology with FAD- or NADPH-binding

regions of other flavoproteins confirms the concept of an overlap between the FAD and the NADPH domains in the NADPH–cytochrome *P*-450 reductase.

(Pyro)phosphoryl binding sites for flavocoenzymes and NADPH in the hydrophilic domain of the NADPH–cytochrome *P*-450 reductase

NADPH–cytochrome *P*-450 reductase contains in the hydrophilic domain 1 molecule of FMN and FAD each. In addition, NADPH–cytochrome *P*-450 reductase binds the cosubstrate NADPH with a molar ratio of 1:1 (Lumper *et al.*, 1980). Therefore three (pyro)phosphoryl group binding sites are to be expected per reductase molecule: two of the dinucleotide (FAD, NADPH) and one of the mononucleotide (FMN) type. The sequence Val^{528}–Ile^{544} of the NADPH–cytochrome *P*-450 reductase shows a strong homology (11/17 matches) with the so-called glycine-rich region (positions 165–180) of the spinach ferredoxin–$NADP^+$ oxidoreductase (Karplus *et al.*, 1984) (Fig. 3). The conserved sequence:

$$\text{Gly-Pro-Gly-Thr-Gly-}\begin{matrix}\text{Ile}\\\text{Val}\end{matrix}\text{-Ala-Pro-Ile-}\begin{matrix}\text{Met}\\\text{Ile}\end{matrix}\text{-Gly-Phe}$$

(positions 532–540) of the NADPH–cytochrome *P*-450 reductase and:

Gly-Thr-Gly-Thr-Gly-Ile-Ala-Pro-Phe-Arg-Ser-Glu

(positions 169–180) of the spinach ferredoxin–$NADP^+$ oxidoreductase can be adapted to the pyrophosphoryl binding consensus sequence of the dinucleotide type:

$$\overset{1}{\text{Gly}}\text{-Xaa-}\overset{3}{\text{Gly}}\text{-Xaa-Xaa-}\overset{5}{\text{Gly}}\overset{6}{}$$

(Möller & Amons, 1985) by replacing residue 5 by a gap. The sequence Gly^{532}–Phe^{540} is most likely to be part of the $NADP^+$-binding site in the NADPH–cytochrome *P*-450 reductase, since the homologous sequence of the spinach ferredoxin–$NADP^+$ oxidoreductase has been identified as an NADPH-binding site by X-ray studies (Sheriff & Herriot, 1981). A distinctly lower degree of similarity (9/36 homology) is observed between the segments Ala^{524}–Thr^{560} in the NADPH–cytochrome *P*-450 reductase and Ser^{259}–Thr^{295} of the human glutathione reductase, which however does not contain the pyrophosphate loop (Val^{191}–Glu^{201}) of the NADPH domain (Krauth-Siegel *et al.*, 1982).

Lys^{244} in the spinach ferredoxin–$NADP^+$ oxidoreductase is modified by affinity labelling with periodate-oxidized $NADP^+$ and therefore is postulated to be localized in the NADPH-binding site of the enzyme (Chan *et al.*, 1985). Lys^{602} in the NADPH–cytochrome *P*-450 reductase, assuming five deletions, is located the same distance from the glycine-rich region as is the reactive Lys^{244} in the ferredoxin–$NADP^+$ oxidoreductase (Fig. 3). Lys^{602}–Gln^{606} of the NADPH–cytochrome *P*-450 reductase shows strong sequence identity with Lys^{244}–Gln^{248} in the plant enzyme. Comparison of the *C*-terminal half of the ferredoxin–$NADP^+$ oxidoreductase (Gly^{169}–Tyr^{314}) with the sequence Gly^{538}–Ser^{678} in the NADPH–cytochrome *P*-450 reductase revealed an obviously extensive homology between both proteins in the NADPH-binding domain. Apparently the relationship between other flavoproteins and the NADPH–cytochrome *P*-450 reductase is specific either to the FMN- or the FAD/NADPH-domains. This observation does not support development from a common ancestor but sustains the alternative mechanism of gene fusion as suggested by Porter & Kasper (1985). Conservation of the backbone conformation in the homologous regions of the NADPH–cytochrome *P*-450 reductase and the ferredoxin–$NADP^+$ oxidoreductase is however unlikely, since all glycine and proline residues in the area 528–678 of the NADPH–cytochrome *P*-450 reductase outside the glycine-rich sequence are exchanged and the construction of gaps is necessary to obtain homology.

Comparative studies revealed the sequence Val^{82}–Ala^{91} (Fig. 2) as a second phosphoryl group binding site (Porter & Kasper, 1985) on the basis of the strong homology with the binding sequence Ile^{6}–Thr^{15} for FMN phosphate in *Desulfovibrio vulgaris* flavodoxin (Dubourdieu *et al.*, 1977). Thr^{15} is conserved in all flavodoxins studied, but replaced by Ala^{91} in the corresponding sequence of the NADPH–cytochrome *P*-450 reductase.

Prediction of secondary structure

The results of secondary structure prediction by using the procedure of Chou & Fasman (1974) and Garnier *et al.* (1978) are presented in Figs. 2 and 4. The repeating pattern of predicted α-helices, β-strands, β-turns and irregular structures in the hydrophilic domain of the NADPH–cytochrome *P*-450 reductase clearly supports its classification (Levitt & Chothia, 1976) as an α/β-protein. The c.d. spectrum of the hydrophilic domain (pig liver) showing a single broad minimum skewed to 220 nm ($\Delta\epsilon$ approx. 27 $M^{-1}\cdot cm^{-1}$) (Ehrig, 1974) agrees, according to Manavalan & Johnson (1983), with the predicted secondary structure. However, the helix content of 19.5% calculated from θ_{220} is lower than the percentage of total amino acid residues assigned consistently as helical (26%) by the predictive methods. Better agreement has been achieved for detergent-solubilized NADPH–cytochrome *P*-450 reductase (rat) by c.d. spectroscopy in the presence of 20% glycerol and 0.1% sodium deoxycholate (Knapp *et al.*, 1977).

The prediction of secondary structure allows comparison with the folding structure associated with nucleotide-binding sequences. The 'ADP $\beta\alpha\beta$-fold' (Wierenga *et al.*, 1985) is distinguished by a glycine residue at the *N*-terminus of the pyrophosphate-binding helix. According to the secondary structure calculations the glycine residues in the NADPH-binding sequence:

$$\overset{532}{\text{Gly-Pro-Gly-Thr-Gly-}}\begin{matrix}\text{Val}\\\text{Ile}\end{matrix}$$

are however localized within a β-turn between two β-strands. The first α-helix following this region begins at Gln^{545}. The proposed folding structure of the *C*-terminal region Phe^{522}–Ser^{678} in the NADPH–cytochrome *P*-450 reductase corresponds on the other hand to the expected topology of a NADPH-binding domain [e.g. ferredoxin–$NADP^+$ oxidoreductase (Sheriff & Herriott, 1981)] containing a pleated sheet with interconnecting α-helices or non-repetitive secondary structure.

Prediction of the chain flexibility in the hydrophilic domain of the NADPH–cytochrome *P*-450 reductase

Using the procedure proposed by Karplus & Schulz (1985) segments of high flexibility are predicted in the regions containing the β-turns. High flexibility is also calculated for areas not consistently predicted as secondary structure by the statistical methods and possibly representing stretches of irregular structure

(e.g. Thr^{260}–Pro^{281} and Ser^{308}–Asp^{318}) and also for α-helices with high charge density (Ser^{397}–Leu^{402} and Gln^{545}–Gly^{558}). Regions predicted to have high flexibility include the preferred cleavage site for trypsin (peptide bond Lys^{56}/Ile^{57}) in the region connecting the membrane and the hydrophobic domain, the accessible cysteine residues (Cys^{472}, Cys^{545} and Cys^{565} in the porcine reductase) which border a region of high flexibility, and the prospective nucleotide-binding sequences Val^{82}–Ala^{91} and Gly^{532}–Gly^{535}.

The plot of hydrophilicity versus sequence positions (Hopp & Woods, 1981) (Fig. 2) shows four points with maximal hydrophilicity (Lys^{46}, Lys^{190}, Asp^{208} and Glu^{270}) which are positioned in flexible regions but do not correspond to the residues of numerically highest predicted flexibility ($B_{norm.}$ value $\leqq 1.110$) (Ser^{355}, Gly^{409} and Gln^{552}). Hydropathy profiles and flexibility plots apparently furnish complementary information on possible antigenic sites in proteins (Karplus & Schulz, 1985).

We are indebted to Professor Dr. S. Stirm and Dr. D. Linder for the use of their laboratory equipment, and to Professor Dr. G. Schulz for an introduction to the computerized secondary structure prediction of proteins. We thank Mr. F. Busch and Mrs. M. Wagner for expert technical assistance.

REFERENCES

Black, S. S. & Coon, M. J. (1982) J. Biol. Chem. **257**,5929–5938
Black, S. D., French, J. S., Williams, C. H., Jr. & Coon, M. J. (1979) Biochem. Biophys. Res. Commun. **91**, 1528–1535
Chan, R. L., Carrillo, N. & Vallejos, R. H. (1985) Arch. Biochem. Biophys. **240**, 172–177
Chou, P. Y. & Fasman, G. D. (1974) Biochemistry **13**, 222–245
Cid, H., Bunster, M., Arriagada, E. & Campos, M. (1982) FEBS Lett. **150**, 247–254
Dubourdieu, M. & Fox, J. L. (1977) J. Biol. Chem. **252**, 1453–1463
Ehrig, H., (1974) Thesis, Universität Gießen, pp. 116–119
Garnier, J., Osguthorpe, D. J. & Robson, B. (1978) J. Mol. Biol. **120**, 97–120
Gilette, J. R., Davis, D. C. & Sasame, H. A. (1972) Annu. Rev. Pharmacol. **12**, 57–84
Haniu, M., Iyanagi, T., Legesse, K. & Shively, J. E. (1984) J. Biol. Chem. **259**, 13703–13711
Haniu, M., Iyanagi, T., Miller, P. & Shively, J. E. (1985) Biochem. Biophys. Res. Commun. **127**, 94–98
Hopp, T. P. & Woods, K. R. (1981) Proc. Natl. Acad. Sci. U.S.A. **78**, 3824–3828
Karplus, P. A. & Schulz, G. E. (1985) Naturwissenschaften **72**, 212–213
Karplus, P. A., Walsh, K. A. & Herriott, J. R. (1984) Biochemistry **23**, 6576–6583
Knapp, J. A., Dignam, J. D. & Strobel, H. W. (1977) J. Biol. Chem. **252**, 437–443
Krauth-Siegel, R. L., Blatterspiel, R., Saleh, M., Schiltz, E., Schirmer, R. H. & Untucht-Grau, R. (1982) Eur. J. Biochem. **121**, 259–267
Levitt, M. & Chothia, C. (1976) Nature (London) **261**, 552–558
Lumper, L., Busch, F., Dzelic, S., Henning, J. & Lazar, T. (1980) Int. J. Peptide Protein Res. **16**, 83–96
Manavalan, P. & Johnson, W. C., Jr. (1983) Nature (London) **305**, 831–832
Masters, B. S. S. & Okita, R. T. (1980) Pharmacol. Ther. **9**, 227–241
Möller, W. & Amons, R. (1985) FEBS Lett. **186**, 1–7
Porter, T. D. & Kasper, C. B. (1985) Proc. Natl. Acad. Sci. U.S.A. **82**, 973–977
Sasagawa, T., Titani, K. & Walsh, K. A. (1983) Anal. Biochem. **128**, 371–376
Sheriff, S. & Herriott, J. R. (1981) J. Mol. Biol. **145**, 441–451
Schwartz, R. M. & Dayhoff, M. O. (1978) Atlas of Protein Sequence and Structure, vol. 5, pp. 353–358, National Biomedical Research Foundation, Washington
Vogel, F. & Lumper, L. (1983) Biochem. J. **215**, 159–166
Vogel, F. & Lumper, L. (1984) Hoppe-Seyler's Z. Physiol. Chem. **365**, 1074
Vogel, F. & Lumper, L., (1985) in Modern Methods in Protein Chemistry (Tschesche, H., ed.), vol. 2, pp. 185–190, Walter de Gruyter, Berlin and New York
Vogel, F., Kaiser, C., Witt, I. & Lumper, L. (1985) Biol. Chem. Hoppe-Seyler **366**, 577–588
Watenpaugh, K. D., Sieker, L. C. & Jensen, L. H. (1973) Proc. Natl. Acad. Sci. U.S.A. **70**, 3857–3860
Wierenga, R. K., De Maeyer, M. C. H. & Hol, W. G. J. (1985) Biochemistry **24**, 1346–1357
Wood, D., Darlison, M. G., Wilde, R. J. & Guest, J. R. (1984) Biochem. J. **222**, 519–534
Yoshinaga, T., Sassa, Sh. & Kappas, A. (1982) J. Biol. Chem. **257**, 7786–7793

Received 8 August 1985/17 January 1986; accepted 19 February 1986