

Molecular cloning and sequence analysis of cDNA encoding Δ^4 -3-ketosteroid 5β -reductase of rat liver

Yoshiaki Onishi¹, Mitsuhide Noshiro², Tsunehiro Shimosato¹ and Kyuichiro Okuda²

¹Department of Oral and Maxillofacial Surgery II and ²Department of Biochemistry, Hiroshima University School of Dentistry, Hiroshima 734, Japan

Received 4 March 1991

A cDNA clone encoding Δ^4 -3-ketosteroid 5β -reductase was isolated from rat liver cDNA libraries using antibodies specific for the enzyme and oligonucleotides as probes. The cDNA contained 981-base pair open reading frame encoding 327 amino acid residues (M_r 37 376) and an unusually long 3'-untranslated region rich in AT sequence in the total length of 3189 base pairs. The predicted amino acid sequence contains the sequences similar to the putative NADPH- and steroid-binding regions.

Δ^4 -3-Ketosteroid 5β -reductase; cDNA cloning

1. INTRODUCTION

In bile acids synthesis and steroid hormone metabolism, Δ^4 -3-ketosteroid 5β -reductase plays an important role to catalyze reduction of the Δ^4 -double bond to give A/B-*cis* conformation [1]. The enzyme was recently purified to homogeneity in this laboratory [2,3]. Subsequent studies have shown that the N-terminal amino acid of this enzyme is blocked, prompting us to determine the amino acid sequences of peptides obtained by peptidase-treatment and prepare the specific monoclonal and polyclonal antibodies against the enzyme (Onishi et al., to be published). In this paper, we describe the isolation of a cDNA clone encoding Δ^4 -3-ketosteroid 5β -reductase from rat liver cDNA libraries using the specific antibodies and synthetic oligonucleotides corresponding to the partial amino acid sequences as probes.

2. MATERIALS AND METHODS

Δ^4 -3-Ketosteroid 5β -reductase was purified from male rat liver cytosol as described previously [2]. Specific polyclonal antibodies were prepared by immunizing BALB/c female mice with the purified protein mixed with Ribi adjuvant as described previously [4]. Oligonucleotides were synthesized based on the amino acid sequence of a peptide fragment of the enzyme ((Lys)-Thr-Phe-Ile-Ala-Val-Lys) as follows 5'-TT IAC IGC G(A/T)AT A(G)AA IGT T(C)TT-3'.

The cDNA libraries were prepared from liver poly(A)⁺RNA of male rats [5], using λ gt 11 and λ ZAP vectors. A λ gt 11 oligo(dT)-primed cDNA library was screened with the specific polyclonal antibodies and a ³²P-labeled family of twelve 20-mer oligonucleotides.

The cDNA obtained was used as a probe to isolate a complete cDNA clone from these libraries [6]. Positively reacted clones isolated through several rounds of screening were subcloned into pBluescript SK(-) plasmid. DNA sequencing was performed by using Exo III/Mung bean nuclease deletion system (Takara Co.) [7] and sequenase kit (United States Biochemical Corp.). Southern hybridization and Northern hybridization were performed by the method of Maniatis et al. [6].

3. RESULTS AND DISCUSSION

Specific polyclonal antibodies were prepared against rat liver Δ^4 -3-ketosteroid 5β -reductase. The antibodies specifically reacted to the enzyme as examined by Western blotting and therefore were used for screening. Out of 1×10^5 clones of λ gt 11 oligo(dT)-primed cDNA library, 8 immunoreactive clones were isolated and subjected to Southern hybridization using a mixture of ³²P-labeled oligonucleotides corresponding to the amino acid sequence of the internal peptide as a probe. Seven clones (λ 2, 5, 7, 9, 11, 12 and 13) which hybridized positively with the oligonucleotides were subjected to restriction mapping. As shown in fig. 1, λ 5, 7, 9, 12 and 13 clones showed a common size (1.8 kbp) and the same restriction map, and λ 2 and 11 clones which had 3.0 kbp long insert contained the longer 3'-end. The mRNA size of Δ^4 -3-ketosteroid 5β -reductase was estimated to be about 3.2 kb long by Northern hybridization using λ 2 cDNA clone insert as a probe (Fig. 2). The insert of λ 2 cDNA clone was then subcloned into pBluescript (p5 β -2) and subjected to nucleotide sequencing. The nucleotide sequence of p5 β -2 contained 903 bp long open reading frame consisting of 301 amino acid residues (M_r 34311). The deduced amino acid sequence contained the sequences of all the peptides obtained by

Correspondence address: K. Okuda, Department of Biochemistry, Hiroshima University School of Dentistry, Kasumi 1-2-3, Minami-ku, Hiroshima 734, Japan

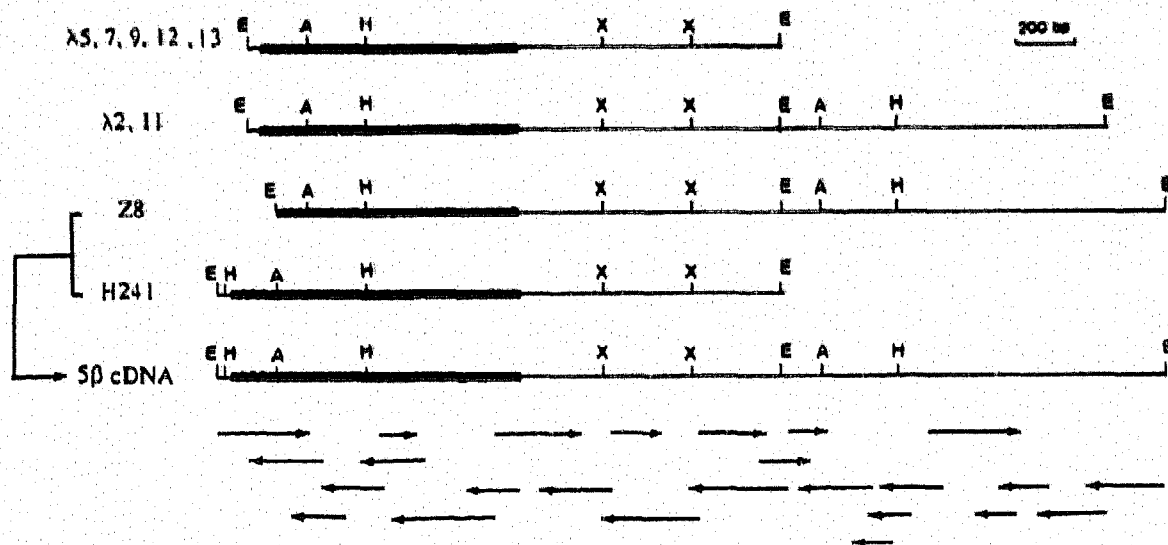


Fig. 1. Restriction map and sequencing strategy of Δ^4 -3-ketosteroid 5β -reductase cDNA which constructed from Z8 and H241. Arrows indicate the directions and extents of sequencing. In the diagram, the restriction sites are abbreviated: A is *Acl*I, E is *Eco*RI, H is *Hind*III, and X is *Xba*I.

peptidase-treatment of the purified enzyme (Onishi et al., to be published) except one (SNCATWEYLEACK) (Fig. 3). Furthermore, the molecular weight calculated from the deduced sequence was much smaller than that of the purified protein (M_r 37000). These results indicated that the isolated clones might be incomplete. In order to isolate a full size clone therefore, oligo(dT)-primed λ ZAP cDNA library and random-primed λ gt 11 cDNA library were rescreened using 32 P-labeled p5 β -2 cDNA insert as a probe. A number of positive clones were isolated and two of them had the region missing in the clones isolated previously. The restriction map (Fig. 1) and nucleotide sequence (Fig. 3) revealed that Z8 clone isolated from the oligo(dT)-primed λ ZAP cDNA library contained the polyadenylation signal and

H241 clone isolated from the random-primed λ gt 11 cDNA library had a larger coding region which codes the amino acid sequence missing in p5 β -2. The overlapping portion of the two sequences was identical. Consequently, the entire sequence for Δ^4 -3-ketosteroid 5β -reductase was constructed from Z8 and H241. The nucleotide sequence thus constructed was 3189 bp long and contained 978 bp long open reading frame which codes 326 amino acid residues. The calculated molecular weight (M_r 37376) was in agreement with that of the purified enzyme (M_r 37000). It is not known at this moment why many clones missing the amino acid sequence such as p5 β -2 are abundant in the liver cDNA libraries.

As shown in Fig. 3, 5 β cDNA clone contained a long 3'-noncoding region which was rich in AT nucleotides and often contained ATTTA motifs, 5'-AAT-3' or 5'-TAA-3' trinucleotides in the single strand region of the secondary structure. Such unique structures in 3'-noncoding region are known to exist in rapidly degrading mRNA [8,9]. Recently, a similar unique structure was observed in 3'-noncoding region of P-450_{ch7 α} which plays an important role in the conversion of cholesterol to bile acids whose mRNA showed rapid degradation and circadian rhythm [10,11]. Although these unique structures are suggestive of the rapid turnover of the mRNA of Δ^4 -3-ketosteroid 5β -reductase, circadian rhythm was not observed in the mRNA level (data not shown).

Comparison of the amino acid sequence with the NADPH-binding enzyme and sex hormone-binding globulin revealed that the two unique amino acid regions exist in Δ^4 -3-ketosteroid 5β -reductase. One region from residue 86-101 in Fig. 3 seems to be involved in steroid-binding. Although this sequence is not homologous to the common sequence of steroid-

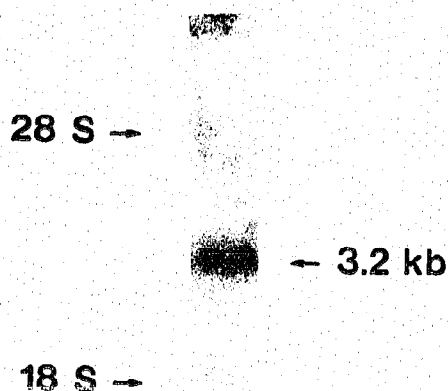


Fig. 2. Northern hybridization of rat liver poly(A)⁺RNA. Five μ g of poly(A)⁺ RNA was electrophoresed on agarose gel containing formaldehyde [6]. A 32 P-labeled insert of p5 β -2 was used as a probe.

1	1	TTCAGAAATCACTACTTCAAGCTTCAATATCTCTCTCTACA	ATG AAG CTC AAG AAT GGA AAG CAG CAG ATA	70
			Met Arg Leu Ser Thr Ala Arg His His Ile	10
71	11	CCC CTA AAT GAT GGT AAG AGC ATT CCG ATC ATC GGG GTT GGA ACC TAC TCA GAG CTT AGA CCG GTA CCT GGC		142
		Pro Leu Arg Arg Gly Arg Ser Ile Pro Ile Ile Gly Leu Gly Thr Tyr Ser Asp Pro Arg Pro Val Pro Gly		34
143	35	AAG ACC TTT ATA GGA GTG AAG ACA GGT ATT GAG GAG GGG TAC GGG CAT ATT GAT GGG GGC TAC CTC TAC CGA		214
		Lys Thr Phe Ile Ala Val Lys Thr Ala Ile Asp Gly Tyr Arg His Ile Asp Gly Ala Tyr Val Tyr Arg		58
215	59	AAT GAA GAT GAA GTG GGT GAG GGC ATC AGA GAA AAG GTG GCA GAA GGG AAG GTA AAG AGG GAA GAG ATT TTC		286
		Asn Glu His Glu Val Gly Glu Ala Ile Arg Gly Lys Val Ala Gly Gly Lys Val Lys Arg Glu Glu Ile Phe		82
287	83	TAC TGT GGA AAG TTA TGG AGT ACA GAG CAT GAT CCA GAG ATG GTG GGC GGA GGC GTG GAA AGG ACC GTG CAG		358
		Tyr Cys Gly Lys Leu Trp Ser Thr Asp His Asp Pro Glu Met Val Arg Pro Ala Leu Glu Arg Thr Leu Glu		106
359	107	ACC CTC AAG CTA GAT TAC ATA GAC CTT TAT ATC ATT GAG ATG CCC ATG GGC TTT AAU GCT GGA GAG GAA TTT		430
		Thr Leu Lys Leu Asp Tyr Ile Asp Leu Tyr Ile Glu Met Pro Met Ala Phe Lys Pro Gly Glu Glu Phe		130
431	131	TAT GCT AAA GAT GAG AAT GGC GGA GTG ATA TAC CAC AAA TGA AAT GTG TGT GGC AAG TGG GAG GGA GTG GAA		502
		Tyr Pro Lys Asp Glu Asn Gly Arg Val Ile Tyr His Lys Ser Asn Leu Cys Ala Thr Trp Glu Ala Leu Glu		154
503	155	GCT TGG AAA GAG GCT GGC TTG GTG AAA TGG CTT GGG GTG TGT AAT TTT AAG GGC AGG CAG GTG GAG GTG ATC		574
		Ala Cys Lys Asp Ala Gly Leu Val Lys Ser Leu Gly Val Ser Asn Phe Asn Arg Arg Glu Leu Glu Val Ile		178
575	179	TTG AAG AAG CCA GGA CTC AAG TAC AAG CCT GTC ACC AAG CAG GTG GAG TGC CAC CCG TAT TTC ACC CAG ACA		646
		Leu Arg Lys Pro Gly Leu Lys Tyr Lys Pro Val Thr Asn Ser Val Glu Cys His Pro Tyr Phe Thr Glu Thr		202
647	203	AAA CTC CTT GAA GTT TCT CCC AGC AGC ATG ACA TCG TTC ATT GTC GCA TAC AGC CCC TTA GGG ACC TGT CCG		718
		Lys Leu Leu Glu Val Ser Ala Ser Ser Met Thr Ser Phe Ile Val Ala Tyr Ser Pro Leu Gly Thr Cys Arg		226
719	227	AAC CCG TTA TGG GTG AAT GTA TCT TCT CCA CCC TTG TTA AAG GAT GAA CTC CTA ACC TCG GTG GGG AAA AAG		790
		Asn Pro Leu Trp Val Asn Val Ser Ser Pro Pro Leu Leu Lys Asp Gly Leu Leu Thr Ser Leu Gly Lys Lys		250
791	251	TAC AAT AAG ACA GAA GCT CAA ATT GTG TCG CTT GTC GAC ATT CAG CAG GAG GTG GTT GTC ATC CCT AAA AGT		862
		Tyr Asn Lys Thr Glu Ala Glu Ile Val Leu Arg Phe Asp Ile Glu Arg Gly Leu Val Val Ile Pro Lys Ser		274
863	275	ACT ACC CCG GAA AGG ATC AAA GAA AAG TTT CAG ATC TTT GAG TTC TCT CTC ACC AAA GAA GAA ATG AAG CAC		934
		Thr Thr Pro Glu Arg Ile Lys Glu Asn Phe Glu Ile Phe Asp Phe Ser Leu Thr Lys Glu Glu Met Lys Asp		298
935	299	ATT GAA GGC TTG AAT AAA AAG GTG CCG TTT GTG GAG ATG CTC ATG TGG AGT GAT CAT CCT GAA TAC CCA TTT		1006
		Ile Glu Ala Leu Asn Lys Asn Val Arg Phe Val Glu Met Leu Met Trp Ser Asp His Pro Glu Tyr Pro Phe		322
1007	323	CAT GAC GAA TAC TGA ACATGCAAAATTTCTTCAGTGGAGTTTCTTTTGTCTTTTCATTCGCGATCTTGAGCTGATTATCCCTCCAGTC		1078
		His Asp Glu Tyr ...		326
1097		CTAGAAATTAATTCGGCTCTTCTATCTTCAGATTAAAGTTGAGCAGAAATACACCATGCTTAAGCTTGTGTATAGTATTAATTCAACTTAGCTCTGAT		1191
1192		GAAGAGAAATTTAAATATATCTAAATAGTTTCTTGGAAATGCTTTACTAGTATCTTTACCTCAATGTGTCTAGATTTGTTGATGGAATTAACC		1266
1287		GAAGCTTTAATTCGAAATTAACAAATATATATAAAATACAAATATATAAAATATATAAAATACAAATATATAAAATACAAATATATAAAATACAAAT		1381
1382		CTGAGACTACCATAGGAGTTTGAATTTGTTTAGCAATTTAGCAGGACAAACAAAGAGGTTTGGCATGCTTCCAGGTGCTGACCATACCATGACCCA		1476
1477		AACAAAGTCATGGAACACAGATGGACACAGCCAGACACAAATTCACATACCAATATTAATGCACCTTGAATATAGACAGAACTGAAAGTCTAGA		1571
1572		AAGTCTAAAGAAACATTAAGCAGGATGTCCCGGGCTTAGAGTTCTCAGCTGAGGACCTTGGCTGCCAGACTGGAGACCCAGGCTGAGGCTGTG		1666
1667		GGTGAAGACTATCAGCAAGGCCATAACGATTTCTGTCTTTACCGCTTTGTTTTCATCCAGTGTGCCCATTCAGCTTTTAATGTGCTGAAAGACAGA		1761
1762		CAGTTACCGGATTCCTTCTTAATTTGTTTGTGTGAGGAGCTGGGAAATGAGACAGGCTGTCAGCTAGTCTGCTCATGTAGCTGCTTACAACT		1856
1857		TGCTGTGACGCTGAGACCACTTTTGAATTTCTGAACCCCTGGTCTCTCTTCCATTTTTCATGGCTGAGCTTCCACCTGACTGGTAAGTCCCACT		1951
1952		TGTACTGTGCTGAATCATTTCTTCCCACTTTCATCAAGGAAACCACTGTAAAGTGGCAAGAAACAATCAATATATACAAACCAAGATTTGTATA		2046
2047		CAATATGCTGGAAGATGTTGATTTTAAAGAAATGATATTTAAATAGCCATTTCTTTCTTTCTTCTGAATCTTTAAATTTATTTTATTTATG		2141
2142		TGCAATGCTGTGATGCTGATGCTATGAGGAATTTGTTGAACCTCATGGAATGGAAGTACAGACAGCTGCTGAGCTGTAGCTGAGCTGCTGAG		2236
2237		AATTGAACCCCAAGTCATGCGAAGAGCTGCAAGTGTCTTAACTGCTGAGCCATCTCCCGAGCCGCCAAATGAATTTGGGAAATAGTAGAAGATA		2331
2332		AAGCTTTGTAAAGTAGTTTAAAGATCTTAAGAAATTTGAGGTACGTATCTGAATTTTGACAAAGCTGCAAGTCTTTTTTTAAATTTGGAATTT		2426
2427		TTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT		2521
2522		AGTGTGACCCACCTACCCACTCCTTCCCT		2616
2617		CCCATTTGCTGCCCAACAAAGGCATCCTCTGCTACATAGGCAGCTGGAGCCATGGGTCTGTCCATGTGTCTCTCTGGATGGTGGTTTACTCCCTGG		2711
2712		GAGCTCTGCTTGGTTGGTTCTGTTGTTATGGGTTGCAACCCCTTCAGCTCCTTTAATCTTTCTCTAATTTCTCCATTTGGGACCCCTGTCTCA		2806
2807		GTCTTAAGGTTGGCTGTGAGCATCCACCTGTGTATTTGTCTGCTCTGAGCAGGCTCTCAGGAGAAATCAGGCTCCTGTGAGCAGGACACAAGTA		2901
2902		ATTTTACCGTATATAAATTTTAAAGAAATGATAAATGATATAGAACGTATCTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT		2996
2997		TCACTTCCAAAGGGTGAAGATCTCTAATATATTTAATTCATAGCAATTTCAATGGCCGAGAGAAATAGTTTTCAGAAATTTATCACITTTAAATTTG		3091
3092		GATCAACATGAACCATTTTGTGCAATAGTGTCTCAATTTGTTAACAATGCTGTTCTCTGAAAATATTTTATTTACATGAATTTAATAAATATA		3186
3187		ATG		

Fig. 3. Nucleotide sequence of Δ^4 -3-ketosteroid 5β -reductase cDNA and the predicted amino acid sequence of the protein. Amino acids determined by peptide sequence analysis (Onishi et al., to be published) are overlined. The nucleotide sequences missing in p5 β -2 are boxed with broken lines. ATTTA motif and similar sequences to it within 3'-noncoding region are marked by dot lines. A sequence of AATAAA indicates a polyadenylation signal.

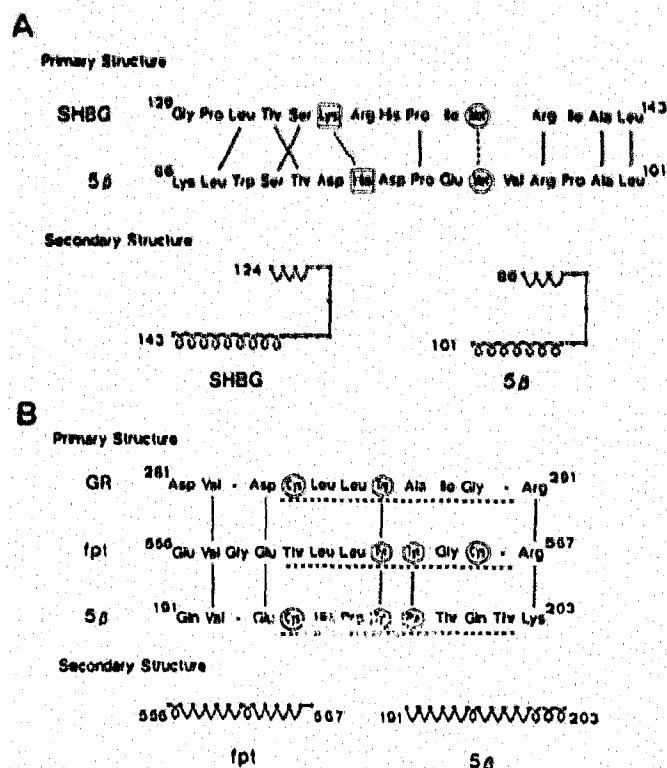


Fig. 4. Two unique regions of Δ^4 -3-ketosteroid 5 β -reductase, steroid-binding (A) and NADPH-binding (B) regions. (A) Residues 86-101 of Δ^4 -3-ketosteroid 5 β -reductase (5 β) and the steroid-binding region of sex hormone-binding globulin (SHBG) are compared. Methionine residues are circled. Positively charged residues which are postulated to bind to the D-ring are boxed. Homologous residues are denoted with vertical lines. In the figures of secondary structure, α -helical regions are denoted by coils, β -sheets by zig-zag lines, and random coils by straight lines. (B) Residues 191-203 of Δ^4 -3-ketosteroid 5 β -reductase (5 β) and NADPH-binding regions of glutathione reductase (GR) and NADPH cytochrome P-450 reductase (fpt) are compared. Cysteine residues are circled and aromatic residues are encircled hexagonally. Homologous residues are denoted with vertical lines, and hydrophobic or nonpolar regions are shown by broken lines. Secondary structures were drawn as indicated above.

binding in the sex hormone-binding globulin [12], the predicted secondary structure of this region calculated by Chou - Fasman methods [13] was very similar to that of the sex hormone-binding globulin (Fig. 4A) and moreover, there exists a methionine residue which is postulated to be required for binding of protein to the A-ring of the steroid nucleus. However, lysine residue which is reported to bind to the D-ring was not found. Instead, positively charged histidine residue was found near the position corresponding to it. These observations suggest that this region may function for substrate binding.

As Δ^4 -3-ketosteroid 5 β -reductase requires NADPH as a cofactor, we have surveyed if there is any similar structure to the NADPH-binding site reported in glutathione reductase [14] or NADPH-cytochrome P-450 reductase [15-18]. As a result, it was found that the region from residue 191-203 in Fig. 3 retains the

similar charged groups and hydrophobic sequences. The location of the aromatic amino acid which is present close to the adenine moiety of NADPH (Fig. 4B) was also acknowledged. The secondary structure around this region calculated by Chou - Fasman method [13] also showed the similarity to the structure of NADPH-binding site of rat hepatic NADPH-cytochrome P-450 reductase [15]. From these results we postulate that this region may be NADPH-binding site of this enzyme.

Curiously, however, the entire amino acid sequence of Δ^4 -3-ketosteroid 5 β -reductase did not show any significant homology to that of Δ^4 -3-ketosteroid 5 α -reductase [19] which works toward the same or similar substrates and requires NADPH as a cofactor.

In conclusion, we have isolated a cDNA clone and determined the primary structure of Δ^4 -3-ketosteroid 5 β -reductase, which catalyzes an important reaction of 5 β -reduction of Δ^4 -3-ketosteroid in the catabolism of cholesterol and metabolism of steroid hormones. The availability of the cDNA probe should lead to insights into the nature and the detailed regulatory mechanism of this important enzyme.

Acknowledgements: We are grateful to Dr. M. Muramatsu (Tokyo University, Tokyo) for his advice and for his allowing us to prepare a cDNA library in his laboratory.

REFERENCES

- [1] Berséus, O., Danielsson, H. and Kallner, A. (1965) *J. Biol. Chem.* 240, 2396-2401.
- [2] Okuda, A. and Okuda, K. (1984) *J. Biol. Chem.* 259, 7519-7524.
- [3] Furuebisu, M., Deguchi, S. and Okuda, K. (1987) *Biochim. Biophys. Acta* 912, 110-114.
- [4] Noshiro, M. and Omura, T. (1987) *J. Biochem.* 83, 61-77.
- [5] Koike, S., Sakai, M. and Muramatsu, M. (1987) *Nucl. Acid Res.* 15, 2499-2513.
- [6] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- [7] Henikoff, S. (1984) *Gene* 28, 351-359.
- [8] Shaw, G. and Kamen, R. (1986) *Cell* 46, 659-667.
- [9] Binder, R., Hwang, S.-P.L., Ratnasabapathy, R. and Williams, D.L. (1989) *J. Biol. Chem.* 264, 16910-16918.
- [10] Noshiro, M., Nishimoto, M. and Okuda, K. (1990) *J. Biol. Chem.* 265, 10039-10041.
- [11] Noshiro, M. and Okuda, K. (1990) *FEBS Lett.* 268, 137-140.
- [12] Namkung, P.C., Kumar, S., Walsh, K.A. and Petra, P.H. (1990) *J. Biol. Chem.* 265, 18345-18350.
- [13] Chou, H.A. and Fasman, G.D. (1978) *Annu. Rev. Biochem.* 47, 251-276.
- [14] Thieme, R., Pai, E.F., Schirmer, R.H. and Schulz, G.E. (1981) *J. Mol. Biol.* 152, 763-782.
- [15] Haniu, M., Iyanagi, T., Legesse, K. and Shively, J.E. (1984) *J. Biol. Chem.* 259, 13703-13711.
- [16] Haniu, M., Iyanagi, T., Miller, P., Lee, T.D. and Shively, J.E. (1986) *Biochemistry* 25, 7906-7911.
- [17] Haniu, M., McManus, M.E., Birkett, D.J., Lee, T.D. and Shively, J.E. (1989) *Biochemistry* 28, 8639-8645.
- [18] Sutter, T.R., Sangard, D. and Loper, J.C. (1990) *J. Biol. Chem.* 265, 16428-16436.
- [19] Andersson, S., Bishop, R.W. and Russell, D.W. (1989) *J. Biol. Chem.* 264, 16249-16255.