Molecular Cloning of the cDNA for the Large Subunit of the High-Ca^{2+}-Requiring Form of Human Ca^{2+}-Activated Neutral Protease†

Shinobu Imajoh,‡ Kazumasa Aoki, Shigeo Ohno, Yasufumi Emori, Hiroshi Kawasaki, Hidemitsu Sugihara, and Koichi Suzuki*

Department of Molecular Biology, The Tokyo Metropolitan Institute of Medical Science, Honkomagome, Bunkyo-ku, Tokyo 113, Japan

Received November 10, 1987; Revised Manuscript Received June 14, 1988

ABSTRACT: A nearly full-length cDNA clone for the large subunit of high-Ca^{2+}-requiring Ca^{2+}-activated neutral protease (mCANP) from human tissues has been isolated. The deduced protein, determined for the first time as an mCANP, has essentially the same structural features as those revealed previously for the large subunits of the low-Ca^{2+}-requiring form (μCANP) [Aoki, K., Imajoh, S., Ohno, S., Emori, Y., Koike, M., Kosaki, G., & Suzuki, K. (1986) *FEBS Lett. 205*, 313–317] and chicken CANP [Ohno, S., Emori, Y., Imajoh, S., Kawasaki, H., Kisaragi, M., & Suzuki, K. (1984) *Nature (London) 312*, 566–570]. Namely, the protein, comprising 700 amino acid residues, is characterized by four domains, containing a cysteine protease like domain and a Ca^{2+}-binding domain. The overall amino acid sequence similarities of the mCANP large subunit with those of human μCANP and chicken CANP are 62% and 66%, respectively. These values are slightly lower than that observed between μCANP and chicken CANP (70%). Local sequence similarities vary with the domain, 73–78% in the cysteine protease like domain and 48–65% in the Ca^{2+}-binding domain. These results suggest that CANPs with different Ca^{2+} sensitivities share a common evolutionary origin and that their regulatory mechanisms are similar except for the Ca^{2+} concentrations required for activation.

Calcium-activated neutral protease (CANP), an intracellular nonlysosomal cysteine protease, mediates various cellular processes: activation and down-regulation of protein kinases and hormone receptors (Tapley & Murray, 1985; Gates & King, 1985; Melloni et al., 1986; Lynch et al., 1986; Suzuki et al., 1987), turnover of myofibrillar proteins (Goll et al., 1983; Zeman et al., 1985), modification of cytoskeletal proteins (Pontremoli et al., 1987), etc. CANP requires Ca^{2+} for expression of proteolytic activity. Two types of CANP with different Ca^{2+} requirements have been found in tissues of higher animals (Murachi, 1983). They are designated as μCANP and mCANP after their Ca^{2+} sensitivities (Imahori et al., 1983). μCANP is active at micromolar concentrations of Ca^{2+}, but mCANP requires higher concentrations (0.5–1 mM) of Ca^{2+} for activity. Both μCANP and mCANP are composed of a large and a small subunit with molecular weights of 80 000 and 28 000, respectively. The small subunit is common to the two CANPs (Kawasaki et al., 1986). Thus, the difference in the Ca^{2+} sensitivities of CANPs is attributed to the difference in the structures of the large subunits. The large subunits of the two types of CANP are similar but apparently distinguishable from each other judging from various studies such as exchange and reconstitution of subunits between μCANP and mCANP (Kikuchi et al., 1984), immunological analyses with polyclonal and monoclonal antibodies (Sasaki et al., 1983; Kasai et al., 1986), and comparison of enzymatic and chemical properties between μCANP and mCANP (Yumoto et al., 1984; Inomata et al., 1984). Comparison of the primary structures of the large subunits between μCANP and mCANP will provide clues to understand the structure–function relationship of CANP and the biological roles of the CANP isozymes.

The primary structure of the large subunit of chicken CANP was first described by Ohno et al. (1984). The sequence of the large subunit can be divided into four domains (I–IV) on the basis of sequence similarity to other proteins. The second domain from the amino terminus (domain II) is highly homologous to cysteine proteases such as papain and cathepsins B and H and bears the catalytic activity of CANP. The carboxyl-terminal domain (domain IV) comprises four putative Ca^{2+}-binding sequences, each composed of about 30 amino acid residues potentially forming helix–loop–helix (E–F hand) structures. The amino-terminal domain (domain I) is important for the regulation of activity, but the function of domain III is unknown, because it has no sequence similarity with other proteins.

We next isolated cDNA clones for the large subunits of mammalian μCANP and mCANP using the cDNA insert for chicken CANP as a probe. A nearly full-length cDNA for human μCANP (Aoki et al., 1986) and cDNA clones corresponding to carboxyl-terminal parts of rabbit μCANP and mCANP (Emori et al., 1986) have been isolated and analyzed thus far. Here we report the complete amino acid sequence of the large subunit of human mCANP deduced from the nucleotide sequence of a cDNA clone. This is the first complete structure determination of the large subunit of CANP classified as the m-type. The amino acid sequence is compared with those of other CNAPs, and the structure–function relationship is discussed.

Materials and Methods

Materials. The following materials were purchased from the sources indicated: sodium *N*-lauroylsarcosinate, guanidinium isothiocyanate, and reverse transcriptase (avian mye-

† This work was supported in part by research grants from the Ministry of Education, Science and Culture of Japan, the Ministry of Health and Welfare of Japan, the Naito Foundation, and Taisho Pharmaceutical Co.

‡ Present address: Department of Cell Chemistry, Institute of Medical Science, University of Tokyo, Shirokanedai, Minato-ku, Tokyo 108, Japan.

0006-2960/88/0427-8122$01.50/0

loblastosis virus), Wako Pure Chemicals; oligo(dT)–cellulose (type III), DNA ligase (*Escherichia coli*), and RNase H (*E. coli*), P-L Biochemicals; cesium chloride, Nakarai Chemicals; [γ-^{32}P]ATP (3000 Ci/mmol), Amersham; *Pst*I-cut and oligo(dT)-tailed pBR322 and [α-^{32}P]dCTP (3000 Ci/mmol), New England Nuclear Research Products; nuclease S1, Seikagaku Kogyo; DNA polymerase I, Boehringer; *Eco*RI methylase, New England Biolabs; Klenow fragment of *E. coli* DNA polymerase I and T4 DNA polymerase I, Takara Shuzo. Restriction enzymes were obtained from Toyobo and Takara Shuzo. A λgt10 cloning system was donated by Drs. Y. Sukenaga and Y. Shibazaki. Lysylendopeptidase was a kind gift from Dr. T. Masaki, Ibaraki University.

Construction of cDNA Libraries. Two cDNA libraries, one from skeletal muscle and the other from spleen, were constructed with poly(A)+ RNA from human tissues. Total RNA was extracted from frozen tissues with guanidinium isothiocyanate plus 0.5% (w/v) sodium *N*-lauroylsarcosinate and banded in cesium chloride (Chirgwin et al., 1979). RNA was enriched in poly(A)+ RNA by oligo(dT)–cellulose chromatography (Maniatis et al., 1982). Poly(A)+ RNA from skeletal muscle was further fractionated by sucrose density gradient centrifugation (longer than 2 kb[1]) (Suzuki et al., 1985). A cDNA library from skeletal muscle was made according to the method of Maniatis et al. (1982): single-stranded DNA synthesis with a reverse transcriptase, second strand synthesis with the Klenow fragment of *E. coli* DNA polymerase I followed by treatment with nuclease S1, poly(dC) tailing, and annealing with the poly(dG)-tailed *Pst*I site of pBR322. Transformation of *E. coli* K12 MM294 was performed according to the method of Hanahan (1983). Procedures for construction of another cDNA library from spleen, detailed by Ohno et al. (1987), were the same as described above up to the first strand synthesis except that the size fractionation of poly(A)+ RNA by sucrose density gradient centrifugation was omitted. The complementary strand was synthesized according to the method of Gubler and Hoffman (1983) with DNA polymerase I, *E. coli* DNA ligase, and *E. coli* RNase H. The double-stranded cDNA was treated with *Eco*RI methylase (37 °C, 15 min), ligated to *Eco*RI linkers after a fill-in reaction with T4 DNA polymerase I, and digested with *Eco*RI (Huynh et al., 1985). The DNAs were then fractionated in a 1% low-melting-temperature agarose gel and ligated with *Eco*RI-cut λgt10.

RNA Blot Hybridization. Poly(A)+ RNA (0.5 μg) from human skeletal muscle selected by oligo(dT)–cellulose chromatography was electrophoresed on a 1.0% agarose–formaldehyde gel (Maniatis et al., 1982). RNAs were transferred to a nitrocellulose membrane (Schleicher & Schuell). ^{32}P-Labeled fragments of λHS21 and λ31 (Aoki et al., 1986), prepared with an Amersham nick-translation kit, were used as probes. Hybridization was carried out at 42 °C for 20 h in 50 mM sodium phosphate buffer, pH 7.0, containing 50% (v/v) formamide, 1 mg/mL bovine serum albumin, 0.1% (w/v) poly(vinylpyrrolidone), 0.1% (w/v) Ficoll 400, 0.2% (w/v) SDS, 0.75 M sodium chloride, 75 mM sodium citrate, and 100 μg/mL heat-denatured salmon sperm DNA. The filter was washed with a solution containing 15 mM sodium chloride, 2 mM sodium citrate (0.1 × SSC), and 0.1% (w/v) SDS at 50 °C and then exposed to a Fuji RX film.

Preparation of mCANP. mCANP was prepared from human liver as described previously (Suzuki et al., 1979). The preparation was further chromatographed on phenyl-Sepharose CL-4B as described by Inomata et al. (1983). The purified mCANP was homogeneous as judged by SDS–polyacrylamide gel electrophoresis (sp act. 390 units/mg of protein). Two milligrams of mCANP was obtained from 400 g of human liver.

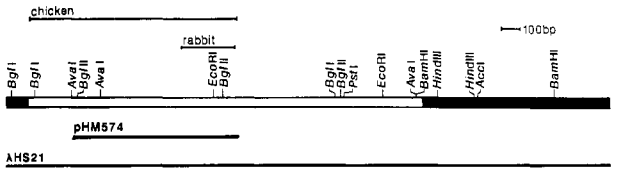

FIGURE 1: Restriction endonuclease map of cDNA clones for the large subunit of human mCANP. The open and closed bars indicate the coding and noncoding regions, respectively. Locations of the probes used are indicated above the map: chicken CANP, a *Pst*I fragment of pCP521 (1097 bp) (Ohno et al., 1984); rabbit mCANP, a *Pst*I fragment of pLM1006 (285 bp) (Emori et al., 1986).

Digestion of mCANP with Trypsin and Isolation of Tryptic Peptides. Purified mCANP (220 μg) was digested with lysylendopeptidase (1.5 μg) at 37 °C for 19 h. The digests were fractionated by high-performance liquid chromatography on a Hitachi gel, no. 3063 (C18, pore size 3 μm), with a linear gradient of 0–60% acetonitrile in 0.1% (v/v) trifluoroacetic acid.

Other Methods. Sequencing of DNA was performed according to the method of Maxam and Gilbert (1980). Automated Edman degradation of tryptic peptides was carried out with an Applied Biosystems 470A protein sequencer. PTH-amino acids released were identified by high-performance liquid chromatography on a Senshu-pak Aquasil-Seq-4 column (Senshu Kagaku).

Results and Discussion

Isolation and Identification of cDNA for the Human mCANP Large Subunit. A cDNA library from human skeletal muscle was screened by using ^{32}P-labeled cDNA inserts for CANPs thus far analyzed, chicken CANP large subunit (Ohno et al., 1984) and rabbit mCANP large subunit (Emori et al., 1986) (see Figure 1). One recombinant clone (pHM574) that hybridized with both probes under mild washing conditions (2 × SSC, 65 °C) was isolated from 20000 transformants. pHM574 was apparently distinguished from cDNA clones for the μCANP large subunit (Aoki et al., 1986), since by DNA blot analysis it did not cross-react with clones for μCANP under stringent washing conditions (0.1 × SSC, 65 °C). Furthermore, RNA blot hybridization analysis showed a band apparently distinct from the μCANP mRNA in size (see Figure 3). As pHM574 had an insert of only 900 bp, a λgt10 cDNA library from human spleen was further screened with pHM574. About 50 clones out of 400000 independent plaques were hybridized with pHM574 insert. Ten positive clones were picked up randomly, and their inserts were analyzed by electrophoresis. λHS21, containing the longest insert, was chosen for further analyses.

The restriction endonuclease map of λHS21 is shown in Figure 1. The complete nucleotide sequence of λHS21 (3213 bp) is shown in Figure 2. Polyadenylation signals (AATAAA) are seen at residues 3058–3063 and 3064–3069, but no poly(A) tract is observed in the downstream region. RNA blot hybridization analysis probed with the λHS21 insert showed a single band of 3–3.5 kb, indicating that λHS21 corresponds to a nearly full-length region of this mRNA species (Figure 3).

A long open reading frame initiating at residue −1 and encoding 700 amino acids (M_r 80005) is indicated under the

[1] Abbreviations: SDS, sodium dodecyl sulfate; SSC, standard saline citrate; bp, base pairs; kb, kilobases; PTH, phenylthiohydantoin.

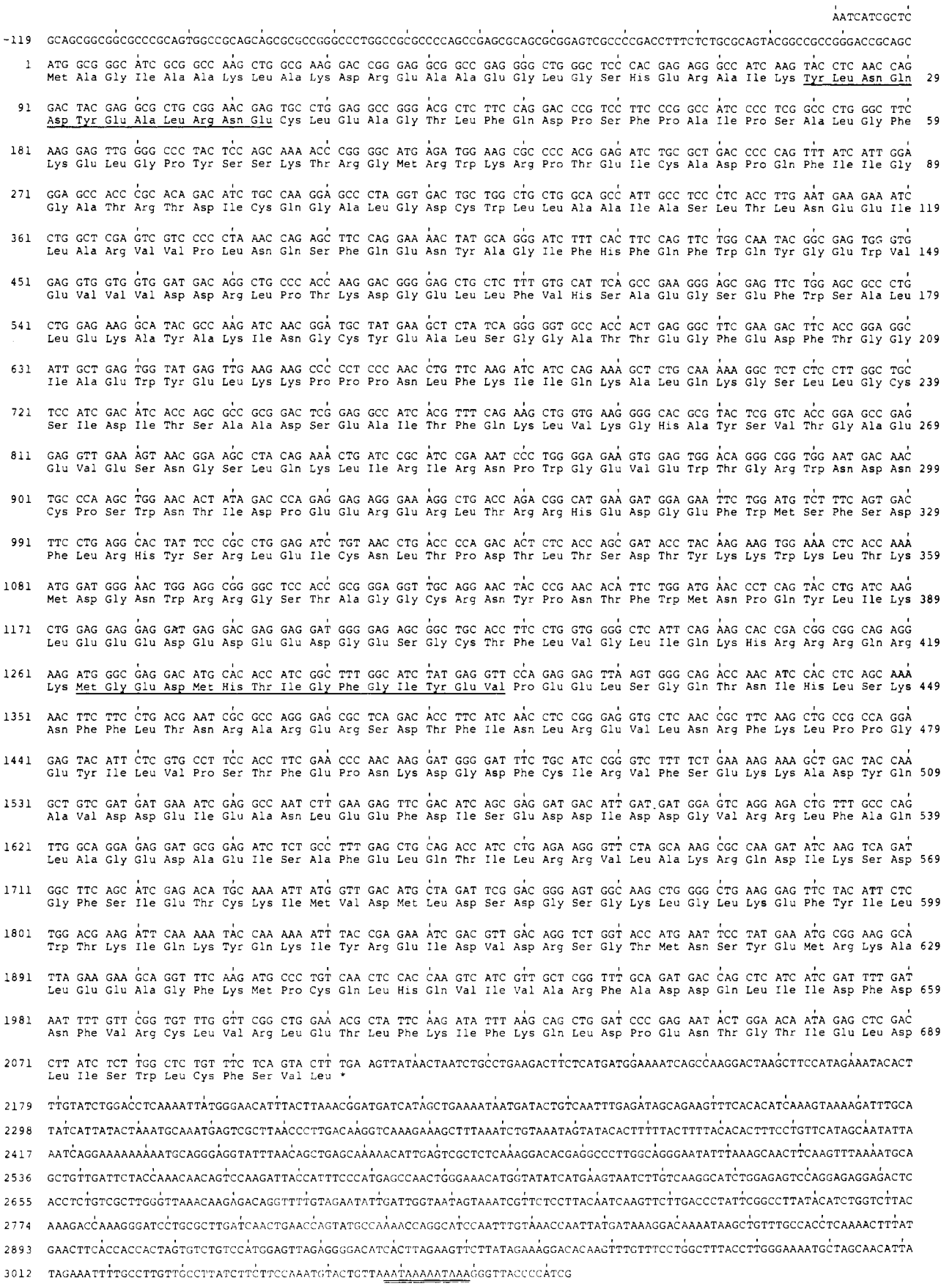

FIGURE 2: Nucleotide and deduced amino acid sequences of the cDNA for the large subunit of human mCANP. Nucleotide numbers are given on the left starting with the A of the initiating ATG, and amino acid numbers are given on the right starting with +1 for the second residue (Ala), the amino-terminal residue of the mature large subunit. Amino acid sequences underlined were also determined at the protein level (see Table I). The polyadenylation signals are double underlined.

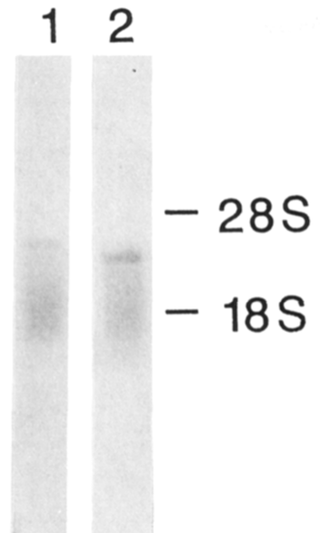

FIGURE 3: Blot hybridization of poly(A)+ RNA from human skeletal muscle. Poly(A)+ RNA (0.1 μg) from human skeletal muscle was electrophoresed on an agarose gel under denaturing conditions. RNA transferred to a nitrocellulose membrane was analyzed by hybridization with ^{32}P-labeled inserts of cDNAs for human CANPs: lane 1, *Eco*RI fragment (31–1128) of λHS21 for mCANP; lane 2, *Pst*I fragment (153–1252) of λ31 for μCANP. Positions of human ribosomal RNAs (18S and 28S) that coelectrophoresed are indicated.

Table I: Amino-Terminal Analysis of Tryptic Fragments of the Human mCANP Large Subunit

cycle no.	PTH-amino acid released/recovery (pmol)[a] peptide 1	peptide 2
1	Tyr 94	Met 362
2	Leu 55	Gly 154
3	Asn 43	Glu 152
4	Gln 33	Asp 114
5	Asp 53	Met 108
6	Tyr 43	His 13
7	Glu 19	Thr 12
8	Ala 30	Ile 49
9	Leu 22	Gly 36
10	Arg +[b]	Phe 29
11	Asn 12	Gly 31
12	Glu 7	Ile 33
13		Tyr 28
14		Glu 5
15		Val 17

[a] The released PTH-amino acid and its recovery are indicated in each sequence cycle. Recoveries of PTH-Arg, -His, and -Thr are 15–20%. [b] Not more than 4 pmol.

nucleotide sequence in Figure 2. The amino acid composition of the deduced protein agrees well with those reported for mammalian mCANP (Yumoto et al., 1984; Inomata et al., 1984). The carboxyl-terminal half (residues 278–699, see

```
                                    10        20        30        40        50        60        70    I ─┐
Human mCANP                 AGIAAKLAKDREAAEGLGSHERAIKYLNQDYEALRNECLEAGTLFQDPSFPAIPSALGFKELGPYSSKTRGMRWKR
Human μCANP       SEEIITPVYCTGVSAQVQKQRARELGLGRHENAIKYLGQDYEQLRVRCLQSGTLFRDEAFPPVPQSLGYKDLGPNSSKTYGIKWKR
Chicken CANP            MP-F-GGIAARLQRDRLRAEGVGEHNNAVKYLNQDYEALKQECIESGTLFRDPQFPAGPTALGFKELGPYSSKTRGVEWKR
                          *  *     *    * * *  * *** **** *   *    **** *  **  *  ** * *** **** *  ***

───── II
│ 80      90       100       110       120       130       140       150       160       170
PTEICADPQFIIGGATRTDICQGALGDCWLLAAIASLTLNEEILARVVPLNQSFQENYAGIFHFQFWQYGEWVEVVVDDRLPTKDGELLFVHSAEGSEFW
PTELLSNPQFIVDGATRTDICQGALGDCWLLAAIASLTLNDTLLHRVVPHGQSFQNGYAGIFHFQLWQFGEWVDVVVDDLLPIKDGKLVFVHSAEGNEFW
PSELVDDPQFIVGGATRTDICQGALGDCWLLAAIGSLTLNEELLHRVVPHGQSFQEDYAGIFHFQIWQFGEWVDVVVDDLLPTKDGELLFVHSAECTEFW
* *    ****  ********************* *****    * ****  ****  ******** ** **** ***** ** *** * ******  ***

180       190       200       210       220       230       240       250       260       270
SALLEKAYAKINGCYEALSGGATTEGFEDFTGGIAEWYELKKPPPNLFKIIQKALQKGSLLGCSIDITSAADSEAITFQKLVKGHAYSVTGAEEVESNGS
SALLEKAYAKVNGSYEALSGGSTSEGFEDFTGGVTEWYELRKAPSDLYQIILKALERGSLLGCSIDISSVLDMEAITFKKLVKGHAYSVTGAKQVNYRGQ
SALLEKAYAKLNGCYESLSGGSTTEGFEDFTGGVAEMYDLKRAPRNMGHIIRKALERGSLLGCSIDITSAFDMEAVTFKKLVKGHAYSVTAFKDVNYRGQ
********** ** ** **** * *********  * * *   *     ** ***  ********** *  * ** ** ***********    *   *

280       290       300       310  II ┐┌320  III  330       340       350       360       370
LQKLIRIRNPWGEVEWTGRWNDNCPSWNTIDPEERERLTRRHEDGEFWMSFSDFLRHYSRLEICNLTPDTLTSDTYKKWKLTKMDGNWRRGSTAGGCRNY
VVSLIRMRNPWGEVEWTGAWSDSSSEWNNVDPYERDQLRVKMEDGEFWMSFRDFMREFTRLEICNLTPDALKSRTIRKWNTTLYEGTWRRGSTAGGCRNY
QEQLIRIRNPWGQVEWTGAWSDGSSEWDNIDPSDREELQLKMEDGEFWMSFRDFMREFSRLEICNLTPDALTKDELSRWHTQVFEGTWRRGSTAGGCRNN
   *** ***** ***** * *     *    **  *  *    ********* ** *  ********** *       *      * ***********

380       390        400       410       420       430       440       450       460       470
PNTFWMNPQYLIKLEEEDE-DE-EDGESGCTFLVGLIQKHRRRQRKMGEDMHTIGFGIYEVPEELSGQTNIHLSKNFFLTNRARERSDTFINLREVLNRF
PATFWVNPQFKIRLDETDDPDDYGDRESGCSFVLALMQKHRRRERRFGRDMETIGFAVYEVPPELVGQPAVHLKRDFFLANASRARSEQFINLREVSTRF
PATFWINPQFKIKLLEEDD-D-PGDDEVACSFLVALMQKHRRRERRVGGDMHTIGFAVYEVPEEAQGSQNVHLKKDFFLRNQSRARSETFINLREVSNQI
* *** ***  * * * * *   * *  * *   * ****** *  * ** ****  **** *          **   *** *  * * **  *******

 480       490       500       510        520       530       540       550 III ─┐┌ 560 ─IV  570
KLPPGEYILVPSTFEPNKDGDFCIRVFSEKKADYQAVDDEIEANL-EEFDISEDDIDDGVRRLFAQLAGEDAEISAFELQTILRRVLAKRQDIKSDGFSI
RLPPGEYVVVPSTFEPNKEGDFVLRFFSEKSAGTVELDDQIQANLPDEQVLSEEEIDENFKALFRQLAGEDMEISVKELRTILNRIISKHKDLRTKGFSL
RLPPGEYIVVPSTFEPHKEADFILRVFTEKQSDTAELDEEISADLADEEEITEDDIEDGFKNMFQQLAGEDMEISVFELKTILNRVIARHKDLKTDGFSL
 ******  ******* *  **  * * **         *  * * *   *    *   *     * ****** ***  ** *** *     *    ****

  580       590       600       610       620       630       640       650       660       670
ETCKIMVDMLDSDGSGKLGLKEFYILWTKIQKYQKIYREIDVDRSGTMNSYEMRKALEEAGFKMPCQLHQVIVARFADDQLIIDFDNFVRCLVRLETLFK
ESCRSMVNLMDRDGNGKLGLVEFNILWNRIRNYLSIFRKFDLDKSGSMSAYEMRMAIESAGFKLNKKLYELIITRYSEPDLAVDFDNFVCCLVRLETMFR
DSCRNMVNLMDKDGSARLGLVEFQILWNKIRSWLTIFRQYDLDKSGTMSSYEMRMALESAGFKLNNKLHQVVVARYADAETGVDFDNFVCCLVKLETMFR
  *  **   * **  *** ** ***   *     * *  * * ** *  **** *  *  ****    *    *      ****** *** *** * *

  680       690
IFKQLDPENTGTIELDLISWLCFSVL
FFKTLDTDLDGVVTFDLFKWLQLTMFA
FFHSMDRDGTGTAVMNLAEWLLLTMCG
 *   *    *     *  **
```

FIGURE 4: Comparison of the primary sequences of the mature large subunits of CANPs. Amino acids are abbreviated by their one-letter codes. Residues are numbered according to human mCANP as indicated in Figure 2. The four domains (I–IV) first predicted for the chicken CANP large subunit (Ohno et al., 1984) are indicated. Gaps are introduced to maximize sequence similarity. Identical residues among all the CANPs compared are indicated by asterisks under the sequence. Four putative Ca^{2+}-binding sites are underlined by solid lines (helix regions) and dotted lines (loop regions).

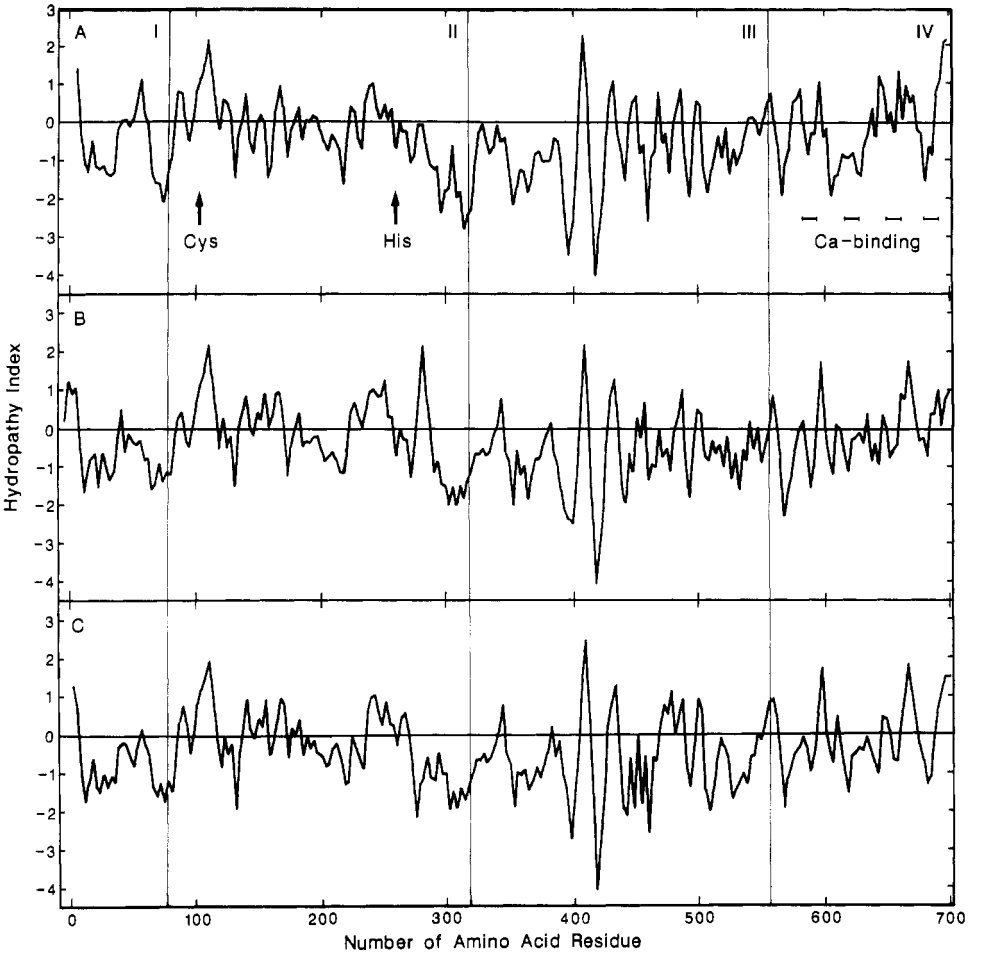

FIGURE 5: Hydropathy plot of the large subunit of CANP: (A) human mCANP, (B) human μCANP, and (C) chicken CANP. The hydropathy index was calculated according to the method of Kyte and Doolittle (1982). The upper direction denotes greater hydrophobicity. The span setting was 9 including gaps. Average hydrophobicity indices for the three CANPs are as follows: human mCANP, −0.43; human μCANP, −0.36; and chicken CANP, −0.41. Boundaries of the four domains are indicated by vertical lines. Positions of the active site of the cysteine protease (Cys and His) and the four putative Ca^{2+}-binding sites (domain IV) are indicated in panel A.

Figure 4) is highly homologous to the corresponding region of rabbit mCANP thus far identified (96% identical for translated protein sequences) (Emori et al., 1986). Furthermore, partial protein sequences of tryptic fragments of human mCANP (Table I) completely coincide with the deduced amino acid sequences (underlined parts). Thus, we concluded that λHS21 is a cDNA clone for the large subunit of human mCANP.

An in-phase termination codon is not found in the region upstream of the presumed initiation methionine (residue −1). Further, the amino terminus of the mature human mCANP large subunit could not be determined. However, we assigned the Met at −1 as the initiation site and Ala-2 as the amino terminus of the mature human mCANP large subunit for the following reasons. In the case of the chicken CANP large subunit, the initiation methionine is removed by processing and the second residue of the primary translation product becomes the amino terminus of the mature protein, though its amino group is blocked, presumably by an acetyl group (Ohno et al., 1984). Amino-terminal sequence analyses of rabbit mature mCANP and μCANP indicate that mammalian CANPs undergo similar posttranslational modifications, i.e., removal of the initiation methionine during processing (Kawasaki et al., 1987). The amino-terminal sequence of the large subunit of rabbit mCANP, Ala-Gly-Thr-Ala-Ala-Arg, corresponds to the deduced sequence of residues 1–6 of human mCANP. On the other hand, no processing occurs in the carboxyl-terminal region of CANP as demonstrated by carboxyl-terminal analysis of mature CANP (Ohno et al., 1984; Imajoh et al., 1986, 1987). Thus, we concluded that the mature form of the human mCANP large subunit corresponds to the sequence of residues 2–700 of the primary translation product and the amino terminus (Ala-2) is probably acetylated. The deduced molecular weight (79 900) matches a reported value (Suzuki et al., 1979).

Comparison of Amino Acid Sequences of the Large Subunits of CANPs. The predicted amino acid sequence of the human mCANP large subunit is aligned with those of the chicken CANP and human μCANP large subunits in Figure 4. The three sequences are different in their total length, and three and four gaps must be introduced in mCANP and chicken CANP, respectively, to maximize sequence similarity. In a previous report (Ohno et al., 1984), the large subunit of chicken CANP was divided into four domains. High sequence homology apparently indicates that the m-type large subunit of human CANP can also be divided into four similar domains: I, 1–76; II, 77–316; III, 317–555; and IV, 556–699. Furthermore, the active site cysteine residue in the protease domain (domain II, Cys-104) and four putative Ca^{2+}-binding sites in domain IV (residues 584–595, 614–625, 649–660, and 679–690) can be assigned easily. A hydropathy plot of the human mCANP large subunit is similar to those of the chicken CANP and human μCANP large subunit (Figure 5). Thus, the domain structure first defined for the chicken CANP large subunit is also applicable to human mCANP. However, the terminal regions of the three CANPs, especially the amino-

Table II: Amino Acid Sequence Similarity among the Large Subunit of Three CANPs

CANP species compared	similarity[a] (%) for domains[b]				
	I	II	III	IV	total
human mCANP/human μCANP	51	73	62	48	62
human mCANP/chicken CANP	70	75	64	53	66
human μCANP/chicken CANP	51	78	71	65	70

[a] Amino acid sequence similarities are indicated for four domains (see Figure 4). Gaps of residues inserted in domains I, III, and IV are considered as mismatches. [b] Domains are defined as indicated in Figure 4.

terminal regions, differ in length, resulting in various total residue numbers; the primary translation products of the large subunits of human mCANP, chicken CANP, and human μCANP are composed of 700, 705, and 714 residues, respectively.

The overall amino acid sequence similarities between pairs of the large subunits are 62%, 66%, and 70% for human mCANP/human μCANP, human mCANP/chicken CANP, and human μCANP/chicken CANP, respectively, assuming gaps as mismatches. The degree of sequence similarity in each domain is summarized in Table II. The catalytic domain (domain II) is the most highly conserved among the four domains, especially around the active site Cys. The high degree of sequence similarity in domain II is reflected in similar enzymatic properties of μCANP and mCANP, though they are significantly different in their Ca^{2+} sensitivities. Domain III, with unknown function, is an intervening domain between the catalytic and the Ca^{2+}-binding domains and is fairly well conserved. Several regions of 6–12 residues (319–327, 336–345, 412–417, 464–470, 476–481, and 484–490) are conserved among the three CANPs. We presume that domain III mediates the regulatory effects of the Ca^{2+}-binding domain (domain IV) to the catalytic domain (domain II), because a fragment of CANP cleaved in domain III expresses a constitutive proteolytic activity that is no longer controlled by Ca^{2+} (unpublished observation). Domain III may also be important for the interaction between the large and small subunits, and the regions conserved among species may serve as interaction sites.

Variation of amino acid sequences in domain IV is reasonable because the structure of domain IV apparently determines the Ca^{2+} sensitivity of CANP. Actually, the carboxyl-terminal parts containing domain IV of the large subunits of chicken CANP, rabbit μCANP, and mCANP, expressed in *E. coli*, bind Ca^{2+} with average dissociation constants of 60, 35, and 150 μM, respectively (Minami et al., 1987), values proportional to their Ca^{2+} requirements for activity. However, the precise relationship between Ca^{2+}-binding affinity and amino acid sequence is difficult to presume only by comparing the sequences of μ- and m-types of CANP. The relation is now being analyzed in our laboratory by utilizing site-directed mutagenesis.

Domain I is not as homologous among CANP species as domains II and III except for the case of human mCANP vs chicken CANP (70% identical; cf. Table II). Low sequence similarity in domain I is not due to the difference in residue numbers. When the regions of residues 1–76 of human mCANP, where three sequences can be aligned without gaps, are considered, numbers of identical residues are as follows: 44 for mCANP and μCANP, 43 for μCANP and chicken CANP, and 55 for mCANP and chicken CANP. The amino-terminal region of domain I is cleaved off during autocatalytic activation in the presence of Ca^{2+}, which is observed in common with chicken CANP (Ohno et al., 1984) and rabbit mCANP (Imajoh et al., 1986). Thus, the similarity in domain I between chicken CANP and mCANP may be related to a common activation mechanism. However, μCANP also undergoes a similar autocatalytic activation (unpublished observation). Further investigation is necessary to clarify the precise function of domain I.

ACKNOWLEDGMENTS

We thank Drs. Y. Sukenaga and Y. Shibazaki for kindly providing us with the λgt10 cloning system and for their helpful comments and Drs. M. Mukai, H. Sugita, M. Koike, and G. Kosaki for their support in obtaining human tissue samples.

REFERENCES

Aoki, K., Imajoh, S., Ohno, S., Emori, Y., Koike, M., Kosaki, G., & Suzuki, K. (1986) *FEBS Lett. 205*, 313–317.

Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J., & Rutter, W. J. (1979) *Biochemistry 18*, 5294–5299.

Emori, Y., Kawasaki, H., Sugihara, H., Imajoh, S., Kawashima, S., & Suzuki, K. (1986) *J. Biol. Chem. 261*, 9465–9471.

Gates, R. E., & King, L. E., Jr. (1985) *Biochemistry 24*, 5209–5215.

Goll, D. E., Otsuka, Y., Nagainis, P. A., Shannon, J. D., Sathe, S. K., & Muguruma, M. (1983) *J. Food Biochem. 7*, 137–177.

Gubler, U., & Hoffman, B. J. (1983) *Gene 25*, 263–269.

Hanahan, D. (1983) *J. Mol. Biol. 166*, 557–580.

Huynh, T., Young, R., & Davis, R. W. (1985) in *DNA Cloning. A Practical Approach* (Glover, D., Ed.) Vol. 1, pp 49–78, IRL Press, London.

Imahori, K., Suzuki, K., & Kawashima, S. (1983) in *Proteinase Inhibitors: Medical and Biochemical Aspects* (Katunuma, N., Umezawa, H., & Holzer, H., Eds.) pp 173–179, Academic, New York.

Imajoh, S., Kawasaki, H., & Suzuki, K. (1986) *J. Biochem. (Tokyo) 100*, 633–642.

Imajoh, S., Kawasaki, H., & Suzuki, K. (1987) *J. Biochem. (Tokyo) 101*, 447–452.

Inomata, M., Hayashi, M., Nakamura, M., Imahori, K., & Kawashima, S. (1983) *J. Biochem. (Tokyo) 93*, 291–294.

Inomata, M., Nomoto, M., Hayashi, M., Nakamura, M., Imahori, K., & Kawashima, S. (1984) *J. Biochem. (Tokyo) 95*, 1661–1670.

Kasai, Y., Inomata, M., Hayashi, M., Imahori, K., & Kawashima, S. (1986) *J. Biochem. (Tokyo) 100*, 183–190.

Kawasaki, H., Imajoh, S., Kawashima, S., Hayashi, H., & Suzuki, K. (1986) *J. Biochem. (Tokyo) 99*, 1525–1532.

Kawasaki, H., Imajoh, S., & Suzuki, K. (1987) *J. Biochem. (Tokyo) 102*, 393–400.

Kikuchi, T., Yumoto, N., Sasaki, T., & Murachi, T. (1984) *Arch. Biochem. Biophys. 234*, 639–645.

Kyte, J., & Doolittle, R. F. (1982) *J. Mol. Biol. 157*, 105–132.

Lynch, C. J., Sobo, G. E., & Exton, J. H. (1986) *Biochim. Biophys. Acta 885*, 110–120.

Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) in *Molecular Cloning. A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Maxam, A. M., & Gilbert, W. (1980) *Methods Enzymol. 65*, 499–560.

Melloni, E., Pontremoli, S., Michetti, M., Sacco, O., Sparatore, B., & Horecker, B. L. (1986) *J. Biol. Chem. 261*, 4101–4105.

Minami, Y., Emori, Y., Kawasaki, H., & Suzuki, K. (1987) *J. Biochem. (Tokyo) 101*, 889–895.

Murachi, T. (1983) in *Calcium and Cell Function* (Cheung, W. Y., Ed.) Vol. IV, pp 377–410, Academic, New York.
Ohno, S., Emori, Y., Imajoh, S., Kawasaki, H., Kisaragi, M., & Suzuki, K. (1984) *Nature (London) 312*, 566–570.
Ohno, S., Emori, Y., Sugihara, H., Imajoh, S., & Suzuki, K. (1987) *Methods Enzymol. 139*, 363–379.
Pontremoli, S., Melloni, E., Michetti, M., Sparatore, B., Salamino, F., Sacco, O., & Horecker, B. L. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 398–401.
Sasaki, T., Yoshimura, N., Kikuchi, T., Hatanaka, M., Kitahara, A., Sakihama, T., & Murachi, T. (1983) *J. Biochem. (Tokyo) 94*, 2055–2061.
Suzuki, K., Ishiura, S., Tsuji, S., Katamoto, T., Sugita, H., & Imahori, K. (1979) *FEBS Lett. 104*, 355–358.
Suzuki, K., Ohno, S., Imajoh, S., Emori, Y., & Kawasaki, H. (1985) *Biomed. Res. 6*, 323–327.
Suzuki, K., Imajoh, S., Emori, S., Kawasaki, H., Minami, Y., & Ohno, S. (1987) *FEBS Lett. 220*, 271–277.
Tapley, P. M., & Murray, A. W. (1985) *Eur. J. Biochem. 151*, 419–423.
Yumoto, N., Kikuchi, T., Sasaki, T., & Murachi, T. (1984) *J. Biochem. (Tokyo) 96*, 1531–1537.
Zeman, R. J., Kameyama, T., Matsumoto, K., Bernstein, P., & Etlinger, J. D. (1985) *J. Biol. Chem. 260*, 13619–13624.

Interaction of Chloroquine with Linear and Supercoiled DNAs. Effect on the Torsional Dynamics, Rigidity, and Twist Energy Parameter†

Pengguang Wu, Lu Song, James B. Clendenning, Bryant S. Fujimoto, A. S. Benight, and J. Michael Schurr*
Department of Chemistry, BG-10, University of Washington, Seattle, Washington 98195
Received March 25, 1988; Revised Manuscript Received June 3, 1988

ABSTRACT: The magnitude and uniformity of the torsion elastic constant (α) of linear pBR322 DNA and supercoiled pBR322 DNAs with high-twist ($\sigma = -0.083$) and normal-twist ($\sigma = -0.48$) are measured in 0.1 M NaCl as a function of added chloroquine/base-pair ratio (chl/bp) by studying the fluorescence polarization anisotropy (FPA) of intercalated ethidium dye. The time-resolved FPA is measured by using a picosecond dye laser for excitation and time-correlated single-photon counting detection. A general theory is developed for the binding of ligands that unwind superhelical DNAs, and the simultaneous binding of two different intercalators is treated in detail. The equilibrium constant (K) for binding chloroquine to linear pBR322 DNA and the number (r) of bound chloroquines per base pair are determined from the relative amplitude ratio of the slow (normally intercalated) and fast (free) components in the decay of the (probe) ethidium fluorescence intensity as a function of chl/bp. For chloroquine binding to supercoiled pBR322 DNAs, the intrinsic binding constant is assumed to be the same as for the linear DNA, but the twist energy parameter E_T (N times the free energy to change the linking number from 0 to 1 in units of k_BT) is regarded as adjustable. Using the best-fit E_T, the binding ratios r are calculated for each chl/bp ratio. Twist energy parameters are also determined for ethidium binding to these supercoiled DNAs by competitive dialysis. For chloroquine binding, we obtain $E_T = 360$ and 460 respectively for the normal-twist and high-twist supercoiled DNAs. For ethidium binding the corresponding values are $E_T = 280 \pm 70$ and 347 ± 50. Like other dye-binding values, these are substantially lower than those obtained by ligation methods. In the absence of chloroquine, the torsion constants of all three DNAs are virtually identical, $\alpha = (5.0 \pm 0.4) \times 10^{-12}$ dyn·cm. For linear pBR322 DNA, the magnitude and uniformity of α remain unaltered by intercalated chloroquine up to $r = 0.19$. This finding argues that the FPA is not significantly relaxed by diffusion of any kinks or solitons. If α_d denotes the torsion constant between a dye and a base pair and α_0 that between two base pairs, then our data imply that α_d/α_0 lies in the range 0.65–1.64, with a most probable value of 1.0. For the supercoiled DNAs, α remains uniform, but its magnitude decreases significantly with increasing intercalated chloroquine through the point where the effective superhelix density vanishes and even up to substantial positive superhelix densities. The significant difference in observed FPA dynamics between the relaxed (by intercalator) supercoiled DNAs and the linear DNA with the same amount of bound chloroquine profoundly contradicts any notion that the local structures and dynamics of such species are equivalent. At very high chl/bp ratios both linear and supercoiled DNAs exhibit evidence of substantial structural changes. The torsion constant increases for the linear DNA but drops for the supercoiled DNAs, which are substantially positively twisted ($\sigma \geq 0.06$) at that point.

Time-resolved fluorescence polarization anisotropy (FPA) measurements on intercalated ethidium dye have been employed to study the rotational dynamics of various DNAs from 10^{-10} to 1.5×10^{-7} s (Thomas et al., 1980a; Wilcoxon et al., 1982; Millar et al., 1982; Madge et al., 1983; Ashikawa et al., 1983, 1984; Thomas & Schurr, 1983; Shibata et al., 1984, 1985; Fujimoto et al., 1985). Previous work in this laboratory (Thomas et al., 1980a; Shibata et al., 1985) demonstrated that the FPA from 1 to 120 ns follows closely the theoretical predictions for a filament with uniform torsional rigidity (Barkley & Zimm, 1979; Allison & Schurr, 1979; Schurr, 1984). For linear DNAs, it was also shown that the apparent torsion constant (α) between base pairs is essentially unaffected by temperature (T) from 0 to 70 °C (Thomas & Schurr, 1983; Wilcoxon & Schurr, 1983), GC content from 34 to 100% (Fujimoto et al., 1985), and spermidine binding in 10 mM

† This work was supported in part by Grants DMB-8403118 from the National Science Foundation (linear DNAs) and by Grant R01-GM29338 from the National Institutes of Health (supercoiled DNAs).

0006-2960/88/0427-8128$01.50/0