Accelerated Publications

A New Form of Guanylyl Cyclase Is Preferentially Expressed in Rat Kidney[†,‡]

Peter S. T. Yuen,[§] Lincoln R. Potter,[∥] and David L. Garbers*[,§,∥]

Howard Hughes Medical Institute and Departments of Pharmacology and of Molecular Physiology and Biophysics, Vanderbilt University, Nashville, Tennessee 37232

Received September 14, 1990; Revised Manuscript Received October 15, 1990

ABSTRACT: On the basis of the conserved amino acid sequences of the catalytic domain of both soluble and plasma membrane forms of guanylyl cyclase, we have used the polymerase chain reaction to identify a new form of guanylyl cyclase that is expressed principally in kidney. The cDNA for this new form (GC-$S_{\beta 2}$) codes for a 76.3-kDa protein, which most closely resembles a 70-kDa subunit (GC-$S_{\beta 1}$) of the lung soluble guanylyl cyclase. The mRNA for GC-$S_{\beta 1}$ is preferentially expressed in lung and brain, whereas GC-$S_{\beta 2}$ mRNA is more abundant in kidney and liver. An 86 amino acid carboxyl-terminal region extends beyond the C-terminus of GC-$S_{\beta 1}$ and contains a consensus sequence (-C-V-V-L) for isoprenylation/carboxymethylation. This is the first demonstration of heterogeneity among the heterodimeric forms of guanylyl cyclase and suggests differential regulation.

Guanylyl cyclase activity, unlike that normally seen with adenylyl cyclase, is generally found in both the particulate and soluble fractions of tissue homogenates (Kimura & Murad, 1974; Chrisman et al., 1975). Various early studies suggested that particulate and soluble enzyme activities were due to different proteins as opposed to translocation of the same enzyme between cellular compartments (Kimura & Murad, 1974; Chrisman et al., 1975). Eventually, it also became clear that the primary regulatory molecules were different, in that nitric oxide, nitrosamines, and other nitrovasodilators stimulated the soluble form whereas small peptides such as atrial natriuretic peptide stimulated a particulate from (Garbers, 1989).

The mRNA encoding a plasma membrane form of guanylyl cyclase was first cloned from sea urchin testis (Singh et al., 1988), which then led to the cloning of a mRNA for a mammalian plasma membrane form of the enzyme. An expressed mammalian guanylyl cyclase was demonstrated to serve as a receptor for atrial natriuretic peptide (Chinkers et al., 1989; Lowe et al., 1989). Later, the plasma membrane guanylyl cyclase family was shown to contain multiple members whose general topology was similar, but whose primary amino acid sequence varied considerably within the extracellular, putative ligand binding region (Chang et al., 1989; Schulz et al., 1989). The plasma membrane forms appear to be imprinted with a single transmembrane domain that is followed on the carboxyl side by a protein kinase like domain that appears important for proper signaling (Chinkers & Garbers, 1989). Distal to the protein kinase domain is the cyclase catalytic region (Chinkers & Garbers, 1989; Thorpe & Morkin, 1990), which is highly conserved among guanylyl cyclases, and to a lesser extent among adenylyl cyclases (Krupinski et al., 1989).

The soluble form of guanylyl cyclase, unlike the plasma membrane forms, appears to exist as a heterodimer[1] (Kamisaki et al., 1986); the mRNA encoding both the α- (Koesling et al., 1990; Nakane et al., 1990) and β-subunits (Koesling et al., 1988; Nakane et al., 1988) have been recently cloned from bovine and rat lung. Diversity within this part of the guanylyl cyclase family has not been suggested, and general models depicting soluble enzyme regulation have not considered possible multiple forms.

The polymerase chain reaction (Mullis & Faloona, 1987) has been successfully used to clone partial-length cDNAs by use of degenerate oligodeoxynucleotide primers that correspond to known amino acid sequences (Lee et al., 1988; Knoth et al., 1988; Gautam et al., 1989). This strategy has been extended to identify new members of a multigene family that share conserved amino acid sequences with known members of this family: for example, tyrosine kinases and G-protein-coupled receptors (Wilks et al., 1989; Libert et al., 1989). We designed degenerate oligonucleotide primers on the basis of highly conserved amino acid sequences within the putative catalytic domain that are distinct from the corresponding adenylyl cyclase sequences (Krupinski et al., 1989) to specifically amplify partial-length guanylyl cyclase cDNAs from rat tissues. By use of this technology we demonstrate the existence of a unique guanylyl cyclase β-subunit that is preferentially expressed in kidney. This is the first evidence of diversity among heterodimeric forms of guanylyl cyclase. General models of guanylyl cyclase regulation, therefore, must now account for potential differences in these forms across cell types.

Materials and Methods

Amplification of Guanylyl Cyclase cDNA by PCR. All known guanylyl cyclase amino acid sequences were compared

[†] This work was partially supported by NIH Grant 5T32DK 07563.

[‡] The nucleic acid sequence in this paper has been submitted to GenBank under Accession Number J05308.

* To whom correspondence should be addressed at his current address: Howard Hughes Medical Institute and Department of Pharmacology, University of Texas Southwestern Medical Center, Dallas, TX 75235.

[§] Howard Hughes Medical Institute and Department of Pharmacology.

[∥] Department of Molecular Physiology and Biophysics.

[1] The identification of multiple forms of guanylyl cyclase creates the need for consistent nomenclature. We propose that the heterodimeric forms of guanylyl cyclase be named GC-S. To reflect the subunit structure, the enzyme characterized from lung is designated GC-$S_{\alpha 1\beta 1}$, where $\alpha 1$ refers to the 82- or 73-kDa subunit and $\beta 1$ refers to the 70-kDa subunit.

0006-2960/90/0429-10872$02.50/0

ANTISENSE DEGENERATE PRIMER

```
MPRYCLFG      rat/human brain membrane A and B guanylyl cyclases
MPRYCLFG      rat/bovine lung soluble guanylyl cyclase β subunit
MPRYCLFG      S. purpuratus membrane guanylyl cyclase

KWQYDVWS      bovine brain adenylyl cyclase (415-422)
RPQYDIWG      bovine brain adenylyl cyclase (996-1003)

                    G   F   L   C   Y   R   P   M
          Xba I Bam HI
5'-GAATTCTCGAGGAT CC.G/AAA.IA G/A G/A CA.G/A TA.ICI.IGG.CAT-3'
```

SENSE DEGENERATE PRIMER

```
VYKVETIGDAYM  rat/human brain particulate A and B guanylyl cyclases
VYKVETVGDKYM  rat/bovine lung soluble guanylyl cyclase β subunit
VYKVETIGDAYM  S. purpuratus membrane guanylyl cyclase

CRRIKILGDCYY  bovine brain adenylyl cyclase (346-357)
LEKIKTIGSTYM  bovine brain adenylyl cyclase (916-927)

                      V   Y   K   V   E   T  V/I  G   D  A/K  Y   M
          Kpn I
     Sma I
5'-GAATTCCCGGGTACC GTI.TAT/C.AAG/A.GTI.GAG/A.ACI.ITI.GGI.GAT/C.III.TAT/C.ATG-3'
```

FIGURE 1: Degenerate oligodeoxynucleotide primers used to amplify partial-length guanylyl cyclase cDNA. Highly conserved amino acid sequences from guanylyl cyclases were aligned and compared to the corresponding sequences from two conserved regions in bovine brain adenylyl cyclase. Sense and antisense degenerate primers were synthesized as shown; deoxyinosine was used for 3- and 4-fold degenerate codon positions. Restriction sites for subcloning the PCR product were also introduced at the 5′ ends.

to identify highly conserved regions. Two peptide sequences in the putative catalytic domain (Chinkers & Garbers, 1989) are almost completely conserved among guanylyl cyclases but are distinguishable from corresponding sequences of adenylyl cyclase (Krupinski et al., 1989). Degenerate primers were designed to specifically amplify any guanylyl cyclase containing such conserved sequences (Figure 1).

RNA (75 μg) prepared from rat brain, heart, kidney, or lung (Cathala et al., 1983) was used to synthesize cDNA by incubation at 37 °C with murine moloney leukemia virus reverse transcriptase (2000 units, BRL) and 4 μg of oligo(dT)$_{12-18}$ primer (Pharmacia). Five percent of this cDNA product was used as a template for PCR[2] in the presence of 10 mM Tris-HCl, pH 8.3, 1.5 mM $MgCl_2$, 50 mM KCl, 0.01% gelatin (w/v), and 0.2 μM of each degenerate primer (sense and antisense, Midland Certified Reagent Co.). This mixture was heated to 95 °C for 7 min, followed by 5 min at 72 °C. At this time 2.5 units of AmpliTaq (Perkin-Elmer/Cetus) was added, and the PCR proceeded for 30 cycles of 94 °C (1 min), 40 °C (2 min), and 72 °C (3 min). A final 7-min extension step (72 °C) was performed before termination of the reaction. The PCR products were digested with *Kpn*I and *Bam*HI and directionally cloned into M13 mp18. Clones were sequenced (Sanger et al., 1977) with Sequenase (USB).

Isolation of Guanylyl Cyclase cDNA. An oligo(dT)-primed, size-selected (>1.6 kb) cDNA library was prepared from rat kidney poly(A+) RNA (Chirgwin et al., 1979; Aviv & Leder, 1972) by use of the λ-Zap II vector (Stratagene) and hemiphosphorylated *Not*I/*Eco*RI adaptors (Clontech). Approximately 600 000 recombinants were screened with the kidney-specific PCR product that was radiolabeled ([α-^{32}P]dCTP) by random-primed DNA synthesis. Duplicate lifts with nylon filters (Hybond-N, Amersham) were denatured with NaOH, neutralized, and fixed by being baked for 2 h at 80 °C. Filters were prehybridized in 5× SSC, 40% formamide (v/v), 2.25 mM sodium pyrophosphate, 50 mM sodium phosphate, and 5× Denhardt's solution for 2 h at 42 °C, followed by addition of 5% dextran sulfate (w/v), radiolabeled probe, and 100 μg/mL salmon sperm DNA, and hybridized overnight at 42 °C. Filters were washed three times at 70 °C in 0.1× SSC, 0.2% SDS (w/v), and 2.25 mM sodium pyrophosphate for 20 min. Fifteen clones were rescued into Bluescript II with R408 helper phage (Stratagene) and isolated; the largest were 2.14 (clone 7) and 2.35 kb (clone 15). The two clones were digested with *Eco*RI or *Pst*I/*Bam*HI and subcloned into M13 mp18 and mp19. Internal DNA sequence of both strands was obtained by shotgun cloning with *Sau*3AI or specific 17-mer synthetic primers.

Northern Blot Analysis. A total of 10 μg of poly(A+) RNA was electrophoresed on a 1% agarose (w/v) gel containing 2.2 M formaldehyde and blotted onto an Optibind membrane (Schleicher & Schuell) with 20× SSC. RNA was fixed by being baked for 2 h in an 80 °C vacuum oven. Random-primed, radiolabeled probes were synthesized with either the full-length cDNA from clone 15 or rat lung GC-$S_{\beta 1}$ (Nakane et al., 1988). The membrane was prehybridized at 42 °C in 5× SSPE, 5× Denhardt's solution, 50% formamide, 0.25% SDS, and 100 μg/mL salmon sperm DNA for 30 min and hybridized overnight with the addition of 10% dextran sulfate and clone 15 radiolabeled probe. The membrane was washed twice with 2× SSPE and 0.1% SDS for 10 min and twice with 0.1× SSPE and 0.1% SDS for 15 min at 20 °C. After autoradiography the membrane was rehybridized with a probe

[2] Abbreviations: Denhardt's solution, 0.2% poly(vinylpyrrolidone), 0.2% Ficoll, and 0.2% bovine serum albumin; G_i, guanine nucleotide binding protein that inhibits adenylyl cyclase; G_o, guanine nucleotide binding protein of unknown function; NMDA, *N*-methyl-D-aspartate; PCR, polymerase chain reaction; SSC, 150 mM NaCl and 15 mM sodium citrate, pH 7.0; SSPE, 150 mM NaCl, 1 mM EDTA, and 10 mM sodium phosphate, pH 7.4.

specific for GC-$S_{\beta 1}$ and treated as described above, except that washes with 0.1× SSPE and 0.1% SDS were performed at 65 °C.

Results

The PCR products from oligo(dT)-primed cDNA from rat brain, heart, kidney, and lung were approximately 260–270 bp (not shown), which corresponded to the predicted sizes for the plasma membrane guanylyl cyclases, GC-A (Chinkers et al., 1989) and GC-B (Schulz et al., 1989), and the low molecular weight soluble enzyme subunit, GC-$S_{\beta 1}$ (Nakane et al., 1988). The PCR product from kidney yielded three identical clones containing a sequence distinct from GC-A (15 clones), GC-B (5 clones), and GC-$S_{\beta 1}$ (4 clones). The unique PCR product clone was then used as a radiolabeled probe to isolate a full-length cDNA from a library of rat kidney cDNA.

The complete nucleotide sequence of a subsequently isolated clone (clone 15; 2.35 kb) contained an open reading frame of 2046 bp, beginning with a potential consensus initiation site at nucleotide position 147. The G at position 150 met one of the two requirements for optimal initiation of translation, as outlined by Kozak (1989); however, the T at position 144 did not agree with the consensus purine (Figure 2). A termination codon at position 15 is in the same reading frame. If the predicted initiation site is correct, there are 2046 nucleotides encoding a protein containing 682 amino acids, which corresponds to a molecular mass of 76.3 kDa (Figure 2).

The amino acid sequence of the kidney-derived guanylyl cyclase (GC-$S_{\beta 2}$) is homologous to that of GC-$S_{\beta 1}$, the subunit obtained from rat or bovine lung (Figure 3). The first 62 amino acids of GC-$S_{\beta 1}$ do not overlap with GC-$S_{\beta 2}$, and the final 86 amino acids of GC-$S_{\beta 2}$ do not overlap with GC-$S_{\beta 1}$. The intervening sequences are 34% identical, which can be divided into four regions. Region I (4–91 for GC-$S_{\beta 2}$ and 65–153 for GC-$S_{\beta 1}$) is 54% identical, whereas region II (92–208 for GC-$S_{\beta 2}$ and 155–211 for GC-$S_{\beta 1}$) contains almost no identity; there are also two gaps within the GC-$S_{\beta 1}$ sequence. Region III (209–372 for GC-$S_{\beta 2}$ and 212–384 for GC-$S_{\beta 1}$) is 40% identical. Region IV (373–588 for GC-$S_{\beta 2}$ and 386–612 for GC-$S_{\beta 1}$) is highly conserved among guanylyl cyclases and is a part of the catalytic domain (Chinkers & Garbers, 1989; Thorpe & Morkin, 1990). In this region GC-$S_{\beta 2}$ resembles GC-A (52% identity) more closely than GC-$S_{\beta 1}$ (43% identity). In contrast, over the entire overlapping sequences GC-$S_{\beta 2}$ and GC-$S_{\alpha 1}$ are 27% identical.

The final 86 amino acids in GC-$S_{\beta 2}$ extend beyond the carboxyl termini of homologous guanylyl cyclase catalytic domains. A search of GenBank did not reveal substantial matches between this region and any coding sequence. This region also contains the sequence -C-V-V-L at the carboxyl terminus, which meets the consensus sequence requirement -C-A-A-X (where A represents an aliphatic amino acid and X is any amino acid) for isoprenylation/carboxymethylation (Clarke et al., 1988).

The relative distributions of GC-$S_{\beta 2}$ and GC-$S_{\beta 1}$ mRNA, as determined by Northern blot analysis, further distinguish these forms of guanylyl cyclase. The 2.5-kb mRNA corresponding to GC-$S_{\beta 2}$ was prominent in kidney, with lower amounts detectable in liver. When GC-$S_{\beta 1}$ was used as a probe, however, a 3.3-kb band was seen mainly in lung and brain (Figure 4), confirming the findings of Nakane et al. (1988). These investigators also detected low amounts of GC-$S_{\beta 1}$-hybridizing mRNA in kidney, heart, liver, and muscle, which may not be detectable on our autoradiograph. In addition to the 3.3-kb band, a band at 4.3 kb is present in lung (Figure 4); this has not been described previously. It is possible that the 4.3-kb band represents an α-subunit, but this band is not seen in brain, and under the stringent conditions of hybridization used, the α- and β-subunits would not be similar enough to favor cross-hybridization.

Discussion

The presence of at least two different guanylyl cyclase subunits in both kidney and liver raises questions about regulatory mechanisms. It has been previously assumed that because GC-$S_{\alpha 1 \beta 1}$ is found in most tissues (Lewicki et al., 1983; Nakane et al., 1988, 1990), its regulation by nitric oxide and other related oxidants is universal. This assumption is apparently valid in the case of endothelium-dependent vasorelaxation, which is a model for the action of nitric oxide (Rapoport & Murad, 1983; Waldman & Murad, 1987). Several vasodilators, including acetylcholine, substance P, and bradykinin, require the presence of vascular endothelium. These agents stimulate endothelial cells to produce a diffusible substance known as endothelium-derived relaxing factor, or EDRF, that then acts directly on vascular smooth muscle cells (Furchgott, 1984). The structure of EDRF is not entirely clear, but it is believed that EDRF is either nitric oxide or a derivative (such as *S*-nitrosocysteine) of this highly reactive molecule (Palmer et al., 1987; Ignarro et al., 1987; Tracey et al., 1990; Myers et al., 1990). Non-EDRF roles for nitric oxide also have been described in macrophages, where nitric oxide appears to be a cytotoxic agent (Marletta, 1989), and in cerebellum, where nitric oxide may mediate the NMDA-induced increases in cGMP content (Garthwaite et al., 1988, 1989; Bredt & Snyder, 1989). The presence of GC-$S_{\alpha 1 \beta 1}$ in lung, kidney, brain, heart, liver, and testis (Lewicki et al., 1983) may be associated with the vasculature that is extensive in all of these tissues.

In addition to nitric oxide several other endogenous agents are capable of stimulating soluble guanylyl cyclase. Among these agents are fatty acid hydroperoxides (Hidaka & Asano, 1977; Graff et al., 1978), prostaglandin endoperoxides (Graff et al., 1978), and dehydroascorbate (Haddox et al., 1978). The multiplicity of different agents capable of activating soluble forms of guanylyl cyclase coupled with the newly established diversity within the soluble enzyme family suggests that general models of soluble enzyme regulation are still premature.

Unlike the other guanylyl cyclases characterized, GC-$S_{\beta 2}$ described here contains a carboxyl-terminal region (86 additional amino acids when compared to GC-$S_{\beta 1}$) that terminates with the four amino acids -C-V-V-L (Figure 2). This sequence corresponds to the consensus sequence -C-A-A-X (where A represents an aliphatic amino acid and X is any amino acid) required for the posttranslational modification of members of the ras family (Clarke et al., 1988; Gutierrez et al., 1989; Hancock et al., 1989; Gibbs et al., 1989). This consensus sequence is also found at the C-terminus of the guanine nucleotide binding proteins transducin, G_i, and G_o (Clarke et al., 1988). There are at least three steps involved in the processing of ras: proteolysis of the three terminal amino acids, isoprenylation of the newly exposed terminal cysteine, and carboxylmethylation of this cysteine (Gutierrez et al., 1989; Hancock et al., 1989). The distribution of ras to the plasma membrane is required for its function, and this translocation is dependent on the modifications of these four amino acids (Willumsen et al., 1984; Gibbs et al., 1989). However, addition of these four amino acids to unrelated proteins, although creating a site for these modifications, does not necessarily result in translocation of the protein to the membrane. Another cysteine near the C-terminus of ras is acylated by palmitate, which increases the affinity of ras for

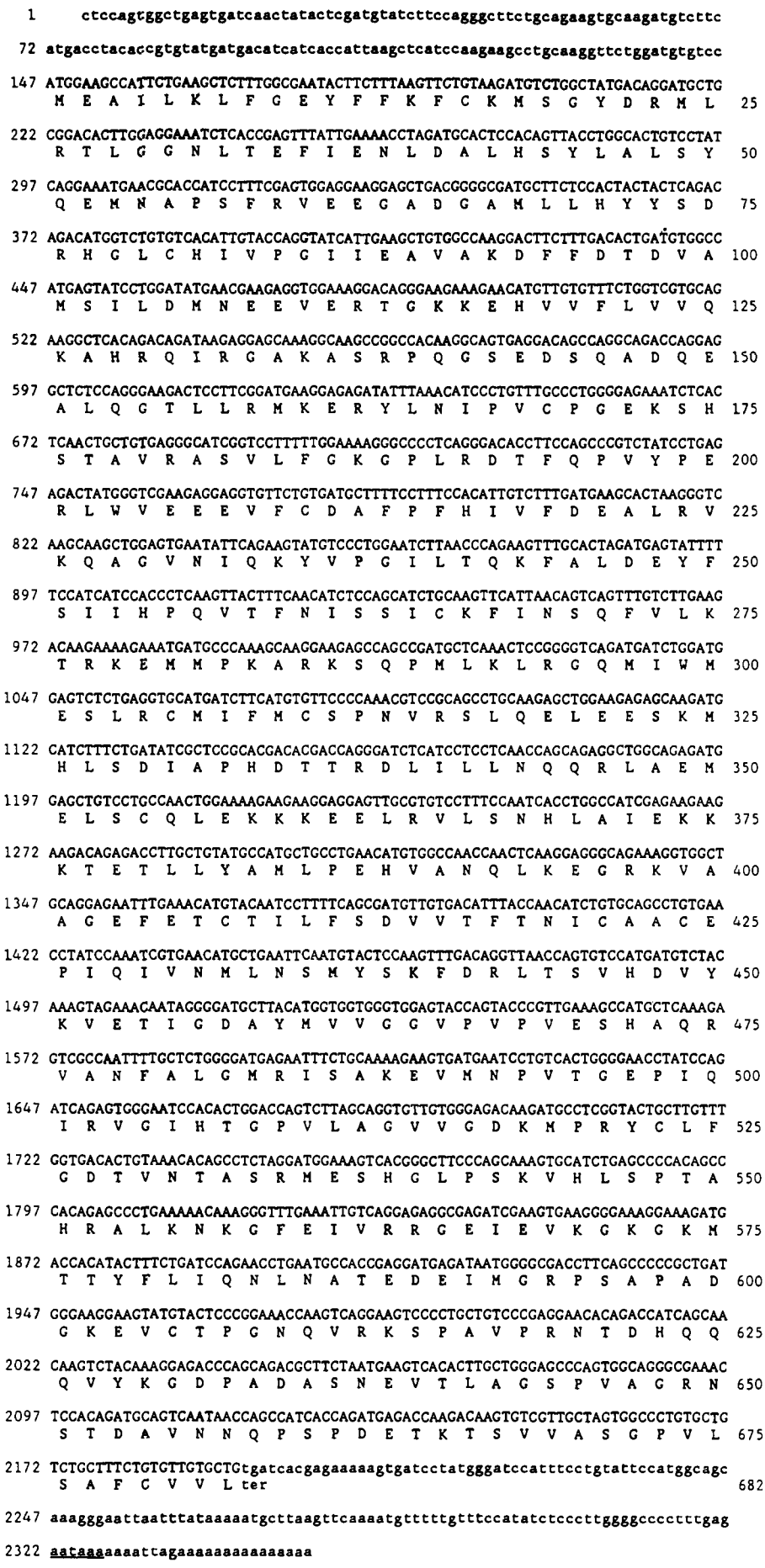

FIGURE 2: Nucleotide and predicted amino acid sequences of kidney-derived guanylyl cyclase. Positions for nucleotides are provided on the left, and positions for amino acids are on the right. Noncoding nucleotides are in lower-case letters, and the signal for polyadenylation is underlined.

```
1           MFCAKLKDLQITGDCPFSLLAPGQVPREPLGEATGSGPA  ALPHA1

1                                   MYGFVNHALELLVIRNYG  BETA1
40      STPGQPGVCPGVPDKNPPGRLPRRKTSRSRVYLHTLAESICKLIF  ALPHA1

18      PEVWEDIKKEAQLDEEGQFLV-RIIYDDSKTYDLVAAASKVLNLN  BETA1
85      PEFERLNLALQRTLAKHKIKENRKSLEREDFEKIVVDQAIAAGVP  ALPHA1

1       MEAILKLFGEYFFKFCKMSGYDRMLRTLGGNLTEFIE---NLDAL  BETA2
62      AGEILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQ---NLDAL  BETA1
129     VEIIKESLGEELFKICYEEDEY-ILGVVGGTLKDFLNSFSTLLKQ  ALPHA1

43      HSYLALSYQEMNAPSFRVEEGADG-AMLLHYYSDRHGL-CHIVPG  BETA2
104     HDHLATIYPGMRAPSFRCTDADKGKGLILHYYSEREGL-QDIVIG  BETA1
173     SSHCQEAEKKGRFEDASILCLDKDPDVLYVYYFFPKRITSLILPG  ALPHA1

86      IIEAVAKDFFDTDVAMSILDMNEEVERTGKKEHVVFLVVQKAHRQ  BETA2
148     IIKTVAQQIHGTEIDMKVIQ-------------------------  BETA1
218     IIK-----------------------------------------  ALPHA1

131     IRGAKASRPQGSEDSQADQEALQGTLLRMKERYLNIPVCPGEKSH  BETA2
168     ---------QRNEECDHTQFLIEEKESKEEDFYEDLDRFEENGTQ  BETA1
222     -------AAARILYETEVEVSSTPSRFHQDCREFVDQPCELYSVH  ALPHA1

176     STAVRASVLFGKGPLRDTFQPVYPERLWVEEEVFCDAFPFHIVFD  BETA2
204     ESRISPYT-------------------------FCKAFPFHIIFD  BETA1
260     IRSARPHPPPGKPVSSL----VIPASL------FCKTFPFHFMLD  ALPHA1

221     EALRVKQAGVNIQKYVPG-ILTQKFALDEYFSIIHPQVTFNISSI  BETA2
224     RDLVVTQCGNAIYRVLPQ-LQPGNCSLLSVFSLVRPHIDISFHGI  BETA1
295     RDMSILQLGHGIRRLMSRRDVQGKPHFDEYFEILTPKISQTFSGI  ALPHA1

265     CKFINSQFVLKTRKEMM---------PKARKSQPMLKLRGQMIWM  BETA2
268     LSHINTVFVLRSKEGLLDVEKSECEDELTGTEISCLRLKGQMIYL  BETA1
340     MTMLNMQFLVRVRR----------WDNSMKKSSRVMDLKGQMIYM  ALPHA1

301     ESLRCMIFMCSPNVRSLQELEESKMHLSDIAPHDTTRDLILLNQQ  BETA2
313     PEADSILFLCSPSVMNLDDLTRRGLYLSDIPLHDATRDLVLLGEQ  BETA1
380     VESSSILFLGSPCVDRLEDFTGRGLYLSDIPIHNALRDVVLIGEQ  ALPHA1

346     RLAEMELSCQLEKKKEELRVLSNHLAIEKKKTETLLYAMLPEHVA  BETA2
358     FREEYKLTQELEILTDRLQLTLRALEDEKKKTDTLLYSVLPPSVA  BETA1
425     ARAQDGLKKRLGKLKATLEQAHQALEEEKRKTVDLLCSIFPSEVA  ALPHA1

391     NQLKEGRKVAAGEFETCTILFSDVVTFTNIC----AACEPIQIVN  BETA2
403     NELRHKRPVPAKRYDNVTILFSGIVGFNAFCSKHASGEGAMKIVN  BETA1
470     RQLWQGHAVQAKRFGNVTMLFSDIVGFTAICS----QCSPLQVIT  ALPHA1

432     MLNSMYSKFDRLTSVH---DVYKVETIGDAYMVVGGVPVPVESHA  BETA2
448     LLNDLYTRFDTLTDSRKNPFVYKVETVGDKYMTVSGLPEPCIHHA  BETA1
506     MLNALYTRFDRQCGEL---DVYKVETIGDAYCVAGGLHKESDTHA  ALPHA1

474     QRVANFALGMRISAKEVMNPVTGEPIQIRVGIHTGPVLAGVVGDK  BETA2
493     RSICHLALDMMEIAGQV--QVDGESVQITIGIHTGEVVTGVIGQR  BETA1
548     VQIALMALKMMELSHEVVSP-HGEPIKMRIGLHSGSVFAGVVGVK  ALPHA1

519     MPRYCLFGDTVNTASRMESHGLPSKVHLSPTAHRAL-----KNKG  BETA2
536     MPRYCLFGNTVNLTSRTETTGEKGKINVSEYTYRCLMTPENSDPQ  BETA1
592     MPRYCLFGNNVTLANKFESCSVPRKINVSPTTYRLLKDCPGFVFT  ALPHA1

559     FEIVRRGEIEVKG-KGKMTTYFLIQNLNATEDEIMGRPSAPADGK  BETA2
581     FHLEHRGPVSMKGKKEPMQVWFLSRKNTGTEETEQDEN         BETA1
637     PRSREELPPNFPSDIPGICH-FLEAYQQGTTSKPWFQKKDVEEAN  ALPHA1

603     EVCTPGNQVRKSPAVPRNTDHQQQVYKGDPADASNEVTLAGSPVA  BETA2
681     ANFLGKASGID                                    ALPHA1

648     GRNSTDAVNNQPSPDETKTSVVASGPVLSAFCVVL            BETA2
```

FIGURE 3: Amino acid comparison of rat kidney-derived guanylyl cyclase (β2) and 73-kDa (α1) and 70-kDa (β1) subunits of bovine lung soluble guanylyl cyclase. Amino acid identities are shaded, and gaps are represented by dashes.

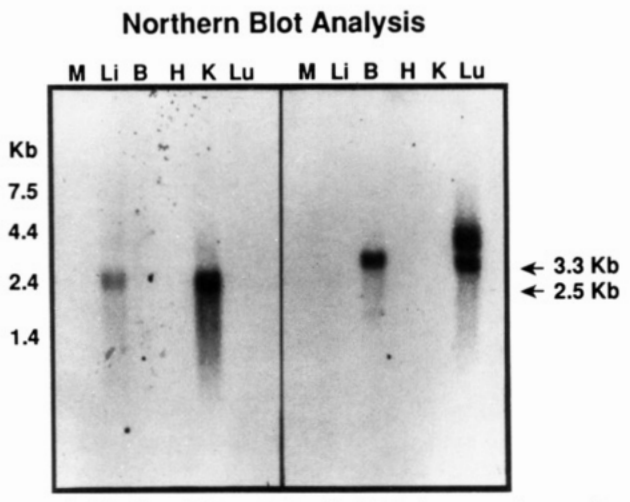

FIGURE 4: Northern blot analysis of mRNA from rat tissues using probes specific for the rat kidney-derived guanylyl cyclase (GC-$S_{\beta 2}$) and the 70-kDa subunit (GC-$S_{\beta 1}$) of rat lung soluble guanylyl cyclase. 10 μg of poly(A+) mRNA from rat skeletal muscle (M), liver (Li), brain (B), heart (H), kidney (K), or lung (L) was electrophoresed, blotted and fixed to a nitrocellulose membrane, and hybridized with a random prime labeled probe from clone 15 (left panel). After autoradiography the membrane was rehybridized with random prime labeled probe for GC-$S_{\beta 1}$ (right panel). The left panel represents a 14-day exposure, and the right panel is a 5-h exposure. The mobilities of RNA markers were determined by ethidium bromide staining and are indicated on the left.

the membrane (Hancock et al., 1989). A second cysteine close to the carboxyl terminus does not exist in GC-$S_{\beta 2}$ as is the case with ras, and the absolute requirements for fatty acid acylation are not known. Therefore, it is unclear whether or not GC-$S_{\beta 2}$ can be associated with membranes. The high level (50%) of hydrophilic amino acids in this C-terminal region, however, may not favor such an association.

We expressed the kidney cDNA for GC-$S_{\beta 2}$ in COS-7 cells (Cullen, 1987), but no guanylyl cyclase activity was detected (not shown); these results are predictable, on the basis of the lack of enzyme activity reported for the lung soluble enzyme subunit GC-$S_{\beta 1}$ (Nakane et al., 1988). Data presented by Nakane et al. (1990) suggest that lung soluble guanylyl cyclase activity requires coexpression of GC-$S_{\alpha 1}$ and GC-$S_{\beta 1}$. Therefore, GC-$S_{\beta 2}$ and an α-subunit yet to be identified are likely required for the expression of guanylyl cyclase activity.

ACKNOWLEDGMENTS

We thank Stephanie Schulz for providing rat tissue RNA, Sujay Singh for rat lung GC-$S_{\beta 1}$, and Laura Jo Rutledge for technical assistance. We also thank Mike Tamkun for the use of a thermal cycler and helpful suggestions regarding PCR.

REFERENCES

Aviv, H., & Leder, P. (1972) *Proc. Natl. Acad. Sci. U.S.A. 69*, 1408–1412.
Bredt, D. S., & Snyder, S. H. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 9030–9033.
Cathala, G., Savouret, J. F., Mendez, B., West, B. L., Karin, M., Martial, J. A. & Baxter, J. D. (1983) *DNA 2*, 329–335.
Chang, M. S., Lowe, D. G., Lewis, M., Hellmiss, R., Chen, E., & Goeddel, D. V. (1989) *Nature 341*, 68–72.
Chinkers, M., & Garbers, D. L. (1989) *Science 245*, 1392–1394.
Chinkers, M., Garbers, D. L., Chang, M. S., Lowe, D. G., Chin, H., Goeddel, D. V., & Schulz, S. (1989) *Nature 338*, 78–83.
Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J., & Rutter, W. J. (1979) *Biochemistry 18*, 5294–5299.
Chrisman, T. D., Garbers, D. L., Parks, M. A., & Hardman, J. G. (1975) *J. Biol. Chem. 250*, 374–381.
Clarke, S., Vogel, J. P., Deschenes, R. J., & Stock, J. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 4643–4647.
Cullen, B. R. (1987) *Methods Enzymol. 152*, 684–704.
Furchgott, R. F. (1984) *Annu. Rev. Pharmacol. Toxicol. 24*, 175–197.
Garbers, D. L. (1989) *J. Biol. Chem. 264*, 9103–9106.
Garthwaite, J., Charles, S. L., & Chess-Williams, R. (1988) *Nature 336*, 385–388.
Garthwaite, J., Garthwaite, G., Palmer, R. M., & Moncada, S. (1989) *Eur. J. Pharmacol. 172*, 413–416.
Gautam, N., Baetscher, M., Aebersold, R., & Simon, M. I. (1989) *Science 244*, 971–974.
Gibbs, J. B., Schaber, M.D., Schofield, T. L., Scolnick, E. M., & Sigal, I. S. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 6630–6634.
Graff, G., Stephenson, J. H., Glass, D. B., Haddox, M. K., & Goldberg, N. D. (1978) *J. Biol. Chem. 253*, 7662–7676.
Gutierrez, L., Magee, A. I., Marshall, C. J., & Hancock, J. F. (1989) *EMBO J. 8*, 1093–1098.
Haddox, M. K., Stephenson, J. H., Moser, M. E., & Goldberg, N. D. (1978) *J. Biol. Chem. 253*, 3143–3152.
Hancock, J. F., Magee, A. I., Childs, J. E., & Marshall, C. J. (1989) *Cell 57*, 1167–1177.
Hidaka, H., & Asano, T. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 3657–3661.
Ignarro, L. J., Byrns, R. E., Buga, G. M., & Wood, K. S. (1987) *Circ. Res. 61*, 866–879.
Kamisaki, Y., Saheki, S., Nakane, M., Palmieri, J. A., Kuno, T., Chang, B. Y., Waldman, S. A., & Murad, F. (1986) *J. Biol. Chem. 261*, 7236–7241.
Kimura, H., & Murad, F. (1974) *J. Biol. Chem. 249*, 6910–6916.
Knoth, K., Roberds, S., Poteet, C., & Tamkun, M. (1988) *Nucleic Acids Res. 16*, 10932.
Koesling, D., Herz, J., Gausepohl, H., Niroomand, F., Hinsch, K. D., Mulsch, A., Bohme, E., Schultz, G., & Frank, R. (1988) *FEBS Lett. 239*, 29–34.
Koesling, D., Harteneck, C., Humbert, P., Bosserhoff, A., Frank, R., Schultz, G., & Bohme, E. (1990) *FEBS Lett. 266*, 128–132.
Kozak, M. (1989) *J. Cell Biol. 108*, 229–241.
Krupinski, J., Coussen, F., Bakalyar, H. A., Tang, W. J., Feinstein, P. G., Orth, K., Slaughter, C., Reed, R. R., & Gilman, A. G. (1989) *Science 244*, 1558–1564.
Lee, C. C., Wu, X., Gibbs, R. A., Cook, R. G., Muzny, D. M., & Caskey, C. T. (1988) *Science 239*, 1288–1291.
Lewicki, J. A., Chang, B., & Murad, F. (1983) *J. Biol. Chem. 258*, 3509–3515.
Libert, F., Parmentier, M., Lefort, A., Dinsart, C., VanSande, J., Maenhaut, C., Simons, M. J., Dumont, J. E., & Vassart, G. (1989) *Science 244*, 569–572.
Lowe, D. G., Chang, M. S., Hellmiss, R., Chen, E., Singh, S., Garbers, D. L., & Goeddel, D. V. (1989) *EMBO J. 8*, 1377–1384.
Marletta, M. A. (1989) *Trends Biochem. Sci. 14*, 488–492.
Mullis, K. B., & Faloona, F. (1987) *Methods Enzymol. 155*, 335–350.
Myers, P. R., Minor, R. L., Jr., Guerra, R., Bates, J. N., & Harrison, D. G. (1990) *Nature 345*, 161–163.
Nakane, M., Seheki, S., Kuno, T., Ishii, K., & Murad, F. (1988) *Biochem. Biophys. Res. Commun. 156*, 1000–1006.

Nakane, M., Arai, K., Saheki, S., Kuno, T., Buechler, W., & Murad, F. (1990) *J. Biol. Chem. 265*, 16841-16845.
Palmer, R. M. J., Ferrige, A. G., & Moncada, S. (1987) *Nature 327*, 524-526.
Rapoport, R. M., & Murad, F. (1983) *J. Cyclic Nucleotide Protein Phosphorylation Res. 9*, 281-296.
Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463-5467.
Schulz, S., Singh, S., Bellet, R. A., Singh, G., Tubb, D. J., Chin, H., & Garbers, D. L. (1989) *Cell 58*, 1155-1162.
Singh, S., Lowe, D. G., Thorpe, D. S., Rodriguez, H., Kuang, W. J., Dangott, L. J., Chinkers, M., Goeddel, D. V., & Garbers, D. L. (1988) *Nature 334*, 708-712.
Thorpe, D. S., & Garbers, D. L. (1989) *J. Biol. Chem. 264*, 6545-6549.
Thorpe, D. S., & Morkin, E. (1990) *J. Biol. Chem. 265*, 14717-14720.
Tracey, W. R., Linden, J., Peach, M. J., & Johns, R. A. (1990) *J. Pharmacol. Exp. Ther. 252*, 922-928.
Waldman, S. A., & Murad, F. (1987) *Pharmacol. Rev. 39*, 163-195.
Wilks, A. F., Kurban, R. R., Hovens, C. M., & Ralph, S. J. (1989) *Gene 85*, 67-74.
Willumsen, B. M., Christensen, A., Hubbert, N. L., Papageorge, A. G., & Lowy, D. R. (1984) *Nature 310*, 583-586.

A Tetrahedral Zinc(II)-Binding Site Introduced into a Designed Protein†

Lynne Regan*,‡ and Neil D. Clarke§

The Laboratory of Molecular Biology, Medical Research Council, Hills Road, Cambridge CB2 2QH, U.K., The Experimental Station, E. I. du Pont de Nemours and Company, Wilmington, Delaware 19880, and Howard Hughes Medical Institute and Department of Molecular Biology and Genetics, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205

Received August 14, 1990; Revised Manuscript Received October 3, 1990

ABSTRACT: The ultimate goal of protein engineering is to create novel proteins which will adopt predetermined structures, bind specified ligands, and catalyze new reactions. Here we describe the successful introduction of metal-binding activity into a model four helix bundle protein. The designed binding site is tetrahedral and is formed by two Cys and two His ligands on adjacent helices. We have introduced this site into the protein and characterized the binding activity. Using $^{65}Zn(II)$, we have shown that the protein binds Zn(II), that the sulfhydryls are essential for binding, and that binding occurs to the protein monomer. The designed protein binds metals with high affinity: we estimate the dissociation constants as 2.5×10^{-8} M for Zn(II) and 1.6×10^{-5} M for Co(II). The characteristic absorption spectrum of the Co(II)-substituted protein fully supports the model of a tetrahedral binding site comprised of two Cys and two His ligands. Circular dichroism studies indicate that no significant changes in secondary structure occur between the metal-bound and metal-free forms of the protein. However, the metal-bound form is substantially stabilized toward denaturation by GuHCl compared to the metal-free form.

The challenge of protein engineering is 2-fold: first, to design and create novel proteins which will fold to give desired structures; second, to incorporate novel activities onto these structural frameworks. In this paper we describe the successful accomplishment of an example of the second stage of the design challenge: the introduction of metal-binding activity onto a designed four helix bundle framework. The site we have introduced is a tetrahedral binding site for metal ions such as Zn(II), Cd(II), and Co(II).

Among natural proteins, tetrahedral Zn(II) sites are by far the most common, serving both structural and catalytic roles. The geometry of such sites in proteins and small molecules is well established from X-ray crystallographic studies. The creation of a tetrahedrally liganded metal-binding site, requiring the precise positioning of four amino acid side chains, represents a very specific design challenge whose success is readily testable by solution spectroscopic techniques.

The designed protein into which the metal-binding site was introduced has been described previously (Regan & DeGrado, 1988). It is a simplified version of a motif found in several naturally occurring proteins (Weber & Salemme, 1980; Presnel & Cohen, 1989). The essential elements of the protein's structure are four α-helices which are packed nearly antiparallel, at an angle of about 20°, and which are connected by three loops. In natural four helix bundle proteins, four helices of varying length and sequence are connected by three loops which vary in length, sequence, and connectivity. In the model protein the design has been simplified to give minimal sequence complexity: four identical helices are connected by three identical loops. The X-ray crystal structure of the protein has not yet been determined; however, solution studies are consistent with the proposed model of the protein as a stable, compact bundle structure (Regan & DeGrado, 1988; Ho & DeGrado, 1987). We will refer to this model protein as wild-type α_4, and its sequence is shown in Figure 1A.

We chose to design sites by use of combinations of His and Cys ligands. Natural examples of such tetrahedral Zn(II) sites are found in alcohol dehydrogenase (Vallee & Auld, 1990) and in the TFIIIA-like Zn-finger proteins (Daikun et al., 1986; Parraga et al., 1988; Lee et al., 1989). The site was designed

† This research was supported by E. I. du Pont de Nemours and Company and a British Ramsay Memorial Fellowship in Chemistry (L.R.) and by support from the Howard Hughes Medical Institute to Prof. Carl O. Pabo (N.D.C.)

* To whom correspondence should be addressed at the Department of Molecular Biophysics and Biochemistry, Yale University, 260 Whitney, Ave., P.O. Box 6666, New Haven, CT 06511.
‡ Medical Research Council and E. I. du Pont de Nemours and Company.
§ Howard Hughes Medical Institute and Johns Hopkins University School of Medicine.

0006-2960/90/0429-10878$02.50/0