Proc. Natl. Acad. Sci. USA
Vol. 78, No. 2, pp. 848–851, February 1981
Biochemistry

Complete amino acid sequence of the catalytic subunit of bovine cardiac muscle cyclic AMP-dependent protein kinase

(regulatory enzyme/phosphoserine/phosphothreonine/metabolic control)

SHOZO SHOJI*†, DAVID C. PARMELEE‡, ROGER D. WADE*, SANTOSH KUMAR‡, LOWELL H. ERICSSON*, KENNETH A. WALSH*, HANS NEURATH*, GEORGE L. LONG*, JACQUES G. DEMAILLE*§, EDMOND H. FISCHER*, AND KOITI TITANI‡

‡The Howard Hughes Medical Institute Laboratory and *Department of Biochemistry, SJ-70, University of Washington, Seattle, Washington 98195

Contributed by Edmond H. Fischer, October 27, 1980

ABSTRACT **The complete amino acid sequence of the 349-residue catalytic subunit of cyclic AMP-dependent protein kinase from bovine cardiac muscle is presented. The sequence of the subunit (M_r 40,580 including phosphate groups at threonine-196 and serine-337) was derived largely by automated Edman degradation of nine fragments generated from the carboxymethylated protein by cleavage of methionyl bonds with cyanogen bromide. These fragments were aligned along the polypeptide chain by analysis of methionine-containing tryptic peptides isolated from protein radiolabeled *in vitro* by [^{14}C]methyl exchange at methionyl residues. The molecule contains only two cysteinyl residues, at positions 198 and 342. It is relatively polar, containing clusters of cationic residues toward the amino terminus and anionic residues towards the carboxyl terminus. Predictions of secondary structure suggest the presence of three major domains with approximately half of the residues occurring in α-helices and 12% in β-strands.**

Cyclic AMP-dependent protein kinases (ATP:protein phosphotransferase, EC 2.7.1.37) are widely distributed and play a central role in the regulation of energy metabolism and other physiological functions (1–5). In eukaryotic systems, the actions of cyclic AMP are mediated exclusively through the activation of this protein kinase. Two forms of the enzyme with slightly different regulatory properties have been isolated (6, 7): type I, found predominantly in rabbit skeletal muscle, and type II, found in bovine cardiac muscle. Both types consist of two regulatory (R) and two catalytic (C) subunits. Differences between the two are attributed to differences in the structure of the regulatory subunits; the catalytic subunits are said to be identical (8). In both types, cyclic AMP acts as a positive effector resulting in dissociation of the oligomeric protein and activation of the enzyme according to the following scheme:

$$R_2C_2 \text{ (inactive)} + 4 \text{ cyclic AMP} \rightleftarrows (\text{R-cyclic AMP}_2)_2 + 2 \text{ C (active)}.$$

As a first step toward elucidating the primary structure of the enzyme, Peters *et al.* (9) developed a large-scale preparation of the catalytic subunit from bovine cardiac muscle and determined its molecular weight, isoelectric point, amino acid composition, and phosphate content and the reactivity of its thiol groups. We recently described the amino acid sequence around the two sites that are phosphorylated (10). In the present communication, the sequence of the 349 amino acid residues in the catalytic subunit is presented and discussed. The experimental details of the sequence determination will be published elsewhere.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

METHODS

The catalytic subunit of cyclic AMP-dependent protein kinase was prepared from bovine heart as described (9). TPCK-Trypsin, pepsin, and carboxypeptidase A were purchased from Worthington. *Myxobacter* AL-I protease II (11) was a generous gift of H. G. Wood (Case Western Reserve University). Sulfhydryl groups were carboxymethylated with iodoacetic acid (Sigma) or its ^{14}C-labeled form (New England Nuclear).

Amino acid analyses were performed with a Dionex amino acid analyzer (model D-500). Automated sequence analyses were performed with a Beckman sequencer (model 890) according to Edman and Begg (12) as modified by either Hermodson *et al.* (13) or Brauer *et al.* (14). Small peptides were analyzed in the sequencer in the presence of Polybrene (15). Some peptides were analyzed with a Sequemat by the method of Laursen (16) after attachment to the solid phase via carboxyl-terminal homoseryl or lysyl residues (17, 18). Phenylthiohydantoin derivatives of the amino acids were identified by two complementary systems of reversed-phase high-performance liquid chromatography (19, 20).

Secondary structure was predicted by using a computerized program according to Chou and Fasman (21), with conformational parameters based upon 29 globular proteins. To the original algorithm was added the rule that two periodic secondary structures must be separated by a minimum of four amino acid residues. Boundary regions and β-turns were assigned manually on the basis of the frequency hierarchies given by Chou and Fasman (21).

Homology with other protein sequences was sought by applying the comparison program of deHäen *et al.* (22) to a collection of the sequences of about 750 proteins (from a 1978 data tape obtained from the National Biomedical Research Foundation).

The seven methionyl bonds in the catalytic subunit were cleaved with cyanogen bromide (Eastman) and eight primary fragments (CB 1–CB 8) were separated by gel filtration and ion exchange chromatography. The four smallest fragments were obtained directly by gel filtration through a column of Sephadex G-25 (superfine) in 9% formic acid. A pool of larger fragments from this same column was treated with citraconic anhydride (Pierce) (23) to minimize aggregation and separated into three

Abbreviation: DTNB, 5,5′-dithiobis(2-nitrobenzoic acid).

† Present address: Department of Biochemistry, Faculty of Pharmaceutical Sciences, Kumamoto University, Kumamoto 862, Japan.
§ Present address: Centre National de la Recherche Scientifique and Institut National de la Santé et de la Recherche Médicale–Centre de Recherches de Biochimie Macromoleculaire, Boite PostaPe 5051-34033, Montpellier Cedex, France.

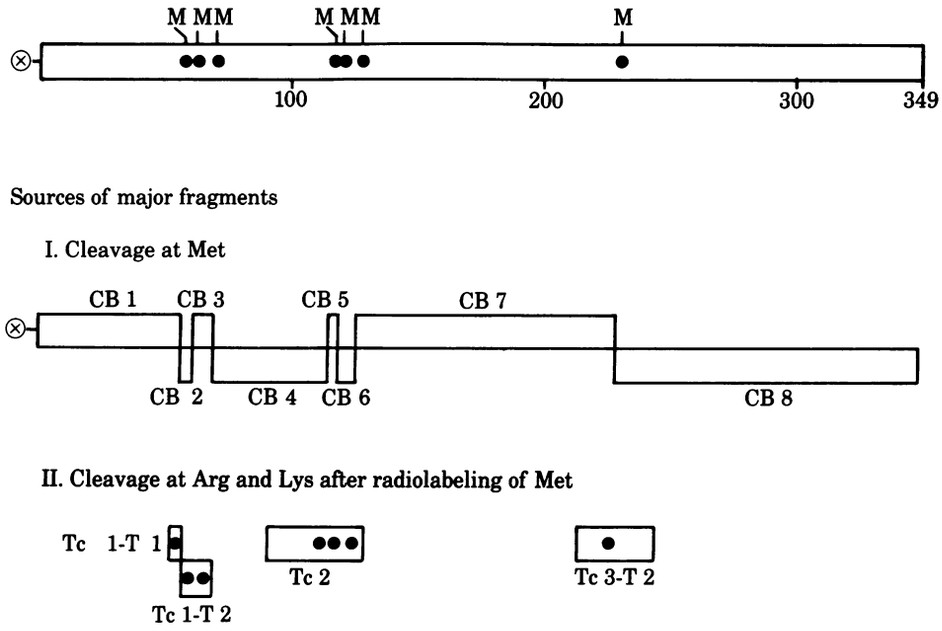

FIG. 1. Diagram indicating the origin of the primary fragments of the catalytic subunit and their relationship to each other. The length of each bar is proportional to the length of the peptide. X, Unidentified blocking group; letter M and ●, positions of methionyl residues. Fragments generated by cleavage with CNBr are identified with the prefix CB; overlapping tryptic peptides have the prefix T.

fractions on a column of Sephadex G-75 (superfine) in 0.1 M NH_4HCO_3 at pH 8. The first fraction still contained aggregated material but was finally separated (into CB 7 and a fragment that overlapped CB 7 and CB 8) by rechromatography on the same column in the presence of 6 M guanidine·HCl. The second fraction contained pure CB 8. After decitraconylation, the third fraction was resolved into CB 1 and CB 4 by applying a linear gradient of sodium formate buffers in 7 M urea to a column of SP-Sephadex C-25.

Methionyl residues of the carboxymethylated protein were labeled with [^{14}C]methyl iodide (New England Nuclear) *in vitro* by the method of Link and Stark (24). The product was citraconylated and treated with trypsin; these products were separated into two radioactive fractions on a column of Sephadex G-50 (superfine) in 0.1 M NH_4HCO_3 at pH 8. The first fraction was further purified on Sephadex G-75 (superfine) in 0.1 M NH_4HCO_3/6 M guanidine·HCl, pH 8, to yield peptide Tc2 (Fig. 1). The second fraction, after removal of citraconyl groups, was again digested with trypsin, yielding three radioactive peptides [Tc1-T1, Tc1-T2, and Tc3-T2 (Fig. 1)] by a combination of gel filtration (Sephadex G-50) and preparative high-voltage paper electrophoresis.

```
                5        10        15        20        25        30
 1  (X)G N A A A K K G S E Q E S V K E F L A K A K E D F L K K W E
31    N P A Q N T A H L D Q F E R I K T L G T G S F G R V M L V K
61    H M E T G N H Y A M K I L D K Q K V V K L K Q I E H T L N E
91    K R I L Q A V N F P F L V K L E F S F K D N S N L Y M V M E
121   Y V P G G E M F S H L R R I G R F S E P H A R F Y A A Q I V
151   L T F E Y L H S L D L I Y R D (P)L K P E N L L I D Q Q G Y I Q
181   V T D F G F A K R V K G R T W (P)T L C G T P E Y L A P E I I L
211   S K G Y N K A V D W W A L G V L I Y E M A A G Y P P F F A D
241   Q P I Q I Y E K I V S G K V R F P S H F S S D L K D L L R N
271   L L Q V D L T K R F G N L K D G V N D I K N H K W F A T T D
301   W I A I Y Q (P)R K V E A P F I P K F K G P G D T S N F D D Y E
331   E E E I R V S I N E K C G K E F S E F
```

FIG. 2. Amino acid sequence of the catalytic subunit of bovine cardiac muscle cyclic AMP-dependent protein kinase. P, *O*-phosphothreonine-196 and *O*-phosphoserine-337; X, unidentified blocking group. See Table 1 for composition and code.

RESULTS AND DISCUSSION

Only two different methods of primary fragmentation of the catalytic subunit were required (CNBr and tryptic cleavage) to yield a complete set of overlapping peptides (Fig. 1). An aliquot of each fragment was subjected to Edman degradation and the remainder was digested with appropriate proteases to provide small peptides to complete the analysis. The sequence so derived (Fig. 2) places the previously characterized O-phosphothreonine at residue 196 and O-phosphoserine at residue 337.

The sequence of the amino- and carboxyl-terminal portions of the molecule were the most difficult to determine. The amino-terminal segment presented special problems because it is blocked and because citraconylated CB 1 (the amino-terminal fragment) comigrated with citraconylated CB 4 on both sodium dodecyl sulfate gels and Sephadex G-75. Because CB

Table 1. Composition of catalytic subunit

Amino acid		Residues/mol	Amino acid		Residues/mol
Ala	(A)	22	Leu	(L)	32
Arg	(R)	15	Lys	(K)	34
Asn	(N)	16	Met	(M)	7
Asp	(D)	19	Phe	(F)	25
CysH	(C)	2	Pro	(P)	14
Gln	(Q)	14	Ser	(S)	16
Glu	(E)	27	Thr	(T)	14
Gly	(G)	22	Trp	(W)	6
His	(H)	9	Tyr	(Y)	14
Ile	(I)	21	Val	(V)	20

Molecular weight without phosphate, 40,420; with phosphate, 40,580. Total residues, 349.

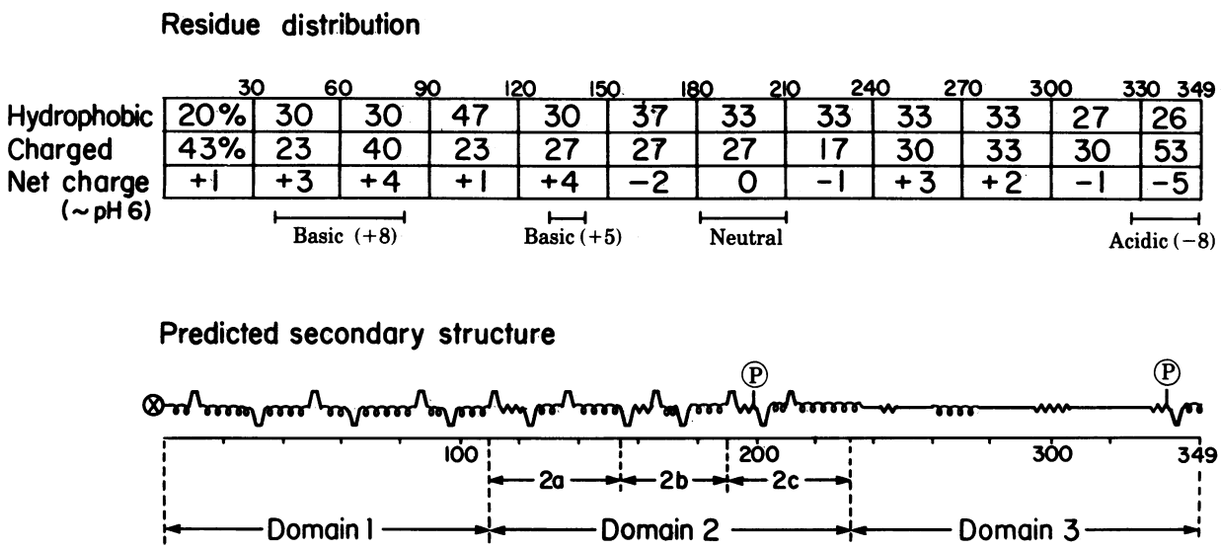

FIG. 3. Distribution of hydrophobic residues (Met, Val, Ile, Leu, Phe, and Trp) and charged groups (Arg, Lys, His, Asp, Glu, phosphoryl, and α-carboxyl) in segments of the sequence of the catalytic subunit. Amino acids are grouped in 30-residue segments (except the 19 carboxyl-terminal residues) as indicated. The lower portion summarizes the prediction of secondary structure: coil, α-helix; sawtooth pattern, β-strand; straight line, aperiodic structure; arch (sometimes inverted), β-turn. Hypothetical domains (see text) are delineated at the bottom of the diagram.

1 is blocked, the mixture of CB 1 and CB 4 behaved like a single homogeneous fragment on the basis of both Edman degradation and sodium dodecyl sulfate/polyacrylamide gel electrophoresis. Furthermore, amino acid analysis yielded integral values. Only after additional digestion of this mixture did it become clear that there were more peptides than could be placed in a single fragment of ≈52 residues. When CB 1 and CB 4 were separated, the peptides could readily be grouped into two consistent sets. Even then, the blocked amino terminus of CB 1 offered unusual difficulty. Two small blocked peptides were isolated from CB 1 by high-performance liquid chromatography on a cyanopropyl column. One, from a peptic digest of CB 1, proved to be X-Gly-Asn-Ala and the other, from a tryptic digest, was X-Gly-Asn-Ala-Ala-Ala-Lys. The nature of group X remains to be elucidated. From the mobility of the blocked tripeptide on high-performance liquid chromatography, group X appears to be much more hydrophobic than the acetyl group found at the amino termini of many muscle proteins. The sequence of the carboxyl-terminal region of the protein was difficult to determine because of the repeating dipeptide in the sequence Lys-Glu-Phe-Ser-Glu-Phe. Initially, this led to confusion in the interpretation of data obtained by both carboxypeptidase digestion and amino acid analysis of the carboxyl-terminal tryptic peptide.

Another difficulty in the proof of structure was encountered in the region of residues 116–120 (Tyr-Met-Val-Met-Glu), where CB 5 (Val-Met) is interposed between CB 4 and CB 6. The peptide necessary to overlap the three CNBr fragments could not be generated by tryptic digestion because aspartic acid-111 prevented cleavage at lysine-110. This sequence was finally established in a peptide derived from Tc 2 (Fig. 1) by cleavage of the Phe_{109}-Lys_{110} bond with the enzyme from *Myxobacter* (11).

The amino acid composition and the molecular weight derived from the sequence data are in good agreement with those previously obtained by amino acid analysis and sodium dodecyl sulfate/polyacrylamide gel electrophoresis of the whole protein (9). One notable exception is the number of cysteinyl residues. This was reported as 3 by titration with 5,5′-dithiobis(2-nitrobenzoic acid) (DTNB) (9, 25). In the present experiments, carboxymethylation of the catalytic subunit with iodo[^{14}C]acetate led to incorporation of 538 × 10^9 cpm/mol of protein, corresponding to 1.5 equivalents of carboxymethylcysteine. More than 90% of the radioactivity was recovered in fragments CB 7 (253 × 10^9 cpm/mol) and CB 8 (237 × 10^9 cpm/mol). Further digestion yielded one small radioactive peptide from each fragment without change in specific radioactivity. On this basis, it was concluded that only two cysteinyl residues were present; they occur at residues 198 and 342 (Fig. 2). This conclusion is in accord with the DTNB titration data of Armstrong and Kaiser (26) and with an alternate analytical method of Hartl (27) but is at variance with the DTNB titration data of Bechtel *et al.* (28), Peters *et al.* (9), and Kupfer *et al.* (25) who reported values between 2.3 and 3.0 cysteinyl residues per molecule. It is difficult to explain these discrepancies unless one questions the specificity of DTNB under the different experimental conditions used. Peters *et al.* (9) also cleaved the catalytic subunit at cysteinyl residues with 2-nitro-5-thiocyanobenzoic acid and observed three fragments, 23,000, 16,500, and 15,000 daltons, respectively. These results are consistent with the present observations of cysteinyl residues 198 and 342 if one assumes incomplete cleavage at cysteine 342.

On the basis of electrofocussing experiments, several investigators have reported that preparations of the bovine catalytic subunit contain at least three isozymes with identical substrate specificities (9, 28, 29). However, no primary structure heterogeneity was detected in the present study of the isozyme mixture. Similarly, no difference in the state of phosphorylation was previously detected (9). The most likely explanations of the apparent heterogeneity involve either differential ampholyte binding to the protein or microheterogeneity in the state of amidation.

Recently, Hixson and Krebs (30) and Zoller and Taylor (31) reported affinity labeling of the ATP-Mg binding site of the catalytic subunit by *p*-fluorosulfonylbenzoyl-5′-adenosine. Using the radioactive reagent, the latter investigators labeled a single lysyl residue of the porcine skeletal type II enzyme and isolated a radioactive peptide with a sequence nearly identical to that of residues 63–75 (Fig. 2) of the bovine cardiac enzyme but containing phenylalanine in place of tyrosine-68. The label

could be fitted in a position corresponding to lysine-71¶ in spite of the species difference.

No information is yet available on the protein substrate binding site or on the site(s) of interaction with the regulatory subunit. Because the specificity of the kinase (32–39) demands that a cluster of basic charges occur on the NH_2-terminal side of the substrate residue to be phosphorylated (5), it is tempting to speculate that the acidic region found at the carboxyl end of the enzyme (residues 327–333) is involved in substrate recognition.

The distribution of charged residues on the enzyme (Fig. 3) shows a marked polarity along the peptide chain—i.e., clusters of positive charges in the amino-terminal half of the molecule, a neutral region in the middle around phosphothreonine-196, and a negatively charged region near the carboxyl terminus. Hydrophobic residues occur at a frequency about 28% higher than the average found in 185 other proteins (40); a high proportion of these occur in the region of residues 93–129.

The pattern of predicted secondary structure indicates that the molecule may be composed of three distinct structural domains approximately equal in size (Fig. 3). The first domain (residues 1–110) is highly ordered and consists of α-helices (79%) and β-turns. The second domain (residues 111–232) is also ordered and appears to contain three subdomains of similar length: 2a, residues 111–154; 2b, 155–189; and 2c, 190–232. Each of these would have a secondary structure consisting of a β-strand followed by two α-helices and separated by two β-turns. Computer analysis for internal sequence homologies among the three suggested subdomains indicates similarity of 2b and 2c but not of either with 2a. The third domain comprises the carboxy-terminal portion of the molecule (residues 233–349); it is highly aperiodic (62%), containing only a small porportion of α-helix (18%) and β-strand (20%).

Secondary structure analysis places both of the phosphoryl groups within β-strands. Both groups are very resistant to dephosphorylation by protein phosphatases; prolonged attack with phosphorylase phosphatase released approximately half of the phosphoryl groups serine-337 and essentially none from threonine-196. Serine-337 can be rephosphorylated in an autocatalytic process (David Brautigan, personal communication).

Preliminary comparison of the amino acid sequence of the catalytic subunit with the sequences of 750 other proteins shows some similarity with the family of dehydrogenases and with glycogen phosphorylase. Although the predicted secondary structure of the kinase subunit is inconsistent with structures characterizing other nucleotide-binding domains (41), it is interesting that the indications of distant homology are found in the amino-terminal domain where the ATP-binding site¶ is located.

The authors are grateful to Dr. David C. Teller for the analyses of homology and to Richard R. Granberg for the amino acid analyses. This work has been supported in part by grants from the National Institutes of Health (GM15731, GM27335, and AM7902), from the Centre National de la Recherche Scientifique, Institut National de la Santé et de la Recherche Médicale, and Direction Générale à la Recherche Scientifique et Technique (to J.G.D.), and from a collaborative North Atlantic Treaty Organization grant (no. 1688) to J.G.D. and E.H.F.

¶ Taylor, S. S., Kerlavage, A. R., Zoller, M. J., Nelson, N. C. & Potter, R. L. (1980) *Second Chemical Congress of the North American Continent*, Division of Biological Chemistry, San Francisco, August 1980, Abstr. 238.

1. Krebs, E. G. (1972) *Curr. Top. Cell. Regul.* **5**, 99–133.
2. Rubin, C. S. & Rosen, O. M. (1975) *Annu. Rev. Biochem.* **44**, 831–887.
3. Rosen, O. M., Rangel-Aldao, R. & Erlichmann, J. (1977) *Curr. Top. Cell. Regul.* **12**, 39–71.
4. Cohen, P. (1978) *Curr. Top. Cell. Regul.* **14**, 118–192.
5. Krebs, E. G. & Beavo, J. A. (1979) *Annu. Rev. Biochem.* **48**, 923–959.
6. Reimann, E. M., Walsh, D. A. & Krebs, E. G. (1971) *J. Biol. Chem.* **246**, 1986–1995.
7. Rubin, C. S., Erlichmann, J. & Rosen, O. M. (1972) *J. Biol. Chem.* **247**, 36–44.
8. Hofmann, F., Beavo, J. A., Bechtel, P. J. & Krebs, E. G. (1975) *J. Biol. Chem.* **250**, 7795–7801.
9. Peters, K. A., Demaille, J. G. & Fischer, E. H. (1977) *Biochemistry* **16**, 5691–5697.
10. Shoji, S., Titani, K., Demaille, J. G. & Fischer, E. H. (1979) *J. Biol. Chem.* **254**, 6211–6214.
11. Wingard, M., Matsueda, G. & Wolfe, R. S. (1972) *J. Bacteriol.* **112**, 940–949.
12. Edman, P. & Begg, G. (1967) *Eur. J. Biochem.* **1**, 80–91.
13. Hermodson, M. A., Ericsson, L. H., Titani, K., Neurath, H. & Walsh, K. A. (1972) *Biochemistry* **11**, 4493–4502.
14. Brauer, A. W., Margolies, M. N. & Haber, E. (1975) *Biochemistry* **14**, 3029–3035.
15. Tarr, G. E., Beecher, J. F., Bell, M. & McKean, D. J. (1978) *Anal. Biochem.* **84**, 622–627.
16. Laursen, R. A. (1971) *Eur. J. Biochem.* **20**, 89–102.
17. Horn, M. J. & Laursen, R. A. (1973) *FEBS Lett.* **36**, 285–288.
18. Laursen, R. A., Horn, M. J. & Bonner, A. G. (1972) *FEBS Lett.* **21**, 67–70.
19. Bridgen, P. J., Cross, G. A. M. & Bridgen, J. (1976) *Nature (London)* **263**, 613–614.
20. Ericsson, L. H., Wade, R. D., Gagnon, J., McDonald, R. M., Granberg, R. R. & Walsh, K. A. (1977) in *Solid Phase Methods in Protein Sequence Analysis*, eds. Previero, A. & Coletti-Previero, M. A. (North-Holland, Amsterdam), pp. 137–142.
21. Chou, P. Y. & Fasman, G. D. (1978), in *Advances in Enzymology*, ed. Meister, E. (Wiley, New York), Vol. 47, pp. 45–148.
22. deHäen, C., Swanson, E. & Teller, D. C. (1976) *J. Mol. Biol.* **106**, 639–661.
23. Habeeb, A. F. S. A. & Atassi, M. Z. (1972) *Biochemistry* **9**, 4939–4952.
24. Link, T. P. & Stark, G. R. (1967) *J. Biol. Chem.* **243**, 1082–1088.
25. Kupfer, A., Gani, V., Jimenez, J. S. & Shaltiel, S. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 3073–3077.
26. Armstrong, R. N. & Kaiser, E. T. (1978) *Biochemistry* **17**, 2840–2845.
27. Hartl, F. T. (1980) *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **39**, 2094.
28. Bechtel, P. J., Beavo, J. A. & Krebs, E. G. (1977) *J. Biol. Chem.* **252**, 2691–2697.
29. Sugden, P. H., Holladay, L. A., Reimann, E. M. & Corbin, J. D. (1976) *Biochem. J.* **159**, 409–422.
30. Hixson, C. S. & Krebs, E. G. (1979) *J. Biol. Chem.* **254**, 7509–7514.
31. Zoller, M. J. & Taylor, S. S. (1979) *J. Biol. Chem.* **254**, 8363–8368.
32. Langan, T. A. (1971) *Ann. N.Y. Acad. Sci.* **185**, 166–180.
33. Hjelmquist, G., Andersson, J., Edlund, B. & Engström, L. (1974) *Biochem. Biophys. Res. Commun.* **61**, 559–563.
34. Huang, T. S., Bylund, D. B., Stull, J. T. & Krebs, E. G. (1974) *FEBS Lett.* **42**, 249–252.
35. Kemp, B. E., Bylund, D. B., Huang, T. S. & Krebs, E. G. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 3448–3452.
36. Daile, P., Carnegie, P. R. & Young, J. D. (1975) *Nature (London)* **257**, 416–418.
37. Zetterqvist, O., Ragnarsson, U., Humble, E., Berglund, L. & Engström, L. (1976) *Biochem. Biophys. Res. Commun.* **70**, 696–703.
38. Yeaman, S. J., Cohen, P., Watson, D. C. & Dixon, G. H. (1976) *Biochem. Soc. Trans.* **4**, 1027–1030.
39. Kemp, B. E., Graves, D. J., Benjamini, E. & Krebs, E. G. (1977) *J. Biol. Chem.* **252**, 4888–4894.
40. Dayhoff, M. O., Dayhoff, R. E. & Hunt, L. T. (1976) in *Atlas of Protein Sequence and Structure*, ed. Dayhoff, M. O. (National Biomedical Research Foundation, Silver Spring, MD), Vol. 5, Suppl. 2, pp. 301–310.
41. Rossman, M. G., Liljas, A., Branden, C. I. & Banaszak, L. J. (1975) in *The Enzymes*, ed. Boyer, P. D. (Academic, New York), 3rd Ed., Vol. 11, pp. 61–102.