Biochem. J. (1992) 287, 791–795 (Printed in Great Britain)

A protonated histidine residue in a phosphorylation site for cyclic AMP-dependent protein kinase

Comparison of a synthetic peptide with the exposed linking region in the multienzyme polypeptide CAD

Elizabeth A. CARREY
Department of Biochemistry, University of Dundee, Dundee DD1 4HN, Scotland, U.K.

The multienzyme polypeptide CAD is phosphorylated at two sites by cyclic AMP (cAMP)-dependent protein kinase. Site 2 has two interesting features: it is located in a 'linking region' between two discretely folded enzyme domains, and a histidine, instead of the more usual arginine, is found three positions *N*-terminal to the phosphorylated serine. A synthetic peptide corresponding to the sequence around site 2 has an extended or random structure in solution, and the proton n.m.r. chemical shift of the histidine residues can be titrated against pH in the range 6.0–8.0. The peptide is phosphorylated more rapidly by cAMP-dependent protein kinase at lower pH values, indicating that the protonated histidine side chain corresponds to the arginine in the consensus recognition sequence for the kinase. Kemptide, a specific synthetic substrate for the kinase, was phosphorylated with a higher affinity and at a similar rate at all pH values. CAD was a better substrate than the synthetic peptide, and labelling was not affected by the pH of the incubation conditions. The results indicate that the phosphorylation site in the interdomain linker is sufficiently exposed to the solvent to ensure accessibility to the kinase, but that secondary or tertiary structure in the intact protein allows the histidine residue to remain protonated at physiological pH and enhances recognition of the phosphorylatable serine residue.

INTRODUCTION

The mammalian multienzyme protein CAD, which catalyses the first three reactions of pyrimidine biosynthesis, is phosphorylated on two serine residues *in vitro* by cyclic AMP (cAMP)-dependent protein kinase [1]. The enzyme activities of glutamine-dependent carbamoyl phosphate synthetase (CPSase), aspartate transcarbamylase (ATCase) and dihydro-orotase (DHOase) are found in discretely folded domains of this large (molecular mass 240 kDa) protein. When CAD is phosphorylated at site 1, which is at the *C*-terminal end of the CPSase domain, that enzyme reaction is released from feedback inhibition by UTP [2]. The protein kinase also phosphorylates a serine residue in a 'linking region' of 120 residues between the DHOase and ATCase domains, but this modification does not affect the activity of the CPSase or the ATCase. The phosphorylated protein, however, is more susceptible to attack by proteolytic enzymes such as trypsin and elastase [3], suggesting that there has been a generalized alteration in the conformation of CAD, exposing the interdomain regions to proteases.

The serine residue at site 2 is found in an unusual sequence (His-Arg-Ala-Ser) compared with the consensus residues specifying phosphorylation by cAMP-dependent kinase [4]. The specificity of the kinase is limited to residues very close to the phosphorylatable serine, usually with the sequence Arg-Arg-Xaa-Ser. I have proposed and tested the idea that the nearby histidine side chain must be positively charged for the kinase to recognize it in the place of the arginine residue. Using a 20-residue synthetic peptide corresponding to the region surrounding phosphorylation site 2, I compared the initial rate of phosphorylation of the serine residue by cAMP-dependent protein kinase over a range of pH values from 6 to 8, and compared the labelling of the equivalent site in the protein CAD. The results show that a protonated histidine side chain is essential for modification of the nearby serine. I believe that this is the first report of the effect of a protonated histidine residue in a site recognized by cAMP-dependent protein kinase.

EXPERIMENTAL

Materials

9-Fluorenylmethoxycarbonyl (Fmoc) amino acids and h.p.l.c.-grade solvents were obtained from Applied Biosystems, Warrington, Cheshire, U.K., and Rathburn Chemicals, Walkerburn, Peeblesshire, Scotland, U.K., respectively. Buffer compounds Hepes, Bis-Tris, Tricine and Tris, and the synthetic peptide kemptide (Leu-Arg-Arg-Ala-Ser-Leu-Gly) were from Sigma. Optiscint T was from Pharmacia LKB, and [γ-^{32}P]ATP was from Amersham.

Synthetic peptide

The synthetic peptide SP 28 was synthesized by Dr. D. G. Campbell on an Applied Biosystems 431A peptide synthesizer using Fmoc chemistry. The peptide was removed from the resin and unblocked using established methods. Reverse-phase h.p.l.c. was used to establish the purity of the peptide; it was 100% pure by this criterion. The peptide concentration was determined by amino acid analysis on a Waters Pico-Tag system. The peptide was readily soluble in water. Its sequence is NH_2-His-Leu-Pro-Pro-Arg-Ile-His-Arg-Ala-Ser-Asp-Pro-Gly-Leu-Pro-Ala-Glu-Glu-Pro-Lys-CO_2H.

Protein purification

The multienzyme polypeptide CAD was prepared from an overproducing hamster cell line as before [5], and stored at −70 °C in storage buffer [100 mM-Hepes, pH 7.0, 1 mM-

Abbreviations used: CAD, multienzyme polypeptide containing the first enzymes dedicated to pyrimidine biosynthesis in higher organisms; cAMP, cyclic AMP; CPSase, glutamine-dependent carbamoyl phosphate synthetase II (EC 6.3.5.5); ATCase, aspartate transcarbamylase (aspartate carbamoyltransferase; EC 2.1.3.2); DHOase, dihydro-orotase (EC 3.5.2.3); Fmoc, 9-fluorenylmethoxycarbonyl; SP28, synthetic peptide corresponding to a section of the CAD sequence, as described in this paper; NOESY, nuclear Overhauser enhancement spectroscopy.

dithiothreitol, 5% (w/v) glycerol and 30% (v/v) dimethyl sulphoxide].

The catalytic subunit of cAMP-dependent protein kinase was prepared from beef heart [6] and stored at −20 °C in 100 mM-Hepes (pH 7.0)/50% (w/v) glycerol. During the preparation, the kinase activity was assayed as described below, using kemptide and peptide SP28 as substrates. The purified kinase labelled 1 mM-kemptide at the rate of 900 pmol of phosphate groups/min per μl, i.e. the preparation contained 900 enzyme units/ml.

Phosphorylation assays

The phosphorylation of peptide substrates was assayed using the paper method [7]. The reaction mixture of 30 μl contained 50 mM-Hepes buffer, pH 7.0 (or another buffer as described below), 1 mM-EGTA, 1 mM-dithiothreitol, 5 mM-$MgCl_2$, 0.54 unit of kinase, 0.3 mM-peptide and [γ-^{32}P]ATP (0.2 mM; 300 c.p.m./pmol). The assay was started by the addition of ATP. After 30 min (or a shorter interval) at 30 °C, samples of 20 μl were removed from the tubes and pipetted on to 1 cm^2 pieces of phosphocellulose paper (P81; Whatman), which were immediately dropped into about 200 ml of 1% phosphoric acid solution. The paper squares were washed in three changes of the acid solution and two changes of water before a final rinse in acetone. After drying in air, the papers were placed in vials containing 5 ml of Optiscint T for scintillation counting. The specific radioactivity of the ATP was estimated by applying a small aliquot of [γ-^{32}P]ATP solution to a P81 paper square which was directly placed in the scintillant; the concentration of the ATP was estimated from the A_{260} of a similar aliquot.

The phosphorylation of CAD was assayed in a similar reaction mixture and used a similar method, except that the reaction was stopped by pipetting samples on to squares of 3MM filter paper, which were then dropped into 25% (w/v) trichloroacetic acid. This method for precipitating protein on to paper has been shown [7] to be equivalent to the use of phosphoric acid. Alternatively, the labelling of protein was quenched by precipitation with 1 ml of 25% (w/v) trichloroacetic acid and 30 μg of BSA added to each tube. The pellet was washed twice with acid and twice with water, before counting the Cerenkov radiation. Some of these pellets were exhaustively digested with trypsin for the isoelectric focusing/peptide mapping procedure [8].

Similar labelling experiments used a series of buffers adjusted to the required pH values with HCl or NaOH, and giving a final concentration of 50 mM-buffer in the incubation with kinase. Buffers were used within 1 pH unit either side of their pK values. The effective buffering capacities of Bis-Tris, Hepes and Tricine overlap, allowing comparison of pairs of buffers at similar pH over the range 6.0–8.0.

Circular dichroism

The c.d. spectra of 0.27 mM synthetic peptide were recorded in several of the buffer solutions used for phosphorylation experiments, using the JASCO J-600 spectropolarimeter at the Scottish SERC Circular Dichroism Facility, Stirling University. The potential for forming α-helical structure was tested by adding trifluoroethanol to the solutions. The CONTIN program [9] was used to estimate the secondary structure content of the peptide in solution.

Proton n.m.r.

The synthetic peptide (3 mM) was dissolved in potassium phosphate buffers (20 mM) across the relevant pH range, dried down and then redissolved to the same concentrations in 99.96% ^2H_2O. Proton chemical shifts were referenced to internal acetone (2.225 p.p.m.), and spectra were recorded at a probe temperature of 300 K. Proton–proton nuclear Overhauser enhancement spectroscopy (NOESY) spectra were recorded with sweep widths of 5000 Hz in each dimension and with 64 scans per t_1 increment.

RESULTS

Secondary structure of the synthetic peptide in solution

The c.d. spectrum of peptide SP28 is shown in Fig. 1. The same spectrum was obtained at four different pH values in the various buffers, and in the presence of 50% trifluoroethanol, a well-known helix-promoting solvent. The estimates of secondary structure shown in Table 1 were obtained from the CONTIN curve-fitting program, which was developed for analysis of proteins, and may therefore be inappropriate for application to a peptide of 20 residues. Nevertheless, the peptide in solution appears to have an extended or random structure at all the pH

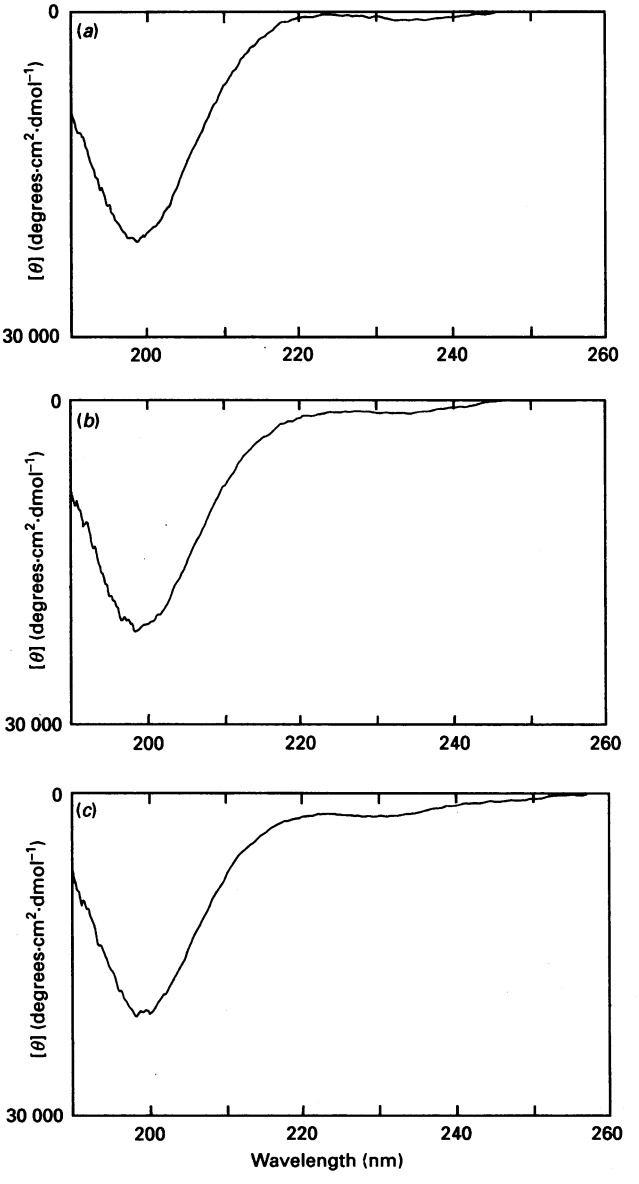

Fig. 1. C.d. spectra of the synthetic peptide

Spectra in u.v. wavelengths are shown for the synthetic peptide SP28 at 0.27 mM in 10 mM-Hepes buffer at pH 6.8 (*a*) and pH 8.0 (*b*). The spectrum of peptide in 10 mM-Hepes (pH 6.8)/50% (v/v) trifluoroethanol (*c*) is also shown. The mean residue ellipticity was calculated from the dichroism using a mean residue weight of 111.

Table 1. Calculated secondary structure content of SP28 in solution

The spectra shown in Fig. 1 were analysed by the CONTIN curve-fitting program. TFE, 50% (v/v) trifluoroethanol in the solvent.

Buffer	Content (%)		
	α-Helix	β-Sheet	Remainder
Hepes (pH 6.8)	6±1.5	44±1.5	50±2.6
Hepes (pH 8.0)	13±1.8	41±1.8	46±3.1
Hepes (pH 6.8)+TFE	9±0.7	37±1.0	54±1.2

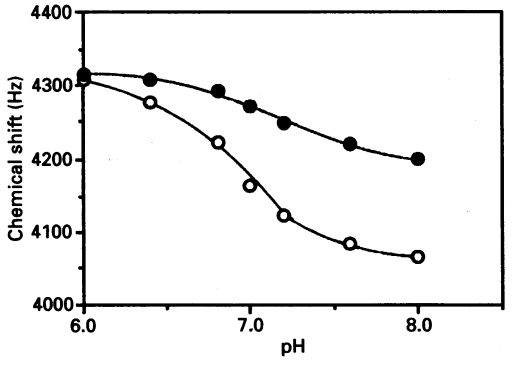

Fig. 2. The chemical shift of the C-2 protons in SP28 is altered by the pH of the buffer

The chemical shift is plotted for each histidine residue in SP28. The peptide was 3 mM in phosphate buffers, as described in the text.

values studied, and no additional α-helix is induced by trifluoroethanol.

Chemical shift of histidine residues

Fig. 2 shows the titration of the chemical shifts of the C-2 histidine protons in the peptide across the pH range 6.0–8.0. The chemical shifts of the C-4 protons altered only slightly. Extrapolation of the (non-horizontal) plateaux at each end of the titration curves gave mid-point pH values of 6.90 and 7.05 for the two C-2 protons, similar to the pK_a values expected for the protonation equilibrium of the free amino acid side chains. It proved impossible, using two-dimensional NOESY experiments, to assign the resonances to protons from one or other of the two histidine residues, indicating the lack of interaction of the side chains with other parts of the molecule in solution.

Phosphorylation of the peptide in the pH range 6.0–8.0

Phosphorylation of the peptide, measured after a short incubation, decreased as the pH of the buffer increased (Fig. 3). Phosphorylation in Hepes buffer was consistently greater than in the other buffers at about pH 7.0, but the general trend was the same in all buffers. There was no definite plateau at the lowest pH values used here, but a plateau of lowest phosphorylation was reached at pH 7.8–8.0. The titration of the initial labelling rate thus corresponded to the protonation of the histidine side chain as shown in Fig. 2 by the chemical shift of the C-2 proton, and supported the idea that two positively charged residues are required *N*-terminal to the phosphorylatable serine.

Fig. 4 shows another experiment in which the initial velocity of the kinase was measured at different pH values in different buffers. The rate of SP28 labelling fell by nearly 3-fold over the pH range used, while the rate with kemptide showed a rise of about 50% (Table 2). The kinase is not inactivated at pH values

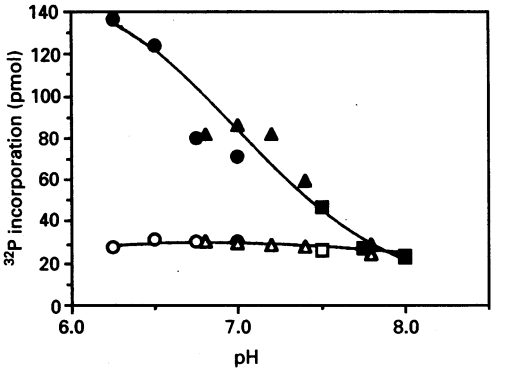

Fig. 3. Phosphorylation of SP28 is decreased by raising the pH of the buffer

The synthetic peptide (●, ▲, ■) was 0.14 mM in the incubation; CAD (○, △, □) was 0.017 mM. Assays were carried out in duplicate incubations, and each point is the mean of two determinations. The assays were quenched after 15 min with cAMP-dependent protein kinase by spotting samples on to squares of P81 or 3MM paper respectively. The incorporation corresponds to 0.1 mol of ^{32}P/mol of peptide at pH 6.2 (0.2 mol of ^{32}P/mol of CAD). The buffers are Bis-Tris (○, ●), Hepes (△, ▲) and Tricine (□, ■).

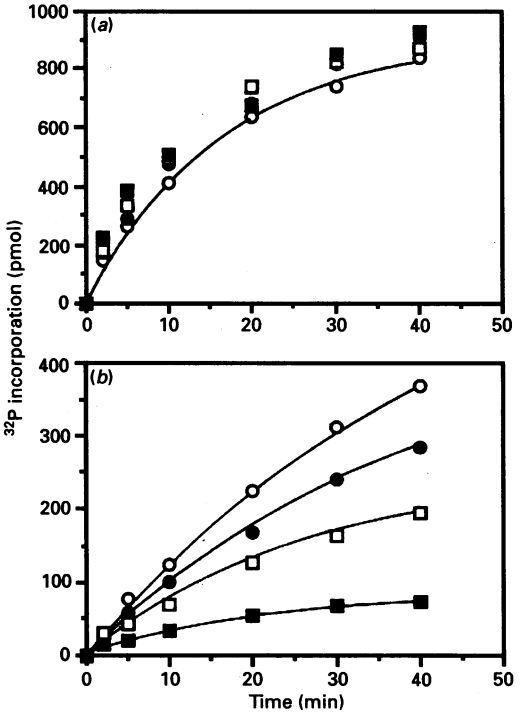

Fig. 4. The initial rate of phosphorylation varies with the pH of the incubation

Incorporation of label was measured over 40 min, taking samples on to P81 paper, as described in the text and in the caption to Fig. 3. The peptides were: (*a*) kemptide (0.17 mM) and (*b*) SP28 (0.14 mM). The buffers were: ○, Bis-Tris, pH 6.5; ●, Hepes, pH 7.0; □, Hepes, pH 7.4; ■, Tricine, pH 8.0. The apparent initial velocities are given in Table 2.

so close to neutrality, so the difference in sensitivity to pH arises from the respective substrates. The two arginine side chains in kemptide will be protonated at all the pH values used in the experiment, while the protonation of the histidine residue has already been shown to vary.

Table 2. Apparent initial velocity of phosphorylation by cAMP-dependent protein kinase

Initial velocities were obtained by taking the tangent to the progress curves in Fig. 4.

		Rate of labelling (pmol of ^{32}P/min)	
Buffer	pH	Kemptide	SP28
Bis-Tris	6.5	71.5	14.5
Hepes	7.0	78.0	10.8
Hepes	7.4	85.0	8.4
Tricine	8.0	103.0	4.0

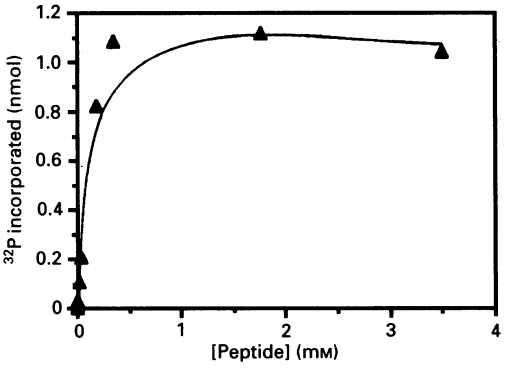

Fig. 5. Concentration-dependence of the phosphorylation of SP28 by cAMP-dependent protein kinase

Incorporation of radioactive label was measured by removing samples on to P81 squares after 5 min of incubation at 30 °C with [γ-^{32}P]ATP and cAMP-dependent protein kinase in 50 mM-Hepes buffer, pH 7.0.

Affinity of the kinase for the synthetic peptide

The initial velocity of labelling by cAMP-dependent protein kinase in Hepes buffer, pH 7.0, was assayed with a range of concentrations of the synthetic peptide (Fig. 5). From a hyperbolic plot of the data, the K_m was calculated as 92 μM. The synthetic peptide was fully labelled after 20 min in these conditions, while kemptide was labelled within the first 8 min. The K_m for kemptide is about 10 μM [10].

Phosphorylation of CAD in the pH range 6.0–8.0

Fig. 3 shows that the phosphorylation of CAD in short incubations was similar at all pH values in the range 6.0–8.0. The concentration of CAD in this experiment was much lower than that of the peptide, but the incorporation of label corresponded to 0.1 mol/mol of peptide at pH 6.2, and 0.2 mol/mol of CAD (or 0.1 mol/site). Comparison of the labelling at pH 7.0 with the equivalent concentration of the peptide (Fig. 5) showed that CAD was phosphorylated 3–4 times more rapidly by the kinase. Tryptic peptides corresponding to the two sites in CAD were analysed by isoelectric focusing of an exhaustive digest followed by autoradiography [8]. These data, which are not shown, demonstrated a similar distribution of label between the two sites at all the pH values studied.

DISCUSSION

This study used a synthetic peptide whose sequence, containing a serine recognized by cAMP-dependent protein kinase, corresponds to 20 residues from the centre of the region that links ATCase and DHOase in CAD [11]. The results show that the peptide has a fully exposed or extended structure in solution, with no tendency to form α-helix, but it can be recognized and phosphorylated by cAMP-dependent protein kinase. Although a two-dimensional n.m.r. analysis of the 20-residue peptide indicated no intramolecular interactions in solution, the extended structure and β-turns predicted [12] for the whole linking region of 120 residues are supported by analysis of the c.d. spectrum.

Model peptide studies have shown that basic side chains near the target residue are important in the recognition and phosphorylation of substrates by protein kinases [13–15]. Most good substrates for cAMP-dependent protein kinase, with kemptide as the paradigm in model compound studies, have basic residues (Arg, Lys) two and three positions *N*-terminal to the serine that is phosphorylated. Enzymes from various organisms [15] or mammalian tissues [16] may tolerate substitution of arginines by lysines, or replacement of the basic residues in the substrate, because of differences in the recognition by their active sites. In the synthetic peptide described here, the labelling of the serine by cAMP-dependent protein kinase clearly titrates in parallel with the protonation of the His residue at the -3 position, replacing the Arg that would be expected in a more conventional substrate. Because the kinase recognizes only residues near the consensus site, protonation of the *N*-terminal His residue is not likely to influence the labelling of this peptide. Kemptide is a shorter synthetic peptide, less likely to have structure in solution, but it is a good substrate for the kinase because the two arginine residues are protonated at all physiological pH values.

Genetic engineering to replace Arg with His at the -3 position in other protein substrates could be used to demonstrate the extent to which the nearby sequence is exposed to the solvent. Many phosphorylation sites in enzymes are found in the *N*- or *C*-terminal regions of the sequences, which might be expected to contain less secondary structure than the core of the protein. If the histidine side chain were fully exposed to solvent, it would titrate with a pK_a of about 6.8 and the site would be readily phosphorylated at the lower pH values, as shown with the synthetic peptide SP28. A histidine side chain would have a pK_a of 5.5–6.0 if it were found in the protein interior, or a pK_a of approx. 8.0 if it were hydrogen-bonded to a carboxylate on a nearby side chain. Each of these pK_a values is outside the range that could reasonably be explored in these experiments without pH-induced unfolding of protein structures, but if a His-containing site is readily phosphorylated in a protein, it is more likely that the latter condition applies. The phosphorylation site should be accessible to kinase, on the surface of the protein, with the nearby serine residue exposed to the solvent while the hydrogen-bonded histidine side chain at the -3 position is maintained in the protonated form at physiological pH.

Comparison with the labelling of the synthetic peptide SP28 provides some hints about the extent to which the equivalent sequence in CAD is exposed to solvent. The linking region in CAD is rich in proline residues and not homologous to the relevant protein sequences [11] for the ATCase or DHOase domains, it is accessible to proteinases that release the ATCase domain as an active enzyme moiety [3], and it is accessible to at least one protein kinase [1]. The hydrophilic nature of the side chains together with the other properties suggests an exposed situation, either along the outside of a folded domain or possibly fully extended in solvent and acting as a spacer arm. A domain of twisted β-sheet has been predicted from the sequence [12] but its orientation between the ATCase and DHOase domains is not known. The length, and hence the structure, of the linking

sequence is crucial to the DHOase activity of expression constructs containing both the enzyme domains [17].

The fact that both sites in CAD can be fully phosphorylated after longer incubations with cAMP-dependent protein kinase shows that they are each accessible to the kinase. Site 1 (Arg-Arg-Leu-Ser) is a classical recognition sequence for cAMP-dependent protein kinase, so it may have a lower K_m than site 2. We found that site 1 was labelled more rapidly than site 2 in intact CAD [2], but it is not possible to measure the K_m for individual sites because CAD solutions more concentrated than 20 μM are not attainable. On the other hand, the data from Figs. 3 and 5 show that CAD is phosphorylated more rapidly than the same concentration of the synthetic substrate at pH 7.0. Neither the distribution of label between the two sites nor the extent of label incorporated into CAD is altered in the pH range 6.0–8.0. Site 2, as found in the intact protein, thus appears to be a better substrate than SP28 for cAMP-dependent protein kinase, and is not sensitive to pH changes. Secondary or tertiary structure, by allowing the hydrogen bonding of the -3 His side chain, may enhance the recognition of the site by the protein kinase. This structure may arise from the alignment of the linking segment on the surface of the ATCase or the DHOase domain, or within a stable discretely folded domain comprising the 120 residues of the linker itself.

I am grateful for assistance from the following people: Dr. Nick Price and Sharon Kelly at the SERC Circular Dichroism Facility, University of Stirling; Dr. Steve Homans, Department of Biochemistry, University of Dundee, for the n.m.r. spectra; Dr. David Campbell and Barry Caudwell, MRC Protein Phosphorylation Unit, University of Dundee, for the synthetic peptide. This work was supported by the Wellcome Trust.

REFERENCES

1. Carrey, E. A. & Hardie, D. G. (1988) Eur. J. Biochem. **171**, 583–588
2. Carrey, E. A., Campbell, D. G. & Hardie, D. G. (1985) EMBO J. **4**, 3735–3742
3. Carrey, E. A. (1986) Biochem. J. **236**, 327–335
4. Cohen, P. (1985) Eur. J. Biochem. **151**, 439–448
5. Coleman, P. F., Suttle, D. P. & Stark, G. R. (1977) J. Biol. Chem. **252**, 6379–6385
6. Reimann, E. M. & Beham, R. A. (1983) Methods Enzymol. **99**, 51–55
7. Roskoski, R. (1983) Methods Enzymol. **99**, 3–6
8. Carrey, E. A. & Hardie, D. G. (1986) Anal. Biochem. **158**, 431–435
9. Provencher, S. W. & Glöckner, J. (1981) Biochemistry **20**, 33–37
10. Kemp, B. E., Graves, D. A. & Krebs, E. G. (1976) Fed. Proc. Fed. Am. Soc. Exp. Biol. **35**, 1384
11. Simmer, J. P., Kelly, R. E., Scully, J. L., Grayson, D. R., Rinker, A. G., Bergh, S. T. & Evans, D. R. (1989) Proc. Natl. Acad. Sci. U.S.A. **86**, 4382–4386
12. Williams, N. K., Simpson, R. J., Moritz, R. L., Peide, Y., Crofts, L., Minasian, E., Leach, S. J., Wake, R. G. & Christopherson, R. I. (1990) Gene **94**, 283–288
13. Pearson, R. B., Woodgett, J. R., Cohen, P. & Kemp, B. E. (1985) J. Biol. Chem. **260**, 14471–14476
14. House, C., Wettenhall, R. E. H. & Kemp, B. E. (1987) J. Biol. Chem. **262**, 772–777
15. Denis, C. L., Kemp, B. E. & Zoller, M. J. (1991) J. Biol. Chem. **266**, 17932–17935
16. Takhar, S. & Munday, M. R. (1992) Biochem. Soc. Trans. **20**, 307S
17. Musmanno, L. A., Maley, J. A. & Davidson, J. N. (1991) Gene **99**, 211–216

Received 9 March 1992/21 April 1992; accepted 20 May 1992