

Complete nucleotide sequence of the gene encoding the regulatory subunit of 3',5'-cyclic AMP-dependent protein kinase from the yeast *Saccharomyces cerevisiae*

Riyo Kunisawa, Trisha N. Davis, Mickey S. Urdea¹ and Jeremy Thorner*

Department of Biochemistry, University of California, Berkeley, CA 94720 and ¹Chiron Research Laboratories, Chiron Corporation, Emeryville, CA 94608, USA

Submitted December 1, 1986

The amino acid sequence of twenty residues from the N-terminus of the regulatory subunit of cAMP-dependent protein kinase purified from the yeast *Saccharomyces cerevisiae* was reported several years ago (1). We prepared three corresponding oligonucleotides: one 23-base mixed sequence (128-fold degenerate) probe and two unique (45- and 60-base) probes that accommodate the apparent codon usage bias of *S. cerevisiae* (2,3). The three synthetic probes were used to screen a yeast genomic DNA library in the vector YCp50 (kindly provided by M. Rose). Two plasmids were obtained that, by restriction endonuclease cleavage site mapping, contained a common segment of genomic DNA. Digests of this common region (BglII, EcoRV, MspI, PstI, Sau3A, SphI, XbaI) were sub-cloned into M13mp18 and M13mp19 (4) and were sequenced on both strands by standard dideoxy chain termination methods (5). The open reading frame present encodes a 416-residue polypeptide that, aside from the initiator methionine, commences with a perfect match (underlined) to the N-terminal sequence of the purified protein. Hybridization of the cloned gene to Southern blots of genomic DNA digests (6), to whole yeast chromosomes separated by orthogonal field-alternation gel electrophoresis (7), and to polyA⁺ RNA species fractionated by gel electrophoresis (8), demonstrated that the gene is a unique single-copy locus, resides on chromosome IX, and encodes a transcript of 1595 bases. Two mutations (bcy1 and sral) known to affect the level of regulatory subunit (9,10) also map on chromosome IX and probably represent lesions in REG1. Computer comparison (ALIGN program of Intelligenetics, Inc.) indicates that the C-terminal two-thirds of the yeast REG1 protein shares 40% identity with both mammalian type I and type II regulatory subunits (11,12). Homology is especially conserved around the residues known to be involved in cAMP binding (asterisks). Like the type II subunit, the predicted yeast protein has a potential phosphorylation site (P) in the domain corresponding to the "hinge region" thought to be the major site of interaction with the catalytic subunit (13). The yeast REG1 protein can be phosphorylated in vitro by both the yeast (14) and bovine (1) catalytic subunit and appears to be phosphorylated in vivo (15).

110

GCTTCTOCTGA CTACTCAAGC GATGCGTATC ATCAACATCCA TTACAAGCATT TTCACCCATA CATAAATAGG AAGATCATTT ACAACAAGCA GATTATTTCCT AAAAGACAAC
 1 5 10 15 20 25
 Met Val Ser Ser Leu Pro Lys Glu Ser Gln Ala Gln Leu Gln Leu Phe Gln Asn Glu Ile Asn Ala Ala Asn Pro Ser
 AGTAAGAATA AACG ATG GTA TCT TCT TTG CCG AAG GAA TCG CAA GCC GAA TTG CAA CTG TTC CAG AAC GAA ATC AAC GCC GCT AAT CCG TCC
 30 35 40 45 50 55
 Asp Phe Leu Gln Phe Ser Ala Asn Tyr Phe Asn Lys Arg Leu Glu Gln Gln Arg Ala Phe Leu Lys Ala Arg Glu Pro Glu Phe Lys Ala
 GAC TTT CTT CAG TTC TCC GCC AAC TAT TTC AAT AAA AGC CTG GAA CAA CAG AGA GCG TTC CTC CAG GCC AGG GAG CCT GAA TTT AAG GCA
 60 65 70 75 80 85
 Lys Asn Ile Val Leu Phe Pro Glu Pro Glu Glu Ser Phe Ser Arg Pro Gln Ser Ala Gln Ser Gln Ser Arg Ser Arg Ser Ser Val Met
 AAG AAC ATT GTT CTA TTT CCG GAA CCA GAG GAG TCA TTT TCC AGA CCT CAA TCA GCT CAA TCT CAA TCA AGA TCC AGA TCG AGT GTT ATG
 90 95 100 105 110 115
 Phe Lys Ser Pro Phe Val Asn Glu Asp Pro His Ser Asn Val Phe Lys Ser Gly Phe Asn Leu Asp Pro His Glu Gln Asp Thr His Gln
 TTC AAA TCC CCC TTT GTG AAC GAG GAC CCA CAC TCC AAC CTG TTT AAA ACT GGG TTT AAT TTA GAC CCG CAC GAA CAG GAC ACT CAC CAG

	120		125		130		135		140		P
Gln Ala Gln	Glu Gln Gln	Gln His Thr	Arg Glu Lys Thr	Ser Thr Pro	Pro Leu Pro	Met His Phe Asn	Ala Gln Arg	Arg Thr Ser	Val		
CAA GCA CAG	GAA GAA CAA	CAG CAT ACT	AGA GAA AAG ACA	TCA ACT CCT	CCA CTC CCA	ATG CAC TTC AAC	GCC CAA AGG	CGT ACT TCT	GTT		
	150		155		160		165		170		175
Ser Gly Glu	Thr Leu Gln	Pro Asn Asn	Phe Asp Asp Trp	Thr Pro Asp	His Tyr Lys	Glu Lys Ser	Glu Gln Gln	Leu Gln Arg	Leu Glu		
AGT GGT GAG	ACC TTA CAA	CAA AAT TTT	GAC GAT TGG ACT	CCA CAT CAC	TAT AAG GAA	AAG TCC GAG	CAG CAA TTG	CAA AGA CTC	GAA		
	180		185		190		195		200		205
Lys Ser Ile	Arg Asn Asn	Phe Leu Phe	Asn Lys Leu Asp	Ser Asp Ser	Lys Arg Leu	Val Ile Asn	Cys Leu Glu	Glu Lys Ser	Val Pro		
AAA TCG ATC	CGT AAT AAC	TTT CTG TTC	TTC AAC AAG	CTG GAT TCC	GAC TCA AAA	AGG CTG	GTC ATA AAT	TGT CTG	GAG GAG	AAG	TCC
	210		215		220		225		230		235
Lys Gly Ala	Thr Ile Ile	Lys Gln Gly	Asp Gln Gly Asp	Tyr Phe Tyr	Val Val Glu	Lys Gly Thr	Val Asp Phe	Tyr Val Asn	Asp Asn		
AAA GGT GCT	ACG ATA ATC	AAG CAA GGT	GAC CAA GGG	GAC TAC TTC	TAT GTC GTC	GAA AAG	GGT ACT GTT	GAC TTC	TAC TAC	AAC	GAC
	240		245	*	*	*	*		260		265
Lys Val Asn	Ser Ser Gly	Pro Gly Ser	Ser Phe Gly Glu	Leu Ala Leu	Met Tyr Asn	Ser Pro Arg	Ala Ala Ala	Thr Val Val	Val Ala Thr	Ser	
AAG GTC AAC	TCT TCC	GGG CCA	GGC TCC	AGT TTC	GGG GAA	CTT GCT	CTT ATG	TAC AAC	AGC	CGT	GCT
	270		275		280		285		290		295
Asp Cys Leu	Leu Trp Ala	Leu Asp Arg	Leu Thr Phe	Arg Lys Ile	Leu Leu Gly	Ser Ser Phe	Lys Lys Arg	Leu Met Tyr	Asp Asp	Leu	
GAC TGT TTG	TTG TCG	GCT CTA	GAC AGG	CTC ACC	TTC AGA	AAA ATA	CTT TTG	GGC AGC	TCT TTC	AAG	AGA
	300		305		310		315		320		325
Leu Lys Ser	Met Pro Val	Leu Lys Ser	Leu Thr Thr	Tyr Asp Arg	Ala Lys Leu	Ala Asp Ala	Leu Asp Thr	Lys Ile Tyr	Gln Pro	Gly	
TTG AAG AGC	ATC CCA	GTT TTG	AAG ACT	TTG ACT	AGG TAC	GAC GCT	GCC AAA	CTT GGC	GAT CCA	CTG	GAT
	330		335		340		345		350		355
Glu Thr Ile	Ile Arg Glu	Gly Asp Gln	Gly Glu Asn	Phe Tyr Leu	Ile Glu Tyr	Gly Ala Val	Asp Val Ser	Lys Lys Gly	Gln Gln	Gly	Val
GAA ACA ATC	ATT CAC	GAG GGT	GAT CAA	GGG GAG	AAC TTT	TAT TTA	ATT GAC	TAC GGA	GCT GTG	GAC	TCT
	360		365	*	*	*	*		380		385
Ile Asn Lys	Leu Lys Asp	His Asp Tyr	Phe Gly Glu	Val Ala Leu	Leu Asn Asp	Leu Pro Arg	Gln Ala Thr	Val Thr Ala	Thr Lys	Arg	
ATA AAT AAA	CTC AAA	GAC CAT	GAT TAT	TTC GGT	GAA GTC	GCC TTG	CTA AAC	GAT TTG	CCG	ACA	GAG
	390		395		400		405		410		415
Thr Lys Val	Ala Thr Leu	Gly Lys Ser	Gly Phe Gln	Arg Leu	Leu Gly Pro	Ala Val Asp	Val Leu	Lys Leu	Asn Asp	Pro Thr	Arg
ACC AAA GTT	GCC ACA	TTG GGG	AAA AGT	GCT TTT	CAA GCT	TTA CTG	GCT CCT	GCA GTA	GAC	GTA	TTA
	420		425		430		435		440		445
TAA GTAAAA	AGGGGAAGCC	ATCTTAATCT	CACATGAATT	TCCTTCTCT	TTTCTTCCT	CTTCGTCC	CTTTTCTTT	CTTCCTCT	CAAAGTCTAC	GTAATATAA	1481

*To whom correspondence and reprint requests should be addressed at Department of Biochemistry, University of California, Berkeley, Room 401, Biochemistry Building, Berkeley, CA 94720, USA

- Hixson, C.S. and Krebs, E.G. (1980) *J. Biol. Chem.* **255**, 2137-2145.
- Davis, T.N. and Thorner, J. (1986) *Mets. Enzymol.* **139**, 248-262.
- Maruyama, T., Gojobori, T., Aota, S.-I. and Ikemura, T. (1986) *Nucleic Acids Res.* **14** (Suppl.), 151.
- Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) *Gene* **33**, 103-119.
- Biggin, M.D., Gibson, T.J. and Hong, G.F. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 3963-3965.
- Southern, E.M. (1975) *J. Molec. Biol.* **98**, 503-517.
- Carle, G.F. and Olson, M.V. (1984) *Nucleic Acids Res.* **12**, 5647-5664.
- Thomas, P.S. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 5201-5205.
- Matsumoto, K., Uno, I., Oshima, Y. and Ishikawa, T. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 2355-2359.
- Cannon, J.F., Gibbs, J.B. and Tatchell, K. (1986) *Genetics* **113**, 247-264.
- Takio, K., Smith, S.B., Krebs, E.G., Walsh, K.A. and Titani, K. (1982) *Proc. Natl. Acad. Sci. USA* **79**, 2544-2548.
- Titani, K., Sasagawa, T., Ericsson, L.H., Kumar, S., Smith, S.B., Krebs, E.G. and Walsh, K.A. (1984) *Biochemistry* **23**, 4193-4199.
- Hoppe, J. (1985) *Trends Biochem. Sci.* **10**, 29-31.
- Wingender-Drissen, R. (1983) *FEBS Letts.* **163**, 28-32.
- Baum, P.R. (1981) Ph.D. Thesis, University of California, Berkeley, California, USA, 209pp.