The EMBO Journal vol.4 no.10 pp.2643–2648, 1985

Nucleotide sequences of *STE2* and *STE3*, cell type-specific sterile genes from *Saccharomyces cerevisiae*

N.Nakayama[1,2], A.Miyajima[1] and K.Arai[1]

[1]DNAX Research Institute of Molecular and Cellular Biology, 901 California Avenue, Palo Alto, CA 94304, USA, and [2]Department of Chemistry, The Institute of Medical Science, University of Tokyo, Takanawa, Minato-ku, Tokyo, Japan

Communicated by A.Kornberg

The nucleotide sequences of *STE2* and *STE3*, cell type-specific sterile genes of *Saccharomyces cerevisiae*, were determined; major open reading frames encode 431 and 470 amino acids, respectively. *STE2* and *STE3* proteins seem to be folded in a similar fashion and are likely to be membrane-bound. Both consist of seven hydrophobic segments in each NH_2-terminal region with a long hydrophilic domain in each COOH-terminal region. However, the two putative gene products do not exhibit extensive sequence homology. The *STE2* protein has no obvious hydrophobic signal peptide; the NH_2 terminus of the *STE3* protein might serve as a signal peptide. The *STE2* transcript, 1.7 kb, was detected in *MATa* strains but not in *MATα* strains, while the *STE3* transcript, also 1.7 kb, was detected only in *MATα* cells. In *STE2*, two canonical TATA sequences are located 18 and 27 bp upstream of the mRNA start site, which has been mapped 32 bp before the initiator ATG codon, while *STE3* contains a similar sequence (TATAGA), which is preceded by a long AT sequence, 140 bp upstream of the initiator ATG codon. Transcription of *STE2* in *a* cells seems to be enhanced by exogenous α-factor.

Key words: DNA sequence/molecular cloning/receptor/*S. cerevisiae*/mating pheromone

Introduction

During the mating process of yeast *Saccharomyces cerevisiae*, the cell division cycle is regulated by peptide pheromones called mating factors (reviewed by Thorner, 1981; Sprague *et al.*, 1983a). Two different types of haploid cells, α and *a*, produce α-factor and *a*-factor, respectively. Each factor acts on the opposite type of haploid cells, that is, α-factor interacts with *a* cells and *a*-factor acts on α cells, leading to cell cycle arrest of the target cells at the G1 phase. Arrested α and *a* cells can then fuse to form diploid cells which are no longer sensitive to either of the mating factors. In addition, these factors induce cell-surface agglutinin formation that facilitates aggregation of opposite cell types. Alpha-factor also elicits 'Schmoo' formation, an aberrant shape of the target cell, which may provide the fusion site with an α cell. These responses are thought to be mediated by binding of *a*- or α-factor to specific receptors on the surface of the α or *a* cells. At least eight genes unlinked to the *MAT* locus (*STE2*, *STE4*, *STE5*, *STE7*, *STE8*, *STE9*, *STE11* and *STE12*) are involved in the response to α-factor in *a* cells (MacKay and Manney, 1974a, 1974b; Manney and Woods, 1976; Hartwell, 1980). Mutations in any one of these genes can prevent the expression of α-factor-induced phenotypes such as growth arrest at G1 phase, agglutinin induction and induction of *a*-factor. Alpha-factor-resistant mutants are also sterile. Similar results were obtained in α cells, except that the *ste2* mutant affects α-factor-inducible phenotype in an *a* cell-specific manner, while other functions for mating, such as production of *a*-factor, are not affected. Furthermore, the observation that the *MATa ste2* mutant is unable to bind α-factor (Jenness *et al.*, 1983) strongly suggests that *STE2*

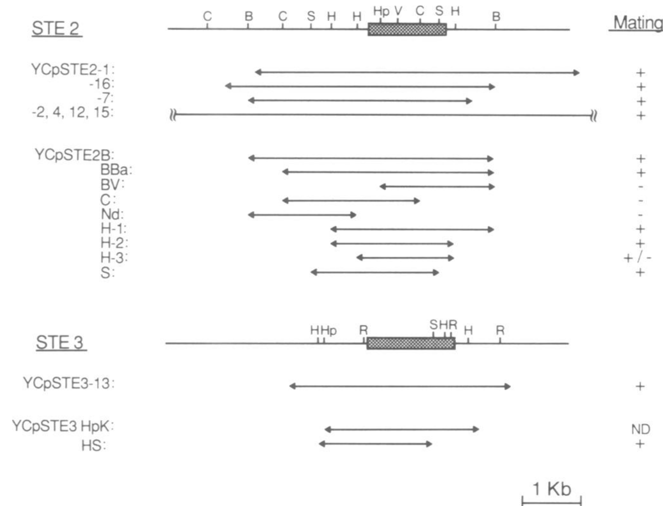

Fig. 1. Restriction maps and deletion analyses of the DNA segments containing *STE2* and *STE3*. Restriction enzyme sites *Bam*HI (B), *Cla*I (C), *Eco*RI (R), *Eco*RV (V), *Hin*dIII (H), *Hpa*I (Hp) and *Sal*I (S) are drawn to physical scale; the open reading frames for (A) *STE2* and (B) *STE3* are indicated by boxes. All the clones obtained from the initial screening and the deletion analyses were listed. Each arrow indicates the position of DNA segments cloned in the YCpN1 vector. Complementing ability of each DNA fragment is listed. The 1.6-kb *Hin*dIII fragment of *STE2* exhibits weak complementation (YCpSTE2H-3). ND: not determined.

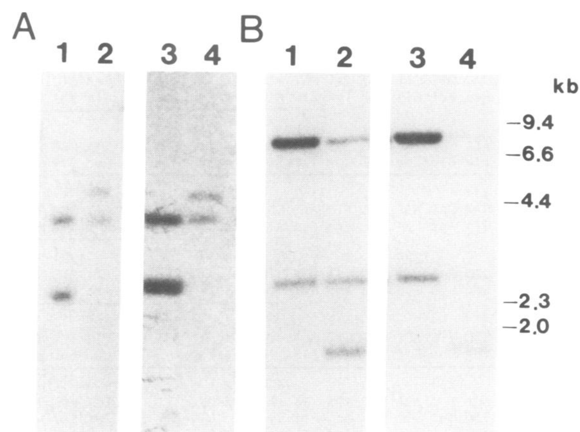

Fig. 2. Genomic Southern blotting of wild-type and mutated cells. Genomic DNAs isolated from wild-type *a* (YP45) cells, α (YP47) cells and from cells in which *STE2* had been disrupted with the *URA3* fragment (YAM10 and NNY111) were digested with restriction enzymes **(A)** *Cla*I and **(B)** *Pst*I, separated on a 0.8% agarose gel and transferred to nitrocellulose membrane. The DNAs were probed with nick-translated 1.6-kb *Hin*dIII fragment containing the *STE2* coding region. **Lane 1:** YP45; **lane 2:** YAM10 (*MATa ste2:URA3*); **lane 3:** YP47; **lane 4:** NNY111 (*MATα ste2:URA3*).

© IRL Press Limited, Oxford, England.

may encode the *a* cell surface-receptor specific for α-factor. Likewise, another mutation, *ste3*, leads to *a* cell-specific sterility, probably by a lack of response to *a*-factor (MacKay and Manney, 1974a, 1974b). *STE3* has been cloned (Sprague *et al.*, 1983b) and transcription of the gene is known to be inducible by *a*-factor in α cells (Hagen and Sprague, 1984). *STE3* might encode a receptor for *a*-factor.

Receptor-mediated transmembrane signalling is of key importance in understanding the mechanism of the mating factor-induced G1 arrest. Since little is known of the structure and function of the mating factor receptors, *STE2* and *STE3* were cloned and the primary structures deduced from their nucleotide sequences were compared. Our results strongly indicate that both genes appear to encode integral membrane proteins, which may be involved in the response to mating factors on the cell membrane.

Results

Cloning and physical mapping of STE2 and STE3

We have isolated *STE2* and *STE3* by complementation of the mating defects of *ste2* and *ste3* mutants. Strains NNY110 and NNY128 were transformed with a plasmid library carrying an average 8-kb insert derived from chromosomal DNA of wild-type haploid cells (DBY746) in the YCpN1 vector. Plasmids carrying *LEU2* or *HIS3* were isolated at a frequency of 1 per 7500 Trp^+ transformants. Approximately 4.5×10^4 Trp^+ transformants were screened for their ability to mate with the opposite type of cells by a replica plating procedure (MacKay, 1983). In this manner, 11 independent clones which complemented *ste2* were isolated, seven of which were analyzed further. One of the five isolated clones which complemented *ste3* was subjected to further analysis. In all cases, the ability to complement the mating defects co-segregate with the *TRP1* marker on the vector, indicating that the genomic DNA segment carried on the vector does, in fact, complement the mutational phenotype.

Seven plasmid DNAs for *STE2* recovered in an *Escherichia coli* strain, MC1061, were re-introduced into NNY110 and NNY124 cells which carry two different *ste2* mutations. All of them complemented the *ste2-3*^{ts} mutation, as well as the original *ste2-1* mutation. Physical mapping of these seven clones (YCpSTE2-1, -2, -4, -7, -12, -15, -16) with restriction enzymes revealed that they share a 2.3-kb *Cla*I restriction fragment and a 1.6-kb *Hind*III fragment (Figure 1). Therefore, the inserts are probably derived from the same locus on the chromosome.

The minimum region required to complement *ste2-1* and *ste2-3*^{ts} mutations was mapped more precisely by subcloning various restriction fragments into the YCp vector (Figure 1). YCpSTE2H-2, the plasmid bearing the 2.0-kb partial *Hind*III fragment, complemented the mating defects of both mutant strains, while YCpSTE2H-3 carrying shorter 1.6-kb *Hind*III fragment derived from the 2.0-kb region only partially complemented the mutations. A 2.2-kb *Sal*I fragment in YCpSTE2S, which overlaps with the 2.0-kb *Hind*III fragment, also complemented

A

```
        10         20         30         40         50         60         70         80         90        100        110        120        130        140
TTTTTAATAC ACCAAAGATT CAAGATAAGA GCATAGAACG AACTGTAGAA TAGTCCGGAT ATGTTATCCA ATGCCTGCCA AAATGCATTG TCACACGCTG TAGTGCTCGA ATAGGTGTTG CAATCCGTCA ATATACGTCT

       150        160        170        180        190        200        210        220        230        240        250        260        270        280
TGCTCTGTGG GTAAATGTCT CGTGCATTAA GACAGGCTAG TATAAACGAG AAGAAGTATC CTGCTTTGCA ATGAAACAAT AGTATCCGCT AAGAATTTAA GCAGGCCAAC GTCCATACTG CTTAGGACCT GTGCCTGGCA

       290        300        310        320        330        340        350        360        370        380        390        400        410        420
AGTCGCAGAT TGAAGTTTTT TCAACCATGT AAATTTCCTA ATTGGGTAAG TACATGATGA AACACATATG AAGAAAAAAG CTTTCCTACA TATTCAAGAT TTTTTTCTGT GGGTGGAATA CTATTTAAGG AGTGCTATTA

       430        440        450        460        470        480        490        500   **   510        520        530                     549
GTATCTTATT TGACTTCAAA GCAATACGAT ACCTTTTCTT TTCACCTGCT CTGGCTATAA TTATAATTGG TTACTTAAAA ATGCACCGTT AAGAACCATA TCCAAGAATC AAAA ATG TCT GAT GCG GCT CCT
                                                                                                                             MET Ser Asp Ala Ala Pro

            564             579             594             609             624             639             654
TCA TTG AGC AAT CTA TTT TAT GAT CCA ACG TAT AAT CCT GGT CAA AGC ACC ATT AAC TAC ACT TCC ATA TAT GGG AAT GGA TCT ACC ATC ACT TTC GAT GAG TTG CAA GGT TTA
Ser Leu Ser Asn Leu Phe Tyr Asp Pro Thr Tyr Asn Pro Gly Gln Ser Thr Ile Asn Tyr Thr Ser Ile Tyr Gly Asn Gly Ser Thr Ile Thr Phe Asp Glu Leu Gln Gly Leu

669             684             699             714             729             744             759             774
GTT AAC AGT ACT GTT ACT CAG GCC ATT ATG TTT GGT GTC AGA TGT GGT GCA GCT GCT TTG ACT TTG ATT GTC ATG TGG ATG ACA TCG AGA AGC AGA AAA ACG CCG ATT TTC ATT
Val Asn Ser Thr Val Thr Gln Ala Ile MET Phe Gly Val Arg Cys Gly Ala Ala Ala Leu Thr Leu Ile Val MET Trp MET Thr Ser Arg Ser Arg Lys Thr Pro Ile Phe Ile
                                                     M1
        789             804             819             834             849             864             879             894
ATC AAC CAA GTT TCA TTG TTT TTA ATC ATT TTG CAT TCT GCA CTC TAT TTT AAA TAT TTA CTG TCT AAT TAC TCT TCA GTG ACT TAC GCT CTC ACC GGA TTT CCT CAG TTC ATC
Ile Asn Gln Val Ser Leu Phe Leu Ile Ile Leu His Ser Ala Leu Tyr Phe Lys Tyr Leu Leu Ser Asn Tyr Ser Ser Val Thr Tyr Ala Leu Thr Gly Phe Pro Gln Phe Ile

                    909      M2     924             939             954             969             984             999
AGT AGA GGT GAC GTT CAT GTT TAT GGT GCT ACA AAT ATA ATT CAA GTC CTT CTT GTG GCT TCT ATT GAG ACT TCA CTG GTG TTT CAG ATA AAA GTT ATT TTC ACA GGC GAC AAC
Ser Arg Gly Asp Val His Val Tyr Gly Ala Thr Asn Ile Ile Gln Val Leu Leu Val Ala Ser Ile Glu Thr Ser Leu Val Phe Gln Ile Lys Val Ile Phe Thr Gly Asp Asn

   1014            1029            1044            1059            1074  M3        1089            1104            1119
TTC AAA AGG ATA GGT TTG ATG CTG ACG TCG ATA TCT TTC ACT TTA GGG ATT GCT ACA GTT ACC ATG TAT TTT GTA AGC GCT GTT AAA GGT ATG ATT GTG ACT TAT AAT GAT GTT
Phe Lys Arg Ile Gly Leu MET Leu Thr Ser Ile Ser Phe Thr Leu Gly Ile Ala Thr Val Thr MET Tyr Phe Val Ser Ala Val Lys Gly MET Ile Val Thr Tyr Asn Asp Val

            1134            1149            1164  M4       1179            1194            1209            1224
AGT GCC ACC CAA GAT AAA TAC TTC AAT GCA TCC ACA ATT TTA CTT GCA TCC TCA ATA AAC TTT ATG TCA TTT GTC CTG GTA GTT AAA TTG ATT TTA GCT ATT AGA TCA AGA AGA
Ser Ala Thr Gln Asp Lys Tyr Phe Asn Ala Ser Thr Ile Leu Leu Ala Ser Ser Ile Asn Phe MET Ser Phe Val Leu Val Val Lys Leu Ile Leu Ala Ile Arg Ser Arg Arg

1239            1254            1269            1284            1299 M5         1314            1329            1344
TTC CTT GGT CTC AAG CAG TTC GAT AGT TTC CAT ATT TTA CTC ATA ATG TCA TGT CAA TCT TTG TTG GTT CCA TCG ATA ATA TTC ATC CTC GCA TAC AGT TTG AAA CCA AAC CAG
Phe Leu Gly Leu Lys Gln Phe Asp Ser Phe His Ile Leu Leu Ile MET Ser Cys Gln Ser Leu Leu Val Pro Ser Ile Ile Phe Ile Leu Ala Tyr Ser Leu Lys Pro Asn Gln

    1359            1374            1389            1404          M6 1419            1434            1449            1464
GGA ACA GAT GTC TTG ACT ACT GTT GCA ACA TTA CTT GCT GTA TTG TCT TTA CCA TTA TCA TCA ATG TGG GCC ACG GCT GCT AAT AAT GCA TCC AAA ACA AAC ACA ATT ACT TCA
Gly Thr Asp Val Leu Thr Thr Val Ala Thr Leu Leu Ala Val Leu Ser Leu Pro Leu Ser Ser MET Trp Ala Thr Ala Ala Asn Asn Ala Ser Lys Thr Asn Thr Ile Thr Ser

            1479            1494            1509 M7         1524            1539            1554            1569
GAC TTT ACA ACA TCC ACA GAT AGG TTT TAT CCA GGC ACG CTG TCT AGC TTT CAA ACT GAT AGT ATC AAC AAC GAT GCT AAA AGC AGT CTC AGA AGT AGA TTA TAT GAC CTA TAT
Asp Phe Thr Thr Ser Thr Asp Arg Phe Tyr Pro Gly Thr Leu Ser Ser Phe Gln Thr Asp Ser Ile Asn Asn Asp Ala Lys Ser Ser Leu Arg Ser Arg Leu Tyr Asp Leu Tyr

    1584            1599            1614            1629            1644            1659            1674            1689
CCT AGA AGG AAG GAA ACA ACA TCG GAT AAA CAT TCG GAA AGA ACT TTT GTT TCT GAG ACT GCA GAT GAT ATA GAG AAA AAT CAG TTT TAT CAG TTG CCC ACA CCT ACG AGT TCA
Pro Arg Arg Lys Glu Thr Thr Ser Asp Lys His Ser Glu Arg Thr Phe Val Ser Glu Thr Ala Asp Asp Ile Glu Lys Asn Gln Phe Tyr Gln Leu Pro Thr Pro Thr Ser Ser

            1704            1719            1734            1749            1764            1779            1794
AAA AAT ACT AGG ATA GGA CCG TTT GCT GAT GCA AGT TAC AAA GAG GGA GAA GTT GAA CCC GTC GAC ATG TAC ACT CCC GAT ACG GCA GCT GAT GAG GAA GCC AGA AAG TTC TGG
Lys Asn Thr Arg Ile Gly Pro Phe Ala Asp Ala Ser Tyr Lys Glu Gly Glu Val Glu Pro Val Asp MET Tyr Thr Pro Asp Thr Ala Ala Asp Glu Glu Ala Arg Lys Phe Trp

1809            1824                 1840       1850       1860       1870       1880       1890       1900       1910       1920       1930       1940
ACT GAA GAT AAT AAT AAT TTA TGA TCAAAATTTA CGGCTTTGAA AAAGTAATTT CGTGACCTTC GGTATAAGGT TACTACTAGA TTCAGGTGCT CATCAGATGC ACCACATTCT CTATAAAAAA AAATGGTATC
Thr Glu Asp Asn Asn Asn Leu  .

       1950       1960       1970       1980       1990       2000       2010       2020
TTTCTTATTT GATAATATTT AAACTCCTTT ACATAATAAA CATCTCGTAA GTAGTGGTAG AAACCACCTT TGCTTTTACG AGTCAA
```

B

```
        10         20         30         40         50         60         70         80         90        100        110        120        130        140
AGCTTTACGG ATTTTTTCAT CTCTTGTTCA TTATTGTAAT GCAGGTTAAC GTTTTCTCCT TCTTTACATG TTTAATATAT TCCAAGTTAC CTAAGAGGTG TACGATATTT TTTTCTTTTA TATATATGAT TTTCTATTCA

       150        160        170        180        190        200        210        220        230        240        250        260        270        280
TTTTTTAGTT TTTTTTGATA CATAAGCGAA TCGCACATTG CGCAACTTCA ATTTGTTGAT TCGCCAAAGT ATTCTTACCA TAAAACAACC ATTCGTTGCT TTACCCTTTC GTAGTCATTT ACCGTGATAA CCATAATCAG

       290        300        310        320        330        340        350        360        370        380        390        400        410        420
AAACTTATTA TTTCAGCCTA GTAGACCGGC CAAGCAGGCC TTGTAATGTT TCTCTTGATT GCTTGAATCT TTTAAGCAGC CAAATCTTTC CAAAAAAATG CAATTATCAG AACAAAACTA TTTAAGGTGA CTTCTCCGTA

       430        440        450        460        470        480        490        500        510        520        530        540        550        560
TTTACACCAC CAGAAGCGTT CTGGCTCCCC TTTTCTCTAA ACGTTAAACA TTTTACAATT GAAATGTTAC CAATCCTATA TTATTGTACC ACATTGCCAG ATTTATGAAC TCTGGGTATG GGGTGCTAAT TTTCGTTAGA

       570        580        590        600        610        620        630        640        650        660        670        680        690        700
AGCGCTGGTA CAATTTTCTC TGTCATTGTG ACACTAATTA GGAAACTTCT CGACCATCAA TGTGTAAATG AAGGAATAAT GGCGGAAACT TTGAAACTTT GTCAATAATT GCATCATTGG ATGCGTTTCA TTTGGCCGTT

       710        720        730        740        750        760        770        780        790        800        810        820        830        840
ATCACGGAGA GGCAGAGTTC TCTCCACAAT TTGGGCAGAA GTCTTTTGAA AAGACATATA TATATATATA TATATGTATA TGAGTGGATG CTTAAGGTAA GAATAATTTC TGAATTCCCA AGTATTCATT TTGTGCAGTA

       850        860        870        880        890        900        910        920            937                   952                   967
TTCACATATT CTATTTTATT GCTTTTTAAC TTTAGAGGCA ATTAAATTTG TGTAGGAAAG GCAAAATACT ATCAAAATTT TC ATG TCA TAC AAG TCA GCA ATA ATA GGG CTT TGT TTG CTA GCT GTG
                                                                                           MET Ser Tyr Lys Ser Ala Ile Ile Gly Leu Cys Leu Leu Ala Val

                982                 997                1012                1027                1042                1057       N1       1072
ATA CTA TTA GCT CCC CCT TTA GCA TGG CAT TCA CAT ACC AAG AAT ATT CCA GCA ATC ATT TTG ATA ACA TGG CTT CTT ACA ATG AAT TTA ACG TGT ATT GTA GAT GCG GCA ATA
Ile Leu Leu Ala Pro Pro Leu Ala Trp His Ser His Thr Lys Asn Ile Pro Ala Ile Ile Leu Ile Thr Trp Leu Leu Thr MET Asn Leu Thr Cys Ile Val Asp Ala Ala Ile

1087                1102                1117                1132                1147                1162 N2             1177                1192
TGG AGT GAC GAC GAT TTC CTC ACG AGA TGG GAT GGT AAA GGT TGG TGT GAT ATT GTC ATC AAG TTG CAA GTT GGT GCG AAT ATT GGC ATA TCA TGT GCC GTT ACC AAC ATC ATT
Trp Ser Asp Asp Asp Phe Leu Thr Arg Trp Asp Gly Lys Gly Trp Cys Asp Ile Val Ile Lys Leu Gln Val Gly Ala Asn Ile Gly Ile Ser Cys Ala Val Thr Asn Ile Ile

        1207                1222                1237                1252                1267      N3      1282                1297
TAC AAC TTG CAT ACA ATT TTG AAG GCA GAT AGT GTT TTA CCG GAT CTT TCA TCA TGG ACG AAA ATC GTC AAG GAC CTT GTG ATT AGC TTG TTC ACA CCT GTC ATG GTC ATG GGA
Tyr Asn Leu His Thr Ile Leu Lys Ala Asp Ser Val Leu Pro Asp Leu Ser Ser Trp Thr Lys Ile Val Lys Asp Leu Val Ile Ser Leu Phe Thr Pro Val MET Val MET Gly

1312            1327                1342                1357                1372                1387                1402                1417       N4
TTT TCA TAT CTG TTA CAA GTA TTT AGA TAT GGT ATC GCT CGC TAC AAC GGT TGC CAA AAC TTA TTG TCT CCG ACG TGG ATT ACC ACT GTT TTG TAT ACC ATG TGG ATG CTT ATA
Phe Ser Tyr Leu Leu Gln Val Phe Arg Tyr Gly Ile Ala Arg Tyr Asn Gly Cys Gln Asn Leu Leu Ser Pro Thr Trp Ile Thr Thr Val Leu Tyr Thr MET Trp MET Leu Ile

        1432                1447                1462                1477                1492                1507                1522      N5       1537
TGG TCA TTT GTG GGC GCT GTT TAT GCC ACC TTA GTA CTG TTC GTG TTT TAT AAA AAA CGC AAG GAC GTT AGG GAT ATT TTA CAC TGT ACC AAT TCA GGT TTA AAC CTG ACA AGG
Trp Ser Phe Val Gly Ala Val Tyr Ala Thr Leu Val Leu Phe Val Phe Tyr Lys Lys Arg Lys Asp Val Arg Asp Ile Leu His Cys Thr Asn Ser Gly Leu Asn Leu Thr Arg

        1552                1567                1582                1597                1612                1627                1642
TTC GCA AGG CTG TTA ATA TTC TGT TTC ATT ATT ATT TTA GTC ATG TTC CCT TTT TCT GTT TAC ACC TTT GTT CAA GAT TTA CAG CAG GTA GAA GGA CAC TAT ACT TTT AAA AAT
Phe Ala Arg Leu Leu Ile Phe Cys Phe Ile Ile Ile Leu Val MET Phe Pro Phe Ser Val Tyr Thr Phe Val Gln Asp Leu Gln Gln Val Glu Gly His Tyr Thr Phe Lys Asn

    1657                1672        N6      1687                1702                1717                1732                1747                1762
ACC CAT TCC AGC ACT ATC TGG AAT ACC ATT ATT AAA TTT GAC CCT GGC AGA CCA ATT TAT AAT ATA TGG CTT TAT GTT TTG ATG TCT TAC CTA GTA TTT CTA ATC TTT GGC TTA
Thr His Ser Ser Thr Ile Trp Asn Thr Ile Ile Lys Phe Asp Pro Gly Arg Pro Ile Tyr Asn Ile Trp Leu Tyr Val Leu MET Ser Tyr Leu Val Phe Leu Ile Phe Gly Leu

        1777                1792                1807                1822                1837                1852           N7 1867
GGT TCT GAT GCT TTG CAT ATG TAC TCT AAA TTC CTG CGT TCC ATC AAA CTA GGA TTT GTA CTT GAC ATG TGG AAA AGA TTC ATT GAT AAG AAT AAG GAA AAA CGA GTA GGC ATA
Gly Ser Asp Ala Leu His MET Tyr Ser Lys Phe Leu Arg Ser Ile Lys Leu Gly Phe Val Leu Asp MET Trp Lys Arg Phe Ile Asp Lys Asn Lys Glu Lys Arg Val Gly Ile

1882                1897                1912                1927                1942                1957                1972                1987
TTG CTA AAC AAG CTG TCC TCA CGC AAA GAG AGT CGT AAC CCA TTT TCT ACA GAC TCT GAG AAC TAT ATC TCC ACG TGT ACA GAA AAC TAT TCT CCC TGT GTA GGT ACA CCA ATA
Leu Leu Asn Lys Leu Ser Ser Arg Lys Glu Ser Arg Asn Pro Phe Ser Thr Asp Ser Glu Asn Tyr Ile Ser Thr Cys Thr Glu Asn Tyr Ser Pro Cys Val Gly Thr Pro Ile

        2002                2017                2032                2047                2062                2077                2092                2107
TCA CAA GCG CAT TTC TAT GTC GAC TAT AGG ATT CCA GAT GAT CCT AGA AAA TCT CAA AAT AAA AGC AAA AAA TAT TTG TTT GCT GAT AAA GAA ACA GAT GAT ATT CTT GAT GAA
Ser Gln Ala His Phe Tyr Val Asp Tyr Arg Ile Pro Asp Asp Pro Arg Lys Ser Gln Asn Lys Ser Lys Lys Tyr Leu Phe Ala Asp Lys Glu Thr Asp Asp Ile Leu Asp Glu

        2122                2137                2152                2167                2182                2197                2212
ATA GAC CTA AAA GAA AGT AGG CAC ATC CCT TAC GTC ACG CAA GGA CAG AGC TTT GAC GAC GAA ATA TCA CTT GGA GGA TTC TCA AAA GTT ACT CTC GAT TAT TCA GAA AAG CTT
Ile Asp Leu Lys Glu Ser Arg His Ile Pro Tyr Val Thr Gln Gly Gln Ser Phe Asp Asp Glu Ile Ser Leu Gly Gly Phe Ser Lys Val Thr Leu Asp Tyr Ser Glu Lys Leu

    2227                2242                2257                2272                2287                2302                2317                2332
CAT AAT TCT GCA AGC TCC AAT TTT GAA GGG GAA AGT CTT TGC TAC TCT CCA GCT TCA AAA GAA GAG AAT TCA AGC TCA AAC GAA CAT AGT TCA GAA AAT ACT GCA GGC CCT TAA
His Asn Ser Ala Ser Ser Asn Phe Glu Gly Glu Ser Leu Cys Tyr Ser Pro Ala Ser Lys Glu Glu Asn Ser Ser Ser Asn Glu His Ser Ser Glu Asn Thr Ala Gly Pro  .

      2345       2355       2365       2375       2385       2395       2405       2415       2425       2435       2445       2455       2465       2475
CACAAGAGTG TCGCATTATA TTTACTGGAC TAGGAGTATT TTATTTTTAC AGGACTAGGA TTGAAATACT GCTTTTTAGT GAATTGTGGC TCAAATAATG TAACGATGAG CTCATCAGCT AATATGTGGC TTAGCGGTAA

      2485       2495       2505       2515       2525       2535       2545       2555       2565       2575       2585       2595       2605       2615
AAATGACGAA TTGTGTGTAA ACTTTACTTT AATATTTACT GCTTTTTTGC TACTTTGGTT TCTATTTTTT CTATAGAAAA GCAATAACGT CTGTATTATA TATAAATATA AGGAGAAATT GATACAAGTT CCACACAAGA

      2625       2635       2645
AAAAATTTTT AAGCTCGAAA CAAAAGTAAG CTT
```

Fig. 3. Nucleotide sequences of *STE2* (**A**) and *STE3* (**B**) and predicted primary structures of these gene products. Sequences homologous to the canonical 'TATA' sequence are denoted by boxes. The putative α2 protein binding site (Miller *et al.*, 1985) is indicated by double underlining. The mRNA start sites determined by primer extension experiments are indicated by asterisks. Amino acid sequences underlined correspond to predicted hydrophobic segments (Figure 5). Nucleotide sequences with wavy underscores indicate the homologous regions (70% homology) within the 5'-non-coding region of *STE2* and *STE3*; those with dotted lines denote the sequence of the chemically synthesized primer for the primer extension experiment.

the two alleles of the *ste2* mutations. Therefore, the minimum complementation unit is probably located within the overlapping *Hind*III-*Sal*I region.

The *ste3-1* mutant transformed with YCpSTE3-13, one of the plasmids carrying *STE3*, was mating proficient (Figure 1). The restriction map for the inserted DNA and the minimum complementation unit in the plasmid was consistent with those previously described for *STE3* (Sprague *et al.*, 1983b). YCpSTE3HS, the plasmid carrying the 2.0-kb *Hind*III-*Sal*I fragment, in fact complemented the mating deficiency of the *ste3-1* mutation (Sprague *et al.*, 1983b; Figure 1).

Disruption of STE2

The 1.2-kb *Hind*III fragment containing *URA3* from pRB45 (Rose *et al.*, 1981) was inserted between *Hpa*I and *Eco*RV sites, located in the coding region of *STE2* (see following section) on the YCpSTE2B plasmid, to yield plasmid pSTE2:URA3. The *Bam*HI fragment containing *STE2* disrupted with *URA3* (*ste2:URA3*), was then introduced into wild-type *MAT*a and *MAT*α strains and stable Ura^{+} transformants were selected (Rothstein, 1983). The integration site was determined by Southern blotting after digesting chromosomal DNA with restriction endonucleases *Cla*I or *Pst*I (Southern, 1975). A longer *Cla*I fragment which hybridized with both *STE2*

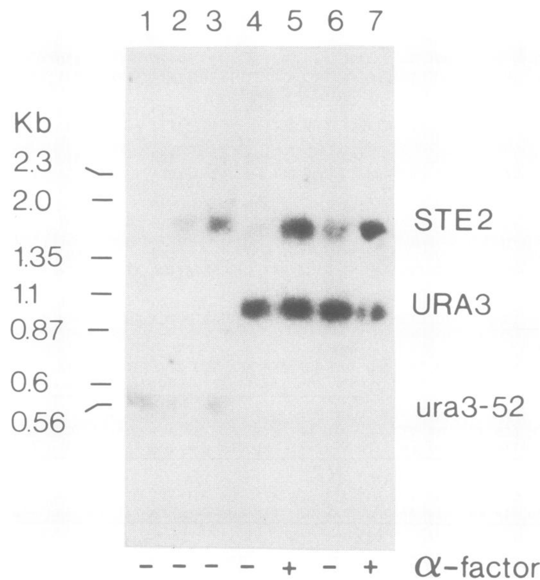

Fig. 4. Northern blotting of *STE2* and *STE3* mRNAs. 5 μg of poly(A)$^+$ from Y47 **(lane 1)**, YP45 **(lane 2)**, RC629 **(lane 4)**, RC629 treated with synthetic α-factor **(lane 5)**, RC631 **(lane 6)** and RC631 treated with synthetic α-factor **(lane 7)** and 10 μg of poly(A)$^+$ RNA from YP45 **(lane 3)** were separated on a 1% agarose gel and transferred to nitrocellulose membrane. The *Bam*HI fragment containing *URA3* within the structural gene of *STE2* (*ste2:URA3* fragment) was used as a probe.

(Figure 2A) and *URA3* (data not shown) probes were generated at the expense of the original 2.3-kb *Cla*I fragment containing both *Hpa*I and *Eco*RV sites. An additional small *Pst*I fragment appeared which also hybridized with both *STE2* (Figure 2B) and *URA3* probes (data not shown). This result is consistent with the presence of a unique *Pst*I site within the *URA3* fragment introduced into the 8-kb *Pst*I region. The difference between the observed size (2 kb) and the expected size (1.2 kb) of the insert was due to the insertion of an additional DNA fragment (0.8 kb), which was next to the *URA3* fragment during construction of pSTE2:URA3. These results indicated the integration sites of the fragment in all stable transformants were at the same locus from which *STE2* clones were derived and that chromosomal *STE2* was disrupted. Although the *MATa ste2:URA3* strains were defective in mating ability, *MATα ste2:URA3* strains were able to mate as wild-type α cells. The insertion mutation thus causes *a* cell-specific mating defect.

Nucleotide sequence of STE2 and STE3

The complete nucleotide sequences of the 2.0-kb partial *Hin*dIII fragment which complements the mating defects of the *ste2-1* and *ste2-3*^{ts} mutations and of the 2.6-kb partial *Hin*dIII fragment containing the minimum complementation unit (2.0-kb *Hin*dIII-*Sal*I region) for the *ste3-1* mutation have been determined (Figure 3). Both fragments have an apparently simple protein coding structure, consisting of long, unique, open reading frames, surrounded by AT-rich non-coding sequences. *STE2* contains an open reading frame within the 1.6-kb *Hin*dIII fragment coding for a polypeptide of 431 amino acids; the open reading frame of *STE3*, which extends beyond the *Sal*I site of the *Hin*dIII-*Sal*I region, encodes for a protein of 470 amino acids.

Expression of STE2 and STE3

Genetic analyses suggest that *STE2* and *STE3* seem to be expressed specifically in *a* cells and α cells, respectively, and the expression of *STE3* is regulated at the level of transcription (Sprague *et al.*, 1983b). Therefore cloned DNA was used to examine

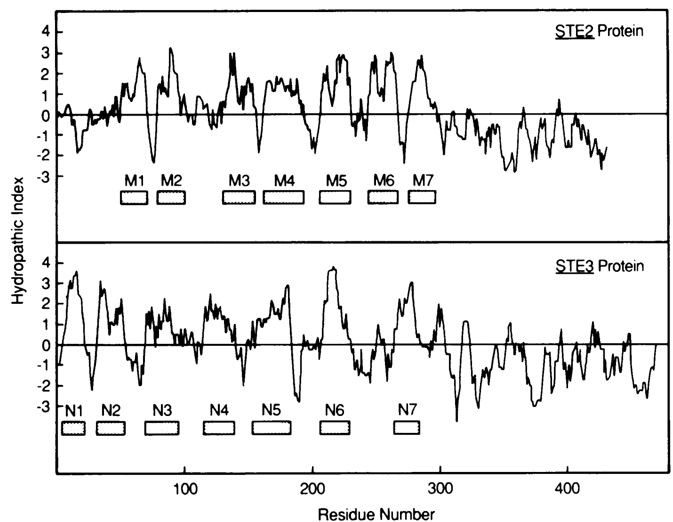

Fig. 5. Hydrophobicity profiles of the *STE2* and *STE3* gene products. The ordinate is the average of hydropathic index (Kyte and Doolittle, 1982) of a stretch of seven residues and the abscissa is the residue number at the centre of the stretch. The locations of the predicted hydrophobic segments of *STE2* (M1 – M7) and *STE3* (N1 – N7) gene products are indicated by boxes.

whether or not cell type-specific transcription might be the case for *STE2*. Poly(A)$^+$ RNA isolated from isogenic *MATa* and *MATα* strains was subjected to Northern blotting analyses (Thomas, 1983). The *STE2* transcript (1.6 – 1.7 kb) was present only in *MATa* strains and the *STE3* transcript (1.6 – 1.7 kb) was present only in *MATα* strains (data not shown), whereas the *ura3-52* transcript (~0.6 kb) was present at a similar level in both cell types (Figure 4, lanes 1 – 3).

The effect of α-factor on expression of *STE2* was also examined. *MATa sst1* strain, RC629, which lacks α-factor-specific protease (Chan and Otte, 1982) was used to prevent α-factor destruction during the incubation with wild-type *a* cells. A supersensitive strain to α-factor, RC631 (*MATa* sst2-1) was also employed for the analysis. As shown in Figure 4 (lanes 4 – 7), 1 h incubation with α-factor seems to increase the steady-state level of *STE2* mRNA. By contrast, wild-type *URA3* transcript (0.9 – 1.0 kb) was not affected by α-factor.

The transcription start site of *STE2* was determined by the primer extension method. Two major bands and two minor bands appeared on the gel (data not shown). The 5′ end of the major transcripts were mapped 32 and 31 bp before the initiator ATG (Figure 3A).

Discussion

Based on the following observations, we concluded that the cloned genomic segment which complemented the *ste2* mutations encodes *STE2*: (i) a genomic library was established in the low copy number YCpN1 vector to minimize the possibility of cloning gene(s) other than *STE2* which complements *ste2* phenotype by gene dosage effect; (ii) all the positive clones analyzed were derived from the same chromosomal locus; (iii) the disruption of the chromosomal *STE2* causes *a* cell-specific mating defects; and (iv) the cloned gene was expressed only in *a* cells. Since the restriction map and the minimum complementation unit (2.0-kb *Hin*dIII-*Sal*I region) for *ste3* are the same as those described previously (Sprague *et al.*, 1983b), we also concluded that our genomic clone encodes *STE3*.

Structure of the STE2 and STE3 gene products

The predicted primary structures of the *STE2* and *STE3* products,

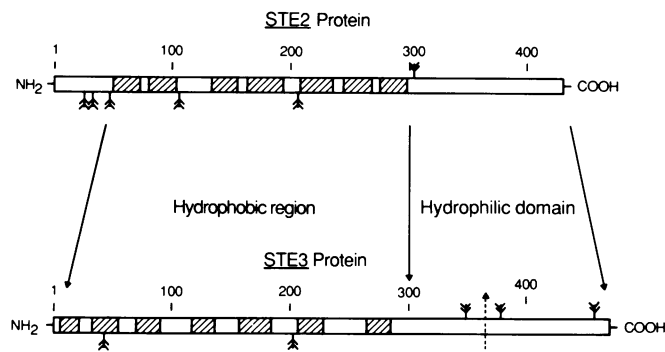

Fig. 6. Predicted protein structure. The primary structures of *STE2* and *STE3* were drawn to physical size. Dotted regions indicate hydrophobic domains predicted by the computer program developed by Kyte and Doolittle (1982) (Figure 5). The doubled Y indicates possible N-glycosylation sites [Asn-x-Ser(Thr)]. The dotted arrow denotes the position of the *Sal*I site in the coding region of *STE3* (106 amino acid residues before the COOH terminus of the *STE3* protein).

which are thought to act as receptor molecules for *a*- and α-factors, are not strikingly homologous. Comparison of the hydrophobicity of these two gene products suggested similar protein folding (Figures 5 and 6). They have seven hydrophobic segments in the NH_2-terminal regions and a long hydrophilic domain (130 or 170 amino acids) in the COOH-terminal regions.

The seven strongly hydrophobic segments of *STE2* and *STE3* products (M1 – M7 and N1 – N7) consist of 17 – 31 amino acids including many non-polar amino acids (Figures 2 and 5). These segments are bounded by a number of charged residues. The average hydrophobicity of each segment is equal to or above the average value of the membrane-spanning regions of other proteins (Kyte and Doolittle, 1982) (data not shown). The minimum number of amino acids required for an α-helix to span the 30 Å thickness of the hydrophobic space in a bilayer is 21. Therefore, all the hydrophobic segments of *STE2* and *STE3* products, except for the NH_2-terminal hydrophobic segment of *STE3* product (N1), could possibly traverse a lipid bilayer. Because these are typical characteristics of integral membrane proteins, for instance rhodopsins (Henderson and Unwin, 1975; Nathans and Hogness, 1983) and acetylcholine receptors (Numa *et al.*, 1983; Fairclough *et al.*, 1983), it is likely that *STE2* and *STE3* products could be membrane proteins. The *STE2* product does not seem to have a hydrophobic signal peptide at the NH_2-terminal end, which is often found in membrane-bound or secreted proteins. The relatively short hydrophobic region close to the NH_2 terminus of the *STE3* product (N1) possibly serves as a signal peptide. There are several possible N-glycosylation sites (Asn-x-Ser[Thr]) in each product (Figure 6). In the NH_2-terminal region of the *STE2* product, three N-glycosylation sites were predicted. Three sites were also predicted in the long hydrophilic domain of the *STE3* product. It is tempting to speculate that these regions could be exposed to the outside of the cell. Deletion analyses were performed by subcloning various fragments of these genes into the YCpN1 vector (Figure 1). YCpSTE2S carrying a 2.2-kb *Sal*I fragment, which lacks 24 amino acids from the COOH-terminal region of the *STE2* product, complemented both *ste2-1* and *ste2-3*^{ts}. These results suggested that at least the end of the COOH-terminal domain is dispensable. Likewise, 106 amino acid residues from the COOH-terminal end of the *STE3* product are probably not essential for the mating function; the YCpSTE3HS carries the minimum complementation unit (2.0-kb *Hin*dIII-*Sal*I region) lacking these residues and complements the mating defect of the *ste3-1* mutation as previously described

Table I. Yeast strains

Name	Genotype	Source
YP45	*MATa ade2-101 lys2-801 trp1-Δ ura3-52*	P.Hieter
YP47	*MATα ade2-101 lys2-801 trp1-Δ ura3-52*	P.Hieter
DBY747	*MATa his3-Δ1 leu2-3 leu2-112 trp1-289 ura3-52*	YGSC[a]
DBY746	*MATα his3-Δ1 leu2-3 leu2-112 trp1-289 ura3-52*	YGSC
XH9-5C-5C	*MATa ste2-1 ade2-1 his3* and/or *his2 gal2 can1*	YGSC
50B	*MATa ste2-3* in 381G-STE+[b]	YGSC
VQ3	*MATα ste3-1* in XT112-S245C[c]	YGSC
LL20	*MATα can1 his3-11 his3-15 leu2-3 leu2-112*	J.W.Szostak
RC618	*MATa ade2-1 ural his6 met1 can1 cyh2 rme GAL*	R.Chan
NNY110	*MATa ste2-1 ade2 lys2-801 trp1-Δ ura3-52*	This work
NNY124	*MATa ste2-3 ade2 lys2 trp1-Δ ura3-52*	This work
YAM10	*MATa ste2:URA3 ade2-101 lys2-801 trp1-Δ*	This work
NNY111	*MATα ste2:URA3 ade2-101 lys2-801 trp1-Δ*	This work
NNY128	*MATα ste3-1 trp1-289 met1 ade6 leu his*	This work
RC629	*MATa sst1-2* in RC618	R.Chan
RC631	*MATa sst2-1* in RC618	R.Chan

[a]Yeast Genetic Stock Center.
[b]*MATa SUP4-3(t.s.) cry-1 his4-580 trp1 ade2-1 tyr1 lys2.*
[c]*MATα ade6 his6 leu1 met1 trp5-1 can1 rme1 gal1.*

(Figures 1 and 6; Sprague *et al.*, 1983b).

Genomic organization of STE2 and STE3 genes

In the 1.6-kb *Hin*dIII fragment of *STE2*, there are two adjacent sequences homologous to the canonical TATA sequence, TATAAA (Sentenac and Hall, 1982), which is located 49 – 40 bp upstream from the putative initiation codon. Another sequence TATGAA is found 114 – 120 bp before the initiator ATG. The former two sequences are more likely to be the canonical sequence, because the 5′ end of *STE2* mRNA was mapped predominantly at 32 bases in front of the ATG codon (Figure 3A).

In the 5′-non-coding region of *STE3*, no sequence homologous to TATAAA was found near the initiator ATG codon. There is an AT-rich sequence, $(AT)_{10}$GTA, followed by TATAGA, which is homologous to the canonical sequence, 165 – 140 bp upstream from the initiator ATG. A similar TATA sequence, TATATAA, was found 128 – 122 bp before the ATG codon in the promoter region of the *MFα1* gene, another α cell-specific gene (Kurjan and Herskowitz, 1982; Singh *et al.*, 1983). Since the canonical TATA sequence is often found as far as 100 bp from the transcription initiation site (Sentenac and Hall, 1982), this sequence may be a candidate for the canonical sequence.

The 3′ end of both genes contains the sequences TAGT or TATGT, followed by TTT, which are found at the 3′ ends of many yeast genes and are thought to be signals for transcription terminators (Zaret and Sherman, 1982).

Expression of STE2 and STE3

Since transcription of *STE and STE3* are regulated by the *MAT* locus, these genes may share sequence homology with other cell type-specific sterile genes. *STE2* is an *a* cell-specific gene whose expression is thought to be repressed by α2 protein in α cells (Strathern *et al.*, 1981). In fact, transcription of *STE6*, another *a* cell-specific gene, was shown to be regulated by the *MATα2* product (Wilson and Herskowitz, 1984). Recently, the consensus sequence for the putative binding site of α2 protein was reported (Miller *et al.*, 1985). This sequence is located at 200 bp upstream of the initiator ATG codon (Figure 3A). By contrast, expression of the *STE3* and α-factor genes (*MFα1* and *MFα2*) are restricted in α cells and depend on the function of the *MATα1* product which acts positively on the α cell-specific genes (Sprague

et al., 1983b; Strathern *et al.*, 1981). However, the nucleotide sequence of the 230–280 bp 5′-non-coding region of *MFα1* and *MFα2* (Singh *et al.*, 1983) contains no extensive homology with that of *STE3*. It is possible that comparison with a region further upstream may reveal conserved features.

Alpha-factor is known to enhance the production of *a*-factor and the *BAR1* product (Strazdis and MacKay, 1983; Manney, 1983) and *a*-factor induces transcription of *STE3* (Hagen and Sprague, 1984). We found that transcription of *STE2* is also induced by α-factor (Figure 4). However, further experiments are necessary to conclude whether or not α-factor directly causes the induction. Comparison of the promoter region of *STE2* and *STE3* reveals no extensive homology except for a similar sequence (~70% homology) located in a similar position in the 5′-non-coding region of both genes (Figure 2, wavy underscores).

Materials and methods

Chemicals and enzymes

Deoxy and dideoxy nucleotides were purchased from P-L Biochemicals. [α-^{32}P]dATP was from Amersham. Klenow fragment of *E. coli* DNA polymerase I was purchased from Boehringer Mannheim. All the primers were synthesized by the 380A DNA synthesizer (Applied Biosystems). Reverse transcriptase was obtained from Life Science and further purified by gel filtration on Sephacryl S-200 (Pharmacia) in 0.2 M potassium phosphate, pH 7.2, 2 mM dithiothreitol, 0.2% Triton X-100, and 20% glycerol. Synthetic α-factor, restriction enzymes and T4 DNA ligase were purchased from Sigma, Bethesda Research Laboratories and New England BioLabs, respectively.

Yeast strains and plasmid vectors

NNY110, NNY124 and NNY128 were constructed by crossing XH9-5C-5C and YP47, 50B and YP47, and VQ3 and DBY747, respectively (Table I). The vector YCpN1 was constructed by deleting the 2.4-kb *Bam*HI fragment carrying the *ADH1* promoter-R·*dhfr-ADH1* terminator from the pADA4 plasmid (Miyajima *et al.*, 1984a).

Genomic DNA library

Chromosomal DNA of DBY746 was extracted by a conventional phenol method (Maniatis *et al.*, 1982), partially digested with the restriction enzyme *Mbo*I, and separated by sucrose density gradient centrifugation. Fractions containing 7–15 kb fragments were combined and cloned into the *Bam*HI site of the YCpN1 vector. The ligation mixtures were used to transform MC1061 by the $CaCl_2$ method (Maniatis *et al.*, 1982). 1.1×10^5 independent colonies were obtained. 75% contained inserts of an average length of 8 kb.

Yeast methods

Yeast cells were transformed by lithium acetate methods (Itoh *et al.*, 1983). All the procedures for mating assays were basically the same as described by MacKay (1983). Testers were LL20 for *ste2* mutants and RC618 for *ste3* mutants. DNA from yeast cells was prepared according to Davis *et al.* (1980).

DNA sequencing

Both the dideoxy chain termination method (Smith, 1980) and the modified chemical method (Maxam and Gilbert, 1980; Rubin and Schmid, 1980) were employed for sequencing DNAs. Acrylamide gels (33 cm × 90 cm) were dried prior to autoradiography. Analyses of the nucleotide sequences were carried out using the programs of Intelligenetics (Brutlag *et al.*, 1981). All the sequences presented here have been determined on both strands.

Poly(A)$^+$ RNA preparation

Yeast cells YP45, YP47 and RC618 were harvested at midlog phase (OD_{630} 0.8). Half of the RC618 cells were further incubated at 30°C with α-factor (0.05 μg/ml) and harvested after 60 min. Poly(A)$^+$ RNA was prepared as described previously (Miyajima *et al.*, 1984b) and dissolved in RNAse-free water at 1 mg/ml.

5′ End mapping of the STE2 transcript

5′ Ends of the *STE2* transcript were determined by the primer extension method. A primer for *STE2* was chemically synthesized (from nucleotide number 612 to 644, Figure 3A); 2–5 μg of poly(A)$^+$ RNA was mixed with the kinased primer (~10^6 c.p.m. of ^{32}P) and transferred into a 20 μl capillary; final concentration of RNA was 1 mg/ml. The sealed capillary was heated at 90°C for 2 min, then at 55°C for 1 h to hybridize the mRNA and the primer. This hybridization mixture was directly transferred into buffer containing 50 mM Tris-HCl (pH 8.3), 80 nmol of $MgCl_2$, 0.3 μmol of KCl, 20 nmol each of four deoxyribonucleotides, 10 nmol of dithiothreitol and 15 units of RNasin (Promega Biotec). Five units of purified reverse transcriptase was added, the final volume was adjusted to 10 μl and the reaction was carried out at 40°C for 1 h. The transcribed complementary DNA was precipitated with ethanol and dissolved in the formamide-dye solution for sequencing gels. The same primers were used for M13-sequencing and both samples were applied to the same gel to determine the start site.

Other procedures

Southern blotting was by the standard procedure described by Southern (1975). Northern blotting was essentially the same as described by Thomas (1983).

E. coli strain MC1061 [*araD139, Δ(ara,leu)7697, ΔlacX74, galU, galK, hsr, strA*] was routinely used for transformation and preparation of plasmid DNA.

Acknowledgements

We thank Karl Pope for synthesis of oligonucleotides, Shigeru Sakonju for primer extension experiments and Philip Heiter for providing useful strains. We are also grateful to Kunihiro Matsumoto for helpful discussions. NN is a recipient of a DNAX Research Institute Graduate Fellowship.

References

Brutlag,D.L., Clayton,J., Friedland,P. and Kedes,L.H. (1981) *Nucleic Acids Res.*, **10**, 279-294.
Chan,R.K. and Otte,C.A. (1982) *Mol. Cell. Biol.*, **2**, 21-29.
Davis,R.W., Thomas,M., Cameron,J., St.John,T.P., Scherer,S. and Padgett,R.A. (1980) *Methods Enzymol.*, **65**, 404-411.
Fairclough,R.H., Finer-Moore,J., Love,R.A., Kristofferson,D., Desmeules,P.J. and Stroud,R.M. (1983) *Cold Spring Harbor Symp. Quant. Biol.*, **48**, 9-20.
Hagen,D.C. and Sprague,G.F., Jr. (1984) *J. Mol. Biol.*, **178**, 835-852.
Hartwell,L.H. (1980) *J. Cell Biol.*, **85**, 811-822.
Henderson,R. and Unwin,P.N.T. (1975) *Nature*, **257**, 28-32.
Itoh,H., Fukuda,Y., Murata,K. and Kimura,A. (1983) *J. Bacteriol.*, **153**, 163-168.
Jenness,D.D., Burkholder,A.C. and Hartwell,L.H. (1983) *Cell*, **35**, 521-529.
Kurjan,J. and Herskowitz,I. (1982) *Cell*, **30**, 933-943.
Kyte,J. and Doolittle,R.F. (1982) *J. Mol. Biol.*, **157**, 105-132.
MacKay,V.L. (1983) *Methods Enzymol.*, **101**, 325-343.
MacKay,V.L. and Manney,T.R. (1974a) *Genetics*, **76**, 255-271.
MacKay,V.L. and Manney,T.R. (1974b) *Genetics*, **76**, 273-288.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning. A Laboratory Manual*, published by Cold Spring Harbor Laboratory Press, NY.
Manney,T.R. (1983) *J. Bacteriol.*, **155**, 291-301.
Manney,T.R. and Woods,V. (1976) *Genetics*, **82**, 639-644.
Maxam,A.M. and Gilbert,W. (1980) *Methods Enzymol.*, **65**, 499-560.
Miller,A.M., MacKay,V.L. and Nasmyth,K.A. (1985) *Nature*, **314**, 598-603.
Miyajima,A., Miyajima,I., Arai,K-I. and Arai,N. (1984a) *Mol. Cell Biol.*, **4**, 407-414.
Miyajima,A., Nakayama,N., Miyajima,I., Arai,N., Okayama,H. and Arai,K-I. (1984b) *Nucleic Acids Res.*, **12**, 6397-6414.
Nathans,J. and Hogness,D.S. (1983) *Cell*, **34**, 807-814.
Numa,S., Noda,H., Takahaski,T., Tanabe,M., Toyosato,Y., Furutani,I. and Kikyotani,S. (1983) *Cold Spring Harbor Symp. Quant. Biol.*, **48**, 57-69.
Rose,M., Casadaban,M.J. and Botstein,D. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 2460-2464.
Rothstein,R.J. (1983) *Methods Enzymol.*, **101**, 202-211.
Rubin,C.M. and Schmid,C.W. (1980) *Nucleic Acids Res.*, **8**, 4613-4619.
Sentenac,A. and Hall,B. (1982) in Strathern,J.N., Jones,E. and Broach,J.R. (eds.), *The Molecular Biology of the Yeast Saccharomyces*, Cold Spring Harbor Laboratory Press, NY, pp. 561-606.
Singh,A., Chen,E.Y., Lugovoy,J.M., Chang,C.N., Hitzeman,R.A. and Seeburg, P.H. (1983) *Nucleic Acids Res.*, **11**, 4049-4063.
Smith,A.J.H. (1983) *Methods Enzymol.*, **65**, 560-580.
Southern,E.M. (1975) *J. Mol. Biol.*, **98**, 503-517.
Sprague,G.F., Jr., Blair,L.C. and Thorner,J. (1983a) *Annu. Rev. Microbiol.*, **37**, 623-660.
Sprague,G.F. Jr., Jensen,R. and Herkowitz,I. (1983b) *Cell*, **32**, 409-415.
Strathern,J., Hicks,J. and Herskowitz,I. (1981) *J. Mol. Biol.*, **147**, 357-372.
Strazdis,J.R. and MacKay,V.L. (1983) *Nature*, **305**, 543-545.
Thomas,P. (1983) *Methods Enzymol.*, **100**, 255-266.
Thorner,J. (1981) in Strathern,J.N., Jones,E. and Broach,J.R. (eds.), *The Molecular Biology of the Yeast Saccharomyces*, Cold Spring Harbor Laboratory Press, NY, pp. 143-180.
Wilson,K. and Herskowitz,I. (1984) *Mol. Cell. Biol.*, **4**, 2420-2427.
Zaret,K.S. and Sherman,F. (1982) *Cell*, **28**, 563-573.

Received on 21 June 1985