Volume 291, number 2, 245–248 FEBS 10295 October 1991

ADONIS 0014579391010050

Molecular cloning and sequence analysis of cDNA and genomic DNA for the human cone transducin α subunit

Mitsumasa Kubo[1], Teiichi Hirano[2] and Mitsuaki Kakinuma[1]

[1]*Section of Bacterial Infection, Institute of Immunological Science* and [2]*The Third Department of Internal Medicine, Hokkaido University, Sapporo 060, Japan*

Received 12 August 1991

A novel GTP binding protein (G protein) α subunit cDNA was isolated from a T cell leukemia cell line, Jurkat, utilizing polymerase chain reaction (PCR) The predicted amino acid sequence of this G protein α subunit showed the highest identity (96 6%) to bovine cone cell-specific transducin ($T_c\alpha$) The organization of the coding region of this G protein α subunit gene was composed of 8 exons and 7 introns Northern hybridization revealed the presence of this G protein message in a retinoblastoma cell line, Y79 In Jurkat, however, the message was detectable only by reverse transcription/PCR Taken together, this novel G protein α subunit must be human T_c

GTP binding protein, Polymerase chain reaction, Genomic organization, Human T cell leukemia cell line

1. INTRODUCTION

Heterotrimeric GTP binding proteins (G proteins) play important roles in trans-signaling systems of eukaryotic cells [1] G proteins are composed of α, β and γ subunits and their functional specificities as signal transducers reside in the α subunits [1,2], while the βγ complex directly activates various downstream effectors [3–5] To date, 11 different kinds of α subunits of G proteins in mammals (G_s, G_{olf}, G_{i1}, G_{i2}, G_{i3}, G_o, G_x, G_q, G_{11}, T_r and T_c) have been reported for their complete amino acid sequences and some more on partial sequences [2,6–9]

T cell receptor (TCR)-mediated lymphocyte activation also implicated G protein(s) in phosphatidylinositol bisphosphate hydrolysis and such G protein(s) appears to be different from G_s although it is cholera toxin sensitive [10]

To clarify the TCR specific G protein(s), portions of cDNA of G protein α subunits were amplified from mRNA of a human T cell leukemia cell line, Jurkat, by polymerase chain reaction (PCR) During the course of the experiment, a novel human G protein α subunit which showed extremely high homology to bovine cone cell transducin α subunit ($T_c\alpha$) [11] was cloned This G protein must be a counterpart of bovine T_c and termed then as human T_c Genomic clones containing human $T_c\alpha$ were also isolated. Here we report the predicted amino acid sequence and the genomic organization of human $T_c\alpha$

Correspondence address M Kubo, Section of Bacterial Infection, Institute of Immunological Science, Hokkaido University, Kita-15, Nishi-7, Kita-ku, Sapporo 060, Japan Fax (81) (11) 758 7568

2 MATERIALS AND METHODS

2 1 *Materials*

Jurkat cells were cultured in the medium of RPMI 1640 with 10% fetal calf serum Oligonucleotides were synthesized on an Applied Biosystems 380B DNA synthesizer Taq DNA polymerase was purchased from Promega cDNA was synthesized using the kits of Boehringer Mannheim and/or Gibco/BRL T4 DNA ligase and Klenow fragment were obtained from TaKaRa T7 sequenase was obtained from USB [^{32}P]-dCTP (111 TBq/mmol) was purchased from NEN

2 2 *Oligonucleotides for PCR* (see Fig 1)

GMO1 5′GAATTCGGNAARWSNACNATHGTNAARCARAT3′, GMO2 5′GAGCTCTYNBHNCKYTGNCCNCCNACRTCRAA3′, TMO1 5′GGTCGACCATGGGNWSNGGNGC3′, TSO1 5′CCAGGTTGTTGAGCTGTCGC3′, TSO2 5′GGGATCCAGAGTCAAAACCACAGGCAT3′, TMO2, 5′GGTCGACTARAANARNCCRCARTCYTT3′, TSO3 5′TCAAGGCTATCATCTATGGA3′, TSO4 5′CATCCTGAAATTCAAGTCTT3′ where W=A/T, R=A/G, S=C/G, Y=C/T, K=G/T, H=A/C/T, B=C/G/T, N=A/C/G/T

Primers GMO1 and GMO2 correspond to the conserved sequences among the α subunits of various G proteins of GKSTIVKQM and FDVGGQ R(S/D)E, respectively, and TMO1 and TMO2 correspond to the sequences of MGSGA (bovine $T_c\alpha$ amino acids 1 to 5) and KDCGLF (bovine $T_c\alpha$ amino acids 349 to 354), respectively Primers TSO1, TSO2, TSO3, and TSO4 are specific oligonucleotides to human $T_c\alpha$

2 3 *Isolation and sequencing of cDNAs of human G proteins*

Total RNA was extracted from Jurkat cells [12] and poly(A)$^+$ RNA was prepared by oligo(dT) column chromatography [13] One μg of poly(A)$^+$ RNA was reverse-transcribed using oligo(dT)$_{17}$ as a primer according to the manufacturer's conditions PCR between primers GMO1 and GMO2 was performed using the resultant single-stranded cDNAs as templates under the condition of 94C 1 5 min, 52C 1 5 min, 72C 3 min Following the amplification of 30 cycles, the PCR products were reamplified in the same conditions The final PCR products were electrophoresed on a 2 5% agarose gel and 0 3–0 6 kb DNAs were isolated by use of DE81 papers and digested with *Sac*I and *Eco*RI Then the DNAs were subcloned in pUC18 plasmid and double-

```
   1 ATGGGAAGTGGAGCCAGTGCTGAGGACAAAGAACTGGCCAAGAGGTCCAACGAGCTAGAA
   1 MetGlySerGlyAlaSerAlaGluAspLysGluLeuAlaLysArgSerLysGluLeuGlu
     c

  61 AAGAAGCTGCAGGAGGATGCTGATAAGGAAGCCAACACTGTCAAGCTGCTACTGCTGGGT
  21 LysLysLeuGlnGluAspAlaAspLysGluAlaLysThrValLysLeuLeuLeuLeuGly

 121 GCTGGGGAGTCAGGAAAGAGCACCATCGTCAAACAGATGAAGATCATTCACCAGGATGGC
  41 AlaGlyGluSerGlyLysSerThrIleValLysGlnMetLysIleIleHisGlnAspGly
          a

 181 TATTCACCAGAAGAATGCCTGGAGTTCAAGGCTATCATCTATGGAAATGTGCTGCAGTCC
  61 TyrSerProGluGluCysLeuGluPheLysAlaIleIleTyrGlyAsnValLeuGlnSer

 241 ATCCTGGCTATCATCCGGGCCATGACCACACTGGGCATCGATTATGCTGAACCAAGCTGT
  81 IleLeuAlaIleIleArgAlaMetThrThrLeuGlyIleAspTyrAlaGluProSerCys

 301 GCGGATGACGGGCGACAGCTCAACAACCTGGCTGACTCCATTGAGGAGGGAACCATGCCT
 101 AlaAspAspGlyArgGlnLeuAsnAsnLeuAlaAspSerIleGluGluGlyThrMetPro

 361 CCTGAGCTCGTGGAGGTCATTAGGAGGTTGTGGAAGGATGGTGGGGTGCAAGCCTGCTTC
 121 ProGluLeuValGluValIleArgArgLeuTrpLysAspGlyGlyValGlnAlaCysPhe

 421 GAGAGAGCTGCAGAATACCAGCTTAATGACTCCGCATCTTACTACCTGAACCAATTAGAA
 141 GluArgAlaAlaGluTyrGlnLeuAsnAspSerAlaSerTyrTyrLeuAsnGlnLeuGlu

 481 CGAATTACAGACCCTGAGTACCTCCCTAGTGAGCAAGTGTGCTCCGATCCAGAGTCAAA
 161 ArgIleThrAspProGluTyrLeuProSerGluGlnAspValLeuArgSerArgValLys

 541 ACCACAGGCATCATTGAAACCAAGTTTTCCGTCAAAGACTTGAATTTCAGGATGTTTGAT
 181 ThrThrGlyIleIleGluThrLysPheSerValLysAspLeuAsnPheArgMetPheAsp
                                                          b

 601 GTGGGAGGGCAGAGATCCGAGAGAAAGAAGTGGATCCACTGCTTCGAGGGAGTCACCTGC
 201 ValGlyGlyGlnArgSerGluArgLysLysTrpIleHisCysPheGluGlyValThrCys

 661 ATCATTTTCTGTGCAGCCCTCAGTGCCTATGATATGGTGCTGGTGGAAGATGACGAAGTG
 221 IleIlePheCysAlaAlaLeuSerAlaTyrAspMetValLeuValGluAspAspGluVal

 721 AATCGTATGCATGAGTCTTTGCATCTGTTCAACAGCATATGTAACCACAAATTCTTTGCG
 241 AsnArgMetHisGluSerLeuHisLeuPheAsnSerIleCysAsnHisLysPhePheAla

 781 GCTACTTCCATTGTACTCTTTCTCAACAAGAAGGACCTCTTTGAGGAAAAAATCAAGAAA
 261 AlaThrSerIleValLeuPheLeuAsnLysLysAspLeuPheGluGluLysIleLysLys

 841 GTCCATCTCAGCATTTGTTTTCCAGAGTATGATGGTAACAACTCCTATGATGATGCGGGG
 281 ValHisLeuSerIleCysPheProGluTyrAspGlyAsnAsnSerTyrAspAspAlaGly

 901 AATTACATAAAGAGCCAGTTCCTTGACCTCAATATGCGAAAAGATGTCAAAGAAATCTAC
 301 AsnTyrIleLysSerGlnPheLeuAspLeuAsnMetArgLysAspValLysGluIleTyr

 961 AGTCACATGACCTGTGCTACAGATACACAGAATGTCAAATTTGTATTTGATGCAGTTACA
 321 SerHisMetThrCysAlaThrAspThrGlnAsnValLysPheValPheAspAlaValThr

1021 GATATTATCATCAAAGAAAACCTCAAGGACTGCGGCCTCTTCTAA
 341 AspIleIleIleLysGluAsnLeuLysAspCysGlyLeuPheEND
                                  f
```

Fig 1 Nucleotide and predicted amino acid sequences of human $T_c\alpha$ cDNA coding region Underlinings a, b, c and f represent synthesized fully degenerate oligonucleotides corresponding to each amino acid sequence of bovine $T_c\alpha$ (a, GMO1 and c, TMO1, sense b, GMO2 and f, TMO2, antisense) Overlinings d, e, g and h represent synthesized specific oligonucleotides to human $T_c\alpha$ cDNA (e, TSO2 and g, TSO3, sense d, TSO1 and h, TSO4, antisense) (▼) denotes the position of an intervening sequence The nucleotide sequence data of exon 1–8 reported in this paper appear in the DDBJ, EMBL and GeneBank Nucleotide Sequence Database under the accession numbers D90438, D90439, D90440, D90441, D90442, D90443, D90444, and D90445.

stranded plasmid sequencing was performed The sequence analyses of PCR products obtained using primers TMO1-TSO1 and TSO2-TMO2 were performed in the same way as that for the amplified DNAs using primers GMO1 and GMO2

2 4 *Isolation and sequencing of human genomic $T_c\alpha$*

Human placental genomic libraries in EMBL3 and λ DASH were screened with the probes (EX345, nucleotide position 237–594 EX45, 433–594, EX78, 633–1065, in Fig 1) prepared from the clones of PCR products which were labeled by random priming Positive clones were restriction mapped and subcloned in pUC18 plasmid after digestion with suitable restriction endonucleases and double-stranded plasmid sequencing was performed

3 RESULTS

Forty-one clones were obtained after transfection of *E coli* with pUC18 plasmid containing PCR products between primers GMO1 and GMO2. By sequencing two clones of these and homology searching, each clone showed extremely high homology to bovine $T_c\alpha$ One clone had the insert of 5′ sequence of *Sac*I site (nucleotide position 160–368) and another 3′ sequence (position 369–593) from *Sac*I site (Fig 1) Because of the remarkable identity of the predicted amino acid sequence of these two clones to bovine $T_c\alpha$, these clones were assumed to be contiguous and have been cloned separately due to *Sac*I digestion They appeared to be a part of the human $T_c\alpha$ cDNA.

In order to obtain the longer cDNA clone of human $T_c\alpha$, the fully degenerate oligonucleotides (bovine $T_c\alpha$ amino acid residues 1–4 TMO1, and 340–345: TMO2) and human $T_c\alpha$-specific oligonucleotides (TSO1 and TSO2) which were based on the sequence of the two clones described above were synthesized The PCR products using primers TMO1-TSO1 and TSO2-TMO2 were subcloned and sequenced. Resulting 5′- and 3′-fragments were found to be overlapped with the portion of the cDNA initially cloned as above and showed again extremely high homology to bovine $T_c\alpha$ In this way we determined the cDNA sequence of human $T_c\alpha$ from the nucleotide position 16–1050 (Fig 1)

Next, we screened human placental libraries with

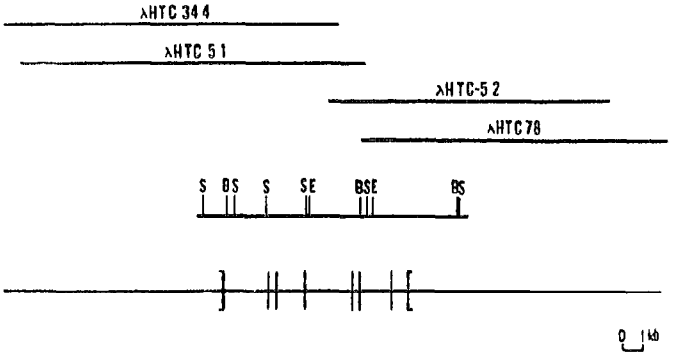

Fig 2 Organization and restriction map of human $T_c\alpha$ gene Four lines over the restriction map indicate λ clones containing human $T_c\alpha$ gene Boxes represent coding regions of the $T_c\alpha$ gene. B, *Bam*HI, E, *Eco*RI, S, *Sac*I

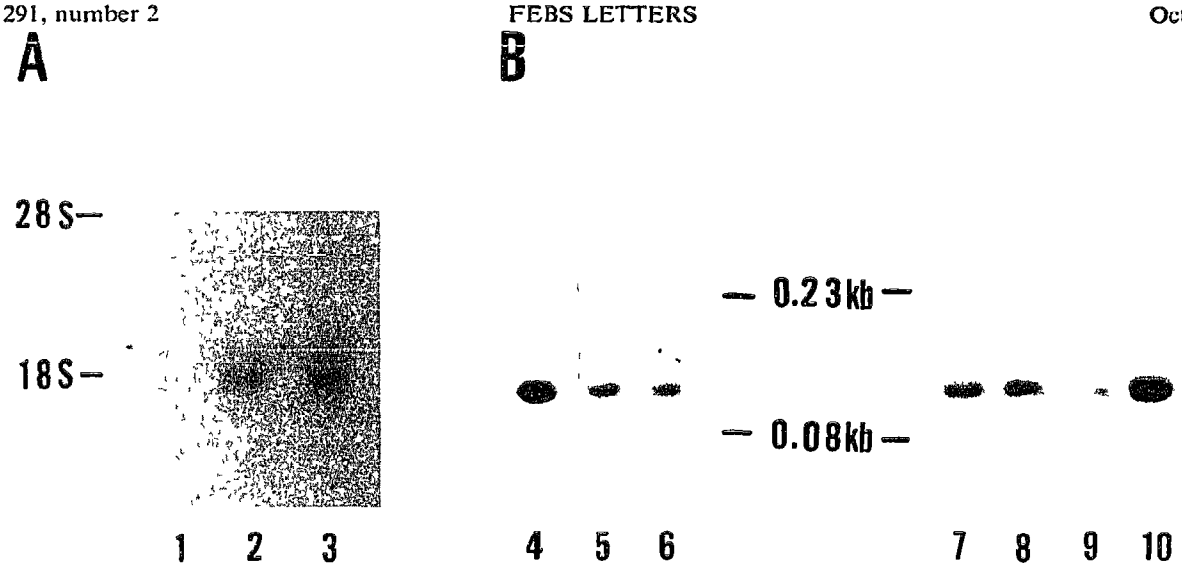

Fig 3 Expression of human $T_c\alpha$ gene (A) Northern hybridization Twenty μg of total RNA from Jurkat (lane 1) and Y79 (lane 2), and 30 μg of total RNA from Y79 (lane 3) were electrophoresed on a 1% formaldehyde agarose gel, transferred to a nitrocellulose membrane and hybridized with the mixed probe, EX345 and EX78 (B) RT/PCR-Southern hybridization Two μg of total RNA from Jurkat (lane 4), HL60 (lane 5), THP-1 (lane 6), normal human lymphocytes (lane 7), HeLa (lane 8), JEG-3 (lane 9), and NCCIT (lane 10) were reverse-transcribed using the synthesized oligonucleotide TSO4 as a primer, followed by PCR between primers TSO3 and TSO1 The resultant DNAs were electrophoresed on a 2 5% agarose gel, transferred to nitrocellulose membranes and hybridized with the probe, EX345

three different probes prepared from the clones of PCR products Totally 5 clones were isolated By the analyses of these genomic clones, it was found that human $T_c\alpha$ gene spanned approximately 10 kb and was composed of 8 exons and 7 introns (Fig. 2) All exon–intron junction sequences were compatible to a conserved GT-AG rule (Table I) [14] Two nucleotides within the sequence of the human $T_c\alpha$ cDNA amplified by Taq DNA polymerase were corrected from the genomic sequence Because the sequences of the regions corresponding to the degenerate primers (TMO1 and TMO2) were ambiguous, these sequences were deduced from the genomic sequence By PCR using two synthesized oligonucleotides corresponding to 5′- and 3′-untranslated regions (each about 10–20 bases upstream of initiation codon and downstream of termination codon) based on the genomic sequence, we obtained a human $T_c\alpha$ cDNA which contained all of the coding region (data not shown) The sequence analysis of this cDNA clone showed that the nucleotide sequence 1–15 and 1051–1063 were the same as those of the genomic sequence. The deduced amino acid sequence of human $T_c\alpha$ is shown in Fig 1, and it was composed of 354 amino acids and possessed Arg^{178} and Cys^{351} which would be ADP-ribosylated by cholera toxin and pertussis toxin, respectively

By comparison with other G proteins using GENETYX program, the amino acid sequence of human $T_c\alpha$ was 96 6% identical to bovine $T_c\alpha$ and approximately 81% to rod cell transducin α subunits ($T_r\alpha$) of bovine, mouse and human. The identity to the α subunits of human G_s, G_{i1}, G_{i2}, G_{i3}, G_o and G_x was less than 70%. Nucleotide sequence identity of human $T_c\alpha$ to bovine $T_c\alpha$ was 93 3%

Northern hybridization revealed the message of human $T_c\alpha$ in Y79 cells (a human retinoblastoma cell line) (Fig 3A) but not in other tumor cell lines. By reverse transcription (RT)/PCR and Southern hy-

Table I

Sequence of exon–intron junctions in human $T_c\alpha$

	Splice junction sequence and intron size		
Intron	Exon	Intron	Exon
1	TGCTGG gtgagt----	2 3kb ----atttttccttccag	GTGCTG
2	GATGAA gtgagt----	0 4kb ----tctgccccttgcag	GATCAT
3	TGTGCG gtatgt----	1 4kb ----ttttctgcattacag	GATGAC
4	ATCTTA gtaaga----	2 3kb ----tcacttttctctag	CTACCT
5	TTTCAG gtaagt----	0 3kb ----gcttcttcccctcag	GATGTT
6	GAAGTG gtggcc----	1 6kb ----ttttccttcttactag	AATCGT
7	ATGATG gtaagt----	0 8kb ----tctggaaaacccag	GTAACA

bridization, we, however, detected the $T_c\alpha$ message in normal human lymphocytes, Jurkat, HL60, THP-1 (a human monocytic leukemia cell line), JEG 3 (a human chorionic carcinoma cell line), NCCIT (a human embryonic carcinoma cell line) (Fig 3B)

4. DISCUSSION

Recently several novel cDNAs encoding for G proteins were isolated by PCR method [7,8,15]

The novel G protein α subunit which we reported in this paper is thought to be human $T_c\alpha$ for the following reasons This G protein has the highest homology to bovine $T_c\alpha$ among all of the G proteins reported so far The notable characteristic of G proteins is that the interspecies identity of amino acid sequences among a given class of mammalian G protein α subunits is higher than the identity among different G proteins in one species Moreover, we detected this G protein message in a retinoblastoma cell line, Y79, by Northern hybridization (Fig 3A) Boegenmann et al reported that retinoblastoma cells expressed cone cell-specific genes including $T_c\alpha$ using bovine $T_c\alpha$ cDNA as a probe [16] They detected two transcripts of $T_c\alpha$ and the lower molecular weight transcript was expressed predominantly It was almost the same molecular weight as reported here

The coding region of human $T_c\alpha$ gene is composed of 8 exons and 7 introns, the feature quite analogous to human $G_{i2}\alpha$, $G_{i3}\alpha$, $G_o\alpha$ and mouse $T_r\alpha$ genes [17–19] Since we did not obtain a full-length cDNA of this gene, it is possible that there might exist untranslated exon(s) as found in human G_{i2} [17], as suggested from the presence of two transcripts of $T_c\alpha$ gene in retinoblastoma cells [16] However, we found only one band in this study by Northern hybridization analysis (Fig 3A). Although there has been no report on the alternative splicing for bovine $T_r\alpha$ and $T_c\alpha$, it is also possible in the case of human $T_c\alpha$ as in human $G_s\alpha$ and $G_o\alpha$ [18,20,21]

Does human $T_c\alpha$ protein function in T cells? We detected the message of human $T_c\alpha$ in Y79 cells but not in other cell lines by Northern hybridization However, in some tumor cell lines including Jurkat and normal human lymphocytes, the presence of the $T_c\alpha$ message was detected by RT/PCR and Southern hybridization (Fig. 3B). The presence of very few molecules of the $T_c\alpha$ message may be the reason for undetectability by conventional Northern hybridization analysis in Jurkat cells Lerea et al detected the mRNA of bovine $T_c\alpha$ in bovine retina by Northern hybridization but not in bovine brain, liver, kidney, spleen or heart [22] They also reported that immunoreactive $T_c\alpha$ protein was seen exclusively in bovine cone cells using a specific antibody against bovine $T_c\alpha$ Although it is obvious that $T_c\alpha$ protein functions as the cone cell-specific phototransducer, immunochemical examinations using a specific antibody against human $T_c\alpha$ protein are necessary to clarify the significance of human $T_c\alpha$ in T cells Such investigations are now underway.

Acknowledgements We thank the Japanese Cancer Research Resources Bank for providing TPH-1 and HL60, Dr M Noda for Y79 (through RGB Cell Bank), and Dr S Teshima for NCCIT This investigation was partly supported by the special Grant-in-Aid for the promotion of Education and Science in Hokkaido University provided by the Ministry of Education, Science and Culture

REFERENCES

[1] Gilman A G (1987) Annu Rev Biochem 56, 615–649
[2] Lochrie, M A and Simon, M I (1988) Biochemistry 27, 4957–4965
[3] Jelsema, C L and Axelrod, J (1987) Proc Natl Acad Sci USA 84, 3623–3627
[4] Logothesis, D E, Kurachi, Y, Galper, J, Neer, E J and Clapham, D E (1988) Nature 325, 321–326
[5] Logothetis, D E, Kim, D, Northup, J K, Neer, E J and Clapham, D E (1988) Proc Natl Acad Sci USA 85, 5814–5818
[6] Jones, D T and Reed, R R (1989) Science 244, 790–795
[7] Strathmann, M, Wilkie, T H and Simon, M I (1989) Proc Natl Acad Sci USA 86, 7407–7409
[8] Strathmann, M and Simon, M I (1990) Proc Natl Acad Sci USA 87, 9113–9117
[9] Matsuoka, M, Itoh, H, Kozawa, T and Kaziro, Y (1988) Proc Natl Acad Sci USA 85, 5384–5388
[10] Imborden J B Shoback, D M, Pattison, G and Stobo, J D (1986) Proc Natl Acad Sci USA 83, 5673–5677
[11] Lochrie, M A, Hurley, J B and Simon, M I (1985) Science 228 96–99
[12] Chomczynski, P and Sacchi, N (1987) Anal Biochem 162, 156–159
[13] Sambrook, J, Fritsch, E F and Maniatis, T (1989) Molecular Cloning A Laboratory Manual, second edition, pp 7–26, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
[14] Padgett, R A, Grabowski, P J, Konarski, M M, Seiler, S and Sharp, P A (1986) Annu Rev Biochem 55, 1119–1150
[15] Gautam, N, Northup, J, Tamir, H and Simon M I (1990) Proc Natl Acad Sci USA 87, 7973–7977
[16] Bogenmann, E Lochrie, M A and Simon, M I (1988) Science 240, 76–78
[17] Itoh, H, Toyama, R, Kozawa, T, Tsukamoto, T, Matsuoka, M and Kaziro, Y (1988) J Biol Chem 263, 6656–6664
[18] Bertrand, P, Sanford, J, Rudolph, U, Codina, J and Birnbaumer, L (1990) J Biol Chem 265 18576–18580
[19] Raport, C J, Dere, B and Hurley J B (1989) J Biol Chem 264, 7122–7128
[20] Kozawa, T, Itoh, H, Tsukamoto, T and Kaziro Y (1988) Proc Natl Acad Sci USA 85, 2081–2085
[21] Bray, P, Carter, A, Simon, C, Guo, V, Puckett, C, Kamholz, J, Spiegel, A and Nirenberg, M (1986) Proc Natl Acad Sci USA 83, 8893-8897
[22] Lerea, C L, Somers, D E, Hurley, J B, Klock, I B and Bunt-Milam, A H (1986) Science 234, 77–80