Pure & Appl. Chem., Vol. 59, No. 3, pp. 437–444, 1987.
Printed in Great Britain.

Chemoenzymatic synthesis of a gene for the interleukin-2 receptor

Joachim W. Engels, Wolfgang Becker *, Thomas Beckers, Dieter Hüsken and Waldemar Wetekam*

Institut für Organische Chemie der Universität Frankfurt,
6000 Frankfurt/Main 50, FRG
*Hoechst AG, P.O.Box 800320, 6230 Frankfurt/Main 80, FRG

Abstract
The human interleukin-2-receptor is a glycopeptide of an apparent molecular weight of 55kD. By chemoenzymatic gene synthesis we constructed the underlying gene consisting of 753 nucleotide base pairs. 38 oligonucleotides varying in length from 27 to 51 bases were joined at their free overlapping ends of 6 to 9 bases by T4 ligase. The resulting gene, subcloned into three pieces was sequenced for its correctness. For subsequent modifications several restriction enzyme cutting sites were introduced synthetically. For expression of this gene in E. coli different ß-galactosidase fusion constructions were tested. In addition the synthesis of the glycoprotein will be studied by introducing this gene into higher cells where it can be directly compared to the natural sequence in structure function studies.

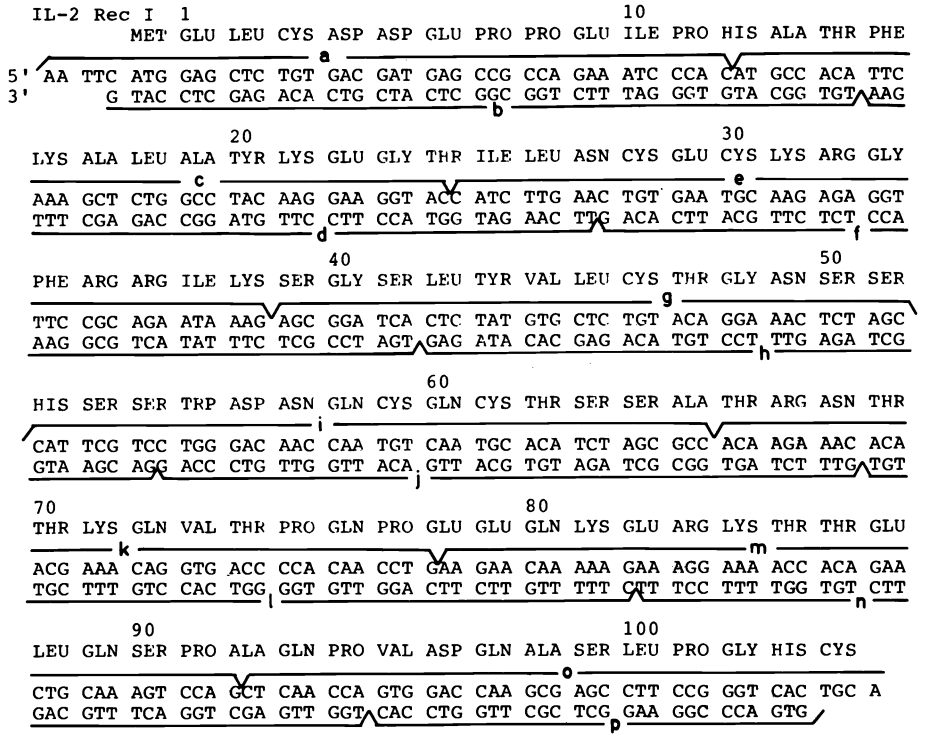
INTRODUCTION

The human immune system is prepared to recognize an almost unlimited variety of antigens. Within this framework thymus-derived lymphocytes (T-cells) do get in contact with foreign antigens via the T-cell-receptor. A consequence of this event is the liberation of a T-cell growth factor, the so called lymphokine interleukin-2 (IL-2) (Ref.1,2). In order to be biologically active this IL-2 must interact with a specific membrane receptor on the surface of T-cells, the interleukin-2-receptor (IL-2-Rec) (Ref.3,4). This receptor is not present on resting T-cells but is rapidly expressed after activation with an antigen (Ref.5). Since neither IL-2 nor IL-2-Rec are expressed constitutively the induction of these two genes is of vital importance to the immune response. Recently by the help of anti-Tac (Ref.6) monoclonal antibody and recombinant DNA-techniques, the cDNA coding for IL-2-Rec was identified (Ref.7,8,9). This gene consits of 753 base-pairs which code for a glycopeptide of 55kD. The precursor protein a 33kD peptide is co- or posttranslationally N- and O-glycosylated, phosphorylated (Ref.10) and sulfated (Ref.11) to the apparent molecular weight. In addition the genomic DNA-structure (Ref.12,13) consisting of eight exons on chromosome 10 was elucidated. This DNA includes a leader sequence of 63 base pairs which is cleaved off during maturation of the protein. In order to get some insight into the complicated events which govern the interaction of IL-2 and IL-2-Rec we have devised a synthetic gene allowing easy modifications by mutations to elucidate structure function relationsships of modified proteins.

CHEMOENZYMATIC GENE SYNTHESIS

Our gene synthesis is based on the chemical synthesis of the appropriate oligonucleotides which are joined enzymatically (Ref.14). On the basis of

the published DNA and protein sequence (Ref. 7,8,9) we constructed our gene but added some especially designed features for future handling. This results in a module system which allows for future modifications including exploratory and functional ones. The total gene consits of 753 base pairs and is divided into three DNA-segments (s.Fig.1) I, II and III.



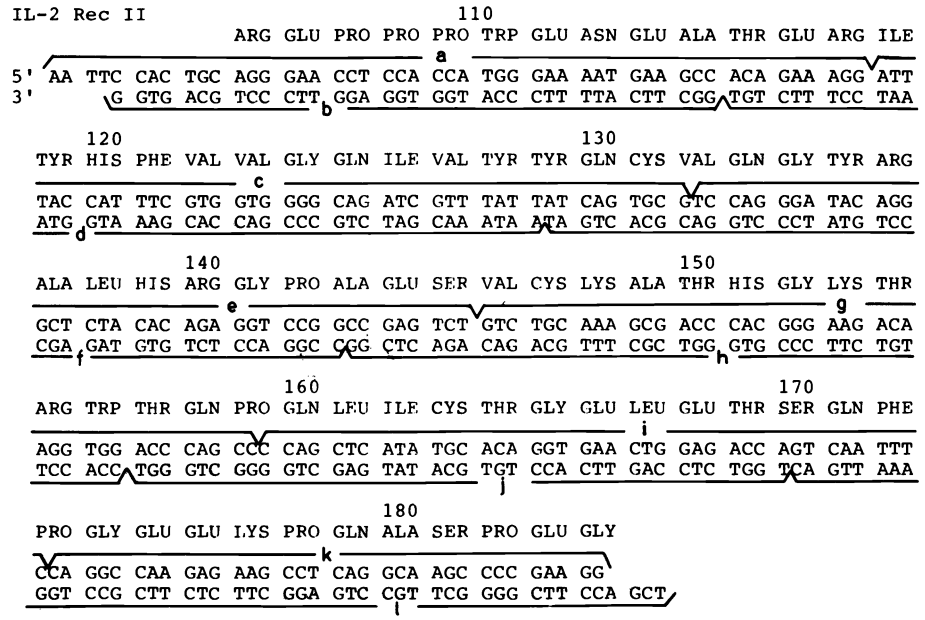


Fig.1 (continued on p.439)

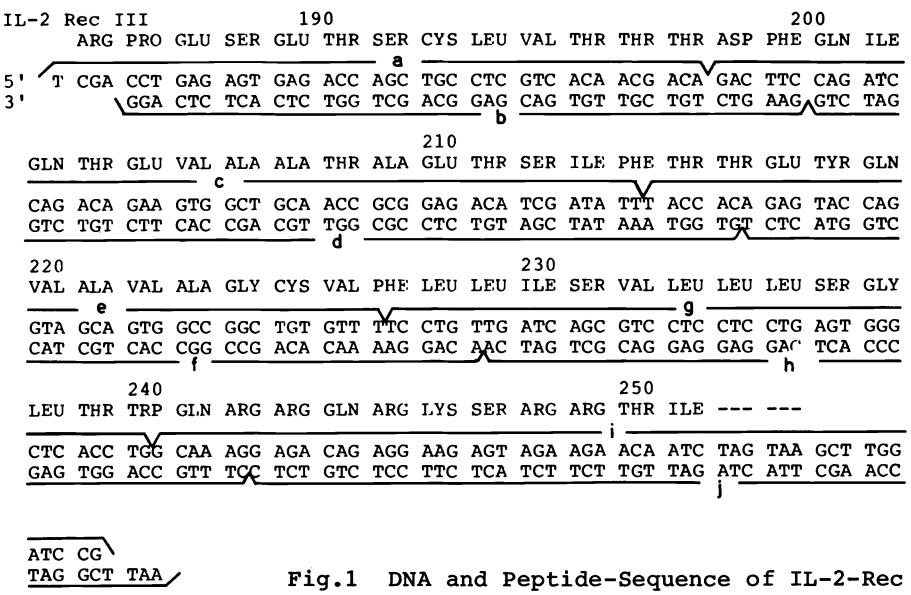


Fig.1 DNA and Peptide-Sequence of IL-2-Rec

They divide the IL-2-Rec gene in three segments, defined by the individual restriction endonuclease sites Eco RI-Pst I, Pst I-Sal I, Sal I-Eco RI (Hind III,Bam H I). Each subfragment is constructed from 10 to 16 individual oligonucleotides ranging in length from 27 to 51 bases. The retrosynthetic approach cuts the double stranded DNA in such a way, that there are free overlapping ends of 6 to 9 bases. In order to guarantee the exact finding of the individual complementary oligonucleotides the DNA duplex was devised by a computer program. This program checks wether the opposite strands a/b, c/d, c/f, do pair preferentially over e.g. a/d or a/f (Fig.1). Especially the free overlapping ends of 6 - 9 bases within the whole segment I, II and III have to be checked for their unambiguous matchings. Furthermore the segments I - III were modified on the DNA-level due to its degeneracy in order to introduce additional unique restriction enzyme cutting sites. This could be accomplished either by introducing new sites as e.g. Kpn I, Sal I, Cla I, or by deleting existing double or triple sites as e.g. Nco I, Bgl I. The chemical synthesis of the oligonucleotides Ia-IIIj is shown in (Fig.1). They were synthesized by the phosphoramidite chemistry (Ref.15) outlined in (Fig.2) with an additional chemical modification (Ref.16). The efficiency of

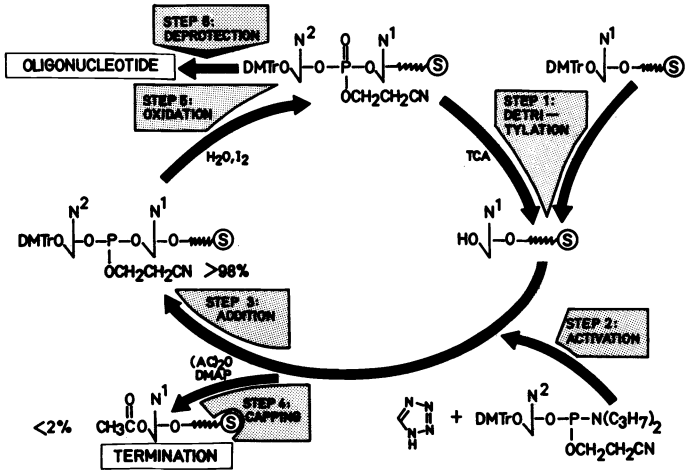


Fig.2 Synthesis cycle for the phosphoramidite chemistry

the coupling steps was estimated by measuring the dimethoxytrityl cation, liberated in each cycle (step one in Fig.2) at 498 nm and was routinely >96%. With the help of the ß-cyanoethyl protecting group at phosphorus the mild ammonia cleavage directly results in the fully unprotected oligonucleotides Ia - IIIj. The crude reaction mixtures were liberated from ammonia by evaporation and directly loaded on a preparative 12% polyacrylamide gel (3 mm thick). The bands were visualized by uv shadowing, cut out and the oligonucleotides eluted with water, desalted by gel chromatography and freeze dried. The yields range from 10-30%. In order to join these oligonucleotides enzymatically a phosphate group has to be introduced at the 5'-hydroxyl end. This can either be done enzymatically with T4-Polynucleotide kinase or chemically (Ref.17). The kinase reaction additionally allows the incorporation of ^{32}P for detection by autoradiography. 1 nmol of each oligonucleotide Ib-Io, IIb-IIk, IIIb-IIIi was kinased in Hepes-buffer pH 7.6 at 37°C with 5-10 units T4-polynucleotide kinase for 1 h. Only the 5'-ends of the restriction enzyme recognition sites were left unphosphorylated in order to avoid ligation of these palindromic sequences. By this procedure the enzymatic ligation can be accomplished in an one step test tube reaction by annealing the individual oligonucleotides at elevated temperature followed by joining with T4-ligase in a fashion similar to the one introduced by Khorana (Ref.18). For example oligonucleotides IIIa-IIIj were heated in Hepes-buffer pH 7.6, 0.05 M NaCl for 5 min at 98°C, cooled down to ambient temperature, and joined at 14°-20°C within 5 to 20 hours with 1000-1500 units T4-ligase added. The course of the reaction was followed by gel electrophoresis, checking the ligation mixture against DNA-duplexes of known size. For fragments I-III the expected DNA-duplex could be identified isolated by gel electrophoresis, desalted and lyophilized. The three segments IL-2-Rec I-III were then cloned in a pUC plasmid (Ref.19) in an E. coli K12 derivative (Fig.3). Due to a frame shift mutation and an additional stop codon in pUC8 and 9 the colonies harbouring the plasmide with the IL-2-Rec I - III inserts were white.

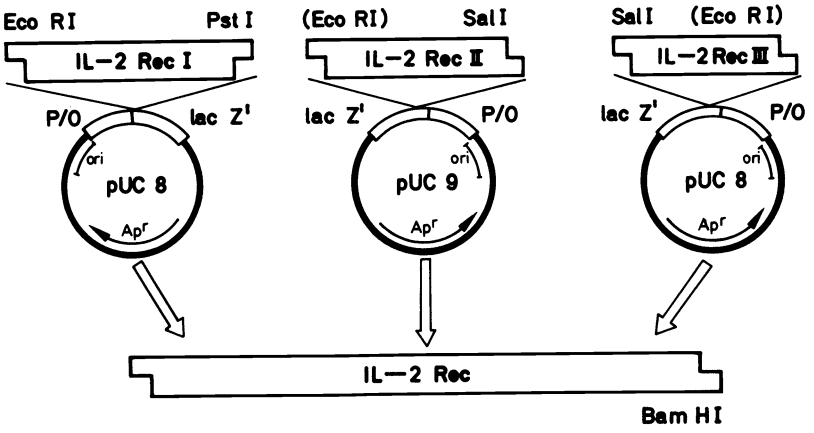


Fig.3 Subcloning of the IL-2-Rec fragments I-III in pUC 8 and 9

Several of these individual colonies were isolated, grown in liquid cultures and their recombinant pUC plasmids isolated according to Birnboim (Ref.20). We first checked their correct size on a gel against known size standards pBR 322 (Hae III) and secondly verified the synthetic DNA-sequence (Fig.1) by sequencing according to Maxam - Gilbert (Ref.21). For example the DNA sequence for the last 50 amino acids of the C-terminus is shown in Fig. 4.

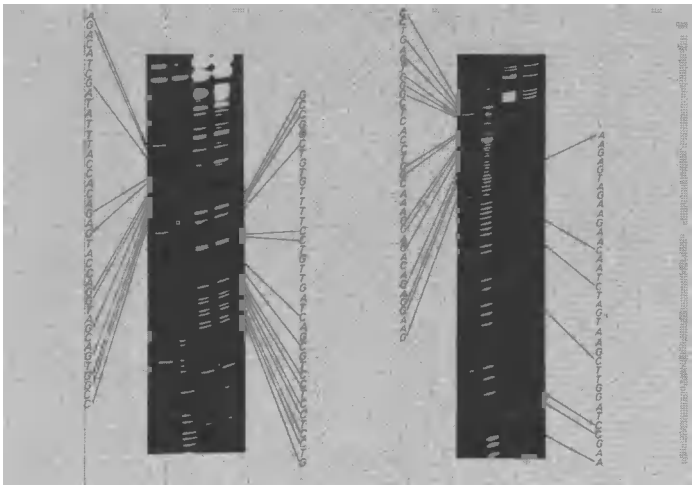


Fig.4 16% polyacrylamide sequencing gel long and short run

Four out of five clones had the correct sequence. The subfragments IL-2-Rec I-III were individually sequenced, isolated and finally joined enzymatically with ligase to the total gene (Fig.5).

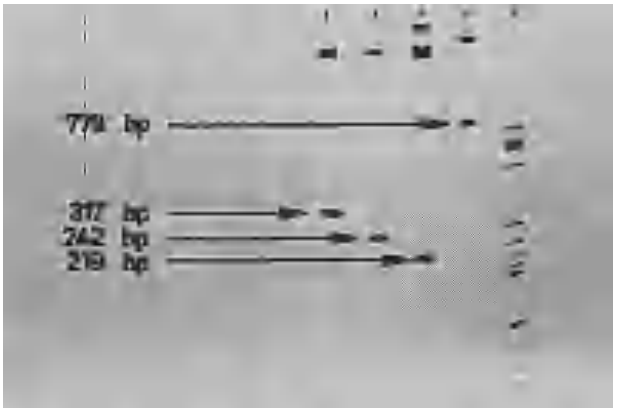


Fig.5 10% polyacrylamide gel electrophoresis: lane 1-3 IL-2 Rec I-III, lane 4 total IL-2 Rec, lane 5 pBR 322 (Hae III)

EXPRESSION-CONSTRUCTIONS OF IL-2-REC

Having confirmed the correct synthetic gene it has to be reconstructed in such a way as to synthesize the receptor protein within a cell. Our first approach was a fusion construction with ß-galactosidase from E. coli a protein which is synthesized after induction by ß-galactosidase in large amounts (Ref.22). This protein precipitates within the cell in a form of inclusion bodies (Ref.23) which can be isolated by centrifugation. In order to cleave the desired receptor protein from the ß-galactosidase fusion moiety, we envisaged a chemical cleavage reaction by cyanogen bromide

(Ref.24). Since cyanogen bromide specifically cleaves in a high yield after each methionine we had eliminated in our synthetic gene the original methionines. The replacements - originally selected on a secondary structure prediction (Ref.25,26,27) (Fig.6) were done with the aim to stabilize the

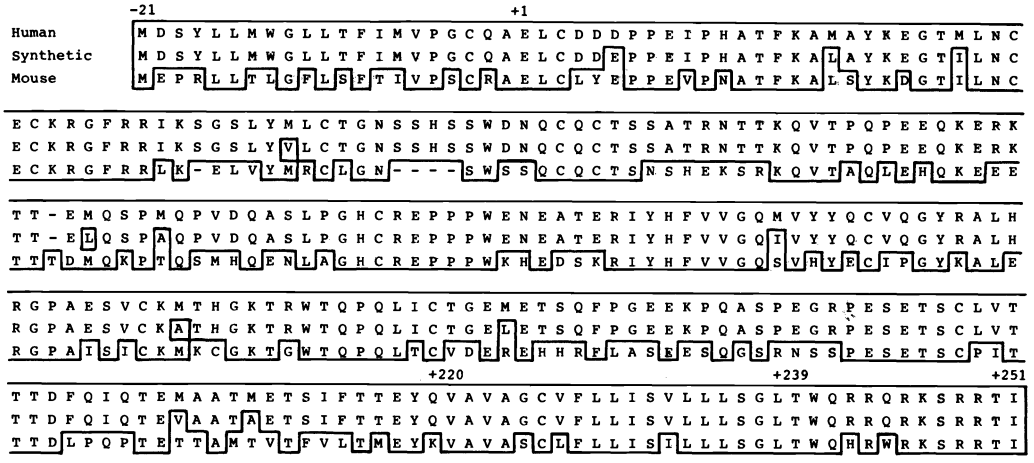
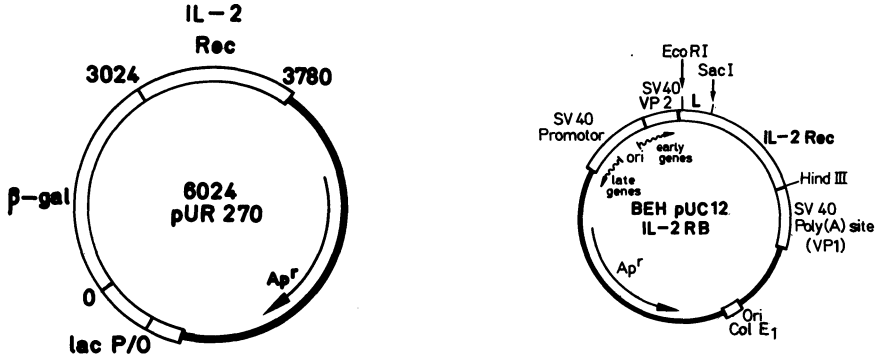


Fig.6 Amino acid sequence in one letter symbols of IL-2-Rec human, synthetic human and mouse (Ref.28), homologous sequences are boxed.



E. coli expression plasmid (pUR) Eucaryotic expression plasmid (BEH)

Fig.7

predicted secondary structure e.g. Met→Ala (ß-sheet), Met→Leu (α-Helix). For the expression of our synthetic gene in E. coli the construction was chosen as shown in (Fig.7).

In this plasmid the IL-2 Rec gene is joined carboxyterminally to either the total ß-galactosidase gene or part of it. We have chosen this fusion construction in order to mask the hydrophobic transmembrane region, of the IL-2 Rec by the ß-galactosidase moiety. This approach tries to circumvent the potential detrimental incorporation of the transmembrane region into the E. coli membrane. All these new constructions had to be verified by sequencing (Ref.20) and showed the correct DNA-sequence.

The above fusion constructions were transformed into an E. coli wild-type strain and the ß-galactosidase production was induced by Isopropylthio-ß-D-galactoside. This reaction could be followed by lysing the cells with ultrasonification and directly applying an aliquot of the protein mixture onto a 10% SDS-polyacrylamide gel, using the appropriate size standards. Here we observed a total lack of any detectable fusion protein in the expected size range. Based on our assumption, that the highly hydrophobic membrane region was responsible for this phenomenon, we also constructed a receptor ß-galactosidase gene fusion which only consisted of the assumed extracellular receptor part.
Beside the expression in E. coli which should result in the naked receptor protein we also initiated the gene construction for the expression of the gene in mammalian cells. Since the interleukin 2-receptor in mammalian cells is synthesized as a precursor having a leader sequence we first added the authentic leader (Ref.7,8,9) sequence to our synthetic gene (Fig.8).

```
          MET ASP SER TYR LEU LEU MET TRP GLY LEU LEU THR PHE ILE MET VAL
5' AA TTC ATG GAT TCA TAC CTG CTG ATG TGG GGA CTG CTC ACG TTC ATC ATG GTG
3'      G TAC CTA AGT ATG GAC GAC TAC ACC CCT GAC GAG TGC AAG TAG TAC CAC

  PRO GLY CYS ASN ALA
  CCT GGC TGC CAG GCA GAG CT
  GGA CCG ACG GTC CGT C
```

Fig.8 - Leader sequence of the IL-2-Rec gene

This sequence was constructed of two oligonucleotides 65 and 73 bp long. After phosphorylation and hybridisation they were joined directly to the IL-2-Rec gene and cloned in pUC 19 (Ref.29). Their sequence was checked by Maxam-Gilbert sequencing (Ref.20). For the expression of this gene in mammalian cells the shuttle vector (BEH) shown in Fig.7 will be used.
Work is in progress to study the various fusion constructions in E. coli and to monitor the synthesis of the IL-2 receptor in mouse L-cells.
The goal of our work is ultimately to study ligand receptor interaction in vitro including artificial membrane vesicles. We hope to achieve this by synthesising the receptor gene as an ample reservoir of the receptor protein in its natural and mutated form. The next step will include the synthesis of modified receptor proteins in E. coli and the synthesis of authentic receptor protein in mammalian cells. Comparison between these two receptors in their binding efficiency to IL-2 will give us an insight into the structural prerequisites for high efficiency binding. Mosaic structures between modified and authentic receptors are under construction to further elucidate the structure function relationship of IL-2 receptor and IL-2.

Acknowledgement
We are grateful for the financial support given by DFG (SFB 169), Hoechst AG and Fonds der chemischen Industrie. We are indebted to H.J. Hauser, GBF, Stöckheim, for a generous gift of the expression plasmid (BEH).

REFERENCES

1. Gallo, R.C., Ruscetti, F.W., Morgan, D.A., Science, 193, 1007, (1976).
2. Taniguchi, T., Matsui, H., Fujita, T., Takaoka, C., Kashima, N., Yoshimoto, R., Hamuro, J., Nature, 302, 305 (1983).
3. Leonard, W.J., Depper, J.M., Uchiyama, T., Smith, K.A., Waldmann, T.A., Greene, W.C., Nature, 300, 267 (1982).
4. Leonard, W.J., Depper, J.M., Robb, R.J., Waldmann, T.A., Greene, W.C., Proc. Natl. Acad. Sci. USA, 80, 6957 (1983).
5. Cantrell, D.A., Smith, K.A., J. exp. med. 158, 1895 (1983).
6. Uchiyama, T., Broder, S., Waldmann, T.A., J. Immunol. 126, 1393 (1981).
7. Leonard, W.J., Depper, J.M., Crabtree, G.R., Rudikoff, S., Pumphrey, J., Robb, R.J., Krönke, M., Svetlik, P.B., Peffer, N.J., Waldmann, T.A., Greene, W.C., Nature, 311, 626 (1984).
8. Nikaido, T., Shimizu, A., Ishida, N., Sabe, H., Teshigawara, K., Maeda, M., Uchiyama, T., Yodoi, J., Honjo, T., Nature, 311, 631 (1984).
9. Cosman, D., Cerretti, D.P., Larsen, A., Park, L., March, C., Dower, S., Gillis, S., Urdal, D., Nature, 312, 768 (1984).
10. Gallis, B., Lewis, A., Wignall, J., Alpert, A., Mochizuki, D.Y., Cosman, D., Hopp, T., Urdal, D., J. Biol. Chem., 261, 5075 (1986).
11. Depper, J.M., Krönke, M., Robb, R.J., Waldmann, T.A., Greene, W.C., J. Biol. Chem., 260, 1872 (1985).
12. Ishida, N., Kanamori, H., Noma, T., Nikaido, T., Sabe, H., Suzuki, N., Shimizu, A., Honjo, T., Nucl . Acids Res., 13, 7579 (1985).
13. Leonard, W.J., Depper, J.M., Kanehias, M., Krönke, M., Peffer, N.J., Svetlik, P.B., Sullivan, M., Greene, W.C., Science, 230, 633 (1985).
14. Engels, J., Leineweber, M., Uhlmann, E., Proc. of the 5th IUPAC Symp. on Org. Synth., 205, Streith, Prinzbach, Schill Eds., Blackwell Sci. Oxford (1985).
15. Beaucage, S.L., Caruthers, M.H., Tetrahedron Letters, 22, 1859 (1982).
16. Köster, H., Sinha, N.D., Biernat, J., McManus, J., Nucl . Acid Res., 12, 4539, (1984).
17. Uhlmann, E., Engels, J., Tetrahedron Lett., 27, 1023 (1986).
18. Khorana, H.G., Science, 203, 614 (1979).
19. Vieira, J., Messing, J., Gene, 19, 259 (1982).
20. Birnboim, H.C., Doly, I., Nucl. Acids Res., 7, 1513 (1979).
21. Gilbert, W., Maxam, A., Methods Enzymol., 65, 499 (1980).
22. Itakura, K., Bolivar, F., Hirose, T., Boyer, H.W., Crea, R., Riggs, A.D., Heyneker, H.L., Science, 198, 1056 (1977).
23. Taylor, G., Marston, F.A.O., Hoare, M., Gray, D.R., Bio/Technology, 4, 553 (1986).
24. Gross, E., Witkop, B., J. Am. Chem. Soc., 83, 1510 (1961).
25. Hopp, T.P., Woods, K.R., Proc. Natl. Acad. Sci. USA, 78, 3824 (1981).
26. Kyte, J., Doolittle, A.F., J. Mol. Biol., 157, 105 (1982).
27. Chou, P.Y., Fasman, G.D., Biochemistry, 13, 222 (1974).
28. Shimuzu, A., Kondo, S., Takeda, S.I., Yodoi, J., Ishida, N., Sabe, H., Osawa, H., Diamantstein, T., Nikaido, T., Honjo, T., Nucl. Acids Res., 13, 1505 (1985).
29. Yanisch-Perron, C., Vieira, J., Messing, J., Gene, 33, 103 (1985).