Vol. 154, No. 1, 1988 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
July 15, 1988 Pages 372-379

STRUCTURAL ANALYSIS OF RECOMBINANT SOLUBLE HUMAN INTERLEUKIN-2 RECEPTOR
Primary Structure, Assignment of Disulfide Bonds and Core IL-2 Binding Structure

May C. Miedel, Jeffrey D. Hulmes, David V. Weber, Pascal Bailon and Yu-Ching E. Pan*

Department of Protein Biochemistry, Roche Research Center, Hoffmann-La Roche Inc., Nutley, NJ

Received June 3, 1988

A purified soluble and functional form of recombinant human interleukin-2 receptor, engineered and expressed in Chinese hamster ovary cells, was structurally characterized. The primary sequence of this 224 amino acid recombinant protein which lacks most of the carboxy-terminal transmembrane and cytoplasmic portions of the intact protein was established by sequence analyses. The disulfide bonds were assigned by comparative peptide mapping of the reduced and non-reduced peptide digests. As in the case of natural interleukin-2 receptor they occur between cysteines 3-147, 46-104, 131-163, and 28/30-59/61. Based on assignment of the disulfide bonds, a structural model of the interleukin-2 receptor for interleukin-2 binding is proposed.

The stimulation of T cells with antigen or mitogen induces the synthesis of interleukin-2 (IL-2) and IL-2 receptor (IL-2R) (1,2). The binding of IL-2 to the IL-2R on human T cells is a key regulatory event during an immune response. Recent advances in the understanding of the low affinity IL-2R (Tac) and its encoding gene provided insights into the structural features of this receptor (for review, see reference 3). Briefly, at the genomic level, this protein is encoded by 8 exons (4). The mature protein consists of 251 amino acids including an extracellular domain of 219 amino acids, a transmembrane region of 19 amino acids, a short carboxy-terminal cytoplasmic tail of 13 amino acids and several glycosylation sites (5-7). The amino acid sequence of the IL-2R molecule reveals two homologous regions corresponding to residues 1-64 and 102-174 encoded by exons 2 and 4, respectively (4). The N-terminal region of the IL-2R has been shown to contain important contact sites for IL-2 interaction (8). Furthermore, truncated IL-2R, lacking the C-terminal transmembrane and cytoplasmic regions via proteolysis still

*To whom correspondence should be addressed.

Abbreviations: IL-2, interleukin-2; IL-2R, interleukin-2 receptor; rIL-2R, recombinant interleukin-2 receptor; CHO, Chinese hamster ovary; PTH, phenylthiohydantoin; NCS, N-chlorosuccinimide; TFA, trifluoroacetic acid.

0006-291X/88 $1.50

exhibits IL-2 binding (9,10). To facilitate the study of the interaction of IL-2 with IL-2R, a recombinant soluble human IL-2R (rIL-2R) which lacks the carboxyterminal transmembrane and cytoplasmic portions of the intact IL-2R has been expressed in Chinese hamster ovary (CHO) cells (11,12). This protein has an apparent M_r of 43 kDa and still binds IL-2. Here we report the determination of the primary structure and the assignment of disulfide bonds for this protein. Based on the disulfide bond arrangement, we propose a structural model of IL-2R which is of importance for IL-2 binding.

MATERIALS AND METHODS

Recombinant soluble human IL-2R, cloned and expressed in CHO cells (11), was purified by recombinant IL-2 affinity gel (13).

Amino acid analysis was performed on an instrument using postcolumn reaction with fluorescamine (14). Sequence analysis was performed using an Applied Biosystem gas phase sequencer model 470A (Foster City, CA) (15). Phenylthiohydantoin (PTH) amino acid derivatives were identified "on-line" with an ABI model 120A PTH analyzer.

Proteolytic digestions of rIL-2R with a variety of proteases were performed. Trypsin-TPCK was purchased from Cooper Biomedical (Philadelphia, PA). S. aureus V8 protease, chymotrypsin, thermolysin, and proteinase K were purchased from Boehringer Mannheim (Indianapolis, IN). Digestion with trypsin and V8 protease were performed in 0.2 M ammonium bicarbonate at a substrate-to-enzyme ratio of 30:1 (w/w) at 37°C for 20 hours. Digestion with thermolysin was performed in 0.2 M ammonium bicarbonate in the presence of calcium chloride with a ratio of 50:1 (w/w) at 37°C for 22 hours. Digestion with proteinase K was performed in 50 mM Tris-HCl, pH 8.0, with a ratio of 100:1 (w/w) at room temperature for one hour. Triple digestion with trypsin, S. aureus V8 protease, and chymotrypsin was performed simultaneously in 0.2 M ammonium bicarbonate with a ratio of 25:1 for each enzyme at 37°C for 20 hours. Chemical cleavage of rIL-2R with N-chlorosuccinimide (NCS) was performed in urea (16). All other reagents were of purest commercial grade.

HPLC peptide mapping of the enzymatic digests was performed with either a low-pressure mixing system equipped with a postcolumn fluorescamine detection system (17), or a Hewlett-Packard (H-P) 1090A system (Avondale, PA). Peptides were eluted with an acetonitrile gradient in 0.1% trifluoroacetic acid (TFA) and peptide-containing fractions were collected for further analysis. For reduced peptide maps, digests were treated with 2-mercaptoethanol at 100°C for 5 minutes before injection onto the column.

RESULTS AND DISCUSSION

The complete amino acid sequence, as given in Fig. 1, for this rIL-2R of 224 amino acids was established by N-terminal sequencing of the intact protein and overlapping peptide sequences from various HPLC peptide maps (profiles in Fig. 2). NCS, which cleaves specifically at the C-terminal side of Trp, was used to identify the location of the three Trp residues. The peptide mixture was desalted by reversed-phase HPLC followed by N-terminal sequencing. Three new N-terminal sequences generated by cleavage at the C-terminal side of Trp^{55}, Trp^{110} and Trp^{156} confirmed the presence of the Trp residues (Fig. 1). Cys residues were identified by separating and sequencing peptides under reducing and non-reducing conditions (see below) and/or by comparison to the sequence of the IL-2R determined by cDNA. The recovery of C-terminal peptides from the V8 and proteinase K maps (V7 and PK6) confirmed the identity of the C-terminus of the polypeptide chain. During the course of peptide sequence analysis, the two putative N-linked glycosylation sites (49-51 and 68-70) and several O-linked glycosylation sites were

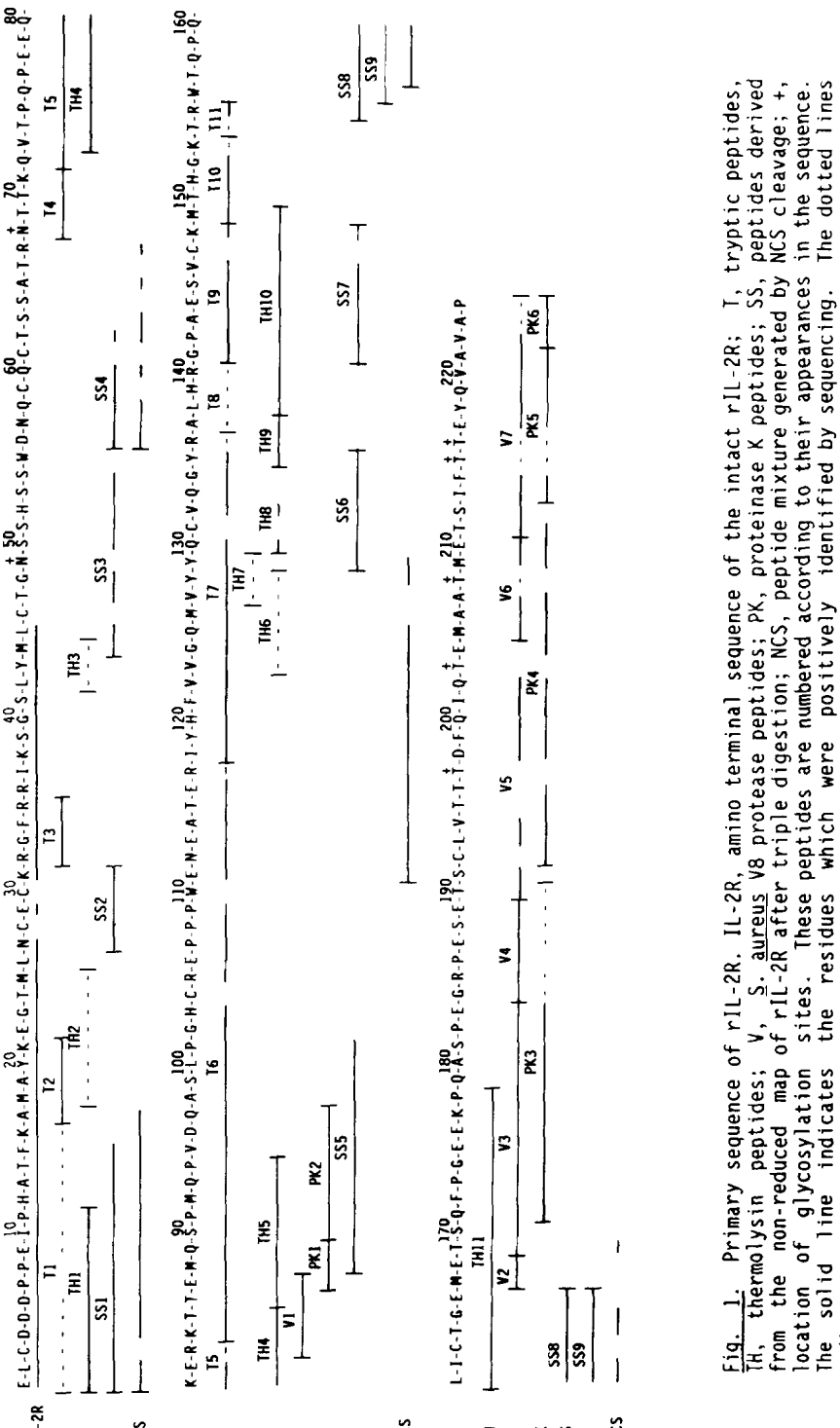

Fig. 1. Primary sequence of rIL-2R. IL-2R, amino terminal sequence of the intact rIL-2R; T, tryptic peptides, TH, thermolysin peptides; V, S. aureus V8 protease peptides; PK, proteinase K peptides; SS, peptides derived from the non-reduced map of rIL-2R after triple digestion; NCS, peptide mixture generated by NCS cleavage; +, location of glycosylation sites. These peptides are numbered according to their appearances in the sequence. The solid line indicates the residues which were positively identified by sequencing. The dotted lines indicate residue assignments based on amino acid compositional analysis of the isolated peptides.

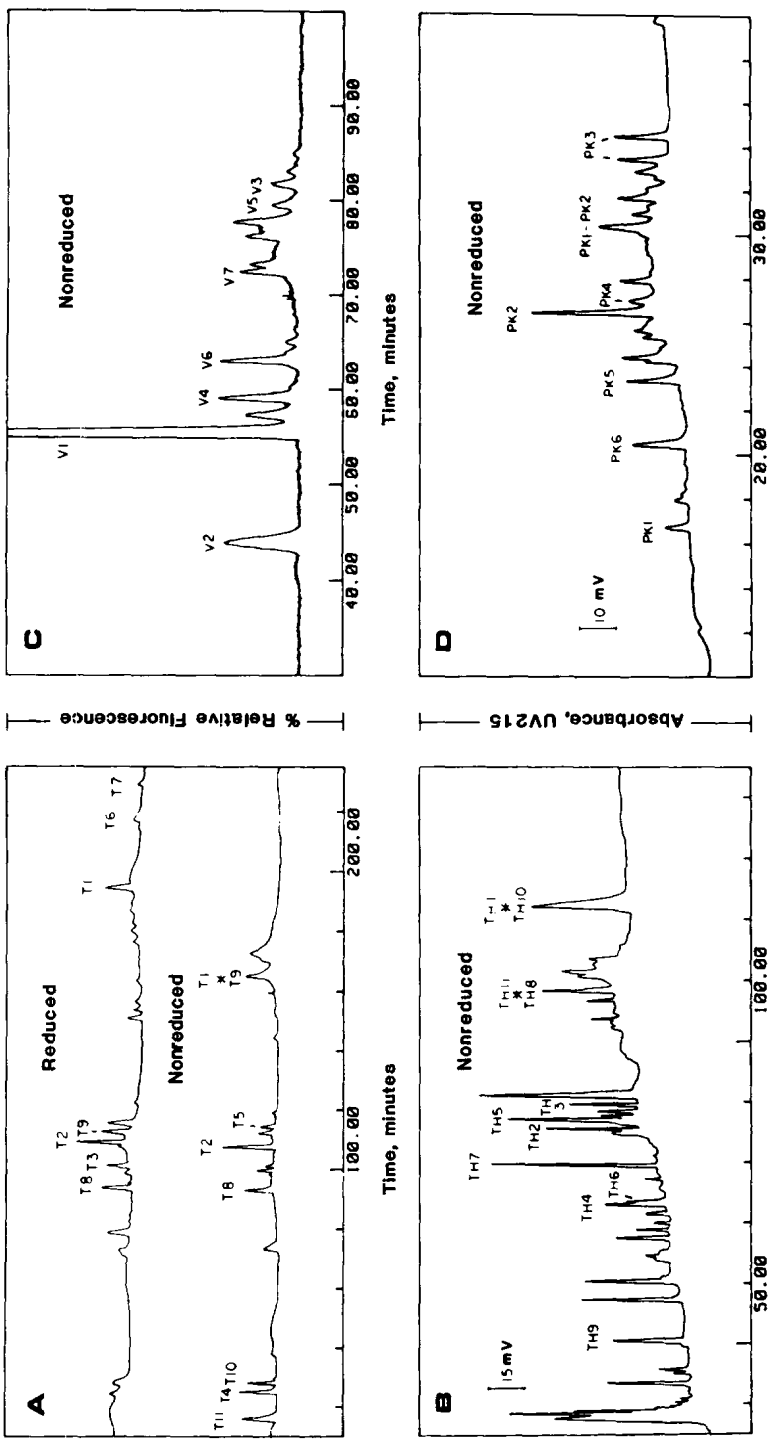

Fig. 2. HPLC peptide maps of rIL-2R after digestion by trypsin (A), thermolysin (B), S. aureus V8 protease (C) and proteinase K (D). Peptide digests generated by tryptic, thermolysin, and V8 were analyzed by a reversed-phase C_{18} column (4.6 x 250 mm, Beckmann/Altex, Berkley, CA) using post column fluorescamine or UV215 detection. Peptide digests generated by proteinase K were analyzed by a reversed-phase C_{18} column (2 x 150 mm, Phase Separation Inc., Norwalk, CT) using the H-P 1090A. The sequenced peptides are marked and numbered according to their appearances in the sequence (Fig. 1).

also encountered. Glycosylated Asn, Ser and Thr cannot be detected directly in the PTH amino acid form by HPLC due to the fact that their anilinothiazolinone derivatives are insoluble in chlorobutane from the sequence filter (18). The assignment of glycosylated Asn^{49}, Asn^{68}, Thr^{197}, Thr^{203}, Thr^{208}, Thr^{215} and Thr^{216} in the peptides SS3, T4, V5, V6, V7, PK4 and PK5 (Figs. 1 and 2) was made based on the absence of the PTH amino acids in the particular Edman cycles and by comparison to the predicted sequence of IL-2R. The unambiguous identification of other unmodified residues (e.g. Asn^{112} and Thr^{150}) suggested that the inability to detect certain Asn and Thr residues is due to the post-translational modification events. No Ser residues were found to be glycosylated. The predicted M_r of the rIL-2R protein from its sequence is 25,261 daltons. This is much lower than the apparent M_r of 43 kDa observed by SDS-PAGE probably as a result of the post-translational modifications. The calculated amino acid composition of the rIL-2R is in good agreement with the composition derived from sequence analysis (in parentheses) as follows: $Asx_{11.7}$ (Asp_6, Asn_5), $Thr_{19.7}$ (25), $Ser_{14(16)}$, $Glx_{41.7}$ (Glu_{23}, Gln_{18}), $Pro_{18.2(18)}$, $Gly_{13.7(12)}$, $Ala_{14.9(13)}$, $Cys_{10.4(11)}$, $Val_{10.9(10)}$, $Met_{9.8(10)}$, $Ile_{6.1(6)}$, $Leu_{8.8(8)}$, $Tyr_{6.5(7)}$, $Phe_{5.7(6)}$, $His_{6.5(6)}$, $Lys_{9.7(10)}$, $Arg_{12.3(11)}$, $Trp_{1.6(3)}$.

The first step taken toward an understanding of the overall conformation of IL-2R was to determine disulfide bonds among the eleven Cys residues in the protein. Triple digestion of the rIL-2R was performed with trypsin, chymotrypsin and V8 protease to generate small peptide fragments. Peptide digests with and without mercaptoethanol treatment were analyzed by reversed-phase HPLC mapping, and the chromatograms were compared (Fig. 3). The peptide peaks which were present in the non-reduced map and absent in the reduced map were likely to be involved in disulfide bonds and therefore were subjected to sequence analyses (marked in Fig. 3). Each of these peaks yielded two Cys containing sequences in approximately equimolar amounts corresponding to disulfide linked peptides. The following pairs of disulfide bonds were identified: Cys^3-Cys^{147}, Cys^{46}-Cys^{104}, and Cys^{131}-Cys^{163}. The disulfide binding relationship between Cys^{28} and Cys^{30} with Cys^{59} and Cys^{61} cannot be assigned with confidence. Since the cysteines are separated by one amino acid residue, it was difficult to assign the correct disulfide bond(s) to the paired fragments. In either case, because of the proximity of the Cys residues (whichever combination of disulfide bonds prevails), the overall conformation of the IL-2R molecule should be maintained. Recently, the arrangement of most of the disulfide bonds in the natural IL-2R (Tac) has been determined by analyzing enzymatic fragments obtained from double digestion with trypsin and chymotrypsin (19). Basically, the same results were obtained with the exception that the fragments containing Cys^{46}-Cys^{104} were impure and recovered in low yield. In the present study, this disulfide bond was identified without ambiguity due to the presence of a sensitive V8 cleavage site at Glu^{87} which released a paired fragment with good yield (Fig. 3, SS3*SS5). Several other minor reducible peptide peaks were identified as being derived from the peptides involving Cys^{131}-Cys^{163} and Cys^{46}-Cys^{104}. The presence of these minor peptides was

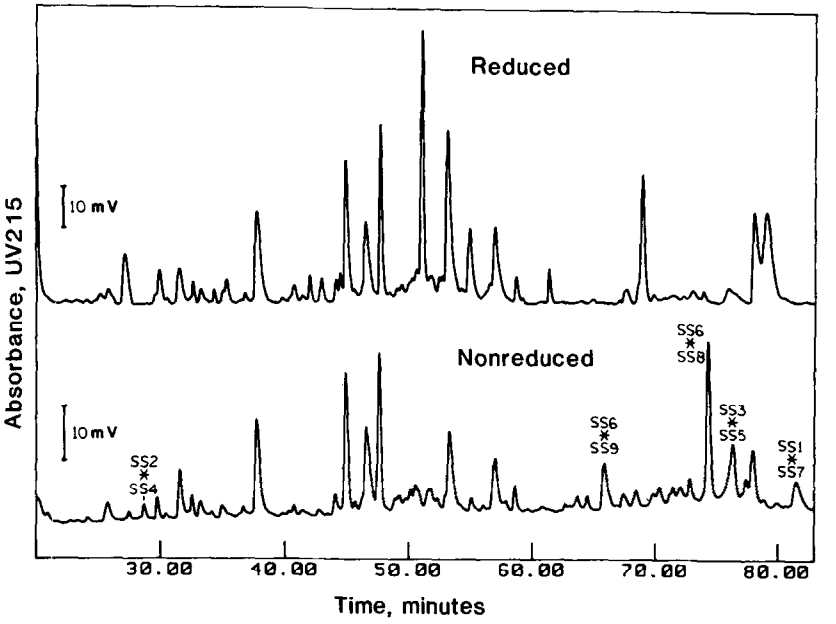

Fig. 3. Reducing and non-reducing HPLC peptide maps of rIL-2R after triple digestion with trypsin, chymotrypsin and S. aureus V8 protease. The analyses were performed with a C_{18} reversed-phase column (2x 150 mm, Phase Separation Inc.).

due to the incomplete cleavage of several chymotryptic sites. The assignment of the disulfide bonds suggested that the remaining Cys at position 192 must be present as a free sulfhydride. The results revealed that the natural and rIL-2R share the same disulfide configuration; therefore, rIL-2R can be used to elucidate the structure/function relationships of IL-2R.

In the absence of x-ray data of IL-2R, a structural model is proposed based on the determined disulfide bonds (Fig. 4). It can be used to describe some of the structural features regarding IL-2/IL-2R interaction. In this model, the two homologous domains encoded by exons 2 and 4, connected by the hydrophilic segment (65-102) encoded by exon 3, are linked by two inter- and three (or two) intra-domain disulfide bonds leaving the peptide region encoded by exons 5 and 6 (174-224) dangling. Several lines of evidence in both the natural (8,10,20) and this rIL-2R[1] have indicated that the hydrophilic segment and the C-terminal region encoded by exons 5 and 6 are sensitive to proteolytic digestion. Under mild conditions, they can be released without affecting IL-2 binding. Upon removal of these two peptide regions from the rIL-2R molecule, a "core" structural model for IL-2 binding, as shown in Fig. 4, is obtained. In this model, the two homologous domains form two nearly identical configurations oriented in an inverse relationship. This is consistent with the notion that disulfide bonding has an

1 Pan, Y.-C. E., et al., International Meeting on Biochemistry, Beijing, China (1986); and unpublished results.

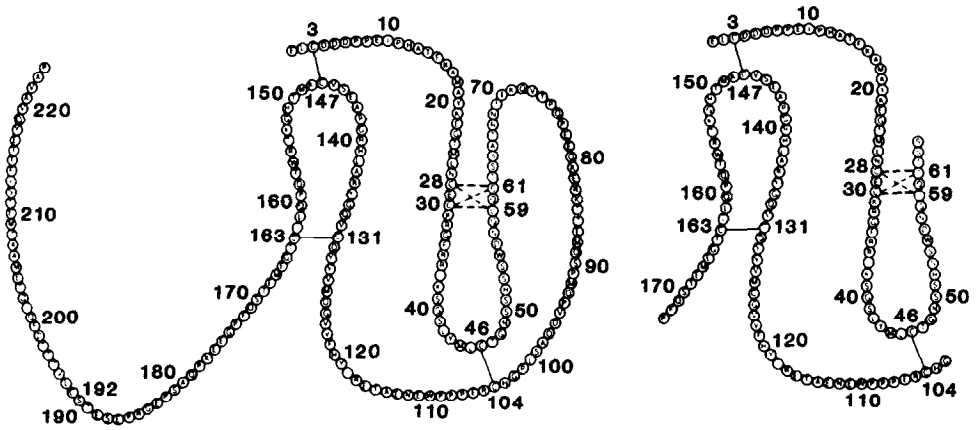

Fig. 4. Structural models of the soluble rIL-2R (left) and the core IL-2 binding structure in the IL-2R molecule (right). The disulfide bonds are denoted by lines linking the Cys pairs. The possible pairing between Cys^{28}/Cys^{30} - Cys^{59}/Cys^{61} are denoted by dotted lines.

essential role in the function of IL-2R binding IL-2 (19,21). Using limited proteolytic digestion, we were able to generate a "core" protein molecule which resembles the structural model shown in Fig. 4. This molecule is the smallest protein moiety of the IL-2R (approximately 135 amino acids) so far known to bind IL-2 (manuscript in preparation). The availability of this "core" rIL-2R molecule has provided us with the opportunity to further analyze the detailed structure-function relationship of the IL-2/IL-2R interaction.

ACKNOWLEDGEMENTS

The authors thank Drs. J. Jenson and J. Smart for their advice and support. Review of this manuscript by Mr. W. Levin and Dr. T. Mowles was greatly appreciated. We also thank Cathy Michaud for her assistance in the preparation of the manuscript.

REFERENCES

1. Morgan, D. A., Ruscetti, F. W., and Gallo, R. C. (1976) Science 193, 1007-1008.
2. Robb, R. J., Munck, A., and Smith, K. A. (1981) J. Exp. Med. 154, 1455-1474.
3. Green, W. C. and Leonard, W. J. (1986) Ann. Review Immuno. 4, 69-95.
4. Leonard, W. J., Depper, J. M., Kanehisa, M., Kronke, M., Peffer, N. J., Svetlik, P. B., Sullivan, M. and Greene, W. C. (1985) Science, 230, 633-639.
5. Leonard, W. J., Depper, J. M., Crabtree, G. R., Rudikoff, S., Pumphrey, J., Robb, R. J., Kronke, M., Svetlik, P. B., Peffer, N. J., Waldmann, T. A., and Greene, W. C. (1984) Nature 311, 626-631.
6. Nikaido, T., Shimizu, A., Ishida, N., Sabe, H., Teshigawara, K., Maeda, M., Uchiyama, T., Yodoi, J., and Honjo, T. (1984) Nature 311, 631-635.
7. Cosman, D., Cerritti, D. P., Larsen, A., Park, L., March, C., Dower, S., Gillis, S., and Urdal, D. (1984) Nature 312, 768-771.
8. Kuo, L.-M., Rusk, C. M., and Robb, R. (1986) J. Immunol. 137, 1544-1551.
9. Rubin, L. A., Kurman, C. C., Fritz, M. E., Biddison, W. E., Boutin, B., Yarchoan, R., and Nelson, D. L. (1985) J. Immunol. 135, 3172-3177.
10. Robb, R. J., and Kutny, R. M. (1987) J. Immunol. 139, 855-862.
11. Hakimi, J., Seals, C., Anderson, L. E., Podlaski, F. J., Lin, P., Danho, W., Jenson, J. C., Perkins, A., Donadio, P. E., Familletti, P. C., Pan, Y.-C. E., Tsien, W.-H., Chizzonite, R. A., Casabo, L., Nelson, D. L., and Cullen, B. R. (1987) J. Biol. Chem. 262, 17336-17341.

12. Cullen, B. R., Podlaski, F. J., Peffer, N. J., Hosking, J. B., and Greene, W. C. (1988) J. Biol. Chem. 263, 4900-4906.
13. Weber, D. V., Keeney, R. F., Familletti, P. C., and Bailon, P. (1988) J. Chromatography, Biomedical Application, in press.
14. Pan, Y.-C. E., and Stein, S. (1986) In Methods of Protein Microcharacterization (J. E. Shively, ed.), pp. 105-119, The Humana Press, Inc., Clifton, NJ.
15. Hewick, R. M., Hunkapillar, M. W., Hood, L., and Dryer, W. J. (1981) J. Biol. Chem. 256, 7990-7997.
16. Shechter, Y., Patchornik, A., and Burstein, Y. (1976) Biochem. 15, 5071-5075.
17. Stein, S., and Morschera, J. (1981) Methods Enzymol. (S. Pestka, ed.) 79, 7-16.
18. Hunkapillar, M. W. (1985) ABI Users Bulletin, Issue No. 14.
19. Rusk, C. M., Neeper M. P., Kuo, L,-M., Kutny, R. M., and Robb, R. (1988) J. Immunol. 140, 2249-2259.
20. Shackelford, D., and Trowbridge, I. S. (1986) The EMBO J. 5, 3275-3280.
21. Neeper, M. P., Kuo, L-M., Kiefer, M. C., and Robb, R. J. (1987) J. Immunol. 138, 3532-3538.