Automated Chemical Synthesis of a Protein Growth Factor for Hemopoietic Cells, Interleukin-3

IAN CLARK-LEWIS, RUEDI AEBERSOLD, HERMAN ZILTENER, JOHN W. SCHRADER, LEROY E. HOOD, STEPHEN B. H. KENT

Interleukin-3 (IL-3), a protein of 140 amino acids, was chemically synthesized by means of an automated peptide synthesizer and was shown to have the biological activities attributed to native IL-3. Assays of synthetic analogues established that an amino terminal fragment has detectable IL-3 activity, but that the stable tertiary structure of the complete molecule was required for full activity. The results demonstrate that automated peptide synthesis can be applied to the study of the structure and function of proteins.

A T LYMPHOCYTE–DERIVED LYMPHOKINE WITH POTENT growth-promoting activity for multiple hemopoietic cell lineages has recently been characterized (*1–3*). The murine lymphokine is active in several assays and as a result has been variously termed persisting cell-stimulating factor, mast cell growth factor, burst-promoting activity, hemopoietic cell growth factor and multi colony-stimulating factor (*4*). This factor has become more widely termed interleukin-3 (IL-3) (*3, 5*). It is thought to play a key role in the regulation of hemopoiesis by the immune system, and a role in the oncogenesis of hemopoietic cells has been proposed (*6*). Thus there is great interest in this glycoprotein and the structural basis and mechanisms of its activities. Recent characterization of isolated complementary DNA (cDNA) clones (*1*) has shown that IL-3 is encoded as a precursor of 166 amino acids. Amino terminal protein sequence analysis indicates that the mature protein has either 134 (*3*) or 140 (*2*) amino acids. However, lack of availability of purified IL-3 has hindered further structural analysis and questions concerning the relation between IL-3 structure and function have not yet been addressed. For example, does the deduced sequence code for a molecule with all the activities of IL-3; are the carbohydrate moieties necessary for activity; does the amino terminal heterogeneity have any functional significance; do fragments of the molecule possess any of the biological activities of the intact protein; which cysteines are involved in disulfide bridges and what role do they play in stabilizing the active folded structure of the protein; and how does IL-3 interact with its cell surface receptor?

In order to address these questions it is necessary to systematically vary the primary structure of IL-3 and assay the effects of these changes on biological function. In the past, there have been several approaches to the study of protein structure-function relations. (i) Protein structure can be modified by chemical reagents (*7*). (ii) Gene structure can be modified, for example, by site-directed mutagenesis and the resulting altered protein produced in appropriate expression systems (*8*). (iii) Modified proteins can be produced by total or partial chemical synthesis of the peptide chain. The first method is not adequate because of the limited types of changes that are possible and the lack of specificity of those changes, and the second is not always adequate because of the time-consuming nature of the methods and the fact that expression and purification of the altered product is not always straightforward. The third method has been applied only in a very limited number of cases, usually by linking synthetic peptides with natural protein fragments (*9*).

In this article we show that automated chemical synthesis can be a practical approach to structure-function studies of a protein. Our strategy takes advantage of recent improvements in the chemistry of solid-phase peptide synthesis together with the development of a fully automated peptide synthesizer. We have chemically synthesized the entire IL-3 molecule of 140 amino acids from the sequence predicted by translation of the corresponding cDNA clones and amino terminal protein sequence data (*1, 2*). The synthetic IL-3 had the spectrum of biological activities attributed to the native molecule. We also synthesized a number of IL-3 analogues, and comparison of their activities suggests that important structure-function correlations can be determined with the peptide synthesis approach and that, in particular, the integrity of a stable tertiary structure is crucial for IL-3 function.

Chemical synthesis of IL-3. The primary structure of IL-3 deduced from the cDNA nucleotide sequence (*1*) and amino terminal protein sequence analysis of purified native IL-3 (*2*) is shown in Fig. 1A. The protected peptide was assembled by solid-phase synthesis (*10*) on a cross-linked polystyrene support starting from the carboxyl terminal residue and adding amino acids in a stepwise fashion until the entire 140-residue chain had been formed. The synthesis was performed on a fully automated peptide synthesizer (Applied Biosystems 430A) (*11*), by means of chemical methods that incorporate cumulative improvements in the technology of chain assembly on the solid phase (*12*).

For the successful stepwise synthesis of long peptide chains (such as IL-3) it is important that each amino acid be added with close to 100 percent efficiency. Key to the high yield of our synthesis was the elimination of side reactions that occur with standard chemical methods (*13*) by the use of resins free of undesirable functional groups and stable to the conditions of chain assembly (*14*). Difficulties can occur in stepwise synthesis because of the aggregating properties of some sequences of protected amino acids (*15*). The effects of this were minimized by keeping the peptide chain solvated in dimethylformamide during the neutralization and amino acid

Ian Clark-Lewis, Ruedi Aebersold, Leroy Hood, and Stephen Kent are members of the Biology Division of the California Institute of Technology, Pasadena 91125. Herman Ziltener and John Schrader are members of the Walter and Eliza Hall Institute of Medical Research, Melbourne, Australia.

coupling steps (*11*). The amino acids were coupled as the highly reactive symmetric anhydrides (*16*), automatically formed immediately prior to use. A substitution of about 1 mmol of reactive groups per gram of resin was used to minimize volumes and achieve high concentrations of reactants. This strategy gave high yields with short (less than 30 minutes) reaction times (*17*).

In preliminary experiments we found a small increase in efficiency of amino acid addition (average 0.3 percent) when a second coupling was performed. Because of the cumulative effect of this increase in yield over the 140 residues of IL-3, this double coupling protocol was used routinely. N^{α}-*t*-butyloxycarbonyl (tBoc)-amino acids were used with appropriate side chain–protecting groups stable to the conditions of chain assembly but labile to strong acids (*18*). Both the amino acids and all the reagents and solvents used were of high purity to minimize side reactions (*19*). After assembly of the protected peptide chain, the protecting groups were removed and the peptide-resin anchoring bond was cleaved by the use of low and then high concentrations of anhydrous hydrogen fluoride in the presence of a thioether scavenger (*20*). Disulfide bridges were formed by means of the glutathione–oxidized glutathione protein refolding method (*21*). After refolding, no free thiols were detected with the Ellman assay (*22*). Three separate total syntheses of IL-3 were carried out with similar results. In each case, about 500 mg of the refolded product was obtained within 12 days of initiation of the chain assembly.

Chemical characterization of synthetic IL-3. The synthetic product was characterized by a number of highly discriminating complementary chemical and analytical techniques. First, the fidelity of the chain assembly was assessed by quantitative ninhydrin monitoring of each amino acid coupling (*23*). This assay determined the level of residual α-amino groups after each cycle during the synthesis as measure of the coupling efficiency at each step. In addition, the overall efficiency of the chain assembly was rigorously assessed by quantitative sequence analysis of the protected peptide chain while still bound to the resin (*24*). This method quantifies the cumulative efficiency of both removal of the N^{α}-protecting group and the extent of the amino acid coupling at each cycle of the synthesis from the degree of premature appearance, "preview," of amino acid derivatives during successive cycles of the Edman degradation (*24*). The quantitative sequence data for the synthesis of IL-3 are summarized in Table 1. Cumulative levels of preview were determined from sequencing runs on four intermediate samples of peptide-resin covering the 140–amino acid chain. Of the synthesized molecules 41 percent had the target sequence, corresponding to an average synthetic yield of 99.4 percent per amino acid residue. The remaining 59 percent is a mixture of closely related molecules, consisting primarily of peptides with a single internal amino acid missing.

After cleavage and deprotection, the synthetic IL-3 was first characterized by amino acid analysis which gave the expected composition. The primary structure of the refolded synthetic IL-3 was further analyzed by tryptic peptide mapping (Fig. 2A). After separation by reverse phase high-performance liquid chromatography (HPLC) tryptic peptides were identified by their amino acid compositions. Every fragment could be accounted for as a tryptic cleavage product of the target sequence (Fig. 2A). Automated

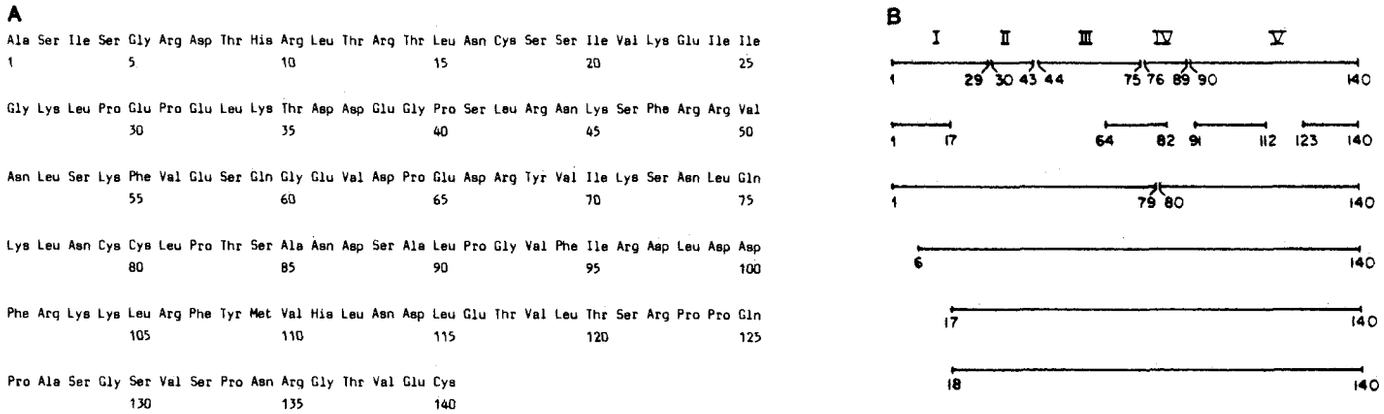


Fig. 1. (A) The amino acid sequence predicted for IL-3. The amino acid sequence was that deduced by translation of the nucleotide sequences of cDNA clones derived from WEHI-3B and T-cell messenger RNA (*1*). The nucleotide sequence for the T cell–derived cDNA indicated an alanine at position 120, whereas the sequence of the IL-3 gene (*28*) and WEHI-3B cDNA indicated a threonine at this position. The sequence with Thr^{120} was synthesized. The amino terminus of the mature protein (designated number 1) was established by amino acid sequence analysis of purified IL-3 (*2*). (B) Diagram of the fragments of IL-3 that were synthesized. The regions corresponding to the five exons of the IL-3 gene are indicated by Roman numerals. The chemical syntheses were performed with a fully automated peptide synthesizer (Applied Biosystems model 430A). The synthesizer can be reprogrammed (*11*) and allows the user to define the synthetic chemistry for each amino acid. We used *t*-butyloxycarbonyl (tBoc)-N^{α}–protected amino acids that were coupled sequentially to tBoc-cysteine-OCH_2-phenylacetoamidomethyl-(polystyrene resin) (*14*) (Applied Biosystems). The loading of the starting resin was 1.0 mmol per gram of polystyrene. IL-3(1–79) was synthesized with the use of 4-methylbenzhydrylamine-polystyrene resin (*14*) (United States Biochemicals) with 0.45 mmol of primary amine per gram of polystyrene. After complete assembly of the protected peptide chain (*35*), the peptide-resin was reacted twice with 5 percent thiophenol in DMF for 20 minutes each time to remove the DNP group from the histidine. The N^{α}-tBoc group was removed with 65 percent trifluoroacetic acid (TFA), and the resin was neutralized, dried, and cleaved and deprotected by the "low-high" hydrogen fluoride method of Tam *et al.* (*20*). The cleaved deprotected peptide was precipitated and washed with ethyl acetate, then dissolved in 5 percent acetic acid and lyophilized. The crude peptide (50 mg) was treated with 2*M* 2-mercaptoethanol in 6*M* guanidine-HCl, 0.05*M* tris acetate, *p*H 8.0, at 37°C for 1 hour to remove residual DNP groups and exchanged into 0.1*M* acetic acid on Sephadex G-25. The peptide was then refolded by the glutathione–oxidized glutathione method (*21*) except that 2*M* guanidine-HCl was added to enhance solubility. After acidification, concentration, and dialysis at 4°C, the refolded peptide was lyophilized and used without further purification. The initial scale of synthesis was 0.5 mmol. Samples of peptide-resin were automatically removed after each cycle for documentation of the synthesis (see below). In addition, at several points, larger samples of peptide-resin were removed to provide intermediates for further studies. The final recovery was 2.2 g of peptide-resin. Deprotection and cleavage of a 500-mg sample gave 280 mg of crude product corresponding to 106 mg of protein, as determined by hydrolysis and amino acid analysis. The recovery after refolding of 50-mg samples of crude product was about 75 percent, indicating an overall total yield of crude refolded IL-3 of 470 mg (35 percent of theoretical).

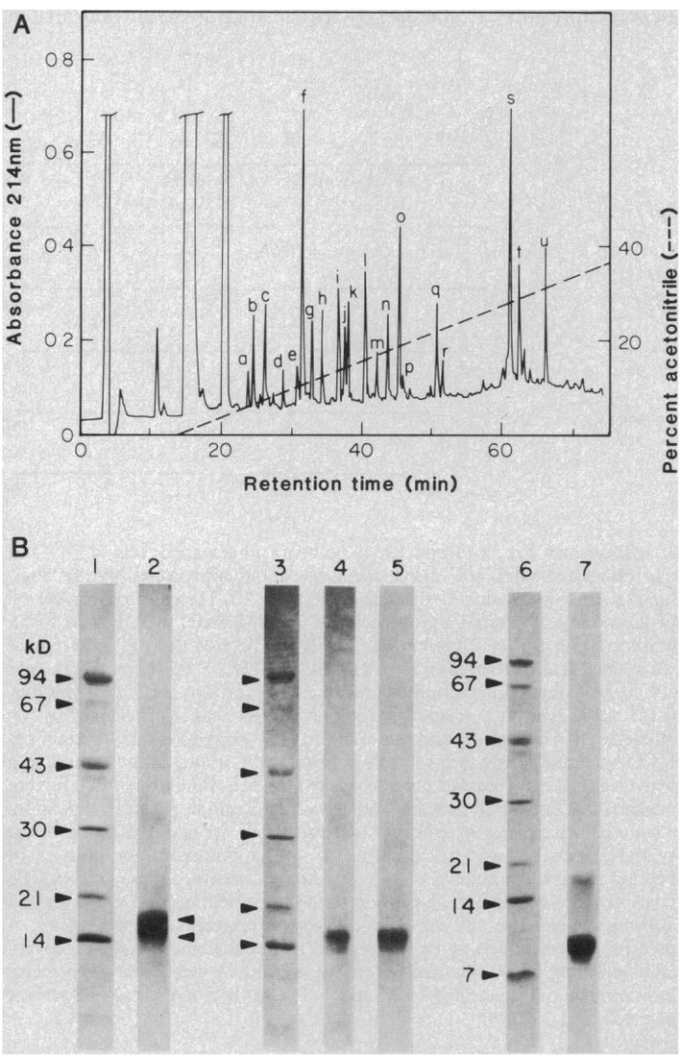
Table 1. Quantitative sequence analysis of synthetic IL-3.

Sampled at residue	Preview at residue	Cycles (No.)	Preview* (%)	Yield (%)	
				Correct chains†	Average per step‡
96	139	43	15.5	84.5	99.6
56	95	39	14.8	85.2	99.6
27	57	30	17.4	82.6	99.4
1	29	28	33.9	66.1	98.5
	Overall yield§			41.0	99.4

*Peptide-resins were sampled at the indicated (first column) position in the IL-3(1–40) synthesis (see Fig. 1) and subjected to automated Edman degradation on a gas-phase microsequencer (*39*). The derivatized amino acids were analyzed by HPLC (*40*). The premature appearance (preview) of the amino acids indicated in the second column was calculated as the ratio of the peak height of amino acid *N* at cycle *N*-1 to the sum of the peak height of amino acid *N* at cycle *N*-1 and the peak height of amino acid *N* at cycle *N*. *N* is the amino acid for which preview was being determined. The peak heights were corrected for background. †The yield of correctly assembled chain over the number of cycles indicated in the third column was calculated as (100 − preview). ‡The average yield per step was calculated as (percent yield/100)$^{1/n}$ × 100 where n is the number of cycles over which preview was determined (third column). §The yield of the complete 140 amino acid chain on the resin was calculated from the average yields per step as follows:

$$(99.6/100)^{45} \times (99.6/100)^{39} \times (99.4/100)^{27} \times (98.5/100)^{26}$$

The two couplings which were not included in the preview data, that is, Glu^{139} and Ile^{95} were assumed to have coupled with the same efficiency as the average from 96 to 138. Data from quantitative ninhydrin monitoring were consistent with this assumption.



Edman degradation of refolded synthetic IL-3 gave the expected NH_2-terminal sequence of amino acids. Moreover, there was no background of amino acids that would result from random internal cleavage of the peptide chain during synthesis or deprotection. Our results show that the crude product, while clearly not homogeneous, contains a high proportion of molecules with the primary structure of IL-3.

The refolded synthetic material had an apparent molecular mass (M_r) of 16,000 by sodium dodecyl sulfate gradient–polyacrylamide gel electrophoresis (SDS-PAGE) under reducing and nonreducing conditions (Fig. 2B) which was consistent with the M_r of 15,834 calculated from the IL-3 sequence (Fig. 1). Under nonreducing conditions the refolded synthetic IL-3 appeared as a doublet (Fig. 2B, lane 2, arrows). However, only the slower moving band is present when the chemically oxidized material was separated under similar conditions (Fig. 2B, lane 4). The faster moving band probably corresponds to more compact refolded forms of IL-3. The absence of significant amounts of high molecular weight material in the nonreducing gel of the refolded protein indicates that extensive polymerization did not occur during the chemical synthesis or refolding. In addition, the absence of low molecular weight material in the reduced gel of crude, unrefolded IL-3 demonstrates the absence of significant termination of the peptide chain during the synthesis (Fig. 2B, lane 5). The refolded fragment 1 to 79 was also examined by SDS-PAGE under nonreducing conditions (Fig. 2B, lane 7). The results indicate an apparent M_r of 9000, which is consistent with its theoretical M_r of 8948. As with the IL-3 (1–140), a major band with the expected M_r of the monomer was seen. The other analogues that were synthesized (see below) were also characterized by sequence analysis, amino acid analysis, tryptic peptide mapping, and SDS-PAGE with similar results. Evaluated together, these data indicate that the synthetic products contain a high proportion of molecules with the correct chemical structure and with the expected physicochemical properties.

Biological activities of synthetic IL-3. When added to IL-3–dependent cell lines, for example W/W.12 (a mast cell line) (Fig. 3A) or R6XE4, a bipotential mast cell–megakaryocyte line, the synthetic material stimulated cell growth in a concentration-dependent manner. The dose-response was characteristic of native IL-3

Fig. 2. Characterization of synthetic IL-3 and analogues. (A) Reverse phase high-performance liquid chromatography (HPLC) profile of a tryptic digest of synthetic IL-3. IL-3(1 mg) was dissolved in 150 μl of 8*M* urea and diluted to 1 ml in a sodium bicarbonate buffer, *p*H 8.0. Trypsin (20 μg) was added, and the mixture was incubated at 37°C for 2 hours. The digested sample was reduced with dithiothreitol and analyzed on a Vydac 5-μm C18 silica column equilibrated in 0.1 percent trifluoroacetic acid in water. The fragments were eluted with a linear gradient to 36 percent acetonitrile, 0.1 percent trifluoroacetic acid in water during a 60 minute period. The peaks were collected, hydrolyzed in 6*N* HCl at 110°C for 24 hours, and the amino acid composition was determined by reverse phase HPLC of the DABS-Cl (dimethylaminoazobenzenesulfoxyl chloride) derivatized amino acids (*36*). All the peaks collected were identified as fragments predicted from the target sequence (Fig. 1). They were : a, 1 to 6; b, 46 to 48; c, 136 to 140; d, 50 to 54; e, 49 to 54; f, 123 to 134; g, 35 to 43; h, 68 to 71; i, 55 to 67; j, 123 to 140; k, 97 to 103; l, 97 to 102; m, 28 to 34; n, 14 to 22; o, 55 to 71; p, 28 to 43; q, 23 to 34; r, 23 to 43; s, 77 to 96; t, 108 to 135; u, 107 to 122. (B) Sodium dodecyl sulfate gradient gel electrophoresis of unpurified, refolded, unreduced synthetic IL-3 (lane 2); unpurified, refolded material which had been treated with performic acid to oxidize the cysteines and thus block disulfide formation (lane 4); reduced, crude material which had not been refolded (lane 5) and unreduced, refolded IL-3(1–79) fragment (lane 7). The 12.5 to 20 percent polyacrylamide gradient gel was established with the use of the Laemmli buffer system (*37*). The molecular weight of standard proteins (lanes 1, 3, and 6) in kilodaltons is indicated. The standards were: 94 kD, phosphorylase B; 67 kD, bovine serum albumin; 43 kD, ovalbumin; 30 kD, carbonic anhydrase; 21 kD, soya bean trypsin inhibitor; 14 kD, lysozyme; and 7 kD, insulin.

with a maximal response followed by a logarithmic reduction in [^{3}H]thymidine incorporation with decreasing concentrations of the factor. The maximal thymidine incorporation was similar to that induced by WEHI-3B conditioned medium, a natural source of IL-3, tested as a positive control (Fig. 3A). The synthetic IL-3 was also tested for several other biological activities that have been attributed to native IL-3. Thus, the synthetic IL-3 also supported the growth of bone marrow cells including cells capable of giving rise to macroscopic colonies of erythroid and myeloid cells in the spleens of irradiated mice (Fig. 4A). In addition, it induced the Thy-1 antigen on cultured bone marrow cells (Fig. 4A). The synthetic material also stimulated the growth of colonies of cells containing granulocytes, macrophages, and, in about 30 percent of cases, megakaryocytes from single bone marrow progenitor cells (Figs. 4B and 5). The concentration of synthetic material required for a 50 percent of maximum response in these assays was higher than that required in the IL-3–dependent cell line assay, an observation consistent with previous findings with purified natural IL-3 (*25*). Synthetic IL-3 also promoted the growth of cultured lines of cells that structurally and functionally resembled in vivo mast cells (*26*). Taken together

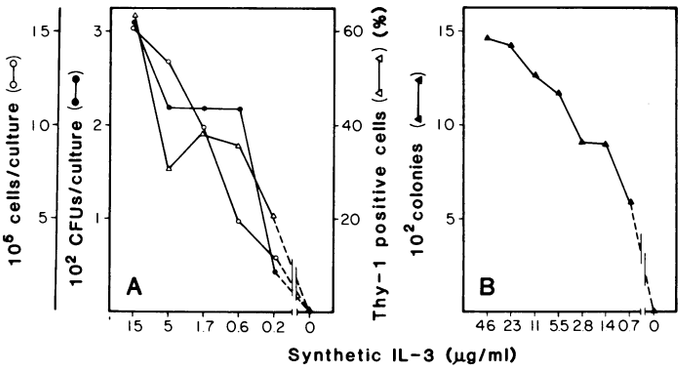


Fig. 4. Multiple biological activities of synthetic IL-3. (A) The effect of synthetic IL-3 (1 to 140) on stimulation of the growth of bone marrow cells in liquid culture (–○–), the maintenance of splenic colony-forming units (CFUs) in cultures of bone marrow cells (–●–), and the induction of the Thy 1 antigen on bone marrow cells (–△–). (B) Stimulation by the synthetic material of the growth of colonies from bone marrow cells cultured in agar. These assays were carried out as described (*38*). At the highest concentrations of IL-3(1–140), 30 percent of the colonies contained megakaryocytes as identified by the presence of acetylcholinesterase (*6*).

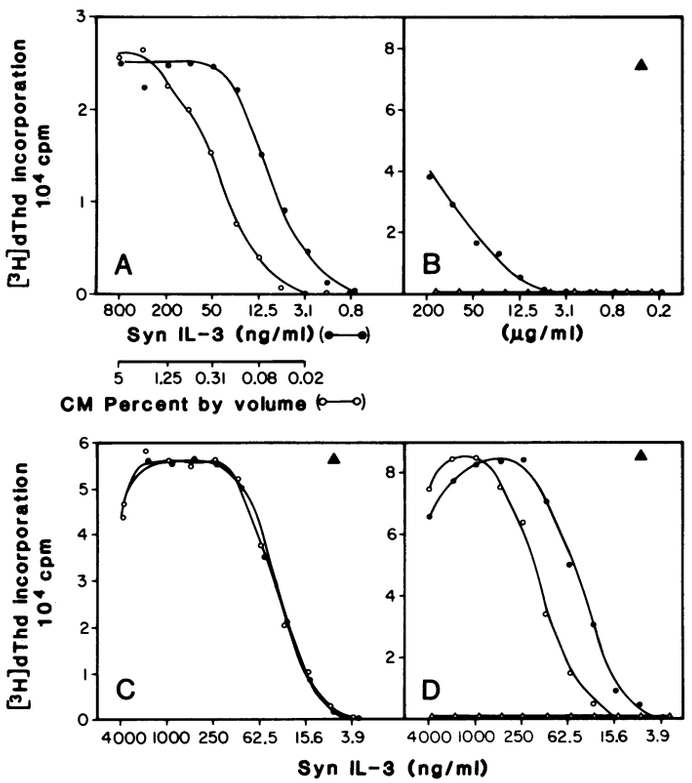


Fig. 3. Biological activity of synthetic IL-3 and various analogues. (A) Synthetic IL-3(1–140) (–●–) and a control preparation of tenfold concentrated WEHI-3 conditioned medium (CM) (–○–). (B) Synthetic IL-3(1–79) (–●–) and IL-3(80–140) (–△–). (C) Synthetic IL-3(7–140) (–○–) and IL-3(1–140) (–●–). (D) Synthetic IL-3(18–140) (–△–), IL-3(17–140) (–○–), and IL-3(1–140) (–●–). The results shown in panels A, C, and D are from separate IL-3 syntheses. The fragments shown in panels C and D are intermediates in the IL-3(1–140) synthesis indicated in that panel. The closed triangles in panels B, C, and D indicate the maximal activity obtained with WEHI-3 conditioned medium, a source of IL-3. The lyophilized refolded synthetic peptides were dissolved in 2*M* guanidine-HCl and desalted on a Sephadex G-25 column equilibrated in phosphate-buffered saline. The weights were calculated from the molarity of the final preparation as determined by acid hydrolysis of the protein and amino acid analysis. The IL-3 activity was assayed with the IL-3–dependent cell line W/W.12 as described (*2*), except that 2×10^3 cells per 10 μl of culture were used and the cells were treated with [^{3}H]thymidine after 40 hours.

these data show that this synthetic product has the multiple activities attributed to IL-3 (*2, 3, 6, 27*).

Crude refolded synthetic IL-3 had a specific activity of 9 ng/ml (50 percent maximal response in the assay measuring the growth of IL-3–dependent cell lines) (Fig. 3A). The specific activity of this material could be increased by further purification as indicated by the reverse phase HPLC gradient analysis shown in Fig. 6. A single sharp peak of IL-3 activity was obtained at an acetonitrile concentration of 35 percent. However, the profile of ultraviolet absorbance was broader with material having low activity eluting later than the peak of IL-3 activity. Previous experience suggested that some of this late-eluting material results from shortcomings of the cleavage and side chain deprotection steps (*20*) and probably in part represents incompletely deprotected or chemically damaged products. The specific activity of the fraction containing the most activity (fraction 39 in Fig. 6) was 700 pg/ml, as determined by the IL-3 assay and amino acid analysis. Highly purified preparations of native IL-3 have reported specific activities ranging from 4 pg/ml (*2*) to 200 pg/ml (*3*) in the same assay, reflecting difficulties in the estimation of the specific activity of purified natural factor, as discussed previously (*2*). Thus, the synthetic IL-3 may be 0.5 percent to 30 percent as active as the native IL-3. Additional

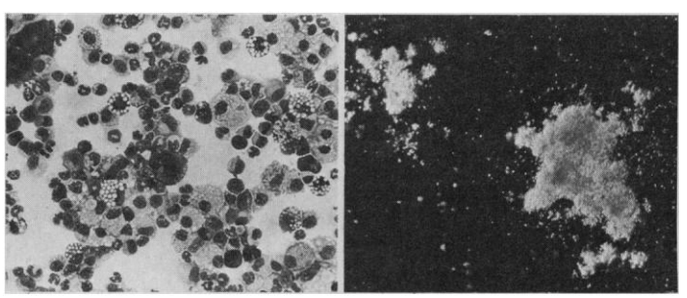


Fig. 5. Photomicrographs of (left) a cytocentrifuged preparation of cells from liquid cultures of bone marrow cells cultured with IL-3(1–140) (15 μg/ml) from the experiment in Fig. 4A, stained with May-Grunwald Giemsa and showing immature myeloid cells, macrophages, granulocytes, and megakaryocytes; (right) macroscopic colonies derived from single bone marrow cells cultured in agar with IL-3 (15 μg/ml) from the experiment in Fig. 4B.

purification could probably be achieved to eliminate the remaining difference. This is supported by preliminary studies that indicate the presence of improperly refolded molecules in the partially purified material. Other possible explanations for these differences in specific activities include increased susceptibility to degradation of the synthetic IL-3 in the bioassay either due to incorrect refolding or the absence of carbohydrate side chains.

Functional analysis of IL-3 analogues. The chemical synthesis of IL-3 allowed us to undertake a series of structure-function studies. The high activity of the crude synthetic product together with the sensitivity of the IL-3 assay made it possible to detect activity in analogues with 10^5 times less potency than the synthetic material corresponding to the predicted amino acid sequence of IL-3.

In one series of experiments we synthesized a number of relatively short peptides from the IL-3 sequence to attempt to identify an active fragment. The fragments and analogues of IL-3 that were synthesized are shown diagrammatically in Fig. 1B. These included the following regions initially chosen for preparing antipeptide antibodies: 1 to 29, 1 to 17, 64 to 82, 91 to 112, and 123 to 140. Because for some proteins functional domains are encoded by gene exons, we also synthesized the sequences corresponding to the five exons of the IL-3 gene (*28*), that is: 1 to 29, 30 to 43, 44 to 76, 77 to 89, and 90 to 140. None of these nine fragments showed detectable activity nor did they inhibit IL-3 activity when added in larger molar excess to IL-3–dependent cells.

We then divided the IL-3 sequence into two parts, 1 to 79 and 80 to 140, each containing two cysteine residues, to test the possibility that pairing of Cys^{17} with Cys^{79} and Cys^{80} with Cys^{140} may form two domains in the native molecule, one of which is responsible for the functional activity. As shown in Fig. 3B, IL-3(80–140) had no detectable activity, although the same material extended to the full length molecule by addition of the 79 NH_2-terminal amino acids had the expected high activity (see Fig. 3A). In contrast, IL-3(1–79) stimulated the growth of IL-3–dependent cells, although the concentration required for half maximal activity was at least 10^4 times higher than that required for the the complete molecule (compare Fig. 3A with Fig. 3B). This result suggests that a 1 to 79 fragment of IL-3 contains sufficient structural information to stimulate growth of IL-3–dependent cells. The low activity compared to the full length material indicates that stimulation by the fragment is inefficient, but the reason for this remains to be determined.

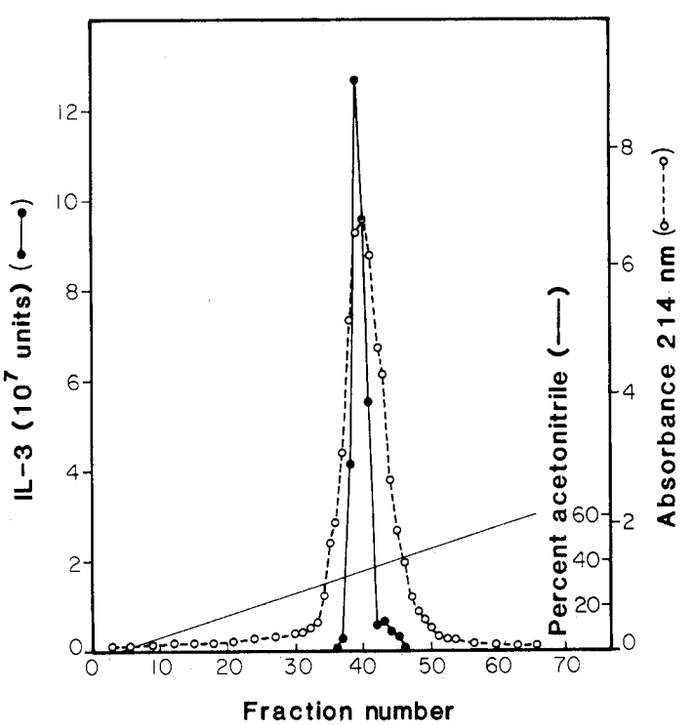


Fig. 6. Reverse phase HPLC of synthetic IL-3. The refolded synthetic material (5 mg) was dissolved in 150 μl of 5 percent acetic acid and loaded on to a Vydac C4 5-μm column equilibrated in 0.1 percent TFA in water. A gradient to 60 percent acetonitrile and 0.1 percent TFA was applied over 60 minutes at a flow rate of 0.7 ml/min. Fractions were collected at 1-minute intervals and samples were diluted 1000-fold in culture medium and titrated in the IL-3 assay. Units of IL-3 were calculated as described (*2*). Samples of each fraction were diluted in water to determine the absorbance at 214 mm.

Following our results with the 1 to 79 fragment, we examined more closely the importance of the amino-terminal region of the molecule for biological activity. One initial question related to the functional significance of the two forms of IL-3 isolated from WEHI-3 cells which differed in their NH_2-terminal amino acid sequence: the IL-3(1–140) form isolated by Clark-Lewis *et al.* (*2*) as P-cell–stimulating factor and the shorter, 7 to 140, form isolated as IL-3 by Ihle *et al.* (*3*). Comparison of synthetic peptides corresponding to the 1 to 140 and 7 to 140 forms indicated that there was no significant difference in the IL-3 activity of the two forms in stimulating the growth of mast cells (Fig. 3C). Furthermore, both molecules showed the full range of biological activities (compare with Fig. 4). Clearly, IL-3 activity is not dependent on the NH_2-terminal seven residues. However, the first 29 amino acids did appear to incorporate structures important for biological activity as preliminary experiments showed that a synthetic peptide corresponding to residues 30 to 140 was inactive. This peptide is missing the cysteine at position 17, which could potentially be important in stabilizing the tertiary structure of IL-3. To test the importance of this cysteine we synthesized, in a single experiment, peptides corresponding to 18 to 140, 17 to 140, and the entire 1 to 140 sequence by sampling the peptide-resin at the appropriate steps. After cleavage and deprotection, all three products were refolded and assayed. The 18 to 140 peptide lacked detectable activity up to a concentration of 4 μg/ml (Fig. 3D). In fact, further assays indicated that peptide 18 to 140 was inactive up to a concentration of 200 μg/ml, 10^5 times higher than the concentration at which synthetic IL-3(1–140) had detectable activity. In contrast the 17 to 140 peptide had high activity, about one-third that of the 1 to 140 peptide, which served as a positive control. These results show that the cysteine at residue 17 is essential for IL-3 activity and suggest that it may be involved in a critical disulfide bond. Experiments with analogues where pairs of cysteines have been replaced by alanine residues support this conclusion. In addition, the results show that residues 7 to 16 are required for maximal biological activity.

Potential for structure-function studies. Our results with chemically synthesized IL-3 and several analogues allow us to draw some important conclusions regarding the relation between IL-3 structure and function. We have directly demonstrated that the amino acid sequence predicted from the cDNA corresponding to IL-3 does indeed specify a molecule with IL-3 activity. Because the source of the synthetic IL-3 is nonbiological, it rules out the possibility that one or more different molecules, or differently processed products of the same gene, are responsible for the multiple biological activities attributed to IL-3. We have also shown directly that, as suggested earlier (*29*), glycosylation (or indeed other possible post-translational modification) is not essential for IL-3 function. Experiments with fragments have shown that while the complete IL-3 molecule is required for maximal activity, an amino terminal fragment (1 to 79) was sufficient to stimulate growth of IL-3–dependent cell lines. Furthermore, we have shown that a disulfide bond incorporating the cysteine at position 17 is

essential for IL-3 activity, that the difference of six amino acid residues between the two isolated forms of IL-3 (*2, 3*) does not affect the biological activities, and that residues 7 to 16 are required for maximal biological activity. These results form a basis for further studies aimed at determining the structural requirements for the functional activity of this lymphokine.

Our data were obtained with a new technology that incorporates recent improvements in solid-phase chemistry and full automation of the synthetic process (*11*). There have been a number of successful total chemical syntheses and structure-function studies of peptides up to about 50 amino acids in length, such as insulin (*30*), transforming growth factor alpha (*31*), and epidermal growth factor (*32*). Relatively few larger peptides or proteins have been synthesized with high biological activity. While these chemical syntheses of longer molecules were major achievements (*33*), the methods used did not lend themselves readily to structure-function or other studies, and, in particular, they were highly specialized and time-consuming. In contrast, the methods we have used in the synthesis of IL-3 are rapid, fully automated, and readily available. This approach has some clear advantages for the study of protein structure and function relationships. (i) The chemistry is applicable to a wide range of peptides and is independent of peptide chain length (*34*). (ii) Relatively large amounts of a peptide or protein of known composition can be produced in a matter of days in yields and purity that allow direct functional analysis. (iii) Appropriate sampling of synthetic intermediates allows multiple variants of a given sequence to be made in a relatively short period of time. (iv) Because the synthetic process is fully automated, reproducibility can be achieved in multiple syntheses, thus allowing meaningful comparisons between different analogues. (v) Variants containing non-naturally occurring amino acids can be synthesized.

The successful synthesis of IL-3 suggests that automated peptide synthesis could be applied to the study of other large peptides and proteins. For example, the approach described here may be extended to other lymphokines and hemopoietic regulatory factors, the protein moieties of which are about the same size as IL-3 (*6*). In addition, functional domains of larger proteins are typically 100 to 200 amino acids in length, and it may be possible to study these domains in vitro, provided that suitable functional assays are available. Syntheses of proteins by the methods described here, in combination with additional purification techniques, may allow further characterization of the structural basis of biological activity by, for example, analysis of the three-dimensional structure of crystallized synthetic products. The successful application of this approach will lead to new insights into protein structure and biological functions.

REFERENCES AND NOTES

1. M. C. Fung *et al.*, *Nature (London)* **307**, 233 (1984); T. Yokota *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 1070 (1984).
2. I. Clark-Lewis, S. B. H. Kent, J. W. Schrader, *J. Biol. Chem.* **259**, 7488 (1984).
3. J. N. Ihle *et al.*, *J. Immunol.* **131**, 282 (1983).
4. Y. P. Yung, R. Eger, G. Tertian, M. A. S. Moore, *J. Immunol.* **127**, 794 (1981); N. N. Iscove, C. A. Roitsch, N. Williams, L. J. Guilbert, *J. Cell Physiol.* **1** (suppl.), 65 (1982); G. W. Bazill, M. Haynes, J. Garland, T. M. Dexter, *Biochem J.* **210**, 747 (1983); R. L. Cutler, D. Metcalf, N. A. Nicola, G. R. Johnson, *J. Biol. Chem.* **260**, 6579 (1985); J. W. Schrader, S. J. Lewis, I. Clark-Lewis, J. G. Culvenor, *Proc. Natl. Acad. Sci. U.S.A.* **78**, 323 (1981).
5. J. N. Ihle, L. Pepersack, L. Rebar, *J. Immunol.* **126**, 2184 (1981).
6. J. W. Schrader, *Crit. Rev. Immunol.* **4**, 197 (1983); ——— and R. M. Crapper, *Proc. Natl. Acad. Sci. U.S.A.* **80**, 6892 (1983).
7. A. N. Glazer, in *The Proteins*, H. Neurath and R. L. Hill, Eds. (Academic Press, New York, 1976), p. 1.
8. M. J. Zoller and M. Smith, *Methods Enzymol.* **100**, 468 (1983).
9. M. Wieczorek and M. Laskowski, Jr., *Biochemistry* **22**, 2630 (1983).
10. R. B. Merrifield, *J. Am. Chem. Soc.* **85**, 2149 (1963).
11. S. B. H. Kent, L. E. Hood, H. Beilan, S. Meister, T. Geiser, in *Peptides 1984*, U. Ragnarsson, Ed. (Almqvist and Wiksell, Stockholm, Sweden, 1984), p. 185; S. B. H. Kent *et al.*, *Peptide Chemistry* **84**, N. Izumiya, Ed. (Protein Research Foundation, B. H. Osaka, Japan, 1985), p. 217.
12. S. B. H. Kent, in *Proceedings of the Ninth American Peptide Symposium*, K. Kopple and C. Deber, Eds. (Pierce Chemical Co., Rockford, IL, in press); S. B. H. Kent, in *Biomedical Polymers*, E. P. Goldberg and A. Nakajima, Eds. (Academic Press, New York, 1980), p. 213.
13. A. R. Mitchell, B. W. Erickson, M. N. Ryabstev, R. S. Hodges, R. B. Merrifield, *J. Am. Chem. Soc.* **98**, 7357 (1976); S. B. H. Kent, A. R. Mitchell, M. Engelhard, R. B. Merrifield, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 2180 (1979); S. B. H. Kent, in *Proceedings of the Eighth American Peptide Symposium*, V. J. Hruby and D. H. Rich, Eds. (Pierce Chemical Co., Rockford, IL, 1983), p. 99.
14. A. R. Mitchell, S. B. H. Kent, M. Engelhard, R. B. Merrifield, *J. Org. Chem.* **43**, 2845 (1978); S. A. Gaehde and G. A. Matsueda, *Int. J. Pept. Protein Res.* **18**, 551 (1981).
15. S. M. Meister and S. B. H. Kent, in *Proceedings of the Eighth American Peptide Symposium*, V. J. Hruby and D. H. Rich, Eds. (Pierce Chemical Co., Rockford, IL, 1983), p. 103.
16. H. Hagenmaier and H. Frank, *Hoppe-Seylers Z. Physiol. Chem.* **353**, 1973 (1972).
17. L. Corley, D. H. Sachs, C. B. Anfinsen, *Biochem. Biophys. Res. Commun.* **47**, 1353 (1972).
18. G. Barany and R. B. Merrifield, in *The Peptides*, E. Gross and J. Meienhofer, Eds. (Academic Press, New York, 1980), vol. 2, p. 1.
19. S. B. H. Kent and R. B. Merrifield, *Int. J. Pept. Protein Res.* **22**, 57 (1983).
20. J. P. Tam, W. F. Heath, R. B. Merrifield, *J. Am. Chem. Soc.* **105**, 6442 (1983).
21. A. K. Ahmed, S. W. Schaffer, D. B. Wetlaufer, *J. Biol. Chem.* **250**, 8477 (1975).
22. G. L. Ellman, *Arch. Biochem. Biophys.* **82**, 70 (1959).
23. V. K. Sarin, S. B. H. Kent, J. P. Tam, R. B. Merrifield, *Anal. Biochem.* **117**, 147 (1981).
24. S. B. H. Kent, M. Riemen, M. LeDoux, R. B. Merrifield, in *Methods in Protein Sequence Analysis*, M. Elzinga, Ed. (Humana Press, Clifton, NJ, 1982), p. 205.
25. I. Clark-Lewis, W. R. Thomas, J. W. Schrader, *Exp. Hematol.* **13**, 304 (1985); I. Clark-Lewis *et al.*, unpublished observations.
26. R. M. Crapper, W. R. Thomas, J. W. Schrader, *J. Immunol.* **133**, 2174 (1984).
27. D. M. Rennick *et al.*, *ibid.* **134**, 910 (1985); J. W. Schrader, I. Clark-Lewis, R. M. Crapper, G. H. W. Wong, *Lymphokine Res.* **2**, 83 (1983).
28. S. Miyatake, T. Yokota, F. Lee, K. Arai, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 316 (1985).
29. I. Clark-Lewis and J. W. Schrader, *J. Immunol.* **127**, 1941 (1981).
30. P. Sieber *et al.*, *Helv. Chim. Acta* **57**, 2617 (1974).
31. J. P. Tam, H. Marquardt, D. F. Rosberger, T. W. Wong, G. J. Todaro, *Nature (London)* **309**, 376 (1983).
32. D. Hagiwara, M. Neya, Y. Miyazaki, K. Hemmi, M. Hashimoto, in *Peptide Chemistry 84*, N. Izumiya, Ed. (Protein Research Foundation, Osaka, Japan, 1985), p. 27.
33. B. Gutte and R. B. Merrifield, *J. Biol. Chem.* **246**, 1922 (1971); H. Yajima and N. Fujii, in *Frontiers in Protein Chemistry*, T. Y. Liu, G. Mamiya, K. T. Yasunbo, Eds. (Protein Research Foundation, Osaka, Japan, 1980), p. 159.
34. S. B. H. Kent and I. Clark-Lewis, in *Synthetic Peptides in Biology and Medicine*, K. Alitalo, P. Partanen, A. Vaheri, Eds. (Elsevier, Amsterdam, in press).
35. Side chain protecting groups used were Asp(OBzl), Glu(OBzl), Ser(Bzl), Thr(Bzl), Lys(Cl-Z), Tyr(Br-Z), Arg(N^GTos), Cys(4-MeBzl), and His(ImDNP). (Bzl, benzyl; Tos, toluene sulfoxyl; DNP, dinitrophenyl; Im, imidazole; Z, benzyloxgycarbonyl. The remaining amino acids had no side chain protecting groups. All the amino acids were obtained from Peninsula Laboratories, except the tBoc-His(ImDNP), which was from Chemical Dynamics and was crystallized from ethanol before use. For each cycle the tBoc N^{α}-protected peptide-resin was exposed to 65 percent trifluoroacetic acid (from Eastman-Kodak) (distilled before use) in dichloromethane (DCM), (Mallinckrodt): first for 1 minute and then for 13 minutes to remove the N^{α}-protecting group. The peptide-resin was washed in DCM, neutralized twice with 10 percent diisopropylethylamine (DIEA) (Aldrich) in dimethylformamide (DMF) (Applied Biosystems), for 1 minute each. Neutralization was followed by washing with DMF. Coupling was performed with the preformed symmetric anhydride of the amino acid in DMF for 16 minutes. The preformed symmetric anhydride was prepared on the synthesizer by dissolving 2 mmol of amino acid in 6 ml of DCM and adding 1 mmol of dicyclohexylcarbodiimide (Aldrich) in 2 ml of DCM. After 5 minutes, the activated amino acid was transferred to a separate vessel and the DCM was evaporated by purging with a continuous stream of nitrogen gas. The DCM was replaced by DMF (6 ml, total) at various stages during the purging. After the first coupling, the peptide-resin was washed with DCM, 10 percent DIEA in DCM, and then with DCM. For recoupling, the same amino acid and the activating agent, dicyclohexylcarbodiimide, were transferred sequentially to the reaction vessel. After activation in situ and coupling for 10 minutes, sufficient DMF was added to make a 50 percent DMF-DCM mixture, and the coupling was continued for 15 minutes. Arginine was coupled as a preformed hydroxybenzotriazole (Aldrich) ester in DMF for 60 minutes and then recoupled in the same manner as the other amino acids. Asparagine and glutamine were coupled twice as preformed hydroxybenzotriazole esters in DMF, 40 minutes for each coupling. For all residues, the resin was washed after the second coupling and a sample was automatically taken for monitoring residual uncoupled α-amine by quantitative ninhydrin reaction (*23*).
36. J. Y. Chang, R. Knecht, D. G. Braun, *Biochem. J.* **199**, 547 (1981).
37. U. K. Laemmli, *Nature (London)* **227**, 680 (1970).
38. J. W. Schrader and I. Clark-Lewis, *J. Immunol.* **129**, 30 (1982); J. W. Schrader, F. Battye, R. Scollay, *Proc. Natl. Acad. Sci. U.S.A.* **79**, 4161 (1982).
39. R. M. Hewick, M. W. Hunkapiller, L. E. Hood, W. J. Dreyer, *J. Biol. Chem.* **256**, 7990 (1981).
40. M. W. Hunkapillar and L. E. Hood, *Methods Enzymol.* **91**, 486 (1983).

41. Supported by research grants from the National Institutes of Health, the Monsanto Company, and Upjohn Pharmaceuticals. Additional support was obtained from the National Health and Medical Research Council, Canberra, Australia; NIH grant R01-CA38684-01 (J.W.S.); a C. J. Martin Traveling Fellowship from the National Health and Medical Research Council, Canberra, Australia (I.C.L.), an EMBO long-term fellowship (R.A.) and fellowships from the Swiss National Foundation (R.A. and H.Z.). We thank C. Davis and J. Kim for expert technical assistance.

7 October 1985; accepted 2 December 1985