# Construction of interleukin-1 $\alpha$  mutants using unequal contamination of synthetic oligonucleotides

Kurt Poindexter, Rita Jerzy and Richard B.Gayle III\* Department of Molecular Biology, Immunex Corporation, Seattle, WA 98101, USA

Received October 12, 1990; Revised and Accepted March 10, 1991

EMBL accession no. X55445

# ABSTRACT

Proteins without readily available three-dimensional structural data present a difficult problem in the exploration of structure/function relationships. Saturation mutagenesis using contaminated oligonucleotides can identify potentially interesting regions of such a protein. This technique, in which synthesized oligonucleotides contain low-level base substitutions, allows random mutations to be placed throughout a gene sequence. Using double-stranded cassettes, a region of the human interleukin-1 $\alpha$  gene has been altered using such mutagenic oligonucleotides. However, instead of contaminating both strands of the gene sequence at the same level, each strand of the insert was contaminated at a different level. Several recombinants were sequenced and the effects of the mutations on the activity of the proteins were examined. Contaminating the two oligonucleotides at different levels produced a significantly different distribution of nucleotide changes from that seen if both strands were contaminated at the same level. The observed distribution followed the average of the distributions for each of the two contamination levels. This resulted in roughly equal frequencies of <sup>1</sup> to 5 nucleotide changes per clone with very few clones containing the wild-type nucleotide sequence. This helped overcome the redundancy in the genetic code, resulting in a high frequency of amino acid changes, and allowed changes at every amino acid to be sampled in a small number of mutants. This procedure can allow a gene sequence to be screened rapidly by removing most wild-type sequences from analysis while making sure that there are many amino acid changes in the resultant mutants.

# INTRODUCTION

Mutagenesis allows the relationship between the structure of a protein and its function to be explored much more rapidly than is otherwise possible. This relationship has been actively explored in proteins for which there are X-ray crystallographic data regarding 3-dimensional structure (1). Site-directed mutagenesis has been used to change residues which appear important. This technique, combined with computer modeling, has greatly expanded the knowledge base of protein structure (2,3).

A potential drawback to most mutagenic approaches arises from the need to choose a particular amino acid residue to change. For a large protein the possibilities are daunting. In order to realistically examine any of the large number of proteins for which there is little structural information, some sort of random mutagenesis should be used. This obviates the need to make a rational hypothesis regarding the choice of important residues. Chemical mutagenic techniques have been employed in order to generate such random mutations (1,4). These suffer, however, from the inherent specificity of the chemical reactions used in the mutagenesis and the subsequent need to generate several libraries of mutants using different chemicals in order to saturate the coding region with changes. A more precise method of random mutagenesis, saturation mutagenesis, uses contaminated oligonucleotides.

Saturation mutagenesis has been used successfully in several investigations (5,6,7). Briefly, this approach involves synthesizing deoxyoligonucleotides in which each of the four phosphoramidites has been contaminated with a small amount of the other three. The absolute frequency of a substitution is a consequence of the level of contamination. The optimal mutation frequency is usually determined empirically (6). Many applications have examined non-coding DNA sequences, with contamination levels chosen which largely produce single base changes. However, this becomes problematic when studying a protein. Because the mutation frequency follows a Poisson distribution, a contamination level that results in a large percentage of single nucleotide changes will also produce a large number of unchanged (i.e. wild-type) nucleotide sequences. In addition, since most codons will only be changed at one base in the triplet, many of the altered nucleotide sequences will contain silent codon changes, adding to the number of recombinants producing a wild-type protein. A large background of wild-type proteins means that a large number of mutants must be examined in order to identify changes at each amino acid. To overcome these problems, higher levels of contamination can be used. This lowers the amount of unaltered sequences present, increases the chance that two nucleotides will be changed in a single codon, and lowers the number of mutants that must be sampled but the larger number

<sup>\*</sup> To whom correspondence should be addressed

of mutations present in each clone will increase the probability that the mutated protein will be unstable. It is hard to empirically determine what the optimal contamination level should be which will result in low numbers of unaltered proteins but will not produce too many sequences which are unstable. The choice of any one contamination level may result in too few or too many amino acid changes.

This paper describes an approach which helps to surmount these problems. Using human interleukin- $1\alpha$  (huIL- $1\alpha$ ) as a model protein, oligonucleotides corresponding to the two complementary strands of a region of the gene were synthesized using different levels of contamination for each strand. The two oligonucleotides were then annealed and ligated as a cassette into a vector containing the huIL-1 $\alpha$  gene. Use of two different levels of contamination resulted in a distribution of mutations which is the additive average of the Poisson distribution for each individual strand. As a consequence, the expected number of recombinants with unchanged nucleotide sequences was quite low, while the frequencies of those with  $1-5$  nucleotide changes were approximately equivalent. This means that in a single construction a range of mutants could be generated with roughly equal chances that there would be single or multiple nucleotide changes. Changes at all possible amino acids could be sampled in a relatively small number of mutants. The effect of several of these mutations on the biological activity and on the ability of the protein to bind IL-1 receptor was examined.

# MATERIALS AND METHODS

#### Enzymes and vectors

All restriction enzymes, DNA ligase and T4 polynucleotide kinase were obtained from Boehringer Mannheim or New England Biolabs. BBG1, a plasmid bearing a synthetic gene for huIL- $1\alpha$ was purchased from British Biotechnology. pPL3 has been previously described (8). pPLIL-1 $\alpha$ , a pPL3 derivative containing the wild-type IL-1 $\alpha$  gene, was provided by Doug Cerretti. Plasmid DNA was purified by the alkaline lysis method (9). Restriction fragments of plasmid DNA were electrophoresed on a 0.8% low-melt agarose gel (AMC Products), eluted and purified using Geneclean (Bio101).

#### DNA synthesis

Two oligonucleotides (61 bases), corresponding to the huIL-1 $\alpha$ gene sequence between the EcoRI and PstI sites of pPLIL-1 $\alpha$ (Fig. 1B), were synthesized on an Applied Biosystems Model 380A DNA synthesizer, using ABI reagents. The phosphoramidites were contaminated at two different levels, 4.2%  $(1.4\%$  each of the other 3 phosphoramidites) for the sense strand and 8.4% (2.8% each of the other 3 phosphoramidites) for the antisense strand. All contaminations were performed under nitrogen in order to reduce inactivation of the phosphoramidites. After entering the wild-type sequence, a standard cycle was used for a <sup>1</sup> micromole synthesis.

## Oligonucleotide Purification

The oligonucleotides were purified on <sup>a</sup> 40 cm 8% polyacrylamide-7 M urea gel. Following visualization by ultraviolet light shadowing, bands of the correct size were removed and eluted from the gel slice. Care was taken to excise full length oligonucleotides so that oligonucleotides with base deletions were minimized. The oligonucleotides were deprotected and resuspended in TE (10 mM Tris-Cl, pH 7.4, <sup>1</sup> mM EDTA). DNA concentration was determined using <sup>a</sup> LKB UltraspecdI.

#### Assembly and cloning of mutagenic oligonucleotides

Twenty picomoles of each oligonucleotide were mixed in 20  $\mu$ l TE and placed at 65°C for <sup>15</sup> minutes. The mixture was allowed to cool slowly to room temperature and then placed on ice. The oligonucleotides have EcoRI and PstI overhanging ends. The oligonucleotides were ligated into the appropriately cleaved vector using different vector:oligonucleotide molar ratios ranging from 1:1 to 1:5.

## Construction of pPLBBGIL-1 $\alpha$  and pPLRPI

The mature wild-type huIL- $1\alpha$  gene was cloned into the vector pPL3, which contains the  $\lambda P_L$  promoter-operator, a multiple cloning site and downstream transcriptional terminators (8). The resulting plasmid, pPLIL-1 $\alpha$ , can directly express the IL-1 $\alpha$  gene when placed in GMl (10) carrying the compatible plasmid, pRK248 (11). This plasmid carries a temperature-sensitive cI repressor which blocks expression from the PL promoter. Shifting the growth temperature to 42°C allows transcription of the IL-1 $\alpha$  gene to take place, resulting in protein production. pPLIL-1 $\alpha$  was cleaved with EcoRI and HinDIII, followed by ligation with an *EcoRI-HinDIII* fragment from BBG1 containing the synthetic gene for huIL-1 $\alpha$  (Fig. 1A). This plasmid,  $pPLBBGIL-1\alpha$ , allows the easy introduction of mutagenic cassettes into the huIL-1 $\alpha$  gene. In order to facilitate the identification of recombinant clones, a vector was constructed which had an irrelevant fragment of DNA inserted between the EcoRI (R) and PstI (P) sites. This plasmid, pPLRPI, has no DNA sequences corresponding to pPLBBGIL-1 $\alpha$  between the EcoRI and PstI sites, but does carry an ApaI site between them.

## Colony hybridizations

Following cloning and transformation of GMl [pRK248] with the recombinant plasmids, colonies were lifted from the plates onto nitrocellulose. The filters were lysed with 0.5 M NaOH for <sup>10</sup> minutes, washed with 1.5 M NaCl, 0.5 M Tris, pH 7.5, for <sup>10</sup> minutes. This was followed by a 10 minute incubation in  $2 \times SSC$ . The filters were baked at 80°C for 2 hours and were placed in a sealed plastic bag with  $2-5$  mls of oligonucleotide prehybridization buffer (5 <sup>x</sup> SSC, 0.5% SDS, <sup>1</sup> mM EDTA, pH 8.0) at 55°C for <sup>1</sup> hour. 32P-labeled probe (sense strand oligonucleotide) was then added and the mix was incubated at  $55^{\circ}$ C for 2 hours followed by a 30 minute wash in  $2 \times$ SSC at 55°C. The filter was removed and placed on film for overnight autoradiography.

#### Sequencing

Double stranded sequencing of the vectors with inserts was performed using the dideoxy method (12).

#### Protein expression and analysis

Mutant proteins can be produced by growing GMl[pRK248] cells containing recombinant plasmids at  $30^{\circ}$ C to an OD<sub>600</sub> = 1, then placing them at 42°C in order to inactivate the temperaturesensitive  $\lambda$  repressor (11). After 3 hours, the cells were pelleted and lysed with sample buffer (125 mM Tris, pH 7.5, 2% SDS). Each lysate was diluted 1:100 in PBS (137 mM NaCl, 2.7 mM KCl, 4.3 mM  $Na<sub>2</sub>HPO<sub>4</sub>·7H<sub>2</sub>O$ , 1.4 mM  $KH<sub>2</sub>HPO<sub>4</sub>$ , pH 7.4) and radioimmunoassays (13) were performed using a rabbit

polyclonal antibody to huIL-1 $\alpha$  in order to determine recombinant protein concentration. Mutant IL-1 $\alpha$  binding activity was determined by the capacity of E. coli cell lysates to inhibit the binding of radiolabeled IL-1 $\alpha$  to EL4 6.1 cells (14). Biological activity was examined using an ELA conversion assay (15). The data were then analyzed by a nonlinear least squares fitting program using values determined from a standard curve derived from purified IL-1 $\alpha$  tested in each biological assay and each binding assay. Wild-type IL-1 $\alpha$  produced in E. coli was used as an internal control for all the assays.



# B



481 TAATAAgatctaagctt 497

Figure 1. (A) Strategy for the construction of mutants in huIL-1 $\alpha$ . Using a naturally occurring EcoRI site in pPLIL-1 $\alpha$ , a synthetic gene fragment from BBG1 was inserted. The resulting plasmid, pPLBBGIL-1 $\alpha$ , then had an irrelevant fragment of DNA inserted between the EcoM and PstI sites in order to facilitate cloning of the mutagenic oligonucleotides. This plasmid, pPLRPI, was used in the actual construction of mutants.  $P_L - \lambda P_L$  promoter; A-ApaI; R-EcoRI; H-HinDIII;  $P-PstI.$  (B) Sequence for coding region of pPLBBGIL-1 $\alpha$  (Accession # X55445). The nucleotides used for mutagenesis are boxed. Some of the restriction enzyme sites present in the sequence are also shown.

# **RESULTS**

#### Isolation of a mutagenic library

The intermediate plasmid, pPLRPI, contained a unique ApaI site (Fig. 1A) which was not present in any of the mutagenic cassettes. The mutagenic cassettes were ligated into pPLRPI which had been cleaved with EcoRI and PstI. Digestion with ApaI after ligation should linearize any pPLRPI which simply recircularized. Isolation of the collection of plasmids bearing mutated huIL- $1\alpha$ genes was simplified further by performing colony hybridizations. This procedure allowed identification of vectors which contained a mutagenic insert. Although the best probe for these hybridizations was one of the mutagenic oligonucleotides, there was some concern regarding the sensitivity of the procedure. As a test, colony hybridizations were done using pPLRPI and plasmids bearing  $0-5$  mutations, as determined by sequencing. The filters were probed with the sense strand oligonucleotide which had been end-labeled with <sup>32</sup>P. As can be seen in figure 2, the hybridizations were very specific, with no signal being seen with pPLRPI and sufficient signal present even with 5 base changes. Over 700 colonies were isolated by using this approach with no apparent selectivity for the number of mutations present in the sequence (see below).

# Sequence Analysis of Recombinants

Thirty-two recombinant plasmids were sequenced. There was one sequence found with a deletion. There were no base changes found in the recognition sequences of the restriction enzyme sites. The lack of changes in these sequences was most likely due to the need for complementary overhanging ends for ligation. This was the only selectivity found in the nucleotide sequences. In the 31 mutants that have been sequenced, only 8 bases out of 57 possible were not altered and several residues had all 3 possible nucleotide changes (Fig. 3A). On average, there were 3.6 nucleotide changes per clone. Most of the changes alter a single residue in a codon, although there are a few codons with 2 nucleotide changes (Fig. 3B). However, as expected due to the degeneracy of the genetic code, only 79 changes actually alter an amino acid residue, an average of 2.5 amino acid changes per mutant protein. At the levels of contamination used, every available amino acid is changed at least once in this group of 31 mutants, indicating that a good representation of sequence changes can be produced in a small set of clones.



Figure 2. Colony hybridizations using a mutagenic oligonucleotide as probe. Colonies bearing plasmids with  $0-5$  base pair (bp) changes were hybridized with <sup>32</sup>P-labeled probe. 1-0 bp changes (pPLBBGIL-1 $\alpha$ ); 2-pPLRPI; 3-5 bp changes;  $4-3$  bp changes;  $5-1$  bp change;  $6-4$  bp changes;  $7-2$  bp changes.

A)

<b>AATT</b>	C ATT		CTG AAC	<b>GAT</b>	<b>GCA</b>		TTG AAC		CAG TCT		ATT ATT	CGT GCA		AAC						GAC CAG TAC CTG AC TGCA
	T	т	T T	TT	T T		TT	TT					T T	TT	T $\mathbf T$		T			
		A		A A A		A A			A AA	Α	Α	AAA			A		$\boldsymbol{\mathrm{A}}$	$\, {\bf A}$		
	G	G	GG	G	G	GG	G		G		GGG GGG			G GG	G		${\mathbb G}$		G	
		c		c c	$\mathbf c$	c <sub>c</sub>	c	C	c		$\mathbf C$	cc		cc		C	$\mathbf C$	$\mathbf C$	$\mathbf C$	
B)																				
	I	L	N	D	Α	L	N	Q	s	I	$\mathbf I$	$\, {\bf R}$	A	N	D	o	Y	L	т	
RP249						$\mathbf{1}$	$\bf H$	q			$\mathbf v$						T	$\mathbf N$		
RP256		$\mathbf R$						L	Y		T	P								
<b>RP259</b>				N	R		n			N	T	H		I			у			
RP265				E																
RP268		ı		Y	$\mathbf{s}$		n	P		i										
RP271	S		n												Y					
RP277	$\mathbf F$											r								
RP279																$\bf H$		$\mathbf N$		
RP293						ı		P		$\mathbf v$		r	a	Т	E					
RP294			Y	н		$\mathbf{1}$	I	$\bf H$									D	Q		
RP313		ı				v						S								
RP345			D			M											D			
RP402					P				s	i	T			S						
RP403				E							N							ı		
RP405		R	S											$\bf H$						
RP406		$\mathbf{P}$									S									
RP459				E		W														
RP470																			${\bf P}$	
RP494		$\mathbf 1$									M	P	S					N		
RP501																		S		
<b>RP502</b>					T						N		$\bf a$		E					
<b>RP504</b>							n		$\, {\bf R}$				$\mathbf a$	Ŀ						
<b>RP505</b>										M			a							
<b>RP506</b>					а		s					r								
RP539					$\star$															
RP562				$\mathbf{v}$										D	d					
RP581				d				L		S										
RP621				E	a	ı			T											
RP683				d				L		S										
RP706				$\mathbf E$							M								$\, {\bf A}$	
RP754		L													d					
RP761		$\overline{P}$																		

Figure 3. (A). Nucleotide changes of mutant IL-la genes. The nucleotide changes are shown below the wild-type sequence. The 8 bases corresponding to the overhanging restriction enzyme ends are at either end of the sequence. (B) Protein sequence of IL-1 $\alpha$  mutants. Single letter codes are used. Amino acids in small letters represent silent codon changes. Capital letters represent actual amino acid substitutions. Underlined amino acids indicate more than 1 base change occurred in the codon. The asterisk indicates a frameshift which results in a stop codon in the reading frame.

The Poisson distributions for contamination levels of 4.2% and 8.4% are shown in figure 4. The average of these two curves produces a distribution with a low frequency of wild-type sequences (O mutations), a plateau of roughly equivalent frequencies of 1- 5 nucleotide changes, and detectable amounts of  $6-8$  changes. The observed distribution is not significantly different from the averaged distribution  $(\chi_2$  of 11.7; 7 dof), while there is significant difference from the distributions for contamination levels of 4.2% and 8.4%. There were 112 mutations in 1767 total bases giving a mutation frequency of 6.3%. There does not seem to have been any selection for sequences with low numbers of changes. In fact, in the 32 sequences looked at, there were none with <sup>a</sup> wild-type DNA sequence.

## Protein Stability

Initial analysis of the crude lysates of several of the mutant proteins produced some interesting insights into the mutability of huIL-1 $\alpha$ . It has been demonstrated for other proteins that structurally unstable mutants do not accumulate in  $E$ . *coli* because they are rapidly proteolyzed (5). Proteins which can fold into a stable structure may be resistant to degradation and can be

isolated from the cells. Using a rabbit polyclonal antibody to huIL- $1\alpha$ , the amount of recombinant protein found in cell lysates for several of the mutants could be determined (Table 1). Most of these mutants produced substantial amounts of immunoreactive material. RP259 has almost 1/3 of the amino acids in this region changed, yet still produced large amounts of protein. Other mutants, such as RP501 and RP506, have never produced protein as measured by reaction with anti-IL-1 $\alpha$  antibody or by observation on silver-stained polyacrylamide gels (data not shown). Each possesses a single amino acid change (Fig. 3B). The apparent lack of protein may be due to changes which alter the conformation of the protein such that anti-IL-1 $\alpha$  antibodies no longer recognize it, however, the absence of any protein on a polyacrylamide gel suggests that structural instability may be the reason.

## Effect of mutations on the activity of huIL-1 $\alpha$

The ability of some of the mutant proteins to inhibit the binding of huIL-1 $\alpha$  to its receptor was examined. Using the protein concentration of each mutant as determined by the RIA, the relative affinity of each mutant could be compared to wild-type IL-1 $\alpha$  (Table 1). Every mutant which produced protein that could



Figure 4. Distribution of observed and expected nucleotide changes. For saturation mutagenesis, the number of mutations is expected to follow a Poisson distribution. The Poisson distribution is described by the equation:

$$
P = \frac{n!}{x!(n-x)!} C^{x} (1-C)^{n-x}
$$

where P is the probability of observing a specific number of changes, X, for an oligonucleotide n long with C being the amount of contamination. In calculating the appropriate distributions for the synthesized oligonucleotides, n was 57 ( 65 bases-the 8 base overhangs), X, was from  $0-8$ , and C was either  $4.2\%$  or  $8.4\%$ . The Poisson distributions for the RP region with C = 4.2% ( $\diamond$ ) and 8.4% ( $\triangle$ ) are shown, as well as the average of these two distributions  $(①)$ . The histogram displays the observed frequency of nucleotide changes per insert.

Table 1. Analysis of mutant proteins

	changes	Mutant # of amino acid Protein Concentration <sup>1</sup> (ng/ml)	<b>Binding</b> Activity <sup>2</sup>	<b>Biological</b> Activity <sup>3</sup>
RP <sub>249</sub>	4	792	0.8	<b>ND</b>
RP259	6	354	0.85	ND
<b>RP345</b>	-3	678	6	0.007
RP459	2	162	$\overline{c}$	0.8
<b>RP494</b>	4	1555	4.5	0.02
<b>RP501</b>		<b>ND</b>	<b>ND</b>	ND
<b>RP504</b>	-2	45	0.7	0.8
<b>RP506</b>	-1	<b>ND</b>	<b>ND</b>	<b>ND</b>
<b>RP562</b>	2	144	0.92	1.1
<b>RP581</b>	2	219		0.1
RP621	2	672	4.9	0.2
<b>RP683</b>	$\mathfrak{D}$	747	0.6	0.05
<b>RP754</b>		174	6.6	<b>ND</b>
RP761		375	2.8	0.4

protein concentration was determined in an RIA using rabbit polyclonal antibodies raised against huIL-1 $\alpha$ ; <sup>2</sup>expressed as a fraction of wild-type IL-1 $\alpha$ activity. Binding activity was determined using a binding inhibition assay (14). Activity was normalized to internal wild-type IL-1 $\alpha$  controls which were expressed in E. coli; <sup>3</sup>expressed as a fraction of wild-type IL-1 $\alpha$  activity. Biological activity was determined using the ELA conversion assay (15). Activity was normalized to internal wild-type IL-1 $\alpha$  controls which were expressed in E. coli; ND-none detected.

be detected by RIA also had significant activity in the binding assay. This ability to bind to the IL-1 receptor with an affinity similar to wild-type IL-1 $\alpha$  would suggest that these mutants did not display a large alteration in protein structure.

The mutations had a much greater effect on the biological activity. Only RP459, RP504 and RP562 have biological activities close to wild-type levels. The rest all have reduced activity on EL4 cells with RP249, RP259 and RP754 having no detectable biological activity. Several of the mutants with reduced biological activity still retain substantial binding activity. Thus, while the proteins apparently have the structural specifications allowing interaction with the receptor, they appear deficient in interactions which determine biological activity.

## **DISCUSSION**

When attempting to understand protein structure-function relationships, a useful approach is to generate random nucleotide changes in the coding sequence for the protein. However, while it is the nucleotide sequence that is being altered, it is the amino acid sequence which is being examined. It is difficult to predict beforehand how many nucleotide changes need to be made in order to produce <sup>a</sup> useful number of amino acid changes. A low frequency of nucleotide changes results in a large proportion of the recombinants having no amino acid changes, with many sequences having no changes or having silent mutations. Thus, a large number of mutant proteins must be examined in order to sample changes at each amino acid. Simply increasing the number of nucleotide changes would not satisfactorily solve the problem, since the increased number of mutations might have a large effect on protein structure. The ideal distribution would be one which resulted in low levels of unchanged nucleotide sequences, and equal probabilities of one or more nucleotide changes. There would be no need to empirically determine the best level of contamination for every protein. Saturation mutagenesis using unequally contaminated oligonucleotides adequately satisfies these criteria.

Sequence analysis of 31 mutants generated a distribution of mutations which fit the averaged distribution of the two contamination levels. Altered nucleotides were found throughout the region synthesized with the only selectivity being found in the 8 bases making up the EcoRI and PstI overhanging restriction enzyme ends. No mutations were found here, presumably due to the need for complementarity in order to properly join the ends of the cassette to the plasmid. The contamination levels chosen resulted in approximately equal probabilities of having one to five nucleotide changes. In addition, several clones were found with 7 or 8 nucleotide changes. Approaches using only one contamination level can result in differences of 5-fold between the fraction of clones with low numbers of nucleotide changes and those with higher numbers. Using unequal contamination should also be useful when mutating non-coding regions. If a powerful selection is available, unequal contamination could quickly pick out a wide range of single and multiple changes from a single construction.

The observed mutation frequency using only one contamination level is often substantially lower than expected (6, 7). This may arise from systematic errors in the synthesis of the oligonucleotides, due to inactivation of the added phosphoramidites. Additionally, there may be selection for the wild-type sequence since the lack of mismatches produces a more stable DNA duplex. Using two different contamination levels helps ameliorate these effects. The presence of two different contamination levels will buffer the effect of systematic errors. While the strand with the lower contamination level may have less mutations than expected, the presence of the complementary strand with a higher contamination level should still prevent a large increase in the presence of wild-type nucleotide sequences.

The important factor in any type of mutagenesis which examines <sup>a</sup> coding sequence is not simply the rate of DNA changes, but the frequency of actual amino acid changes. In this study, there were on average 2.5 amino acid changes per protein. This compares with the nucleotide rate of 3.6 changes. If roughly 30% of the nucleotide changes are silent, then generating single nucleotide changes will result in many unchanged amino acid sequences, increasing the presence of wild-type protein

sequences. The fraction of silent changes will vary between proteins because different sequences will have different levels of degeneracy. Increasing the number of nucleotide changes overcomes this problem by raising the probability of multiple nucleotide changes, while also raising the likelihood of generating two nucleotide changes in a single codon. However, the increased chance of creating multiple amino acids changes could adversely affect protein stability. The number of changes that can be accommodated before this happens will vary from protein to protein, making it difficult to predict what contamination level to use. Using unequal contamination levels for the two oligonucleotides produces both low numbers of substitutions, useful for generating single amino acid changes, and high numbers of substitutions, useful for overcoming silent mutations or for generating two or more nucleotide changes per codon. Since the relative frequencies of <sup>1</sup> to 5 changes are similar, there is no need to worry that an inappropriate level of mutagenesis will be found. This technique yielded <sup>1</sup> to 6 amino acid changes from a single construction, resulting in a wide range of substitutions in just the 31 mutants examined. All of the amino acids which could be altered were mutated at least once. In a pool of 300 mutants one would expect that each amino acid residue would be altered 10 times in conjunction with a wide range of other amino acids. It should be then be possible to quickly identify an area of the protein sequence which affects function.

The mutants generated in the EcoRI-PstI region of the huIL- $1\alpha$ gene are being screened for structural and functional effects. Even in the small number examined there are some interesting effects. RP501 and RP506, which have single amino acid changes, appear to have large disruptions in their structure. No protein has been observed for either mutant, either on polyacrylamide gels or by RIA. The lack of detectable protein may result from proteolytic degradation in the cell (5) or a change in binding epitopes such that polyclonal antibodies no longer recognize the protein. Either case indicates that these amino acid changes have had an effect on the structural integrity of the protein. The binding affinities for the IL-I receptor of many of the other mutants appear to be similar to wild-type IL-1 $\alpha$ . This seems to indicate that altering many of the amino acids in this region of the protein does not alter the structure of the protein greatly. As many as 6 amino acid changes in the molecule has little effect on the ability of the protein to bind to the IL-1 receptor.

Comparing the effects of the mutations between the EcoRI and PstI sites on binding activity and biological activity helps separate structural effects from functional ones. Several reports indicate that altering amino acids in this region in IL-1 $\alpha$  or IL-1 $\beta$  affects biological activity without changing binding affinities (16, 17). Of the mutants examined in this paper, only RP459, RP504 and RP562 have activities close to wild-type values. In contrast, RP683 has a greatly reduced biological activity with only a slight decrease in binding activity. The other mutations also show decreased biological activity while maintaining significant ability to bind to receptor. RP754, with a single amino acid change (Fig. 3B), has no detectable biological activity. RP761 also has a single amino acid change at same this location. This Leu residue is very highly conserved in molecules from 4 different species (8). The other six mutants with lowered biological activities have more than one amino acid altered. One of the residues, Asp26, is altered in several mutants that display low biological activity. This residue has been reported to be involved in determining different biological activities for IL-1 $\alpha$  (16). Site-directed mutagenesis

should clarify the role of this amino acid. Altering the entire huIL- $1\alpha$  gene using this approach should allow amino acids which are important for activity to be quickly identified and studied.

## ACKNOWLEDGMENTS

We wish to thank David Cosman and Steve Dower for insights into lymphokine biology. We also wish to thank the DNA Synthesis group for producing the oligonucleotides and Norman Boiani for performing the radioimmunoassays.

## **REFERENCES**

- 1. Smith, M. (1985) Annu. Rev. Genet.., 19, 423-462.
- van Gunsteren, W.F. (1988) Protein Engineering, 2, 5-13.
- Blundell, T.L., Sibanda, B.L., Sternberg, M.J.E. and Thornton, J. M. (1987) Nature, 326, 347-352.
- 4. Shortle, D. and Botstein, D. (1983) Methods Enzymol.,  $100$ ,  $457 468$ .<br>5. Bowie, J.U. and Sauer, R.T. (1989) Proc. Natl. Acad. Sci. USA, 86
- 5. Bowie, J.U. and Sauer, R.T. (1989) Proc. Natl. Acad. Sci. USA, 86, 2152-2156.
- 6. McNeil, J. B. and Smith, S. (1985) Mol. Cell. Biol., 5, 3645-3551.
- 7. Murray, R., Pederson, K., Prosser, H., Muller, D., Hutchison, C.A. III and Frelinger, J.A. (1988) Nucleic Acids Res., 16, 9761-9773.
- 8. Maliszewski, C.R., Baker, P.E., Shoenborn, M.A., Davis, B.S., Cosman, D., Gillis, S. and Cerretti, D.P. (1988) Mol. Immunol., 25, 429-437.
- 9. Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.J., Smith, J.A. and Struhl, K. (1988) Current Protocols in Molecular Biology. John Wiley and Sons, New York.
- 10. Coulondre, C. and Miller, J. (1977) J. Mol. Biol., 117, 525-575.
- 11. Bernard, H. and Helinski, D.R. (1979) Methods Enzymol., 68, 482-492.
- 12. Sanger, F.S., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA., 74, 5463-5467.
- 13. Chard, T. (1987) An Introduction to Radioimmunoassay and Related Techniques. Elsevier, New York.
- 14. Mosley, B., Urdal, D.L., Larsen, A., Cosman, D., Conlon, P.J., Gillis, S. and Dower, S.K. (1987) J. Biol. Chem., 262, 2941-2944.
- 15. Mosley, B., Dower, S.K., Gillis, S. and Cosman, D. (1987) Proc. Natl. Acad. Sci. USA., 84, 4572-4576.
- 16. Yamayoshi, M., Ohue, M., Kawashima, H., Kotani, H., IIda, M., Kawata, S., and Yamada, M. (1990) Lyphokine Res. 9, 405-413.
- 17. Gehrke, L., Jobling, S.A., Paik, L.S.K., McDonald, B., Rosenwasser, L.J. and Auron, P.E. (1990). J. Biol. Chem. 265, 5922-5925.