DOE/ER/60533--3

DE91 002373

Progress Report on Grant DE-FG02-87ER60533
entitled, "Studies of Human Mutation Rates"
for period November, 1989-October, 1990

To: Mr. Dana R. Dixon, Acquisition and Assistance Unit, Dept. of Energy, Chicago Operations Office, 9800 S. Cass Avenue, Argonne, IL 60439

From: Prof. James V. Neel, Principal Investigator, Dept. of Human Genetics, University of Michigan Medical School, Ann Arbor, MI 48109-0618

Date: 13 September 1990

November, 1989, marked the beginning of a new three-year cycle of DoE grant support, in connection with which the program underwent a major reorganization. The three objectives of the present program are:

1) To isolate by the technique of two-dimensional polyacrylamide gel electrophoresis (2-D PAGE), proteins of special interest because of the relative mutability of the corresponding gene, establish the identity of the protein, and, for selected proteins, move to a characterization of the corresponding gene;

2) To develop a more efficient approach, based on 2-D PAGE, for the detection of variants in DNA, with special reference to the identification of mutations in the parents of the individual whose DNA is being examined; and,

3) To continue an effective interface with the genetic studies on the children of atomic bomb survivors in Japan, with reference to both the planning and implementation of new studies at the molecular level.

We will present the progress under four headings:

Project I. Isolation of Proteins of Special Genetic Interest from Two-Dimensional Polyacrylamide Gel Electrophoresis (Strahler).

This section of the progress report outlines our efforts to isolate polypeptides of interest identified by two-dimensional polyacrylamide gel electrophoresis in sufficient quantities for studies of their amino acid sequences. The basic approach to obtaining purified protein utilizes the high resolution of 2-D PAGE as the final step in the purification. Two fundamental problems have been addressed in the progress presented below. One is the ability to recover protein from 2-D gels in amounts sufficient to yield reliable sequence information for a sufficient number of amino acid residues to unambiguously identify the polypeptide in a sequence database or, if the protein represents a novel sequence, to obtain sufficient material that sequence information enabling the construction of an olignonucleotide probe having low redundance for cloning purposes can be constructed. Since the majority

MASTER

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

of the proteins of interest, from the point of view of mutation and genetic variability, are known only from their detection on 2-D gels, and represent moderate to low abundance proteins in human lymphoid cells, enrichment protocols which allow unambiguous identification of the target protein(s) in the 2-D gel "assay" are essential. A second obstacle to sequencing is that of N-terminal blockage, whether artifactually produced during purification or naturally occurring *in vivo*. The majority of soluble proteins in mammalian cells are believed to be N-terminally blocked, most frequently as a result of N-acetylation. Thus, a strategy has been implemented which will generate internal amino acid sequences. Several of the steps in the production of preparative 2-D gels have been investigated in some detail and are presented below.

1) Increased loading of protein onto 2-D gels: As a first approach to obtaining the material for sequencing polypeptides of interest, as determined from our prior experiences with chemical mutagenesis and studies of genetic variation, we examined our ability to purify polypeptides on 2-D gels in sufficient quantities to obtain direct N-terminal sequence information and to determine the amount of material necessary with existing "state of the art" microsequencing capability. Two approaches to the generation of preparative 2-D gels were followed; these differed in the methodology used for the first dimension isoelectric focusing separation.

A. Conventional carrier ampholyte (CA)-based isoelectric focusing was used (250 μg total protein per gel), essentially as we have done in prior sequencing efforts. Protein spots visualized with Coomassie Blue at this relatively high "load" retained overall resolution comparable to that obtained for "analytical" gels. About 100 spots are visualized in preparations of lymphocytoid cells, of which <50 are of sufficient intensity to warrant attempts at isolation for sequencing. Increasing the amount of protein per gel (400-650 μg) results in drastic deterioration of the first dimension separation. Some pI regions of the gel, however, do maintain the required resolution. We have obtained material for sequencing using a protein load of 250 μg soluble cytosolic protein per gel. Proteins were electroblotted onto Glassybond membranes (siliconized glass fiber sheet, Biometra) using a semi-dry electroblotting apparatus.

During the past year, prior to the arrival on campus of Dr. Andrews, the protein sequencing aspects of this project have been contracted with the Max Planck Institute of Biochemistry in Martinsried, West Germany. Employing the above-referenced methodology, we have made two different sets of polypeptides available to them. In a first exercise, 28 major polypeptides were isolated for direct N-terminal sequence analysis, 8 of which represented the products of loci which had mutated in prior ENU mutagenesis studies or which were loci at which spontaneous genetic variation had been observed. Sequence information was obtained for 7 of the 28 polypeptides. Five of them were known proteins while 2 represented novel sequences. Twenty-one, or 75%, of the polypeptides yielded no sequence information, presumably due to N-terminal blockage. In another group of polypeptides isolated from CA-based 2-D gels of soluble nuclear polypeptides, 18 were selected for

analysis. Of these, 8 were the products of loci at which protein variants had been recognized (either ENU induced or spontaneous). Eleven polypeptides yielded sequence information. Eight of the 11 were known proteins, and 3 yielded novel sequences. Three of the 11 were in the special interest category, i.e., 3 had exhibited genetic variation in population surveys, and, of these 3, 2 had exhibited induced variation in ENU studies (sequence information was obtained from the "normal" allele in all cases), and all 3 were subsequently found to be known proteins. Two polypeptides were common to both the cytosolic and nuclear sets and both yielded sequence information. Thus, of the 46 different polypeptides analyzed by direct N-terminal sequence analysis, and which were in the moderate to high abundance class of polypeptides, sequence information was obtained for 16 of them. These results are summarized in Appendix A, which illustrates the form in which the data on these proteins is stored in a computer. (We will be happy to supply the code to anyone interested.)

B. As an alternative approach, we have used immobilized pH gradient gels for polypeptide separation in the isoelectric focusing axis. About 1-1.2 mg. total protein per gel can be loaded on such gradient gels resulting in 4-5-fold more protein per gel. About 200 proteins are visualized with Coomassie Blue, increasing substantially the number of "sequencable" proteins from the same number of gels as contrasted with carrier ampholyte-based separations. While the vast majority of spots identified on CA-based 2-D gels can be identified also on IPG-based gels, occasionally a spot of interest is not readily identifiable on the IPG-based pattern. We have utilized known structural variants of 5 different proteins (identified on CA-based 2-D gels) to verify the location of these proteins on IPG-based gels. To assess the number of gels required to obtain sequence information, IPG-based 2-D gels were prepared from whole cell extracts of a lymphoid cell line (Jurkat). Polypeptides were isolated by electroblotting onto PVDF membranes (Immobilon P), using a semi-dry blotting apparatus. Fourteen polypeptide spots were excised from 6 to 10 blots and used in a collaborative effort with Dr. Jack Watson, Director of the Mass Spectrophotometry Facility at Michigan State University. Twenty or more blots derived from CA-based 2-D gels would have been needed to harvest the amount of protein resulting in these sequencing results. The increased protein capacity and maintenance of high resolution separations of IPG-based 2-D preparative gels significantly reduces the number of gels needed for a given separation.

<u>Electroblotting onto "Glassybond" and "Immobilon P" membranes</u>: Comparison of "sequencability" of selected polypeptide spots electroblotted onto either Glassybond (a siliconized glass fiber sheet) or Immobilon P (a polyvinylidene difluoride) membranes have given comparable results when equivalent amounts of polypeptide were studied. Four polypeptides yielded sequence information useful in this comparison. For high abundance proteins such as cyclin [also termed proliferating cell nuclear antigen (PCNA)], triose phosphate isomerase, and HuCha60 (human homologue to the chaperonin family), as few as 2 electroblots from either type of membrane provided the protein from which sequence information could be obtained, whereas 6 blots (IPG-based 2-D gels blotted onto PVDF) or 20 blots (CA-based gels blotted onto

Glassybond) were required for the moderately abundant protein glutathione S-transferase. Conversely several polypeptides which did not yield any sequence data from PVDF blots also did not result in sequence information from Glassybond.

Rotophor enrichment: The value of a first step partial purification of a very low abundance protein from a complex protein mixture, utilizing liquid phase isoelectric focusing on a large scale, has been explored. The target protein in this study was a low abundance 28 kD nuclear protein (np28) which is induced in the Jurkat T cell line following activation with phorbol esters and calcium ionophor or is constitutitively expressed in tat III transfected Jurkat cells. The protein did not yield a sequence by N-terminal sequencing. To obtain sufficient material for generation of internal peptide sequences, a nuclear extract derived from 4.5×10^9 cells was fractionated under denaturing conditions (urea, non-ionic detergent and dithioerithritol) similar to the conditions of the first dimension of 2-D gel separations, using the Rotophor apparatus (BioRad) operated at 12 w constant power for 6 h. The preparation was fractionated according to the manufacturer's instructions into 20 aliquots. Np28 was enriched about five-fold in two acidic Rotophor fractions, which were then used to purify the protein on preparative 2-D gels. Gels were stained with Coomassie Blue, and np28 excised from 40 gels. Following extensive washing and drying of the gel pieces, protein was digested *in situ* with trypsin (~1:10; enzyme:protein) and the peptides recovered by sequential elution with 75% trichloroacetic acid (TFA) in water followed by 50% TFA in acetonitrile. Tryptic peptides were separated by reverse phase HPLC (Vydac 204 TP column). Reliable sequence information was obtained for 2 of the isolated peptides, allowing identification of homology with a known highly conserved nuclear phosphoprotein B23.

Project II. Preparation of Partial Amino Acid Sequences of the Proteins Isolated from Gels (Andrews).

Andrews did not assume his position at the University of Michigan until May 1, 1990, i.e., six months into this grant year. In the interim, the protein-sequencing needs of the project were met by a contractual arrangement with the Max Planck Institute for Biochemistry in Martinsried, West Germany. Strahler played a key role in expediting this contract, and in his (preceding) section has already summarized those results.

The effort in protein sequencing here at Michigan has since May 1, 1990, proceeded along four lines.

1) The first area of concern has been to establish a microbore HPLC system for the isolation of proteolytic fragments of proteins in 50 to 250 pmol quantities, a facility previously lacking. This effort has involved the modification of an ABI model 130A to allow large volumes of dilute protein solutions to be injected onto 1 mm diameter, reversed-phase columns. This configuration was first used to obtain the sequence of several proteolytic fragments from two DNA-binding proteins of previously unknown sequence at approximately 60 pmol quantity. These first samples were in solution so methods had to be developed fo

isolation of fragments from proteins adsorbed to PVDF membranes. Previously published methods for this purpose were found to have significant drawbacks. The technique now in use in this facility involves chemical reduction and cleavage of the protein on the PVDF membrane followed by elution and proteolysis. This technique has been applied to the protein anticipated to be chaperonin, transferred to PVDF membrane from a two-dimensional gel of a whole cell extract. The HPLC profile of the tryptic digest contained several well-resolved peaks. These have been submitted for sequence analysis to verify that the protein is human chaperonin. The first tier of "candidate" proteins (i.e., proteins of special interest) is currently undergoing similar treatment.

The second area of progress has been the acquisition and operation of an ABI model 473 pulsed-liquid phase protein sequencer. The sequencer has since been upgraded by addition of a specialized reaction cartridge designed for sequencing proteins directly from PVDF membranes. A modified sequencing protocol in conjunction with this cartridge has significantly enhanced sequencing yields of proteins adsorbed directly to PVDF. This development will be particularly useful for obtaining amino-terminal sequence from proteins resolved using 2-D gels in this study.

The third area has been the hiring and training of a technical support person for working with picomole quantities of proteins. Progress in this area has been very successful as evidenced by the results cited above.

The fourth area has been the re-evaluation of candidate proteins on the basis of their degrees of probable biological significance and of feasibility. Several particularly interesting proteins have been selected and sufficient quantities collected for proteolytic digestion and sequence analysis.

Project III. Preparation of DNA Probes and Development of an Improved Method to Identify Mutations in the Human Genome (Kurachi).

In the past one year, this laboratory has focused its DoE-sponsored efforts on two projects. One is to develop a better method to detect gene mutations efficiently. The other is to identify and clone new genes of particular interest in terms of our past mutation rate studies.

A) Development of an Improved Method to Detect New Mutations.

Applications of both the polymerase chain reaction (PCR) and denaturing gradient gel electrophoresis (DGGE) in detecting mutation have been well described (Takahashi *et al*, Mutation Res. 1990; 234:61; Trystman *et al*, Genomics 1990; 6:293). The approach which utilizes PCR and DGGE still appears to be the most practical and useful one. However, it must be modified and optimized to best serve our objective of efficiently detecting single nucleotide mutations.

In our initial approach, we have used a combination of multiplex PCR and 2-D DNA gel electrophoresis to analyze the factor IX gene as a model. Multiplex PCR has already been successfully used to provide analytic quantities of some genes, such as the Duchenne muscular dystrophy gene. In our experiments with the factor IX gene, all eight exons could be amplified in one PCR reaction. However, some exons, including exon 7 and 8, presented some problems. These included read-through amplification by the 5' most and 3' most primers, generating some of the background amplification bands. This may be avoided if we include a significant segment of intron sequence as well as the flanking sequence. At the present, however, we are carrying out two to three separate multimeric PCRs to avoid this problem and to amplify all the exons and the 5' end flanking sequence containing the regulatory region. The optimal conditions for amplifying all regions at about the same yield were worked out. The PCR products of the normal gene (reference DNA) as well as of the abnormal genes isolated from hemophilia B patients were mixed at equimolar amounts and heat denatured prior to loading onto the non-denaturing polyacrylamide gel (6.5%). All the PCR amplified fragments were separated in the electrophoresis. The gel strips containing the separated DNAs were cut out and subjected to two-dimensional DGGE (6.5% polyacrylamide with a gradient of 0% to 80% formamide). PCR-amplified DNA fragments were well separated in the denaturing gradient gel, forming easily identifiable DNA spots. This method efficiently located a mutation to exon 8 in a hemophilia B patient. If the DNA to be analyzed is known to be heterozygous (e.g., in the normal but carrier mother of hemophilia B patients), then no mixing of the normal reference DNA and the patient's DNA is necessary. This approach is significantly more efficient in detecting unknown mutations in a specific gene than a combination of separate PCRs and single lane DGGE analysis. We are also extending this approach so that we can directly analyze restriction fragments of the whole genomic DNA instead of PCR products. If this latter approach is successful, we can substantially speed up the initial screening of mutations in both well-characterized and partially-characterized genes.

B) Cloning and Characterization of New Genes.

In a separate line of experiments, we are also carrying out a series of protein 2-D gel electrophoresis analyses to detect capillary endothelial cell genes involved in the process of angiogenesis. So far, we have detected several protein spots in the 2-D gel analysis which apparently change their phosphorylation status upon exposure to known angiogenic factors such as angiogenin, FGF, and others. We anticipate that this series of experiments may also provide us with a number of very interesting genes to be employed in mutation rate studies.

IV. The Japanese Interface (Neel).

A major objective the past several years has been an analysis of all the genetic data accumulated in Japan over the past 42 years in the light of the recently available DS86 radiation exposure schedule, and the estimation of the most probable genetic doubling dose of radiation suggested by these data. This analysis and estimate, a major paper, appeared in June, 1990, and has attracted a great deal of attention. We

suggest that the doubling dose to gonial cells of acute radiation of the Hiroshima-Nagasaki type is in the neighborhood of 2.0 Sv, and that the doubling dose for chronic radiation should be in the neighborhood of 4.0 Sv. This latter is about four times higher than the usual extrapolation from the mouse experimental data to the human situation, but in an extensive review of the human and mouse data, undertaken in collaboration with S. E. Lewis, completed this fiscal year and now in press in the *Annual Review of Genetics*, we suggest that when <u>all</u> the murine data are considered in perspective, the doubling dose following the chronic irradiation of mice is also in the neighborhood of 4.0 Sv. These two papers, then, suggest a major revision in thinking concerning the genetic risks of radiation for humans.

During the past year, in collaboration with W. J. Schull of the University of Texas, plans have matured to republish an anthology of the 13 most important papers concerning the genetic effects of the atomic bomb experience, together with a (new) introductory chapter providing historical perspective and a concluding chapter discussing possible future studies. This material has been reviewed for quality through the Board on Radiation Effects Research of the National Research Council, and, as of this writing, a decision has just been reached that the material will be published as a book by the National Academy of Sciences Press.

Unexpected personal developments prevented Neel from making his annual trip to Japan during the course of this fiscal year, but two trips are now projected within the upcoming fiscal year. One of these, in November, 1990, in addition to the usual consulting concerning the ongoing genetics program, will include participation in a workshop concerned with the future research projects of the Radiation Effects Research Foundation. The other, projected for May, 1991, will be for participation in a workshop focussing on the design of future genetic studies at the DNA level with respect to the children of survivors and of controls.

Finally, Neel has participated during this year in two symposia concerning the biological effects of radiation, namely, a Brookhaven National Laboratory Symposium on "DNA Damage and Repair in Human Tissues" and a UCLA Symposium on "Ionizing Radiation Damage to DNA: Molecular Approaches." In both instances, he delivered papers which will be published in the proceedings of these symposia.

The publication of these various papers will essentially mark the end of the morphological-cytogenetic-biochemical era in the study of radiation genetics. Any major new round of studies will almost certainly concentrate on DNA. The issues involved in developing an appropriate DNA system are extremely complex. In this connection, Neel has published this past year a paper acknowledging DoE support, in which the hypothesis is developed that the genes responsible for the proteins visualized on two-dimensional polyacrylamide gel electrophoresis may as a class be less mutable than the genes responsible for the erythrocyte enzymes and serum transport proteins which have been the basis for most of the recent studies on human mutation rates. Furthermore, in collaboration with Ranajit Chakraborty, Neel has within the past year

published an analysis suggesting that there may be as much as a 20-fold difference in the mutation rates of the genes responsible for erythrocyte enzymes and serum proteins. These two papers illustrate the great care that must be exercised in selecting the DNA segments to be employed in any study of DNA mutation rates.

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Appendix A. Storage format for data on proteins being analyzed in the course of this Project.

	A	B	C	D	E	F	G	H	I
1	STRUCTURAL STUDIES OF 2D POLYPEPTIDES - DoE								
2									
3	Spot#:	2	22	26	28	34	35	58	59
4	Master gel image:	389y	389y	389y	389y	389y	389y	389y	389y
5	Trivial name(s):	np9	np15	np8	np5, allele of spot29	np10	np11	L25	L24
6	Species:	human	human	human	human	human	human	human	human
7	Source (tissue/cell line):	TK6	TK6	TK6	TK6	TK5	TK6	TK6	TK6
8	Features:								
9	Genetic variation:	monomorphic	monomorphic	monomorphic	variable	monomorphic	monomorphic	monomorphic	monomorphic
10	Induced mutations:	no	no	no	mutation (deletion)	mutation	mutation	no	mutation
11	Variant gel# and/or clone:	–	–	–	662hh(E4-10E84)	(E4-10D45)	A21,E4-10E88)	–	)B89,E5-50E91)
12	Normal MW	90	73	62	58	67	69	56	55
13	pI=f(x);[x, pixels]	172	341	130	253	441	487	864	797
14	post-trans mod								
15	Variant MW	–	–	–	58	68	69	–	55
16	pI=f(x);[x, pixels]	–	–	–	262	432	504	–	763&812
17	post-trans mod								
18	Sequencing Results								
19	source:	Jurkat	Jurkat	Jurkat	Jurkat	Jurkat	Jurkat	Jurkat	Jurkat
20	enrichment:	ne	ne	ne	ne	ne	ne	cyto	cyto
21	facility:	mp	mp	mp	mp	mp	mp	mp	mp
22	sample type:	sgf	sgf	sgf	sgf	sgf	sgf	sgf	sgf
23	N-term/internal	N-terminal	N-terminal	N-terminal	N-terminal	N-terminal	N-terminal	N-terminal	N-terminal
24	residues sequenced:	13/16r	17/17r	9/9r	15/15r	0	0	0	0
25	sequence	(V)/(G)EVKR(P)XGV(S)LTNX(P)F	EEEDKKEDVGTVVG(I)DL	EPAVYFKEQ	ADAPEEEDHVLVLRK				
26	db search	genbank/mipsx	genbank	genbank	genbank				
27	identity/homology (ref)	new	#A29821	B31345	A29787/A26632				
28	useful probe	no	yes	yes	yes				
29	n mer		any length	any length	any length				
30	degeneracy		0	0	0				
31	other								

	J	K	L	M	N	O	P	Q	R	S	T
1											
2											
3	85	85	85	85	91	96	136	155	159	161	162
4	389y	389y	389y	389y	389y	389y	389y	389y	389y	389y	389y
5	HuCha 60	HuCha 60	HuCha 60	HuCha60	np4	L26	L7	L8	L11	L18	L15
6	human	human	human	human	human	human	human	human	human	human	human
7	TK6	TK6	TK6	TK6	TK6	TK6	TK6	TK6	TK6	TK6	TK6
8											
9	variable	variable	variable	variable	monomorphic	monomorphic	monomorphic	monomorphic	monomorphic	monomorphic	monomorphic
10	mutation	mutation	mutation	mutation	no	mutation	no	no	no	mutation	mutation
11	568ll (E5-50D44)				_	E13,E5-50C100)	_	_	_	2hh (E5-50B20)	12hh (E5-50B8)
12	58	58	58	60	52	48	42	39	38	37	37
13	421	421	421		238	309	544	867	818	669	559
14											
15	58	58	58	58	_	50/48	_	_	_	37	37
16	454	454	454	454	_	332/289	_	_	_	591	544
17											
18											
19	Jurkat	Jurkat	Jurkat	TK6	Jurkat	Jurkat	Jurkat	Jurkat	Jurkat	Jurkat	Jurkat
20	ne	cyto	cyto	wc/IPG	ne	cyto	cyto	cyto	cyto	cyto	cyto
21	mp	msu	msu	um	mp	mp	mp	mp	mp	mp	mp
22	sgf	pvdf	pvdf	pvdf(10)	sgf	sgf	sgf	sgf	sgf	sgf	sgf
23	N-terminal	N-terminal	N-terminal	internal	N-terminal	N-terminal	N-terminal	N-terminal	N-terminal	N-terminal	N-terminal
24	4/4r	15/15r	20/20r	pending	0	0	0	too low	9/9r	0	0
25	AKDV	AKDVKFGADARALML	AKDVKFGADARALMLQGVDL						IDLE(D)(L)EEG		
26	genbank/mipsx	genbank/mipsx	genbank/mipsx						genbank/mipsx		
27	E33060;E34671;J35030	060;E34671;J35030	_/E33060;E34671;J35030						NEW		
28	yes	yes	yes						no		
29	any length	any length	any length								
30	0	0	0								
31											

	U	V	W	X	Y	Z	AA	AB	AC	AD
1										
2										
3	163	166	174	174	176	177	177	178	178	178
4	389y	389y	389y	389y	389y	389y	389y	389y	389y	389y
5	L20	L21	cyclin	cyclin	np3			L30	L30	L30
6	human	human	human	human	human	human	human	human	human	human
7	TK6	TK6	TK6	TK6	TK6	TK6	TK6	TK6	TK6	TK6
8										
9	monomorphic	monomorphic	monomorphic	monomorphic	monomorphic	monomorphic	monomorphic	monomorphic	monomorphic	monomorphic
10	mutation	no	no	no	no	no	no	mutation	mutation	mutation
11	(E4-10C61)	_	_	_	_			(E5-50E101)		
12	36	35	36kD	36kD	33	33	33	33	33	33
13	503	466	212	212	157			232	232	232
14										
15	36	_	_	_	_	_	_	33	33	33
16	503	_	_	_	_	_		239	239	239
17										
18										
19	Jurkat	Jurkat	Jurkat	Jurkat	Jurkat	Hester	TK6	Jurkat	Hester	TK6
20	cyto	cyto	NE	cyto	ne	wc	wc	cyto	wc	wc
21	mp	mp	mp	msu	mp	um	um	mp	um	um
22	sgf	sgf	SGF	pvdf	sgf	pvdf(11)	pvdf(15)	sgf	pvdf(13)	pvdf(15)
23	N-terminal	N-terminal	N-terminal	N-terminal	N-terminal	internal	internal	N-terminal	internal	internal
24	0	0	6/6r	15/15r	28/30r	pending	pending	0	pending	pending
25			MFEARL	MFEARLVQGSILK(K)V	LHTDGDKAFVDFLSDEIKE(E)R/(G)KIQIH(K)TXL					
26			genbank/mipsx	genbank/mipsx	genbank/mipsx					
27			_/A54532	_/A54532	new					
28			yes	yes	yes					
29			any length	any length	14/17					
30			0	0	32/48					
31										

	AE	AF	AG	AH	AI	AJ	AK	AL	AM	AN
1										
2										
3	179	189	195	208	209	210	223	234	235	235
4	389y	389y	389y	389y	389y	389y	389y	389y	389y	389y
5	poly71,np1	L16	L27	L12, TPI	L6	L5, TPI	L17	L14	L13, GTST	I11, GTST
6	human	human	human	human	human	human	human	human	human	human
7	TK6	TK6	TK6	TK6	TK6	TK6	TK6	TK6	TK6	TK6
8										
9	variable	monomorphic	variable	monomorphic	monomorphic	monomorphic	monomorphic	monomorphic	monomorphic	monomorphic
10	no	mutation	no	no	no	no	no	no	no	no
11	ref 1	(E4-10E9)	–	–	–	–	–	–	–	–
12	31	31	31	28	30	29	26	25	26	
13	177	577	669	802	905	923	656	499	506	
14										
15	31	31	31	–	–	–	–	–	–	–
16	162	597	598	–	–	–	–	–	–	–
17										
18										
19	Jurkat	Jurkat	Jurkat	Jurkat	Jurkat	Jurkat	Jurkat	Jurkat	Jurkat	Jurkat
20	ne	cyto	cyto	cyto	cyto	cyto	cyto	cyto	cyto	wc
21	mp	mp	mp	mp	mp	mp	mp	mp	mp	msu
22	sgf	sgf	sgf	sgf	sgf	sgf	sgf	sgf	sgf	pvdf
23	N-terminal	N-terminal	N-terminal	N-terminal	N-terminal	N-terminal	N-terminal	N-terminal	N-terminal	N-terminal
24	15/17r	0	0	11/11r	0	8/8r	0	0	9/9r	9/12R
25	P/(G)FGDLKS(P)AGLQVXN/(L)D(R)			APSRKFFVGGN		APSRKFFV			PPYTVVYFP	XXYTVVYFXVRG
26	genbank			genbank/mipsx		genbank/mipsx			genbank/mipsx	genbank/mipsx
27	#Y00322			_/A52728		_/A52728			_/B59286	_/B59286
28	no			yes		yes			yes	yes
29	14			any length		any length			any length	any length
30	64			0		0			0	0
31										

	AO	AP	AQ	AR	AS	AT	AU	AV	AW
1									
2									
3	241	242	243	248	255	256	257	800	800
4	389y	389y	389y	389y	389y	389y	389y	367yy	367yy
5	L9	np6	np7	np2	np14	np13	np12	np28	np28
6	human	human	human	human	human	human	human	human	human
7	TK6	TK6	TK6	TK6	TK6	TK6	TK6	PBMC	PBMC
8									
9	monomorphic	monomorphic	monomorphic	monomorphic	monomorphic	monomorphic	monomorphic	monomorphic	monomorphic
10	no	mutation	mutation	no	no	no	mutation	-	-
11	-	(E4-10E49)	6hh(E5-50A14)	-	-	-	(E5-50E114)	-	-
12	27	29	30	26	19	19	17	28	28
13	356	363	384	279	33	368	374	130	130
14									
15	-	29	30	-	-	-	17	-	-
16	-	384	365	-	-	-	430	-	-
17									
18									
19	Jurkat	Jurkat	Jurkat	Jurkat	Jurkat	Jurkat	Jurkat	Jurkat-tatIII	Jurkat-tatIII
20	cyto	ne	ne	ne	ne	ne	ne	ne	ne/rotophor
21	mp	mp	mp	mp	mp	mp	mp	mp	mp
22	sgf	sgf	sgf	sgf	sgf	sgf	sgf	sgf	cb gel
23	N-terminal	N-terminal	N-terminal	N-terminal	N-terminal	N-terminal	N-terminal	N-terminal	internal
24	too low	0	0	21/23R	10/12r	4/7r	0	0	5/10r
25				(M)IIYRDLIS(S?(H)DEMF(G)/(S)DIYKI(R)(E)I	ATFPMQX(S)(A)LXK	A/STF(D)/(P)(R)X(D)			XXEL(G)I/(S)XLXE
26				genbank/mipsx	genbank/mipsx	genbank/mipsx			genbank/mipsx
27				#S00775(mouse)/B60756	#H30463/E35863	new			homology;E33978
28				yes	yes	no			no
29				15	any length				
30				16	0				
31									

	AX	AY	AZ	BA	BB	BC
1						
2						
3	800	314	L19	L22	L23	L29
4	367yy	367yy	p389y	p389y	p389y	p389y
5	np28	L10	L19	L22	L23	L29
6	human	human	human	human	human	human
7	PBMC	PBMC	TK6	TK6	TK6	TK6
8						
9	monomorphic			variable		
10	-			mutation		
11	-		-			
12	28	49		38	43	5J
13	130					
14						
15	-	-	-	38	-	-
16	-	-	-	513	-	-
17						
18						
19	Jurkat-tatIII	Jurkat	Jurkat	Jurkat	Jurkat	Jurkat
20	ne/rotophor	cyto	cyto	cyto	cyto	cyto
21	mp	mp	mp	mp	mp	mp
22	cb gel	sgf	sgf	sgf	sgf	sgf
23	internal	N-terminal	N-terminal	N-terminal	N-terminal	N-terminal
24	8/8r	0	6/6r	0	0	0
25	SAPGGGSK		X(G)EYYL			
26	genbank/mipsx		genbank/mipsx			
27	homology;E33978		new			
28	.io		no			
29						
30						
31						

END

DATE FILMED

12 / 27 / 90