

(19) BUNDESREPUBLIK DEUTSCHLAND

PATENTSCHRIFT



(12) Ausschließungspatent

(11) **DD 297 840 A5**

Erteilt gemäß § 17 Absatz 1
Patentgesetz der DDR
vom 27. 10. 1983
in Übereinstimmung mit den entsprechenden
Festlegungen im Einigungsvertrag

5(51) C 12 N 15/12
C 12 P 21/00
C 12 Q 1/70

DEUTSCHES PATENTAMT

In der vom Anmelder eingereichten Fassung veröffentlicht

(21)	DD C 12 N / 331 993 0	(22)	22.08.89	(44)	23.01.92
(31)	88113793.9	(32)	24.08.88	(33)	GB
	88119602.6		24.11.88		

(71)	siehe (73)
(72)	Tadatsugu, Taniguchi, Dr.; Takashi, Fujita, Dr., JP
(73)	Dr. Tadatsugu Taniguchi, Ibaraki-shi, Osaka 567, JP
(74)	Boehringer Ingelheim, Zentrale GmbH, ZA Patente, W - 6507 Ingelheim am Rhein, DE

(54) **Verfahren zur Herstellung eines Vektors mit einer DNA-Sequenz, die ein Protein mit Wirkung einer Interferon-Regulatorsequenz (IRF-1) codiert und dessen Verwendung**

(55) Vektorherstellung; DNA-Sequenz codiert Protein mit Interferon-Regulatorsequenz (IRF-1)-Wirkung

(57) Die Erfindung betrifft ein Verfahren zur Herstellung eines Vektors mit einer DNA-Sequenz, die ein Protein codiert, das die Wirkung einer Interferon-Regulatorsequenz (IRF-1) hat, gekennzeichnet dadurch, daß es die Spaltung eines Vektors mit einem geeigneten Restriktionsenzym und die Bindung einer DNA-Sequenz, die das genannte Protein codiert, an die Spaltstelle umfaßt.

ISSN 0433-6461

51 Seiten

Patentansprüche:

1. Verfahren zur Herstellung eines Vektors mit einer DNA-Sequenz, die ein Protein codiert, das die Wirkung einer Interferon-Regulatorsequenz (IRF-1) hat, **gekennzeichnet dadurch**, daß es die Spaltung eines Vektors mit einem geeigneten Restriktionsenzym und die Bindung einer DNA-Sequenz, die das genannte Protein codiert, an die Spaltstelle umfaßt.
2. Verfahren gemäß Anspruch 1, **gekennzeichnet dadurch**, daß die genannte DNA-Sequenz ein menschliches oder Mäuse-IRF-1 codiert.
3. Verfahren gemäß den Ansprüchen 1 oder 2, **gekennzeichnet dadurch**, daß die DNA-Sequenz ein Protein codiert, das zur Bindung an die sich wiederholende Oligomersequenz AAGTGA (CI-Sequenz) und das „Stromaufwärts-Regulator-Element“ des menschlichen IFN-β-Gens fähig ist.
4. Verfahren gemäß einem der Ansprüche 1–3, **gekennzeichnet dadurch**, daß die DNA-Sequenz mit der Promotor-Region an eine Spaltstelle in dem Vektor gebunden ist, wodurch sie in dem resultierenden Vektor operabel an die DNA-Sequenz, die das IRF-1-aktive Protein codiert, gebunden ist.
5. Verfahren gemäß einem der vorstehenden Ansprüche, **gekennzeichnet dadurch**, daß der Promotor ein konstitutiver Promotor oder ein induzierbarer Promotor ist.
6. Verfahren gemäß einem der vorstehenden Ansprüche, **gekennzeichnet dadurch**, daß worin die genannte DNA-Sequenz, die ein IRF-1-aktives Protein codiert, eine Sequenz der folgenden Formel oder eine Degenerat-Variante davon ist.

```

      10      20      30      40      50
ATGCCCATCACTTGGATGCGCATGAGACCCTGGCTAGAGATGCAGATTAA

      60      70      80      90     100
TTCCAACCAATCCCGGGGCTCATCTGGATTAATAAAGAGGAGATGATCT

      110     120     130     140     150
TGGAGATCCCATGGAAGCATGCTGCCAAGCATGGCTGGGACATCAACAAG

      160     170     180     190     200
GATGCCTGTTTGTTCGGAGCTGGGCCATTACACAGGCCGATACAAAGC

      210     220     230     240     250
AGGGGAAAAGGAGCCAGATCCCAAGACGTGGAAGGCCAACTTTCGCTGTG

      260     270     280     290     300
CCATGAACTCCCTGCCAGATATCGAGGAGGTGAAAGACCAGAGCAGGAAC

      310     320     340     340     350
AAGGGCAGCTCAGCTGTGCGAGTGTACCGGATGCTTCCACCTCTCACCAA

      360     370     380     390     400
GAACCAGAGAAAAGAAAGAAAGTCAAGTCCAGCCGAGATGCTAAGAGCA

      410     420     430     440     450
AGGCCAAGAGGAAGTCATGTGGGGATTCCAGCCCTGATACCTTCTCTGAT

      460     470     480     490     500
GGACTCAGCAGCTCCACTCTGCCTGATGACCACAGCAGCTACACAGTTCC

      510     520     530     540     550
AGGCTACATGCAGGACTTGGAGGTGGAGCAGGCCCTGACTCCAGCACTGT

      560     570     580     590     600
CGCCATGTGCTGTCAGCAGCACTCTCCCGACTGGCACATCCCAGTGGAA

      610     620     630     640     650
GTTGTGCCGGACAGCACCAGTGATCTGTACAACCTCCAGGTGTCACCCAT

```

```

        660      670      680      690      700
GCCCTCCATCTCTGAAGCTACAACAGATGAGGATGAGGAAGGGAAATTAC

        710      720      730      740      750
CTGAGGACATCATGAAGCTCTTGGAGCAGTCGGAGTGGCAGCCAACAAAC

        760      770      780      790      800
GTGGATGGGAAGGGGTACCTACTCAATGAACCTGGAGTCCAGCCCACCTC

        810      820      830      840      850
TGTCTATGGAGACTTTAGCTGTAAGGAGGAGCCAGAAATTGACAGCCCAG

        860      870      880      890      900
GGGGGGATATTGGGCTGAGTCTACAGCGTGTCTTCACAGATCTGAAGAAC

        910      920      930      940      950
ATGGATGCCACCTGGCTGGACAGCCTGCTGACCCCAGTCCGGTTGCCCTC

        960      970
CATCCAGGCCATTCCCTGTGCACCG

```

7. Verfahren gemäß Anspruch 6, gekennzeichnet dadurch, daß die genannte DNA-Sequenz, die ein IRF-1-aktives Protein codiert, eine Sequenz der folgenden Formel oder eine Degenerat-Variante davon ist.

```

CGAGCCCCGCCGAACCGAGGCCACCCGGAGCCGTGCCCAGTCCACGC
CGGCCGTGCCCGGGCGGCTTAAGAACCAGGCAACCACTGCCTTCTTCCCT
CTTCCACTCGGAGTCGCGCTTCGCGCGCCCTCACTGCAGCCCTGCGTCG
CCGGGACCCCTCGCGCGGACCCAGCCGAATCGCTCCTGCAGCAGAGCCAAC

        10      20      30      40      50
ATGCCCATCACTTGGATGCGCATGAGACCCCTGGCTAGAGATGCAGATTAA

        60      70      80      90      100
TTCCAACCAAAATCCCGGGGCTCATCTGGATTAAATAAGAGGAGATGATCT

        110     120     130     140     150
TGGAGATCCATGGAAGCATGCTGCCAAGCATGGCTGGGACATCAACAAG

        160     170     180     190     200
GATGCTGTGTTTGTTCGGAGCTGGGCCATTACACAGGCCGATACAAAGC

        210     220     230     240     250
AGGGGAAAAGGAGCCAGATCCCAAGACGTGGAAGGCCAACTTTCGCTGTG

        260     270     280     290     300
CCATGAACCTCCCTGCCAGATATCGAGGAGGTGAAAGACCAGAGCAGGAAC

        310     320     340     340     350
AAGGGCAGCTCAGCTGTGCGAGTGTACCGGATGCTTCCACCTCTCACCAA

        360     370     380     390     400
GAACCAGAGAAAAGAAAGAAAGTGAAGTCCAGCCGAGATGCTAAGAGCA

        410     420     430     440     450
AGGCCAAGAGGAAGTCATGTGGGGATTCCAGCCCTGATACCTTCTCTGAT

        460     470     480     490     500
GGA CTCAGCAGCTCCACTCTGCCTGATGACCACAGCAGCTACACAGTTCC

```

510 520 530 540 550
 AGGCTACATGCAGGACTTGGAGGTGGAGCAGGCCCTGACTCCAGCACTGT
 560 570 580 590 600
 CGCCATGTGCTGTCAGCAGCACTCTCCCCGACTGGCACATCCCAGTGGAA
 610 620 630 640 650
 GTTGTGCCGGACAGCACCAGTGATCTGTACAACCTCCAGGTGTCACCCAT
 660 670 680 690 700
 GGCCTCCATCTCTGAAGCTACAACAGATGAGGATGAGGAAGGGAAATTAC
 710 720 730 740 750
 CTGAGGACATCATGAAGCTCTTGGAGCAGTCGGAGTGGCAGCCAACAAAC
 760 770 780 790 800
 GTGGATGGGAAGGGGTACCTACTCAATGAACCTGGAGTCCAGCCCACCTC
 810 820 830 840 850
 TGTCTATGGAGACTTTAGCTGTAAGGAGGAGCCAGAAATTGACAGCCCAG
 860 870 880 890 900
 GGGGGGATATTGGGCTGAGTCTACAGCGTGTCTTCACAGATCTGAAGAAC
 910 920 930 940 950
 ATGGATGCCACCTGGCTGGACAGCCTGCTGACCCCAGTCCGGTTGCCCTC
 960 970
 CATCCAGGCCATTCCCTGTGCACCGTAGCAGGGCCCCTGGGCCCCCTCTTA
 TTCCTCTAGGCAAGCAGGACCTGGCATCATGGTGGATATGGTGCAGAGAA
 GCTGGACTTCTGTGGGCCCCCTCAACAGCCAAGTGTGACCCCACTGCCAAG
 TGGGGATGGGCCTCCCTCCTTGGGTCATTGACCTCTCAGGGCCTGGCAGG
 CCAGTGTCTGGGTTTTTCTTGTGGTGTAAGCTGGCCCTGCCTCCTGGGA
 AGATGAGGTTCTGAGACCAGTGTATCAGGTCAGGGACTTGGACAGGAGTC
 AGTGTCTGGCTTTTTCTCTGAGCCCAGCTGCCTGGAGAGGGTCTCGCTG
 TCACTGGCTGGCTCCTAGGGGAACAGACCAGTGACCCCAGAAAAGCATAA
 CACCAATCCCAGGGCTGGCTCTGCTACTAAGAGAAAATTGCACTAAATGAA
 TCTCGTTCCAAAGAACTACCCCTTTTCAGCTGAGCCCTGGGGACTGTTCC
 AAAGCCAGTGAATGTGAAGGAAAGTGGGGTCCTTCGGGGCAATGCTCCCT
 CAGCCTCAGAGGAGCTCTACCCTGCTCCCTGCTTTGGCTGAGGGGCTTGG
 GAAAAAACTTGGCACTTTTTCTGTGTGGATCTTGCCACATTTCTGATCAG
 AGGTGTACACTAACATTTCCCCGAGCTCTTGGCCTTTGCATTTATTTAT
 ACAGTGCCTTGCTCGGGGGCCACCACCCCTCAAGCCCCAGCAGCCCTCA
 ACAGGGCCAGGGAGGGAAGTGTGAGCGCCTTGGTATGACTTAAAATTGGA
 AATGTCATCTAACCATTAAAGTCATGTGTGAACACATAAGGACGTGTGTAA
 ATATGTACATTTGTCTTTTTATAAAAAGTAAAATTGTT

8. Verfahren gemäß einem der Ansprüche 1-5, **gekennzeichnet dadurch**, daß die DNA-Sequenz, die ein IRF-1-aktives Protein codiert, eine Sequenz der folgenden Formel oder eine Degenerat-Variante davon ist.

```

ATG CCA ATC ACT CGA ATG CGG ATG AGA CCC TGG CTA GAG ATG CAG ATT
AAT TCC AAC CAA ATC CCA GGG CTG ATC TGG ATC AAT AAA GAA GAG ATG
ATC TTC CAG ATT CCA TGG AAG CAC GCT GCT AGG CAC GGC TGG GAC ATC
AAC AAG GAT GCC TGT CTG TTC CGG AGC TGG GCC ATT CAC ACA GGC CGA
TAC AAA GCA GGA GAA AAA GAG CCA GAT CCC AAG ACA TGG AAG GCA AAC
TTC CGT TGT GCC ATG AAC TCC CTG CCA GAC ATC GAG GAA GTG AAG GAT
CAG AGT AGG AAC AAG GGC AGC TCTGCT GTG CGG CTG TAC CGG ATG CTG
CCA CCC CTC ACC AGG AAC CAG AGG AAA GAG AGA AAG TCC AAG TCC AGC
CGA GAC ACT AAG AGC AAA ACC AAG AGG AAG CTG TGT GGA GAT GTT AGC
CCG GAC ACT TTC TCT GAT GGA CTC AGC AGC TCT ACC CTA CCT GAT GAC
CAC AGC AGT TAC ACC ACT CAG GGC TAC CTG GGT CAG GAC TTG GAT ATG
GAA AGG GAC ATA ACT CCA GCA CTG TCA CCG TGT GTC GTC AGC AGC AGT
CTC TCT GAG TGG CAT ATG CAG ATG GAC ATT ATA CCA GAT AGC ACC ACT
GAT CTG TAT AAC CTA CAG GTG TCA CCC ATG CCT TCC ACC TCC GAA GCC
GCA ACA GAC GAG GAT GAG GAA GGG AAG ATA GCC GAA GAC CTT ATG AAG
CTC TTT GAA CAG TCT GAG TGG CAG CCG ACA CAC ATC GAT GGC AAG GGA
TAC TTG CTC AAT GAG CCA GGG ACC CAG CTC TCT TCT GTC TAT GGA GAC
TTC AGC TGC AAA GAG GAA CCA GAG ATT GAC AGC CCT CGA GGG GAC ATT
GGG ATA GGC ATA CAA CAT GTC TTC ACG GAG ATG AAG AAT ATG GAC TCC
ATC ATG TGG ATG GAC AGC CTG CTG GGC AAC TCT GTG AGG CTG CCG CCC
TCT ATT CAG GCC ATT CCT TGT GCA CCA TAG

```

9. Verfahren gemäß Anspruch 8, **gekennzeichnet dadurch**, daß die genannte DNA-Sequenz, die ein IRF-1-aktives Protein codiert, eine Sequenz der folgenden Formel oder eine Degenerat-Variante davon ist.

```

1  GGACGTGCTTTTACAGTCTAAGCCGAACCGAACCGAACCGAACCGAACCGAACCGGGCC
60  GAGTTGCGCCGAGGTCAGCCGAGGTGGCCAGAGGACCCAGCATCTCGGGCATCTTTCC
119,CTTCGTGCGCGCATCGCGTACCTACACCGCAACTCCGTGCCTCGCTCTCCGGCACCCTC
178 TGGGAATCGCTCCTGCAGCAA   AGCCACC ATG CCA ATC ACT CGA ATG CGG
227 ATG AGA CCC TGG CTA GAG ATG CAG ATT AAT TCC AAC CAA ATC CCA
272 GGG CTG ATC TGG ATC AAT AAA GAA GAG ATG ATC TTC CAG ATT CCA
317 TGG AAG CAC GCT GCT AAG CAC GGC TGG GAC ATC AAC AAG GAT GCC
362 TGT CTG TTC CGG AGC TGG GCC ATT CAC ACA GGC CGA TAC AAA GCA
407 GGA GAA AAA GAG CCA GAT CCC AAG ACA TGG AAG GCA AAC TTC CGT
452 TGT GCC ATG AAC TCC CTG CCA GAC ATC GAG GAA GTG AAG GAT CAG
497 AGT AGG AAC AAG GGC AGC TCT GCT GTG CCG GTG TAC CGG ATG CTG
542 CCA CCC CTC ACC AGG AAC CAG AGG AAA GAG AGA AAG TCC AAG TCC
587 AGC CGA GAC ACT AAG AGC AAA ACC AAG AGG AAG CTG TGT GGA GAT

```

632 GTT AGC CCG GAC ACT TTC TCT GAT GGA CTC AGC AGC TCT ACC CTA
677 CCT GAT GAC CAC AGC AGT TAC ACC ACT CAG GGC TAC CTG GGT CAG
722 GAC TTG GAT ATG GAA AGG GAC ATA ACT CCA GCA CTG TCA CCG TGT
767 GTC GTC AGC AGC AGT CTC TCT GAG TGG CAT ATG CAG ATG GAC ATT
812 ATA CCA GAT AGC ACC ACT GAT CTG TAT AAC CTA CAG GTG TCA CCC
857 ATG CCT TCC ACC TCC GAA GCC GCA ACA GAC GAG GAT GAG GAA GGG
902 AAG ATA GCC GAA GAC CTT ATG AAG CTC TTT GAA CAG TCT GAG TGG
947 CAG CCG ACA CAC ATC GAT GCC AAG CGA TAC TTG CTC AAT GAG CCA
992 GGG ACC CAG CTC TCT TCT GTC TAT GGA GAC TTC AGC TGC AAA GAG
1037 GAA CCA GAG ATT GAC AGC CCT CGA GGG GAC ATT GGG ATA GGC ATA
1082 CAA CAT GTC TTC ACG GAG ATG AAG AAT ATG GAC TCC ATC ATG TGG
1127 ATG GAC AGC CTG CTG GGC AAC TCT GTG AGG CTG CCG CCC TCT ATT
1172 CAG GCC ATT CCT TGT GCA CCA TAG TTTGGGTCTCTGACCCGTTCTTGCCC
1222 TCCTGAGTGAGTTAGGCCTTGGCATCATGGTGGCTGTGATACAAAAAAGCTAGACTCC
1281 TGTGGGCCCCCTTGACACATGGCAAAGCATAGTCCCACTGCAAACAGGGGACATCCTCC
1340 TTGGGTGAGTGGGCTCTCAGGGCTTAGGAGGCAGAGTCTGAGTTTCTTGTGAGGTGAA
1399 GCTGGCCCTGACTCCTAGGAAGATGGATTGGGGGTCTGAGGTGTAAGGCAGAGGCCAT
1458 GGACAGGAGTCATCTTCTAGCTTTTTTAAAAGCCTTGTTCATAGAGAGGGTCTTATCGC
1517 TGGGCTGGCCCTGAGGGGAATAGACCAGCGCCACAGAAGAGCATAGCACTGGCCCTAG
1576 AGCTGGCTCTGTACTAGGAGACAATTGCACTAAATGAGTCCTATTCCTAAAGAACTGCT
1635 GCCCTTCCCAACCGAGCCCTGGGATGGTTCCCAAGCCAGTGAAATGTGAAGGGAAAAAA
1694 AATGGGGTCCCTGTGAAGGTGGCTCCCTTAGCCTCAGAGGGAATCTGCCTCACTACCTG
1753 CTCCAGCTGTGGGGCTCAGGAAAAAAATGGCACTTTCTCTGTGGACTTTGCCACATT
1812 TCTGATCAGAGGTGTACACTAACATTTCTCCCCAGTCTAGGCCTTTGCATTATTATTA
1871 TAGTGCCTTGCTGGTGCCTGCTGTCTCCTCAGGCCTTGGCAGTCCTCAGCAGGCCAG
1930 GGAAAAGGGGGTGTGAGCGCCTTGGCGTGACTCTTGACTATCTATTAGAAACGCCAC
1989 CTAAGTCTAAATGGTGTGTTGGTCATGTGGTGGACCTGTGTAATATGTATATTTGTCT
2048 TTTTATAAANATTTAAGTTGTTTACAAAAA 2082

10. Verfahren gemäß einem der Ansprüche 1-5, **gekennzeichnet dadurch**, daß die DNA-Sequenz, die ein IRF-1-aktives Protein codiert, eine Sequenz der folgenden Formel oder eine Degenerat-Variante davon ist.

10	20	30	40
ATGCCGGTGAACGGATGCGAATGCCCGCTGGCTGGAGGAGCAG			
50	60	70	80
ATAAATTCCAATACGATACCAGGGCTAAAGTGGCTGAACAAGGAG			
100	110	120	130
AAGAAGATTTTCCAGATCCCCTGGATGCATGCGGCTCGGCACGGA			

140 150 160 170 180
 TGGGACGTGGAAAAGGATGCTCCGCTCTTCAGAAACTGGGCGATC
 190 200 210 220
 CATACAGGAAAGCATCAACCAGGAATAGATAAACCCAGATCCAAAA
 230 240 250 260 270
 AGATGGAAAGCAAATTTTCGATGTGCCATGAATTCCCTGCCCGAC
 280 290 300 310
 ATTGAGGAAGTGAAGGACAGAAGCATAAAGAAAGGAAACAACGCC
 320 330 340 350 360
 TTCAGAGTCTACCGGATGCTGCCCTTATCCGAACGACCTTCCAAG
 370 380 390 400
 AAAGGAAAGAAACCAAAGACAGAAAAAGAAGAGAGAGTTAAGCAG
 410 420 430 440 450
 ATCAAGCAAGAACCAGTTGAGTCATCTTTGGGGCTTAGTAATGGA
 460 470 480 490
 GTAAGTGGCTTTTCTCCTGAGTATGCGGTCTGACTTCAGCTATA
 500 510 520 530 540
 AAAAATGAAGTGGATAGTACGGTGAACATCATAGTTGTAGGACAG
 550 560 570 580
 TCCCATCTGGACAGCAACATTTGAAGATCAAGAGATCGTCACTAAC
 590 600 610 620 630
 CCGCCAGACATCTGCCAGGTTGTAGAAGTGACCACTGAGAGTGAT
 640 650 660 670
 GACCAGCCAGTCAGCATGAGTGAGCTCTACCCTCTACAGATTTCT
 680 690 700 710 720
 CCTGTGTCTTCCTACGCAGAAAGCGAAACTACCGACAGTGTGGCC
 730 740 750 760
 AGTGATGAAGAGAACGCAGAGGGGAGACCACACTGGAGGAAGAGG
 770 780 790 800 810
 AGCATCGAAGGCAAGCAGTACCTCAGCAACATGGGGACACGGAAC
 820 830 840 850
 ACCTATCTGCTGCCCAGCATGGCGACCTTTGTCACCTCCAACAAG
 860 870 880 890 900
 CCAGATCTGCAGGTCACCATCAAAGAGGATAGCTCTCCGATGCCT
 910 920 930 940
 TACAACAGCTCCTGGCCCCCATTTACAGACCTTCCCCTTCCTGCC
 950 960 970 980 990
 CCAGTGACCCCCACGCCAGCAGCAGTGGGCCAGACCGGGAGACC
 1000 1010 1020 1030
 CGGGCCAGTGTTCATCAAGAAGACATCTGATATCACCCAGGCCCGT
 1040
 GTCAAGAGCTGT

11. Verfahren gemäß Anspruch 10, gekennzeichnet dadurch, daß die genannte DNA-Sequenz, die ein IRE-1-aktives Protein codiert, eine Sequenz der folgenden Formel oder eine Degenerat-Variante davon ist.

TCTCAGGCAAGCCGGGA

CTAACTTTTAGTTTGGCTCCTGCGATTATTCAACTGACGGGCTTT

CATTTCCATTTTACAGACCCTAACAACACTCACACCTTGCGGGAT

TGTATTGGTAGCGTGGAAAAAAAAAAGCACATTGAGAGGGTACC

10 20 30 40
ATGCCGGTGGAAACGGATGCGAATGCGCCCGTGGCTGGAGGAGCAG

50 60 70 80 90
ATAAATTCCAATACGATACCAGGGCTAAAGTGGCTGAACAAGGAG

100 110 120 130
AAGAAGATTTTCCAGATCCCCTGGATGCATGCGGCTCGGCACGGA

140 150 160 170 180
TGGGACGTGGAAAAGGATGCTCCGCTCTTCAGAACTGGCGATC

190 200 210 220
CATACAGGAAAGCATCAACCAGGAATAGATAAACCAGATCCAAA

230 240 250 260 270
ACATGGAAAGCAAATTTTCGATGTGCCATGAATTCCTGCCCGAC

280 290 300 310
ATTGAGGAAGTGAAGGACAGAAGCATAAAGAAAGGAAACAACGCC

320 330 340 350 360
TTCAGAGTCTACCGGATGCTGCCCTTATCCGAACGACCTTCCAAG

370 380 390 400
AAAGGAAAGAAACCAAGACAGAAAAAGAAGAGAGAGTTAAGCAC

410 420 430 440 450
ATCAAGCAAGAACCAGTTGAGTCATCTTTGGGGCTTAGTAATGGA

460 470 480 490
GTAAGTGGCTTTTCTCCTGAGTATGCGGCTCCTGACTTCAGCTATA

500 510 520 530 540
AAAAATGAAGTGGATAGTACGGTGAACATCATAGTTGTAGGACAG

550 560 570 580
TCCCATCTGGACAGCAACATTGAAGATCAAGAGATCGTCACTAAC

590 600 610 620 630
CCGCCAGACATCTGCCAGGTTGTAGAAGTGACCACTGAGAGTGAT

640 650 660 670
GACCAGCCAGTCAGCATGAGTGAGCTCTACCCTCTACAGATTTCT

680 690 700 710 720
CCTGTGTCTTCCCTACGCAGAAAGCGAAACTACCGACAGTGTGGCC

730 740 750 760
AGTGATGAAGAGAACGCAGAGGGGAGACCACACTGGAGGAAGAGG

770 780 790 800 810
AGCATCGAAGGCAAGCAGTACCTCAGCAACATGGGGACACGGAAC

820 830 840 850
ACCTATCTGCTGCCCAGCATGGCGACCTTTGTACCTCCAACAAG

860 870 880 890 900
CCAGATCTGCAGGTCACCATCAAAGAGGATAGCTGTCCGATGCCT

910 920 930 940
TACAACAGCTCCTGGCCCCCATTTACAGACCTTCCCCTTCTGCC

950 960 970 980 990
CCAGTGACCCCCACGCCCAGCAGCAGTCGGCCAGACCGGGAGACC

1000 1010 1020 1030
CGGGCCAGTGTCAATCAAGAAGACATCTGATATCAGCCAGGCCCGT

1040
GTCAAGAGCTGTTAAGCCTTTGACTCTCCCTGGTGGTTGTTGGGA

TTTCTTAGCTTTGTGTTGTTCTTTGTTTGTATTATATTATTTTTT

TTCTCTATGATACCTATCTTAGACACATCTAAGGGAGAAAGCCTT

GACGATAGATTATTGATTGCTGTGTCCAACCTCCAGAGCTGGAGCT

TCTTCTTAACTCAGGACTCCAGCCCCCCCCCCCCCTCGGTAGATG
CGTATCTCTAGAACCTGCTGGATCTGCCAGGGCTACTCCCTCAAG
TTCAAGGACCAACAGCCACACGGGCAGTGGAGGTGCTGCGTTGCC
TACGGTCAAGGCCAGCATGGTGGAGTGGATGCCTCAGAACGGAGG
AGAAAATGTGAACTAGCTGGAATTTTTTTATTCTTGTGAATATGT
ACATAGGGCAGTACGAGCAATGTCGCGGGCTGCTTCTGCACCTTA
TCTTGAAGCACTTACAATAGGCCTTCTTGTAATCTTGCTCTCCTT
CACAGCACACTCGGCGACCCCTTCTGTGTCCACTACCCCACTACC
CACCCCTCCCTCCTCAACCCCTCCATCCCGGTCTCTATGCGCCC
CTTCCCCCAACCAATCCCATCACAACCTCTTACCTATCCTTTCC
CTCCCAACCCCTTCTATCCCAGCCCACCACCTACCCCACTCCTCC
CCAACTCCTCCATTCTAGCCCATTACCCACGCCTCTCTCCTCAGC
CCAGCCTACCCCATCCCACCCTGTTCTTTCTCCAGTTTCCTCT
CCTCAAAGGCAAGGCTCTACATCTTGGAGGAGGAGGAGGAGAAGA
AAATGAGTTTCTTCACCGCTGTCCCATTTTAAGACTGCTTGAATA

ATAAAAAAAAAATCTTTCTAATCTGCTATGCTTGAATGGCACGCGG

TACAAAGGAAAACGTGCATGGAAATATTATGCAAATCCCAGATC

TGAAGACGGAAAATACTCTAATTCTAACCAGAGCAAGCTTTTTTA

TTTTTTTATACAAGGGGAATATTTTATTCAAGGTAAAAAATTCT

AAATAAAATATAATTGTTTTTTATCTTTTCTACAGCAAATTTATA

ATTTTAAGATTCCTTTTCTGTTTCATCAGCAGTTGTTATTACATC

CCTTGTGGCACATTTTTTTTTTAATTTTGTAAGGTGAAAAAAA

ACTTTTATGAGCTCATGTAGCAATCAAATTATCCTGTGGATTGAT

AATAAATGAATATGGTATATAGTTAAAGATTTTAAAAA

12. Verfahren gemäß einem der vorstehenden Ansprüche, **gekennzeichnet dadurch**, daß die DNA-Sequenz, die ein IRF-1-aktives Molekül codiert, eine Sequenz ist, die sich zu einer DNA-Sequenz hybridisiert, wie sie in einem der Ansprüche 6-11 definiert ist.
13. Verfahren gemäß einem der Ansprüche 1-12, **gekennzeichnet dadurch**, daß eine Promotor- und Regulatorsequenz, enthalten in der nachstehenden Formel, an eine Spaltstelle gebunden ist, wodurch in dem resultierenden Vektor die genannte Promotor- und Regulatorsequenz in operablen Assoziationen mit der DNA-Sequenz, die die IRF-1-aktiven Proteine codiert, vorhanden ist.

Pst I

CTGCAGAAAGAGGGGGACGGTCTCGGCTTTCCAAGACAGGCAAGGGGG

-299 GC Box 1 GC Box 2

CAGGGGAGTGGAGTGGAGCAAGGGGCGGGCCCGGGTAGCCCCGGGGCGGTGGCGCGG

-251

GCCCGAGGGGGTGGGGAGCACAGCTGCCTTGTACTTCCCTTCGCGGCTTAGCTCTAG

-193

AACAGCCTGATTTCCTCGAAATGATGAGGCCGAGTGGCCCAATGGGCGCGCAGGAGCG

-135

CAAT Box

Minor Cap site

GCGCGGCGGGGGCGTGGCCGAGTCCGGGCGGGGAATCCCGCTAAGTGTTTAGATTTC

-77

-21

Major Cap site

plRF-L

TTCGCGGGCGCGCGGACTCGCCAGTGGCGACCACTCCTTCGTCGAGGTAGGACGTGCT

-19 +1

```

      TTCACAGTCTAAGCCGAACCGAACCGAACCGAACCGAACCGGGCCGAGTTGCG
+39
      CCGAGGTCAGCCGAGGTGGCCAGAGGACCCCAGCATCTCGGGCATCTTTCGGCTTCGTG
+97

      CGCGCATCGCGTACCTACACCGCAACTCCGTGCGCTCTCCGGCACCCCTCTGCGAA
+155

      Pst I
      TCGCTCCTGCAG
+213      +225

```

14. Verfahren gemäß einem der Ansprüche 1–13, **gekennzeichnet dadurch**, daß die DNA-Sequenz, die ein strukturelles Gen für ein gewünschtes, pharmazeutisch wirksames Protein umfaßt, an eine Spaltstelle des Vektors gebunden ist, wodurch in dem resultierenden Vektor das genannte Gen unter der Kontrolle des IRF-1-Proteins durch die entsprechende DNA-Sequenz codiert wurde.
15. Verfahren gemäß den Ansprüchen 3 und 14, **gekennzeichnet dadurch**, daß in dem resultierenden Vektor das Gen, das das IRF-1-aktive Protein codiert, unter Kontrolle des konstitutiven oder des induziblen Promotors ist.
16. Verfahren gemäß Anspruch 14 oder 15, **gekennzeichnet dadurch**, daß das Gen für das gewünschte, pharmazeutisch wirksame Protein eine IRF-1-Bindungsstelle umfaßt, die eine Vielzahl von sich wiederholenden AAGTGA-Sequenzen enthält.
17. Verfahren gemäß einem der Ansprüche 14–16, **gekennzeichnet dadurch**, daß das gewünschte, pharmazeutisch wirksame Protein ein Cytokin- oder Plasmogen-Aktivator ist.
18. Verfahren gemäß einem der vorstehenden Ansprüche, **gekennzeichnet dadurch**, daß der Vektor ein Plasmid ist.
19. Verwendung eines Vektors, wie er in einem der vorstehenden Ansprüche definiert ist, **gekennzeichnet dadurch**, daß er zur Transformation einer Wirtszelle eingesetzt wird.
20. Verwendung gemäß Anspruch 19, **gekennzeichnet dadurch**, daß die Wirtszelle mit einem Plasmid transformiert wird, das eine DNA-Sequenz enthält, die ein Protein mit der Aktivität eines IRF-1 codiert, und durch ein separates DNA-Molekül mit einem Gen, das ein gewünschtes pharmazeutisch wirksames Protein unter der Kontrolle des IRF-1-Protein-Codes für das genannte Plasmid codiert.
21. Verwendung gemäß Anspruch 20, **gekennzeichnet dadurch**, daß das Gen, das das IRF-1-aktive Molekül codiert, unter Kontrolle eines konstitutiven oder eines induziblen Promotors ist.
22. Verwendung gemäß Anspruch 20 oder 21, **gekennzeichnet dadurch**, daß das Gen für das gewünschte, pharmazeutisch wirksame Protein eine Bindungsstelle für das IRF-1-aktive Protein umfaßt, die eine Vielzahl von sich wiederholenden AAGTGA-Sequenzen einschließt.
23. Verwendung gemäß einem der Ansprüche 19–22, **gekennzeichnet dadurch**, daß die Wirtszelle eine Bakterien-, Hefe- oder Säugetierzelle ist.
24. Verfahren gemäß Anspruch 1, **gekennzeichnet dadurch**, daß der Vektor ein Plasmid ist, und das resultierende Plasmid ist
 pHIRF31
 pIRF-L oder
 pIRF2–5.

Hierzu 16 Seiten Zeichnungen

Anwendungsgebiet der Erfindung

Die vorliegende Erfindung betrifft allgemein die Regulation der Genexpression. Insbesondere betrifft die Herstellung eines rekombinanten DNA-Molekül, das für ein Protein codiert, das die Aktivität eines Interferon-Regulationsfaktors-1 (IRF-1) besitzt; sowie die Verwendung derartiger DNA Moleküle zur Transformation von Wirtszellen, die auch durch DNA-Moleküle transformiert werden, die für ein gewünschtes Protein unter der Kontrolle dieses IRF-1-aktiven Proteins codieren.

Charakteristik des bekannten Standes der Technik

Die Transkription von Genen in Säugetierzellen wird durch komplexe Mechanismen reguliert, in denen die Wechselwirkungen der Regulations-DNA-Sequenzen mit den trans-wirkenden DNA-Bindungsproteinen eine zentrale Rolle spielen. Im Zusammenhang mit der Regulation der Transkription stellen die für Interferone (IFN) codierenden Gene ein Merkmal dar, das vielen Cytokingenen eigen ist: die Transkription dieser Gene wird in einem transienten Zustand nach verschiedenen extrazellulären Signalen induziert. Es ist bereits gut dokumentiert worden, daß die Transkription der Gene für IFN- α und IFN- β in einer Vielzahl von Zellen durch Viren wirksam induziert wird, während die Transkription des für IFN- γ codierenden Gens in T-Lymphozyten (T-Zellen) nach mitogener Stimulierung induziert wird (siehe Weissmann und Weber, 1986; Taniguchi, 1988). IFN- β , ein Cytokin, das ursprünglich wegen seiner potenten antiviralen Wirksamkeit identifiziert wurde, scheint auch eine wichtige Rolle bei der Steuerung des Zellwachstums und der Zelldifferenzierung zu spielen. In diesem Zusammenhang scheinen neben Viren und poly(rI):poly(rC), die die bekannten Induktoren des IFN- β -Gens sind, auch viele Cytokine wie der koloniestimulierende Faktor-1 (CSF-1) (Moore u. a., 1984; Warren und Ralf, 1986; Resnitzky u. a., 1986), der Tumornekrose-Faktor (TNF) (Onozaki u. a., 1988); der plättchenstammige Wachstumsfaktor (PDGF) (Zullo u. a., 1985) und die IFN (Kohase u. a., 1987) IFN- β in bestimmten Zellen zu induzieren, was vermuten läßt, daß sie ähnliche oder identische Signale in die Zielzellen übertragen können.

Es wurde auch gezeigt, daß die IFN- β -Geninduktion durch Viren und poly(rI):poly(rC) auf der Transkriptionsebene stattfindet (Raji und Pitham 1983; Ohno und Taniguchi, 1983; Dinter u. a., 1983; Zinn u. a., 1983). So wurden cis-wirkende DNA-Sequenzen, die als induzierbare Verstärker fungieren, in der 5'-Flankenregion des Human-IFN- β -Gens identifiziert (Fujita u. a., 1985, 1987; Goodbourn u. a., 1985; Dinter und Hauser, 1987). Die induzierbare Verstärkerregion (d. h. -65 bis -105 in bezug auf die CAP-Stelle) enthält sich wiederholende Hexanucleotideinheiten, von denen einige tatsächlich in der induzierten Aktivierung der Transkription bei der Multimerisierung wirken (Fujita u. a., 1987). In Säugetierzellen wie z. B. Maus-L 929-Zellen und Human-Zellen wurden ein Faktor, IRF-1, der spezifisch an die IFN- β -Regulationssequenzen bindet, sowie die funktionellen, sich wiederholenden Hexanucleotidsequenzen, (AAGTGA)₄, identifiziert. Es wurde festgestellt, daß der IRF-1 eine wesentliche Rolle bei der virusinduzierten Aktivierung der IFN- β -Gentranskription durch Wechselwirkung mit den identifizierten cis-Elementen spielt.

Ziel der Erfindung

Durch das erfindungsgemäße Verfahren wird ein pharmazeutisch wertvoller Vektor mit einer DNA-Sequenz, die ein Protein codiert, das die Wirkung einer Interferon – Regulatorsequenz (IRF-1) hat, zur Verfügung gestellt.

Darlegung des Wesens der Erfindung

Der vorliegenden Erfindung liegt die Aufgabe zugrunde, ein Verfahren zur Herstellung eines rekombinanten DNA-Moleküls zur Verfügung zu stellen, das für ein Protein codiert, das als ein Interferon-Regulationsfaktor-1 (IRF-1) wirkt.

Entsprechend dem breiten Aspekt der Erfindung wird ein Verfahren zur Herstellung einer rekombinanten DNA-Moleküls zur Verfügung gestellt, das für ein Protein codiert, das als ein Interferon-Regulationsfaktor-1 (IRF-1) wirkt.

Vorzugsweise codiert das DNA-Molekül für oder ist hybridisierbar an das DNA-Molekül, das für Human-IRF-1 oder Maus-IRF-1 codiert.

Das DNA-Molekül ist vorzugsweise eines, das für ein Protein codiert, das an die wiederholte Oligomersequenz AAGTGA und die „upstream“-Regulationselemente des Human-IFN- β -Gens bindet.

Das DNA-Molekül kann eine Promotorregion enthalten, die konstitutiv oder induzierbar ist, beispielsweise virusinduzierbar z. B. durch das Newcastle-Virus oder durch mitogene Stimulierung, z. B. mittels Concanavalin A, induzierbar.

Ein bevorzugtes DNA-Molekül (das für Human-IRF-1 codiert) ist durch ein Strukturgemisch gekennzeichnet, das die folgende Formel I hat:

Formel I

```

      10      20      30      40      50
ATGCCCATCACTTGGATGCGCATGAGACCCTGGCTAGAGATGCAGATTAA

      60      70      80      90     100
TTCCAACCAAATCCCGGGGCTCATCTGGATTAATAAAGAGGAGATGATCT

     110     120     130     140     150
TGGAGATCCCATGGAAGCATGCTGCCAAGCATGGCTGGGACATCAACAAG

     160     170     180     190     200
GATGCCTGTTTGTTCGGAGCTGGGCCATTACACAGGCCGATACAAAGC

     210     220     230     240     250
AGGGGAAAAGGAGCCAGATCCCAAGACGTGGAAGGCCAACTTTCGCTGTG

     260     270     280     290     300
CCATGAACCTCCCTGCCAGATATCGAGGAGGTGAAAGACCAGAGCAGGAAC

```

```

      310      320      340      340      350
AAGGGCAGCTCAGCTGTGCGAGTGTACCGGATGCTTCCACCTCTCACCAA

      360      370      380      390      400
GAACCAGAGAAAAGAAAGAAAGTCTGAAGTCCAGCCGAGATGCTAAGAGCA

      410      420      430      440      450
AGGCCAAGAGGAAGTCATGTGGGGATTCCAGCCCTGATACCTTCTCTCTAT

      460      470      480      490      500
GGACTCAGCAGCTCCACTCTGCCTGATGACCACAGCAGCTACACAGTTCC

      510      520      530      540      550
AGGCTACATGCAGGACTTGGAGGTGGAGCAGGCCCTGACTCCAGCACTGT

      560      570      580      590      600
CGCCATGTGCTGTCAGCAGCACTCTCCCCGACTGGCACATCCCAGTGGAA

      610      620      630      640      650
GTTGTGCCGGACAGCACCAGTGATCTGTACAACCTCCAGGTGTCACCCAT

      660      670      680      690      700
GCCCTCCATCTCTGAAGCTACAACAGATGAGGATGAGGAAGGGAAATTAC

      710      720      730      740      750
CTGAGGACATCATGAAGCTCTTGGAGCAGTCGGAGTGGCAGCCAACAAAC

      760      770      780      790      800
GTGGATGGGAAGGGGTACCTACTCAATGAACCTGGAGTCCAGCCCACCTC

      810      820      830      840      850
TGTCTATGGAGACTTTAGCTGTAAGGAGGAGCCAGAAATTGACAGCCCAG

      860      870      880      890      900
GGGGGGATATTGGGCTGAGTCTACAGCGTGTCTTCACAGATCTGAAGAAC

      910      920      930      940      950
ATGGATGCCACCTGGCTGGACAGCCTGCTGACCCAGTCCGGTTGCCCTC

      960      970
CATCCAGGCCATTCCCTGTGCACCG

```

oder eine degenerierte Variante davon.

Das DNA-Molekül kann beispielsweise durch ein Strukturgen gekennzeichnet sein, das die im vorangegangenen erläuterte Formel sowie „upstream“- und „downstream“-Flankensequenzen wie in der folgenden Formel II hat:

Formel II

CGAGCCCCGCCGAACCGAGGCCACCCGGAGCCGTGCCCAGTCCACGC
 CGGCCGTGCCCCGGCGGCCCTTAAGAACCAGGCAACCACTGCCTTCTTCCCT
 CTTCCACTCGGAGTCGCGCTTCGCGCGCCCTCACTGCAGCCCCCTGCGTCG
 CCGGGACCCTCGCGCGCGACCAGCCGAATCGCTCCTGCAGCAGAGCCAAC

 10 20 30 40 50
 ATGCCCATCACTTGGATGCGCATGAGACCCTGGCTAGAGATGCAGATTAA

 60 70 80 90 100
 TTCCAACCAAATCCCGGGGCTCATCTGGATTAATAAAGAGGAGATGATCT
 110 120 130 140 150
 TGGAGATCCCATGGAAGCATGCTGCCAAGCATGGCTGGGACATCAACAAG

 160 170 180 190 200
 GATGCCTGTTTGTCCGGAGCTGGGCCATTACACAGGCCGATACAAAGC
 210 220 230 240 250
 AGGGGAAAAGGAGCCAGATCCCAAGACGTGGAAGGCCAACTTTCGCTGTG

 260 270 280 290 300
 CCATGAACTCCCTGCCAGATAFCGAGGAGGTGAAAGACCAGAGCAGGAAC

 310 320 340 340 350
 AAGGGCAGCTCAGCTGTGCGAGTGTACCGGATGCTTCCACCTCTCACCAA

 360 370 380 390 400
 GAACCAGAGAAAAGAAAGAAAGTCAAGTCCAGCCGAGATGCTAAGAGCA
 410 420 430 440 450
 AGGCCAAGAGGAAGTCATGTGGGGATTCCAGCCCTGATACCTTCTCTGAT

 460 470 480 490 500
 GGACTCAGCAGCTCCACTCTGCCTGATGACCACAGCAGCTACACAGTTCC
 510 520 530 540 550
 AGGCTACATGCAGGACTTGGAGGTGGAGCAGGCCCTGACTCCAGCACTGT

 560 570 580 590 600
 CGCCATGTGCTGTCAGCAGCACTCTCCCCGACTGGCACATCCCAGTGGA
 610 620 630 640 650
 GTTGTGCCGGACAGCACCAGTGATCTGTACAACCTCCAGGTGTCACCCAT

 660 670 680 690 700
 GCCCTCCATCTCTGAAGCTACAACAGATGAGGATGAGGAAGGGAAATTAC

 710 720 730 740 750
 CTGAGGACATCATGAAGCTCTTGGAGCAGTCGGAGTGGCAGCCAACAAAC

 760 770 780 790 800
 GTGGATGGGAAGGGGTACCTACTCAATGAACCTGGAGTCCAGCCCACCTC

 810 820 830 840 850
 TGTCTATGGAGACTTTAGCTGTAAGGAGGAGCCAGAAATTGACAGCCCAG

 860 870 880 890 900
 GGGGGGATATTGGGCTGAGTCTACAGCGTGTCTTCACAGATCTGAAGAAC

```

          910          920          930          940          950
ATGGATGCCACCTGGCTGGACAGCCTGCTGACCCCAGTCCGGTTGCCCTC

          960          970
CATCCAGGCCATTCCCTGTGCACCGTAGCAGGGCCCCTGGGCCCCCTCTTA
TTCCTCTAGGCAAGCAGGACCTGGCATCATGGTGGATATGGTGCAGAGAA
GCTGGACTTCTGTGGGCCCCCTCAACAGCCAAGTGTGACCCCACTGCCAAG
TGGGGATGGGCCTCCCTCCTTGGGTCATTGACCTCTCAGGGCCTGGCAGG
CCAGTGTCTGGGTTTTTCTTGTGGTGTAAGCTGGCCCTGCCTCCTGGGA
AGATGAGGTTCAGAGACCAGTGTATCAGGTCAGGGACTTGGACAGGAGTC
AGTGTCTGGCTTTTTCTCTGAGCCCAGCTGCCTGGAGAGGGTCTCGCTG
TCACTGGCTGGCTCCTAGGGGAACAGACCAGTGACCCAGAAAAGCATAA
CACCAATCCCAGGGCTGGCTCTGCACTAAGAGAAAATTGCACTAAATGAA
TCTCGTTCCAAAGAACTACCCCTTTTCAGCTGAGCCCTGGGGACTGTTCC
AAAGCCAGTGAATGTGAAGGAAAGTGGGGTCCTTCGGGGCAATGCTCCCT
CAGCCTCAGAGGAGCTCTACCTGCTCCCTGCTTTGGCTGAGGGGCTTGG
GAAAAAACTTCGCACTTTTTCGTGTGGATCTTGCCACATTTCTGATCAG
AGGTGTACACTAACATTTCCCCCGAGCTCTTGGCCTTTGCATTTATTTAT
ACAGTGCCTTGCTCGGGGGCCACCACCCCTCAAGCCCCAGCAGCCCTCA
ACAGGGCCAGGGAGGGAAGTGTGAGCGCCTTGGTATGACTTAAAATTGGA
AATGTCATCTAACCATTAAGTCATGTGTGAACACATAAGGACGTGTGTAA
ATATGTACATTTGTCTTTTTATAAAAAGTAAAATTGTT

```

oder eine degenerierte Variante davon.

Ein weiteres bevorzugtes DNA-Molekül (das für Maus-IRF-I codiert) ist durch ein Strukturgen der folgenden Formel III gekennzeichnet:

Formel III

```

ATG CCA ATC ACT CGA ATG CGG ATG AGA CCC TGG CTA GAG ATG CAG ATT
AAT TCC AAC CAA ATC CCA GGG CTG ATC TGG ATC AAT AAA GAA GAG ATG
ATC TTC CAG ATT CCA TGG AAG CAC GCT GCT AAG CAC GGC TGG GAC ATC
AAC AAG GAT GCC TGT CTG TTC CGG AGC TGG GCC ATT CAC ACA GGC CGA
TAC AAA GCA GGA GAA AAA GAG CCA GAT CCC AAG ACA TGG AAG GCA AAC
TTC CGT TGT GCC ATG AAC TCC CTG CCA GAC ATC GAG GAA GTG AAG GAT
CAG AGT AGG AAC AAG GGC AGC TCT GCT GTG CGG GTG TAC CGG ATG CTG
CCA CCC CTC ACC AGG AAC CAG AGG AAA GAG AGA AAG TCC AAG TCC AGC
CGA GAC ACT AAG AGC AAA ACC AAG AGG AAG CTG TGT GGA GAT GTT AGC
CCG GAC ACT TTC TCT GAT GGA CTC AGC AGC TCT ACC CTA CCT GAT GAC
CAC AGC AGT TAC ACC ACT CAG GGC TAC CTG GGT CAG GAC TTG GAT ATG

```

CAA AGG GAC ATA ACT CCA GCA CTG TCA CCG TGT GTC GTC AGC AGC AGT
 CTC TCT GAG TGG CAT ATG CAG ATG GAC ATT ATA CCA GAT AGC ACC ACT
 GAT CTG TAT AAC CTA CAG GTG TCA CCC ATG CCT TCC ACC TCC GAA GCC
 GCA ACA GAC GAG GAT GAG GAA GGG AAG ATA GCC GAA GAC CTT ATG AAG
 CTC TTT GAA CAG TCT GAG TGG CAG CCG ACA CAC ATC GAT GGC AAG GGA
 TAC TTG CTC AAT GAG CCA GGG ACC CAG CTC TCT TCT GTC TAT GGA GAC
 TTC AGC TGC AAA GAG GAA CCA GAG ATT GAC AGC CCG GGA GGG GAC ATT
 GGG ATA GGC ATA CAA CAT GTC TTC ACG GAG ATG AAG AAT ATG GAC TCC
 ATC ATG TGG ATG GAC AGC CTG CTG GGC AAC TCT GTG AGG CTG CCG CCC
 TCT ATT CAG GCC ATT CCT TGT GCA CCA TAG

oder eine degenerierte Variante davon.

Ein derartiges wie im vorhergehenden Absatz genanntes DNA-Molekül kann beispielsweise durch ein Strukturgen gekennzeichnet sein, das die im vorangegangenen erläuterte Formel sowie „upstream“- und „downstream“-Flankensequenzen wie in der folgenden Formel IV hat:

Formel IV

1 GGACGTGCTTTTCACAGTCTAAGCCGAACCGAACCGAACCGAACCGAACCGAACCGGGCC
 60 GAGTTGCGCCGAGGTCAGCCGAGGTGGCCAGAGGACCCCAGCATCTCGGGCATCTTTTCG
 119 CTTCGTGCGCGCATCGCGTACCTACACCGCAACTCCGTGCCTCGCTCTCCGGCACCCCTC
 178 TGCGAATCGCTCCTGCAGCAA AGCCACC ATG CCA ATC ACT CGA ATG CGG
 227 ATG AGA CCC TGG CTA GAG ATG CAG ATT AAT TCC AAC CAA ATC CCA
 272 GGG CTG ATC TGG ATC AAT AAA GAA GAG ATG ATC TTC CAG ATT CCA
 317 TGG AAG CAC GCT GCT AAG CAC GGC TGG GAC ATC AAC AAG GAT GCC
 362 TGT CTG TTC CGG AGC TGG GCC ATT CAC ACA GGC CGA TAC AAA GCA
 407 GGA GAA AAA GAG CCA GAT CCC AAG ACA TGG AAG GCA AAC TTC CGT
 452 TGT GCC ATG AAC TCC CTG CCA GAC ATC GAG GAA GTG AAG GAT CAG
 497 AGT AGG AAC AAG GGC AGC TCT GCT GTG CCG GTG TAC CGG ATG CTG
 542 CCA CCC CTC ACC AGG AAC CAG AGG AAA GAG AGA AAG TCC AAG TCC
 587 AGC CGA GAC ACT AAG AGC AAA ACC AAG AGG AAG CTG TGT GGA GAT
 632 GTT AGC CCG GAC ACT TTC TCT GAT GGA CTC AGC AGC TCT ACC CTA
 677 CCT GAT GAC CAC AGC AGT TAC ACC ACT CAG GGC TAC CTG GGT CAG
 722 GAC TTG GAT ATG GAA AGG GAC ATA ACT CCA GCA CTG TCA CCG TGT
 767 GTC GTC AGC AGC AGT CTC TCT GAG TGG CAT ATG CAG ATG GAC ATT
 812 ATA CCA GAT AGC ACC ACT GAT CTG TAT AAC CTA CAG GTG TCA CCC
 857 ATG CCT TCC ACC TCC GAA GCC GCA ACA GAC GAG GAT GAG GAA GGG
 902 AAG ATA GCC GAA GAC CTT ATG AAG CTC TTT GAA CAG TCT GAG TGG

947 CAG CCG ACA CAC ATC GAT GCC AAG GGA TAC TTG CTC AAT GAG CCA
 992 GGG ACC CAG CTC TCT TCT GTC TAT GGA GAC TTC AGC TGC AAA GAG
 1037 GAA CCA GAG ATT GAC AGC CCT CGA GGG GAC ATT GGG ATA GGC ATA
 1082 CAA CAT GTC TTC ACG GAG ATG AAG AAT ATG GAC TCC ATC ATG TGG
 1127 ATG GAC AGC CTG CTG GGC AAC TCT GTG AGG CTG CCG CCC TCT ATT
 1172 CAG GCC ATT CCT TGT GCA CCA TAG TTTGGGTCTCTGACCCGTTCTTGCCC
 1222 TCCTGAGTGAGTTAGGCCTTGGCATCATGGTGGCTGTGATACAAAAAAGCTAGACTCC
 1281 TGTGGGCCCCCTTGACACATGGCAAAGCATAGTCCCAGTCAAACAGGGGACCATCCTCC
 1340 TTGGGTGAGTGGGCTTCAGGGCTTAGGAGGCAGAGTCTGAGTTTCTTGTGAGGTGAA
 1399 GCTGGCCCTGACTCCTAGGAAGATGGATTGGGGGGTCTGAGGTGTAAGGCAGAGGCCAT
 1458 GGACAGGAGTCATCTTCTAGCTTTTTTAAAAGCCTTGTTCATAGAGAGGGTCTTATCGC
 1517 TGGGCTGGCCCTGAGGGGAATAGACCAGCGCCACAGAAGAGCATAGCACTGGCCCTAG
 1576 AGCTGGCTCTGTACTAGGAGACAATTGCACTAAATGAGTCCTATTCCTAAAGAACTGCT
 1635 GCCCTTCCCAACCGAGCCCTGGGATGGTTCCCAAGCCAGTGAAATGTGAAGGGAAAAAA
 1694 AATGGGGTCCTGTGAAGGTGGCTCCCTTAGCCTCAGAGGGAAATCTGCCTCACTACCTG
 1,53 CTCCAGCTGTGGGGCTCAGGAAAAAAAATGGCACTTTCTCTGTGGACTTTGCCACATT
 1812 TCTGATCAGAGGTGTACACTAACATTTCTCCCCAGTCTAGGCCTTTGCAATTTATTTATA
 1871 TAGTGCCCTGCCTGGTGCCTGCTGTCTCCTCAGGCCTTGGCAGTCTCAGCAGGCCAG
 1930 GGAAAAGGGGGGTGTGAGCGCCTTGGCGTGACTCTTGACTATCTATTAGAAACGCCAC
 1989 CTAAC TGCTAAATGGTGTGTTGTCATGTGGTGGACCTGTGTAAATATGTATATTTGTCT
 2048 TTTTATAAAAATTTAAGTTGTTTACAAAAA 2082

oder eine degenerierte Variante davon.

Erfindungsgemäße DNA-Sequenzen, die für Proteine codieren, die IRF-I-Aktivität besitzen, schließen DNA-Sequenzen eukaryontischen Ursprungs ein, und zwar nicht nur Human- und Maus-, sondern auch Kücken-, Frosch- und Hefe-DNA-Sequenzen, die mit den DNA-Sequenzen des vorgenannten Human- oder Maus-IRF-I (Fig. I bis IV) hybridisieren und die für ein Protein mit IRF-I-Aktivität codieren.

Ein derartiges cDNA-Molekül, das an die cDNA von Maus-IRF-I nach der Definition in Fig. 4A hybridisiert, wird im folgenden in Formel IIIa vorgestellt.

Formel IIIa

10 20 30 40
 ATGCCGGTGGAACGGATGCGAATGCGCCCGTGGCTGGAGGAGCAG
 50 60 70 80 90
 ATAAATTCCAATACGATACCAGGGCTAAAGTGGCTGAACAAGGAG
 100 110 120 130
 AAGAAGATTTTCCAGATCCCCTGGATGCATGCGGCTCGGCACGGA
 140 150 160 170 180
 TGGGACGTGGAAAAGGATGCTCCGCTCTTCAGAACTGGGCGATC
 190 200 210 220
 CATACAGGAAAGCATCAACCAGGAATAGATAAACCAGATCCAAAA
 230 240 250 260 270
 ACATGGAAAGCAAATTTTCGATGTCCATGAATTCCTGCCCGAC

```

      280      290      300      310
ATTGAGGAAGTGAAGGACAGAAGCATAAAGAAAGGAACAACGCC

      320      330      340      350      360
TTCAGAGTCTACCGGATGCTGCCCTTATCCGAACGACCTTCCAAG

      370      380      390      400
AAAGGAAAGAAACCAAAGACAGAAAAAGAAGAGAGAGTTAAGCAC

      410      420      430      440      450
ATCAAGCAAGAACCAGTTGAGTCATCTTTGGGGCTTAGTAATGGA

      460      470      480      490
GTAAGTGGCTTTTCTCCTGAGTATGCGGTCTGACTTCAGCTATA

      500      510      520      530      540
AAAAATGAAGTGGATAGTACGGTGAACATCATAGTTGTAGGACAG

      550      560      570      580
TCCCATCTGGACAGAACATTGAAGATCAAGAGATCGTCACTAAC

      590      600      610      620      630
CCCCAGACATCTGCCAGGTTGTAGAAGTGACCACTGAGAGTGAT

      640      650      660      670
GACCAGCCAGTCAGCATGAGTGAGCTCTACCCTCTACAGATTCT

      680      690      700      710      720
CCTGTGTCTTCTACGCAGAAAGCGAAACTACCGACAGTGTGGCC

      730      740      750      760
AGTGATGAAGAGAACGCAGAGGGGAGACCACACTGGAGGAAGAGG

      770      780      790      800      810
AGCATCGAAGGCAAGCAGTACCTCAGCAACATGGGGACACGGAAC

      820      830      840      850
ACCTATCTGCTGCCCAGCATGGCGACCTTTGTACCTCCAACAAG

      860      870      880      890      900
CCAGATCTGCAGGTCAACATCAAAGAGGATAGCTGTCCGATGCCT

      910      920      930      940
TACAACAGCTCCTGGCCCCCATTTACAGACCTTCCCCTTCTGCC

      950      960      970      980      990
CCAGTGACCCCCACGCCAGCAGCAGTCGGCCAGACCGGGAGACC

      1000      1010      1020      1030
CGGGCCAGTGTTCATCAAGAAGACATCTGATATCACCCAGGCCCGT

      1040
GTCAAGAGCTGT

```

oder eine degenerierte Variante davon.

Dieses DNA-Molekül kann beispielsweise durch ein Strukturgen gekennzeichnet sein, das die im vorangegangenen erläuterte Formel sowie „upstream“- und „downstream“-Flankensequenzen wie in der folgenden Formel IVa hat.

TCTCAGGCAAGCCGGGGA
CTAACTTTTAGTTTGGCTCCTGCGATTATTCAACTGACGGGCTTT
CATTTCCATTTTACACACCCTAACAACTCACACCTTGCGGGAT
TGTATTGGTAGCGTGGAAAAAAAAAAGCACATTGAGAGGGTACC
10 20 30 40
ATGCCGGTGGAACGGATGCGAATGCGCCCGTGGCTGGAGGAGCAG
50 60 70 80 90
ATAAATTCCAATACGATACCAGGGCTAAAGTGGCTGAACAAGGAG
100 110 120 130
AAGAAGATTTTCCAGATCCCCTGGATGCATGCGGCTCGGCACGGA
140 150 160 170 180
TGGGACGTGGAAAAGGATGCTCCGCTCTTCAGAACTGGGCGATC
190 200 210 220
CATACAGGAAAGCATCAACCAGGAATAGATAAACCAGATCCAAAA
230 240 250 260 270
ACATGGAAAGCAAATTTTCGATGTGCCATGAATTCCCTGCCCGAC
280 290 300 310
ATTGAGGAAGTGAAGGACAGAAGCATAAAGAAAGGAAACAACGCC
320 330 340 350 360
TTCAGAGTCTACCGGATGCTGCCCTTATCCGAACGACCTTCCAAG
370 380 390 400
AAAGGAAAGAAACCAAAGACAGAAAAAGAAGAGAGAGTTAAGCAC
410 420 430 440 450
ATCAAGCAAGAACCAGTTGAGTCATCTTGGGGCTTAGTAATGGA
460 470 480 490
GTAAGTGGCTTTTCTCCTGAGTATGCGGTCCTGACTTCAGCTATA
500 510 520 530 540
AAAAATGAAGTGGATAGTACGGTGAACATCATAGTTGTAGGACAG
550 560 570 580
TCCCATCTGGACAGCAACATTGAAGATCAAGAGATCGTCACTAAC

590 600 610 620 630
CCGCCAGACATCTGCCAGGTTGTAGAAGTGACCACTGAGAGTGAT

640 650 660 670
GACCAGCCAGTCAGCATGAGTGAGCTCTACCCTCTACAGATTCT

680 690 700 710 720
CCTGTGTCTTCTACGCAGAAAGCGAACTACCGACAGTGTGGCC

730 740 750 760
AGTGATGAAGAGAAAGCAGAGGGGAGACCACACTGGAGGAAGAGG

770 780 790 800 810
AGCATCGAAGGCAAGCAGTACCTCAGCAACATGGGGACACGGAAC

820 830 840 850
ACCTATCTGCTGCCCAGCATGGCGACCTTTGTCACCTCCAACAAG

860 870 880 890 900
CCAGATCTGCAGGTCACCATCAAAGAGGATAGCTGTCCGATGCCT

910 920 930 940
TACAACAGCTCCTGGCCCCATTTACAGACCTTCCCCTCCTGCC

950 960 970 980 990
CCAGTGACCCCCACGCCCAGCAGCAGTCGGCCAGACCGGGAGACC

1000 1010 1020 1030
CGGGCCAGTGTTCATCAAGAAGACATCTGATATCACCCAGGCCCGT

1040
GTCAAGAGCTGTTAAGCCTTTGACTCTCCCTGGTGGTTGTTGGGA

TTTCTTAGCTTTGTGTTGTTCTTTGTTTGTATTATATTATTTTTT

TTCTCTATGATACCTATCTTAGACACATCTAAGGGAGAAAGCCTT

GACGATAGATTATTGATTGCTGTGTCCAACCTCCAGAGCTGGAGCT

TCTTCTTAACCTCAGGACTCCAGCCCCCCCCCCCCCTCGGTAGATG

CGTATCTCTAGAACCTGCTGGATCTGCCAGGGCTACTCCCTCAAG
TTCAAGGACCAACAGCCACACGGGCAGTGGAGGTGCTGCGTTGCC
TACGGTCAAGGCCAGCATGGTGGAGTGGATGCCTCAGAACGGAGG
AGAAAATGTGAACTAGCTGGAATTTTTTTATTCTTGTGAATATGT
ACATAGGGCAGTACGAGCAATGTCGCGGGCTGCTTCTGCACCTTA
TCTTGAAGCACTTACAATAGGCCTTCTTGTAACTTGTCTCCTT
CACAGCACACTCGGCGACCCCTTCTGTGTCCACTACCCCACTACC
CAGCCCTCCCTCCTCAACCCCTCCATCCCGGTCCTCTATGCGCCC
CTTCCCCCAACCAATCCCATCACAACCTCTTACCTATCCTTTCC
CTCCCAACCCCTTCTATCCCAGCCCACCACCTACCCCACTCCTCC
CCAACCTCCTCCATTCTAGCCCATTACCCACGCCTCTCTCCTCAGC
CCAGCCTACCCCATCCCACCCTGTTCCCTTCCCTCCAGTTTCCTCT
CCTCAAAGGCAAGGCTCTACATCTTGGAGGAGGAGGAGGAGAAGA
AAATGAGTTTCTTCACCGCTGTCCCATTTTAAGACTGCTTGAATA
ATAAAAAAAAAATCTTTCTAATCTGCTATGCTTGAATGGCACGCGG
TACAAAGGAAAACGTGCATGGAAATATTATGCAAATTCCCAGATC

TGAAGACGGAAATACTCTAATTCTAACCAGAGCAAGCTTTTTTA

TTTTTTTATACAAGGGGAATATTTTATTCAAGGTAAAAAATTCT

AAATAAAATATAATTGTTTTTATCTTTCTACAGCAAATTTATA

ATTTTAAGATTCCCTTTTCCTGTTTCATCAGCAGTTGTTATTACATC

CCTTGTGGCACATTTTTTTTTTAATTTTGTAAAGGTGAAAAAAAAA

ACTTTTATGAGCTCATGTAGCAATCAAATTATCCTGTGGATTGAT

AATAAATGAATATGGTATATAGTTAAAGATTTTAAAAAAAAAAAAA

oder eine degenerierte Variante davon.

Im folgenden wird die Isolierung von cDNA-Molekülen aus Humanzellen und Mauszellen beschrieben, die für Human- und Maus-IRF-1 codieren, bzw. aus Mauszellen und Hefe, die an die cDNA von Maus-IRF-1 und Human-IRF-1 hybridisieren.

Das rekombinante DNA-Molekül kann auch eine Promotor- und Regulationssequenz wie in der folgenden Formel V enthalten:

Formel V

PstI
CTGCAGAAAGAGGGGGACGGTCTCGGCTTCCAAGACAGGCAAGGGGG
-299

GC Box 1 GC Box 2
CAGGGGAGTGGAGTGGAGCAAGGGGCGGGCCCGCGGTAGCCCCGGGGCGGGTGGCGCGG
-251

GCCCGAGGGGGTGGGGAGCACAGCTGCCTTGTACTTCCCCTTCGCCGCTTAGCTCTAC
-193

AACAGCCTGATTTCCTCGAAATGATGAGGCCGAGTGGCCCAATGGGCGCCAGGAGCG
-135

kleine CAP-Stelle

GC GCGGCGGGGCGTGGCCGAGTCCGGGCCGGGAATCCCGCTAAGTGT TTAGATTTC
-77 -21

große CAP-Stelle

große CAP-Stelle
 ▼▼▼
 TTCGCGGCGCCGCGGACTCGCCAGTGCGCACCACTCCTTCGTCGAGGTAGGACGTGCT
 -19 +1 pIRF-L

TTCACAGTCTAAGCCGAACCGAACCGAACCGAACCGAACCGGGCCGAGTTGCG
+39

```

CCGAGGTCAGCCGAGGTGGCCAGAGGACCCCAGCATCTCGGGCATCTTTTCGCTTCGTG
+97
CGCGCATCGCGTACCTACACGCAACTCCGTGCCTCGCTCTCCGGCACCCCTCTGCGAA
+155
      PstI
      TCGTCTCCTGCAG
+213      +225

```

oder eine degenerierte Variante davon.

Das rekombinante DNA-Molekül kann auch für die Expression eines pharmazeutisch aktiven Proteins wie beispielsweise eines Cytokins oder eines Plasminogenaktivators vorgesehen sein und wird in dieser Form vorzugsweise ein Strukturgen für ein pharmazeutisch aktives Protein operabel gebunden an eine Promotorregion des Gens für dieses Protein einschließlich einer Bindungsstelle für das IRF-I-aktive Molekül enthalten.

Somit kann ein erfindungsgemäßes rekombinantes DNA-Molekül eine wie im vorangegangenen erläuterte DNA-Sequenz und ein Strukturgen für ein gewünschtes pharmazeutisch aktives Protein unter der Kontrolle des IRF-I-aktiven Proteins, für das diese DNA-Sequenz codiert, enthalten. In einem derartigen rekombinanten DNA-Molekül befindet sich das für das IRF-I-aktive Molekül codierende Gen vorzugsweise unter der Kontrolle eines konstitutiven Promotors oder noch besser eines induzierbaren Promotors. Vorzugsweise schließt das Gen für das gewünschte pharmazeutisch aktive Protein eine IRF-Bindungsstelle ein, die repetitive AAGTGA-Sequenzen enthält.

Die vorliegende Erfindung umfaßt auch eine Wirtszelle, z. B. eine Bakterienzelle z. B. E. coli, oder eine Hefezelle oder eine Säugetierzelle z. B. eine CHO-Zelle oder eine Mauszelle z. B. L929, die durch ein rekombinantes DNA-Molekül nach der im vorangegangenen erläuterten Bedeutung transformiert wurden. Idealerweise wird die Wirtszelle aus einer Zelllinie ausgewählt, die keine oder im wesentlichen keine endogene IRF-I-Aktivität zeigt.

Alternativ für die Produktion eines pharmazeutisch aktiven Proteins kann eine Wirtszelle transformiert sein durch ein erstes DNA-Molekül, das eine für ein Protein mit IRF-I-Aktivität codierende Sequenz enthält, und durch ein zweites separates DNA-Molekül, das ein Gen enthält, das unter der Kontrolle des IRF-I-aktiven Proteins, für das das erste DNA-Molekül codiert, für ein gewünschtes, pharmazeutisch aktives Protein codiert. Vorzugsweise umfaßt das erste, für das IRF-I-aktive Molekül codierende DNA-Molekül einen konstitutiven Promotor oder am besten eine induzierbare Promotorsequenz, die operabel an das für die IRF-I-aktive Verbindung codierende Gen gebunden ist. Das zweite DNA-Molekül enthält vorzugsweise auch eine Bindungsstelle für das die repetitiven AAGTGA-Sequenzen enthaltende IRF-I-aktive Protein.

Das IRF-I-aktive Protein oder das pharmazeutisch aktive Protein kann durch Kultivierung der transformierten Zelle und Isolierung des erzeugten Proteins auf herkömmliche Weise hergestellt werden.

Die Wirtszellen werden geeigneterweise durch Behandlung in einer Art und Weise, die dem Promotor entspricht, der operabel an das für den IRF-I codierende Gen gebunden ist, induziert, wie im folgenden erläutert wird.

Die Erfindung umfaßt auch ein Protein, das die Aktivität eines IRF-I besitzt, das durch die Kultivierung eines Wirts gewonnen wurde, der mit einem DNA-Molekül transformiert wurde, wie im vorangegangenen beschrieben wurde.

Ein bevorzugtes Protein mit einer IRF-I-Aktivität hat die folgende Formel VI:

Formel VI

```

Met Pro Ile Thr Arg Met Arg Met Arg Pro Trp Leu Glu Met Gln Ile
Asn Ser Asn Gln Ile Pro Gly Leu Ile Trp Ile Asn Lys Glu Glu Met
Ile Phe Gln Ile Pro Trp Lys His Ala Ala Lys His Gly Trp Asp Ile
Asn Lys Asp Ala Cys Leu Phe Arg Ser Trp Ala Ile His Thr Gly Arg
Tyr Lys Ala Gly Glu Lys Glu Pro Asp Pro Lys Thr Trp Lys Ala Asn
Phe Arg Cys Ala Met Asn Ser Leu Pro Asp Ala Glu Glu Val Lys Asp
Gln Ser Arg Asn Lys Gly Ser Ser Ala Val Arg Val Tyr Arg Met Leu
Pro Pro Leu Thr Arg Asn Gln Arg Lys Glu Arg Lys Ser Lys Ser Ser
Arg Asp Thr Lys Ser Lys Thr Lys Arg Lys Leu Cys Gly Asp Val Ser
Pro Asp Thr Phe Ser Asp Gly Leu Ser Ser Ser Thr Leu Pro Asp Asp
His Ser Ser Tyr Thr Thr Gln Gly Tyr Leu Gly Gln Asp Leu Asp Met
Glu Arg Asp Ile Thr Pro Ala Leu Ser Pro Cys Val Val Ser Ser Ser
Leu Ser Glu Trp His Met Gln Met Asp Ile Ile Pro Asp Ser Thr Thr
Asp Leu Tyr Asn Leu Gln Val Ser Pro Met Pro Ser Thr Ser Glu Ala
Ala Thr Asp Glu Asp Glu Glu Gly Lys Ile Ala Glu Asp Leu Met Lys
Leu Phe Glu Gln Ser Glu Trp Gln Pro Thr His Ile Asp Gly Lys Gly

```

Tyr Leu Leu Asn Glu Pro Gly Thr Gln Leu Ser Ser Val Tyr Gly Asp
Phe Ser Cys Lys Glu Glu Pro Glu Ile Asp Ser Pro Arg Gly Asp Ile
Gly Ile Gly Ile Gln His Val Phe Thr Glu Met Lys Asn Met Asp Ser
Ile Met Trp Met Asp Ser Leu Leu Gly Asn Ser Val Arg Leu Pro Pro
Ser Ile Gln Ala Ile Pro Cys Ala Pro

Ein weiteres bevorzugtes Protein mit einer IRF-1-Aktivität hat die folgende Formel VII:

Formel VII

MetProIleThrTrpMetArgMetArgProTrpLeuGluMet
GlnIleAsnSerAsnGlnIleProGlyLeuIleTrpIleAsnLysGluGluMetIleLeu
GluIleProTrpLysHisAlaAlaLysHisGlyTrpAspIleAsnLysAspAlaCysLeu
PheArgSerTrpAlaIleHisThrGlyArgTyrLysAlaGlyGluLysGluProAspPro
LysThrTrpLysAlaAsnPheArgCysAlaMetAsnSerLeuProAspIleGluGluVal
LysAspGlnSerArgAsnLysGlySerSerAlaValArgValTryArgMetLeuProPro
LeuThrLysAsnGlnArgLysGluArgLysSerLysSerSerArgAspAlaLysSerLys
AlaLysArgLysSerCysGlyAspSerSerProAspThrPheSerAspGlyLeuSerSer
SerThrLeuProAspAspHisSerSerTyrThrValProGlyTyrMetGlnAspLeuGlu
ValGluGlnAlaLeuThrProAlaLeuSerProCysAlaValSerSerThrLeuProAsp
TrpHisIleProValGluValValProAspSerThrSerAspLeuTyrAsnPheGlnVal
SerProMetProSerIleSerGluAlaThrThrAspGluAspGluGluGlyLysLeuPro
GluAspIleMetLysLeuLeuGluGlnSerGluTrpGlnProThrAsnValAspGlyLys
GlyTryLeuLeuAsnGluProGlyValGlnProThrSerValTyrGlyAspPheSerCys
LysGluGluProGluIleAspSerProGlyGlyAspIleGlyLeuSerLeuGlnArgValPheThrAspLeuLysAsnMetAspAlaThrTrp
LeuAspSerLeuLeuThrProValArgLeuProSerIleGlnAlaIleProCysAlaPro

Noch ein weiteres bevorzugtes Protein, das durch die cDNA-Sequenz der Formel IIIa codiert wird, hat die folgende Formel VIII.

Formel VIII

MetProValGluArgMetArgMetArgProTrpLeuGluGluGln
IleAsnSerAsnThrIleProGlyLeuLysTrpLeuAsnLysGlu
LysLysIlePheGlnIleProTrpMetHisAlaAlaArgHisGly
TrpAspValGluLysAspAlaProLeuPheArgAsnTrpAlaIle
HisThrGlyLysHisGlnProGlyIleAspLysProAspProLys
ThrTrpLysAlaAsnPheArgCysAlaMetAsnSerLeuProAsp
IleGluGluValLysAspArgSerIleLysLysGlyAsnAsnAla
PheArgValTyrArgMetLeuProLeuSerGluArgProSerLys
LysGlyLysLysProLysThrGluLysGluGluArgValLysHis
IleLysGlnGluProValGluSerSerLeuGlyLeuSerAsnGly
ValSerGlyPheSerProGluTyrAlaValLeuThrSerAlaIle
LysAsnGluValAspSerThrValAsnIleIleValValGlyGln
SerHisLeuAspSerAsnIleGluAspGlnGluIleValThrAsn
ProProAspIleCysGlnValValGluValThrThrGluSerAsp

AspGlnProValSerMetSerGluLeuTyrProLeuGlnIleSer
 ProValSerSerTyrAlaGluSerGluThrThrAspSerValAla
 SerAspGluGluAsnAlaGluGlyArgProHisTrpArgLysArg
 SerIleGluGlyLysGlnTyrLeuSerAsnMetGlyThrArgAsn
 ThrTyrLeuLeuProSerMetAlaThrPheValThrSerAsnLys
 ProAspLeuGlnValThrIleLysGluAspSerCysProMetPro
 TyrAsnSerSerTrpProProPheThrAspLeuProLeuProAla
 ProValThrProThrProSerSerSerArgProAspArgGluThr
 ArgAlaSerValIleLysLysThrSerAspIleThrGlnAlaArg
 ValLysSerCys

Im folgenden wird die molekulare Clonierung und Charakterisierung von Maus- und Human-cDNA, die für DNA-bindende Proteine, die die IRF-1-Aktivität besitzen, codieren, beschrieben.

Eine bemerkenswerte Sequenzerhaltung ist zwischen den Maus- und Human-IRF-1-Molekülen ersichtlich, wie sich aus der Analyse der clonierten cDNAs ergab. Weiterhin zeigt sich, daß die Expression des für IRF-1 codierenden Gens durch das Newcastle-Virus (NVD) und durch Concanavalin A (ConA) in Maus-L929-Zellen bzw. Milzlymphozyten induziert wird.

Ausführungsbeispiele

1. Clonierung und Expression von IRF-1-DNA in E. coli

A. Poly(A)⁺ RNA wurde aus nichtinduzierten Maus-L929-Zellen isoliert und für die Synthese von cDNA (nach dem Verfahren von Aruffo und Seed, 1987) verwendet. Die entstandene cDNA wurde anschließend in einen EcoRI-geschnittenen λ gt11-Vektor cloniert, und eine cDNA-Bibliothek wurde nach dem Standardverfahren von Huynh u. a., 1985, mittels E. coli Y1090 als Wirtstamm konstruiert.

Die entstandene λ gt11-Bibliothek wurde anschließend mittels der multimerisierten (4mal) AAGTGA-Sequenz (im folgenden als das CI-Oligomer bezeichnet; Fujita u. a., 1987) als Sonde gescreent.

Bei diesem Screening-Verfahren wurde E. coli Y1090, durch die rekombinanten λ gt11-Phagen infiziert, auf rechteckige Platten von 10cm $\times$ 13cm ausgestrichen. Die Platten wurden dann bei 12°C 4 bis 5 Stunden, wenn ungefähr 20000 Plaques/Platte zu sehen waren, inkubiert.

Die Membranfilter zum Screenen waren entweder aus Nylon (Nytran; Schleicher & Schnell) oder Nitrocellulose (Schleicher & Schnell). Die Filter wurden in 10mM IPTG (zur Induktion der LacZ-Genexpression in den entsprechenden Phagenplaques) getaucht und an der Luft getrocknet, dann über die Platten gelegt und 2,5 Stunden bei 37°C inkubiert. Die Plaques werden nur ein Protein erzeugen, für das die cDNA codiert, wenn die cDNA mit dem LacZ-Gen im Rahmen ist.

Die Filter werden entfernt, 20 Minuten bei 4°C gekühlt und dann ohne Trocknen gescreent.

Die Membranfilter wurden dann wie folgt für den Test vorbereitet: Aus Maus-L929-Zelle wurde der Kernextrakt hergestellt und auf die Membranen getüpfelt (in einer Menge, die etwa 10 μ g Protein entsprach).

Vor der DNA-Bindung wurde ein Bindungspuffer, der aus 10mM Hepes, pH 7,5, 5,0mM NaCl, 1mM DTT, 1mM EDTA, 5% Glycerol bestand, zugesetzt. 5% fettfreies Milchpulver (Yukijurishi Inc.) wurde zu dem Puffer gegeben, und die Filter wurden in diesem Gemisch 1 Stunde lang bei 4°C inkubiert und anschließend 1 Minute im gleichen Puffer, der jedoch kein Milchpulver enthielt, gespült.

Die Filter wurden dann in Bindungspuffer (1 ml) inkubiert, der 350 μ g/ml Lachssperma-DNA mit einer durchschnittlichen Länge von etwa 300bp und die ³²P-markierte Sonde aus C1-Oligomer-DNA (10⁶cpm/ml; spezifische Aktivität 2000cpm/f Mol) enthielt (siehe Fujita u. a., 1987). Die Sonden-DNA wurde mit T4-Kinase an den 5'-Termini [γ -³²P]ATP endmarkiert.

Nach der Bindung wurden die Filter bei Raumtemperatur mit dem Bindungspuffer 1 bis 2 Stunden lang gewaschen (10ml Filter), wobei der Bindungspuffer mehrere Male gewechselt wurde. Die Filter wurden dann an der Luft getrocknet und autoradiographisch untersucht. Bei diesem Test ergaben die Nylonmembranen, jedoch nicht die Nitrocellulosemembranen, ein positives Signal, was durch Zugabe von überschüssigem nichtmarkiertem C1-Oligomer spezifisch inhibiert wurde.

Auf diese Weise wurden etwa 1,4 $\times$ 10⁶ Rekombinanten gescreent. Unter den 32 positiven Phagenclonen, die beim ersten Screenen identifiziert wurden, war ein Clon (der als λ L28-8 bezeichnet wurde), der wiederholt in aufeinanderfolgenden Screeningdurchgängen an die Sonden-DNA band.

2. Herstellung und Reinigung von Protein in transfektiertem E. coli

Lysogene Bakterienclone wurden durch Transfektion von E. coli Y1089 mit λ L28-8 hergestellt. Die über Nacht stehengelassenen Lysogene, die λ L28-8 enthielten, wurden zu 1% in 400ml L-Brühe eingimpft. Die Bakterien wurden bei 31°C gezüchtet, bis OD₆₀₀ gleich 1 war. Die Temperatur wurde dann für 20 Minuten auf 42°C erhöht. IPTG wurde dann zu 10mM zugegeben, und nach weiterer 20minütiger Inkubation bei 38°C wurden die Kulturen schnell pelletiert und in 10ml Lysispuffer suspendiert, der aus 20mM Hepes pH 7,9, 0,2mM EDTA, 0,5mM Spermidin, 0,15mM Spermin, 0,1mM DTT, 10% Glycerol, 0,5mM Phenylmethylsulfonylsulfonylfluorid (PMSF), 1 μ g/ml Pepstatin A, 1 μ g/ml Leupeptin, 500 μ M L-I-Tosylamid-2-phenylethylchloromethylbenzamidin, 10mM Natriummolybdat, 2mM Natriumpyrophosphat und 2mM Natriumorthovanadat bestand. Die Zellsuspensionen wurden drei schnellen Gefrier/Auftau-Zyklen ausgesetzt und anschließend bei 30000 U min⁻¹

1 Stunde lang bei 4°C mittels eines Beckman 50Ti Rotors zentrifugiert. Der Überstand wurde entweder direkt für einen Gelverzögerungstest (siehe unten) verwendet oder weiter gereinigt.

Die weitere Reinigung wurde folgendermaßen ausgeführt:

Etwa 4 ml Überstand wurden auf eine poly(di-C):poly(di-C)-Säule mit einem Bettvolumen von 2 ml und mit Lysepuffer äquilibriert, aufgegeben. Das Durchflußmaterial (4 ml) wurde dann auf eine DE 52 (Whatman) Säule mit einem Bettvolumen von 2 ml und mit Puffer Z (25 mM Hepes, pH 7,8, 12,5 mM MgCl₂, 1 mM DTT, 20% Glycerol, 0,1% NP-40, 0,5 mM PMSF) äquilibriert, aufgegeben. Die DNA-Bindungsaktivität wurde durch den 0,1 M KCl enthaltenden Puffer Z eluiert. Das Eluat (etwa 4 ml) wurde weiter mit Centricon-10 (Amicon) eingeeengt.

Die Endkonzentration des Proteins betrug 28 mg/ml.

3. Charakterisierung des Proteinprodukts des Clons λ L28-8

(1) Gelverzögerungstest

Lysogene Bakterienclone, die λ L28-8 enthielten, wurden durch Transfektion von *E. coli* Y1089 hergestellt und wurden induziert, um die clonierte cDNA in großen Mengen mittels des Verfahrens von Huynh u. a., 1985, zu exprimieren. Lysogene Clone, die in der gleichen Art und Weise hergestellt und behandelt worden waren, denen das cDNA-Insert (als λ 6 bezeichnet) fehlt, wurden als Kontrolle benutzt.

Aus den induzierten Lysogenkulturen wurden Extrakte hergestellt, die je 3 μ g Protein enthielten (vier Präparate, die als λ L28-8a, λ L28-8b und λ 6a und λ 6b bezeichnet wurden).

Kernextrakte (3 μ g Protein) wurde ebenfalls aus Maus-L929-Zellen hergestellt.

Die Extrakte wurden mit 1 f mol gekennzeichnete C1-Oligomer Sonde nach obiger Beschreibung mit einer spezifischen Aktivität von 8000 cpm/f mol inkubiert.

Jeder Extrakt wurde auch einem Vergleichstest unterzogen, in dem eine kompetitive DNA in verschiedenen Konzentrationen zu der Inkubationskultur gegeben wurde.

Die Ergebnisse des Tests sind in Fig. 1 dargestellt, in der die verschiedenen Bahnen folgendem entsprechen:

Bahn	Extrakt
1	nichts
2	λ 6a
3	λ 6b
4	λ 6b mit 1000fachem molarem Überschuß an nichtmarkiertem C1-Oligomer
5	λ 6b mit 1000fachem molarem Überschuß an nichtmarkiertem C5A-Oligomer*
6	λ L28-8a
7	λ L28-8b
8	λ L28-8b mit 1000fachem molarem Überschuß an nichtmarkiertem C1-Oligomer
9	λ L28-8b mit 1000fachem molarem Überschuß an nichtmarkiertem C5A-Oligomer*
10	L929-Zelle
11	L929-Zelle mit 1000fachem molarem Überschuß an nichtmarkiertem C1-Oligomer
12	L929-Zelle mit 1000fachem molarem Überschuß an nichtmarkiertem C5A-Oligomer*

* Das in Fujita u. a., 1985 beschriebene C5A-Oligomer ist eine 6mal wiederholte GAAA-Sequenz.

Aus Fig. 1 wird ersichtlich, daß gebundene Sonden in den Bahnen 6, 7, 9, 10 und 12 nachweisbar sind.

Wie in Fig. 1 gezeigt wird, ergaben Proteinextrakte aus den λ L28-8-Lysogenen verschobene Bahnen (Bahnen 6 und 7), deren Erscheinen durch den Überschuß an nichtmarkierter C1-Oligomer-DNA inhibiert war, jedoch nicht in dem gleichen Maße wie bei C5A-Oligomer (Bahnen 8 und 9).

Im Gegensatz zu den λ L28-8-stämmigen Proteinen ergaben diejenigen, die aus den induzierten λ 6-stämmigen Lysogenen hergestellt worden waren, keine derartigen verschobenen Banden (Bahnen 2–5). Die verschobenen Banden sind dicht bei denen von natürlichem IRF-1 aus Maus-L929-Zellen (Bahnen 10 und 12) zu sehen. Die in den beiden Reihen der verschobenen Banden existierenden Unterschiede sind als eine Folge der unterschiedlichen Mengen an Protein, das an die Sonden-DNA gebunden ist, zu sehen.

Weiter wird festgestellt, daß die verschobenen Banden nur bei den Proteinen aus IPTG-induzierten Y1089-Zellen, die mit λ L28-8 transfiziert waren, nachweisbar waren.

(II) DNAase-Fußspurenanalyse

Die Fußspurenanalyse wurde durchgeführt, um die Bindungseigenschaften des durch λ L28-8-cDNA codierten Proteins an eine DNA, die für die IFN- β -Gen-„upstream“-Region codiert, zu testen.

Das durch λ L28-8-cDNA codierte Protein wurde aus dem induzierten Lysogen extrahiert und teilweise durch Säulenchromatographie wie oben beschrieben gereinigt und auf seine Bindungseigenschaften an eine DNA, die die IFN- β -Gen-„upstream“-Region enthält, geprüft.

Die Sonden-DNAs wurden als Sall-HindIII-Fragmente hergestellt, die aus p-125cat (das das Wildtyp IFN- β -Gen enthält) und p-125DPcat (das ein mutiertes IFN- β -Gen enthält) isoliert wurden. Das Plasmid p-125cat wurde als p-105cat konstruiert (Fujita u. a., 1987), mit der Ausnahme, daß das Fragment BamHI (–125)-Tagl (+19) aus pSE-125 (Fujita u. a., Cell, Bd. 41, S. 489–496, 1985) verwendet wurde. Das Plasmid p-125DPcat, das Punktmutationen innerhalb der IFN- β -Regulationselemente trägt, wurde durch synthetische ortsspezifische Oligonucleotidmutagenese an p-125cat wie bei Hatakeyama u. a., Proc. Natl. Acad. Sci. USA, Bd. 83, S. 9650–9654, 1986) beschrieben, gewonnen. Beide DNAs wurden durch [γ -³²P]ATP mittels T4-Kinase an der HindIII-Stelle markiert.

4 f · Mol Sonden-DNA (spezifische Aktivität 3000 cpm/f · Mol) wurden in 20 μ l Reaktionsgemisch, das 25 mM Tris-HCl, pH 7,9, 6,25 mM MgCl₂, 50 mM KCl, 1 mM EDTA, 0,5 mM DTT, 10% Glycerol, 2% Polyvinylalkohol enthielt, mit oder ohne 280 μ g gereinigtes Protein inkubiert.

In dem Test wurden 5×10^{-4} Einheiten DNAase I (Worthington) zugefügt und 1 Minute bei 25°C inkubiert.

Fig. 2 zeigt Autoradiogramme der aus den Proben gewonnenen DNA-Fragmente, die während der folgenden Verfahren

gewonnen wurden. Auf der linken Seite von Fig. 2 ist der Teil der DNA-Sequenz der Wildtyp-IFN- β -Sonde dargestellt, und auf der rechten Seite der Teil der DNA-Sequenz der mutierten IFN- β -Sonde.

1. Die Wildtyp-IFN- β -Sonden wurde durch A+G-Reaktionen geschnitten (siehe Methods in Enzymology, Bd. 65, S. 499–560) – das Ergebnis wird in Bahn 1 gezeigt.
2. Die Wildtyp-IFN- β -Sonde wurde teilweise durch DNAase I ohne Schutz digeriert – das Ergebnis wird in Bahn 2 gezeigt.
3. Die Wildtyp-IFN- β -Sonde wurde mit dem Protein umgesetzt und dann mit DNAase I digeriert – das Ergebnis wird in Bahn 3 gezeigt.
4. Die Wildtyp-IFN- β -Sonde wurde mit dem Protein in Gegenwart eines 1000fachen molaren Überschusses an nichtmarkiertem C1-Oligomer umgesetzt und anschließend mit DNAase I digeriert – das Ergebnis wird in Bahn 4 gezeigt.
5. Die mutierte IFN- β -Sonde wurde mit dem Protein umgesetzt und anschließend durch DNAase I digeriert – das Ergebnis wird in Bahn 5 gezeigt.
6. Die mutierte IFN- β -Sonde wurde durch A+G-Reaktionen geschnitten – das Ergebnis wird in Bahn 6 gezeigt.

In Fig. 2 werden die in den Bahnen 3 und 5 gefundenen geschützten Regionen entsprechend den auf der linken bzw. rechten Seite des Autoradiogramms dargestellten Sequenzen gezeigt. Die Hexamermotive sind eingerahmt.

Aus den Ergebnissen in Fig. 2 ist ersichtlich, daß die geschützte Region den Nucleotiden –100 bis –64 entspricht und dies die Region ist, die durch aus L929-Zellen gewonnene IRF-1 geschützt wird. Der Schutz wurde durch die Verwendung von überschüssigem nichtmarkiertem C1-Oligomer (Bahn 4) aufgehoben. Bei Verwendung geringerer Proteinkonzentration wurde auch festgestellt, daß die das AAGTGA-Motiv (–80 bis –70) enthaltende Region bevorzugt geschützt war.

Diese Ergebnisse zeigen, daß das Protein eine höhere Affinität zu der das AAGTGA-Motiv enthaltenden Region und eine geringere Affinität zu der Nachbarrregion hat.

Das mutierte IFN- β -Genesegment trägt T-G-Mutationen an den Positionen –106, –100, –73 und –67. In den Vergleichsbahnen 5 und 2 unter den gleichen Testbedingungen erscheint der durch die rekombinanten Proteine gewährte Schutz an der nichtmutierten Region eingeschränkt (Bahn 2). Diese Beobachtung stimmt mit der Beobachtung überein, daß die Einführung dieser Mutationen zu einer drastischen Reduzierung (20fach) der Induzierbarkeit der Transkription durch NDV in L929-Zellen führt und daß die *in vitro* Bindung IRF-1 an das mutierte IFN- β -Gen nur in der nichtmutierten Region feststellbar war. Weiter zeigt die Verwendung des C1-Oligomers, daß das Protein spezifisch die Region schützt, die die Oligomersequenzen enthält. Dieser Schutz entspricht vollkommen dem, der durch nativen IRF-1, der aus L-929-Zellen stammt, gewährt wird.

3. DNA-Konkurrenztest

Dieser Test wird ausgeführt, um die Affinität des rekombinanten Proteins zu verschiedenen DNA-Sequenzen einschließlich einigen bekannten Transkriptionsregulations-DNA-Sequenzen zu untersuchen. Das Verfahren wird folgenderweise durchgeführt:

Das HindIII-Sall-Fragment der IFN- β -Sonde wurde aus p-125cat isoliert (siehe Fujita u. a., 1987); dieses Fragment enthält die Human-IFN- β -Genesequenz von +19 bis –125. Die DNA wurde an den 3'-Termini durch Füllen beider Enden mit [α - ^{32} P] dCTP mittels Klenow-Fragment markiert.

Die spezifische Aktivität der Sonden-DNA betrug 8000 cpm/f · Mol.

Die Gelverzögerungstests wurden unter den oben beschriebenen Bedingungen durchgeführt.

In den Durchläufen des Konkurrenztests wurde, wie in Fig. 3 gezeigt, die DNA mit dem Protein in Gegenwart verschiedener Konzentrationen von Konkurrenz-DNAs in dem Bindungsgemisch umgesetzt.

Der Bildungswert des Komplexes wurde durch densitometrische Analyse des Autoradiogramms quantitativ bestimmt. Die Komplexbildung in Abwesenheit von Konkurrenz-DNA wurde als 100% angenommen.

Die Struktur der Konkurrenz-DNAs sah folgendermaßen aus:

- a) AP-1: eine synthetische DNA mit folgender Sequenz



- b) TNF: 37 bp lange synthetische DNA, die von +1162 bis +2116 (siehe Nedwin u. a., 1985) des ersten TNF- α -Introns reicht;

- c) Maus-H2-D^d: 37 bp lange synthetische DNA, die das von Korber u. a., 1988 beschriebene IRS-Element (–159 bis –123) umfaßt;

- d) Human-IFN- α : 46 bp lange synthetische DNA, die dem Virus-Antwortelement (Virus Response Element) entspricht (siehe Ryals u. a., 1985).

- e) Hexamersequenzen C1, C2, C3, C4 und C5A.

Die Sequenzen C1 und C5A sind die oben beschriebenen. Die Sequenzen C2, C3 und C4 sind die in Cell, 49, 352–367 (1987) beschriebenen. Sie stellen die Sequenzen

AAATGA	– C2
AAGGGA	– C3 und
AAAGGA	– C4

dar.

- f) die IFN- β -Genesequenz von +19 bis –66.

Die linke Tafel in Fig. 3 zeigt die Ergebnisse, die erhalten wurden, wenn die Hexamerwiederholungen als Konkurrenten verwendet wurden. Die mittlere Tafel zeigt die Ergebnisse, wenn Human-IFN-Genesegmente als Konkurrenten eingesetzt wurden, und die rechte Tafel zeigt die entsprechenden Ergebnisse, wenn die in der Tafel angegebenen unterschiedlichen DNA-Segmente verwendet wurden.

Aus den Ergebnissen in Fig. 3 ist ersichtlich, daß das Erscheinen der verschobenen Bande durch die Hexamersequenzen in dieser Reihenfolge der Wirksamkeit C1–C2–C3–C4 verdrängt wurde, jedoch signifikant nicht verdrängt wurde durch C5A.

Es kann auch beobachtet werden, daß die synthetischen DNA-Segmente, die die Regulationselemente von entweder Human-IFN- α 1 oder Maus-H-2D^d-Genen umfassen, zu einer kompetitiven Aktivität Anlaß geben. Dies ist insbesondere interessant, da das DNA-Segment des H-2D^d-Gens die sogenannte IFN-Antwortsequenz (IRS) enthält, die als ein Verstärker fungiert, wenn die Zellen auf IFN reagieren (Sugita u. a., 1987; Israel u. a., 1986; Korber u. a., 1988).

Tatsächlich werden ähnliche oder gleiche Sequenzmotive wie auf dem IFN- β -Gen in vielen Promotorsequenzen der IFN-induzierbaren Gene gefunden, wo Kernfaktoren spezifisch zu binden scheinen (Korber u. a., 1988; Lery u. a., 1988).

Die in Fig. 3 gegebenen Ergebnisse ähneln sehr denen, die gewonnen werden, wenn der Test unter gleichen Bedingungen unter Verwendung von natürlichem, aus L929-Zellen produziertem Protein wiederholt wird.

Struktur der für Maus-IRF-I codierenden cDNA

Die DNA-Sequenz der klonierten DNAs wurde entweder durch die Didesoxymethode (SEQUENASE; United States Biochemical, Inc.) oder durch das Standardverfahren von Maxam und Gilbert, 1980) bestimmt.

Das λ L28-8-Insert in E. coli Y1089 wurde folgendermaßen isoliert: Die Phagen-DNA wurde durch das Standardverfahren hergestellt und die DNA wurde durch EcoRI digeriert, dann wurde die aus der Phagen-DNA ausgeschnittene cDNA mit Hilfe der Didesoxymethode (Sequenase: United States Biochemical Inc.) isoliert und sequenziert. Die Länge des cDNA-Inserts in λ L28-8 betrug 1,8 kb. Die Nucleotidsequenzanalyse zeigte einen großen offenen Leserahmen in Phase mit dem β -Galactosidgen.

Zum Screenen von Clonen, die größere cDNA-Inserts enthalten, wurde doppelsträngige cDNA synthetisiert mit L929-Zellen-stämmiger poly(A)⁺ RNA und in den Vektor CDM8 kloniert nach dem Verfahren von Aruffo und Seed (Proc. Natl. Acad. Sci., USA, Bd. 84, S. 8573, 1987) und Seed (Nature, Bd. 329, S. 840–842, 1987).

Die rekombinanten Plasmide wurden in den E. coli Stamm MC1061/p3 eingeführt (nach dem Verfahren von Aruffo und Seed, siehe oben), und die Clone der cDNA wurden mittels λ L28-8-stämmiger (³²P-markierter)-cDNA-Sonde unter strengen Bedingungen für DNA-DNA-Hybridisierung gescreent (Kashima u. a., Nature, Vol. 313, S. 402–404, 1985), und ein Clon pIRF-L wurde zur weiteren Untersuchung selektiert.

Das gewünschte cDNA-Insert von pIRF-L wurde durch Digestion mit HindIII und XbaI gewonnen und durch die oben beschriebenen Methoden sequenziert. Die Sequenz wird in Formel IV und in Fig. 4 gezeigt.

Es stellte sich heraus, daß die cDNA-Sequenz aus λ L28-8 eine identische Sequenz in der überlappenden Region enthielt, mit der Ausnahme, daß ein A-Rest zwischen den Nucleotiden 1773 und 1781 fehlte.

Die 5'- und 3'-Termini der λ L28-8-stämmigen cDNA sind durch Pfeile in Fig. 4 gekennzeichnet. Die Sequenzen ATTTATTTA und ATTTA, die möglicherweise die mRNA-Instabilität übertragen, sind eingerahmt.

Wie aus Fig. 4A ersichtlich wird, war die cDNA der Maus-cDNA, die von pIRF-L abgeleitet war, 198 bp lang und damit 20 bp länger als die von λ L28-8-cDNA in den 5'- bzw. 3'-Regionen.

Die Analyse der Genom-DNA-Sequenz, die die Promotorregion dieses Gens enthält, zeigt, daß die cDNA von pIRF-L etwa 30 bp von der größeren CAP-Stelle fehlt, siehe unten und Formel V und Fig. 7.

Die unmittelbare „upstream“-Sequenz des ersten ATG-Codons, GGACCATGC, paßt gut in die Kozaksche Consensus-Sequenz GCC₀CCATGG für die Translationsinitiationsstelle.

Eine ATG-Sequenz im Rahmen wurde in der „upstream“-Sequenz von der oben erwähnten ATG-Sequenz aus nicht gefunden, was bestätigt, daß dies tatsächlich das Initiationscodon für die IRF-I-mRNA ist.

Wie bereits erwähnt, wurde in der Nucleotidsequenz keine Differenz zwischen den cDNAs von λ L28-8 und pIRF-L innerhalb der überlappenden Regionen gefunden, mit Ausnahme eines Nucleotids in der 3'-nichtcodierenden Region.

Die Maus-IRF-I bestand damit aus 329 Aminosäuren mit einer berechneten relativen Molekülmasse von 37,3KD. Kanonische N-Glycosylierungsstellen erscheinen nicht in der Sequenz.

Zu anderen bekannten Proteinen wurde bei Recherchen in Protein Sequence Database (Natl. Biomed. Res. Found., Washington, D. C.) und bei erst kürzlich veröffentlichten Sequenzen keine signifikante Homologie nachgewiesen.

Die Hydropathieanalyse nach Kyte und Doolittle, 1982, zeigt, daß das Protein als Ganzes hoch hydrophil ist (Fig. 4B).

Die Untersuchung der abgeleiteten Primärsequenz des Maus-IRF-I zeigt die folgenden Merkmale:

Die endständige Aminohälfte, zur Aminosäure 140 verlängert, ist reich an Lysin (Lys) und Arginin (Arg). Tatsächlich sind 31 der 39 Lys- und Arg-Reste in dieser Region lokalisiert. Im unteren Teil von Fig. 4B ist im Diagramm eine Zusammenfassung der Lokalisierung der basischen Aminosäuren (Arg, Lys) (nach oben gerichtete Säulen) und der sauren Aminosäuren (Asp, Glu) (nach unten gerichtete Säulen) dargestellt.

Wie in Fig. 4B gezeigt, weist diese Region eine starke Hydrophilie auf und wird als die Region angesehen, die hauptsächlich für die IRF-I-Bindung an die spezifischen DNA-Sequenzen verantwortlich ist.

In diesem Zusammenhang waren die für viele DNA-bindenden Proteine charakteristischen Motive wie Zinkfinger und Helix/Helix-Motive (Pabo und Sauer, 1984; Evans und Hollenberg, 1988) in dem IRF-I-Protein nicht nachweisbar.

Im Gegensatz dazu zeigt der Rest des Moleküls (d. h. die endständige Carboxylhälfte einen relativen Überfluß an Asparaginsäure (Asp), Glutaminsäure (Glu), Serin (Ser) und Threonin (Thr). Von 189 Aminosäuren (von Aminosäure 140 bis 329) sind 33 (17%) saure Aminosäuren und 36 (19%) sind Ser und Thr. Beachtenswert ist, daß ein Cluster von 5 aufeinanderfolgenden sauren Aminosäuren in den Aminosäuren 227 bis 231 gefunden wurde. Ser und Thr scheinen viele Cluster zu bilden (Aminosäureregion bei 153–156, 190–192, 206–208, 220–222; die hier als die S-T-Regionen bezeichnet werden). Die S-T-Regionen werden durch kleine offene Rechtecke in der unteren Tafel von Fig. 4B dargestellt.

Struktur der Human-IRF-I codierenden cDNA

Nach der im vorangegangenen beschriebenen Verfahrensweise für Maus-IRF-I wurde Human-IRF-I-cDNA kloniert und sequenziert.

Eine Human-cDNA-Bibliothek wurde durch Synthetisierung von cDNA mittels poly(A)⁺ RNA aus einer Human-T-Zelllinie Jurkat hergestellt. Die doppelsträngige cDNA-Synthese und anschließende Clonierung in den Plasmidvektor CDM8 wurde entsprechend dem Verfahren von Aruffo und Seed (Proc. Natl. Acad. Sci., USA, Bd. 84, S. 8573–8577, 1987) und Seed (Nature, Bd. 329, S. 840–842, 1987) durchgeführt.

Die rekombinanten Plasmide wurden in den E. coli Stamm MC1061/p3 mittels des Verfahrens von Aruffo und Seed (siehe oben) eingeführt.

Die Clone, die mit Maus-IRF-I-cDNA kreuzhybridisieren, wurden zum Screenen mit λ L28-8-cDNA (³²P-markiert) als Sonde unter

wenig strengen Bedingungen für die DNA-DNA-Hybridisierung verwendet. Die angewandten Bedingungen waren genau die gleichen wie bei Kashima u. a. (Nature, Bd. 313, S. 402–404, 1985) beschrieben.

Aus den positiven Clonen wurde Clon pHIRF31, der das längste cDNA-Insert enthielt, selektioniert.

Die gewünschte cDNA-Sequenz von Clon pHIRF31 wurde durch Digestion der Plasmid-DNA durch XhoI isoliert und danach durch die oben beschriebenen Methoden sequenziert. Die Struktur des Human-IRF-I-Gens wird in Formel II und Fig. 8 gezeigt. Die Sequenzen für das abgeleitete Maus- und Human-IRF-I sind zum Vergleich in Fig. 5 untereinander dargestellt.

Die Analyse der Human-DNA zeigte, daß diese IRF-I um vier Aminosäuren kürzer ist als die Maus-IRF-I.

Die starke Erhaltung der Aminosäuresequenzen läßt sich zwischen den beiden IRF-I-Molekülen erkennen. Insbesondere 133 der 140 Aminosäuren (95%) der Aminoendhälfen können als gleich angesehen werden.

Zusammenfassend läßt sich sagen, daß die obigen Beobachtungen zeigen, daß IRF-I eine neue Klasse von DNA-bindendem Protein ist.

Es sei auch darauf hingewiesen, daß die Sequenzen ATTTATTTA und ATTTA, die in vielen Cytokin- und Protoonkogen-mRNAs gefunden werden, in der 3'-nichttranslatierten Region der Maus-IRF-I-cDNA vorhanden sind und daß auch die Sequenz ATTTA in der entsprechenden Region der Human-IRF-I-cDNA gefunden wird. Von diesen Sequenzen nimmt man an, daß sie in der posttranskriptionalen Regulation der Genexpression eine Rolle spielen, indem sie der mRNA Instabilität übertragen (Shaw und Kamen, 1986; Caput u. a., 1986).

Das Plasmid pIRF-L wurde in E. coli MC 1061/p3 transfektiert, das als E. coli MC 106/p3 (pIRF-L) am Fermentation Research Institute Agency of Industrial Science and Technology (FRI), 1–3, Higashi 1-chome, Tsukuba-shi, Ibaraki-ken 305, Japan, entsprechend den Bedingungen des Budapester Vertrages am 19. August 1988 unter Nr. Ferm BP-2005 hinterlegt wurde. Das Plasmid pIRF31 wurde in gleicher Weise in E. coli MC 1061/p3 transfektiert, das als E. coli MC 106/p3 (pIRF-31) am FRI entsprechend den Bedingungen des Budapester Vertrages am 19. August 1988 unter Nr. Ferm BP-2006 hinterlegt wurde.

Regulation des IRF-Gens

1. Expression der IRF-I-mRNA

In Anbetracht der Tatsache, daß IRF-I Affinitäten zu Regulationssequenzen von anderen Genen als dem IFN- β -Gen beweist und somit an der Regulation einer Reihe von Genen in verschiedenen Zelltypen beteiligt ist, wurde die Expression der IRF-I-mRNA in Mauszellen, die aus verschiedenen Geweben und Organen stammten, mit Maus-cDNA als Sonde untersucht. Zur Herstellung dieser Sonde wurde M13 mp10 Phagen-DNA (siehe unten), die den codierenden Strang von IRF-I-Gen-PstI-Fragment enthielt, als Matrize zur Synthetisierung der ^{32} P-markierten „anti-sense“ DNA verwendet. Das Produkt wurde durch EcoRI digeriert, und die Sonden-DNA wurde wie bei Fujita u. a. (1985) beschrieben isoliert.

Die gesamte RNA wurde nach dem bewährten Verfahren von Aruffo und Seed, 1987, isoliert.

Die Blot-Analyse wurde dann im wesentlichen wie bei Thomas (1980) beschrieben durchgeführt, der Röntgenfilm wurde 3 Tage belichtet, und die Ergebnisse sind in Fig. 6A gezeigt. Die verschiedenen Bahnen stellen die Ergebnisse der Durchgänge dar, bei denen Ganzzell-RNA aus den folgenden Geweben verwendet wurde:

Bahn 1	Hirn
Bahn 2	Herz
Bahn 3	Leber
Bahn 4	Lunge
Bahn 5	Milz (nicht stimuliert)
Bahn 6	Thymus
Bahn 7	Niere
Bahn 8	Muskel
Bahn 9	Darm
Bahn 10	Milz (nicht stimuliert)
Bahn 11	ConA – stimulierte Milz

In den Durchgängen für jede Bahn wurden 5 μ g der Ganzzell-RNA verwendet, mit Ausnahme von Bahn 8, bei der nur 1,2 μ g RNA verwendet wurden.

Aus Fig. 6A ist ersichtlich, daß eine Bande, die etwa 2,0kb entsprach, in den meisten RNA-Proben durch diese Blot-Analyse nachgewiesen wurde, obwohl der mRNA-Expressionswert niedrig zu sein schien. Es ist bemerkenswert, daß der Expressionswert von mRNA in den Lymphozyten aus der Milz nach der Stimulierung durch ConA (Bahn 11) drastisch zunahm. In einem weiteren Test wurden Maus-L 929-Zellen durch NDV induziert wie im vorangegangenen beschrieben (Fujita u. a., 1985), und die Zytoplasma-RNA wurde nach dem Verfahren von Aruffo und Seed, 1987, alle drei Stunden nach der Infektion extrahiert. Sonden-DNAs wurden aus den folgenden verschiedenen Sequenzen hergestellt und durch die „multiprime“-Markierungsreaktion (Amersham) markiert, und zwar

(i) ein 1,8kb EcoRI-Fragment aus λ L 28-8 (spezifische Aktivität 2×10^8 cpm/ μ g);

(ii) ein 0,5kb BamHI-BglII-Fragment aus einem Maus-IFN- β -Genomclon (spezifische Aktivität 5×10^8 cpm/ μ g) und

(iii) ein 2,0kb BamHI-PvuII-Fragment eines Clons, das ein Human- β -Actinpseudogen enthält (spezifische Aktivität 5×10^8 cpm/ μ g).

Die Ergebnisse sind in Fig. 6B dargestellt. Die Blot-Analyse wurde wie oben beschrieben durchgeführt, wobei das Verfahren von Thomas (1980) angewandt wurde.

Jede Bahn erhielt 10 μ g Zytoplasma-RNA. Der Röntgenfilm wurde 3 Stunden belichtet. Die densitometrische Analyse zeigte, daß die IRF-I-mRNA 9–12 Stunden nach der NDV-Infektion um das 25fache zunahm.

Obwohl die Zunahme an mRNA drastisch ist, ist es ein transientes Ergebnis, es erreicht seinen Höhepunkt bei 9 bis 12 Stunden und gleicht sich 12 Stunden nach der Induktion aus. Die mRNA-Anreicherung geht der Anreicherung an IFN- β -mRNA voraus; wie aus Fig. 6B ersichtlich ist, läßt sich die Induktion von IRF-I-mRNA bereits 3 Stunden nach der NDV-Infektion beobachten, während die IFN- β -mRNA bei gleichen Blot-Bedingungen für beide RNAs erst 6 Stunden danach nachweisbar ist.

Der IRF-I-Promotor

Wie im vorangegangenen demonstriert wurde, wird das IRF-I-Gen transkriptionell durch verschiedene Agenzien wie Viren und Mitogene reguliert.

Die Southern-Blot-Analyse der Chromosomen-DNA zeigte, daß das IRF-I-Gen gespleißt werden kann und nicht in der Maus „multimembered“ ist.

Eine λPhagenbibliothek, die Baby-Maus-DNA enthielt, wurde nach den Clonen gescreent, die die IRF-I-Promotorsequenz enthalten, wobei die gleiche λL28-8-stämmige cDNA-Sonde wie oben verwendet wurde. Es wurden vier positive Clone identifiziert, von denen bei allen die gleiche Genom-DNA festgestellt wurde und eine davon für die weitere Analyse verwendet wurde (λg 14-2). Ein PstI-Fragment wurde in die PstI-Stelle von pUC 19 subcloniert, um p19IRFP zu konstruieren.

Die gleiche DNA wurde danach in die PstI-Stelle von M13mp10 und M13mp11 cloniert, die zur Erzeugung von DNA für die Sequenzanalyse verwendet wurden.

Die Nucleotidsequenzanalyse des PstI-Fragments aus den obengenannten Clonen wurde wie bereits beschrieben durchgeführt. Große und kleine CAP-Stellen wurden durch SI-Kartierungsanalyse identifiziert.

Die bestimmte Sequenz wird in Fig. 7A gezeigt. Wie daraus ersichtlich ist, paßt die „downstream“-Sequenz der DNA ausgezeichnet zu der von pIRF-L abgeleiteten cDNA.

Die S1-Nucleaseanalyse zeigt das Vorhandensein von zwei CAP-Stellen für die IRF-I-mRNA, in der sich die größere Stelle etwa 20 Nucleotide „downstream“ von der kleineren Stelle befindet. Typische TATA-Boxsequenzen sind in der „upstream“-Region des Gens nicht vorhanden. In Anbetracht der ungewöhnlichen Fülle an CpG-Sequenzen, stellt diese Region wahrscheinlich eine „HTF-Insel“ dar (Bird, 1986).

Die Promotorregion enthält zwei CC-Boxen und eine CAAT-Box (siehe Fig. 7a); die ersteren Boxen müßten Spl binden (Kadogan u. a., 1986) und die letzteren CP-1 oder CP-2 (Chodosh u. a., 1988).

Das die Promotorsequenzen enthaltende PstI-Fragment wurde dann auf folgende Art und Weise auf seine Reaktivität als Antwort auf extrazelluläre Signale, z. B. Virusinduzierbarkeit, getestet.

Ein chimäres Gen wurde konstruiert, in dem ein Reportergen, und zwar ein bakterielles Chloramphenicolacetyltransferase-(CAT)-Gen „downstream“ vom PstI-Segment anstieß. Dies geschah durch Ausschneiden eines PstI-Fragmentes aus P 19IRFP (siehe oben) durch BamHI und HindIII und Clonierung des erhaltenen Fragmentes in das BglII-HindIII-Rückgratfragment von pA₁₀cat₂ (Rosenthal u. a., 1983), um pIRFcat herzustellen.

Verschiedene weitere Konstrukte wurden folgendermaßen hergestellt:

pIRF Δcat wurde durch Digerieren des p 19IRFP-stämmigen BamHI-HindIII-Fragmentes mit HaeIII hergestellt, dessen einzelne Erkennungsstelle bei –30 bis –35 von der größeren CAP-Stelle (Fig. 5A) liegt. Das entstandene HaeIII-HindIII-Fragment wurde mit BglII-HindIII-Rückgratfragment von pA₁₀cat₂ und der folgenden synthetischen DNA ligiert:

5'GATCCTAGATTCTTCGCGGCGC 3'
3'GATCTAAAGAAGCGC 5'

Damit enthielten sowohl pIRFcat als auch pIRFΔcat Sequenzen bis zu –320 bzw. –48 von der größeren CAP-Stelle aus.

p-125cat enthält die Promotorsequenz des Human-IFN-β-Gens wie bei Fujita u. a., 1987, beschrieben.

pSV2cat ist in Gorman u. a., Science, Bd. 221, S. 551–553 beschrieben.

Als Referenzgen wurde pRSVgpt verwendet (siehe Gorman u. a., s. o.).

Die unterschiedlichen Gene wurden mittels der Calcium-Phosphat-Methode (Fujita u. a., 1985) in Maus-L929-Zellen transfektiert. 5 × 10⁶ Zellen wurden mit 7,5 µg des CAT-Reportergen und 2,5 µg pRSVgpt enthaltenden Testplasmids transfektiert. Die Zellen wurden durch NDV induziert oder „mock“-induziert und anschließend dem Enzym-Test wie bei Fujita u. a., 1985, beschrieben unterzogen.

Zur Berechnung der relativen CAT-Aktivität wurde die CAT-Aktivität aus den „mock“-induzierten Zellen, die mit pSV2cat transfektiert waren, als 100% angenommen. Jede CAT-Aktivität wurde durch Ecogpt der entsprechenden Proben normalisiert (Mulligan und Berg, Proc. Natl. Acad. Sci. USA, Bd. 78, S. 2072–2076). Die Proben, in denen CAT unter dem Hintergrundwert lag, wurden mit b. b. (below background level) gekennzeichnet.

Die Ergebnisse sind in Fig. 78 dargestellt.

Es ist erkennbar, daß die Transfektion von pIRFcat in Maus-L929 die Expression der CAT-Aktivität mit niedrigem Wert verursachte. Der CAT-Expressionswert wurde erhöht, wenn die transfektierten Zellen durch NDV stimuliert wurden. Die Deletion der 300-bp-„upstream“-Sequenz des IRF-I-Gens (pIRF Δ cat) hob faktisch sowohl die konstitutive als auch die induzierte Expression des CAT-Gens auf. Dies zeigt, daß die Promotorsequenz innerhalb der 300-bp-„upstream“-Region liegt und virusinduzierbar ist.

Konstruktion von Expressionsplasmiden

1. Phagen-DNA von Clon λL28-8 wurde durch EcoRI digiert, und das Insert wurde gewonnen. Die ecoRI-Stellen der cDNA wurden durch T4-DNA-Polymerase geglättet und anschließend mit synthetischen Adaptor-DNAs, die die Sequenz pGATCCATTGTGCTGG und pCCAGCAACAATG besaßen, ligiert, s. Aruffo und Seed 1987.

Nach Entfernung der synthetischen DNAs durch gradiente Zentrifugierung mit 5–20% Kaliumacetat (Aruffo und Seed, 1987) wurde die IRF-I-cDNA mit den an beide Enden geknüpften Adaptor-DNAs mit BstXI-geschnittener CDM8-Vektor-DNA ligiert (Seed, B, Nature, Bd. 329, S. 840–842, 1987). Die Plasmide pIRF-S und pIRF-A, die die IRF-I-cDNA in der „sense“- und „antisense“-Ausrichtung in bezug auf den CMV-Promotor enthielten, wurden isoliert.

Jede Plasmid-DNA wurde entweder mit p-55cat oder p55CIB (Fujita u. a., 1987) in L929-Zellen co-transfektiert und der CAT-Expressionswert wurde bestimmt.

Die Ergebnisse sind in Tabelle 1 unten dargestellt:

Tabelle 1

	Transfizierte Plasmide	Induktion durch NDV	CAT-Aktivität (% Umwandlung)
Experiment 1	pIRF-S	–	< 1 %
	p-55cat		
	pIRF-S	–	50 %
	p-55CIB		
	pIRF-A	–	< 1 %
	p-55cat		
Experiment 2	pIRF-A	–	< 1 %
	p-55CIB		
	pIRF-S	–	1,7 %
	p-55CIB		
	pIRF-S	+	3,6 %
	p-55CIB		
	pIRF-A	–	< 0,1 %
	p-55CIB		
	pIRF-A	+	< 0,1 %
	p-55CIB		

Die DNA-Transfektionseffizienz hängt vom Zustand der Rezipientenzellen (in diesem Fall Maus-L929-Zellen) ab. Die Effizienz war in Experiment 2 viel niedriger als in Experiment 1, der CAT-Expressionswert ist daher vergleichsweise niedriger als in Experiment 2.

Wie aus der obigen Tabelle ersichtlich ist, war die signifikante CAT-Aktivität nur in durch p-55CIB und p-IRF-S transfizierten Zellen nachweisbar. Sie zeigen, daß der IRF-I an die wiederholten (8mal) AAGTGA-Sequenzen, die „upstream“ des CAT-Gens in p-55CIB vorhanden sind, bindet und daher die Transkription des distalen CAT-Gens fördert.

Die Ergebnisse zeigen weiterhin, daß der CAT-Expressionswert durch Infektion der transfizierten Zellen mit NDV (Tabelle 1) erhöht wird (um mehr als das Doppelte), was zeigt, daß die Kontrolle der Genexpression durch unterschiedliche Stimuli wie beispielsweise Viren möglich ist.

2. Ein Expressionsplasmid für die Produktion eines Proteins, das aus IRF-I-DNA bindender Domäne und Transkriptionsaktivierungsdomäne von Hefe-GAL 4 bestand, wurde folgendermaßen konstruiert:

Das Plasmid pIRF-S wurde durch HindIII und PstI digeriert, und das cDNA-Insert wurde isoliert. Die cDNA wurde durch DraIII digeriert, und das HindIII-DraIII-Fragment (etwa 550 bp) wurde gewonnen und als Fragment A bezeichnet.

Der Expressionsvektor CDM8 wurde durch HindIII und XbaI digeriert, und die Rückgrat-DNA wurde isoliert und als Fragment B bezeichnet.

Die für die Hefe-GAL4-Transkriptionsaktivierungsdomäne codierende DNA wurde aus Plasmid pRB968 folgendermaßen isoliert (Ma und Ptashne, 1987):

Die pRB968-DNA wurde zuerst mit HindIII digeriert, und die Enden wurden durch T4-DNA-Polymerase geglättet. Synthetische Xba-I-Linker-DNA wurde zu der DNA gegeben, und die DNA wurde dann durch PvuII und XbaI digeriert.

Das entstandene ca. 600 bp lange PvuII-XbaI-DNA-Fragment wurde gewonnen und als Fragment C bezeichnet.

Zusätzlich wurde eine synthetische DNA mit der folgenden Sequenz hergestellt:

(5')GTGTCGTCCAG(3')
(3')GCACACAGCAGGTC(5')

und als Fragment D bezeichnet.

Ein Expressionsvektor pIRFGAL4 wurde durch Ligation der Fragmente A, B, C und D konstruiert. Als Kontrollplasmid wurde Plasmid pIRFGAL4 durch Ligation der Fragmente A, B, C und einer synthetischen DNA mit der folgenden Sequenz konstruiert:

(5')GTGTCTGACAG(3')
(3')GCACACAGACTGTC(5')

Da ein Terminatortriplett, TGA, im Rahmen zwischen den IRF-I und GAL4-Sequenzen in pIRFGAL4 vorhanden ist, müßte dem exprimierten Protein die GAL4-Aktivierungsdomäne fehlen.

Um die funktionellen Eigenschaften des Plasmids zu testen, das für den chimären Transkriptionsfaktor codiert, wurden pIRFGAL4 und pIRFGAL4 beide mit p-55CIB in L929-Zellen cotransfiziert, und die CAT-Expression wurde überwacht. Die Ergebnisse sind in Tabelle 2 unten dargestellt.

Tabelle 2

Transfiziertes Plasmid	CAT-Aktivität (% Umwandlung)
pIRF-S	2,0 %
p-55CIB	
pIRF-4	< 0,2 %
p-55CIB	
pIRFGAL	1,4 %
p-55CIB	
pIRFGAL4	< 0,2 %
p-55CIB	

Wirtszellen waren Maus-L929-Zellen. Die Zellen waren nicht durch NDV induziert.

Die Ergebnisse zeigen, daß die Expression eines Zielgens (wie beispielsweise Gene, die für ein Interleukin, ein Interferon (α , β und γ), einen plasminogenen Aktivator, Erythropoietin, Granulozyt-kolonie-stimulierenden Faktor, Insulin, Human-Wachstumshormon oder Superoxiddismutase (oder Mutanten der Humangene) durch IRF-I erhöht werden kann.

Solche Zielgene wie die oben erwähnten, wie Interferongene, z. B. IFN- α , IFN- β , IFN- γ , IFN- ω und plasminogene Aktivatoren, z. B. t-PA pro-Urokinase oder Urokinase usw. können auch wirkungsvoller exprimiert werden, indem Promotoren eingeführt werden, die mit unterschiedlich langen Erkennungssequenzen für IRF-I, z. B. AAGTGA, verschmolzen sind.

Die Zielgene können beispielsweise in unterschiedliche Wirtszellen zusammen mit entweder intakten IRF-I- oder chimären IRF-I-Genen eingeführt werden. Durch Verlängerung der IRF-I-Erkennungsstellen-DNA, z. B. durch Erhöhung der Anzahl der AAGTGA-Wiederholungen und des Expressionswertes des Transkriptionsfaktors, läßt sich ein hoher Expressionswert der Zielgene erreichen.

Beispielsweise können die AAGTGA-Wiederholungssequenzen an einen geeigneten Promotor wie einen IFN- β -Promotor oder einen frühen SV40-Virus-Promotor angrenzen. Ein Zielgen z. B. ein t-PA-Gen oder IFN- β -Gen kann „downstream“ eines solchen Promotors verbunden werden; die Struktur eines derartig konstruierten Gens würde sein

(AAGTGA)_x (Promotor) (Zielgen z. B. t-PA-Ge.)

Ein derartiges Gen könnte dann in CHO-Zellen, z. B. CHO DXBII-Zellen (dhfr-Stamm) eingeführt und in diesen verstärkt werden (Urlaub und Chasin, Proc. Natl., Acad. Sci., USA, Bd. 77, S. 4216–4220, 1980).

Idealerweise wird, wie im vorangegangenen diskutiert wurde, eine Wirtszelle aus einer Zelle, die im wesentlichen keinen Wert einer endogenen IRF-I-Aktivität aufweist, gewählt. Das IRF-I-Gen, vorzugsweise entweder mit einem starken Promotor wie einem CMV-Promotor oder einem induzierbaren Promotor wie einem Metallothionein-Genpromotor, kann auf herkömmliche Art und Weise in die verschiedenen Wirtszellen eingeführt werden.

Das IRF-I-Gen kann co-eingeführt und zusammen mit dem Zielgen verstärkt werden. Auf alternative Weise können das IRF-I-Gen und das Zielgen getrennt in die Wirtszelle eingeführt werden.

In derartigen transfektierten Zellen kann IRF-I entweder konstitutiv (im Falle z. B. eines CMV-Promotors) oder induzierbar (im Fall z. B. des Metallothionein-Promotors wird durch zweiwertige Metalle wie Zink induziert) hergestellt werden. Der exprimierte IRF-I bindet an die AAGTGA-Wiederholungen und verstärkt das distale Zielgen, z. B. das t-PA-Gen oder das IFN-Gen.

Eine derartige Expression kann weiter durch ein Virus, z. B. NDV-Induktion, verstärkt werden, und wie aus Tabelle 1, Experiment 2, ersichtlich ist, erhöht diese Induktion die Aktivität des IRF-I.

Maus-cDNA-Sequenz, die mit Maus-IRF-I-cDNA der Formel III kreuzhybridisiert

Eine Maus-cDNA-Bibliothek wurde nach der im vorangegangenen beschriebenen Verfahrensweise hergestellt. Die cDNA wurde durch das Standardverfahren synthetisiert. Dabei wurde Maus-L929-Zellen-stämmige mRNA verwendet, und die cDNA wurde durch das Standardverfahren in den λ gt11-Vektor inseriert. Die entstandene λ gt11-Bibliothek wurde dann zur Isolierung der cDNA-Clone gescreent, deren Inserte mit der oben beschriebenen Maus-IRF-I-cDNA folgendermaßen kreuzhybridisieren: Nitrocellulosefilter, die die Phagen-Plaques-DNAs enthalten, wurden in den folgenden Stufen inkubiert

(1) in 3 $\times$ SCC bei 65°C 30 Minuten lang und anschließend

(2) 60minütige Inkubation in 3 $\times$ SSC, das Denhartsche Lösung enthält (0,2% Rinderserumalbumin, 0,2% Ficoll, 0,2% Polyvinylpyrrolidon 25), und anschließend

(3) eine vor-Hybridisierungsstufe, die aus der 12stündigen Inkubation bei 65°C in einer Lösung, die 1 M NaCl, 50 mM Tris-HCl, pH 8,0, 10 mM EDTA, 0,1% SDS, 50 μ g/ml einzelsträngige Träger-DNA (z. B. Lacksperma-DNA) und Denhartsche Lösung enthält, besteht und

(4) die Inkubation der Stufe 3 wurde wiederholt, wobei ³²P-markierte Maus-IRF-I-cDNA als Sonde eingeführt wurde. Diese cDNA-Sonde wurde als das von EcoRI geschnittene Insert aus λ L-28-8 hergestellt und durch das Multiprime-Markierungssystem (siehe oben) translatiert. Die Inkubation wurde 12 Stunden lang bei 65°C durchgeführt.

Dann wurden die Filter gewaschen, kurz in 2 $\times$ SCC-Lösung gespült und anschließend in 3 $\times$ SCC-Lösung, die 0,1% SDS enthielt, 30 Minuten lang bei 65°C gewaschen. Diese Prozedur wurde zweimal wiederholt.

Einer der positiven Clone, der als pHH-45 bezeichnet wurde, wurde selektioniert und es zeigte sich, daß er cDNA enthielt, die nur einen Teil der Codiersequenz für einen IRF abdeckte.

Das cDNA-Insert in pHH-45 wurde deshalb isoliert und zum Screenen von Clonen verwendet, die größere Inserte enthielten, wie im vorangegangenen unter der Überschrift „Struktur der für Maus-IRF-I codierenden cDNA“ für die Herstellung von pIRF-L beschrieben wurde.

Von den identifizierten positiven Clonen wurde der als pIRF 2–5 bezeichnete selektioniert und charakterisiert, wobei das im vorangegangenen für Maus-IRF-I beschriebene Verfahren angewandt wurde. Die komplette hybridisierende cDNA-Sequenz wird in Formel IV a gezeigt, und die Aminosäuresequenz des entsprechenden IRF-Proteins wird in Formel VIII dargestellt.

Das Plasmid pIRF 2–5 wurde in E. coli MC 1061/p3 transfektiert, das als E. coli am FRI entsprechend dem Budapest Vertrag am 22. November 1988 unter der Nr. FERM BP-2157 hinterlegt wurde.

cDNA aus dem Hefegenom, die mit Human-IRF-I-cDNA-Sequenz hybridisiert

Hefe-DNA wurde nach dem Standardverfahren hergestellt und mit EcoRI digeriert. 5 μ g der digerierten DNA wurde auf 0,8% Agarosegel gegeben und mittels Standardverfahren einer Elektrophorese und einem DNA-Blot-Test unterzogen.

Das geblottete Filter wurde bis auf folgende Ausnahme genau nach der Beschreibung im vorangegangenen Beispiel zur Isolierung von Maus-DNA, die mit Maus-IRF-I hybridisiert, behandelt:

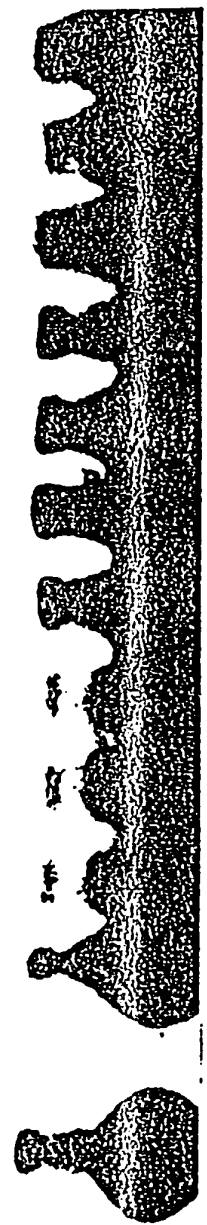
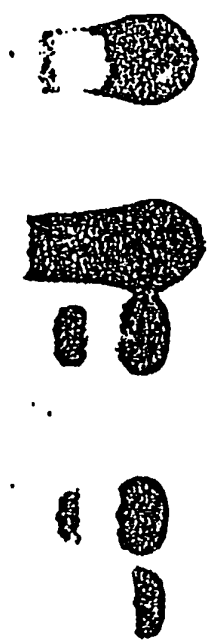
In Stufe (3) betrug die Inkubationstemperatur 55°C, und in Stufe (4) wurde die Inkubation ebenfalls bei 55°C ausgeführt, und die radioaktive Sonde war die aus pHIRF31 durch Digestion des Plasmids mit XhoI isolierte Human-IRF-I-cDNA, wobei diese Sonde wie im vorangegangenen Beispiel für Maus-IRF-I beschrieben markiert wurde. Das Filter wurde bei 55°C in 2 $\times$ SSC gewaschen. Die positiven Clone wurden durch Autoradiographie identifiziert (siehe Fig. 10.).

Literaturangaben

- Abreu, S. L., Bancroft, F. C., und Stewart, II E. W. (1979). Interferon priming. *J. Biol. Chem.* **254**, 414-418.
- Aruffo, A., und Seed, B. (1987). Molecular cloning of a cDNA by a high-efficiency COS cell expression system. *Proc. Natl. Acad. Sci. USA* **84**, 8573-8577.
- Bird, A. P. (1986). CpG-rich islands and the function of DNA methylation. *Nature* **321**, 209-213.
- Caput, D., Beutler, B., Hartog, K., Thayer, R., Brown-Schimer, S., und Cerami, A. (1986). Identification of a common nucleotide sequence in the 3' untranslated region of mRNA molecules specifying inflammatory mediators. *Proc. Natl. Acad. Sci. USA* **83**, 1670-1674.
- Cavalierj, R. L., Havell, E. Z., Vilcek, J., und Pestka, S. (1977). Induction and decay of human fibroblast interferon mRNA. *Proc. Natl. Acad. Sci. USA* **74**, 4415-4419.
- Chodosh, L. A., Baldwin, A. S., Carthew, R. W., und Sharp, P. A. (1988). Human CCAAT-binding proteins have heterologous subunits. *Cell* **53**, 11-24.
- Dinter, H., Hauser, H., Mayr, U., Lammers, R., Bruns, W., Gross, G., und Collins, J. (1983). Human interferon-beta and co-induced genes: molecular studies. In *the biology of the Interferon System 1983*, E. De Maeyer and H. Schellekens, hrag. (Amsterdam: Elsevier Science Publishers), 33-34.
- Dinter, H., und Hauser, H. (1987). Cooperative interaction of multiple DNA elements in the human interferon- β promoter. *Eur. J. Biochem.* **166**, 103-109.
- Evans, R. M., und Hollenberg, S. M. (1988). Zinc Fingers: Gilt by association. *Cell* **52**, 1-3.
- Fujita, T., Saito, S., und Kohno, S. (1979). Priming increases the amount of interferon mRNA in poly(rI):poly(rC)-treated L cells. *J. Gen. Virol.* **45**, 301-308.
- Fujita, T., Ohno, S., Yasumitsu, H., und Taniguchi, T. (1985). Delimitation and properties of DNA sequences required for the regulated expression of human interferon- β gene. *Cell* **41**, 489-496.
- Fujita, T., Shibuya, H., Hotta, H., Yamanishi, K., und Taniguchi, T. (1987). Interferon- β gene regulation: Tandemly repeated sequences of a synthetic 6bp oligomer function as a virus-inducible enhancer. *Cell* **49**, 357-367.
- Galabru, J., und Hovanessian, A. G. (1985). Two interferon- β induced proteins are involved in the protein kinase complex dependent on double-stranded RNA. *Cell* **41**, 685-694.
- Goodbourn, S., Zinn, K., und Maniatis, T. (1985). Human β -interferon gene expression is regulated by an inducible enhancer element. *Cell* **41**, 509-520.
- Huynh, T. V., Young, R. A., und Davis, R. W. (1985). Constructing and screening cDNA libraries in gt 10 and 11. In *DNA cloning-A Practical Approach*, Bd. 1, D. M. Glover, hrag. (Oxford: IRL Press), S. 49-78.
- Israel, A., Kimura, A., Fournier, A., Fellous, M., und Kourilsky, P. (1986). Interferon response sequence potentiates activity of an enhancer in the promoter region of a mouse H-2 gene. *Nature* **322**, 743-746.
- Kadonaga, J. T., Jones, K. A., und Tjian, R. (1986). Promoter-specific activation of RNA polymerase II transcription by Sp 1. *Trends Biochem. Sci.* **11**, 20-23.
- Kakidani, H., und Ptashne, M. (1986). GAL 4 activates gene expression in mammalian cells. *Cell* **52**, 161-167.
- Keller, A. D., und Maniatis, T. (1988). Identification of an Inducible factor that binds to a positive regulatory element of the human β -interferon gene. *Proc. Natl. Acad. Sci. USA* **85**, 3309-3313.
- Kohase, M., May, L. T., Tamm, I., Vilcek, J., und Sehgal, P. B. (1987). A cytokine network in human diploid fibroblasts: interactions of beta interferon, tumor necrosis factor, platelet-derived growth factor and interleukin-1. *Mol. Cell. Biol.* **7**, 272-280.
- Korber, B., Mermod, N., Hood, L., und Stroynowski, I. (1988). Regulation of gene expression by interferons: Control of H-2 Promoter responses. *Science* **239**, 1302-1306.
- Kozak, M. (1987). An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucl. Acids Res.* **15**, 8125-8143.
- Krebs, E., Eisenman, R., Kuenzel, E., Litchfield, D., Lozeman, F., Lüscher, B., und Sommercorn, J. (1988). Casein Kinase II as a potentially important enzyme concerned with signal transduction. In *the Molecular Biology of Signal Transduction* (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory), Kurzfassung S.35
- Kuhl, D., de la Fuente, J., Chaturvedi, M., Parimoo, S., Ryals, J., Mayer, F., und Weissmann, C. (1987). Reversible silencing of enhancers by sequences derived from the human IFN- α promoter. *Cell* **50**, 1057-1069.
- Kyte, J., und Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**, 105-132.
- Levy, D. E., Kessler, D. S., Pine, R., Reich, N., und Darnell, J. E. (1988). Interferon-induced nuclear factors that bind a shared promoter element correlate with positive and negative transcriptional control. *Genes and Development* **2**, 383-393.
- Ma, J., und Ptashne, M. (1987). The carboxy-terminal 30 amino acids of GAL4 are recognised by GAL80. *Cell* **50**, 137-142.
- Maxam, A., und Gilbert, W. (1980). Sequencing end-labeled DNA with base specific chemical cleavages. *Meth. Enzym.* **65**, 499-560.
- Moore, R. N., Larsen, H. S., Horohov, D. W., und Rouse, B. T. (1984). Endogenous regulation of macrophase proliferative expansion by colony-stimulation-factor-induced interferon. *Science* **223**, 178-180.
- Nedwin, G., Naylor, S., Sakaguchi, A., Smith, D., Jarrett-Nedsin, J., Pennica, D., Goeddel, D., und Gray, P. (1985). Human lymphotoxin and tumor necrosis factor genes: structure, homology and chromosomal localisation. *Nucl. Acids Res.* **13**, 6361-6373.
- Nir, U., Cohen, B., Chen, L., und Revel, M. (1984). A human IFN- β gene deleted of promoter sequences upstream from the TATA box is controlled post-transcriptionally by dsRNA. *Nucl. Acids Res.* **12**, 6979-6993.
- Ohno, S., und Taniguchi, T. (1983). The 5'-flanking sequence of human interferon- β gene is responsible for viral induction of transcription. *Nucl. Acids Res.* **11**, 5403-5412.
- Onozaki, K., Urawa, H., Tamatani, T., Iwamura, Y., Hashimoto, T., Baba, T., Suzuki, H., Yamada, M., Yamamoto, S., Oppenheim, J. J., Matsushima, K. (1988). Synergistic interactions of interleukin 1, interferon- β and tumor necrosis factor in terminally differentiating a mouse myeloid leukemic cell line (M1).

- J. Immunol. **140**, 112–119.
- Pabo, C. O., und Sauer, R. T. (1984). Protein-DNA recognition. *Ann. Rev. Biochem.* **53**, 293–321.
- Raji, N. B. K., und Pitha, P. M. (1981). An analysis of interferon mRNA in human fibroblast cells induced to produce interferon. *Proc. Natl. Acad. Sci. USA* **78**, 7426–7430.
- Raji, N. B. K., und Pitha, P. M. (1983). Two levels of regulation of β -interferon gene expression in human cells. *Proc. Natl. Acad. Sci. USA* **80**, 3923–3927.
- Resnitzky, D., Yarden, A., Zipori, D., und Kimchi, A. (1986). Autocrine β -related interferon controls c-myc suppression and growth arrest during hematopoietic cell differentiation. *Cell* **46**, 31–40.
- Rosenthal, N., Kress, M., Gruss, P., und Khoury, G. (1983). BK viral enhancer element and a human cellular homolog. *Science* **222**, 749–755.
- Ryals, J., Dieks, P., Ragg, H., und Weissmann, C. (1985). A 46-nucleotide promoter segment from an INF- α gene renders an unrelated promoter inducible by virus. *Cell* **41**, 497–507.
- Shaw, G., und Kiinen, R. (1986). A conserved AU sequence from the 3' untranslated region of GM-CSF mRNA mediates selective mRNA degradation. *Cell* **46**, 659–667.
- Singh, H., LeBowitz, J. H., Baldwin, Jr. A. S., und Sharp, P. A. (1988). Molecular cloning of an enhancer binding protein: Isolation by screening of an expression library with a recognition site DNA. *Cell* **52**, 415–423.
- Sugita, K., Miyazaki, J. I., Appella, E., und Ozato, K. (1987). Interferons increase transcription of a major histocompatibility class I gene via a 5' interferon consensus sequence. *Mol. Cell. Biol.* **7**, 2625–2630.
- Taniguchi, T., Matsui, H., Fujita, T., Takaoka, C., Kashima, N., Yoshimoto, R., und Hamuro, J. (1983). Structure and expression of a cloned cDNA for human interleukin-2. *Nature* **302**, 305–310.
- Taniguchi, T. (1988), regulation of cytokine gene expression. *Ann. Rev. Immunol.* **6**, 439–464.
- Thomas, P. S. (1980). Hybridisation of denatured RNA and small DNA fragments transferred to nitrocellulose. *Proc. Natl. Acad. Sci. USA* **77**, 5201–5205.
- Tiwari, R. J., Kusari, J., und Sen, G. C. (1987). Functional equivalents of interferon-mediated signals needed for induction of a mRNA can be generated by double-stranded RNA and growth factors. *EMBO J.* **6**, 3373–3378.
- Warren, M. K., und Ralf, P. (1986). Macrophage growth factor CSF-1 stimulates human monocyte production of interferon, tumor necrosis factor and colony stimulating activity. *J. Immunol.* **137**, 2281–2285.
- Webster, N., Jin, J. R., Green, S., Hollis, M., und Chambon, P. (1988). The yeast UAGs is a transcriptional enhancer in human HeLa cells in the presence of the GAL4 trans-activator. *Cell* **52**, 169–178.
- Weissmann, C., und Weber, H. (1986). The interferon genes. *Proc. Natl. Acad. Res. Mol. Biol.* **33**, 251–300.
- Young, R. A., und Davis, R. W. (1983). Yeast RNA polymerase II genes: Isolation with antibody probes. *Science* **222**, 778–782.
- Zinn, K., Dimaio, D., und Maniatis, T. (1983). Identification of two distinct regulatory regions adjacent to the human β -interferon gene. *Cell* **34**, 865–879.
- Zullo, J. N., Cochran, B. H., Huang, A. S., und Stiles, C. D. (1985). Platelet-derived growth factor and double-stranded ribonucleic acids and stimulate expression of the same genes in 3T3 cells. *Cell* **43**, 793–800.

Figure 1



1 2 3 4 5 6 7 8 9 10 11 12

- 36 -

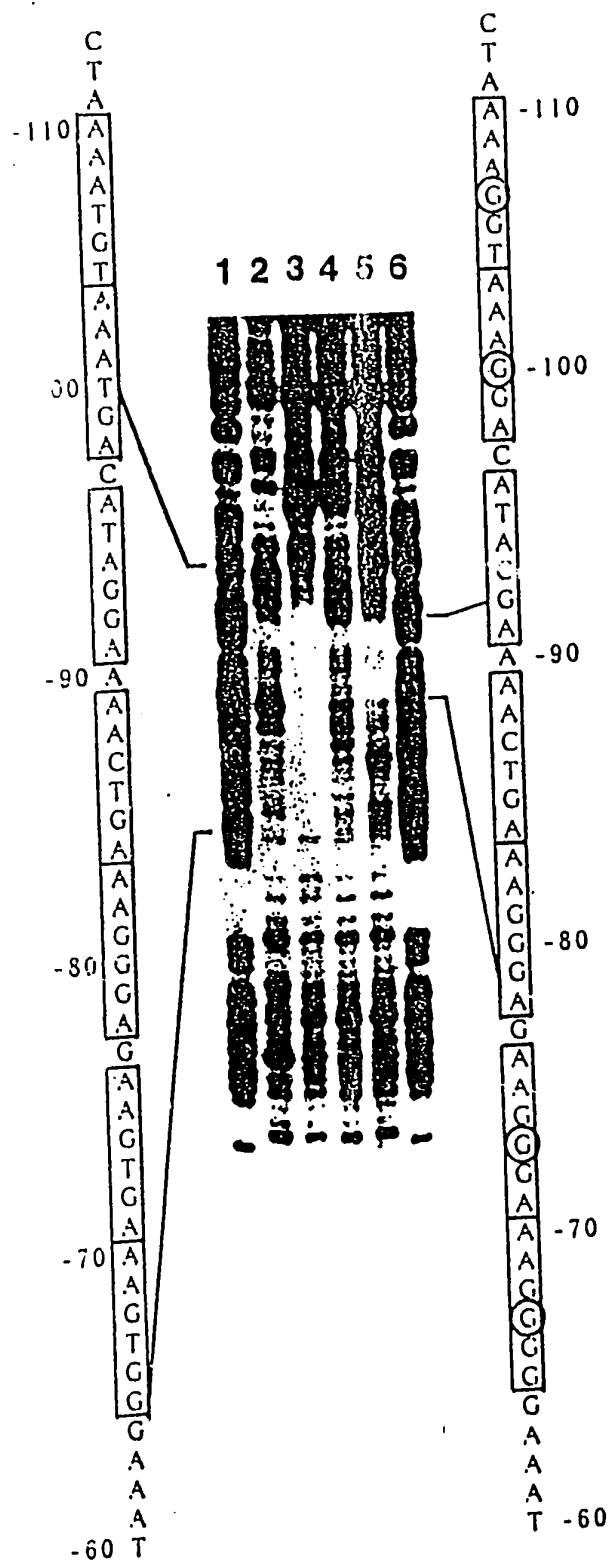


Figure 3.

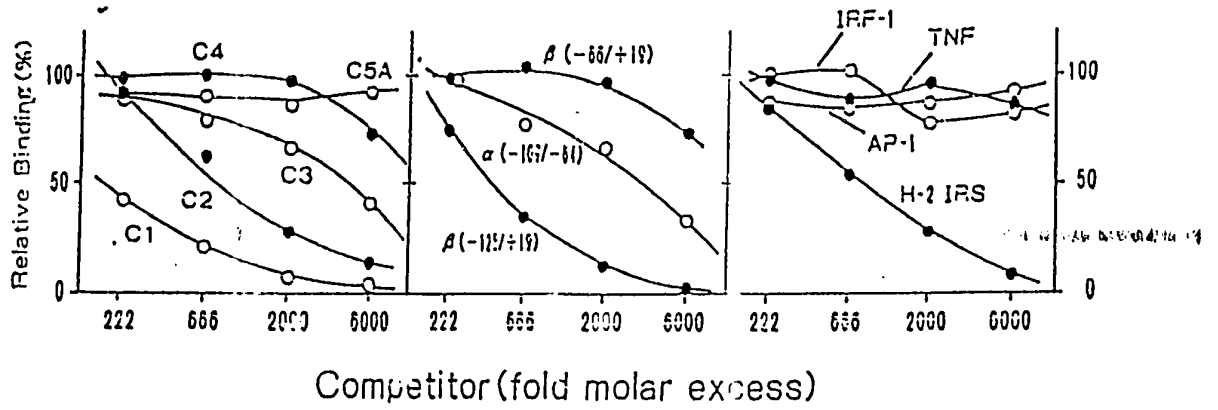


Figure 4 A

Page a

1 GGACGTGCTTTTCACAGTCTAAGCCGAACCGAACCGAACCGAACCGAACCGAACCGGGCC
 60 GAGTTGCGCCGAGGTCAGCCGAGGTGGCCAGAGGACCCAGCATCTCGGGCATCTTTTCG
 119 CTTCGTGCGCGCATCGCGTACCTACACCGCAACTCGTGCTCGCTCTCCGGCACCCCTC
 178 TGCGAATCGCTCCTGCAGCAA $\xrightarrow{\lambda L28-8}$ 1 Met Pro Ile Thr Arg Met Arg
 AGCCACC ATG CCA ATC ACT CGA ATG CGG
 20
 Met Arg Pro Trp Leu Glu Met Gln Ile Asn Ser Asn Gln Ile Pro
 227 ATG AGA CCC TGG CTA GAG ATG CAG ATT AAT TCC AAC CAA ATC CCA
 20
 Gly Leu Ile Trp Ile Asn Lys Glu Glu Met Ile Phe Gln Ile Pro
 272 GGG CTG ATC TGG ATC AAT AAA GAA GAG ATG ATC TTC CAG ATT CCA
 40
 Trp Lys His Ala Ala Lys His Gly Trp Asp Ile Asn Lys Asp Ala
 317 TGG AAG CAC GCT GCT AAG CAC GGC TGG GAC ATC AAC AAG GAT GCC
 60
 Cys Leu Phe Arg Ser Trp Ala Ile His Thr Gly Arg Tyr Lys Ala
 362 TGT CTG TTC CGG AGC TGG GCC ATT CAC ACA GGC CGA TAC AAA GCA
 80
 Gly Glu Lys Glu Pro Asp Pro Lys Thr Trp Lys Ala Asn Phe Arg
 407 GGA GAA AAA GAG CCA GAT CCC AAG ACA TGG AAG GCA AAC TTC CGT
 100
 Cys Ala Met Asn Ser Leu Pro Asp Ile Glu Glu Val Lys Asp Gln
 452 TGT GCC ATG AAC TCC CTG CCA GAC ATC GAG GAA GTG AAG GAT CAG
 120
 Ser Arg Asn Lys Gly Ser Ser Ala Val Arg Val Tyr Arg Met Leu
 497 AGT AGG AAC AAG GGC AGC TCT GCT GTG CCG GTG TAC CGG ATG CTG
 140
 Ser Arg Asp Thr Lys Ser Lys Thr Lys Arg Lys Leu Cys Gly Asp
 587 AGC CGA GAC ACT AAG AGC AAA ACC AAG AGG AAG CTG TGT GGA GAT
 160
 Val Ser Pro Asp Thr Phe Ser Asp Gly Leu Ser Ser Ser Thr Leu
 632 GTT AGC CCG GAC ACT TTC TCT GAT GGA CTC AGC AGC TCT ACC CTA

Figure 4 A

Page b

160
 Pro Asp Asp His Ser Ser Tyr Thr Thr Gln Gly Tyr Leu Gly Gln
 677 CCT GAT GAC CAC AGC AGT TAC ACC ACT CAG GGC TAC CTG GGT CAG

180
 Asp Leu Asp Met Glu Arg Asp Ile Thr Pro Ala Leu Ser Pro Cys
 722 GAC TTG GAT ATG GAA AGG GAC ATA ACT CCA GCA CTG TCA CCG TGT

200
 Val Val Ser Ser Ser Leu Ser Glu Trp His Met Gln Met Asp Ile
 767 GTC GTC AGC AGC AGT CTC TCT GAG TGG CAT ATG CAG ATG GAC ATT

Ile Pro Asp Ser Thr Thr Asp Leu Tyr Asn Leu Gln Val Ser Pro
 812 ATA CCA GAT AGC ACC ACT GAT CTG TAT AAC CTA CAG GTG TCA CCC

220
 Met Pro Ser Thr Ser Glu Ala Ala Thr Asp Glu Asp Glu Glu Gly
 857 ATG CCT TCC ACC TCC GAA GCC GCA ACA GAC GAG GAT GAG GAA GGG

240
 Lys Ile Ala Glu Asp Leu Met Lys Leu Phe Glu Gln Ser Glu Trp
 902 AAG ATA GCC GAA GAC CTT ATG AAG CTC TTT GAA CAG TCT GAG TGG

260
 Gln Pro Thr His Ile Asp Gly Lys Gly Tyr Leu Leu Asn Glu Pro
 947 CAG CCG ACA CAC ATC GAT GCC AAG GGA TAC TTG CTC AAT GAG CCA

Gly Thr Gln Leu Ser Ser Val Tyr Gly Asp Phe Ser Cys Lys Glu
 992 GGG ACC CAG CTC TCT TCT GTC TAT GGA GAC TTC AGC TGC AAA GAG

280
 Glu Pro Glu Ile Asp Ser Pro Arg Gly Asp Ile Gly Ile Gly Ile
 1037 GAA CCA GAG ATT GAC AGC CCT CGA GGG GAC ATT GGG ATA GGC ATA

300
 Gln His Val Phe Thr Glu Met Lys Asn Met Asp Ser Ile Met Trp
 1082 CAA CAT GTC TTC ACG GAG ATG AAG AAT ATG GAC TCC ATC ATG TGG

320
 Met Asp Ser Leu Leu Gly Asn Ser Val Arg Leu Pro Pro Ser Ile
 1127 ATG GAC AGC CTG CTG GGC AAC TCT GTG AGG CTG CCG CCC TCT ATT

329
 Gln Ala Ile Pro Cys Ala Pro
 1172 CAG GCC ATT CCT TGT GCA CCA TAG TTTGGGTCTCTGACCCGTTCTTGCCC

1222 TCCTGAGTGAGTTAGGCCTTGGCATCATGGTGGCTGTGATACAAAAAAGCTAGACTCC

1281 TGTGGGCCCTTGACACATGGCAAAGCATAGTCCCACTGCAAACAGGGGACCATCCTCC

Figure 4 A

Page c

1340 TTGGGTCAGTGGGCTCTCAGGGCTTAGGAGGCAGAGTCTGAGTTTCTTGTGAGGTGAA .
1399 GCTGGCCCTGACTCCTAGGAAGATGGATTGGGGGGTCTGAGGTGTAAGGCAGAGGCCAT
1458 GGACAGGAGTCATCTTCTAGCTTTTTTAAAAGCCTTGTTGCATAGAGAGGGTCTTATCGC
1517 TGGGCTGGCCCTGAGGGGAATAGACCAGCGCCACAGAAGAGCATTCGACTGGCCCTAG
1576 AGCTGGCTCTGTACTAGGAGACAATTGCACTAAATGAGTCCTATTCCCAAAGAACTGCT
1635 GCCCTTCCAACCGAGCCCTGGGATGGTTCCCAAGCCAGTGAAATGTGAAGGGAAAAAA
1694 AATGGGGTCCTGTGAAGGTTGGCTCCCTTAGCCTCAGAGGGAATCTGCCTCACTACCTG
1753 CTCCAGCTGTGGGGCTCAGGAAAAAAAAAATGGCACTTTCTCTGTGGACTTTGCCACATT
1812 TCTGATCAGAGGTGTACACTAACATTTCTCCCCAGTCTAGGCCTTTGCATTTATTTATA
1871 TAGTGCCTTGCCTGGTGCCTGCTGTCTCCTCAGGCCTTGGCAGTCCTCAGCAGGCCCAG
1930 GGAAAAGGGGGGTTGTGAGCGCCTTGGCGTGACTCTTGACTATCTATTAGAAACGCCAC
1989 CTAAGTGTAAATGGTGTTTGGTCATGTGGTGGACCTGTGTAAATATGTATATTTGTCT
2048 TTTTATAAAATTTAAGTTGTTTACAAAAAAAAA 2082

Fig. 4 B

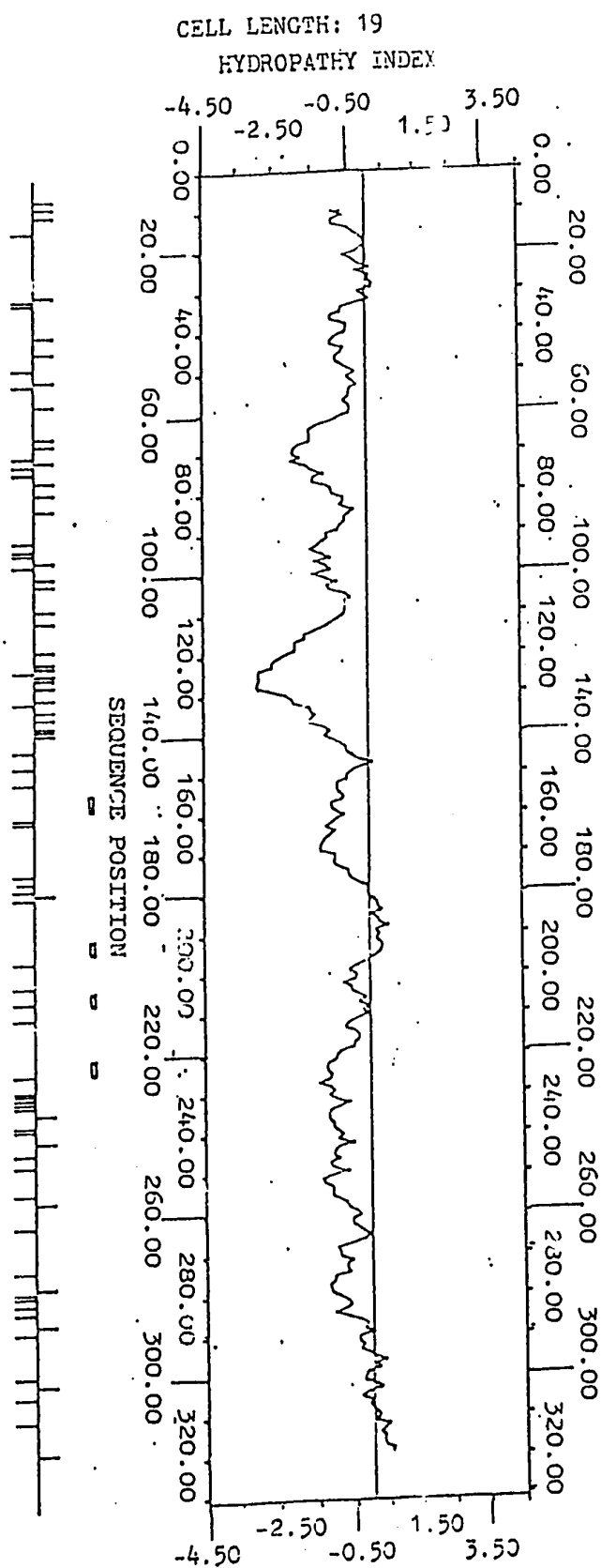


Figure 5/1

	1	25	50
MOUSE	MPITRMRMRPWLEMQINSNQIPGLIWINKHEMIFQIPWKHAAKHGWDINK		
	++++ ++++++		
HUMAN	MPITWMRMRPWLEMQINSNQIPGLIWINKHEMILEIPWKHAAKHGWDINK		

	75	100
	DACLFERSWAIHTGRYKAGEKEPDPKTKANFRFCAMNSLPDIEEVKDQSRN	
	++++ ++++++	
	DACLFERSWAIHTGRYKAGEKEPDPKTKANFRFCAMNSLPDIEEVKDQSRN	

	125	150
	KGSSAVRVYRMLPPLTRNQKQKSKSSRDTKSKTKRKLCDVSPDTFSD	
	++++ ++++++	
	KGSSAVRVYRMLPPLTKNQKQKSKSSRDAKSKAKRLSCGDSSPDTFSD	

	175	200
	GLSSSTLPDDHSSYTTQGYLGQDLMDERDITPALSPCVVSSSLSEWHMQM	
	++++ ++++++	
	GLSSSTLPDDHSSYTVPGYM-QDLEVEQALTPALSPCAVSSTLPDWHIPV	

	225	250
	DIIPDSTTDLYNLQVSPMPSTSEAATDEDEEGKIAEDLMKLFQSEWPQT	
	++++ ++++++	
	EVVPDSTSDLYNFQVSPMPSISEATDEDEEGKLPEDIMKLLEQSEWQPT	

	275	300
	HIDGKGYLLNEPGTQLSSVYGDVSCKEEPEIDSPRGDIGIGIQHVFTMK	
	++++ ++++++	
	NVDGKGYLLNEPGVQPTSVYGDVSCKEEPEIDSPGGDIGLSLQRVFTDLK	

	329
	NMDSIMWMDSLLGNSVRLPPSIQAIPCAP
	++++ ++++++
	NMDAT-WLDSLLTP-VRLP-SIQAIPCAP

Figur 5/2

Die erhaltenen Aminosäuren sind durch Sternchen gekennzeichnet. Die Sequenzen sind mit den Ein-Buchstaben-Abkürzungen für Aminosäuren dargestellt, in denen folgendes bedeutet: A - Alanin, C - Cystein, D - Asparaginsäure, E - Glutaminsäure, F - Phenylalanin, G - Glycin, H - Histidin, I - Isoleucin, K - Lysin, L - Leucin, M - Methionin, N - Asparagin, P - Prolin, Q - Glutamin, R - Arginin, S - Serin, T - Threonin, V - Valin, W - Tryptophan und Y - Tyrosin.

Figure 6

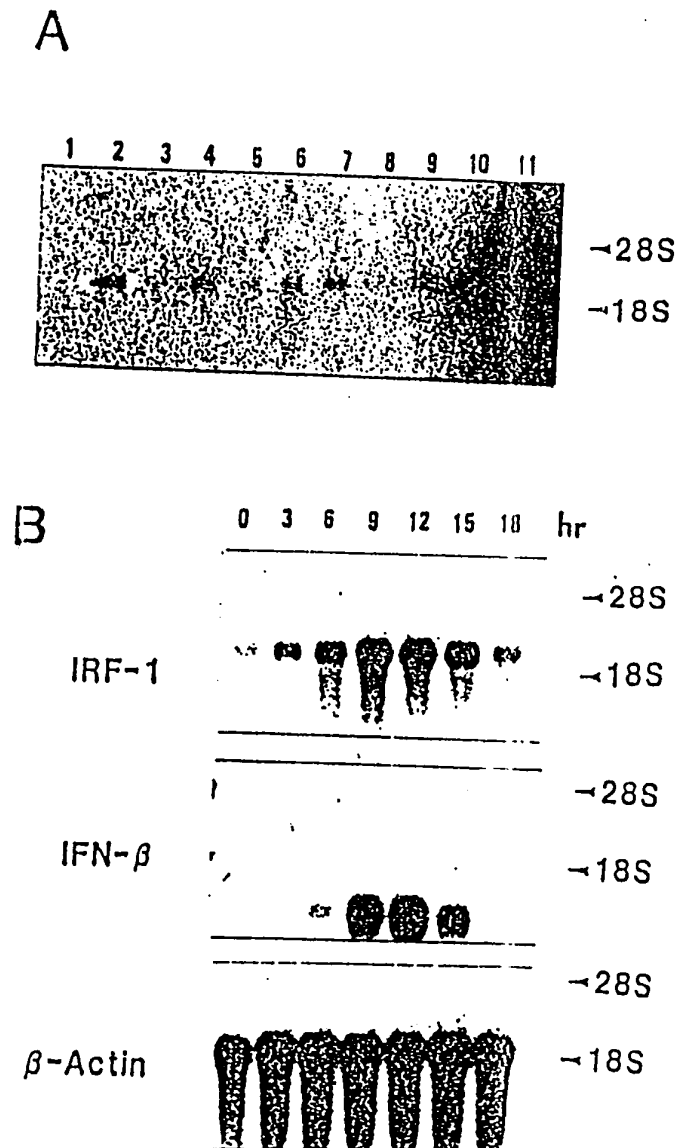


Figure 7 A

PstI
CTGCAGAAAGAGGGGGACGGTCTCGGCTTTCCAAGACAGGCAAGGGGG
 -299

GC Box 1 GC Box 2
 CAGGGGAGTGGAGTGGAGCAAGGGGCGGGCCCGCGGTAGCCCCGGGGCGGTGGCGCGG
 -251

GCCCGAGGGGGTGGGGAGCACAGCTGCCTTGTACTTCCCCTTCGCCGCTTAGCTCTAC
 -193

CAAT Box
 AACAGCCTGATTTCCCCGAAATGATGAGGCCGAGTGGGCAATGGGCGCGCAGGAGCG
 -135

Minor Cap site
 |||
 GCGCGGCGGGGGCGTGGCCGAGTCCGGGCCGGGGAATCCCGCTAAGTGTTTAGATTC
 -77 -21

Major Cap site
 ||| plRF-L
 TTCGGGCGCCCGGACTCGCCAGTGCGCACCACCTCCTTCGTCGAGGTAGGACGTGCT
 -19 +1

TTCACAGTCTAAGCCGAACCGAACCGAACCGAACCGAACCGAACCGGGCCGAGTTGCG
 +39

CCGAGGTCAGCCGAGGTGGCCAGAGGACCCAGCATCTCGGGCATCTTTCGCTTCGTG
 +97

CGCGCATCGCGTACCTACACCGCAACTCCGTGCCTCGCTCTCCGGCACCTCTGCGAA
 +155

PstI
 TCGCTCCTGCAG
 +213 +225

Figure 7 B

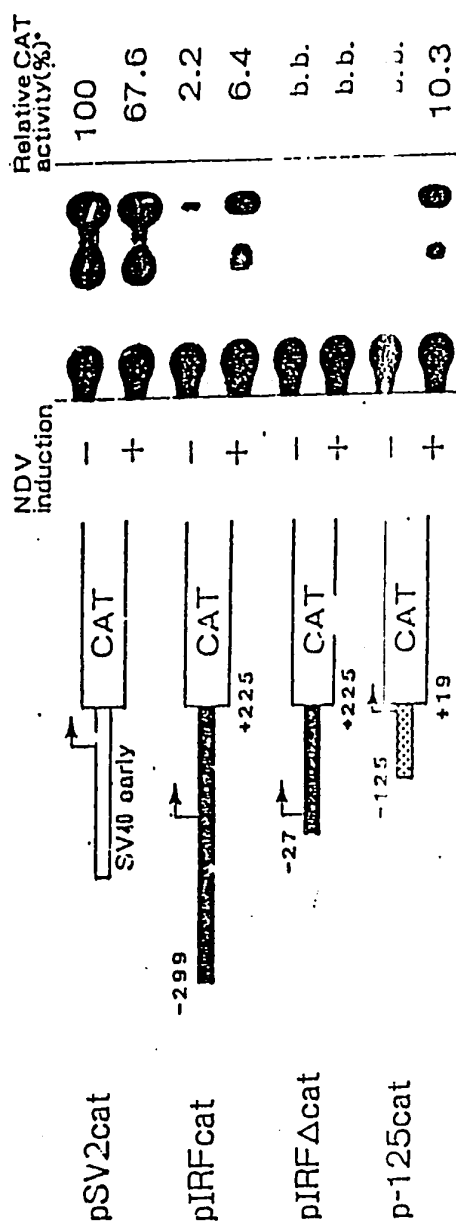


Figure 8

Page a

CGAGCCCCGCCGAACCGAGGCCACCCGGAGCCGTGCCAGTCCACGC
CGGCCGTGCCCCGGCGGCCTTAAGAACCAGGCAACCACTGCCTTCTCCCT
CTTCCACTCGGAGTCGCGCTTCGCGCGCCCTCACTGCAGCCCCTGCGTCG
CCGGGACCCCTCGCGCGGACCAGCCGAATCGCTCCTGCAGCAGAGCCAAC

10 20 30 40 50
ATGCCCATCACTTGGATGCGCATGAGACCCTGGCTAGAGATGCAGATTAA
MetProIleThrTrpMetArgMetArgProTrpLeuGluMetGlnIleAsn

60 70 80 90 100
TTCCAACCAATCCCGGGGCTCATCTGGATTAATAAACAGGAGATGATCT
SerAsnGlnIleProGlyLeuIleTrpIleAsnLysGluGluMetIleLeu

110 120 130 140 150
TGGAGATCCCATGGAAGCATGCTGCCAAGCATGGCTGGGACATCAACAAG
GluIleProTrpLysHisAlaAlaLysHisGlyTrpAspIleAsnLys

160 170 180 190 200
GATGCCTGTTTGTTCGGAGCTGGGCCATTACACAGGCCGATACAAAGC
AspAlaCysLeuPheArgSerTrpAlaIleHisThrGlyArgTyrLysAla

210 220 230 240 250
AGGGGAAAAGGAGCCAGATCCCAAGACGTGGAAGGCCAACTTTCGCTGTG
GlyGluLysGluProAspProLysThrTrpLysAlaAsnPheArgCysAla

260 270 280 290 300
CCATGAATCCCTGCCAGATATCGAGGAGGTGAAAGACCAGAGCAGGAAC
MetAsnSerLeuProAspIleGluGluValLysAspGlnSerArgAsn

310 320 330 340 350
AAGGGCAGCTCAGCTGTGCGAGTGTACCGGATGCTTCCACCTCTCACCAA
LysGlySerSerAlaValArgValTyrArgMetLeuProProLeuThrLys

360 370 380 390 400
GAACCAGAGAAAAGAAAGAAAGTCCGAAGTCCAGCCGAGATGCTAAGAGCA
AsnGlnArgLysGluArgLysSerLysSerSerArgAspAlaLysSerLys

410 420 430 440 450
AGGCCAAGAGGAAGTCATGTGGGGATTCCAGCCCTGATACCTTCTCTGAT
AlaLysArgLysSerCysGlyAspSerSerProAspThrPheSerAsp

460 470 480 490 500
GGACTCAGCAGCTCCACTCTGCCTGATGACCACAGCAGCTACACAGTTCC
GlyLeuSerSerSerThrLeuProAspAspHisSerSerTyrThrValPro

Figure 6

Page b

510 520 530 540 550
AGGCTACATGCAGGACTTGGAGGTGGAGCAGGCCCTGACTCCAGCACTGT
GlyTyrMetGlnAspLeuGluValGluGlnAlaLeuThrProAlaLeuSer

560 570 580 590 600
CGCCATGTGCTGTCAGCAGCACTCTCCCCGACTGGCAGATCCCAGTGGAA
ProCysAlaValSerSerThrLeuProAspTrpHisIleProValGlu

610 620 630 640 650
GTTGTGCCCGACAGCACCAGTGATCTGTACAACCTCCAGGTGTCAACCAT
ValValProAspSerThrSerAspLeuTyrAsnPheGlnValSerProMet

660 670 680 690 700
GCCCTCCATCTCTGAAGCTACAACAGATGAGGATGAGGAAGGCAAATTAC
ProSerIleSerGluAlaThrThrAspGluAspGluGluGlyLysLeuPro

710 720 730 740 750
CTGAGGACATCATGAAGCTCTTGGAGCAGTCGGAGTGCCAGCCAACAAAC
GluAspIleMetLysLeuLeuGluGlnSerGluTrpGlnProThrAsn

760 770 780 790 800
GTGGATGGGAAGGGGTACCTACTCAATGAACCTGGAGTCCAGCCCACCTC
ValAspGlyLysGlyTyrLeuLeuAsnGluProGlyValGlnProThrSer

810 820 830 840 850
TGTCTATGGAGACTTTAGCTGTAAGGAGGAGCCAGAAATTGACAGCCCAG
ValTyrGlyAspPheSerCysLysGluGluProGluIleAspSerProGly

860 870 880 890 900
GGGGGGATATTGGGCTGAGTCTACAGCGTGTCTTCACAGATCTGAAGAAC
GlyAspIleGlyLeuSerLeuGlnArgValPheThrAspLeuLysAsn

910 920 930 940 950
ATGGATGCCACCTGGCTGGACAGCCTGCTGACCCCAGTCCGGTTGCCCTC
MetAspAlaThrTrpLeuAspSerLeuLeuThrProValArgLeuProSer

960 970
CATCCAGGCCATTCCCTGTGCACCGTAGCAGGGCCCCCTGGGCCCCCTCTTA
IleGlnAlaIleProCysAlaPro***

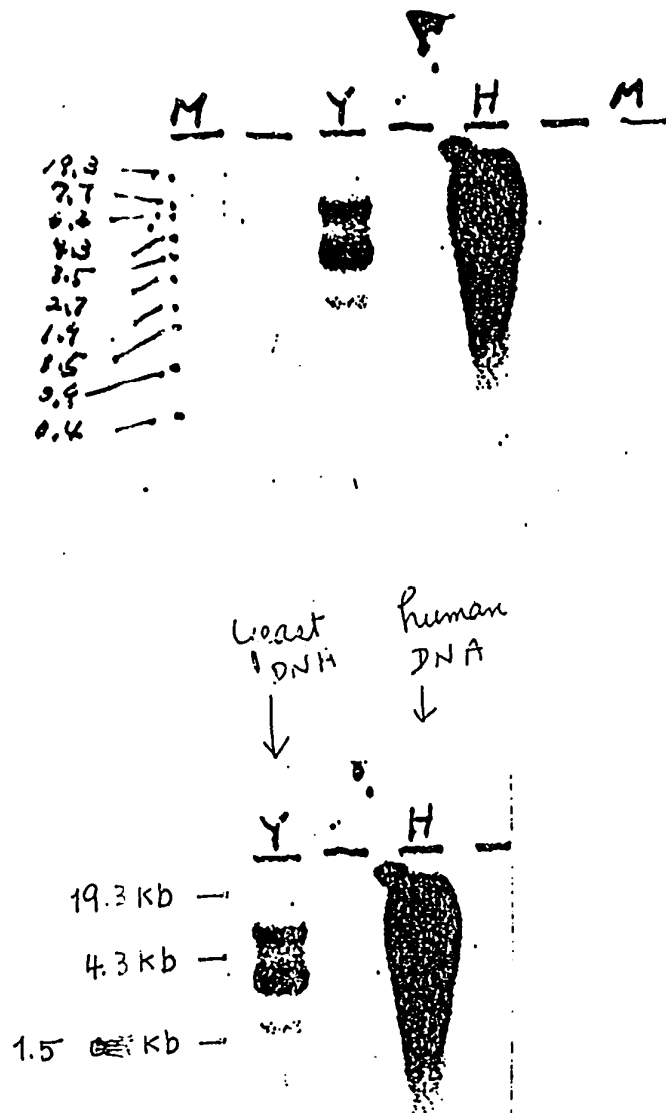
TTCCTCTAGGCAAGCAGGACCTGGCATCATGGTGGATATGGTGCAGAGAA
GCTGGACTTCTGTGGGCCCCCTCAACAGCCAAGTGTGACCCCACTGCCAAG
TGGGGATGGGCCTCCCTCCTTGGGTCATTGACCTCTCAGGGCCTGGCAGG
CCAGTGCTGGGTTTTTCTTGTGGTGTAAAGCTGGCCCTGCCTCCTGGGA
AGATGAGGTTCTGAGACCAGTGTATCAGGTCAGGGACTTGGACAGGAGTC

Figure 8

Page c

AGTGTCTGGCTTTTCCTCTGAGCCCAGCTGCCTGGAGAGG TCTCGCTG
TCACTGGCTGGCTTCCTAGGGGAACAGACCAGTGACCCAGAAAAGCATAA
CACCAATCCCAGGGCTGGCTCTGCACTAAGAGAAAATTGCACTAAATGAA
TCTCGTTCCAAAGAACTACCCCTTTTCAGCTGAGCCCTGGGACTGTTC
AAAGCCAGTGAATGTGAAGGAAAGTGUGGTCCTTCGGGGCAATGCTCCCT
CAGCCTCAGAGGAGCTCTACCCTGCTCCCTGCTTTGGCTGAGGGGCTTGG
GAAAAAACTTGGCACTTTTTCGTGTGGATCTTGCCACATTTCTGATCAG
AGGTGTACACTAACATTTCCCCCGAGCTCTTGGCCTTGCATTTATTTAT
ACAGTGCCTTGCTCGGGGGCCACCACCCCTCAAGCCCCAGCAGCCCTCA
ACAGGCCCAGGGAGGGAAGTGTGAGCGCCTTGGTATGACTTAAAATTGGA
ATGTCATCTAACCATTAAGTCATGTGTGAACACATAAGGACGTGTGTAA
ATAAGTACATTTGTCTTTTATAAAAAGTAAAATTGTT

Fig. 10



DNA sequences present in yeast that cross-hybridize with the human IRF-1 cDNA.