



(19)  
Bundesrepublik Deutschland  
Deutsches Patent- und Markenamt

(10) **DE 600 24 451 T2 2006.08.17**

(12) **Übersetzung der europäischen Patentschrift**

(97) **EP 1 192 268 B1**

(51) Int Cl.<sup>8</sup>: **C12Q 1/00 (2006.01)**

(21) Deutsches Aktenzeichen: **600 24 451.2**

(86) PCT-Aktenzeichen: **PCT/EP00/05582**

(96) Europäisches Aktenzeichen: **00 942 090.2**

(87) PCT-Veröffentlichungs-Nr.: **WO 2001/000864**

(86) PCT-Anmeldetag: **09.06.2000**

(87) Veröffentlichungstag  
der PCT-Anmeldung: **04.01.2001**

(97) Erstveröffentlichung durch das EPA: **03.04.2002**

(97) Veröffentlichungstag  
der Patenterteilung beim EPA: **30.11.2005**

(47) Veröffentlichungstag im Patentblatt: **17.08.2006**

(30) Unionspriorität:  
**99202068 25.06.1999 EP**

(84) Benannte Vertragsstaaten:  
**AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT,  
LI, LU, MC, NL, PT, SE**

(73) Patentinhaber:  
**Vlaams Interuniversitair Instituut voor  
Biotechnologie vzw., Zwijnaarde, BE**

(72) Erfinder:  
**HUYLEBROECK, Danny, B-1770 Liedekerke, BE;  
VERSCHUEREN, Kristin, B-3078 Everberg, BE;  
REMACLE, Jacques, B-4280 Hannut, BE**

(74) Vertreter:  
**Dehmel & Bettenhausen, Patentanwälte, 80331  
München**

(54) Bezeichnung: **BINDUNG MEHRERER ZINKFINGER TRANSKRIPTIONSFAKTOREN AN NUKLEINSÄUREN**

Anmerkung: Innerhalb von neun Monaten nach der Bekanntmachung des Hinweises auf die Erteilung des europäischen Patents kann jedermann beim Europäischen Patentamt gegen das erteilte europäische Patent Einspruch einlegen. Der Einspruch ist schriftlich einzureichen und zu begründen. Er gilt erst als eingelegt, wenn die Einspruchsgebühr entrichtet worden ist (Art. 99 (1) Europäisches Patentübereinkommen).

Die Übersetzung ist gemäß Artikel II § 3 Abs. 1 IntPatÜG 1991 vom Patentinhaber eingereicht worden. Sie wurde vom Deutschen Patent- und Markenamt inhaltlich nicht geprüft.

## Beschreibung

**[0001]** Die Erfindung betrifft ein Verfahren zur Identifizierung von Transkriptionsfaktoren umfassend die Bereitstellung von Zellen mit einer Nukleinsäuresequenz, umfassend mindestens eine CACCT-Sequenz als Köder für das Durchmusteren einer Bibliothek, die für potentielle Transkriptionsfaktoren kodiert, und Durchführung eines Spezifitätstests, um die genannten Faktoren zu isolieren. Vorzugsweise umfasst der Köder zwei CACCT-Sequenzen, noch bevorzugter umfasst der Köder eine der Sequenzen CACCT-N-CACCT, CACCT-N-AGGTG, AGGTG-N-CACCT oder AGGTG-N-AGGTG, wobei N eine Zwischensequenz ist.

**[0002]** Der (die) mittels des erfindungsgemäßen Verfahrens identifizierte(n) Transkriptionsfaktor(en) umfasst (umfassen) getrennte Cluster von Zinkfingern wie beispielsweise einen zueinander zweihändigen Zinkfinger-Transkriptionsfaktor. Die vorliegende Erfindung offenbart weiterhin, dass wenigstens ein solcher Zinkfinger-Transkriptionsfaktor, bezeichnet als SIP1, Tumormetastasierung durch Herunterregulieren der Expression von E-Cadherin induziert. Daher können Verbindungen, die die SIP1-Aktivität stören, verwendet werden, um Tumordinvasion und -metastasierung zu verhindern.

**[0003]** Zinkfinger sind unter den häufigsten DNA-Bindungsmotiven, die man in Eukaryoten findet. Man schätzt, dass vom Hefegenom 500 Zinkfingerproteine kodiert werden und dass vielleicht ein Prozent aller Säugerogene für Zinkfinger-enthaltende Proteine kodiert. Diese sind gemäß der Anzahl und Position der Cystein- und Histidinreste, die für die Zink-Koordination verfügbar sind, klassifiziert. Die CCHH-Klasse, die vom Xenopus-Transkriptionsfaktor IIIA beispielhaft repräsentiert wird (19), ist die größte. Diese Proteine enthalten zwei oder mehr Finger in Tandem-Wiederholungen. Im Gegensatz dazu enthalten Steroidrezeptoren nur Cysteinreste, die zwei Arten von Zink-kooordinierten Strukturen mit vier ( $C_4$ ) und fünf ( $C_5$ ) Cysteinen bilden (28). Die dritte Klasse an Zinkfingern enthält die CCHC-Finger. Die CCHC-Finger, die man in Drosophila, und in Säuger- und retroviralen Proteinen findet, weisen die Konsensussequenz  $C-X_2-C-X_4-H-X_4-C$  auf (7, 21, 24). Kürzlich wurde eine neuartige Konfiguration des CCHC-Fingers, vom  $C-X_5-C-X_{12}-H-X_4-C$ -Typ, in der neuronalen Zinkfingerfaktor/Myelintranskriptionsfaktorfamilie gefunden (11, 12, 36). Schließlich enthalten mehrere Hefetranskriptionsfaktoren wie GAL4 und CHA4 eine atypische  $C_6$ -Zinkfingerstruktur, die zwei Zinkionen koordiniert (9, 32).

**[0004]** Zinkfinger findet man normalerweise in multiplen Kopien (bis zu 37) pro Protein. Diese Kopien können in Tandem angeordnet sein, einen oder multiple Cluster bilden, oder sie können über das ganze Protein verteilt sein. Mehrere Transkriptionsfaktorfamilien teilen die gleiche Gesamtstruktur, nämlich indem sie zwei (oder drei) weit voneinander getrennte Cluster an Zinkfingern in ihrer Proteinsequenz haben. Die erste, die MBPs/PRDII-BF1-Transkriptionsfaktorfamilie schließt die Gene Schnurri und Spult aus Drosophila mit ein (1, 3, 6, 14, 33). Sowohl MBP-1 (auch bekannt als PRDII-BF1) als auch MBP-2 enthalten zwei weit voneinander getrennte Cluster von zwei CCHH-Zinkfingern. Die insgesamt Ähnlichkeit zwischen MBP-1 und MBP-2 beträgt 51%, aber die Konservierung ist sowohl für die N-terminalen als auch für die C-terminalen Zinkfingercluster viel höher (über 90%) (33). Das zeigt eine wichtige Rolle für beide Cluster bei der Funktionsweise dieser Proteine an. Zusätzlich sind die N-terminalen und C-terminalen Zinkfingercluster von MBP-1 sehr homolog zueinander (3). Der neuralspezifische Zinkfingerfaktoren 1 und 3 (NZF-1 und NZF-3) sowie der Myelintranskriptionsfaktor 1 (MyT1, auch bekannt als NZF-2) gehören zu einer weiteren Proteinfamilie, die zwei weit voneinander getrennte Cluster an CCHC-Zinkfingern enthält (11, 12, 36). Wie die MBP-Proteine weisen verschiedene NZF-Faktoren einen hohen Grad an Sequenzidentität (über 80%) zwischen den entsprechenden Zinkfingerclustern auf, wohingegen die Sequenzen außerhalb des Zinkfingerbereichs weitgehend divergieren (36). Zusätzlich kann jeder dieser Cluster unabhängig an DNA binden und erkennt ähnliche Kernkonsensussequenzen (11). NZF-3 bindet an ein DNA-Element, das eine einzelne Kopie dieser Konsensussequenz enthält, weist aber nachweislich eine ausgeprägte Verstärkung der relativen Affinität zu einem zweiteiligen Element auf, das zwei Kopien dieser Sequenz enthält (36). Das liegt nahe, dass die NZF-Faktoren auch an die wiederholten Sequenzen binden können. Der kooperativen Bindung von NZF-3 an das zweiteilige Element zugrunde liegende Mechanismus ist jedoch momentan unbekannt. Das Drosophila Protein Zfh-1 und das Vertebratenprotein  $\delta$ EF1 (auch bekannt als ZEB oder AREB6) gehören zu einer dritten Transkriptionsfaktorfamilie. Diese Familie ist dadurch gekennzeichnet, dass zwei voneinander getrennte Cluster an CCHH-Zinkfingern und eine Homöodomänen-artige Struktur vorliegen (siehe Fig. 1A) (4, 5, 35). In  $\delta$ EF1 sind die N-terminalen und C-terminalen Cluster ebenfalls sehr homolog und binden nachweislich unabhängig an sehr ähnliche Kern-Konsensussequenzen (10). Kürzlich wurde gezeigt, dass mutierte Formen von  $\delta$ EF1, denen entweder der N-terminale oder der C-terminale Cluster fehlt, ihre DNA-Bindungskapazität verloren haben, was anzeigt, dass beide Cluster notwendig für die Bindung von  $\delta$ EF1 an DNA sind (31). Für den Evi-1-Transkriptionsfaktor wurde gezeigt, dass er 10 CCHH-Zinkfinger enthält; 7 Zinkfinger liegen im N-terminalen Bereich vor und 3 Zinkfinger liegen im C-terminalen Bereich (22). Bei diesem Faktor ist die Situation anders als bei den oben beschriebenen Trans-

kriptionsfaktoren, weil die zwei Cluster an zwei verschiedene Zielsequenzen binden, die simultan von Evi-1 voller Länge gebunden werden (20). Die Bindung von Evi-1 voller Länge wird vor allem beobachtet, wenn die zwei Zielsequenzen in einer bestimmten relativen Orientierung positioniert sind, aber es gab keine strikte Anforderung für einen optimalen Abstand zwischen diesen zwei Zielsequenzen. Kispert und Herrman (87) offenbaren das T-Protein, das an der Mesoderm-Bildung beteiligt ist und DNA bindet. Ikeda et al. (86) offenbaren eine schnelle Durchmusterung und Charakterisierung von DNA-bindenden oder -interagierenden Proteinen.

**[0005]** Zell-Zell-Adhäsion ist eine vorherreschende Notwendigkeit im Laufe von Zelldifferenzierung, Gewebeentwicklung und Gewebehomöostase. Die Wirkung einer zerstörten Zell-Zell-Adhäsion zeigt sich in vielen Krebsarten, wo Metastasierung und eine schlechte Prognose mit dem Verlust von Zell-Zell-Adhäsion korrelieren. E-Cadherin, ein homophiles  $\text{Ca}^{2+}$ -abhängiges Transmembran-Adhäsionsmolekül, und die assoziierten Catenine gehören zu den Hauptbestandteilen des epithelialen Zell-Verbindungssystems. E-Cadherin spielt eine starke invasionsunterdrückende Rolle in Tumorzelliniensystemen (46, 47) und in in vivo Tumormodellsystemen (48). Der Verlust der E-Cadherin-Expression im Verlauf der fortschreitenden Tumorentwicklung ist für mehr als 15 verschiedenen Karzinomtypen beschrieben worden (49). Extensive Analysen machten klar, dass aberrante E-Cadherin-Expression als Resultat somatischer inaktivierender Mutationen beider E-Cadherin-Allele selten ist und bislang auf die diffuse Magenkarzinome und infiltrative lobuläre Brustkrebskarzinome beschränkt sind (50, 51). Northern-Blot-Analysen und in situ Hybridisierungsstudien zeigten, dass eine reduzierte E-Cadherin-Immunreaktivität in humanen Karzinomen mit reduzierten mRNA-Spiegeln korreliert (52–54). Die Analyse der E-Cadherin-Promotorsequenzen von Maus und Mensch zeigte eine konservierte modulare Struktur mit positiv regulatorischen Elementen einschließlich einer CCAAT-Box und einer GC-Box sowie zwei E-Boxen (CANNTG) mit potentieller Repressorfunktion (55, 56). Eine Mutationsanalyse der zwei E-Boxen im E-Cadherin-Promotor offenbarte eine äußerst wichtige Rolle bei der Regulation der Epithel-spezifischen Expression von E-Cadherin. Mutation dieser zwei E-Box-Elemente führt zur Hochregulierung des E-Cadherin-Promotors in de-differenzierten Krebszellen, wo der Wildtyp-Promotor geringe Aktivität zeigt (55, 56).

#### Kurze Beschreibung der Figuren

**[0006] Fig. 1.** Schematische Darstellung von Zfh-1, SIP1 und  $\delta\text{EF1}$ , und eines Alignments der Zinkfinger von SIP1 und  $\delta\text{EF1}$ . (A) schematische Darstellung des  $\delta\text{EF1}$  der Maus (1117 Aminosäuren) und SIP1 (1214 Aminosäuren). Die gefüllten Kästen stellen CHH-Zinkfinger dar, die leeren Kästen sind CCHC-Zinkfinger. Die homöodomänenartige Domäne (HD) ist als Oval dargestellt. Der Prozentsatz gibt die Homologie zwischen verschiedenen Domänen an. SIP1-Polypeptide, die in dieser Untersuchung verwendet wurden, sind mit ihren Koordinaten angegeben. SBD: Smad-bindende Domäne (Verschueren et al., 1999). (B) Alignments der Aminosäuresequenzen der Zinkfinger von SIP1 und  $\delta\text{EF1}$ . Vertikale Balken zeigen die Sequenzidentität an. Die konservierten Cystein- und Histidinreste, die die Zinkfinger bilden, sind fettgedruckt und durch ein Sternchen angezeigt. Die Reste in den Zinkfingern, die DNA kontaktieren können, sind mit einem Pfeil angezeigt. (C) Alignment der Proteinsequenz von SIP1<sub>NZF3+NZF4</sub> und SIP1<sub>CZF2+CZF3</sub> bzw. von  $\delta\text{EF1}$ <sub>NZF3+NZF4</sub> und  $\delta\text{EF1}$ <sub>CZF2+CZF3</sub>, das die intramolekulare Konservierung der Zinkfinger zeigt.

**[0007] Fig. 2.** Mögliche DNA-Bindemechanismen für SIP1. Modell 1: SIP1 bindet an DNA als Monomer. Modell 2: SIP1 bindet an DNA als Dimer.

**[0008]** Für die meisten zuvor erwähnten Komplexfaktoren versteht man den Mechanismus der DNA-Bindung nur unzureichend. Es ist unsere Erfindung, die DNA-Bindungseigenschaften von Transkriptionsfaktoren von Vertebraten, die zu der immer mehr Gestalt annehmenden Familie von zweihändigen Zinkfinger-Transkriptionsfaktoren wie  $\delta\text{EF1}$  und SIP1 gehören, zu charakterisieren. SIP1 ist Mitglied dieser Transkriptionsfaktorfamilie, und wurde kürzlich als ein mit Smad-wechselwirkendes Protein (34 WO 98/55512) isoliert und charakterisiert. SIP1 und  $\delta\text{EF1}$ , ein transkriptioneller Repressor, der an der Skelettentwicklung und an der Muskelzell-differenzierung beteiligt ist, gehören zur gleichen Transkriptionsfaktorfamilie. Sie enthalten zwei getrennte Cluster an CHH-Zinkfingern, die hohe Sequenzidentität aufweisen (>90%). Die DNA-Bindungseigenschaften dieser Transkriptionsfaktoren wurden untersucht. Die N-terminalen und die C-terminalen Cluster von SIP1 zeigen ebenfalls hohe Sequenzhomologie, und gemäß der Erfindung bindet jeder an eine 5'-CACCT-Sequenz. Darüber hinaus sind hochaffine Bindungsstellen für SIP1 voller Länge und  $\delta\text{EF1}$  in den Promotorbereichen von Kandidatenzielgenen, wie Brachyury,  $\alpha 4$ -Integrin und E-Cadherin, zweiteilige Elemente, die sich aus einer CACCT-Sequenz und einer CACCTG-Sequenz zusammensetzen. Es wurde keine absolute Notwendigkeit für die relative Orientierung beider Sequenzen beobachtet, und der Abstand zwischen ihnen (auch als N bezeichnet) kann von 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 ..., bis mindestens 44 bp variieren. Um an diese zweiteiligen Elemente zu binden, ist die Intaktheit beider SIP1-Zinkfinger-Cluster notwendig, was anzeigt, dass sie beide an der DNA-Bindung beteiligt sind. Außerdem bindet SIP1 als Monomer an eine CACCT- $X_N$ -CACCTG-Stelle, wobei

ein Zinkfinger-Cluster die CACCT-Stelle kontaktiert, und der andere Zinkfinger-Cluster an die CACCTG-Sequenz bindet. Diese neuartige Bindungsart kann auf andere Transkriptionsfaktoren verallgemeinert werden, die voneinander getrennte Zinkfinger-Cluster enthalten und kann auch auf andere Smad-bindende Proteine übertragen werden. Darüber hinaus zeigt das mit Smad-wechselwirkende Protein SIP1 eine hohe Expression in E-Cadherin-negativen humanen Karzinom-Zelllinien, was zur Herunterregulierung der E-Cadherin-Transkription führt. Die konditionale Expression von SIP1 in E-Cadherin-positiven MDCK-Zellen hebt auch die E-Cadherin-vermittelte interzelluläre Adhäsion auf und induziert simultan die Invasion. Daher kann SIP1 als ein stark die Invasion förderndes Molekül betrachtet werden und Verbindungen, wie Anti-SIP1-Antikörper, kleine Moleküle, die spezifisch an SIP binden, Antisense-Nukleinsäuren und Ribozyme, die die Produktion oder Aktivität von SIP1 stören, können Tumorinvasion und Metastasierung verhindern.

**[0009]** Die Erfindung betrifft daher ein Verfahren zur Identifizierung von Transkriptionsfaktoren wie Aktivator(en) und/oder Repressoren, umfassend die Bereitstellung von Zellen mit einer Nukleinsäuresequenz, umfassend mindestens eine CACCT-Sequenz, vorzugsweise eine zweifache CACCT-Sequenz, als Köder für das Durchmusterung einer Bibliothek, die für potentielle Transkriptionsfaktoren kodiert, und Durchführung eines Spezifitätstests, um die genannten Faktoren zu isolieren. In einer weiteren Ausführungsform umfasst der Köder eine der Sequenzen CACCT-N-CACCT, CACCT-N-AGGTG, AGGTG-N-CACCT oder AGGTG-N-AGGTG, wobei N eine Zwischensequenz ist. Letztere Zwischensequenz kann in ihrer Länge variieren und kann jede Anzahl an Basenpaaren (bp) von N = 1 bp bis N = mindestens 44 bp enthalten. Daher kann N zum Beispiel 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 200, 300 oder 400 bp lang sein.

**[0010]** Der (die) unter Verwendung des erfindungsgemäßen Verfahrens identifizierte(n) Transkriptionsfaktor(en) umfasst (umfassen) voneinander getrennte Zinkfinger-Cluster wie beispielsweise zweihändige Zinkfinger-Transkriptionsfaktoren.

**[0011]** Die oben erwähnte Sequenz kann von jedem Promotorbereich abstammen, stammt aber vorzugsweise von der Gruppe ausgewählt aus Brachyury,  $\alpha$ 4-Integrin, Follistatin oder E-Cadherin ab (auch als Zielgene bezeichnet, siehe später).

**[0012]** Die mit dem oben erwähnten Verfahren erhältlichen Transkriptionsfaktoren sind ebenfalls Teil der vorliegenden Erfindung.

**[0013]** In einer weiteren Ausführungsform betrifft die vorliegende Erfindung ein Verfahren zur Identifizierung von Verbindungen, die in der Lage sind, mit Transkriptionsfaktoren, die wie oben erwähnt erhältlich sind, zu interferieren durch

- a) Zusetzen einer Probe, umfassend eine potentielle Substanz, zur Identifizierung zu einem Testsystem, umfassend (i) eine Nukleotidsequenz, die eine der Sequenzen CACCT-N-CACCT, CACCT-N-AGGTG, AGGTG-N-CACCT oder AGGTG-N-AGGTG als Köder umfasst, wobei N eine Zwischensequenz ist, und (ii) ein Protein, das in der Lage ist, die Nukleotidsequenz zu binden,
- b) Inkubieren der Probe in dem System für eine Zeitdauer, die ausreichend ist, um eine Interaktion zwischen der Substanz, einem Derivat der Substanz, oder ihrem Gegenstück mit dem Protein zu ermöglichen,
- c) Vergleichen der Menge und/oder Aktivität des an die Nukleotidsequenz gebundenen Proteins vor und nach dem Zusetzen.

**[0014]** Das Vergleichen der Menge und/oder Aktivität des an die Nukleotidsequenz gebundenen Proteins vor und nach dem Zusetzen der Test-Probe kann beispielsweise erreicht werden durch Verwendung eines Gel-Retardationsassays oder eines Filterbindungsassays. Im nächsten Schritt kann die derart identifizierte Verbindung isoliert werden und gegebenenfalls aufgereinigt und weiter analysiert werden gemäß Verfahren, die dem Fachmann bekannt sind. Das Protein aus Schritt a) (ii) kann jedes Protein sein, das in der Lage ist, an die Nukleotidsequenz zu binden, ist aber vorzugsweise ein Protein, das mit Smad wechselwirkt, wie SIP1. Verbindungen, die mit letzterem Verfahren identifiziert werden, sind auch Teil der vorliegenden Erfindung. Mit den Ausdrücken „Verbindungen die in der Lage sind, mit Transkriptionsfaktoren zu interferieren“ sind Verbindungen gemeint, die in der Lage sind, die Bioaktivität von Transkriptionsfaktoren zu modulieren (= d. h. zu unterbinden, zu schwächen, zu stärken). Insbesondere sind letztere Verbindungen in der Lage, die Produktion und/oder Bioaktivität von SIP1 vollständig oder teilweise zu unterbinden. Beispiele für solche Verbindungen sind kleine Moleküle oder Anti-SIP1-Antikörper oder funktionale Fragmente, die davon abgeleitet sind, die spezifisch an SIP1-Protein binden oder Antisense-Nukleinsäuren oder Ribozyme, die an mRNA binden, die für SIP1 kodiert, oder kleine Moleküle, die an den Promotorbereich binden, der von SIP1 gebunden wird. Diesbezüglich betrifft die vorliegende Erfindung Verbindungen, die die Regulation der E-Cadherin-Expression durch SIP1 modulieren. Insbesondere betrifft die vorliegende Erfindung Verbindungen die, über die Unterbindung der SIP1-Pro-

duktion und/oder -Aktivität, die Herunterregulierung der Expression des Zielgens E-Cadherin verhindern. Mit anderen Worten, die vorliegende Erfindung betrifft Verbindungen, die als ein Medikament verwendet werden können, um eine Tumorerkrankung und/oder Metastasierung, die in der Herunterregulierung der E-Cadherin-Expression durch SIP1 begründet ist, vorzubeugen oder zu behandeln. Verfahren, um letztere Verbindungen herzustellen und zu verwenden, werden später beispielhaft dargestellt.

**[0015]** Im Rahmen der vorliegenden Erfindung liegt auch ein Testkit, um das Verfahren durchzuführen, umfassend mindestens (i) eine Nukleotidsequenz, die eine der Sequenzen CACCT-N-CACCT, CACCT-N-AGGTG, AGGTG-N-CACCT oder AGGTG-N-AGGTG umfasst, wobei N eine Zwischensequenz ist und (ii) ein Protein, das in der Lage ist, an die Nukleotidsequenz zu binden.

**[0016]** In einer weiteren Ausführungsform betrifft die vorliegende Erfindung eine Alternative zu sogenannten Zwei-Hybrid-Durchmusterungsassays, wie sie im Stand der Technik offenbart sind. Mehrere Mittel und Verfahren sind entwickelt worden, um Bindungspartner von Proteinen zu identifizieren. Das führte zur Identifizierung einer Anzahl entsprechender Bindungsproteine. Viele dieser Proteine wurden unter Verwendung sogenannter Zwei-Hybrid-Systeme gefunden. Zwei-Hybrid-Kloniersysteme sind in mehreren Laboren entwickelt worden (Chien et al., 1991; Durfee et al., 1993; Gyuris et al., 1993). Alle haben 3 grundlegende Bestandteile: Hefektoren für die Expression eines bekannten Proteins, das an eine DNA-Bindedomäne fusioniert ist, Hefektoren, die die Expression cDNA-kodierter Proteine, die an eine Transkriptions-Aktivierungsdomäne fusioniert sind, steuern, und Hefereportergene, die Bindungsstellen für die DNA-Bindedomäne enthalten. Diese Bestandteile unterscheiden sich im Detail von einem System zum anderen. Alle Systeme nutzen die DNA-Bindedomäne von entweder Gal4 oder LexA. Die Gal4-Domäne lokalisiert effizient im Hefenuklear, wo sie mit hoher Affinität an gut charakterisierte Bindungsstellen bindet, die stromaufwärts der Reportergene platziert werden können (Silver et al., 1986). LexA hat kein nukleäres Lokalisationssignal, erreicht aber den Hefenuklear und bindet, wenn es in ausreichendem Maße exprimiert wird, effizient LexA-Bindestellen (Operatoren), die stromaufwärts eines Reportergens platziert sind (Brent et al., 1985). Endogene Hefepoteine binden nicht an LexA-Operatoren.

**[0017]** Verschiedene Systeme nutzen auch verschiedene Reporter. Die meisten Systeme verwenden einen Reporter, der einen Hefepromotor hat, entweder vom GAL1-Gen oder vom CYC1-Gen, fusioniert an lacZ (Yocum et al., 1984). Diese lacZ-Fusionen liegen entweder auf Hefepiasmiden mit hoher Kopienzahl vor oder sind in ein Hefechromosom integriert. Um aus den lacZ-Fusionen geeignete Reporter zu machen, wurde die Transkription von GAL1- oder CYC1-regulierenden Bereichen entfernt und ausgetauscht gegen Bindestellen, die von der verwendeten DNA-Bindedomäne erkannt werden. Eine Durchmusterung auf die Aktivität von lacZ-Reportern wird durchgeführt durch Ausplattieren von Hefe auf Indikatorplatten, die X-Gal enthalten (5-Brom-4-Chlor-3-Indolyl- $\beta$ -D-Galaktosid); auf diesem Medium produziert Hefe, in der die Reporter transkribiert werden,  $\beta$ -Galaktosidase und wird blau. Einige Systeme verwenden ein zweites Reportergen und einen Hefestamm, der der Expression dieses Reporters bedarf, um in einem bestimmten Medium zu wachsen. Diese „selektionierbaren Marker“-Gene kodieren normalerweise für Enzyme, die für die Biosynthese einer Aminosäure gebraucht werden. Solche Reporter haben den merklichen Vorteil eine Selektion auf cDNAs zur Verfügung zu stellen, die für wechselwirkende Proteine kodieren, anstelle einer visuellen Durchmusterung auf blaue Hefen. Um geeignete Reporter von den Markergenen zu machen, müssen deren stromaufwärts liegende, Transkriptions-regulierende Elemente ausgetauscht werden gegen Bindestellen für eine DNA-Bindedomäne. Die HIS3- und LEU2-Gene sind beide als Reporter in Verbindung mit geeigneten Hefestämmen, die deren Expression bedürfen, um in einem Medium zu wachsen, dem entweder Histidin oder bzw. Leucin fehlt, verwendet worden. Schließlich verwenden verschiedene Systeme verschiedene Mittel, um aktivierungsmarkierte cDNA-Proteine zu exprimieren. In allen vorliegenden Schemata werden die cDNA-kodierten Proteine mit einer Aktivierungsdomäne am Amino-Terminus exprimiert. Die verwendeten Aktivierungsdomänen schließen die starke Aktivierungsdomäne von Gal4, die sehr starke Aktivierungsdomäne des Herpes Simplex-Virusproteins VP16, oder eine schwächere Aktivierungsdomäne, die von Bakterien abgeleitet wurde, genannt B42, mit ein. Die aktivierungsmarkierten, cDNA-kodierten Proteine werden entweder von einem konstitutiven Promotor, oder von einem konditionalen Promotor wie dem des GAL1-Gens exprimiert. Die Verwendung eines konditionalen Promotors macht es möglich, schnell zu zeigen, dass die Aktivierung des Reportergens abhängig ist von der Expression des aktivierungsmarkierten cDNA-Proteins.

**[0018]** Es wird aus der obigen Diskussion klar, dass Zwei-Hybrid-Systeme zum Auffinden von Bindungsproteinen in der Vergangenheit verwendet wurden. Auch wenn sich das konventionelle Zwei-Hybrid-System als wertvolles Werkzeug beim Auffinden von proteinösen Molekülen erwiesen hat, die an andere Proteine binden können, so ist es jedoch ein (sehr) artifizielles System. Ein Charakteristikum jeden 2-Hybrid-Systems ist, dass ein Fusionsprotein gemacht wird, das aus einem Teil besteht, für den Bindungspartner gesucht werden und

einem Reporterteil besteht, der den Nachweis der Bindung ermöglicht. Um relevante Bindungspartner zu finden, müssen mehrere Kriterien erfüllt werden, von denen natürlich eines die korrekte Wahl des Bereichs im Protein ist, wo die Bindung an andere Proteine auftritt. Ein weiteres Kriterium, das weitaus schwieriger, wenn nicht im Voraus unmöglich, präzise vorherzusagen ist, ist die korrekte Faltung des Bereichs (d. h. eine Faltung des Bereichs, die ausreichend ähnlich zur Faltung des Bereichs im natürlichen Protein ist). Die korrekte Faltung hängt unter anderem von der tatsächlichen Aminosäuresequenz ab, die für die Erzeugung des Fusionsproteins gewählt wird. Ein weiterer Faktor, der die Identifizierung relevanter Bindungspartner bestimmt, ist die Sensitivität, mit der die Bindung nachgewiesen werden kann.

**[0019]** Eine Alternative zu dem oben erwähnten konventionellen Zwei-Hybrid-System wird hiermit in der vorliegenden Erfindung auch zur Verfügung gestellt. Daher ist es ein alternatives Ziel der Erfindung, ein in vivo-Verfahren und einen Kit zum Nachweis von Wechselwirkungen zwischen Proteinen und dem Einfluss anderer Verbindungen auf diese Wechselwirkung als solche durch Wiederherstellung der Aktivität eines transkriptionellen Aktivators zur Verfügung zu stellen. Diese Wiederherstellung macht Verwendung von zwei, sogenannten Hybriden, chimären oder fusionierten Proteinen. Diese zwei fusionierten Proteine zeigen jeweils unabhängig voneinander eine schwache Affinität zu einer Nukleinsäuresequenz, die eine der Sequenzen CACCT-N-CACCT, CACCT-N-AGGTG, AGGTG-N-CACCT oder AGGTG-N-AGGTG umfasst, wobei N eine Zwischensequenz ist. Wenn jedoch beide fusionierten Proteine unabhängig an die Sequenz gebunden werden und die Test-Proteine, jedes verfügbar in jedem der zwei fusionierten Proteine, als Konsequenz daraus in unmittelbare Nähe gebracht werden, dann wird die Bindungsaffinität zu der Nukleinsäuresequenz, die eine der Sequenzen CACCT-N-CACCT, CACCT-N-AGGTG, AGGTG-N-CACCT oder AGGTG-N-AGGTG umfasst, wobei N eine Zwischensequenz ist, viel stärker. Wenn die zwei Test-Proteine tatsächlich in der Lage sind zu wechselwirken, dann bringen sie als Konsequenz daraus die zwei Domänen der transkriptionellen Aktivatoren in die unmittelbare Nähe. Diese Nähe ist ausreichend um Transkription zu bewirken, die über die Aktivität eines Markergens, das benachbart zu der Nukleinsäure lokalisiert ist, die eine der Sequenzen CACCT-N-CACCT, CACCT-N-AGGTG, AGGTG-N-CACCT oder AGGTG-N-AGGTG umfasst, wobei N eine Zwischensequenz ist, nachgewiesen werden kann. In Übereinstimmung damit wird ein Verfahren zum Nachweis einer Wechselwirkung zwischen einem ersten wechselwirkenden Protein und einem zweiten wechselwirkenden Protein zur Verfügung gestellt, umfassend

- a) zur Verfügung stellen einer geeigneten Wirtszelle mit einem ersten Fusionsprotein, umfassend ein erstes wechselwirkendes Protein fusioniert an eine DNA-Bindedomäne, das in der Lage ist, an eine Nukleinsäuresequenz umfassend eine der Sequenzen CACCT-N-CACCT, CACCT-N-AGGTG, AGGTG-N-CACCT oder AGGTG-N-AGGTG zu binden, wobei N eine Zwischensequenz ist,
- b) zur Verfügung stellen einer geeigneten Wirtszelle mit einem zweiten Fusionsprotein umfassend ein zweites wechselwirkendes Protein fusioniert an eine DNA-Bindedomäne, das in der Lage ist, an eine Nukleinsäuresequenz umfassend eine der Sequenzen CACCT-N-CACCT, CACCT-N-AGGTG, AGGTG-N-CACCT oder AGGTG-N-AGGTG zu binden, wobei N eine Zwischensequenz ist,
- c) Halten der Wirtszelle bei Bedingungen, unter denen das erste wechselwirkende Protein und das zweite wechselwirkende Protein in unmittelbare Nähe gebracht werden und
- d) Bestimmen, ob ein nachweisbares Gen, das in der Wirtszelle vorliegt und benachbart zu der Nukleinsäuresequenz lokalisiert ist, zu einem Grad exprimiert wird, der größer ist als in Abwesenheit der Wechselwirkung zwischen dem ersten und dem zweiten wechselwirkenden Protein.

**[0020]** Beispielsweise sollte es klar sein, dass, wenn ein Bindungspartner (Beute) für ein spezifisches Protein (Köder) identifiziert wurde, das erste Fusionsprotein, das den Köder enthält, zum Beispiel an die Sequenz CACCT (oder AGGTG) der Sequenz CACCT-N-AGGTG bindet und dass das zweite Fusionsprotein, das die Beute enthält, die Sequenz AGGTG (oder bzw. CACCT) der Sequenz CACCT-N-AGGTG binden wird, sodass die Transkription eines Markergens stattfindet.

**[0021]** Schließlich betrifft die Erfindung die neuen Sequenzen CACCT-N-CACCT, CACCT-N-AGGTG, AGGTG-N-CACCT oder AGGTG-N-AGGTG, wobei N eine Zwischensequenz wie oben definiert ist, und betrifft die Verwendung der Sequenzen zusätzlich zu jeder anderen Sequenz, die mindestens eine CACCT-Sequenz umfasst, für die Identifizierung neuer Zielgene, die verschieden von den bereits beschriebenen Zielgenen Brachyury,  $\alpha$ 4-Integrin, Follistatin oder E-Cadherin sind, mit jedem dem Fachmann bekannten Verfahren.

**[0022]** Die folgenden Definitionen werden festgelegt, um die Bedeutung und den Rahmen verschiedener Ausdrücke, die hier verwendet wurden um die Erfindung zu beschreiben, zu illustrieren und zu definieren und ihre Bedeutung wird im Folgenden um der Klarheit willen weiter ausgeführt.

**[0023]** „Nukleinsäure“ oder „Nukleinsäuresequenz“ oder „Nukleotidsequenz“ bezeichnet genomische DNA,

cDNA, doppelsträngige oder einzelsträngige DNA, Boten-RNA oder jede Form von Nukleinsäuresequenz, die dem Fachmann bekannt ist.

**[0024]** Die Ausdrücke „Protein“ und „Polypeptid“, die in dieser Anmeldung verwendet werden, sind gegenseitig austauschbar. „Polypeptid“ bezieht sich auf ein Aminosäure-Polymer (Aminosäuresequenz) und bezieht sich nicht auf eine spezifische Länge des Moleküls. Daher sind Peptide und Oligopeptide innerhalb dieser Definition des Polypeptids eingeschlossen. Der Ausdruck bezieht sich auch auf oder schließt posttranslationale Modifikationen des Polypeptids mit ein, beispielsweise Glykosylierungen, Acetylierungen, Phosphorylierungen und dergleichen. In die Definition eingeschlossen sind beispielsweise Polypeptide, die ein Analogon oder mehrere Analoga einer Aminosäure (einschließlich beispielsweise unnatürlicher Aminosäuren, etc.), Polypeptide mit substituierten Bindungen sowie anderen Modifikationen, die im Stand der Technik bekannt sind, sowohl natürlich auftretende als auch nicht natürlich auftretende, enthalten. Die oben beschriebenen Proteine und Polypeptide werden nicht notwendigerweise von einer bestimmten Nukleinsäuresequenz translatiert; die Polypeptide können auf jede Weise erzeugt werden, einschließlich beispielsweise chemischer Synthese, oder der Expression eines rEcombinanten Expressionssystems, oder der Isolierung aus einem geeigneten viralen System. Die Polypeptide können ein Analogon oder mehrere Analoga von Aminosäuren, phosphorylierten Aminosäuren oder unnatürlichen Aminosäuren einschließen. Im Stand der Technik sind Verfahren zum Einsetzen von Aminosäureanaloga in eine Sequenz bekannt. Die Polypeptide können auch eine Markierung oder mehrere Markierungen einschließen, die dem Fachmann bekannt sind. In diesem Zusammenhang versteht es sich auch, dass die Proteine weiter durch im Stand der Technik bekannte konventionelle Verfahren modifiziert werden können. Durch zur Verfügung stellen der Proteine ist es auch möglich, Fragmente zu bestimmen, die ihre biologische Aktivität beibehalten, nämlich die reife, prozessierte Form. Das erlaubt die Konstruktion chimärer Proteine und von Peptiden, die eine Aminosäuresequenz umfassen, die vom reifen Protein abgeleitet ist, die essentiell für seine Bindungsaktivität ist. Die anderen funktionellen Aminosäuresequenzen können entweder physisch, beispielsweise chemisch, mit den Proteinen verknüpft werden oder können mittels rEcombinanter DNA-Techniken, die im Stand der Technik bekannt sind, an die Proteine fusioniert werden.

**[0025]** Der Ausdruck „Derivat“, „funktionelles Fragment einer Sequenz“ oder „funktioneller Teil einer Sequenz“ bezeichnet eine verkürzte Sequenz der als Originalsequenz bezeichneten Sequenz. Die verkürzte Sequenz (Nukleinsäure- oder Proteinsequenz) kann in der Länge stark variieren; die minimale Größe ist eine Sequenz von ausreichender Größe, um eine Sequenz zur Verfügung zu stellen mit wenigstens vergleichbarer Funktion und/oder Aktivität wie die Originalsequenz, während die maximale Größe nicht kritisch ist. Bei einigen Anwendungen ist die maximale Größe normalerweise nicht wesentlich größer als die, die benötigt wird, um die gewünschte Aktivität und/oder Funktionen) der Originalsequenz zur Verfügung zu stellen. Typischerweise liegt die verkürzte Aminosäuresequenz in einem Längenbereich von ungefähr 5 bis ungefähr 60 Aminosäuren. Die Sequenz wird jedoch insbesondere eine maximale Länge von ungefähr 50 Aminosäuren, vorzugsweise eine maximale Länge von ungefähr 30 Aminosäuren haben. Normalerweise ist es wünschenswert, Sequenzen von wenigstens ungefähr 10, 12 oder 15 Aminosäuren bis zu einem Maximum von ungefähr 20 oder 25 Aminosäuren auszuwählen. Die Ausdrücke „Gen(e)“, „Polynukleotid“, „Nukleinsäuresequenz“, „Nukleotidsequenz“, „DNA-Sequenz“ oder „Nukleinsäuremolekül(e)“, wie hier verwendet, bezeichnen eine polymere Form von Nukleotiden jeder Länge, entweder Ribonukleotide oder Desoxyribonukleotide. Dieser Ausdruck bezeichnet nur die Primärstruktur des Moleküls. Daher schließt dieser Ausdruck doppel- und einzelsträngige DNA und RNA mit ein. Er schließt auch alle bekannten Modifikationsarten mit ein, beispielsweise Methylierungen und „CAP“-Substitutionen einer oder mehrerer der natürlich vorkommenden Nukleotide mit einem Analogon.

**[0026]** Eine „kodierende Sequenz“ ist eine Nukleotidsequenz, die in mRNA transkribiert und/oder in ein Polypeptid translatiert wird, wenn es unter die Kontrolle geeigneter regulatorischer Sequenzen gebracht wird. Die Grenzen der kodierenden Sequenz werden durch ein Translationsstartkodon am 5'-Terminus und ein Translationsstoppkodon am 3'-Terminus bestimmt.

**[0027]** Eine kodierende Sequenz kann, ist aber nicht beschränkt auf, mRNA, cDNA, rekombinante Nukleotidsequenzen oder genomische DNA einschließen, wobei Introns unter bestimmten Umständen ebenfalls vorliegen können.

**[0028]** Mit „Transkriptionsfaktor“ ist eine Klasse von Proteinen gemeint, die an einem Promotor oder an eine nahe gelegene DNA-Sequenz binden, um die Transkriptionsinitiation zu erleichtern oder zu verhindern.

**[0029]** Mit „Promotor“ wird eine orientierte DNA-Sequenz gemeint, die vom RNA-Polymerase-Holoenzym erkannt wird, um die Transkription zu initiieren. Mit „RNA-Polymerase“ ist ein Enzym mit vielen Untereinheiten gemeint, das RNA komplementär zur DNA-Matrize synthetisiert. Mit „Holoenzym“ ist die aktive Form eines En-

zyms gemeint, das aus vielen Untereinheiten besteht.

**[0030]** Der Ausdruck „Antikörper“ betrifft einen Antikörper, der dadurch gekennzeichnet ist, das er spezifisch gegen einen Transkriptionsfaktor wie SIP1 oder jedes funktionelle Derivat davon gerichtet ist, wobei die Antikörper vorzugsweise monoklonale Antikörper sind; oder ein Antigen-bindendes Fragment davon vom  $F(ab')_2$ -,  $F(ab)$ - oder einzelkettigen Fv-Typ, oder von jedem Typ an rEcombinantem Antikörper, der davon abgeleitet ist. Die erfindungsgemäßen monoklonalen Antikörper können beispielsweise mittels jeden Hybridoms hergestellt werden, das verlässlich nach klassischen Verfahren aus Milzzellen eines Tieres gebildet werden kann, insbesondere von einer Maus oder Ratte, die mit SIP1 oder jedem funktionellen Derivat davon immunisiert wurde, und aus Zellen einer Myelomzelllinie, und nach der Fähigkeit des Hybridoms ausgewählt wird, monoklonale Antikörper zu produzieren, die SIP1 oder jedes funktionelle Derivat davon, das ursprünglich verwendet wurde für die Immunisierung der Mäuse, erkennen. Die erfindungsgemäßen monoklonalen Antikörper dieser Ausführungsform können humanisierte Versionen der monoklonalen Antikörper der Maus sein, die mittels rEcombinanter DNA-Technologie hergestellt werden, ausgehend von den genomischen DNA-Sequenzen von Maus und/oder Mensch, die für H- und L-Ketten kodieren oder von cDNA-Klonen, die für H- und L-Ketten kodieren. Alternativ können die monoklonalen Antikörper humane monoklonale Antikörper sein. Solche humanen monoklonalen Antikörper werden beispielsweise zubereitet mittels Neu-Ansiedlung humaner Lymphozyten des peripheren Blutes (PBL) von Mäusen mit schwerer kombinierter Immunschwäche (SCID) wie in WO 00/60846 beschrieben oder unter Verwendung transgener nicht-humaner Tiere, die in der Lage sind, humane Antikörper herzustellen wie in US-Patent 5,545,806 beschrieben. Auch Fragmente, die von diesen monoklonalen Antikörpern abstammen wie Fab,  $F(ab)_2$  und ssFv („Einzelkettensvariables Fragment“), sind Teil der vorliegenden Erfindung, vorausgesetzt, sie haben die ursprünglichen Bindungseigenschaften beibehalten. Solche Fragmente werden üblicherweise mittels beispielsweise enzymatischem Verdau der Antikörper mit Papain, Pepsin oder anderen Proteasen erzeugt. Es ist dem Fachmann bekannt, dass monoklonale Antikörper oder Fragmente davon für verschiedene Verwendungen modifiziert werden können. Die Antikörper können auch mit einer geeigneten enzymatischen, fluoreszierenden oder radioaktiven Markierung markiert werden.

**[0031]** Der Ausdruck „kleine Moleküle“ betrifft beispielsweise kleine organische Moleküle, und andere Wirkstoffkandidaten, die beispielsweise aus kombinatorischen und natürlichen Produktbibliotheken über Verfahren, die im Stand der Technik bekannt sind, erhalten werden können. Zufällige Peptidbibliotheken, die aus allen möglichen Kombinationen von Aminosäuren bestehen, die an einen Festphasenträger geheftet sind, können verwendet werden, um Peptide zu identifizieren, die SIP1 oder einen Promotorbereich, der von SIP1 gebunden wird, zu binden. Die Durchmusterung von Peptidbibliotheken kann bei der Entdeckung pharmazeutischer Wirkstoffe, die die biologische Aktivität von SIP1 unterbinden, von therapeutischem Wert sein.

**[0032]** Die Ausdrücke „Antisense-Nukleinsäure“ und „Ribozyme“ bezeichnen Moleküle, die die Translation der SIP1-mRNA unterbinden. Antisense-Nukleinsäuren oder Antisense-RNA- und DNA-Moleküle blockieren direkt die Translation der mRNA durch Binden an die Ziel-mRNA und Verhindern der Proteintranslation. Ribozyme sind enzymatische RNA-Moleküle, die in der Lage sind, spezifische Spaltungen von RNA zu katalysieren. Der Wirkmechanismus von Ribozymen schließt die Sequenz-spezifische Hybridisierung des Ribozymmoleküls mit der komplementären Ziel-RNA ein, gefolgt von einer endonukleolytischen Spaltung. Im Rahmen der Erfindung liegen konstruierte Hammerkopfmotiv-Ribozymmoleküle, die spezifisch und effizient die endonukleolytische Spaltung von SIP1-RNA-Sequenzen katalysieren. Spezifische Ribozym-Spaltungsstellen innerhalb jeder potentiellen Ziel-RNA wurden anfänglich durch Scannen des Zielmoleküls auf Ribozym-Spaltungsstellen identifiziert, die die folgenden Sequenzen einschließen: GUA, GUU und GUC. Sobald sie identifiziert sind, können kurze RNA-Sequenzen im Bereich zwischen 15 und 20 Ribonukleotide, die dem Bereich des Zielgens, das die Spaltungsstelle enthält, entsprechen, bezüglich vorhergesagter struktureller Merkmale wie Sekundärstruktur bewertet werden, die die Oligonukleotidsequenz ungeeignet machen könnten. Die Eignung von Kandidatenzielen kann auch bewertet werden durch Testen der Zugänglichkeit der Hybridisierung mit komplementären Oligonukleotiden unter Verwendung von Ribonuklease-Schutzassays. Sowohl die Antisense-RNA- als auch die DNA-Moleküle als auch die Ribozyme der Erfindung können mit jedem im Stand der Technik für die Synthese von RNA-Molekülen bekannten Verfahren zubereitet werden. Diese schließen allgemein bekannte Techniken für die chemische Synthese von Oligodesoxyribonukleotiden wie beispielsweise die Festphasen-Phosphoramiditsynthese mit ein. Alternativ können die RNA-Moleküle mittels in vitro und in vivo-Transkription von DNA-Sequenzen erzeugt werden, die für das Antisense-RNA-Molekül kodieren. Solche DNA-Sequenzen können in eine große Vielzahl an Vektoren eingebaut werden, die geeignete RNA-Polymerase-Promotoren wie den T7- oder den SP6-Polymerase-Promotor beinhalten. Alternativ können Antisense-cDNA-Konstrukte, die Antisense-RNA in Abhängigkeit vom verwendeten Promotor konstitutiv oder induzierbar synthetisieren, stabil in Zelllinien eingeführt werden.



**[0033]** Die oben beschriebenen Antikörper, kleinen Moleküle, Antisense-Nukleinsäuren und Ribozyme können als „ein Medikament“ verwendet werden, um Tumorinvasion und/oder Metastase zu verhindern und/oder zu behandeln, durch Unterbindung der Herunterregulierung der E-Cadherin-Expression durch SIP1. Bösartigkeit von Tumoren impliziert eine inhärente Tendenz der Tumorzellen zu metastasieren (den Körper im großen Maße zu befallen und durch subtile Mittel verbreitet zu werden) und letztendlich den Patienten zu töten, es sei denn, die bösartigen Zellen können ausgelöscht werden. Die Metastasierung ist daher das herausragende Charakteristikum der Bösartigkeit. Metastasierung ist die Tendenz von Tumorzellen, von ihrer Ursprungsstelle über das Kreislaufsystem und andere Kanäle fortgetragen zu werden, was zu einer Einnistung dieser Zellen in beinahe jedem Gewebe und Organ des Körpers letztlich führt. Im Gegensatz dazu verbleiben die Zellen eines gutartigen Tumors immer in Kontakt miteinander in einer kompakten Masse um den Ursprungsort. Aufgrund der physischen Kontinuität gutartiger Tumorzellen können sie vollständig chirurgisch entfernt werden, wenn die Stelle dazu geeignet ist. Die Ausbreitung bösartiger Zellen, von denen jede einzelne individuell (durch Zellteilung) die Eigenschaft besitzt Ausgangspunkt für neue Zellmassen (neue Tumore) an neuen und entfernten Stellen zu sein, schließt die vollständige Auslöschung durch eine einzelne chirurgische Maßnahme in allen mit Ausnahme den frühesten Wachstumsphasen aus. Es sollte klar sein, dass das erfindungsgemäße „Medikament“ in Kombination mit jeder anderen Tumortherapie, die im Stand der Technik bekannt ist, wie Bestrahlung, Chemotherapie oder Chirurgie, verwendet werden kann.

**[0034]** Bezüglich der oben erwähnten kleinen Moleküle betrifft der Ausdruck „Medikament“ eine Zusammensetzung, die kleine Moleküle wie oben beschrieben und einen pharmazeutisch verträglichen Träger oder Exzipienten (beide Ausdrücke können gegenseitig austauschbar verwendet werden) umfasst, um die oben angegebenen Erkrankungen zu behandeln. Geeignete Träger oder Exzipienten, die dem Fachmann bekannt sind, sind physiologische Kochsalzlösung, Ringer-Lösung, Dextrose-Lösung, Hank-Lösung, fette Öle, Ethyloleat, 5% Dextrose in physiologischer Kochsalzlösung, Substanzen, die die Isotonie und chemische Stabilität verstärken, Puffer und Konservierungsstoffe. Andere geeignete Träger schließen jeden Träger ein, der nicht selbst die Produktion von Antikörpern induziert, die schädlich für das Individuum wären, das die Zusammensetzung erhält, wie Proteine, Polysaccharide, Polymilchsäuren, Polyglykolsäuren, polymere Aminosäuren und Aminosäure-Copolymere.

**[0035]** Das „Medikament“ kann verabreicht werden mit jedem geeigneten Verfahren innerhalb des Wissens des Fachmanns. Der bevorzugte Verabreichungsweg ist parenteral. Bei der parenteralen Verabreichung wird das erfindungsgemäße Medikament in einer als Einheitsdosis injizierbaren Form formuliert wie einer Lösung, Suspension oder Emulsion, in Verbindung mit den pharmazeutisch verträglichen Exzipienten wie oben definiert. Die Dosierung und die Art der Verabreichung werden jedoch vom Individuum abhängen. Im Allgemeinen wird das Medikament so verabreicht, dass das erfindungsgemäße Molekül in einer Dosis zwischen 1 µg/kg und 10 mg/kg vorliegt, noch bevorzugter zwischen 10 µg/kg und 5 mg/kg, am bevorzugtesten zwischen 0,1 und 2 mg/kg. Vorzugsweise wird es als Bolus-Dosis gegeben. Kontinuierliche Infusion kann auch verwendet werden und schließt die kontinuierliche subkutane Zufuhr über eine osmotische Minipumpe ein. In diesem Fall kann das Medikament in einer Dosis zwischen 5 und 20 µg/kg/Minute infundiert werden, noch bevorzugter zwischen 7 und 15 µg/kg/Minute.

**[0036]** Bezüglich der Antikörper, der Antisense-Nukleinsäuren und der Ribozyme der vorliegenden Erfindung ist eine bevorzugte Art der Verabreichung des „Medikaments“ für die Behandlung die Verwendung einer Gentherapie, um die oben erwähnten Moleküle zuzuführen. Gentherapie bedeutet die Behandlung mittels Zufuhr der therapeutischen Nukleinsäuren an Zellen des Patienten. Das wird ausführlich in Lever und Goodfellow 1995; Br. Med Bull., 51, 1–242; Culver 1995; Ledley, F. D. 1995. Hum. Gene Ther. 6, 1129, besprochen. Um eine Gentherapie zu erreichen muss es ein Verfahren geben, um die Gene den Zellen des Patienten zuzuführen und zusätzliche Verfahren, um die effektive Herstellung jedes therapeutischen Gens sicherzustellen. Es gibt zwei allgemeine Herangehensweisen, um eine Genzufuhr zu erreichen; diese sind die nicht-virale Zufuhr und die Virus-vermittelte Genzufuhr.

**[0037]** Die folgenden Beispiele veranschaulichen deutlicher bevorzugte Merkmale der Erfindung, sind aber nicht gedacht, die Erfindung in irgendeiner Art und Weise zu beschränken.

## Beispiele

- Charakterisierung von Nukleinsäuresequenzen, die mindestens eine CACCT-Sequenz umfassen.

## Einleitung und Zusammenfassung

- SIP1 und  $\delta$ EF1 binden an Zielstellen, die eine CACCT-Sequenz und eine CACCTG-Sequenz enthalten

**[0038]** Die vorliegende Erfindung betrifft die DNA-Bindeigenschaften von SIP1. Wie oben bemerkt gehört SIP1, ein kürzlich isoliertes, mit Smad-wechselwirkendes Protein, zu der in Erscheinung tretenden Familie der zwei-händigen Zinkfingertranskriptionsfaktoren (34). Die Organisation von SIP1 ist ähnlich zu der von  $\delta$ EF1, dem Prototyp-Mitglied dieser Familie. Beide Proteine enthalten zwei weit voneinander getrennte Cluster von Zinkfingern, die an der DNA-Bindung beteiligt sind. Die Aminosäuresequenzhomologie ist innerhalb dieser zwei Zinkfingercluster sehr hoch (über 90%), wohingegen sie in den anderen Bereichen weniger deutlich ist. Dieses Ergebnis liegt nahe, dass beide Proteine in analoger Weise an ähnliche DNA-Ziele binden würden. Tatsächlich binden SIP1 als auch  $\delta$ EF1 mit vergleichbaren Affinitäten an viele verschiedene Zielstellen, die immer zwei CACCT-Sequenzen enthalten. Für alle hier getesteten Zielstellen ist die Integrität beider CACCT-Sequenzen für die Bindung durch entweder SIP1 oder  $\delta$ EF1 absolut notwendig.

**[0039]** SIP1<sub>FS</sub> unterbindet die Expression von Xbra2, wenn es im Xenopus-Embryo überexprimiert wird (34), und SIP1<sub>FS</sub> bindet an den Xbra2-Promotor durch Kontaktierung von zwei CACCT-Sequenzen. Kürzlich haben Studien, die transgene Embryos von Xenopus verwendeten, gezeigt, dass 2,1 kb der Xbra2-Promotorsequenzen ausreichen, um ein Reporterprotein in derselben Domäne wie Xbra selbst zu exprimieren (17). Eine einzelne Punktmutation innerhalb der stromabwärts gelegenen CACCT-Stelle (Xbra-D) im Promotor, die die Bindung von SIP1 unterbricht (wie man in Gelretardationsassays sieht), hat jedoch einen deutlichen Effekt. Die Expression des Markerproteins beginnt früher (d. h. im Stadium 9) und findet sich nun an ektopischen Stellen, z. B. in der Mehrheit der ektodermalen, mesodermalen und endodermalen Zellen (17). Das zeigt an, dass dieses Nukleotid, das sich innerhalb der stromabwärts gelegenen CACCT-Stelle befindet, für die korrekte räumliche und zeitliche Expression des Xbra2-Gens benötigt wird. Zusätzlich beobachteten wir, wenn eine Mutation in der stromaufwärts gelegenen CACCT-Sequenz eingeführt wird, dieselbe vorzeitige und ektopische Expression von Xbra2 wie für die Mutation innerhalb der stromabwärts gelegenen CACCT-Stelle. Daher führen Mutationen entweder in der stromabwärts oder der stromaufwärts gelegenen CACCT, von denen bekannt ist, dass sie die Bindung von SIP1 oder  $\delta$ EF1 im EMSA beeinflussen, in vivo zu demselben Phänotyp, was anzeigt, dass ein  $\delta$ EF1-ähnliches Protein in Xenopus an der Regulation des Xbra2-Gen beteiligt ist. Zusätzlich unterstützen dieses in vivo-Daten die Schlussfolgerungen aus den in vitro-Bindungsexperimenten, die hier präsentiert sind: SIP1/ $\delta$ EF1-ähnliche Transkriptionsfaktoren benötigen zwei CACCT-Stellen für die Regulation der Expression des Xbra2-Promotors.

**[0040]** Nicht alle Promotorbereiche, die zwei CACCT-Sequenzen enthalten, stellen Bindungsstellen für SIP1 oder  $\delta$ EF1 dar. Insbesondere ist die Duplikation der Xbra-F-Sonde, die die stromaufwärts gelegene CACCT-Sequenz, die im Xbra-WT-Element vorliegt, enthält, unempfindlich für die Bindung durch entweder SIP1 oder  $\delta$ EF1. Darüber hinaus kann weder SIP1<sub>NZF</sub> noch SIP1<sub>CZF</sub> effizient an diese Stelle (Xbra-F) als Monomer oder als Dimer binden. Daher könnten andere Sequenzen zusätzlich zu CACCT für die Erzeugung einer Bindungsstelle mit hoher Affinität benötigt werden. Es scheint so, als ob CACCTG immer eine bessere Zielsequenz für die Bindung durch diese Zinkfingercluster ist. Tatsächlich wurde für die Hochaffinitäts-CACCTG-Stelle (Xbra-E) gezeigt, dass sie entweder den SIP1<sub>NZF</sub>- oder den SIP1<sub>CZF</sub>-Cluster bindet. Zusätzlich beeinflusst die Modifikation der CACCTG-Stelle nach CACCTA die Bindung von SIP1<sub>FS</sub> und  $\delta$ EF1 an den Xbra-Promotor stark, was die Bedeutung dieses 3'-Guaninrestes bestätigt. Durch Vergleich der Sequenzen aller SIP1- und  $\delta$ EF1-Zielstellen wurde eine minimale Konsensussequenz gefunden, die sich aus einer CACCT-Sequenz und einer CACCTG-Sequenz zusammensetzt, was zeigt, dass diese zwei Sequenzen ausreichend sind, um eine Hochaffinitätsbindungsstelle für SIP1 oder  $\delta$ EF1 zu bilden.

**[0041]** Obwohl die stromaufwärts gelegene CACCT-Sequenz nicht in der Lage ist, SIP1<sub>CZF</sub> oder SIP1<sub>NZF</sub> zu binden, wird diese Sequenz durch SIP1 voller Länge im Kontext der Xbra-WT-Sonde kontaktiert. Die stromaufwärts gelegene CACCT-Sequenz ist eine Voraussetzung für die Bindung von SIP1<sub>FS</sub> an die Xbra-WT-Sonde. Daher steht diese Niedrigaffinitätsstelle (Xbra-F), wenn die stromaufwärts gelegene CACCT-Sequenz mit einer weiteren Hochaffinitäts-CACCTG-Stelle (Xbra-E) kombiniert wird, der Bindung von SIP1<sub>FS</sub> zur Verfügung. Ein Modell wird bevorzugt, in dem SIP1<sub>FS</sub> seinen Zielpromotor über die Bindung eines seiner Zinkfingercluster an eine Hochaffinitäts-CACCTG-Sequenz (z. B. Xbra-E) kontaktiert, gefolgt von der Kontaktierung der Niedrigaffinitäts-CACCT-Stelle (Xbra-F) durch den zweiten Cluster folgt und diese zusätzliche Wechselwirkung stabili-

siert die Bindung von SIP1 stark. Daher kann eine CACCT-Stelle immer noch eine wichtige Wirkung bei der Regulation der Genexpression haben obwohl sie von sich aus weder SIP1<sub>NZF</sub>, SIP1<sub>CZF</sub> noch SIP1<sub>FS</sub> bindet.

**[0042]** Für die DC5-Sonde des  $\delta$ 1-Crystallin-Enhancers wurde früher gezeigt, dass sie spezifisch an  $\delta$ EF1 bindet (31). Diese Sonde enthält jedoch nur eine CACCT-Sequenz. Daher kann trotz des hier erbrachten Nachweises, dass eine Hochaffinitätsbindestelle für  $\delta$ EF1 eine CACCT-Sequenz und eine CACCTG-Sequenz enthalten sollte, nicht ausgeschlossen werden, dass in bestimmten Fällen, wie der DC5-Sonde, eine CACCT-Stelle für die Bindung dieses Typs von Transkriptionsfaktor ausreichend sein würde.

#### Art der DNA-Bindung durch SIP1

**[0043]** Wenn sie unabhängig in EMSAs getestet wurden, binden sowohl die C-terminalen als auch die N-terminalen Zinkfingercluster von SIP1 oder  $\delta$ EF1 an sehr ähnliche CACCT-enthaltende Konsensussequenzen. Sowohl bei SIP1 als auch bei  $\delta$ EF1 teilen NZF3 und NZF4 eine umfangreiche Aminosäuresequenzhomologie mit CZF2 bzw. CZF3. Diese Homologie könnte erklären, warum diese zwei Cluster an ähnliche Konsensussequenzen binden können. Zusätzlich ist gezeigt worden, dass SIP1 oder  $\delta$ EF1 zwei CACCT-Sequenzen für die Bindung an mehrere potentielle Zielstellen benötigt. Ausgehend von diesen Ergebnissen wird vorgeschlagen, dass SIP1 und  $\delta$ EF1 an ihre Zielelemente derart binden würden, dass ein Zinkfingercluster eine der CACCT-Stellen kontaktiert, während der andere Cluster die zweite CACCT-Stelle kontaktiert (siehe [Fig. 2](#), Modell 1). Ein alternatives Modell wäre, dass SIP1 oder  $\delta$ EF1 homodimerisiert, bevor sie in der Lage sind, diese Zielstellen mit hoher Affinität zu binden (Modell 2). Die DNA-Bindekapazität von SIP1<sub>NZF</sub> verschwindet bei Mutationen entweder in NZF3 oder NZF4. In ähnlicher Weise beeinflussen Mutationen innerhalb von CZF2 oder CZF3 die Bindungskapazität von SIP1<sub>CZF</sub>. Wenn diese Mutationen im Rahmen von SIP1 der vollen Länge eingeführt werden, kann keine Bindung von SIP1<sub>FS</sub> mehr beobachtet werden. Das zeigt klar an, dass die Bindungsaktivität beider Zinkfingercluster für die Bindung von SIP1<sub>FS</sub> an sein Zielelement, das eine Dublette der CACCT-Stellen enthält, benötigt wird. Ebenso wurde früher gezeigt, dass die Integrität beider Zinkfingercluster von  $\delta$ EF1 auch für die Bindung von DNA notwendig ist (31). Diese Beobachtungen zeigen an, dass beide Zinkfingercluster direkt die DNA kontaktieren. Daher sollte im Dimer-Modell ([Fig. 2](#), Modell 2), SIP1<sub>NZF</sub> eines SIP1-Moleküls eine CACCT-Sequenz binden und SIP1<sub>CZF</sub> des zweiten SIP1-Moleküls sollte die andere CACCT-Sequenz kontaktieren. Falls solch eine Dimer-Konfiguration existieren würde, dann kann davon ausgegangen werden, dass bestimmte Kombinationen von SIP1-Molekülen voller Länge mit verschiedenen Mutationen innerhalb von CZF bzw. NZF die Bildung eines funktionellen Dimers erlauben sollten, das in der Lage ist, an seine Ziel-DNA zu binden. Keine dieser möglichen Kombinationen der vier getesteten SIP1<sub>FS</sub>-Mutanten (NZF3mut, NZF4mut, CZF2mut und CZF3mut) führte zu einem DNA/SIP1-Komplex in EMSAs. Das spricht gegen die Existenz der SIP1-Dimere. Darüber hinaus war es durch Verwendung verschieden markierter SIP1<sub>FS</sub>-Moleküle nicht möglich, SIP1-Dimere in EMSAs nachzuweisen, noch war es möglich, solche dimeren Komplexe mit verschiedenen Antikörpern zu supershiften. Daher werden Beweise für Modell 1 zur Verfügung gestellt, in dem SIP1 als Monomer an eine Zielstelle bindet, die eine CACCT-Sequenz und eine CACCTG-Sequenz enthält.

**[0044]** Es wurde in dieser Erfindung gezeigt, dass weder die relative Orientierung der zwei CACCT-Sequenzen noch der Abstand zwischen diesen Sequenzen kritisch für die Bindung von SIP1<sub>FS</sub> oder  $\delta$ EF1 ist. Das zeigt, dass diese Transkriptionsfaktoren eine hochgradig flexible Sekundärstruktur aufweisen sollten, um die Bindung an diese verschiedenen Zielstellen zu bewerkstelligen. Der lange Linker-Bereich zwischen den zwei Zinkfingerclustern innerhalb von SIP1 und  $\delta$ EF1 könnte diese Flexibilität in der Sekundärstruktur dieser Proteine erlauben. Diese Transkriptionsfaktoren können an Stellen binden, die CACCT-Sequenzen, die durch mindestens 44 bp (Ecad-WT) getrennt sind, binden, was vermuten lässt, dass ein Bereich von ungefähr 50 bp der Promotorsequenzen bedeckt sein könnte, und daher weniger zugänglich für transkriptionelle Aktivatoren sein könnte, sobald einmal SIP1<sub>FS</sub> oder  $\delta$ EF1 an diesen Promotor gebunden ist. Das zeigt an, dass SIP1 oder  $\delta$ EF1 als transkriptioneller Repressor durch Konkurrenz mit transkriptionellen Aktivatoren, die an diesen von SIP1 oder  $\delta$ EF1 bedeckten Bereich binden, wirken könnte.

– Andere Transkriptionsfaktorfamilien können DNA über einem ähnlichen Mechanismus wie SIP1 binden

**[0045]** Diese neue DNA-Bindungsart kann auch für andere Transkriptionsfaktorfamilien verallgemeinert werden, die, wie SIP1 und  $\delta$ EF1, voneinander getrennte Zinkfingercluster enthalten, wie jene der MBP/PRDII-BF1-Familie (1, 3, 6, 29, 33). Wie bei SIP1 und  $\delta$ EF1 ist die Konservierung dieser Zinkfingercluster sehr stark innerhalb der verschiedenen Mitglieder dieser Familie (1). Zusätzlich ist der C-terminale Cluster sehr homolog zum N-terminalen Cluster und, im Fall von PRDII-BF1, binden diese Cluster dieselben Sequenzen, wenn sie unabhängig voneinander getestet werden (3). Daher könnte diese Art von Transkriptionsfaktor an

zwei wiederholte Sequenzen durch die Kontaktierung eines Zinkfingerclusters einer Sequenz und die Kontaktierung eines zweiten Zinkfingerclusters der zweiten Sequenz binden. Ebenso haben die verschiedenen Mitglieder der NZF-Transkriptionsfaktorfamilie zwei weit voneinander getrennte Zinkfingercluster (11, 12, 36). MyT1, NZF-1 und NZF-3 binden alle an dasselbe Konsenselement, AAAGTTT. Wie bei SIP1 und  $\delta$ EF1, die eine wesentlich höhere Affinität für Elemente zeigen, die zwei CACCT-Sequenzen enthalten, zeigte ein Element, das zwei AAAGTTT-Sequenzen enthält, eine merklich höhere Affinität für NZF-3 (36). Das lässt vermuten, dass zwei AAAGTTT-Sequenzen auch notwendig sind, um eine Hochaffinitätsbindestelle für diese Transkriptionsfaktoren zu erzeugen, und dass sie DNA über einen ähnlichen Mechanismus wie SIP1 und  $\delta$ EF1 binden könnten. Schließlich bindet das Evi-1-Protein, das 7 Zinkfinger am N-Terminus und 3 Zinkfinger am C-Terminus enthält, an zwei Konsensussequenzen. Es bindet an eine komplexe Konsensussequenz (GACAA-GATAAGATAA-N<sub>1-28</sub>-CTCATCTTC) über einen Mechanismus, der die Bindung des N-terminalen Zinkfingerclusters an den ersten Teil und die Bindung des C-terminalen Clusters an den zweiten Teil umfassen könnte (20). Schließlich könnte die Art der DNA-Bindung, wie sie hier beschrieben ist, nicht nur auf die SIP1/ $\delta$ EF1-Transkriptionsfaktorfamilie anwendbar sein, sondern viel allgemeiner.

**[0046]** SIP1 wurde als mit Smad1-wechselwirkendes Protein kloniert, es wurde aber auch gezeigt, dass es mit Smad2, 3 und 5 wechselwirkt (34). Smad-Proteine sind Signal-Weiterleiter, die an der BMP/TGF- $\beta$ -Signalkaskade beteiligt sind (13). Nach Bindung von TGF- $\beta$ -Liganden an den Serin/Threonin-Kinase-Rezeptor-Komplex werden die Rezeptor-regulierten Smad-Proteine durch Typ1-Rezeptoren phosphoryliert und wandern in den Nukleus, wo sie die Transkription von Zielgenen modulieren. Die Wechselwirkung zwischen SIP1 und Smads beobachtet man nur nach Ligandenstimulation, was anzeigt, dass Smads aktiviert sein müssen, bevor sie in der Lage sind, mit SIP1 wechselzuwirken (34). Überraschenderweise ist Evi-1, ein Transkriptionsfaktor, der DNA über einen ähnlichen Mechanismus wie SIP1 bindet, ein mit Smad3 wechselwirkendes Protein (15). Bis dato wurde gezeigt, dass Evi-1 die Bindung von Smad3 an DNA stört, aber sicherlich eine Wirkung auf Ziel promotoren von Evi-1 hat. Schnurri, das das Drosophila-Homolog zum humanen PRD-II-Bf1-Transkriptionsfaktor ist, ist ein Protein, das auch DNA über einen ähnlichen Mechanismus wie das SIP1-Protein binden könnte. Interessanterweise wurde für Schnurri vorgeschlagen, dass es ein nukleäres Zielprotein im dpp-Signalweg ist (1, 6). Dpp ist ein Mitglied der TGF- $\beta$ -Familie. Das macht Schnurri zu einem nukleären Zielkandidaten für das Drosophila Mad-Protein, dem Drosophila-Homolog der Smads aus Vertebraten. Daher kann die von SIP1 verwendete Art der DNA-Bindung für andere Zinkfinger-enthaltende Proteine, die mit Smad wechselwirken, verallgemeinert werden, und stellt ein gemeinsames Merkmal verschiedener Smad-Partner im Nukleus dar.

**[0047]** Aufbauend auf diesen Ergebnissen wird eine neuartige Art der DNA-Bindung für die  $\delta$ EF1-Transkriptionsfaktorfamilie gezeigt. Diese Art der DNA-Bindung ist auch relevant für andere Transkriptionsfaktorfamilien, die voneinander getrennte Zinkfingercluster enthalten.

Material und Methoden, die in diesem Beispiel verwendet wurden

Plasmidkonstruktionen.

**[0048]** Für die Expression in Säugerzellen wurden die cDNAs von SIP1 (34) und  $\delta$ EF1 (5) im pCS3 (27) subkloniert. In diesem Plasmid wurden die offenen Leserahmen von SIP1 und  $\delta$ EF1 an einen (Myc)<sub>6</sub>-Tag am N-Terminus fusioniert. Die SIP1-cDNA wurde auch in pCDNA3 (Invitrogen) als eine N-terminale Fusion mit dem FLAG-Tag kloniert. Für die Expression von SIP1<sub>NZF</sub> und SIP1<sub>CZF</sub> subklonierten wir die cDNA-Fragmente, die für die Aminosäuren 1 bis 398 bzw. 977 bis 1214 kodieren, in pCS3. SIP1<sub>CZF</sub> (als Aminosäuren 957 bis 1156) und SIP1<sub>NZF</sub> (Aminosäuren 90 bis 383) wurden in E. coli als GST-Fusionsprotein (in pGEX-5X-1, Pharmacia) hergestellt und unter Verwendung des GST-Reinigungsmoduls (Pharmacia) aufgereinigt.

**[0049]** Identische Mutationen zu jenen, die in AREB6 (10) gemacht wurden, wurden auch in die SIP1-Zinkfinger eingeführt. Die Mutagenese der Zinkfinger NZF3, NZF4, CZF2 und CZF3 umfasst die Substitution ihres dritten Histidins durch ein Serin. Diese Mutationen wurden mittels eines PCR-basierten Ansatzes mit den folgenden Primern eingeführt: SIP1<sub>NZF3MUT</sub>, 5'-CCACCTGAAAGAATCCCTGAGAATTCACAG; SIP1<sub>NZF4MUT</sub>, 5'-GGGTCCTACAGTTCATCTATCAGCAGCAAG; SIP1<sub>CZF2MUT</sub>, 5'-CACACCTTATCGAGTCCTCGAGGCTGCAC; SIP1<sub>CZF3MUT</sub>, 5'-TCCTACTCGCAGTCCATGAATCACAGGTAC. Die entsprechenden mutierten Cluster wurden in pCS3 in SIP1 voller Länge rekloniert, um in Säugertierzellen die mutierten SIP1-Proteine namens NZF3mut, NZF4mut, CZF2mut bzw. CZF3mut herzustellen. Darüber hinaus wurden diese mutierten Cluster in pGEX-5X2 (Pharmacia) subkloniert und in E. coli als GST-Fusionsprotein hergestellt (GST-NZF3mut, GST-NZF4mut, GST-CZF2mut und GST-CZF3mut). Alle Konstrukte wurden über Restriktionskartierung und Sequenzierung bestätigt.

## Zellkultur und DNA-Transfektion.

**[0050]** COS1-Zellen wurden in mit 10% fötalem Rinderserum ergänzten DMEM kultiviert. Die Zellen wurden mittels Fugene gemäß den Anweisung des Herstellers (Boehringer Mannheim) transfiziert und 30–48 Stunden nach der Transfektion geerntet.

## Gelretardationsassay.

**[0051]** Das Xbra-WT-Oligonukleotid deckt den Bereich von –344 bis –294 des Xbra2-Promotors ab (16). Der Bereich zwischen –412 bis –352 des  $\alpha$ 4-Integrinpromotors liegt innerhalb des  $\alpha$ 4I-WT-Oligonukleotids (26). Die Ecad-WT-Sonde enthält den Bereich zwischen –86 bis –17 des humanen Ecad-Promotors (2). Die Sequenzen des oberen Strangs der Wildtypen und die mutierten doppelsträngigen Sonden sind in Tabelle 1 aufgelistet. Doppelsträngige Oligonukleotide wurde mit [ $^{32}$ P]- $\gamma$ -ATP und T4-Polynukleotidkinase (New England Biolabs) markiert. Von COS1-Zellen (25), die mit verschiedenen pCS3-Vektoren, die die Synthese von SIP1 voller Länge,  $\delta$ EF1 voller Länge und verschiedenen mutanter Formen von SIP1 (25) oder die simultane Produktion von gleichen Mengen an Myc-markierten SIP1 und FLAG-markierten SIP1 erlauben, transfiziert waren, wurden Gesamtzellextrakte zubereitet. GST-SIP1-Fusionsproteine wurden aus einem E. coli-Extrakt mittels GST-Reinigungsmoduls (Pharmacia) aufgereinigt und in der Gelretardation getestet. Der DNA-Bindeassay (20  $\mu$ l) wurde bei 25°C mit 1  $\mu$ g des Gesamtproteins der COS1-Zellen, 1  $\mu$ g Poly-dI-dC, 10 pg  $^{32}$ P-markiertem doppelsträngigen Oligonukleotid (ungefähr  $10^4$  Cerenkov-Ereignisse) im zuvor beschriebenen  $\delta$ EF1-Bindepuffer (30) durchgeführt. Für Supershift-Experimente wurden die Extrakte mit Anti-Myc-(Santa Cruz) oder Anti-FLAG-(Kodak)Antikörpern inkubiert. Für die Kompetition wurde ein Überschuss an unmarkierten doppelsträngigen Oligonukleotiden zusammen mit der markierten Sonde zugesetzt. Die Bindungsreaktion wurde auf ein 4%-iges Polyacrylamid-Gel (Acrylamid/Bis-Acrylamid, 19:1), das in  $0,5 \times$  TBE-Puffer zubereitet wurde, geladen. Nach der Elektrophorese wurden die Gele getrocknet und auf einem Röntgenfilm geliegt. Alle Experimente wurden mindestens 3 Mal wiederholt.

## Methylierungsinterferenz-Assay.

**[0052]** Der obere und der untere Strang der Xbra-WT-Sonde wurden separat voneinander markiert und mit einem Überschuss komplementären DNA-Strangs angelagert. Die Sonden wurden präzipitiert und mit Dimethylsulfat (8) behandelt. Die methylierte Sonde ( $10^5$  Cerenkov-Ereignisse) wurde in einer  $10 \times$  Gelretardationsreaktion (siehe oben) (200  $\mu$ l Endvolumen) mit 10  $\mu$ g Gesamtzellextrakt aus COS1-Zellen, die entweder SIP1<sub>FS</sub> oder SIP1<sub>CZF</sub> exprimieren, inkubiert. Nach 20 minütiger Inkubation bei 25°C wurden die Produkte auf ein 4%-iges Polyacrylamid-Gel geladen und eine Elektrophorese wurde wie für den Gelretardationsassay durchgeführt. Anschließend wurde das Gel auf DEAE-Zellulosemembran geblottet; der Transfer wurde bei 100 V für 30 Min. in  $0,5 \times$  TBE-Puffer durchgeführt. Die Membran wurde dann eine Stunde lang exponiert und die Banden, die SIP1<sub>FS</sub> (oder SIP1<sub>CZF</sub>) entsprachen, und die freie Sonde wurden bei 65°C unter Hochsalzbedingungen (1 M NaCl, 20 mM Tris, pH 7,5, 1 mM EDTA) eluiert. Die eluierte DNA wurde präzipitiert und mit Piperidin behandelt (18). Nach mehreren Zyklen der Auflösung in Wasser und Verdampfen der Flüssigkeit im Vakuum wurde das resultierende DNA-Pellet in 10  $\mu$ l Sequenzierpuffer (97,5% deionisiertes Formamid, 0,3% von Bromphenol Blau und Xylencyanol, 10 mM EDTA) aufgelöst und für 5 Min. bei 85°C denaturiert. Die gleiche Anzahl an Ereignissen (1.500 Cerenkov-Ereignisse) für die freie Sonde und die gebundene Sonde wurden auf ein 20%-iges Polyacrylamid-8M Harnstoff-Sequenziergel geladen. Das Gel lief in  $0,5 \times$  TBE für eine Stunde bei 2.000 V. Danach wurde das Gel in 50% Methanol/10% Essigsäure fixiert und getrocknet. Das Gel wurde dann zur Autoradiographie exponiert.

## Western-Blot-Analyse.

**[0053]** Transfizierte Zellen wurden mit PBS-O (137 mM NaCl, 2,7 mM KCl, 6,5 mM Na<sub>2</sub>HPO<sub>4</sub>, 1,5 mM KH<sub>2</sub>PO<sub>4</sub>) gewaschen, in Ablösepuffer (10 mM Tris pH 7,5, 1 mM EDTA, 10% Glycerin, mit Protease-Inhibitoren (Protease-Inhibitor-Cocktailtablette, Boehringer Mannheim)) gesammelt und mittels Zentrifugation bei niedriger Geschwindigkeit pelletiert. Die Zellen wurden dann in 10 mM Tris, pH 7,4, 125 mM NaCl, 1% Triton X-100 aufgelöst. Für die direkte elektrophoretische Analyse wurden Gelprobenpuffer zu den Zelllysaten gegeben und die Proben wurden gEcocht. Für andere Experimente wurden die Lysate erst einer Immunpräzipitation mit entweder Anti-Myc- oder Anti-FLAG-Antikörpern unterworfen. Antikörper wurden den Aliquots der Zelllysate zugesetzt, die über Nacht bei 4°C inkubiert wurden. Die Antikörper und die gebundenen Proteine (das gebundene Protein) der Zelllysate wurden miteinander als ein Komplex an Proteine A-Sepharose für 2 Stunden bei 4°C gEcopelt. Die Immunpräzipitate wurden 4 Mal mit NET-Puffer (50 mM Tris pH 8,0, 150 mM NaCl, 0,1% NP40, 1 mM EDTA, 0,25% Gelatine) gewaschen, mittels SDS-Polyacrylamid (7,5%) Gel-Elektrophorese aufgetrennt

und elektrophoretisch auf Nitrozellulosemembranen transferiert. Die Membranen wurden 2 Stunden lang in TBST (10 mM Tris pH 7,5, 150 mM NaCl, 0,1% Tween-20) mit 3% (w/v) Magermilch blockiert und mit Primärantikörper (1 µg/ml) für 2 Stunden inkubiert, gefolgt von Sekundärantikörper (0,5 µg/ml) gebunden an Meerrettichperoxidase. Immunreaktive Banden wurden mit verstärktem Chemilumineszenz-Reagenz (NEN) nachgewiesen.

#### Xenopus laevis-Transgene und Whole-mount in situ-Hybridisierung

**[0054]** Xenopus-Embryos, die für Xbra2-GFP transgen sind, wurden wie zuvor beschrieben erzeugt (Kroll und Amaya, 1996), mit den folgenden Modifikationen. Ein Drummond Nanoinject wurde für die Injektion eines festgesetzten Volumens von 5 nl Spermienkern-Suspension pro Ei verwendet, bei einer theoretischen Konzentration von 2 Kernen pro 5 nl. NotI wurde für die Plasmidlinearisierungen und für das Einführen von Einzelstrangbrüchen in den Spermienkern verwendet. Ungefähr 800 Eier wurden pro Eierextraktinkubation injiziert. Die Prozedur führte zu einer erfolgreichen Spaltung des Embryos, die im Bereich zwischen 10% und 30% liegt. Von diesen beendeten 50 bis 80% die Gastrulation und 20 bis 30% entwickelten sich in normale schwimmende Kaulquappen weiter, wenn es gestattet wurde. Die transgene Frequenz, die über die Expression analysiert wurde, variierte zwischen 50 bis 90%. Die Embryos wurden Stadien zugeordnet gemäß Niewkoop und Faber (Niewkoop und Faber, 1967). Ein Minimum von 30 exprimierenden Embryos wurde pro Konstrukt analysiert und die Stadien gezeigt. Whole-mount in situ-Hybridisierung für das GFP-Reportergen wurde bereits früher beschrieben (Latinkic et al., 1997). Nach dem Farbnachweis wurden die Embryos entwässert und mit einem 2:1-Gemisch von Benzylalkohol/Benzylbenzoat geklärt.

Tabelle 1.

Oligo	Sequenz	Zwischenabstand
Xbra-WT	ATCCAGGCCACCTAAAATATAGAATGATAAAAGTGACCAGGTGTCAGTTCT	24
Xbra-D	-----A-----	
Xbra-E	TAAAGTGACCAGGTGTCAGTTCT	
Xbra-F	ATCCAGGCCACCTAAAATATAGAATGA	
Rdm + Xbra-E	CAATTTAGAGTACTGTGACTTGGGAGTAAAGTGACCAGGTGTCAGTTCT	
Xbra-F + AREB5	ATCCAGGCCACCTAAAATATAGAATGAGGCTCAGACAGGTGTAGAATTCGGCG	23
Rdm + AREB5	CAATTTAGAGTACTGTGACTTGGGAGGCTCAGACAGGTGTAGAATTCGGCG	
Xbra-WT	ATCCAGGCCACCTAAAATATAGAATGATAAAAGTGACCAGGTGTCAGTTCT	24
Xbra-J	CGA-----	
Xbra-K	---ACT-----	
Xbra-L	-----TAA-----	
Xbra-M	-----CAA-----	
Xbra-N	-----GCC-----	
Xbra-O	-----CCG-----	
Xbra-P	-----CGC-----	
Xbra-Q	-----TCC-----	
Xbra-R	-----GTC-----	
Xbra-S	-----T-----	
Xbra-Z	-----T-----	
Xbra-WT	ATCCAGGCCACCTAAAATATAGAATGATAAAAGTGACCAGGTGTCAGTTCT	24
Xbra-B	ATCCAGGCCACCTA__TATAGAATGATAAAAGTGACCAGGTGTCAGTTCT	21
Xbra-C	ATCCAGGCCACCTAAAATATAGAATGAT__GTGACCAGGTGTCAGTTCT	21
Xbra-U	ATCCAGGCCACCTAAAATATA__GTGACCAGGTGTCAGTTCT	14
Xbra-EE	TAAAGTGACCAGGTGTCAGTTCTTAAAGTGACCAGGTGTCAGTTCT	18
Xbra-ErE	AGAACTGACACCTGGTCACTTTATAAAAGTGACCAGGTGTCAGTTCT	20
Xbra-ErF	ATCCAGGCCACCTAAAATATAGAATATTCTATATTTTAGGTGGCTGGAT	24
Xbra-V	ATCCAGGCAGGTGTAATATAGAATGATAAAAGTGACCCACCTACAGTTCT	24
Xbra-W	ATCCAGGCAGGTGTAATATAGAATGATAAAAGTGACCAGGTGTCAGTTCT	24
<hr/>		
$\alpha$ 4I-WT	GCAGGGCACACCTGGATTGCATTAGAATGAGACTCACTACCCAGTTCAGGTGTGTTGCGT	34
$\alpha$ 4I-A	-----A-----	
$\alpha$ 4I-B	-----T-----	
<hr/>		
Ecad-WT	TGGCCGGCAGGTGAACCCCTCAGCCAATCAGCGGTACGGGGGGGGTCTCCGGGGCTCACCTGGCTGCAG	44
Ecad-A	-----T-----	
Ecad-B	-----A-----	

**[0055]** Tabelle 1. Liste aller in dieser Untersuchung verwendeten Sonden. Die CACCT-Sequenzen sind in Fettdruck markiert. Der Zwischenabstand (rechte Spalte) ist die Anzahl an Nukleotiden, die zwischen den zwei CACCT-Sequenzen vorliegt. Unterstrichene Lücken entsprechen Nukleotiddeletionen aus den Wildtyp-Sonden. Für viele Sonden sind nur die Reste angegeben, die im Vergleich zu den Wildtyp-Sonden verändert wurden, um die Interpretation der eingeführten Mutationen zu erleichtern.

**[0056]** Die folgenden 8 Absätze enthalten einige zusätzliche – Materialien und Methoden –, um die weiterhin beschriebenen Experimente durchzuführen:

- Gelretardationsassay mit verschiedenen Sonden für den Xbra2-Promotor. Die verschiedenen  $^{32}$ P-markierten Xbra-Sonden (10 pg) wurden mit 1  $\mu$ g Gesamtproteinextrakt von COS1-Zellen, die mit pCS3-SIP1<sub>CZF</sub>, oder mit pCS3-SIP1<sub>FS</sub> transfiziert worden waren oder von Kontroll-transfizierten Zellen, inkubiert.
- Zwei CACCT-Stellen werden nach der Bindung von SIP1<sub>FS</sub> an den Xbra2-Promotor kontaktiert. Nur Mutationen innerhalb der stromaufwärts gelegenen CACCT-Sequenz (wie offenbart durch Scanning-Mutagenese, siehe Tabelle 1) oder innerhalb der stromabwärts gelegenen CACCT-Sequenz (siehe anderenorts in Tabelle 1) von XbraWT beseitigt die SIP1<sub>FS</sub>-Bindung. Der Methylierungsinterferenzassay zeigt an, dass SIP1<sub>FS</sub> beide CACCT-Sequenzen kontaktiert. XbraWT, entweder markiert im oberen oder im unteren Strang, wurde methyliert und mit Gesamtextrakt von COS1-Zellen, die entweder mit pCS3-SIP1<sub>FS</sub> oder

pCS3-SIP1<sub>CZF</sub> transfiziert worden waren, inkubiert. Die DNA retardierte in einem geschifteten Komplex oder die ungebundene DNA (FREI) wurde gereinigt, mit Piperidin gespalten und auf einem Sequenziergel laufen gelassen. Guanin-Reste sind in der freien Sonde methyliert. Die stromaufwärts gelegene und die stromabwärts gelegene CACCT-Stelle des Xbra2-Promotors ist angezeigt.

– Zwei CACCT-Sequenzen sind für die Bindung von SIP1<sub>FS</sub> und  $\delta$ EF1 an die Xbra2-, die  $\alpha$ 4-Integrin- und die E-Cadherin-Promotoren notwendig.  $\delta$ EF1 bindet an den Xbra2-Promotor; SIP1 und  $\delta$ EF1 binden an den  $\alpha$ 4-Integrin-Promotor; die Bindung von SIP1 und  $\delta$ EF1 an den  $\alpha$ 4-Integrin-Promotor einschließlich der Kompetition mit einem Überschuss nicht markierter Wildtyp- und mutierter Bindungsstellen; Bindung von SIP1 und  $\delta$ EF1 an den E-Cadherin-Promotor. In jeder Bindungsreaktion wurden 10 pg markierter Sonden mit 1  $\mu$ g Gesamtzellproteinextrakt, der von COS1-Zellen, die entweder mit pCS3-SIP1<sub>FS</sub> oder mit pCS3- $\delta$ EF1 transfiziert worden waren, zubereitet worden ist, inkubiert. In den Kompetitionsexperimenten wurden 5 ng und 50 ng unmarkierter DNA zur selben Zeit wie die markierte Sonde zugesetzt. Gegen die Myc-Markierung gerichteter Antikörper wurde zur Bindungsreaktion und zum supergeschifteten Komplex zugesetzt.  $\delta$ EF1 und die durch SIP1 retardierten Komplexe wurden gezeigt. Für die Sequenzen aller Sonden siehe Tabelle 1.

– Der Zwischenabstand und die relative Orientierung der CACCT-Sequenzen sind nicht kritisch für die Bindung von SIP1<sub>FS</sub> und  $\delta$ EF1 an den Xbra2-Promotor. Zehn pg markierter Sonden wurden mit 1  $\mu$ g Gesamtzellproteinextrakt, der von COS1-Zellen, die entweder mit pCS3-SIP1<sub>FS</sub> oder mit pCS3- $\delta$ EF1 transfiziert worden waren, zubereitet worden ist, inkubiert. Wir verwendeten 10 pg der Xbra-E-Sonde und 10 pg der Xbra-F-Sonde in der gleichen Bindungsreaktion. Aus Gründen einer klaren und vergleichbaren Darstellung, ließen wir die freie Sonde bei den SIP1-Bindungsreaktionen aus.

– Die Integrität beider SIP1-Zinkfingercluster ist für die Bindung von SIP1<sub>FS</sub> an DNA notwendig. Mutationen innerhalb von NZF3, NZF4, CZF2, CZF3 beseitigen die DNA-Bindeaktivität entweder des SIP1<sub>NZF</sub>- oder des SIP1<sub>CZF</sub>-Zinkfingerclusters. Die Wildtyp- und die mutierten Zinkfingercluster wurden an GST fusioniert und die Fusionsproteine wurden in E. coli hergestellt. Nach der Reinigung wurde eine gleiche Menge eines jeden Fusionsproteins (0,1 ng) mit 10 pg markierter Xbra-E-Sonde inkubiert. Mutationen innerhalb von NZF3, NZF4, CZF2 oder CZF3 beeinflussten die Bindung von SIP1<sub>FS</sub> an die Xbra-WT-Sonde. Zehn pg markierter Xbra-WT-Sonde wurden mit 1  $\mu$ g Gesamtzellproteinextrakt, der von COS1-Zellen, die entweder mit pCS3-SIP1<sub>FS</sub>, pCS3-SIP1<sub>NZF3mut</sub>, pCS3-SIP1<sub>NZF4mut</sub>, pCS3-SIP1<sub>CZF2mut</sub> oder pCS3-SIP1<sub>CZF3mut</sub> transfiziert worden waren, zubereitet worden ist, inkubiert. Alle möglichen Kombinationen von 2 COS-Zellextrakten (jeweils 1  $\mu$ g), die die verschiedenen SIP1-Mutanten exprimieren, wurden getestet. Gegen die Myc-Markierung gerichteter Antikörper wurde zur Bindungsreaktion zugegeben und der supergeschiftete Komplex und der SIP1<sub>FS</sub>-retardierte Komplex sind angezeigt. Mutationen innerhalb von NZF3, NZF4, CZF2 oder CZF3 beseitigen die Bindung von SIP1<sub>FS</sub> an den  $\alpha$ 4-Integrin-Promotor. Zehn pg markierte  $\alpha$ 4I-WT-Sonde wurden mit 1  $\mu$ g Gesamtzellproteinextrakt, der von COS1-Zellen, die entweder mit pCS3-SIP1<sub>FS</sub>, pCS3-SIP1<sub>NZF3mut</sub>, pCS3-SIP1<sub>NZF4mut</sub>, pCS3-SIP1<sub>CZF2mut</sub> oder pCS3-SIP1<sub>CZF3mut</sub> transfiziert worden waren, zubereitet worden ist, inkubiert. Der gegen die Myc-Markierung gerichteter Antikörper wurde zur Bindungsreaktion zugegeben und der supergeschiftete Komplex und der SIP1<sub>FS</sub>-retardierte Komplex sind angezeigt. SIP1-Mutanten werden in vergleichbaren Mengen in COS-Zellen produziert. Zehn  $\mu$ g der Gesamtextrakte aus COS-Zellen wurden mittels Western-Blots unter Verwendung des Anti-Myc-Antikörpers analysiert. Die Expressionsniveaus der SIP1-Mutanten sind in der Tat leicht höher als das Expressionsniveau von SIP1-WT.

– SIP1<sub>FS</sub> bindet als Monomer an die Xbra-WT-Sonde. 10 pg markierter Xbra-WT-Sonde wurden mit 1  $\mu$ g Gesamtzellprotein, zubereitet von COS1-Zellen, die mit einer gleichen Menge an pCS3-SIP1<sub>FS</sub> (Myc-markiert) und pCDNA3-SIP1 (FLAG-markiert) transfiziert worden waren, inkubiert. Anti-FLAG- und Anti-Myc-Antikörper wurden separat zugesetzt oder es wurden sowohl Anti-FLAG- als auch Anti-Myc-Antikörper zum Bindungsassay zugesetzt. Die FLAG- und die Myc-supergeschifteten Komplexe sind angezeigt.

– Die Integrität von CZF oder NZF ist für die SIP1-Repressoraktivität notwendig. SIP1<sub>FS</sub>-Bindung an ein Gel-gereinigtes Fragment, das von dem multiple CACCT-enthaltenden künstlichen Promotor vom Reporterplasmid p3TP-Lux abgeleitet ist. Anti-Myc-Antikörper wurde zugesetzt; der supergeschiftete Komplex ist angezeigt. Kotransfektionsassay von pCS3-SIP1<sub>FS</sub>, pCS3-CZF3-Mut oder pCS3-NZF3-Mut zusammen mit dem p3TP-Lux-Reportervektor. Die Aktivität ist angegeben als Prozentsatz voller SIP1<sub>FS</sub>-Repressoraktivität, die 100% beträgt.

– Ektopische Aktivität der mutierten Xbra2-Promotorvarianten (Xbra2-Mut) in transgenen Froschembryos. SIP1<sub>FS</sub>-Bindung an die Wildtyp- und die mutierten (Xbra-Mut; siehe Tabelle 1) Xbra2-Promotorelemente. Whole-mount in situ-Hybridisierung für GFP-mRNA von Xenopus-Embryos, die transgen für ein Wildtyp- oder punktmutiertes 2,1 kb Xbra2-Promotorfragment sind, dass einen GFP-Reporter antreibt. Alle Embryonen wurden im Stadium 11 fixiert und für eine bessere Visualisierung des Signals geklärt. Prozentangaben sind kennzeichnend für intermediären Phänotyp (d. h. 35% der transgenen Embryos zeigten, dass normale Xbra2-Expressionsmuster und 65% zeigten ektopische Expression).



– SIP1 hat eine zu  $\delta$ EF1 ähnliche Struktur

**[0057]** SIP1 wurde kürzlich als ein mit Smad-wechselwirkendes Protein isoliert und bindet Smad1, Smad5 und Smad2 in Liganden-abhängiger Weise (in BMP und Aktivin-Pathways) (34). SIP1 ist ein neues Mitglied der Familie der zweihändigen Zinkfinger/Homöodomänen-Transkriptionsfaktoren, die auch das Vertebraten  $\delta$ EF1 und das Drosophila Zfh-1 einschließt (4, 5). Wie diese enthält SIP1 zwei weit voneinander getrennte Zinkfingercluster. Ein Cluster der 4 Zinkfinger (3 CHH und 1 CCHC-Finger) liegt im N-terminalen Bereich des Proteins und ein weiterer Cluster der 3 CHH-Zinkfinger liegt im C-terminalen Bereich vor (**Fig. 1A**). Zwischen SIP1 und  $\delta$ EF1 ist ein hoher Grad an Sequenzidentität innerhalb des N-terminalen Zinkfingerclusters (87%) und innerhalb des C-terminalen Zinkfingerclusters (97%) offensichtlich (siehe **Fig. 1B**), wohingegen die zwei Proteine in den Bereichen außerhalb der Zinkfingercluster weniger konserviert sind (34). Daher wird angenommen, dass SIP1 und  $\delta$ EF1 an sehr ähnliche Sequenzen binden. Zusätzlich binden die N-terminalen und C-terminalen Zinkfingercluster von  $\delta$ EF1 an sehr ähnliche Sequenzen, die die CACCT-Kernkonsensussequenz enthalten (10). Innerhalb des N-terminalen Clusters sind sowohl  $\delta$ EF1<sub>NZF3</sub> als auch  $\delta$ EF1<sub>NZF4</sub> die Hauptdeterminanten für die Bindung an die CACCT-Konsensussequenz, und  $\delta$ EF1<sub>CZF2</sub> und  $\delta$ EF1<sub>CZF3</sub> werden für die Bindung an den C-terminalen Cluster benötigt (10). Darüber hinaus zeigt die  $\delta$ EF1<sub>NZF3+NZF4</sub>-Domäne eine hohe Homologie (67%) zur  $\delta$ EF1<sub>CZF2+CZF3</sub>-Domäne und das mag erklären, warum diese zwei Cluster an ähnliche Konsensus-Zielsequenzen auf der DNA binden (**Fig. 1C**). All diese Reste sind für die Bindung essentiell, und die, die zwischen  $\delta$ EF1<sub>NZF3+NZF4</sub> und  $\delta$ EF1<sub>CZF2+CZF3</sub> konserviert sind, sind auch zwischen SIP1<sub>NZF3+NZF4</sub> und SIP1<sub>CZF2+CZF3</sub> konserviert. Zusammengefasst implizieren diese Vergleiche, dass die N- und C-terminalen Zinkfingercluster von SIP1 auch an sehr ähnliche Zielsequenzen binden.

– Zwei CACCT-Stellen sind für die Bindung von SIP1 an den Xbra2-Promotor notwendig.

**[0058]** CACCT-Stellen sind für die Bindung von SIP1 an den Xbra2-Promotor notwendig. CACCT-Stellen sind für die Bindung von SIP1 an den Xbra2-Promotor notwendig. CACCT-Stellen sind für die Bindung von SIP1 an den Xbra2-Promotor notwendig. CACCT-Stellen sind für die Bindung von SIP1 an den Xbra2-Promotor notwendig. SIP1 bindet an den Xenopus-Xbra2-Promotor und unterdrückt die Expression der Xbra2-mRNA, wenn sie im Xenopus-Embryo überexprimiert wird (34). Der Xbra2-Promotor enthält mehrere CACCT-Sequenzen, von denen zwei in einem Bereich (-381 bis -231) lokalisiert sind, der für die Induktion durch Aktivin notwendig ist (16). Diese zwei Stellen, eine stromaufwärts gelegene CACCT-Stelle bzw. eine stromabwärts gelegene AGGTG-Stelle (d. h. 5'-CACCT auf dem anderen DNA-Strang), sind durch 24 bp getrennt. Um die Bindungsvoraussetzungen für SIP1 an diese Stellen weiter aufzuklären, wurde ein entsprechendes 50 bp-langes Oligonukleotid (Xbra-WT; für eine Liste aller Sonden siehe Tabelle 1) als Sonde in elektrophoretischen Mobilitäts-Shift-Assays (EMSAs) verwendet. Die Xbra-D-Sonde, die eine Mutation der stromabwärts gelegenen AGGTG-Stelle nach AGATG enthält, wurde ebenfalls eingeschlossen. Von einer ähnlichen Mutation war zuvor gezeigt worden, dass sie die Bindung von  $\delta$ EF1 an den  $\kappa$ E2-Enhancer beseitigt (30). Zusätzlich testeten wir die stromabwärts gelegene Stelle (Xbra-E Sonde) und die stromaufwärts gelegene Stelle (Xbra-F Sonde) unabhängig als kürzere Sonden. Diese Sonden wurden mit Gesamtextrakten von COS-Zellen, die den Myc-markierten C-terminalen Zinkfingercluster von SIP1 (SIP1<sub>CZF</sub>), den Myc-markierten N-terminalen Zinkfingercluster von SIP1 (SIP1<sub>NZF</sub>), oder Myc-markiertes SIP1 voller Länge (SIP1<sub>FS</sub>) exprimierten, inkubiert.

**[0059]** Wenn Kontroll-transfizierte COS-Zellen als Kontrolle mit der A-Sonde verwendet werden, werden zwei schwache Komplexe und ein starker Komplex sichtbar. Durch Verwendung von Kompetitor-Oligonukleotiden zeigte es sich, dass die zwei schwachen Komplexe nicht-spezifisch sind, wohingegen der starke, schnell wandernde Komplex Spezifität für die Bindung an die Xbra-Sonde aufweist. Letztere Beobachtung liegt nahe, dass die COS-Zellen ein endogenes Protein enthalten, das an die Xbra-WT-Sonde binden kann. Wenn SIP1<sub>CZF</sub> im Extrakt vorliegt, beobachteten wir einen starken und langsam wandernden Komplex zusätzlich zur endogenen Bindeaktivität des COS-Extrakts. Dieser Komplex konnte mit einem Anti-Myc-Antikörper supergeschiftet werden, was bestätigt, dass der Komplex aus der Bindung von SIP1<sub>CZF</sub> an die Xbra-WT-Sonde resultiert. Mutation der stromabwärts gelegenen Stelle (Xbra-D-Sonde) beeinflusste die Bildung dieses SIP1<sub>CZF</sub>-Komplexes stark. Darüber hinaus bindet SIP1<sub>CZF</sub> an die Xbra-E-Sonde, nicht aber an die Xbra-F-Sonde, was anzeigt, dass die stromabwärts gelegene Stelle essentiell für die Bindung von SIP1<sub>CZF</sub> ist, und dass SIP1<sub>CZF</sub> exklusiv an diese Stelle binden kann. Der starke, mit der Xbra-F-Sonde sichtbar gemachte Komplex lag auch in SIP1<sub>FS</sub>-Extrakten und in Kontroll-Extrakten vor, und geht von einem bis dato nicht charakterisierten endogenen Protein aus COS-Zellen aus, das die Xbra-F-Sonde bindet. Zusätzlich zeigten COS-Zellextrakte, die SIP1<sub>NZF</sub> enthielten, ähnliche Bindungsmuster in EMSAs, wie sie mit SIP1<sub>CZF</sub> erhalten wurden. Es ist offensichtlich, dass wie bei

δEF1 (10), beide Zinkfingercluster von SIP1 ähnliche DNA-Bindemerkmale haben.

**[0060]** Ein starker Komplex, der SIP1<sub>FS</sub> entspricht, wird auch mit der Xbra-WT-Sonde erzeugt. Es ist wichtig zu erwähnen, dass die Produktionsniveaus von SIP1<sub>CZF</sub> in COS-Zellen ungefähr 50-fach höher sind als das SIP1<sub>FS</sub>-Niveau. Für jede EMSA-Reaktion verwendeten wir immer dieselbe Menge an Rohproteinen aus COS-Zellen. Die Bindung von SIP1<sub>FS</sub> an die Xbra-WT-Sonde ist genauso stark wie die Bindung von SIP1<sub>CZF</sub>. Interessanterweise zeigt das an, dass die Affinität von SIP1<sub>FS</sub> für Xbra-WT mindestens 50 Mal höher ist als die von SIP1<sub>CZF</sub>.

**[0061]** Der SIP1<sub>FS</sub>-Komplex ist, ähnlich wie bei SIP1<sub>CZF</sub> und SIP1<sub>NZF</sub>, nicht da, wenn die mutierte Xbra-D-Sonde verwendet wird. Daher ist eine intakte stromabwärts gelegene Stelle erneut für die Bindung von SIP1<sub>FS</sub> notwendig. Im Gegensatz zu SIP1<sub>CZF</sub> und SIP1<sub>NZF</sub>, die mit ähnlichen Affinitäten an die Xbra-WT- und Xbra-E-Sonden binden, bindet SIP1<sub>FS</sub> nicht an die Xbra-E-Sonde. Wie SIP1<sub>CZF</sub> und SIP1<sub>NZF</sub> bindet SIP1<sub>FS</sub> nicht an die Xbra-F-Sonde. Wie schließen daraus, dass die stromabwärts gelegene Stelle (AGGTG) notwendig ist für SIP1<sub>FS</sub> um an den Xbra2-Promotor zu binden.

**[0062]** Diese Stelle ist jedoch nicht ausreichend, weil zusätzliche Sequenzen stromaufwärts der Xbra-E-Sonde für die Bindung von SIP1<sub>FS</sub> notwendig sind. Einer der Gründe warum SIP1<sub>FS</sub> nicht in der Lage war, an die Xbra-E-Sonde zu binden, kann einfach die Länge der Xbra-E-Sonde sein, weil sie kürzer ist als die Xbra-WT-Sonde. Um dies zu überprüfen, bereiteten wir eine Sonde zu, die eine zufällige Sequenz (Random Sequence, Rdm) stromaufwärts der Xbra-E-Sonde enthielt (Tabelle 1), um sie auf dieselbe Länge wie Xbra-WT auszuweiten. Im Gegensatz zu SIP1<sub>CZF</sub>, das effizient an die Rdm+Xbra-E-Sonde band, war SIP1<sub>FS</sub> nicht in der Lage, zu binden. Dieses Ergebnis zeigt, dass die Länge der Xbra-E-Sonde per se nicht die Ursache für das Versagen von SIP1<sub>FS</sub> ist, an diese Sonde zu binden.

**[0063]** Um zu untermauern, dass das Xbra-F-Oligonukleotid auch Sequenzen enthält, die für die Bindung von SIP1<sub>FS</sub> notwendig sind, fusionierten wir dieses Oligonukleotid als auch eine zufällige Sequenz stromaufwärts einer weiteren CACCT-Stelle, von der bekannt ist, dass sie stark von dem AREB6-Protein gebunden wird (10) (Xbra-F+AREB6 bzw. Rdm+AREB6 Sonden). SIP1<sub>CZF</sub> bindet mit gleicher Affinität sowohl Xbra-F+AREB6 und Rdm+AREB6-Sonden, was anzeigt, dass die AREB6-Sequenz auch durch SIP1<sub>CZF</sub> erkannt wird. SIP1<sub>FS</sub> bindet jedoch nur an die Xbra-F+AREB6-Sonde, nicht aber an Rdm+AREB6. Das bestätigt, dass das Xbra-F-Oligonukleotid Sequenzen enthält, die notwendig sind für die Bindung von SIP1<sub>FS</sub>. Zusätzlich ist das einzige gemeinsame Merkmal zwischen der Xbra-E- und der AREB6-Sonde die CAGGTGT-Sequenz, was nahe liegt, dass keine anderen Sequenzen als diese CAGGTGT-Sequenz in der Xbra-E-Sonde notwendig sind für die Bindung von SIP1<sub>FS</sub>. Einer der Gründe, warum SIP1<sub>FS</sub> nicht in der Lage ist an die Xbra-E-Sonde zu binden, mag sein, dass die Länge der Xbra-E-Sonde kürzer ist als die Länge der Xbra-WT-Sonde. Um diese Hypothese zu überprüfen, bereiteten wir eine Sonde zu, die eine zufällige Sequenz stromaufwärts der Xbra-E-Sonde enthielt, zu, um dieselbe Länge wie die Xbra-WT-Sonde zu erhalten. Im Gegensatz zu SIP1<sub>CZF</sub>, das effizient an diese Sonde bindet, war SIP1<sub>FS</sub> nicht in der Lage zu binden. Dieses Ergebnis zeigt ganz klar an, dass die Länge der Xbra-E-Sonde nicht der Grund war, warum SIP1<sub>FS</sub> diese Sonde nicht bindet. Um zu untermauern, dass das Xbra-F-Oligonukleotid auch Sequenzen enthält, die für die Bindung von SIP1<sub>FS</sub> notwendig sind, fusionierten wir dieses Oligonukleotid sowie eine zufällige Sequenz stromaufwärts einer weiteren CACCT-Stelle, von der bekannt ist, dass sie stark AREB6-Protein bindet (Xbra-F+AREB6 bzw. Rdm+AREB6). Wir beobachteten, dass SIP1<sub>CZF</sub> mit gleichen Affinitäten sowohl Xbra-F+AREB6 als auch Rdm+AREB6-Sonden bindet, was anzeigt, dass die AREB6-Sequenz auch von SIP1<sub>CZF</sub> erkannt wird. SIP1<sub>FS</sub> bindet jedoch nur die Xbra-F+AREB6-Sonde und nicht an die Rdm+AREB6-Sonde. Das bestätigt, dass das Xbra-F-Oligonukleotid Sequenzen enthält, die für die Bindung von SIP1<sub>FS</sub> notwendig sind. Zusätzlich ist der einzige gemeinsame Nenner zwischen der Xbra-E- und der AREB6-Sonde die AGGTG-Sequenz, was nahe liegt, dass keine anderen Sequenzen als diese AGGTG-Sequenz in der Xbra-E-Sonde notwendig sind für die Bindung von SIP1<sub>FS</sub>. Einer der Gründe, warum SIP1<sub>FS</sub> nicht in der Lage ist an die Xbra-E-Sonde zu binden, mag sein, dass die Länge der Xbra-E-Sonde kürzer ist als die Länge der Xbra-WT-Sonde. Um diese Hypothese zu überprüfen, bereiteten wir eine Sonde zu, die eine zufällige Sequenz stromaufwärts der Xbra-E-Sonde enthielt, zu, um dieselbe Länge wie die Xbra-WT-Sonde zu erhalten. Im Gegensatz zu SIP1<sub>CZF</sub>, das effizient an diese Sonde bindet, war SIP1<sub>FS</sub> nicht in der Lage zu binden. Dieses Ergebnis zeigt ganz klar an, dass die Länge der Xbra-E-Sonde nicht der Grund war, warum SIP1<sub>FS</sub> diese Sonde nicht bindet. Um zu untermauern, dass das Xbra-F-Oligonukleotid auch Sequenzen enthält, die für die Bindung von SIP1<sub>FS</sub> notwendig sind, fusionierten wir dieses Oligonukleotid sowie eine zufällige Sequenz stromaufwärts einer weiteren CACCT-Stelle, von der bekannt ist, dass sie stark AREB6-Protein bindet (Xbra-F+AREB6 bzw. Rdm+AREB6). Wir beobachteten, dass SIP1<sub>CZF</sub> mit gleichen Affinitäten sowohl Xbra-F+AREB6 als auch Rdm+AREB6-Sonden bindet, was anzeigt, dass die AREB6-Sequenz auch von SIP1<sub>CZF</sub> erkannt wird. SIP1<sub>FS</sub> bindet jedoch nur die Xbra-F+AREB6-Sonde und nicht an die

Rdm+AREB6-Sonde. Das bestätigt, dass das Xbra-F-Oligonukleotid Sequenzen enthält, die für die Bindung von SIP1<sub>FS</sub> notwendig sind. Zusätzlich ist der einzige gemeinsame Nenner zwischen der Xbra-E- und der AREB6-Sonde die AGGTG-Sequenz, was nahe liegt, dass keine anderen Sequenzen als diese AGGTG-Sequenz in der Xbra-E-Sonde notwendig sind für die Bindung von SIP1<sub>FS</sub>. Einer der Gründe, warum SIP1<sub>FS</sub> nicht in der Lage ist an die Xbra-E-Sonde zu binden, mag sein, dass die Länge der Xbra-E-Sonde kürzer ist als die Länge der Xbra-WT-Sonde. Um diese Hypothese zu überprüfen, bereiteten wir eine Sonde zu, die eine zufällige Sequenz stromaufwärts der Xbra-E-Sonde enthielt, zu, um dieselbe Länge wie die Xbra-WT-Sonde zu erhalten. Im Gegensatz zu SIP1<sub>CZF</sub>, das effizient an diese Sonde bindet (Fig. 2, Spur 6), war SIP1<sub>FS</sub> nicht in der Lage zu binden (Spur 3). Dieses Ergebnis zeigt ganz klar an, dass die Länge der Xbra-E-Sonde nicht der Grund war, warum SIP1<sub>FS</sub> diese Sonde nicht bindet. Um zu untermauern, dass das Xbra-F-Oligonukleotid auch Sequenzen enthält, die für die Bindung von SIP1<sub>FS</sub> notwendig sind, fusionierten wir dieses Oligonukleotid sowie eine zufällige Sequenz stromaufwärts einer weiteren CACCT-Stelle, von der bekannt ist, dass sie stark AREB6-Protein bindet (Xbra-F+AREB6 bzw. Rdm+AREB6). In den Spuren 4 und 5 beobachteten wir, dass SIP1<sub>CZF</sub> mit gleichen Affinitäten sowohl Xbra-F+AREB6 als auch Rdm+AREB6-Sonden bindet, was anzeigt, dass die AREB6-Sequenz auch von SIP1<sub>CZF</sub> erkannt wird. SIP1<sub>FS</sub> bindet jedoch nur die Xbra-F+AREB6-Sonde (Spur 1) und nicht an die Rdm+AREB6-Sonde. Das bestätigt, dass das Xbra-F-Oligonukleotid Sequenzen enthält, die für die Bindung von SIP1<sub>FS</sub> notwendig sind. Zusätzlich ist der einzige gemeinsame Nenner zwischen der Xbra-E- und der AREB6-Sonde die AGGTG-Sequenz, was nahe liegt, dass keine anderen Sequenzen als diese AGGTG-Sequenz in der Xbra-E-Sonde notwendig sind für die Bindung von SIP1<sub>FS</sub>. Einer der Gründe, warum SIP1<sub>FS</sub> nicht in der Lage ist an die Xbra-E-Sonde zu binden, mag sein, dass die Länge der Xbra-E-Sonde kürzer ist als die Länge der Xbra-WT-Sonde. Um diese Hypothese zu überprüfen, bereiteten wir eine Sonde zu, die eine zufällige Sequenz stromaufwärts der Xbra-E-Sonde enthielt, zu, um dieselbe Länge wie die Xbra-WT-Sonde zu erhalten. Im Gegensatz zu SIP1<sub>CZF</sub>, das effizient an diese Sonde bindet (Fig. 2, Spur 6), war SIP1<sub>FS</sub> nicht in der Lage zu binden (Spur 3). Dieses Ergebnis zeigt ganz klar an, dass die Länge der Xbra-E-Sonde nicht der Grund war, warum SIP1<sub>FS</sub> diese Sonde nicht bindet. Um zu untermauern, dass das Xbra-F-Oligonukleotid auch Sequenzen enthält, die für die Bindung von SIP1<sub>FS</sub> notwendig sind, fusionierten wir dieses Oligonukleotid sowie eine zufällige Sequenz stromaufwärts einer weiteren CACCT-Stelle, von der bekannt ist, dass sie stark AREB6-Protein bindet (Xbra-F+AREB6 bzw. Rdm+AREB6). In den Spuren 4 und 5 beobachteten wir, dass SIP1<sub>CZF</sub> mit gleichen Affinitäten sowohl Xbra-F+AREB6 als auch Rdm+AREB6-Sonden bindet, was anzeigt, dass die AREB6-Sequenz auch von SIP1<sub>CZF</sub> erkannt wird. SIP1<sub>FS</sub> bindet jedoch nur die Xbra-F+AREB6-Sonde (Spur 1) und nicht an die Rdm+AREB6-Sonde. Das bestätigt, dass das Xbra-F-Oligonukleotid Sequenzen enthält, die für die Bindung von SIP1<sub>FS</sub> notwendig sind. Zusätzlich ist der einzige gemeinsame Nenner zwischen der Xbra-E- und der AREB6-Sonde die AGGTG-Sequenz, was nahe liegt, dass keine anderen Sequenzen als diese AGGTG-Sequenz in der Xbra-E-Sonde notwendig sind für die Bindung von SIP1<sub>FS</sub>. Einer der Gründe, warum SIP1<sub>FS</sub> nicht in der Lage ist an die Xbra-E-Sonde zu binden, mag sein, dass die Länge der Xbra-E-Sonde kürzer ist als die Länge der Xbra-WT-Sonde. Um diese Hypothese zu überprüfen, bereiteten wir eine Sonde zu, die eine zufällige Sequenz stromaufwärts der Xbra-E-Sonde enthielt, zu, um dieselbe Länge wie die Xbra-WT-Sonde zu erhalten. Im Gegensatz zu SIP1<sub>CZF</sub>, das effizient an diese Sonde bindet (Fig. 2, Spur 6), war SIP1<sub>FS</sub> nicht in der Lage zu binden (Spur 3). Dieses Ergebnis zeigt ganz klar an, dass die Länge der Xbra-E-Sonde nicht der Grund war, warum SIP1<sub>FS</sub> diese Sonde nicht bindet. Um zu untermauern, dass das Xbra-F-Oligonukleotid auch Sequenzen enthält, die für die Bindung von SIP1<sub>FS</sub> notwendig sind, fusionierten wir dieses Oligonukleotid sowie eine zufällige Sequenz stromaufwärts einer weiteren CACCT-Stelle, von der bekannt ist, dass sie stark AREB6-Protein bindet (Xbra-F+AREB6 bzw. Rdm+AREB6). In den Spuren 4 und 5 beobachteten wir, dass SIP1<sub>CZF</sub> mit gleichen Affinitäten sowohl Xbra-F+AREB6 als auch Rdm+AREB6-Sonden bindet, was anzeigt, dass die AREB6-Sequenz auch von SIP1<sub>CZF</sub> erkannt wird. SIP1<sub>FS</sub> bindet jedoch nur die Xbra-F+AREB6-Sonde (Spur 1) und nicht an die Rdm+AREB6-Sonde. Das bestätigt, dass das Xbra-F-Oligonukleotid Sequenzen enthält, die für die Bindung von SIP1<sub>FS</sub> notwendig sind. Zusätzlich ist der einzige gemeinsame Nenner zwischen der Xbra-E- und der AREB6-Sonde die AGGTG-Sequenz, was nahe liegt, dass keine anderen Sequenzen als diese AGGTG-Sequenz in der Xbra-E-Sonde notwendig sind für die Bindung von SIP1<sub>FS</sub>.

**[0064]** Um die Sequenzen innerhalb von Xbra-F zu kartieren, die, in Verbindung mit der Xbra-E-Sequenz, benötigt werden für die Bindung von SIP1<sub>FS</sub>, bereiteten wir eine Reihe an Sonden mit identischer Länge wie Xbra-WT zu, die benachbarte Dreifach-Mutationen innerhalb des Xbra-F-Teils enthielten (siehe Tabelle 1). Nur drei dieser mutierten Sonden (d. h. Xbra-L, Xbra-M und Xbra-N) beeinflussten die Bindung von SIP1<sub>FS</sub>. Tatsächlich wurde die stromaufwärts gelegene CACCT-Sequenz, die in der Xbra-F-Sonde intakt ist, in den L-, M- und N-Sonden modifiziert. Wir zeigten auch, dass SIP1<sub>FS</sub> nicht an die Xbra-S-Sonde bindet, die eine Punktmutation enthält, die das stromaufwärts gelegene CACCT nach CATCT ändert. Diese Mutation ist ähnlich zur stromabwärts gelegenen AGATG-Mutation, die innerhalb der Xbra-D-Sonde gemacht wurde. Die oben beschriebenen Ergebnisse weisen darauf hin, dass SIP1<sub>FS</sub> beide CACCT-Sequenzen im Xbra-Promotor kontak-

tiert. Um die Bedeutung dieser Stellen weiter zu untersuchen, wurde ein DNA-Methylierungsinterferenzassay ausgeführt. Die Methylierung der drei Gs der stromabwärts gelegenen AGGTG-Stelle (SIP<sub>DO</sub>) und der zwei Gs der stromaufwärts gelegenen CACCT-Stelle (SIP<sub>UP</sub>) war signifikant niedriger in der SIP<sub>FS</sub>-gebundenen gegenüber der ungebundenen Sonde, was nahe legt, dass die Methylierung dieser Gs die Bindung von SIP<sub>FS</sub> stört. Das stützt stark die These, dass diese Reste für die SIP<sub>FS</sub>-Bindung essentiell sind. Es wurde auch beobachtet, dass die Methylierung eines der zwei Gs nahe der SIP<sub>DO</sub> lokalisiert, auch die Bindung von SIP<sub>FS</sub> stört. Konsequenterweise ist daher gezeigt worden, dass SIP<sub>FS</sub> für die DNA-Bindung zwei CACCT-Sequenzen und deren Integrität benötigt.

– SIP1 und  $\delta$ EF1 benötigen zwei CACCT-Sequenzen für die Bindung an verschiedene potentielle Kandidatenstellen.

**[0065]** SIP1 und  $\delta$ EF1 haben eine sehr ähnliche Struktur mit zwei hochgradig konservierten Zinkfingerclustern und es ist wahrscheinlich, dass diese zwei Proteine DNA in ähnlicher Weise binden. Wir liegten dar, ob  $\delta$ EF1 auch an den Xbra2-Promotor durch Kontaktierung beider CACCT-Sequenzen bindet, was zuvor nicht berichtet worden war. Myc-markiertes  $\delta$ EF1 wurde in COS-Zellen exprimiert und die entsprechenden nukleären Extrakte wurden im EMSA mit WT und einer Anzahl an mutierten Xbra-Sonden getestet.  $\delta$ EF1 bindet stark an die Xbra-WT-Sonde, die beide CACCT-Stellen enthält.  $\delta$ EF1 bindet jedoch, wie SIP<sub>FS</sub>, weder an die Xbra-E-Sonde, die nur die stromabwärts gelegene CACCT-Stelle umfasst, noch an die Xbra-F-Sonde, die nur die stromaufwärts gelegene CACCT-Stelle enthält. Zusätzlich beseitigte die Punktmutation entweder der stromaufwärts gelegenen CACCT-Stelle (Xbra-S) oder der stromabwärts gelegenen CACCT-Stelle (Xbra-D) ebenfalls die Bindung von  $\delta$ EF1. Daher benötigt  $\delta$ EF1, wie SIP<sub>FS</sub>, auch die Integrität beider CACCT-Sequenzen für die Bindung an den Xbra-Promotor. Die Tatsache, dass zwei CACCT-Sequenzen für die Bindung von SIP<sub>FS</sub> sowie für  $\delta$ EF1 benötigt werden, mag der Einzelfall für den Xbra-Promotor sein. Daher war die nächste Frage, zu analysieren, ob zwei CACCT-Sequenzen auch für SIP1/ $\delta$ EF1 Bindung an anderen Zielstellen notwendig sind. Potentielle  $\delta$ EF1- und SIP1-Bindeelemente liegen in mehreren Promotoren vor. Ein potentiell  $\delta$ EF1-Bindeelement, das tatsächlich zwei intakte und voneinander getrennte CACCT-Stellen enthält, wurde innerhalb des Promotors des humanen  $\alpha$ 4-Integrin-Gens gefunden (23). Interessanterweise sind beide Stellen innerhalb von E2-Boxen enthalten. Die Mutation dieser zwei CACCT-Stellen führte zur Derepression der  $\alpha$ 4-Integrin-Genexpression in Myoblasten, was nahe liegt, dass  $\delta$ EF1 ein Repressor der  $\alpha$ 4-Integrin-Gentranskription ist (23).

**[0066]** Da diese zwei CACCT-Stellen im Promotor nahe beieinander positioniert sind (Abstand beträgt 34 bp), untersuchten wir, ob beide CACCT-Sequenzen für die Bindung von  $\delta$ EF1 benötigt werden. Zu diesem Zweck wurde eine 60 bp-lange Sonde, die mit beiden CACCT-Stellen des  $\alpha$ 4-Integrin-Promotors überlappt, synthetisiert ( $\alpha$ 4I-WT), sowie zwei mutierte Versionen, d. h. mit einer Punktmutation entweder in der stromaufwärts gelegenen ( $\alpha$ 4I-B) bzw. der stromabwärts gelegenen CACCT-Stelle ( $\alpha$ 4I-A) (siehe Tabelle 1). Diese Sonden wurden auf die Bindung in EMSAs mit COS-Zellextrakten von entweder  $\delta$ EF1 oder SIP<sub>FS</sub>-transfizierten Zellen getestet. Sowohl  $\delta$ EF1 als auch SIP<sub>FS</sub> bilden starke Komplexe mit der  $\alpha$ 4I-WT-Sonde. Der  $\delta$ EF1-Komplex wurde vollständig mit einem Anti-Myc-Antikörper supergeschifft, was seine Spezifität zeigt. Sowohl die Bindung von SIP1 als auch von  $\delta$ EF1 wird beseitigt oder stark beeinflusst durch eine Mutation entweder der stromaufwärts gelegenen oder der stromabwärts gelegenen CACCT-Stelle. Darüber hinaus zeigten Konkurrenzexperimente, dass 50 ng unmarkierter  $\alpha$ 4I-WT-Sonde ausreichen, um die Bindung von SIP1 oder  $\delta$ EF1 an die  $\alpha$ 4I-WT-Sonde zu beseitigen, wohingegen 50 ng der unmarkierten  $\alpha$ 4I-A- oder der  $\alpha$ 4I-B-Sonden nicht ausreichend waren. Wir schlussfolgern, dass SIP<sub>FS</sub> sowie  $\delta$ EF1 die Integrität von zwei CACCT-Stellen für die Bindung an den Promotor des  $\alpha$ 4-Integrin-Gens benötigen.

**[0067]** Wir entdeckten auch zwei nahe beieinander positionierte CACCT-Stellen innerhalb des Promotors des humanen E-Cadherin-Gens. Ein Oligonukleotid, das beide CACCT-Stellen dieses E-Cadherin-Promotors umfasst, wurde als eine Sonde (Ecad-WT) zusammen mit SIP<sub>FS</sub> oder  $\delta$ EF1-Extrakten in EMSAs verwendet. Sowohl SIP<sub>FS</sub> als auch  $\delta$ EF1 bilden einen Komplex mit dieser Sonde. Wenn jedoch entweder die stromaufwärts gelegene (Ecad-A-Sonde) oder die stromabwärts gelegene (Ecad-B-Sonde) CACCT-Stelle mutiert wird (siehe Tabelle 1, unterer Teil), war die Bindung von SIP<sub>FS</sub> und  $\delta$ EF1 beseitigt. Das liegt ebenfalls nahe, dass die zwei CACCT-Stellen in diesem Promotor eine Stelle hoher Affinität für die Bindung von zweihändigen Zinkfinger/Homöodomänen-Transkriptionsfaktoren darstellen.

**[0068]** Durch Alignment der Xbra-WT-, der  $\alpha$ 4I-WT- und der Ecad-WT-Sonden (siehe Tabelle 1) entdeckten wir keine offensichtliche Homologie mit Ausnahme einer CACCTG-Stelle und einer zweiten CACCT-Stelle. Unsere oben beschriebenen Ergebnisse und dieses Alignment deuten an, dass nur jene Sequenzen an der Bindung von entweder SIP<sub>FS</sub> oder  $\delta$ EF1 beteiligt sind. Wir schlussfolgern daher, dass für die Bindung an Zielpro-

motoren SIP1<sub>FS</sub> oder  $\delta$ EF1 mindestens eine CACCT-Stelle und eine CACCTG-Stelle benötigen.

– Variationen im Abstand zwischen den und Orientierung der CACCT-Stellen

**[0069]** Innerhalb der Xbra-WT-,  $\alpha$ 4I-WT- und Ecad-WT-Sonden (Tabelle 1) war der Abstand zwischen den zwei CACCT-Sequenzen 24 bp, 34 bp bzw. 44 bp. Da SIP1<sub>FS</sub> und  $\delta$ EF1 an diese Sonden effizient binden, zeigt dies auch, dass diese Proteine Abstände zwischen den zwei CACCT-Stellen im Bereich von 24 bp bis mindestens 44 bp vertragen. Um weiter zu untersuchen, ob der Abstand zwischen den zwei CACCT-Stellen ein wichtiger Parameter für die Bindung ist, erzeugten wir verschiedene Xbra-Sonden mit Deletionen zwischen diesen Stellen. Zwei mutierte Sonden (Xbra-B und Xbra-C) haben eine Deletion von 3 Adeninen, wohingegen die Sonde Xbra-U eine Deletion von 10 Nukleotiden hat (Tabelle 1). Diese Sonden wurden in einem EMSA mit Zellextrakten von COS-Zellen getestet, die entweder SIP1<sub>FS</sub> oder  $\delta$ EF1 exprimierten. Sowohl SIP1<sub>FS</sub> als auch  $\delta$ EF1 binden mit gleicher Affinität an die Xbra-WT, Xbra-B-, Xbra-C- und Xbra-U-Sonden. Wie bereits durch die für die verschiedenen Promotoren gezeigten Ergebnisse nahe geliegt wurde, zeigt dies an, dass ebenfalls innerhalb desselben Promotorelements der Abstand zwischen den zwei CACCT-Stellen kein kritischer Parameter für die Bindung dieser zwei Transkriptionsfaktoren ist.

**[0070]** Durch ausführlichen Vergleich der Xbra-WT-,  $\alpha$ 4I-WT- und Ecad-WT-Sonden beobachteten wir, dass im Fall der Xbra-WT- und  $\alpha$ 4I-WT-Sonden die Orientierung der zwei CACCT-Stellen CACCT-N-AGGTG ist, wohingegen in der Ecad-WT-Sonde die Orientierung AGGTG-N-CACCT ist. Aufgrund des nicht-palindromischen Charakters der CACCT-Stelle könnten diese zwei Stellen als im Wesentlichen unterschiedlich angenommen werden. SIP1<sub>FS</sub> und  $\delta$ EF1 binden diese unterschiedlich orientierten Stellen jedoch mit vergleichbaren Affinitäten (siehe oben). Das liegt nahe, dass SIP1<sub>FS</sub> und  $\delta$ EF1 unabhängig von der Orientierung der zwei CACCT-Stellen binden können.

**[0071]** Um die Orientierung der zwei CACCT-Stellen bezüglich der DNA-Bindekapazität von SIP1<sub>FS</sub> und  $\delta$ EF1 weiter zu untersuchen, wurden zusätzliche Sonden entworfen. Die Sonde Xbra-EE enthält eine Tandem-Wiederholung der Xbra-E-Sonde, wohingegen die Sonde Xbra-ErE eine invertierte Wiederholung derselben Xbra-E-Sequenz enthält. Zusätzlich synthetisierten wir Xbra-V, in der die stromaufwärts gelegene CACCT-Stelle (plus einem zusätzlichen Basenpaar auf jeder Seite) durch die stromabwärts gelegene AGGTG-Sequenz ersetzt wurde und umgekehrt. Schließlich wurde in der Xbra-W-Sonde nur die stromabwärts gelegene Stelle durch die stromaufwärts gelegene CACCT-Sequenz ersetzt. All diese Sonden wurden erneut in EMSAs mit Extrakten, die von COS-Zellen, die entweder SIP1<sub>FS</sub> oder  $\delta$ EF1 exprimierten, zubereitet worden waren, getestet. Wir beobachteten die stärkste Bindung von SIP1<sub>FS</sub> oder  $\delta$ EF1 an die Xbra-EE-Sonde. Daher können SIP1<sub>FS</sub> und  $\delta$ EF1 nicht an Xbra-E, die eine einzelne CACCT-Stelle enthält, binden, binden aber stark, wenn diese Sequenz dupliziert wird, was erneut die Notwendigkeit von zwei CACCT-Stellen andeutet. Zusätzlich ist offensichtlich, dass die zwei CACCT-Stellen auf demselben DNA-Fragment vorliegen müssen, und nicht auf zwei voneinander getrennten Strängen (siehe unten). SIP1 und  $\delta$ EF1 binden an Xbra-ErE, was auch nahe liegt, dass die entsprechende Orientierung der zwei CACCT-Stellen nicht kritisch für die Bindung ist. Darüber hinaus wirkte sich das Austauschen sowohl der stromaufwärts gelegenen als auch der stromabwärts gelegenen Stellen (Sonde Xbra-V) oder der Austausch nur der stromaufwärts gelegenen Stelle durch eine zweite Kopie der stromabwärts gelegenen Stelle (Sonde Xbra-W) nicht auf die Bindung von SIP1<sub>FS</sub> und  $\delta$ EF1 aus. Aus diesen Experimenten schlussfolgern wir, dass weder der Zwischenabstand zwischen den zwei CACCT-Stellen noch die entsprechende Orientierung dieser zwei Stellen kritisch für die Bindung von zueinander zinkfinger/Homöodomänen-Transkriptionsfaktoren in vitro ist.

**[0072]** Überraschenderweise können nicht alle duplizierten CACCT-Stellen diese Faktoren binden. Tatsächlich ist die Duplikation der Xbra-F-Sequenz in Kombination mit der Xbra-E-Sequenz, von der gezeigt wurde, dass sie notwendig ist für die Bindung von SIP1<sub>FS</sub> und  $\delta$ EF1, nicht empfänglich für die Bindung von SIP1<sub>FS</sub> und  $\delta$ EF1. Das liegt nahe, dass die CACCT-Stelle innerhalb des Xbra-F-Kontexts eine Stelle mit niedriger Affinität ist und dass die Sequenzen benachbart zu dieser CACCT-Stelle die Affinität optimieren können.

**[0073]** Zusätzlich bestätigt die Tatsache, dass weder der C-terminale Cluster noch der N-terminale Cluster unabhängig an die Xbra-F-Sonde binden kann, die Annahme, dass diese Stelle eine niedrige Affinität aufweist. Im Gegensatz dazu kann die CACCTG-Stelle, die in der Xbra-E-Sonde vorliegt, SIP1<sub>CZF</sub> und SIP1<sub>NZF</sub> binden, und eine Duplikation dieses Elements schafft eine Hochaffinitätsbindestelle sowohl für SIP1<sub>FS</sub> als auch  $\delta$ EF1 voller Länge. Das liegt nahe, dass die terminale G-Base in der stromabwärts gelegenen Stelle ebenfalls die Diskriminierung zwischen einer Hoch- und einer Niedrigaffinitätsbindestelle erlauben kann. Die CACCT-Stelle in Xbra-F kann jedoch nur an einen der Zinkfingercluster von SIP1<sub>FS</sub> binden, sobald der andere Cluster durch die benachbarte Hochaffinitäts-CACCTG-Stelle belegt ist (in Xbra-E). Um die Bedeutung dieses terminalen

G-Basenrests für die Bindung von SIP1<sub>FS</sub> und  $\delta$ EF1 zu bestätigen, mutagenisierten wir die stromabwärts gelegene CACCTG-Stelle nach CACCTA (Sonde Xbra-Z). Die Bindung von SIP1<sub>FS</sub> oder  $\delta$ EF1 an die Xbra-Z-Sonde war stark vermindert (im Vergleich mit der Xbra-WT-Sonde), was nahe liegt, dass dieser G-Basenrest wichtig für die Erzeugung einer Hochaffinitätsbindestelle sowohl für SIP1<sub>FS</sub> als auch  $\delta$ EF1 ist.

**[0074]** Schließlich beobachteten wir, wenn die Xbra-E- und Xbra-F-Sonden vor dem Zusatz von SIP1<sub>FS</sub> oder  $\delta$ EF1 gemischt werden, keine Bindung, was erneut anzeigt, dass beide CACCT-Stellen in der Cis-Konfiguration vorliegen müssen, d. h. auf derselben DNA.

– Die zwei Zinkfingercluster von SIP1 werden für die DNA-Bindung benötigt und müssen intakt sein

**[0075]** SIP1 und  $\delta$ EF1 binden an DNA-Elemente, die zwei CACCT-Stellen enthalten, und beide dieser Proteine enthalten zwei Zinkfingercluster, die in der Lage sind, unabhängig an CACCT-Stellen zu binden. In der nachfolgenden Arbeit wollten wir die Bedeutung eines jeden Zinkfingerclusters für die Bindung von SIP1<sub>FS</sub> an DNA bewerten. Von Mutationen, die entweder den dritten oder den vierten Zinkfinger des N-terminalen Clusters von  $\delta$ EF1<sub>NZF</sub> zerstörten, wurde gezeigt, dass sie die Bindung dieses Clusters an die DNA beseitigen. In ähnlicher Weise beseitigte die Mutagenese des zweiten oder des dritten Zinkfingers des C-terminalen Clusters ebenfalls die Bindung von  $\delta$ EF1. Diese mutierten und Wildtyp-Cluster wurden an GST fusioniert und die Fusionsproteine wurden aus Bakterien aufgereinigt. Wir zeigten, dass sowohl Wildtyp SIP1<sub>NZF</sub> als auch SIP1<sub>CZF</sub> stark an die Xbra-E-Sonde binden. Mit derselben Menge an gereinigten Fusionsproteinen aus mutiertem Cluster und GST (GST-NZF3, GST-NZF4, GST-CZF2 und GST-CZF3) konnte jedoch keine Bindung an die Xbra-E-Sonde mit einem dieser Fusionsproteine nachgewiesen werden. Tatsächlich beseitigten diese Mutationen auch die Fähigkeit jedes Clusters (SIP1<sub>NZF</sub> und SIP1<sub>CZF</sub>) unabhängig an eine CACCT-Stelle zu binden.

**[0076]** Wir führten dann ähnliche Mutationen in SIP1 voller Länge ein (NZF3-Mut, NZF4-MUT, CZF2-MUT und CZF3-MUT) und überexprimierten diese SIP1-Mutanten in COS-Zellen als Myc-markierte Proteine. Die Expression der verschiedenen Mutanten wurde durchgeführt und mittels Western-Blot-Analyse unter Verwendung von Anti-Myc-Antikörper normalisiert. Mittels EMSAs beobachteten wir, dass WT-SIP1 stark an die Xbra-WT-Sonde bindet, und dass der SIP1-Komplex nach Inkubation mit einem Anti-Myc-Antikörper supergeschifftet wird. Im Gegensatz dazu war keine der mutierten Formen von SIP1 voller Länge in der Lage, einen SIP1-ähnlichen oder einen SIP1 supergeschiffteten Komplex zu bilden. Die gleichen Beobachtungen wurden gemacht, wenn die  $\alpha$ 4-WT-Sonde als eine Sonde verwendet wurde. Zusammengefasst bedeutet das, dass SIP1 voller Länge die Bindungskapazitäten beider intakter Zinkfingercluster benötigt, um sein Ziel zu binden, das notwendigerweise zwei CACCT-Stellen enthält. Die Auswirkung dieser Mutationen auf die Repressoraktivität von SIP1 wurde in einem Transfektionsassay unter Verwendung des p3TB-Lux-Reporterplasmids getestet. Dieses Plasmid enthält 3 Kopien, von denen jede eine CACCT-Stelle hat, einer Sequenz, die sich über den –73 bis –42-Bereich des humanen Collagenase-Promotors erstreckt (de Groot und Kruijer, 1990). SIP1 band an ein Fragment, das dieses multimerisierte Element enthielt, aber weder NZF3-MUT noch CZF3-MUT war in der Lage zu binden. Die Überexpression von SIP1 in CHO-Zellen führt zu einer starken Repression der basalen Transkriptionsaktivität von p3TP-Lux. Die Repression war jedoch 6- bis 7-fach niedriger nach Überexpression von SIP1-Mutanten, deren DNA-Bindung gestört war (NZF3-MUT oder CZF3-MUT). Daher ist die Integrität beider Zinkfingercluster sowohl für die DNA-Bindung als auch optimale, d. h. Wildtyprepressoraktivität von SIP1 notwendig.

– SIP1 bindet an DNA als Monomer

**[0077]** Die Beobachtung, dass die Integrität beider SIP1-Zinkfingercluster für seine Bindung an zwei CACCT-Sequenzen benötigt wird, veranlasste uns zu testen, ob SIP1 als ein Monomer bindet, bei dem jeder Zinkfingercluster eine CACCT-Stelle bindet. Es kann jedoch auch die Hypothese aufgestellt werden, dass SIP1 seine Zielstellen als ein Dimer bindet. Das kann implizieren, dass eines der SIP1-Proteine des Dimers eine CACCT-Stelle über seinen N-terminalen Zinkfingercluster bindet, während das zweite SIP1-Molekül die DNA über seinen C-terminalen Zinkfingercluster bindet. Konsequenterweise sollten bestimmte Kombinationen an NZF- und CZF-Mutanten im Kontext SIP1 voller Länge (siehe oben) eine dimere Konfiguration erzeugen, die DNA bindet. Wie bereits in **Fig. 5A** gezeigt, konnte in keiner der getesteten Kombinationen von NZF- mit CZF-Mutationen eine Bindung an die Xbra-WT-Sonde nachgewiesen werden. Obwohl wir nicht ausschließen können, dass diese Mutationen auch die Dimerbildung beeinflussen würden, ist es hochgradig unwahrscheinlich, dass dieselbe Mutation sowohl die DNA-Bindekapazität als auch die Monomer-Monomer-Wechselwirkung beeinflusst. Darüber hinaus ist es hochgradig unwahrscheinlich, dass zwei verschiedene Mutanten, d. h. verschiedene Mutationen innerhalb eines Clusters, sich identisch verhalten würden. Daher ziehen wir in Betracht, dass SIP1 an DNA nicht als ein Dimer bindet. Die Beobachtung, dass die Integrität beider Zinkfingercluster für

die von SIP1 an zwei CACCT-Sequenzen benötigt wird, legt nahe, dass SIP1 als ein Monomer bindet, bei dem jeder Zinkfingercluster eine CACCT-Stelle bindet. Es kann jedoch die Hypothese aufgestellt werden, dass SIP1 die Zielstellen als ein Dimer bindet. Das würde implizieren, dass eines der SIP1-Moleküle des Dimers eine CACCT-Stelle über seinen N-terminalen Zinkfingercluster bindet, während das zweite SIP1-Molekül die DNA über seinen C-terminalen Zinkfingercluster bindet. Weil beide Zinkfingercluster für die Bindung notwendig sind, wäre der Zinkfingercluster, der nicht mit der DNA wechselwirkt, dann an der Dimerisierung beteiligt. Konsequenterweise sollten einige Kombinationen an NZF- und CZF-Mutanten (siehe oben) eine Dimer-Konfiguration erzeugen, die DNA bindet. Wie gezeigt, konnte in keiner der Kombinationen von NZF- und CZF-Mutationen eine Bindung an die Xbra-WT-Sonde nachgewiesen werden. Obwohl wir nicht ausschließen können, dass diese Mutationen auch die potentielle Dimerbildung beeinflussen, ist es hochgradig unwahrscheinlich, dass dieselbe Mutation sowohl die DNA-Bindekapazität als auch die Protein-Protein-Wechselwirkung beeinflusst. Darüber hinaus ist es hochgradig unwahrscheinlich, dass zwei verschiedene Mutanten, d. h. mit verschiedenen Mutationen innerhalb eines Clusters, sich gleich verhalten würden. Diese Beobachtungen zeigen an, dass SIP1 an DNA nicht als ein Dimer bindet. Die Beobachtung, dass die Integrität beider Zinkfingercluster für die von SIP1 an zwei CACCT-Sequenzen benötigt wird, legt nahe, dass SIP1 als ein Monomer bindet, bei dem jeder Zinkfingercluster eine CACCT-Stelle bindet. Es kann jedoch die Hypothese aufgestellt werden, dass SIP1 die Zielstellen als ein Dimer bindet. Das würde implizieren, dass eines der SIP1-Moleküle des Dimers eine CACCT-Stelle über seinen N-terminalen Zinkfingercluster bindet, während das zweite SIP1-Molekül die DNA über seinen C-terminalen Zinkfingercluster bindet. Weil beide Zinkfingercluster für die Bindung notwendig sind, wäre der Zinkfingercluster, der nicht mit der DNA wechselwirkt, dann an der Dimerisierung beteiligt.

**[0078]** Konsequenterweise sollten einige Kombinationen an NZF- und CZF-Mutanten (siehe oben) eine Dimer-Konfiguration erzeugen, die DNA bindet. Wie gezeigt, konnte in keiner der Kombinationen von NZF- und CZF-Mutationen eine Bindung an die Xbra-WT-Sonde nachgewiesen werden. Obwohl wir nicht ausschließen können, dass diese Mutationen auch die potentielle Dimerbildung beeinflussen, ist es hochgradig unwahrscheinlich, dass dieselbe Mutation sowohl die DNA-Bindekapazität als auch die Protein-Protein-Wechselwirkung beeinflusst. Darüber hinaus ist es hochgradig unwahrscheinlich, dass zwei verschiedene Mutanten, d. h. mit verschiedenen Mutationen innerhalb eines Clusters, sich gleich verhalten würden. Diese Beobachtungen zeigen an, dass SIP1 an DNA nicht als ein Dimer bindet. Die Beobachtung, dass die Integrität beider Zinkfingercluster für die von SIP1 an zwei CACCT-Sequenzen benötigt wird, legt nahe, dass SIP1 als ein Monomer bindet, bei dem jeder Zinkfingercluster eine CACCT-Stelle bindet. Es kann jedoch die Hypothese aufgestellt werden, dass SIP1 die Zielstellen als ein Dimer bindet. Das würde implizieren, dass eines der SIP1-Moleküle des Dimers eine CACCT-Stelle über seinen N-terminalen Zinkfingercluster bindet, während das zweite SIP1-Molekül die DNA über seinen C-terminalen Zinkfingercluster bindet. Weil beide Zinkfingercluster für die Bindung notwendig sind, wäre der Zinkfingercluster, der nicht mit der DNA wechselwirkt, dann an der Dimerisierung beteiligt. Konsequenterweise sollten einige Kombinationen an NZF- und CZF-Mutanten (siehe oben) eine Dimer-Konfiguration erzeugen, die DNA bindet. Wie in **Fig. 5A** gezeigt, konnte in keiner der Kombinationen von NZF- und CZF-Mutationen eine Bindung an die Xbra-WT-Sonde nachgewiesen werden. Obwohl wir nicht ausschließen können, dass diese Mutationen auch die potentielle Dimerbildung beeinflussen, ist es hochgradig unwahrscheinlich, dass dieselbe Mutation sowohl die DNA-Bindekapazität als auch die Protein-Protein-Wechselwirkung beeinflusst. Darüber hinaus ist es hochgradig unwahrscheinlich, dass zwei verschiedene Mutanten, d. h. mit verschiedenen Mutationen innerhalb eines Clusters, sich gleich verhalten würden. Diese Beobachtungen zeigen an, dass SIP1 an DNA nicht als ein Dimer bindet. Die Beobachtung, dass die Integrität beider Zinkfingercluster für die von SIP1 an zwei CACCT-Sequenzen benötigt wird, legt nahe, dass SIP1 als ein Monomer bindet, bei dem jeder Zinkfingercluster eine CACCT-Stelle bindet. Es kann jedoch die Hypothese aufgestellt werden, dass SIP1 die Zielstellen als ein Dimer bindet. Das würde implizieren, dass eines der SIP1-Moleküle des Dimers eine CACCT-Stelle über seinen N-terminalen Zinkfingercluster bindet, während das zweite SIP1-Molekül die DNA über seinen C-terminalen Zinkfingercluster bindet. Weil beide Zinkfingercluster für die Bindung notwendig sind, wäre der Zinkfingercluster, der nicht mit der DNA wechselwirkt, dann an der Dimerisierung beteiligt. Konsequenterweise sollten einige Kombinationen an NZF- und CZF-Mutanten (siehe oben) eine Dimer-Konfiguration erzeugen, die DNA bindet. Wie gezeigt, konnte in keiner der Kombinationen von NZF- und CZF-Mutationen eine Bindung an die Xbra-WT-Sonde nachgewiesen werden. Obwohl wir nicht ausschließen können, dass diese Mutationen auch die potentielle Dimerbildung beeinflussen, ist es hochgradig unwahrscheinlich, dass dieselbe Mutation sowohl die DNA-Bindekapazität als auch die Protein-Protein-Wechselwirkung beeinflusst. Darüber hinaus ist es hochgradig unwahrscheinlich, dass zwei verschiedene Mutanten, d. h. mit verschiedenen Mutationen innerhalb eines Clusters, sich gleich verhalten würden. Diese Beobachtungen zeigen an, dass SIP1 an DNA nicht als ein Dimer bindet. Die Beobachtung, dass die Integrität beider Zinkfingercluster für die von SIP1 an zwei CACCT-Sequenzen benötigt wird, legt nahe, dass SIP1 als ein Monomer bindet, bei dem jeder Zinkfingercluster eine CACCT-Stelle bindet. Es kann jedoch die Hypothese aufgestellt werden, dass SIP1 die Zielstellen als ein Dimer bindet. Das würde implizieren, dass eines der



SIP1-Moleküle des Dimers eine CACCT-Stelle über seinen N-terminalen Zinkfingercluster bindet, während das zweite SIP1-Molekül die DNA über seinen C-terminalen Zinkfingercluster bindet. Weil beide Zinkfingercluster für die Bindung notwendig sind, wäre der Zinkfingercluster, der nicht mit der DNA wechselwirkt, dann an der Dimerisierung beteiligt. Konsequenterweise sollten einige Kombinationen an NZF- und CZF-Mutanten (siehe oben) eine Dimer-Konfiguration erzeugen, die DNA bindet. Wie in **Fig. 5A** gezeigt, konnte in keiner der Kombinationen von NZF- und CZF-Mutationen eine Bindung an die Xbra-WT-Sonde nachgewiesen werden. Obwohl wir nicht ausschließen können, dass diese Mutationen auch die potentielle Dimerbildung beeinflussen, ist es hochgradig unwahrscheinlich, dass dieselbe Mutation sowohl die DNA-Bindekapazität als auch die Protein-Protein-Wechselwirkung beeinflusst. Darüber hinaus ist es hochgradig unwahrscheinlich, dass zwei verschiedene Mutanten, d. h. mit verschiedenen Mutationen innerhalb eines Clusters, sich gleich verhalten würden. Diese Beobachtungen zeigen an, dass SIP1 an DNA nicht als ein Dimer bindet.

**[0079]** Um diese experimentell anzugehen, verwendeten wir eine Kombination verschieden markierter SIP1 in Supershift-Experimenten in EMSAs. Als erstes produzierten wir Myc-markiertes und/oder FLAG-markiertes SIP1<sub>FS</sub> separat auf vergleichbaren Niveaus in COS-Zellen, und bestätigten, dass beide Proteine an DNA mit ähnlichen Affinitäten binden. Der mit Myc-markiertem SIP1 erzeugte SIP1-Komplex hat eine leicht verringerte Wanderungsgeschwindigkeit im Vergleich zum FLAG-markierten Komplex (die Myc-Markierung ist länger als die FLAG-Markierung). Extrakte, die von COS-Zellen zubereitet worden waren, die ähnliche Mengen sowohl an Myc-markiertem als auch FLAG-markiertem SIP1 exprimierten, wurden mit der Xbra-WT-Sonde inkubiert und in EMSAs verwendet. Wir beobachteten die Bildung eines breiten SIP1-Komplexes, der eine Kombination des schnell wandernden FLAG-markierten SIP1-Komplexes mit dem langsam wandernden Myc-markierten SIP1-Komplex ist. Bei Verwendung eines Anti-FLAG-Antikörpers wird nur der untere Teil des Komplexes, der dem FLAG-markierten SIP1 entspricht, supergeschiftet, wohingegen ungefähr 50% der Radioaktivität innerhalb des Myc-markierten SIP1-Komplex verbleiben. Das zeigt an, dass letzterer SIP1-Komplex mit dem Anti-FLAG-Antikörper nicht supergeschiftet wird. Umgekehrt supershiftete die Inkubation des Extrakts mit einem Anti-Myc-Antikörper nur den unteren Teil des Komplex, der dem Myc-markierten SIP1 entspricht, wohingegen 50% der Radioaktivität innerhalb des FLAG-markierten SIP1-Komplex zurückgehalten werden. Dies deutet erneut an, dass kein FLAG-markiertes SIP1 mit einem Anti-Myc-Antikörper supergeschiftet wird. Wurden beide Antikörper verwendet beobachteten wir diesselben zwei supergeschifteten Banden, die dem Myc-markierten und dem FLAG-markierten supergeschifteten Komplex entsprachen, im oberen Teil des Gels. Wenn SIP1-Dimere gebildet würden, dann hätten wenigstens einige der Heterodimere aus Myc-markiertem SIP1 und FLAG-markiertem SIP1 assembliert werden müssen. Es ist jedoch keine andere supergeschiftete Bande, die einem potentiellen doppelten Supershift entsprechen würde, nämlich supergeschiftet sowohl durch Anti-Myc- als auch Anti-FLAG-Antikörper, nachweisbar. Daher gab dieses Experiment keine nachweisbare Dimerbildung zwischen FLAG-markiertem SIP1 und Myc-markiertem SIP1.

**[0080]** Schließlich wurde FLAG-markiertes SIP1 in einem Extrakt von COS-Zellen in Anwesenheit eines großen Überschusses an DNA-Bindestellen immunpräzipitiert. Die Ko-Immunpräzipitation von Myc-markiertem SIP1 war jedoch nicht machbar. Das umgekehrte Experiment, d. h. Immunpräzipitation mit einem Anti-Myc-Antikörper und Nachweis mit einem Anti-FLAG-Antikörper zeigte ebenfalls keine SIP1-Dimere. Zusammengekommen führte diese Beobachtung uns zu der Schlussfolgerung, dass SIP1 als ein Monomer an die Xbra-WT-Sonde bindet.

– Mutationen in entweder der stromaufwärts gelegenen oder der stromabwärts gelegenen CACCT-Stelle führen zu ektopischer Aktivität des Xbra2-Promotors in transgenen Froschembryos

**[0081]** SIP1 bindet an den Xbra2-Promotor und unterdrückt die Expression endogener Xbra2-mRNA, wenn es in Xenopus-Embryos überexprimiert wird (Verschueren et al., 1999). Um die Bedeutung der CACCT-Sequenzen bei der Regulation des Xbra2-Promotors in vivo zu untersuchen, testeten wir, ob Mutationen dieser die Xbra2-Promotoraktivität in transgenen Embryos beeinflussen würden. Xbra2-Promotorsequenzen wurden stromaufwärts des grünen Fluoreszenz-Protein(GFP)Gens fusioniert und diese Reporter-kassette wurde für die Herstellung der Transgene verwendet. Ein 2,1 kb-langes Xbra2-Promotorfragment erwies sich als ausreichend, um in derselben Domäne des Embryos (85% der Embryos, Stadium 11, n = 57) Synthese des Reporterproteins zu ergeben, wie im Vergleich zu endogener Xbra mRNA (die in der Randzone ist, marginal zone) mit Ausnahme des Organizer-Bereichs, für den ein regulatorisches Element in der hier getesteten Reporter-kassette fehlen könnte.

**[0082]** Eine einzelne Punktmutation innerhalb der stromabwärts gelegenen CACCT-Stelle im Promotor, die die SIP1-Bindung zerstörte (Xbra2-MUT1) und identisch zu XbraD ist, hatte ernsthafte Auswirkungen auf die räumliche Produktion des Reporterproteins. Alle Embryonen (n > 30) zeigten ektopische Expression in der in-



neren Ektodermschicht. Mutationen innerhalb der stromaufwärts gelegenen CACCT-Sequenz (Xbra-MUT4) beeinflusste ebenfalls die SIP1-Bindung: wir beobachteten in allen transgenen Embryos ( $n > 30$ ) dieselbe ektopische Expression wie für die Xbra2-MUT1-Mutation. Mutation der stromabwärtsgelegenen CACCTG zu CACCTA (Xbra2-MUT2) beeinflusst ebenfalls die Bindung von SIP1 an eine solche Sonde. Diese Mutation führte, wenn sie in den Xbra2-2,1 kb-Promotor eingeführt wurde, in allen getesteten transgenen Embryos ( $n > 30$ ) ebenfalls zu ektopischer Expression von GFP mRNA. Wir testeten auch eine Mutation (Xbra2-MUT3), die den ursprünglichen 24 bp Abstand zwischen den zwei CACCT-Sequenzen um 3 bp verringerte. Diese Mutation schwächte die Wechselwirkung einer solchen Sonde mit SIP1. Das zeigte sich auch in den entsprechenden transgenen Embryos ( $n = 37$ ): während 35% der Embryos dasselbe Expressionsmuster wie das Wildtyp Xbra2-2,1 kb Promotorfragment zeigten, hatten 65% entweder Flecken oder eine schwache kontinuierliche Expression in der inneren Ektodermschicht.

**[0083]** Eine schöne Korrelation zwischen der Auswirkung dieser Mutationen auf die SIP1-Bindeaffinität im EMSA und dem Phänotyp (ektopische Expression des Reportergens) und seiner Penetranz in vivo wurde daher erhalten, was die Bedeutung der SIP1-Zielstellen bei der normalen Regulation der Xbra2-Expression in der Entwicklung von *Xenopus* (Stadium 11) zeigt. Sie legt auch nahe, dass ein bis dato unbekannter SIP1-ähnlicher Repressor in *Xenopus* die Xbra2-Genexpression in vivo reguliert. Zusätzlich dazu bestätigt sie, dass SIP1-ähnliche Faktoren zwei intakte CACCT-Stellen für die Regulation von Zielpromotoren wie Xbra2 benötigen.

## 2. SIP1 induziert die Invasion durch Herunterregulierung von E-Cadherin

### Ergebnisse

– SIP1-Bindung unterdrückt die E-Cadherin-Promotoraktivität durch Bindung an zwei konservierte E-Boxen.

**[0084]** Um herauszufinden, ob die SIP1-Bindung sich auf die transkriptionelle Aktivität des humanen E-Cadherin-Promotors (–308/+41) auswirkt, ko-exprimierten wir transient SIP1 voller Länge mit E-Cadherin-Promotor getriebenen Reporterkonstrukten in den E-Cadherin-positiven Zelllinien NMe (Maus), MDCK (Hund) und MCF7/AZ (Mensch). SIP1-Expression führte zu einer 80%-igen Zunahme der humanen E-Cadherin-Promotoraktivität. Um die Bindungsspezifität von SIP1 für die zwei konservierten E-Boxen zu untersuchen, wurde eine Mutagenese der stromaufwärts gelegenen E-Box1 (–75) oder der stromabwärts gelegenen E-Box3 (–25) oder simultan beider E-Boxen durchgeführt. Wenn eine Ko-Transfektion mit SIP1 cDNA und den mutierten E-Cadherin-Promotorkonstrukten durchgeführt wurde (68), wurde durchweg eine De-Repression der humanen E-Cadherin-Promotoraktivität gezeigt. Zusätzlich wurden mutierte SIP1-Konstrukte mit dem humanen E-Cadherin-Promotor ko-transfiziert. Mutationen der N-terminalen oder C-terminalen Zinkfingercluster resultierten in einer leichten De-Repression der E-Cadherin-Promotoraktivität. Interessanterweise führte die Ko-Transfektion des humanen E-Cadherin-Promotors und einer SIP1-Doppelmutante, bei der beide Zinkfingercluster betroffen sind, zu einem beachtlichen Verlust an SIP1-vermittelter Repression der E-Cadherin-Promotoraktivität. Wir können daher schlussfolgern, dass SIP1 die E-Cadherin-Promotoraktivität durch Bindung an die 2 E-Boxen reprimiert, und dass die zwei Zinkfingercluster in der Tat für die volle Repression der E-Cadherin-Promotoraktivität gebraucht werden.

– Induzierbare Expression von SIP1 resultiert in einem Dosis-abhängigen Verlust an E-Cadherin-Protein und -mRNA.

**[0085]** Um herauszufinden, ob SIP1 die endogenen E-Cadherin-Expressionsniveaus beeinflusst, wurden E-Cadherin-positive MDCK-Tetoff-Zellen mit hoher Expression des tTA-Transaktivators stabil mit einem Plasmid transfiziert, das eine cDNA für ein Myc<sub>6</sub>-markiertes SIP1 voller Länge der Maus exprimiert, unter Kontrolle eines reagierenden tTA-Elements. Um SIP1 zu induzieren, wurden die Zellen ohne Tetrazyklin für 3 Tage kultiviert. Die Analyse der E-Cadherin- und SIP1-Expression mittels Immunfluoreszenz eines repräsentativen geklonten Transfektanden ergab induziertes SIP1 im Nukleus bei gleichzeitigem totalen Verlust des typischen wabenartigen E-Cadherin-Expressionsmusters an Zell-Zell-Kontakten. Die Western-Blot-Analyse bestätigte diese Ergebnisse. Die SIP1-Induktion trat bei Tetrazyklinkonzentrationen gleich oder niedriger als 2 g/ml auf. Bei fortschreitender Verringerung der Tetrazyklinkonzentration wurde E-Cadherin stärker reprimiert und das korrelierte invers mit der SIP1-Akkumulation. Weiterhin überprüften wir, ob Catherine, die E-Cadherin an das Aktin-Zytoskelett knüpfen, von der SIP1-Expression beeinflusst wurden. Nach einem Western-Blot erschienen weder  $\alpha$ -Catherin noch  $\beta$ -Catherin beeinflusst zu sein, und das wurde über Immunfluoreszenz bestätigt. Gleiche Mengen an Gesamt-RNA sowohl nicht-induzierter als auch induzierter Zellen wurden mittels Northern-Blots analysiert. Nach Hybridisierung mit einer E-Cadherin-spezifischen Sonde zeigten die SIP1-expri-

mierenden Zellen beinahe keine E-Cadherin-mRNA-Expression, wohingegen die nicht-induzierten Zellen (+tet) normale Mengen an E-Cadherin mRNA exprimierten. Diese Resultate bestätigten jene der Reporterassays insofern, dass die Induktion von SIP1-Expression sich auf die endogene E-Cadherin-Expression durch Herunterregulierung der mRNA auswirkt.

– SIP1-Expression in humanen Karzinomzelllinien.

**[0086]** Um die Expression von SIP1 in einer Anzahl von E-Cadherin-negativen und – positiven Zelllinien zu untersuchen, wurden Northern-Blot-Analysen durchgeführt. Um mögliche Kreuz-Hybridisierungen mit anderen Mitgliedern der  $\delta$ EF1-Familie zu vermeiden, wurden geeignete SIP1 cDNA-Fragmente aus Maus und Mensch als Sonden verwendet. Eine ganz klare starke inverse Korrelation zwischen der SIP1-Expression und der E-Cadherin-Expression wurde festgestellt. Hohe Expression von SIP1 findet sich in humanen Fibroblasten und die am meisten verbreitete Expression von SIP1 fand sich in E-Cadherin-negativen Karzinomzellen, von denen berichtet worden ist, dass sie einen methylierten E-Cadherin-Promotor haben (53). Da das Expressionsniveau von SIP1 in den beschriebenen Zelllinien gleich der Expression von Snail-mRNA-Expression in E-Cadherin-negativen Zelllinien (66) ist, untersuchten wir die Snail-Expressionsniveaus in unserer, konditional SIP1-exprimierenden Zelllinie MDCK-Tetoff-SIP1. Snail-Expression konnte nicht nach SIP1-Induktion nachgewiesen werden. Die E-Cadherin-Repression ist in unserem Zellsystem nicht Snail-bezogen.

– SIP1 verstärkt den bösartigen Phänotyp durch Verstärken des Verlusts der Zell-Zell-Adhäsion und Invasion.

**[0087]** Da E-Cadherin ein bekanntes Invasions-Suppressormolekül ist (47), gingen wir der Frage nach, ob SIP1-Induktion die Zellen auf einen invasiveren Phänotyp umschaltet. Ein Zell-Aggregationsassay von nicht-induzierten gegenüber induzierten MDCK-Tetoff-SIP1-Zellen wurde durchgeführt. Die nicht-induzierten MDCK-Tetoff-SIP1-Zellen zeigten eine signifikante Aggregation nach 30 Min., aber die SIP1-Induktion beseitigte die normale Zell-Zell-Aggregation in einem ähnlichen Ausmaß wie ein E-Cadherin-blockierender Antikörper, DECMA-1. Invasion in Collagen-Typl-Gele wurde durch SIP1 genauso effizient wie durch den DECMA-1-Antikörper induziert.

– SIP1-Expression führt zur Reduktion uni-direktionaler Zellwanderung.

**[0088]** Die Funktion von E-Cadherin bei der Zellwanderung wurde mittels einer Blockierung von E-Cadherin durch einen spezifischen Antikörper gezeigt, was zu einer Reduktion der unidirektionalen Zellwanderung führte (72). Die Wirkung der SIP1-Expression auf unterschiedliche Zellwanderung als Folge der Herunterregulierung von E-Cadherin wurde in einem Wunden-Assay in der induzierbarer MDCK-Tetoff-SIP1-exprimierenden Zelllinie getestet. Wir konnten zeigen, dass die Induktion von SIP1 in einer niedrigeren uni-direktionalen Zellwanderung resultiert. Die Herunterregulation von E-Cadherin-vermitteltem Zell-Zell-Kontakt resultierte in der Störung der uni-direktionalen Wanderung.

## Diskussion

**[0089]** Invasion und Metastasierung sind die wesentlichsten Schritte bei der Tumorprogression. Die Bösartigkeit von Karzinomzellen ist gekennzeichnet vom Verlust von sowohl der Zell-Zell-Adhäsion als auch der zellulären Differenzierung und es ist häufig berichtet worden, dass das negativ korreliert ist mit der E-Cadherin-Herunterregulierung. Der Verlust der E-Cadherin-Expression wurde der transkriptionellen De-Regulierung zugeschrieben (52, 73). Wir zeigen hier für das Zinkfingerprotein SIP1, dass es die E-Cadherin-Expression auf der transkriptionellen Stufe reprimiert durch die Bindung an zwei konservierte E-Boxen, die im minimalen E-Cadherin-Promotor vorliegen. Die spezifische Bindung von SIP1 an die zwei E-Boxen wurde mittels Mutagenese eines der beiden Zinkfingercluster von SIP1 oder der E-Box-Sequenzen im E-Cadherin-Promotor bestätigt. Tatsächlich resultierten solche Mutationen im Verlust der Repression der E-Cadherin-Promotoraktivität durch SIP1. Diese Ergebnisse passen zu dem Befund, dass vergleichbare Mutationen der E-Boxen in der Hochregulierung der E-Cadherin-Promotoraktivität in E-Cadherin-negativen Zelllinien resultierten, wo der Wildtyp-Promotor eine niedrige Aktivität zeigt (56,58). Stabile Transfektion des transkriptionellen Repressors SIP1 induziert die Herunterregulierung von E-Cadherin sowohl auf mRNA- als auch auf Proteinniveau. Der Wunden-Assay zeigt, dass SIP1 die uni-direktionale Wanderung, die durch einen funktionalen E-Cadherin-vermittelten Zell-Zell-Kontakt vermittelt wird, stört. Schwächere Zell-Zell-Kontakte resultieren in einer verstärkt multi-direktionalen Wanderung der Epithelzellen. Eine auffallende Korrelation zwischen herunterregulierter E-Cadherin- und hochregulierter SIP1-Expression wurde in verschiedenen humanen Tumorzellen beobachtet. Schließlich zeigen wir hier, dass die Herunterregulierung von E-Cadherin aufgrund der Expression von SIP1 auch mit einer bemerkenswerten Zunahme der Invasionskapazität zusammenhängt. Daher kann SIP1, auf-

grund seiner Bindung an den E-Cadherin-Promotor, als Invasionsinduktor betrachtet werden. Die Tatsache, dass der transkriptionelle Repressor Snail auch spezifisch an E-Boxen bindet, das in der transkriptionellen Repression von E-Cadherin resultiert (66, 67), brachte die Frage auf, ob die E-Cadherin-Repression in unseren Studien von Snail vermittelt wird. Eine Hochregulierung der Snail-mRNA konnte in der konditional SIP1-exprimierenden MDCK-Tetoff-SIP1-Zelllinie nicht nachgewiesen werden. Diese Daten führten uns dazu, SIP1 als den Effektor der transkriptionellen E-Cadherin-Repression in unserem Zellsystem zu erachten. Diese Idee wurde durch die Tatsache erhärtet, dass Mutationen in den E-Boxen eine breite Wirkung auf die Verringerung der Repression des E-Cadherin-Promotors haben, wenn sie mit SIP1 ko-transfiziert werden. Die De-Repression der E-Cadherin-Promotoraktivität ist, wenn mit SIP1 ko-transfiziert, schon mit einer einzelnen E-Box-Mutation nachweisbar. Für die Snail-Ko-Transfektion war eine klare De-Repressionswirkung nur sichtbar, wenn mehrere E-Boxen in dem humanen E-Cadherin-Promotor mutiert wurden (66). Die hohe Expression von SIP1 in den Brustkrebszelllinien MDA-MB435S und MDA-MB231 ist bemerkenswert. Für diese Tumorzelllinien ist beschrieben worden, dass sie einen hypermethylierten E-Cadherin-Promotor tragen (53). Das sollte jedoch eine wichtige Rolle für SIP1-Repression des endogenen E-Cadherin-Promotors nicht ausschließen. Mutationen der E-Boxen reaktivieren die exogene E-Cadherin-Promotoraktivität in diesen Zelllinien stark. Tatsächlich machten kürzliche Forschungsarbeiten klar, dass viele Transkriptionsfaktoren durch Rekrutierung von Multiproteinkomplexen mit Chromatin-modifizierenden Aktivitäten an spezifische Stellen auf der DNA wirken (74). Es war bereits gezeigt worden, dass der weitere, mit Smad wechselwirkende Transkriptionsfaktor TGIF mit Histon-Deacetylase assoziiert (75). DNA-Methylierung und Chromatin-Kondensierung könnten daher synergistisch mit der Histondeacetylierung zusammenwirken, um die Gentranskription zu reprimieren (76).

## Material und Methoden

### Zellkultur und Reagenzien

**[0090]** Die MDCK-Tetoff-Zelllinie wurde von Cloneteck (Palo Alto, USA) erhalten. Diese Zelllinie ist von der Madin Darby Kaninchen-Nieren(MDCK)-TypII-Epithelzelllinie abgeleitet und exprimiert stabil den Tet-Off-Transaktivator, tTA (77). Die MCF7/AZ-Zelllinie ist ein Zelllinie, die von MCF7 abgeleitet ist, einer humanen Mammakarzinomzelllinie (78). Die NMe-Zelllinie ist ein E-Cadherin-exprimierender Subklon von NMuMG, einer Epithelzelllinie von einer normalen Mausbrustdrüse (47). MDA-MB231 ist eine humane Brustkrebszelllinie (ATCC, Manassas, VA).

### Plasmide

**[0091]** Die SIP1 cDNA-Sequenz voller Länge der Maus wurde in den eine Myc-Markierung enthaltenden pCS3-Expressionsvektor für Eukaryoten kloniert, der von pCS2 abgeleitet ist (69). Das resultierende Plasmid wurde als pCS3-SIP1FS bezeichnet. Mutagenese der Zinkfingercluster von SIP1 ist von Remacle et al. beschrieben (68). Für die Konstruktion des induzierbaren Vektors pUHD10.3SIP1 wurde ein ClaI/XbaI-Fragment von pCS3SIP1FS in den EcoRI/XbaI-geschnittenen pUHD10.3-Vektor kloniert (79). Die ClaI-Stelle des SIP1-Fragments und die EcoRI-Stelle des Vektors wurde mit glatten Enden mittels Pfu-Polymerase (Stratagene; La Jolla, CA) versehen. Die E-Cadherin-Promotorsequenz (-341/+41) wurde mittels PCR von genomischer DNA aus der humanen MCF7/AZ-Zelllinie erhalten. Die verwendeten PCR-Primer waren: 5'-ACAAAAGAACT-CAGCCAAGTG-3' und 5'-CCGCAAGCTCACAGGTGC-3'. Der GC-Melt-Kit (Clontech; Palo Alto, CA) wurde für die effiziente Amplifikation verwendet. Das PCR-Produkt wurde mit glatten Enden versehen, kinasiert und dann in den pGL3basic-Vektor (Promega; Madison, WA) kloniert, der an der SrfI-Stelle geöffnet worden war. Durch Verwendung der KpnI-HindIII-Stellen in diesem Luciferase-Reporterkonstrukt wurde der E-Cadherin-Promotor auch in den pGL3enhancer-Vektor transferiert. Die Mutagenese der E-Boxen im humanen E-Cadherin-Promotor wurde mittels des QuickChange Site-Directed Mutagenesis Kit (Stratagene) durchgeführt, wobei folgenden Primer verwendet wurden: Vorwärtsprimer E-Box1: 5'-gctgtggccggCAGATGaacctcag-3'; Rückwärtsprimer E-Box1: 5'-ctgagggttCAICTGccggccacagc-3'; Vorwärtsprimer E-Box3: 5'-gctccgggctCAICTGgctgcagc-3'; Rückwärtsprimer E-Box3: 5'-gctgcagcCAGATGagccccggagc-3'.

### Stabile Transfektion von Zellen

**[0092]** Für die stabile Transfektion der MDCK-Tetoff-Zelllinie wurde das LipofectAMINE PLUS™ (Gibco BRL, Rockville, USA)-Verfahren verwendet. Zweitausend Zellen wurden in einem 75 cm<sup>2</sup>-Falcon für 24 h kultiviert und dann mit 30 µg pUHD10.3-SIP1-Plasmid plus 3 µg pPHT-Plasmid transfiziert. Letzteres ist ein Derivat von pPNT und vermittelt Hygromycin-Resistenz (80). Stabile MDCK-Tetoff-Transfektanten, MDCK-Tetoff-SIP1, wurden für einen Zeitraum von 2 Wochen mit Hygromycin-B selektioniert (150 Einheiten/ml) (Duchefa Biochemie, Haarlem, Niederlande). Die Induktion von SIP1 wurde durch Zusatz von Tetracyclin (1 µg/µl) verhindert

(Sigma Chemicals, USA). Die Expression von SIP1 wurde durch Auswaschen des Tetrazyklins zum Zeitpunkt des Subklonierens erreicht. Stabile Klone mit verlässlichen Induktionseigenschaften wurden mittels Immunfluoreszenz unter Verwendung von Anti-Myc-Antikörpern identifiziert.

#### Promotor-Reporterassays

**[0093]** MCF7/AZ-Zellen wurden transient unter Verwendung von FuGENE 6 (Roche, Basel, Schweiz) transfiziert. NMe und MDA-MB231 wurden mit LIPOFECTAMINE (Gibco BRL; Rochville, USA)-Verfahren transfiziert und die parentale MCDK-Zelllinie wurde mit LIPOFECTAMINEPLUS (Gibco BRL; Rochville, USA) transient transfiziert. Für transiente Transfektionen wurden ungefähr 200 000 Zellen pro 10 cm<sup>2</sup>-Loch ausgesät. Nach Inkubation für 24 h wurden 600 ng jeder Plasmid-Typ-DNA transfiziert. Das Medium wurde 24 h nach der Transfektion erneuert. Die Zellen wurden nach 3 Tagen in Lyse-Lösung des Galacto-Star™-Kits (Tropix, Bedford, MA) lysiert. Die Normalisierung der Transfektion wurde durch Messung der  $\beta$ -Galactosidase, die von ko-transfizierten pUT651-Plasmid kodiert wird (Eurogentec; Seraing, Belgien), erreicht. Luciferase-Substrat wurde zu jeder Probe zugesetzt. Für den  $\beta$ -Galactosidase-Nachweis wird ein chemiluminisierendes Substrat bereitgestellt (Tropix, Bedford, MA). Luciferase- und  $\beta$ -Galactosidase-Aktivität wurde in einem Topcount Microplate Scintillation-Lesegerät (Packard Instrument, Co., Meriden, CT) ermittelt.

#### Northern-Analyse

**[0094]** Gesamt-RNA wurde mit dem RNeasy-Kit (Qiagen; Chatsworth, CA) nach den Angaben des Herstellers isoliert. Gesamt-RNA (25  $\mu$ g) wurde glyoxyliert, auf einem 1%igen Agarosegel der Größe nach fraktioniert und auf eine Hybond-N<sup>+</sup>Membran (Amersham Pharmacia Biotech, Rainham, UK) transferiert. Hybridisierungen wurden wie vorher beschrieben durchgeführt (81).

**[0095]** Die SIP1-Sonde der Maus (459 bp) wurde durch EcoRI-Verdau der SIP1-cDNA der Maus erzeugt. Die humane SIP1-Sonde (707 bp) wurde durch einen BstEII-NotI-Verdau des Kiaa 0569-Klons (Kazusa DNA Forschungsinstitut) erzeugt. Die E-Cadherin-Sonde der Maus wurde als ein SacI-Fragment (500 bp) der E-Cadherin-cDNA der Maus verwendet. Zwei degenerierte Primer 5'-CTTCCAGCAGCCCTACGAYCARGCNCA-3'; 5'-GGGTGTGGGACCGGATRTGCATYTTNAT-3' wurden verwendet, um ein Fragment der Snail-cDNA des Hundes aus einer Gesamt-cDNA-Population der MDCK-Zelllinie zu isolieren. Klonierung und Sequenzierung der amplifizierten Bande ergab ein 432 bp cDNA-Fragment. Um die Menge an geladener RNA zu kontrollieren, wurde eine GAPDH-Sonde auf demselben Blot verwendet. Die Quantifizierung der radioaktiven Banden wurde mittels eines PhosphorImager 425 (BioRad, Richmond, CA) durchgeführt.

#### Immunfluoreszenz-Assays und Antikörper.

**[0096]** Die Zellen von Interesse wurden auf Deckgläschen kultiviert. Die Fixierung erfolgte über Standardverfahren (82). Die folgenden Antikörper wurden verwendet: der monoklonale Ratten-Antikörper DECMA-1 (Sigma; Irvine, UK), der sowohl E-Cadherin der Maus als auch des Hundes erkennt, und der Maus-Anti-Myc-Antikörper (Oncogene, Cambridge, MA). Die verwendeten Sekundäntikörper waren Alexa 488-gekoppeltes Anti-Ratten Ig und Alexa 594-gekoppeltes Anti-Maus Ig.

#### Zell-Aggregationsassay

**[0097]** Einzelzellsuspensionen wurden in Übereinstimmung mit einer E-Cadherin bewahrenden Prozedur zubereitet (83). Die Zellen wurden in einem isotonischen Puffer, der 1,25 mM Ca<sup>2+</sup> enthält, unter gyrotorischem Schütteln (New Brunswick Scientific, New Brunswick, NJ) bei 80 Upm für 30 Min. inkubiert. Die Partikeldurchmesser wurden in einem Coulter-Partikel-Size-Counter LS200 (Coulter, Lake Placid, NY) bei Beginn (N<sub>0</sub>) und nach 30 minütiger Inkubation (N<sub>30</sub>) gemessen und über dem Prozentsatz der Volumenverteilung aufgetragen.

#### Kollagen-Invasionsassay

**[0098]** Sechs-Loch-Platten wurden pro Loch mit 1,25 ml neutralisiertem Typ I Kollagen gefüllt (Upstate Biotechnology, Lake Placid, NY). Für die Gelbildung war eine Inkubation von mindestens 1 h bei 37°C notwendig. Einzelzellsuspensionen wurden auf dem Kollagen-Gel ausgesät und die Kulturen wurden bei 37°C für 24 h inkubiert. Mittels eines Inversionsmikroskops, das von einem Computerprogramm kontrolliert wurde, wurden die eindringenden und die auf der oberflächliche verbleibenden Zellen in 12 Feldern à 0,157 mm<sup>2</sup> gezählt. Der Invasionsindex drückt den Prozentsatz an Zellen aus, der in das Gel eindringt über der Gesamtzahl an Zellen (84).

## Wunden-Assay

**[0099]** Der Wunden-Assay wurde wie zuvor beschrieben durchgeführt (85). Kurz gesagt, verwundete Mono-Layers wurden für 24 h in Serum-freiem Medium in Anwesenheit oder Abwesenheit von Tetrazyklin kultiviert. Die Zell-Wanderung wurde durch Messung der Distanz von der Wunde festgestellt. Die Wanderungsergebnisse sind ausgedrückt als Durchschnitt der Distanz von der Wunde.

## REFERENZEN

1. Arora, K., H. Dai, S. G. Kazuko, J. Jamal, O. C. MB, A. Letsou, and R. Warrior. 1995. The *Drosophila* schnurri gene acts in the Dpp/TGF beta signaling pathway and encodes a transcription factor homologous to the human MBP family. *Cell* 81: 781–90.
2. Bussemakers, M. J., L. A. Girolidi, A. van Bokhoven, and J. A. Schalkan. 1994. Transcriptional regulation of the human E-cadherin gene in human prostate cancer cell lines: characterization of the human, E-cadherin gene promoter. *Biochem Biophys Res Commun* 203: 1284–90.
3. Fan, C. M., and T. Maniatis. 1990. A DNA-binding protein containing two widely separated zinc finger motifs that recognize the same DNA sequence. *Genes Dev* 4: 29–42.
4. Fortini, M. E., Z. C. Lai, and G. M. Rubin. 1991. The *Drosophila* zfh-1 and zfh-2 genes encode novel proteins containing both zinc-finger and homeodomain motifs. *Mech Dev* 34: 113–22.
5. Funahashi, J., R. Sekido, K. Murai, Y. Kamachi, and H. Kondoh. 1993. Delta-crystallin enhancer binding protein delta EF1 is a zinc finger-homeodomain protein implicated in postgastrulation embryogenesis. *Development* 119: 433–46.
6. Grieder, N. C., D. Nellen, R. Burke, K. Basler, and M. Affolter. 1995. Schnurri is required for *Drosophila* Dpp signaling and encodes a zinc finger protein similar to the mammalian transcription factor PRDII-BF1. *Cell* 81: 791–800.
7. Henderson, L. E., T. D. Copeland, R. C. Sowder, G. W. Smythers, and S. Oroszlan. 1981. Primary structure of the low molecular weight nucleic acid-binding proteins of murine leukemia viruses. *J Biol Chem* 256: 8400–6.
8. Hendrickson, W., and R. Schleif. 1985. A dimer of AraC protein contacts three adjacent major groove regions of the aral DNA site. *Proc Natl Acad Sci USA* 82: 3129–33.
9. Holmberg, S., and P. Schjerling. 1996. Cha4p of *Saccharomyces cerevisiae* activates transcription via serine/threonine response elements. *Genetics* 144: 467–78.
10. Ikeda, K., and K. Kawakami. 1995. DNA binding through distinct domains of zinc-finger-homeodomain protein AREB6 has different effects on gene transcription. *Eur J Biochem* 233: 73–82.
11. Jiang, Y., V. C. Yu, F. Buchholz, O. C. S, S. J. Rhodes, C. Candeloro, Y. R. Xia, A. J. Lusic, and M. G. Rosenfeld. 1996. A novel family of Cys-Cys, His-Cys zinc finger transcription factors expressed in developing nervous system and pituitary gland. *J Biol Chem* 271: 10723–30.
12. Kim, J. G., and L. D. Hudson. 1992. Novel member of the zinc finger superfamily: A C2-HC finger that recognizes a glia-specific gene. *Mol Cell Biol* 12: 5632–9.
13. Kretzschmar, M., and J. Massague. 1998. SMADs: mediators and regulators of TGF-beta signaling. *Curr Opin Genet Dev* 8: 103–11.
14. Kuhnlein, R. P., G. Frommer, M. Friedrich, M. Gonzalez-Gaitan, A. Weber, J. F. Wagner-Bernholz, W. J. Gehring, H. Jackle, and R. Schuh. 1994. spalt encodes an evolutionarily conserved zinc finger protein of novel structure which provides homeotic gene function in the head and tail region of the *Drosophila* embryo. *Embo J* 13: 168–79.
15. Kurokawa, M., K. Mitani, K. Irie, T. Matsuyama, T. Takahashi, S. Chiba, Y. Yazaki, K. Matsumoto, and H. Hiral. 1998. The oncoprotein Evi-1 represses TGFbeta signalling by inhibiting Smad3. *Nature* 394: 92–6.
16. Latinkic, B. V., M. Umbhauer, K. A. Neal, W. Lerchner, J. C. Smith, and V. Cunliffe. 1997. The *Xenopus* Brachyury promoter is activated by FGF and low concentrations of activin and suppressed by high concentrations of activin and by paired-type homeodomain proteins [published erratum appears in *Genes Dev* 1998 Apr 15; 12 (8): 1240]. *Genes Dev* 11: 3265–76.
17. Lerchner W., J. E. Remacle, D. Huylebroeck, and J. C. Smith. Unpublished observations.
18. Maxam, A. M., and W. Gilbert. 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol* 65: 499–560.
19. Miller, J., A. D. McLachlan, and A. Klug. 1985. Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. *Embo J* 4: 1609–14.
20. Morishita, K., K. Suzukawa, T. Taki, J. N. Ihle, and J. Yokota. 1995. EVI-1 zinc finger protein works as a transcriptional activator via binding to a consensus sequence of GACAAGATAAGATAAN1-28 CTCATCTTC. *Oncogene* 10: 1961–7.
21. Mount, S. M., and G. M. Rubin. 1985. Complete nucleotide sequence of the *Drosophila* transposable element copia: homology between copia and retroviral proteins. *Mol Cell Biol* 5: 1630–8.

22. Nucifora, G. 1997. The EVI1 gene in myeloid leukemia. *Leukemia* 11: 2022–31.
23. Postigo, A. A., and D. C. Dean. 1997. ZEB, a vertebrate homolog of *Drosophila* Zfh-1, is a negative regulator of muscle differentiation. *Embo J* 16: 3935–43.
24. Rajavashisth, T. B., A. K. Taylor, A. Andalibi, K. L. Svenson, and A. J. Lusis. 1989. Identification of a zinc finger protein that binds to the sterol regulatory element. *Science* 245: 640–3.
25. Ray, D., R. Bosselut, J. Ghysdael, M. G. Mattei, A. Tavitian, and F. Moreau-Gachelin. 1992. Characterization of Spi-B, a transcription factor related to the putative oncoprotein Spi-1/PU.1. *Mol Cell Biol* 12: 4297–304.
26. Rosen, G. D., J. L. Barks, M. F. Iademarco, R. J. Fisher, and D. C. Dean. 1994. An intricate arrangement of binding sites for the Ets family of transcription factors regulates activity of the alpha 4 integrin gene promoter. *J Biol Chem* 269: 15652–60.
27. Rupp, R. A., L. Snider, and H. Weintraub. 1994. *Xenopus* embryos regulate the nuclear localization of XMyoD. *Genes Dev* 8: 1311–23.
28. Schwabe, J. W., and D. Rhodes. 1991. Beyond zinc fingers: steroid hormone receptors have a novel structural motif for DNA recognition. *Trends Biochem Sci* 16: 291–6.
29. Seeler, J. S., C. Muchardt, A. Suessle, and R. B. Gaynor. 1994. Transcription factor PRDII-BF1 activates human immunodeficiency virus type 1 gene expression. *J Virol* 68: 1002–9.
30. Sekido, R., K. Murai, J. Funahashi, Y. Kamachi, A. Fujisawa-Sehara, Y. Nabeshima, and H. Kondoh. 1994. The delta-crystallin enhancer-binding protein delta EF1 is a repressor of E2-box-mediated gene activation. *Mol Cell Biol* 14: 5692–700.
31. Sekido, R., K. Murai, Y. Kamachi, and H. Kondoh. 1997. Two mechanisms in the action of repressor delta EF1: binding site competition with an activator and active repression. *Genes Cells* 2: 771–83.
32. Todd, R. B., and A. Andrianopoulos. 1997. Evolution of a fungal regulatory gene family: the Zn(II)2Cys6 binuclear cluster DNA, binding motif. *Fungal Genet Biol* 21: 388–405.
33. van 't Veer, L. J., P. M. Lutz, K. J. Isselbacher, and R. Bernards. 1992. Structure and expression of major histocompatibility complex-binding protein 2, a 275-kDa zinc finger protein that binds to an enhancer of major histocompatibility complex class I genes. *Proc Natl Acad Sci USA* 89: 8971–5.
34. Verschueren, K., J. E. Remacle, C. Collart, H. Kraft, B. S. Baker, P. Tylzanowski, L. Nelles, G. Wuytens, M. T. Su, R. Bodmer, J. Smith, and D. Huylebroeck. SIP1, a novel zinc finger/homeodomain repressor, interacts with Smad proteins and binds to 5'-CACCT sequences in candidate target genes. *J. Biol. Chem* (1999).
35. Watanabe, Y., K. Kawakami, Y. Hirayama, and K. Nagano. 1993. Transcription factors positively and negatively regulating the Na,K-ATPase alpha 1 subunit gene. *J Biochem (Tokyo)* 114: 849–55.
36. Yee, K. S., and V. C. Yu. 1998. Isolation and characterization of a novel member of the neural zinc finger factor/myelin transcription factor family with transcriptional repression activity. *J Biol Chem* 273: 5366–74.
37. Brent, R. and Ptashne, M. (1985). A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. *Cell* 43, 729–736.
38. Chien, C. T., Bartel, P. L., Sternglanz, R., and Fields, S. (1991). The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest. *Proc. Natl. Acad. Sci. U.S.A.* 88, 9578–9582.
39. Durfee, T., Becherer, K., Chen, P. L., Yeh, S. H., Yang, Y., Kilburn, A. E., Lee, W. H., and Elledge, S. J. (1993). The retinoblastoma protein associates with the protein phosphatase type 1 catalytic subunit. *Genes Dev.* 7, 555–569.
40. Gyuris, J., Golemis, E., Chertkov, H., and Brent, R. (1993). Cdi1, a human G1 and S phase protein phosphatase that associates with Cdk2. *Cell* 75, 791–803.
41. Silver, P. A., Brent, R., and Ptashne, M. (1986). DNA binding is not sufficient for nuclear localisation of regulatory proteins in *Saccharomyces cerevisiae*. *Mol. Cell Biol.* 6, 4783–4766.
42. Yocum, R. R., Hanley, S., West, R. J., and Ptashne, M. (1984). Use of lacZ fusions to delimit regulatory elements of the inducible divergent GAL1-GAL10 promoter in *Saccharomyces cerevisiae*. *Mol. Cell Biol.* 4, 1985–1998.
43. de Groot, R. P. and Kruijer, W. (1990) Transcriptional activation by TGF beta 1 mediated by the dyad symmetry element (DSE) and the TPA responsive element (TRE). *Biochem. Biophys. Res. Commun.*, 168, 1074–1081.
44. Kroll, K. L. and Amaya, L. (1996) Transgenic *Xenopus* embryos from sperm nuclear transplantations reveal FGF signaling requirements during gastrulation. *Development*, 122, 3173–3183.
45. Niewkoop, P. D. and Faber, J. (1967) Normal Table of *Xenopus laevis* (Daudin). Amsterdam, North Holland.
46. Frixen, U. H. et al. E-cadherin-mediated cell-cell adhesion prevents invasiveness of human carcinoma cells. *Journal of Cell Biology* 113, 173–185 (1991).
47. Vleminckx, K., Vakaet Jr, L., Mareel, M., Fiers, W. & van Roy, F. Genetic manipulation of E-cadherin expression by epithelial tumour cells reveals an invasion suppressor role. *Cell* 66, 107–119 (1991).
48. Perl, A. K., Wilgenbus, P., Dahl, U., Semb, H. & Christofori, G. A causal role for E-cadherin in the transition from adenoma to carcinoma. *Nature (London)* 392, 190–193 (1998).
49. Potter, E., Bergwitz, C. & Brabant, G. The cadherin-catenin system: Implications for growth and differenti-

- ation of endocrine tissues. *Endocrine Reviews* 20, 207–239 (1999).
50. Becker, K. F. et al. E-cadherin gene mutations provide clues to diffuse type gastric carcinomas. *Cancer Research* 54, 3845–3852 (1994).
  51. Berx, G., Noliet, F. & van Roy, F. Dysregulation of the E-cadherin/catenin complex by irreversible mutations in human carcinomas. *Cell Adhesion and Communication* 6, 171–184 (1998).
  52. Brabant, G. et al. E-cadherin – a differentiation marker in thyroid malignancies. *Cancer Research* 53, 4987–4993 (1993).
  53. Graff, J. R. et al. E-cadherin expression is silenced by DNA hypermethylation in human breast and prostate carcinomas. *Cancer Research* 55, 5195–5199 (1995).
  54. Yoshiura, K. et al. Silencing of the E-cadherin invasion-suppressor gene by CpG methylation in human carcinomas. *Proceedings of the National Academy of Sciences of the United States of America* 92, 7416–7419 (1995).
  55. Behrens, J., Löwrick, O., Klein-Hitpaas, L. & Birchmeier, W. The E-cadherin promoter: functional analysis of a G°C-rich region and an epithelial cell-specific palindromic regulatory element. *Proceedings of the National Academy of Sciences of the United States of America* 88, 11495–11499 (1991).
  56. Giroldi, L. A. et al. Role of E boxes in the repression of E-cadherin expression. *Biochemical and Biophysical Research Communications* 241, 453–458 (1997).
  57. Hennig, G. et al. Progression of carcinoma cells is associated with alterations in chromatin structure and factor binding at the E-cadherin promoter in vivo. *Oncogene* 11, 475–484 (1995).
  58. Ji, X. D., Woodard, A. S., Rimm, D. L. & Fearon, E. R. Transcriptional defects underlie loss of E-cadherin expression in breast cancer. *Cell Growth & Differentiation* 8, 773–778 (1997).
  59. Hajra, K. M., Ji, X. D. & Fearon, E. R. Extinction of E-cadherin expression in breast cancer via a dominant repression pathway acting on proximal promoter elements. *Oncogene* 18, 7274–7279 (1999).
  60. Miettinen, P. J., Ebner, R., Lopez, A. R. & Derynck, R. TGF-beta induced transdifferentiation of mammary epithelial cells to mesenchymal cells: involvement of type I receptors. *Journal of Cell Biology* 127, 2021–2036 (1994).
  61. Shiozaki, H. et al. Effect of epidermal growth factor on cadherin-mediated adhesion in a human oesophageal cancer cell line. *British Journal of Cancer* 71, 250–258 (1995).
  62. Reichmann, E. et al. Activation of an inducible c-FosER fusion protein causes loss of epithelial polarity and triggers epithelial-fibroblastoid cell conversion. *Cell* 71, 1103–1116 (1992).
  63. Batsche, E., Muchardt, C., Behrens, J., Hurst, H. C. & Cremisi, C. RB and c-Myc activate expression of the E-cadherin gene in epithelial cells through interaction with transcription factor AP-2. *Molecular and Cellular Biology* 18, 1–12 (1998).
  64. Torban, E. & Goodyer, P. R. Effects of PAX2 expression in a human fetal kidney (HEK293) cell line. *Biochimica et Biophysica Acta – Molecular Cell Research* 1401, 53–62 (1998).
  65. Spath, G. F. & Weiss, M. C. Hepatocyte nuclear factor 4 provokes expression of epithelial marker genes, acting as a morphogen in dedifferentiated hepatoma cells. *Journal of Cell Biology* 140, 935–946 (1998).
  66. Battle, E. et al. The transcription factor Snail is a repressor of E-cadherin gene expression in epithelial tumour cells. *Nature Cell Biology* 2, 84–89 (2000).
  67. Cano, A. et al. The transcription factor Snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression. *Nature Cell Biology* 2, 75–83 (2000).
  68. Remacle, J. E. et al. New mode of DNA binding of multi-zinc finger transcription factors: deltaEF1 family members bind with two hands to two target sites. *EMBO Journal* 18, 5073–5084 (1999).
  69. Verschueren, K. et al. SIP1, a novel zinc finger/homeodomain repressor, interacts with Smad proteins and binds to 5'-CACCT sequences in candidate target genes. *Journal of Biological Chemistry* 274, 20489–20498 (1999).
  70. Derynck, R., Zhang, Y. & Feng, X. H. Smads: transcriptional activators of TGFbeta-responses. *Cell* 95, 737–740 (1998).
  71. Massague, J. TGF-beta signal transduction. *Annual Review of Biochemistry* 67, 753–791 (1998).
  72. Andre, F. et al. Integrins and E-cadherin cooperate with IGF-I to induce migration of epithelial colonic cells. *International Journal of Cancer* 83, 497–505 (1999).
  73. Hirohashi, S. Inactivation of the E-cadherin-mediated cell adhesion system in human cancers. *American Journal of Pathology* 153, 333–339 (1998).
  74. Bird, A. P. & Wolffe, A. P. Methylation-induced repression – Belts, braces, and chromatin. *Cell* 99, 451–454 (1999).
  75. Wotton, D., Lo, R. S., Lee, S. & Massague, J. A Smad transcriptional corepressor. *Cell* 97, 29–39 (1999).
  76. Cameron, E. E., Bachman, K. E., Myohanen, S., Herman, J. G. & Baylin, S. B. Synergy of demethylation and histone deacetylase inhibition in the re-expression of genes silenced in cancer. *Nature Genetics* 21, 103–107 (1999).
  77. Gossen, M. et al. Transcriptional activation by tetracyclines in mammalian cells. *Science* (Washington DC)

258, 1766–1769 (1995).

78. Bracke, M. E., Van Larebeke, N. A., Vyncke, B. M. & Mareel, M. M. Retinoic acid modulates both invasion and plasma membrane ruffling of MCF-7 human mammary carcinoma cells in vitro. *British Journal of Cancer* 63, 867–872 (1991).

79. Gossen, M. & Bujard, H. Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proceedings of the National Academy of Sciences of the United States of America* 89, 5547–5551 (1992).

80. Tybulewicz, V. L. J., Crawford, C. E., Jackson, P. K., Bronson, R. T. & Mulligan, R. C. Neonatal lethality and lymphopenia in mice with a homozygous disruption of the c-abl proto-oncogene. *Cell* 65, 1153–1163 (1991).

81. Bussemakers, M. J. G., Van de Ven, W. J. M., Debruyne, F. M. J. & Schalken, J. A. Identification of High Mobility Group Protein I(Y) as potential progression marker for prostate cancer by differential hybridization analysis. *Cancer Research* 51, 606–611 (1991).

82. van Hengel, J., Vanhoenacker, P., Staes, K. & van Roy, F. Nuclear localization of the p120<sup>ctn</sup> Armadillo-like catenin is counteracted by a nuclear export signal and by E-cadherin expression. *Proceedings of the National Academy of Sciences of the United States of America* 96, 7980–7985 (1999).

83. Bracke, M. E. et al. Insulin-like growth factor I activates the invasion suppressor function of E-cadherin in MCF-7 human mammary carcinoma cells in vitro. *British Journal of Cancer* 68, 282–289 (1993).

84. Bracke, M. E., Boterberg, T., Bruyneel, E. A. & Mareel, M. M. in *Metastasis Methods and Protocols* (eds. Brooks, S. & Schumacher, U.) In press (Humana Press, Totowa, 1999).

85. Andre, F. et al. Protein kinase C-gamma and -delta are involved in insulin-like growth factor I-induced migration of colonic epithelial cells. *Gastroenterology* 116, 64–77 (1999).

86. Ikeda, M. et al. A fusion protein library: on uninvolved method for rapid screening and characterization of DNA binding or interacting proteins. *Gene* 181, 167–171 (1996).

87. Kispert, A. and Herrmann, B. G. The Brachyury gene encodes a novel DNA binding protein. *EMBO* 12, 3211–3220 (1993).

### Patentansprüche

1. Ein Verfahren zur Identifizierung von Transkriptionsfaktoren wie Aktivatoren und/oder Repressoren, umfassend die Bereitstellung von Zellen mit einer Nukleinsäuresequenz, umfassend mindestens eine CACCT-Sequenz, vorzugsweise eine zweifache CACCT-Sequenz, als Köder für das Durchmustern einer Bibliothek, die für potentielle Transkriptionsfaktoren kodiert, und Durchführung eines Spezifitätstests, um die genannten Faktoren zu isolieren.

2. Ein Verfahren zur Identifizierung von Transkriptionsfaktoren wie Aktivatoren und/oder Repressoren, umfassend die Bereitstellung von Zellen mit einer Nukleinsäuresequenz, umfassend eine der Sequenzen CACCT-N-CACCT, CACCT-N-AGGTG, AGGTG-N-CACCT oder AGGTG-N-AGGTG als Köder, wobei N eine Zwischensequenz ist.

3. Das Verfahren nach Anspruch 1 oder 2, dadurch charakterisiert, dass der Transkriptionsfaktor separate Anhäufungen von Zinkfingern umfasst.

4. Das Verfahren nach einem der Ansprüche 1 bis 3, wobei die Sequenz von einer Promotorregion stammt.

5. Das Verfahren nach Anspruch 4, wobei die Promotorregion von Brachyury,  $\alpha$ 4-Integrin, Follistatin oder E-Cadherin ausgewählt ist.

6. Das Verfahren zur Identifizierung von Verbindungen, die in der Lage sind, mit Transkriptionsfaktoren wie in den Ansprüchen 1 bis 5 definiert zu interferieren, durch

a) zusetzen einer Probe, umfassend eine potentielle Substanz, zur Identifizierung zu einem Test-System, umfassend (i) eine Nukleotidsequenz, die eine der Sequenzen CACCT-N-CACCT, CACCT-N-AGGTG, AGGTG-N-CACCT oder AGGTG-N-AGGTG als Köder umfasst, wobei N eine Zwischensequenz ist, und (ii) ein Protein, das in der Lage ist, die Nukleotidsequenz zu binden,

b) inkubieren der Probe in dem System für eine Zeitdauer, die ausreichend ist, um eine Interaktion zwischen der Substanz, einem Derivat der Substanz, oder ihrem Gegenstück mit dem Protein zu ermöglichen,

c) vergleichen der Menge und/oder Aktivität des an die Nukleotidsequenz gebundenen Proteins vor und nach dem Zusetzen und

d) identifizieren und wahlweise isolieren und/oder reinigen der Substanz.

7. Das Verfahren nach Anspruch 6, wobei das Protein ein Smad-interagierendes Protein ist.



8. Das Verfahren nach Anspruch 7, wobei das Smad-interagierende Protein SIP1 ist.

9. Ein Test-Kit zur Durchführung des Verfahrens nach Anspruch 6, umfassend mindestens (i) eine Nukleotidsequenz, die eine der Sequenzen CACCT-N-CACCT, CACCT-N-AGGTG, AGGTG-N-CACCT oder AGGTG-N-AGGTG umfasst, wobei N eine Zwischensequenz ist und (ii) ein Protein, das in der Lage ist, an die Nukleotidsequenz zu binden.

10. Ein Test-Kit zur Durchführung des Verfahrens nach Anspruch 2, umfassend mindestens eine Nukleinsäuresequenz, die eine der Sequenzen CACCT-N-CACCT, CACCT-N-AGGTG, AGGTG-N-CACCT oder AGGTG-N-AGGTG umfasst, wobei N eine Zwischensequenz ist.

11. Eine isolierte Nukleinsäuresequenz, die eine der Sequenzen CACCT-N-CACCT, CACCT-N-AGGTG, AGGTG-N-CACCT oder AGGTG-N-AGGTG umfasst, wobei N eine Zwischensequenz mit einer Länge größer 0 Basenpaare und nicht die Sequenz TCT ist, wenn N gleich 3 ist.

Es folgen 2 Blatt Zeichnungen



Figur 2

