

Proteindesign und -engineering

◆ Proteine besitzen ein enormes Potenzial, komplexe molekulare Prozesse mit hoher Effizienz und Präzision auszuführen. Wenn wir derartige Proteine nachbauen könnten, ergäben sich atemberaubende Möglichkeiten. Seitdem Techniken zur Manipulation des genetischen Codes zur Verfügung stehen, wird die Konstruktion von Proteinen mit veränderten Eigenschaften angestrebt.

Gerichtete Evolution

◆ Die Vorgehensweisen lassen sich in evolutive und rationale Methoden unterteilen. Erstere imitieren im Ablauf die natürliche Evolution und werden als gerichtete Evolution bezeichnet. Dabei führt fehlerhafte Genamplifikation zu zufälligen Austauschen in einem Protein. Anschließend wird ein künstlicher Selektionsdruck ausgeübt, um veränderte, der Umgebung angepasste Proteine zu identifizieren. Veränderungen lassen sich somit ohne mechanistisches Vorwissen und unvoreingenommen einführen.

Bei der gerichteten Evolution ist es wichtig, eine große und ausgewogene Bibliothek an Varianten zu erstellen sowie ein geeignetes Aus-

wahlverfahren zu entwickeln. Lässt sich die gesuchte Funktion an das Überleben eines Organismus koppeln, können aus mehreren Varianten die gewünschten auf elegante Weise selektiert werden. Ist dies nicht möglich, müssen alle Varianten durchsucht und die unerwünschten aussortiert werden (Screening).

Um den Sequenzraum effizient auszunutzen, werden Methoden wie die fehlerhafte Polymerasekettenreaktion (error-prone Polymerase Chain Reaction, epPCR), DNA-Shuffling oder schrittweise Verkürzung verwendet. Der für die epPCR verwendeten Polymerase fehlt die Korrekturleseaktivität, so dass sie falsch eingebaute Nukleotide nicht korrigieren kann.¹⁾ Das DNA-Shuffling²⁾ und die abgestufte Genfragment-Verlängerung (Staggered Extension Process, StEP)³⁾ haben die gerichtete Evolution weiter voran gebracht. In beiden Methoden führt fehlerhaftes Zusammensetzen von Genfragmenten zu chimären Genen.

Um Kreuzungen auch in nicht homologen Regionen zu ermöglichen, wurden Fusionsbibliotheken aus schrittweise verkürzten Genfragmenten gebaut (Incremental Truncation for the Creation of Hybrid enzymes, ITCHY).⁴⁾

Weiterhin elementar für die gerichtete Evolution ist ein Selektions- oder Screeningsystem, das verbesser-

te Varianten identifiziert und die entsprechenden Gene isoliert. Genotyp und Phänotyp müssen gekoppelt sein – eine Eigenschaft, die Zellen von Natur aus besitzen. Ist die Funktion des exprimierten Proteins wichtig für den Metabolismus der Zelle, so kann eine verbesserte Funktion aufgrund von Antibiotikaresistenz oder Komplementation in einem auxotrophen Stamm, der nicht in der Lage ist, bestimmte essentielle Substanzen selbstständig zu synthetisieren, selektiert werden (Abbildung 1a). Das Protein kann auch an einen Reporter wie das grünfluoreszierende Protein gekoppelt werden, welches zur Evaluierung der Löslichkeit eines Proteins dient.⁵⁾

Ein anderer verbreiteter Ansatz ist das Phagendisplay, in dem das Zielprotein mit einem Phagenhüllprotein fusioniert und auf der Zelloberfläche präsentiert wird (Abbildung 1b).⁶⁾ Diese Technik wurde in Kombination mit einer proteolytischen Selektion eingesetzt, um stabile Faltungen aus einer Bibliothek kombinierter Polypeptidsegmente zu isolieren. Die Kristallstruktur eines der selektierten Proteine zeigt, dass das N-terminale Fragment des Kälteschutzproteins CspA seine strukturelle Integrität erhält, aber nun durch Domänen austausch ein Multimer bildet (Abbildung 2).⁸⁾

Ein Schwachpunkt der In-vivo-Techniken ist die Transformation, welche die Größe der Bibliotheken auf höchstens 10^{10} Varianten begrenzt. Außerdem ist die Expression der Proteinvarianten vom System abhängig, so dass mit der Entwicklung der In-vitro-Translation zellfreie Systeme für das Engineering attraktiv wurden. Daraufhin wurden In-vitro-Evolutions-Techniken entwickelt, wie das Ribosom-Display, das mRNA-Display und die In-vitro-Kompartimentierung (IVC). Das Ribosom- und das mRNA-Display koppeln ein mRNA-Molekül an das translatierte Protein. Beim Ribosom-Display wird die mRNA bis kurz vor das 3'-Ende translatiert, so dass das Ribosom nicht hinunterfällt. Auch das Protein bleibt mit dem Ribosom verbunden. Über Ligandenbindung lassen sich Proteine mit der ge-

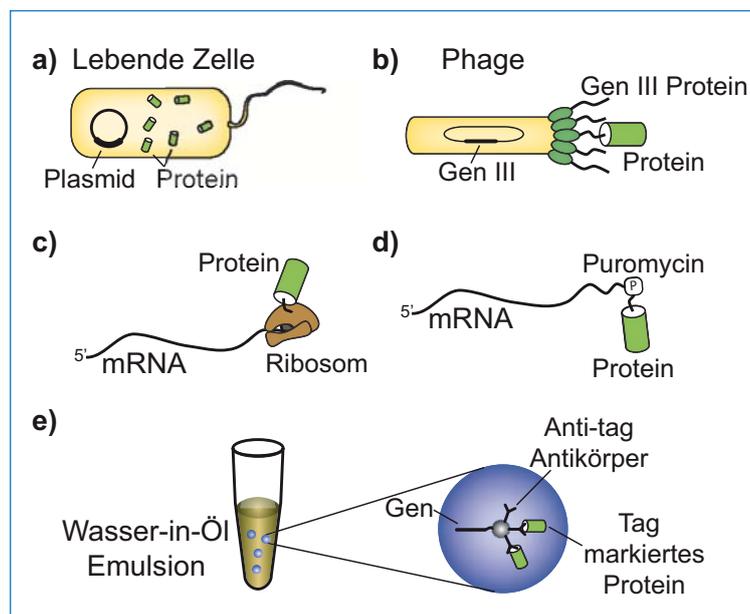


Abb. 1. Verschiedene Selektionssysteme der gerichteten Evolution: a) Bakterienzelle, b) Phagen-Display, c) Ribosom-Display, d) mRNA-Display, e) In-vitro-Kompartimentierung.

wünschten Bindungskapazität anreichern und die daran gebundenen mRNAs mit reverser Transkription und PCR vervielfältigen (Abbildung 1c).⁹⁾

Im mRNA-Display wird das Protein über das Antibiotikum Puromycin stabil mit seiner mRNA verknüpft, da es das Aminoacyl-Ende einer tRNA imitiert und durch die Peptidyl-Transferase-Aktivität des Ribosoms an das Peptid gebunden wird (Abbildung 1d).¹⁰⁾

Bei der IVC entstehen zellähnliche Kompartimente durch Wasser-in-Öl-Emulsionen. Die große Zahl von Tröpfchen pro Milliliter, ihre einfache Herstellung sowie ihre hohe Stabilität ermöglichen biochemische und genetische Ansätze in miniaturisierter und parallelisierter Form. Jedes Tröpfchen ist ein unabhängiger Minireaktor, in dem die Gene transkribiert und translatiert werden und die resultierenden RNAs und Proteine aktiv sind (Abbildung 1e). Verbesserungen in der Mikrofluidik vereinfachen IVC-Anwendungen zum Engineering von Katalyse, Bindung und Regulation.¹¹⁾

Die Effizienz der Methoden der gerichteten Evolution wird weiter durch die Entwicklung neuer Werkzeuge wie substratregulierende Enzyme und Degradations-Tags optimiert, mit denen sich die Stringenz der In-vivo-Selektion von Enzymen genauer abstimmen lässt.¹²⁾ Dennoch sind den evolutionären Ansätzen durch die kombinatorischen Möglichkeiten Grenzen gesetzt, da der Mutationsweg immer über verbesserte Varianten führen muss. Lässt sich die gewünschte Veränderung nur durch mehrere gleichzeitige Austausche erreichen, reduziert sich die Wahrscheinlichkeit, dass das Ziel mit einem zufälligen Ansatz erreicht wird.

Rationales und computergestütztes Design

◆ Im Gegensatz zum evolutiven nimmt das rationale Design auf Basis detaillierter Annahmen über Struktur und Funktion gezielt Austausche

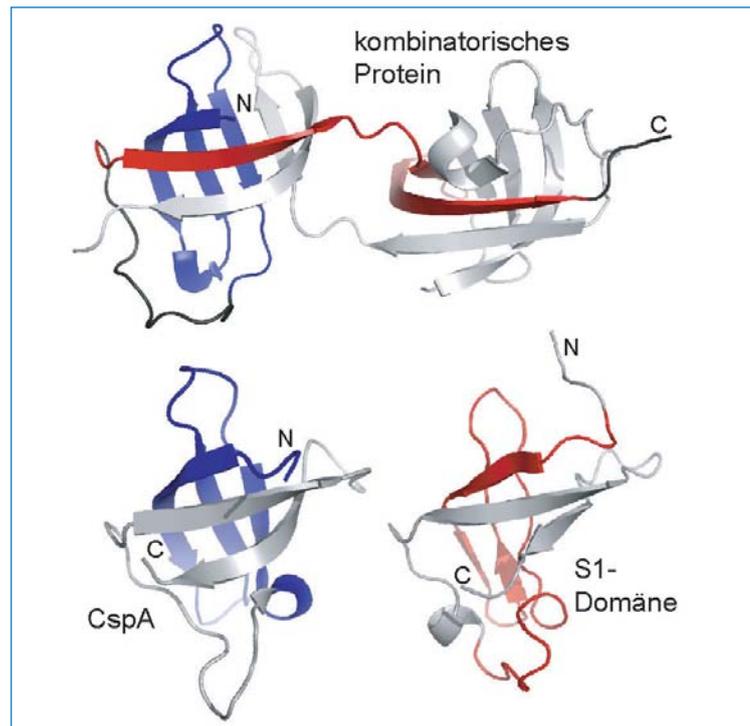


Abb. 2. Stabile Faltung eines kombinatorischen Proteins, entstanden aus Fragmenten von CspA (blau) und der S1-Domäne der Polynukleotid-Phosphorylase (rot).

Das Protein bildet ein Dimer mit einem β -Strang-Austausch. Zum Vergleich sind darunter die Strukturen der Ausgangsproteine abgebildet.

an einem Protein vor. Dies setzt einerseits ein genaues Wissen über das zu verändernde Protein voraus. Es hat aber andererseits den Vorteil, dass nur eine oder wenige Varianten getestet werden müssen. Zudem erlaubt dieses Vorgehen größere Veränderungen als die gerichtete Evolution, der die Zahl der Austausche und die Größe der Bibliothek Grenzen setzt.

Das rationale Design will die Grundsätze von Struktur-Funktions-Beziehungen gut genug verstehen, um sie bei der Konstruktion neuer Eigenschaften oder sogar neuer Proteine anzuwenden. Dazu aufgestellte Regeln und Mechanismen sollen die Faltung, Katalyse und Interaktionen genau beschreiben.

Die für das rationale Design zur Verfügung stehende Datenmenge an Proteinsequenzen und -strukturen ist in den letzten Jahren rasant gewachsen. In den letzten zehn Jahren entwickelte Algorithmen machen die Daten für rationales Proteindesign nutzbar. Um die Proteinsequenzen mit der niedrigsten Energie zu identifizieren, lassen sich Dead-End-

Elimination (DEE), Monte-Carlo-Minimierung und genetische Algorithmen nutzen. Ihre Anwendung und experimentelle Validierung machten computergestütztes Design immer wichtiger (siehe Trendbericht 2004, *Nachr. Chem.* 2005, 53, 273).

Besonders das Jahr 2003 konnte mit zwei Durchbrüchen aufwarten: dem Design einer neuartigen Proteinstruktur und ihrer kristallographischen Validierung¹³⁾ sowie der Konstruktion neuer Bindungstaschen für verschiedene niedermolekulare Liganden die zur Ent-



Birte Höcker, Jahrgang 1974, studierte Biologie in Göttingen und Ottawa. Sie promovierte 2003 an der Universität Köln bei Reinhard Sterner

über die Evolution von $(\beta\alpha)_2$ -barrel-Enzymen. Danach arbeitete sie in Homme Hellingas Labor am Duke University Medical Center in Durham (NC) an computergestütztem Protein-Design. Seit 2006 leitet sie eine Nachwuchsgruppe am Max-Planck-Institut für Entwicklungsbiologie in Tübingen.

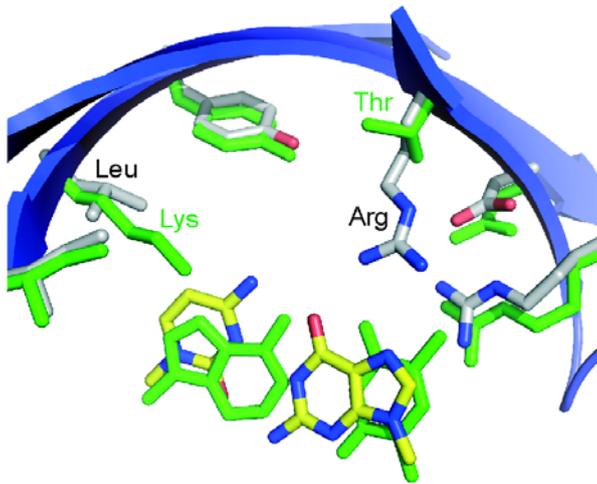


Abb. 3. Vergleich der Spezifitätstasche der wildtypischen (grün) und der neu konstruierten Homing-Endonuklease mit jeweils gebundenem Basenpaar.

wicklung einfacher Biosensoren führte.¹⁴⁾ Was hat sich seither getan?

Es gab Bestrebungen, Aktivitäten und Stabilitäten von Enzymen zu modulieren. So konnte die Thermostabilität der Cytosin-Deaminase aus Hefe verbessert werden, ohne ihre katalytische Effizienz zu beeinträchtigen.¹⁵⁾ Drei von fünf vorhergesagten Mutanten der Chorismat-Mutase besaßen eine enzymatische Aktivität, die mindestens so hoch wie die des Wildtyps war.¹⁶⁾ Eine veränderte Homing-Endonuklease I-MsoI kann eine neue DNA-Erkennungssequenz binden und schneiden (Abbildung 3).¹⁷⁾ Außerdem

wurde erfolgreich an Protein-Protein-Grenzflächen gearbeitet. So führte ein neues Wasserstoffbrücken-Netzwerk zwischen DNase und einem Immunitätsprotein zu einer hoch spezifischen Erkennung zwischen den beiden Partnern.¹⁸⁾ Das Zelladhäsions-Protein CD2 erhielt eine Calcium-Bindestelle ohne die Funktionalität des Konstruktes zu kompromittieren.¹⁹⁾ Schließlich wurden Tetrapeptide konstruiert, die Aggregate aus Serinprotease-Inhibitoren effizient depolymerisieren können. So helfen sie bei der Erforschung von Krankheiten, die durch Protein-Missfaltung hervorgerufen werden.²⁰⁾

Die meisten dieser Ansätze nutzen hochaufgelöste Strukturen als Startgerüst, dem unter Verwendung von Rotamer-Bibliotheken Konformationen der Aminosäureseitenketten aufgepfropft werden. Die daraus entstehenden Vorhersagen beruhen jedoch auf der Voraussetzung, dass sich die Gesamtstruktur nicht verändert. Tatsächlich verzerren Modelle, die kleine Strukturänderungen der Hauptkette nicht berücksichtigen, die Vorhersagen und vernachlässigen Lösungen, in denen kleine Bewegungen zu einer Stabilisierung führen. Georgiev und Donald inte-

grierten in den Dead-End-Eliminations-Algorithmus deshalb ein Kriterium, das die Backbone-Flexibilität berücksichtigt und zeigten, dass dies zu Lösungen mit deutlich niedrigeren Energien führt.²¹⁾

Auch die Strukturvorhersage, die einen starken Einfluss auf die Entwicklung von Proteindesign-Algorithmen hat, erzielte große Fortschritte. Die Struktur eines kleinen Proteins konnte so genau vorhergesagt werden, dass mit diesem Suchmodell die Kristallstruktur gelöst werden konnte.²²⁾

Diese Beispiele zeigen, dass das computergestützte Design in den letzten drei Jahren deutliche Fortschritte gemacht hat. So gab es Entwicklungen in den Algorithmen des Enzym-Designs,^{23,24)} die jedoch noch experimentell getestet werden müssen. Zudem müssen die Anwendungen robuster und für den allgemeinen Nutzer zugänglicher werden. Auch vernachlässigt die vereinfachte Annahme der Proteine als ein vorwiegend rigides Gerüst, dass die Flexibilität oft ein integraler Bestandteil von Proteinen und ihren Funktionen ist.

Kombination der Methoden

◆ Kombinierte Ansätze vereinen die Vorteile der evolutiven und rationalen Methoden. Zunehmend werden Computerprogramme eingesetzt, um aus den großen Datenmengen zu lernen und das Gelernte dann zum rationalen oder fokussierten Design von Proteinen zu nutzen. Das Proteinengineering und -Design integriert mittlerweile viele fachübergreifende Techniken aus der Proteinchemie, Gentechnik, Strukturanalyse und Bioinformatik. Im rational gelenkten Design von Bibliotheken wird die Mutagenese oder Rekombination auf ausgewählte Positionen fokussiert, wie beispielhaft am Cytochrom P450 und am blaufluoreszierenden Protein gezeigt wurde.^{25,26)}

Das Ziel ist es, evolute Methoden durch rationale Fokussierung effektiver zu gestalten. Gleichzeitig vertiefen zufällige Ansätze das Ver-

◆ Glossar

Da nicht klar definiert ist, was ein künstliches Protein ausmacht, lassen sich die Begriffe „Proteinengineering“ und „Proteindesign“ nicht eindeutig voneinander trennen und werden oft synonym gebraucht. Im engeren Sinne bezeichnet **Proteinengineering** jedoch die gezielte Veränderung natürlich vorhandener Proteine zur Herstellung nutzbarer Proteine.

Proteindesign beschreibt dagegen die Konstruktion künstlicher Proteine mit neuartigen Eigenschaften; allerdings wird dieser Begriff nicht nur für de-novo-konstruierte Proteine verwendet.

Die Vorgehensweisen beim Proteindesign und -engineering

unterteilen sich in evolute und rationale Methoden:

Gerichtete Evolution imitiert in vitro die natürliche Evolution. Dabei werden große Genbanken durch Zufallsmutagenese erstellt und Proteine mit verbesserten oder neuen Eigenschaften mittels Selektion oder Screening isoliert.

Rationales Design beschreibt die Anwendung detaillierten Wissens über Struktur und Funktion eines Proteins, um gewünschte Änderungen vorzunehmen.

Computergestütztes Design nutzt Algorithmen, um die Aminosäuresequenz mit der geringsten Energie für eine bestimmte Konformation zu ermitteln.

ständnis der Struktur-Funktions-Beziehungen von Proteinen, die wiederum die Vorhersagetechniken verbessern können.

Birte Höcker
 Max-Planck-Institut für
 Entwicklungsbiologie, Tübingen
 Birte.Hoecker@tuebingen.mpg.de

- 1) R. C. Cadwell, G. F. Joyce, *PCR Methods Appl.* 1994, 3, 136.
- 2) W. P. Stemmer, A. Cramer, K. D. Ha, T. M. Brennan, H. L. Heyneker, *Gene* 1995, 164, 49.
- 3) H. Zhao, L. Giver, Z. Shao, J. A. Affholter, F. H. Arnold, *Nat. Biotechnol.* 1998, 16, 258.
- 4) M. Ostermeier, A. E. Nixon, J. H. Shim, S. J. Benkovic, *Proc. Natl. Acad. Sci. USA* 1999, 96, 3562.
- 5) G. S. Waldo, B. M. Standish, J. Berendzen, T. C. Terwilliger, *Nat. Biotechnol.* 1999, 17, 691.
- 6) G. P. Smith, *Science* 1985, 228, 1315.
- 7) L. Riechmann, G. Winter, *Proc. Natl. Acad. Sci. USA* 2000, 97, 10068.
- 8) S. de Bono, L. Riechmann, E. Girard, R. L. Williams, G. Winter, *Proc. Natl. Acad. Sci. USA* 2005, 102, 1396.
- 9) J. Hanes, A. Plückthun, *Proc. Natl. Acad. Sci. USA* 1997, 94, 4937.
- 10) R. W. Roberts, J. W. Szostak, *Proc. Natl. Acad. Sci. USA* 1997, 94, 12297.
- 11) A. D. Griffiths, D. S. Tawfik, *Trends Biotechnol.* 2006, 24, 395.
- 12) A. C. Kleeb et al., *Proc. Natl. Acad. Sci. USA* 2007, 104, 13907.
- 13) B. Kuhlman et al., *Science* 2003, 302, 1364.
- 14) L. L. Looger, M. A. Dwyer, J. J. Smith, H. W. Hellinga, *Nature* 2003, 423, 185.
- 15) A. Korkegian, M. E. Black, D. Baker, B. L. Stoddard, *Science* 2005, 308, 857.
- 16) J. K. Lassila, J. R. Keefe, P. Oelschlaeger, S. L. Mayo, *Protein Eng. Des. Sel.* 2005, 18, 161.
- 17) J. Ashworth et al., *Nature* 2006, 441, 656.
- 18) L. A. Joachimiak, T. Kortemme, B. L. Stoddard, D. Baker, *J. Mol. Biol.* 2006, 361, 195.
- 19) W. Yang et al., *J. Am. Chem. Soc.* 2005, 127, 2085.
- 20) P. Chowdhury et al., *J. Mol. Biol.* 2007, 369, 462.
- 21) I. Georgiev, B. R. Donald, *Bioinformatics* 2007, 23, i185.
- 22) B. Qian et al., *Nature* 2007, 450, 259.
- 23) A. Zanghellini et al., *Protein Sci.* 2006, 15, 2785.
- 24) J. K. Lassila, H. K. Privett, B. D. Allen, S. L. Mayo, *Proc. Natl. Acad. Sci. USA* 2006, 103, 16710.
- 25) C. R. Otey et al., *PLoS Biol.* 2006, 4, e112.
- 26) M. A. Mena, T. P. Treynor, S. L. Mayo, P. S. Daugherty, *Nat. Biotechnol.* 2006, 24, 1569.

Stammzelltherapie

Allen Stammzellen ist die Fähigkeit gemeinsam, in einzelne oder mehrere Funktionszellen auszureifen sowie bei der Zellteilung eine identische Kopie zu bilden und sich dadurch zu erneuern (Abbildung 1). Deshalb spielen sie eine wichtige Rolle für die regenerative Medizin. Das Vermögen zur unbegrenzten Selbsterneuerung und zur Differenzierung ist in verschiedenen Zelltypen jedoch unterschiedlich stark ausgeprägt.

Entsprechend ihrer Herkunft unterscheidet man gewebespezifische oder adulte Stammzellen von embryonalen Stammzellen. Letztere gewinnt man aus undifferenzierten Zellen früher Embryonalstadien. Adulte Stammzellen entstehen während der Embryonalentwicklung der einzelnen Gewebe- und Organsysteme. Sie bleiben auch beim Erwachsenen zeitlebens erhalten, um zugrundegehende Zellen zu ersetzen.

Knochenmarkstransplantationen zur Erneuerung des Blutzellsystems sind heute klinische Routine. Beim Einsatz von Stammzellen zur Regeneration anderer Organe wie Muskel sind aber noch viele technischen

Probleme zu lösen. Es gibt jedoch Hoffnung, dass Stammzellen in zunehmenden Maße für die Therapie heute noch nicht zu behandelnder Krankheiten dienen werden. Im letzten Jahr sind die durch Reprogrammierung aus Hautzellen hergestellten pluripotenten Stammzellen (iPS) in den Mittelpunkt des Interesses gerückt. Die iPS ähneln embryonalen Stammzellen und können in reifen Zellen einer Vielzahl von Geweben differenzieren. Sie besitzen jedoch ein großes Potential, einen Tumor auszulösen. Auch wenn die Klinikreife noch weit entfernt scheint, sind iPS ein Durchbruch für die Stammzellforschung (siehe *Nachr. Chem.* 2007, 55, 995).

Embryonale Stammzellen

Bei der Verschmelzung von Ei und Samenzelle entsteht ein totipotenter Ein-Zell-Embryo, aus dem sich durch Teilung und Spezialisierung ein eigenständiges Lebewesen entwickeln kann. Nach ungefähr vier Tagen besteht der Embryo aus einem Zellverband von etwa 50 bis 100 Zellen: der Blastozyste, aus deren innerer Zellmasse man embryonale Stammzellen (ES-Zellen) gewinnen kann. ES-Zellen können

Abb. 1.
 Vergleich: Normale Stammzellen und Krebsstammzellen.

