

Das Protein-Puzzle

Der menschliche Körper besteht aus Zigtausenden Proteinen. Diese kommen in unterschiedlichen Varianten vor, zudem kann sich ihre Konzentration im Organismus mit der Zeit ändern.

Matthias Mann vom **Max-Planck-Institut für Biochemie** in Martinsried braucht deshalb schlaue Algorithmen und viel Rechenkraft für seine Forschung. Schließlich will er das menschliche Proteom, also die Gesamtheit der Proteine des Menschen, entschlüsseln und für die Medizin nutzbar machen.

TEXT **TIM SCHRÖDER**

In den Martinsrieder Labors von Matthias Mann ist es so aufgeräumt wie auf der Intensivstation eines Krankenhauses. Wände und Türen aus Glas bieten freie Sicht, in den Räumen stehen mehrere identisch aussehende Arbeitsstationen. An jeder Station hängt ein Roboterarm von der Decke, der kleine Plättchen mit Proben greift und in die Geräte steckt. Diese Apparate enthalten die technische Essenz dessen, was Matthias Mann über die Jahre entwickelt hat: Maschinen, mit denen sich Proteine aufbereiten und in atemberaubender Geschwindigkeit analysieren und bestimmen lassen.

Mann interessiert sich für diese Biomoleküle, weil sie an fast allen Stoffwechselfvorgängen eines Organismus beteiligt sind. Proteine bestehen aus einer Kette von Aminosäure-Molekülen, die sich in komplizierten Mustern dreidimensional auffaltet. Manche – wie das Keratin in Haut- und Haarzellen – dienen als Stützsubstanz. Andere Proteine, die sogenannten Enzyme, beschleunigen Stoffwechselreaktionen: Sie wandeln beispielsweise Fette in Energie um oder machen den Sauerstoff

als Energiequelle für die Zellen nutzbar. Zieht man den Wasseranteil am Körpergewicht ab, besteht ein Lebewesen zur Hälfte aus Proteinen. Kurz: Ohne Proteine gäbe es auf der Erde kein Leben.

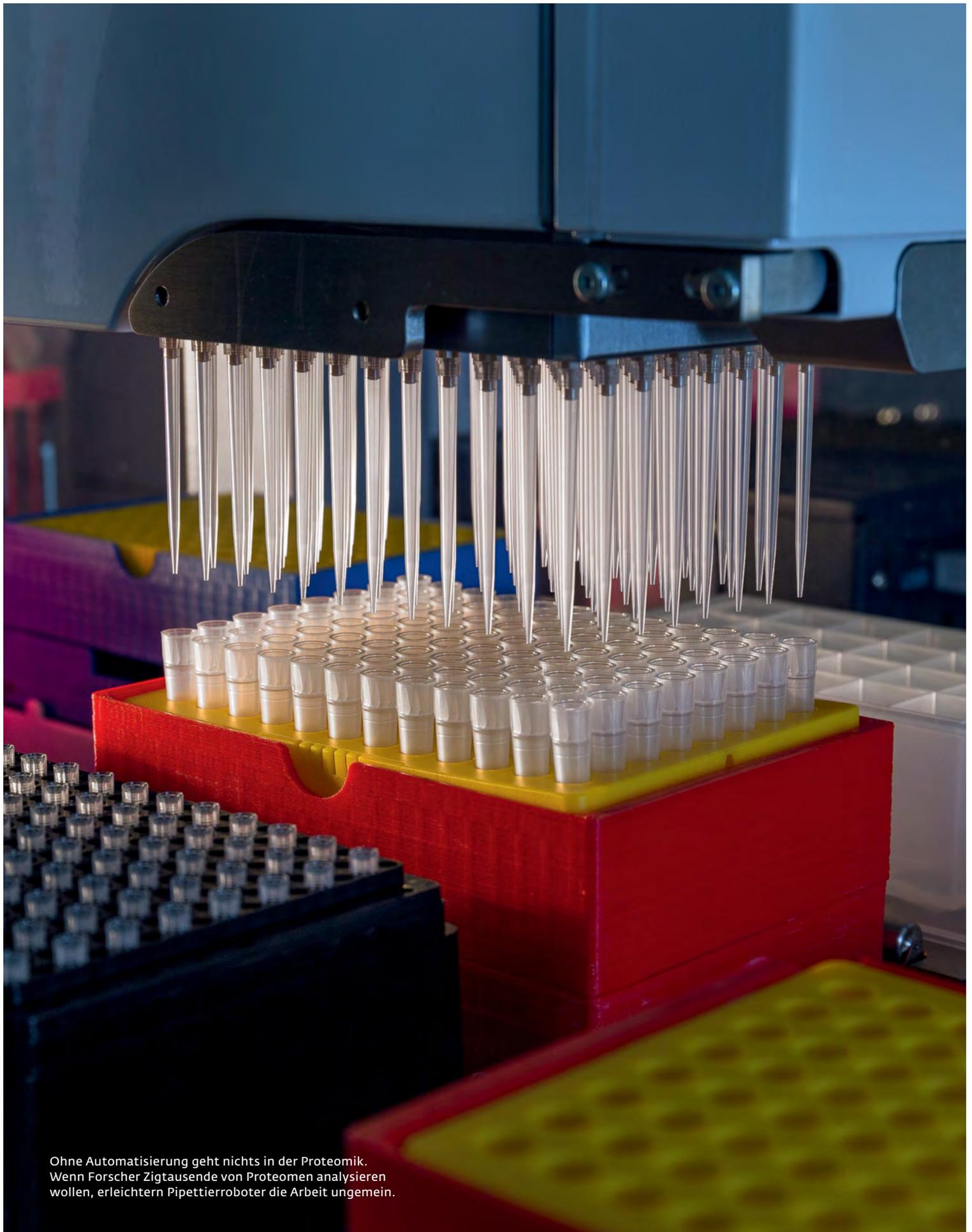
ANWENDUNGSREIFE TECHNIK

Matthias Mann hat zunächst Physik und Mathematik studiert und erforscht die Welt der Proteine seit den 1980er-Jahren. „Zwei Jahrzehnte hat es gedauert, die Technik so weiterzuentwickeln, dass man Proteine überhaupt vernünftig analysieren kann“, sagt er. „Inzwischen sind wir so weit, dass wir die Technik auch wirklich anwenden können – und jetzt wird es richtig spannend!“ Wie komplex die Analyse der Proteine ist, wurde vor allem nach der Entschlüsselung des Erbguts des Menschen im Jahr 2001 deutlich. Die Wissenschaftler des Human Genome Project identifizierten damals etwa 20.000 Gene, welche die Baupläne für etwa ebenso viele Proteine enthalten.

Erst nach und nach wurde klar, dass es eine Fülle von Varianten dieser Proteine gibt. So können nach dem Ablesen

eines Gens Teile aus der so entstandenen Boten-RNA herausgeschnitten werden, sodass die unterschiedlich langen RNA-Moleküle in jeweils unterschiedliche Proteine übersetzt werden. Viele Proteine wiederum müssen als fertige Moleküle erst noch zurechtgeschnitten werden, bevor sie einsatzbereit sind. Andere werden mit chemischen Anhängseln versehen. Alles in allem gibt es Hunderttausende von Proteinvarianten, die in einer fein abgestimmten Choreografie zusammenwirken.

Dazu kommt: Während ein Organismus zeit seines Lebens dieselben Gene besitzt, schwankt die Proteinzusammensetzung je nach Zelltyp. Manche Proteine kommen in großen Mengen und in jeder Zelle vor – andere nur in kleinen Dosen und nur in bestimmten Geweben. Je nachdem, welche Aufgaben der Stoffwechsel gerade erledigt, sind unterschiedliche Proteine aktiv. „Wenn wir wissen wollen, wie der Stoffwechsel arbeitet oder in welchem Zustand er sich befindet, dann müssen wir die Proteine in einem Gewebe mit all ihren Veränderungen analysieren können“, sagt Mann. >



Ohne Automatisierung geht nichts in der Proteomik.
Wenn Forscher Zigtausende von Proteomen analysieren
wollen, erleichtern Pipettierroboter die Arbeit ungemein.



Unter der Aufsicht von Heiner Koch, Florian Meier und Scarlet Beck (von links) analysieren allein in diesem Labor von Matthias Mann (ganz rechts) sechs Massenspektrometer Proteinproben.

Nach und nach wurde klar, wie wichtig die Proteine für die Vorgänge im Körper sind. Immer mehr Wissenschaftler interessierten sich nun für die Gesamtheit aller Proteine eines Organismus. So entstand schließlich der heute als Proteomik bezeichnete Forschungszweig. Bei der Analyse der Proteinzusammensetzung eines Lebewesens fallen jedoch riesige Datenmengen an. Auch für die Auswertung und Interpretation der Ergebnisse müssen unzählige Informationen zusammengeführt werden. Eine ausgeklügelte Datenverarbeitung ist deshalb die Grundlage für die Proteomik.

Ein großes Problem bei der Analyse von Proteinen besteht darin, dass sie so empfindlich sind. Wie jeder weiß, der schon einmal ein Ei gekocht oder geschlagen hat, bricht der dreidimensionale Aufbau von Proteinen zusammen, wenn man diese zu sehr erhitzt oder mechanisch strapaziert – sie verklumpen. Deshalb ließen sie sich zunächst nicht mit der klassischen Massenspektrometrie analysieren. Solche Massen-

spektrometer werden genutzt, um Proben mit unbekanntem Inhaltsstoffen zu untersuchen – um etwa Giftstoffe in Gewebeproben aufzuspüren.

FLUGBAHN IM ELEKTRISCHEN FELD

Bevor sich die Substanzen im Massenspektrometer untersuchen lassen, müssen sie in geladene Teilchen verwandelt werden – zum Beispiel, indem sie mit Elektronen oder anderen geladenen Teilchen beschossen und dadurch selbst in elektrisch geladene Ionen umgewandelt werden. Denn nur wenn Moleküle selbst elektrisch geladen sind, werden sie von ihrer Bahn durch das elektrische Feld des Massenspektrometers abgelenkt – je nach Ladungsstärke und Molekulargewicht unterschiedlich stark. So können die Wissenschaftler auf die Identität eines Moleküls schließen.

Für die empfindlichen Proteine aber war die klassische Massenspektrometrie mitsamt der Ionisierung zu hart. Im Labor seines Doktorvaters John B. Fenn an der Universität in Yale entwickelte

Matthias Mann zusammen mit anderen Kollegen Anfang der 1980er-Jahre eine sanftere Lösung. Dabei zerlegten die Forscher die Proteine mithilfe des Verdauungsenzyms Trypsin in etwa zehn Aminosäuren lange Bruchstücke, sogenannte Peptide. Diese versprühten sie durch ein feines Röhrchen wie ein Spray und versahen sie so mit einer elektrischen Ladung. Mit dieser Elektrospray-Ionisation konnten die Wissenschaftler erstmals empfindliche Peptide in einem Massenspektrometer analysieren – ein Verfahren, für das Fenn im Jahr 2002 den Chemie-Nobelpreis erhielt.

Nun schlug die Stunde der Informatiker, denn für einen Menschen ist es schier unmöglich, aus dem Peptidmix auf die ursprünglichen Proteine zu schließen. Zusammen mit seinem Kollegen Jürgen Cox entwickelte Mann deshalb eine Analysesoftware namens MaxQuant, die viele Tausend Peptidbruchstücke mit internationalen Datenbanken abgleichen kann. Die Datenbanken enthalten neben den Gewichtsangaben für alle möglichen Peptide



Ausschnitt aus dem Proteinspektrum einer Krebszelle: Die Proteine werden zunächst in Peptide mit unterschiedlichen Massen zerlegt (dargestellt in verschiedenen Farben), in Chromatografie-Säulen voneinander getrennt und anschließend im Massenspektrometer nachgewiesen.

auch die Information, zu welchem Protein ein Bruchstück gehört. MaxQuant gleicht die Daten aus dem Massenspektrometer mit den Datenbanken ab und rekonstruiert aus den Ergebnissen die in einer Probe enthaltenen Proteine.

CHROMATOGRAPHIE IM MINIATURFORMAT

Noch konnten Elektrospray-Ionisation und MaxQuant gemeinsam aber nicht alle Hindernisse der Proteinanalyse überwinden. Sie erfassten etwa keine Proteine, die nur in geringen Mengen in einer Probe vorhanden sind. Das lag daran, dass die Forscher im Verhältnis große Flüssigkeitsmengen benötigten, um die Proteine vor der Analyse im Massenspektrometer mittels Chromatografie von den restlichen Bestandteilen einer Probe abzutrennen. Dadurch wurden die seltenen Proteine zu stark verdünnt und waren nicht mehr nachzuweisen.

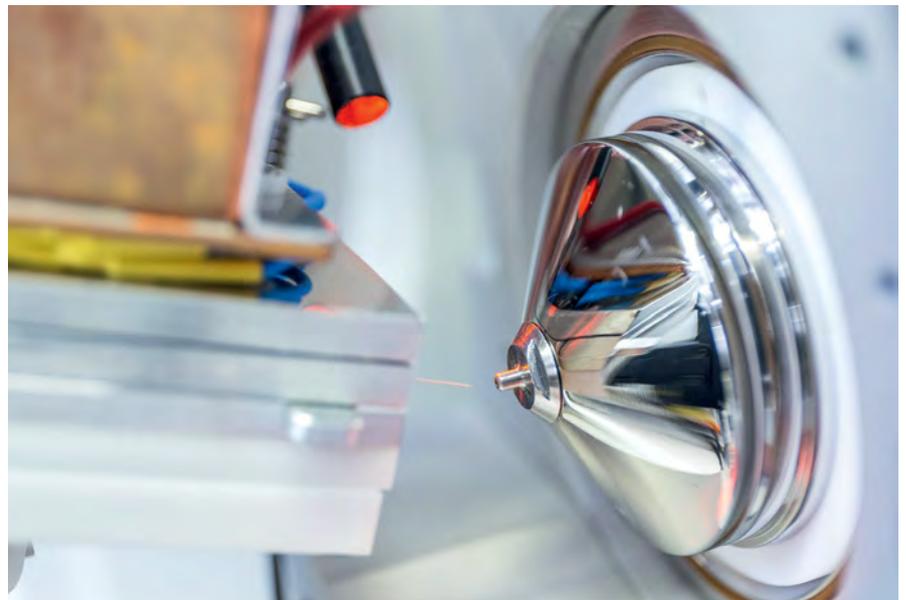
Mann miniaturisierte die bei dieser Form der Chromatografie benutzten Röhren so weit, dass sie nur noch einen Durchmesser von wenigen milliardstel

Metern aufwiesen. Für diese Röhren benötigt er lediglich winzige Probenmengen: „Mit unserer Nano-Chromatografie können wir auch die Proteine in ausreichend hoher Konzentration auffangen, die nur in ganz geringen Mengen vorliegen.“

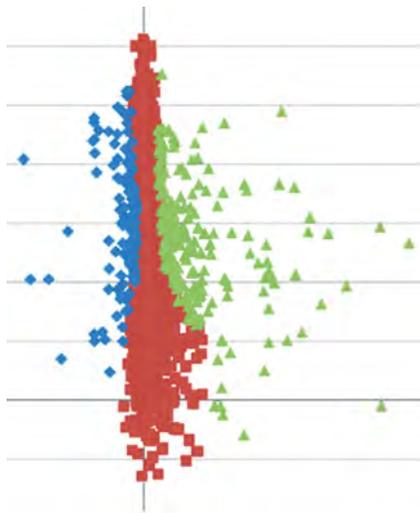
Mit der Kombination aus Elektrospray-Ionisation, MaxQuant und Nano-chromatografie ist Mann gelungen, was zuvor unmöglich schien: das vollständige Proteom eines Organismus zu entschlüsseln. Im Jahr 2008 hat der Forscher so das Proteom eines ganzen Organismus vollständig analysiert und die 4399 Proteine eines Hefepilzes bestimmt. Der

jüngste Coup gelang ihm und seinem Max-Planck-Kollegen Frank Schnorrer 2016 mit der Entschlüsselung des Fruchtfliegen-Proteoms: Rund 10000 Proteine besitzt das winzige Insekt. „Zum Vergleich: Allein im Gehirn einer Maus kommen etwa 13000 Proteine vor“, erklärt Mann. Nur zehn Prozent sind auf bestimmte Zelltypen beschränkt.

Manns Methode hat sich heute weltweit in Proteomiklabors als Standard durchgesetzt. Ohne entsprechende Computerunterstützung ist sie nach wie vor undenkbar. Cox und Mann haben die MaxQuant-Software so weiterentwickelt, dass sie inzwischen



Damit die elektrisch neutralen und in kleinere Peptidfragmente zerstückelten Proteine vom elektrischen Feld des Massenspektrometers auf unterschiedliche Flugbahnen geschickt werden können, müssen sie selbst elektrisch geladen sein. Die Peptide werden deshalb mit dem Elektrospray-Verfahren ionisiert. Dabei werden sie in einer Metallkapillare mit elektrischer Ladung versehen und an der Spitze der Kapillare versprüht.



Die Bäckerhefe war der erste Organismus, dessen Proteom analysiert wurde. Der einzellige Organismus kann sich sowohl geschlechtlich als auch ungeschlechtlich vermehren. Die Grafik zeigt das Mengenverhältnis von Proteinen in beiden Wachstumsphasen: Manche Proteine werden verstärkt während sexueller Vermehrung gebildet (blau), zum Beispiel Lockstoffproteine, die die Zellen für die Paarung benötigen. Einige werden vor allem bei der ungeschlechtlichen Vermehrung produziert (grün). Auf andere wiederum wirkt sich die Art der Vermehrung nicht aus (rot). Obwohl Zellen identische Gene haben, können sie für ihre unterschiedliche Lebensweise ein völlig verschiedenes Proteinrepertoire nutzen.

auch die Menge der in einer Probe enthaltenen Proteine ermitteln kann. So lassen sich mehrere Proben eines Patienten miteinander vergleichen und bestimmen, wie sich die Konzentration eines Proteins mit der Zeit verändert.

Bevor die Proteomik aber in die Krankenhäuser einziehen kann, müssen die Verfahren noch schneller werden. „Im Moment optimieren wir den Arbeitsablauf unserer Anlage, damit wir möglichst viele Proben untersuchen können. In einem Jahr werden wir vermutlich so weit sein, dass wir 100 Proteome pro Tag analysieren können“, sagt Matthias Mann. Dann wer-

den die Forscher etwa untersuchen können, wie sich die Proteinkonzentrationen eines Patienten im Laufe eines Tages oder mit fortschreitender Erkrankung verändert.

Schon jetzt können die Wissenschaftler Gruppen von Menschen miteinander vergleichen, um die Unterschiede im Stoffwechsel kranker und gesunder Patienten zu bestimmen. Zu diesem Zweck hat Mann zusammen mit Medizinern der Klinik der Universität Kopenhagen untersucht, wie sich die Proteine von Menschen mit Fettleibigkeit ändern, wenn diese acht Wochen lang Diät halten.

NICHT JEDER SPRICHT GLEICH AUF EINE DIÄT AN

Auf Übergewicht reagiert der Körper wie auf eine Entzündung und produziert verstärkt Proteine, die für Entzündungsreaktionen typisch sind. Die Forscher wollten nun wissen, ob während der Diät die Menge der Entzündungsproteine bei allen Patienten gleichermaßen abnimmt. Mann und seine Kollegen analysierten dazu mehr als 1000 Proteome und bestimmten mit der MaxQuant-Software die Menge der Entzündungsproteine. Es zeigte sich, dass diese während einer Diät nicht bei allen Menschen sinkt. Nicht jeder spricht also gleichermaßen darauf an.

Erschwert wird die Proteomanalyse bis heute durch die vielen Varianten, in denen Proteine vorliegen können. An altersschwache oder defekte Proteine etwa dockt das Protein Ubiquitin an. Dieses Andocken setzt einen Abbauprozess in Gang, in dem das Protein nach und nach demontiert wird. Viele Proteine werden zudem aktiviert, indem ihnen ein Phosphatmolekül angehängt wird. Dieser Prozess wird Phosphorylierung genannt.

So konnte Mann nachweisen, dass von diesem Phosphorylierungszustand der Tag-Nacht-Rhythmus eines Lebewe-

sens ganz entscheidend abhängt. „Es gibt ungeheuer viele Proteinvarianten, deren Bedeutung wir erst noch verstehen müssen. Ganze Gruppen von Proteinen können zudem unterschiedliche Zustände einnehmen. Für die Behandlung von Krankheiten sind aber gerade solche Veränderungen des Proteoms eines Menschen entscheidend“, erklärt der Wissenschaftler – der deshalb wenig von manchen der heutigen Tests zur Krankheitsdiagnose hält.

Als Beispiel nennt er den sogenannten PSA-Wert, der ein Hinweis auf Prostatakrebs sein kann, aber wegen seiner Unsicherheit umstritten ist. „Mit solchen Tests weist man die Anwesenheit oder Menge eines einzigen Proteins nach. Nach allem, was wir inzwischen wissen, reicht das aber nicht aus. Man wird künftig viel stärker auf das Proteom eines Menschen setzen, um einen Überblick über den Gesundheitsstatus zu bekommen“, sagt Mann.

Dabei soll ihm eine weitere Software seines Kollegen Jürgen Cox helfen: „Perseus“ nutzt die statistisch aufbereiteten Proteindaten von MaxQuant und führt damit eine Big-Data-Analyse durch. Die Software greift unter anderem auch auf internationale Datenbanken zu, in denen das gesammelte Wissen über Proteine gespeichert ist – zum Beispiel, wo manche Proteine vorkommen oder was es bedeutet, wenn der Stoffwechsel bestimmte Proteine vermehrt produziert. Auch die vorhandenen Erkenntnisse über Krankheiten fließen in die Analysen von Perseus mit ein.

Die Proteom-basierte Diagnose und Behandlung von Krankheiten stehen zwar noch ganz am Anfang. So wäre es extrem schwer, im Proteom frühzeitige Hinweise auf einen bösartigen Hautkrebs zu finden, denn der Tumor ist im Anfangsstadium noch sehr klein und die von ihm abgegebenen Proteinmengen entsprechend gering. Selbst mit der Nanochromatografie lassen sich so kleine Mengen nicht sicher aufspüren.



Datenanalyse am Computer (von links vorne im Uhrzeigersinn): Jan Rudolph, Jürgen Cox, Camila Duitama, Pavel Sinitcyn und Art Carlson – und dazwischen ein Schachspiel zur Entspannung.

Deutlich weiter als zu Beginn dieses Jahrtausends ist die Proteomik jedoch allemal. Nach der Entschlüsselung des menschlichen Erbguts 2001 gründeten sich viele Start-up-Unternehmen, die die Proteomanalyse als Dienstleistung für die klinische Forschung anbieten. Angesichts von Methoden wie Nanochromatografie und Elektrospray-Ionisation ist offensichtlich, dass die Technik von damals völlig unzureichend war.

Entsprechend waren die Ergebnisse für den klinischen Alltag nicht zu gebrauchen. Der Katzenjammer folgte auf dem Fuß: Viele Start-ups verschwanden von der Bildfläche, der Begriff Proteomik verkam zum bloßen Modewort. „Mit unseren neuen Verfahren sind wir einen großen Schritt weiter – im Grunde geht es erst jetzt so richtig los“, sagt Matthias Mann. Wie gut, dass er im Studium nicht ausschließlich auf Informatik und Physik gesetzt und sich schon frühzeitig für die biologischen Fragestellungen interessiert hat, die seine Forschung aufwerfen. So kann er heute dazu beitragen, das Potenzial der Proteomik für die Biologie und Medizin zu nutzen. ◀

AUF DEN PUNKT GEBRACHT

- Für eine Proteomanalyse müssen die Proteine zunächst in Peptide zerlegt werden. Nur ausgeklügelte Computer-Algorithmen können aus den dabei anfallenden riesigen Datenmengen wieder die ursprünglichen Proteine rekonstruieren.
- Das Programm MaxQuant greift dafür auf Datenbanken zu, in denen das Wissen über die Peptide und Proteine gesammelt wird.
- Mit der Software Perseus werten die Wissenschaftler Informationen aus Datenbanken zu Vorkommen und Funktion von Proteinen aus. So lässt sich die Rolle einzelner Proteine bei Erkrankungen aufdecken.

GLOSSAR

Proteinmodifikationen: Die Zahl der Proteine eines Organismus kann um ein Vielfaches höher sein als die seiner Gene. Die enorme Vielfalt entsteht durch Veränderungen nach dem Ablesen (Transkription) eines Gens oder der Übersetzung der Boten-RNA in ein Protein (Translation). Beim sogenannten alternativen Spleißen werden etwa Abschnitte der Boten-RNA herausgeschnitten oder umgruppiert, sodass unterschiedliche Genprodukte entstehen. Kleine Anhängsel wie Phosphat- oder Zuckerreste verändern dann nachträglich die Funktion von Proteinen. Wenn in einem Gen die Information für mehrere Proteine enthalten ist oder wenn eine Aminosäurekette nachträglich in mehrere Proteine gespalten wird, entstehen aus einem Gen zudem mehrere Proteine. Beim Menschen können so bis zu zehn verschiedene Proteine auf ein einzelnes Gen zurückgehen.

Proteom: Die Zahl der Proteine im menschlichen Körper wird heute auf 80 000 bis 400 000 geschätzt. Diese werden jedoch nicht von allen Zellen und zu jeder Zeit gebildet. Je nach Zelltyp besitzt eine Zelle ein anderes Proteom. Bei rund 250 verschiedenen Zelltypen des menschlichen Körpers gibt es also mindestens ebenso viele Proteome. Das Proteom hängt von vielen Faktoren ab. So kann ein Organismus je nach Alter, Ernährung oder Gesundheitszustand unterschiedliche Proteine bilden. Auch Umwelteinflüsse wie Medikamente oder Schadstoffe beeinflussen die Proteinzusammensetzung.