

De-novo-Strukturbestimmung mit NMR-Spektroskopie auf der Basis von dipolaren Restkopplungen**

Bernd Simon und Michael Sattler*

Die NMR-Spektroskopie ist neben der Kristallographie die einzige Methode, mit der die Strukturen biologischer Makromoleküle wie Proteine, Nucleinsäuren und deren Komplexe mit hoher Auflösung bestimmt werden können. Als einzige Methode ermöglicht die NMR-Spektroskopie die Bestimmung von Struktur und Dynamik dieser Moleküle in Lösung. Bisher wurden Strukturen NMR-spektroskopisch hauptsächlich über kurzreichweite Proton-Proton-Abstände bestimmt, die aus dem Kern-Overhauser-Effekt (NOE) abgeleitet werden. Hus et al. demonstrierten nun erstmals, dass die Struktur eines Proteins ausschließlich durch Messung der dipolaren Restkopplungen (RDCs; engl.: residual dipolar couplings) ermittelt werden kann, d.h. ohne die Verwendung von NOE-Daten.^[1] Wie wir im Folgenden zeigen, ist dies ein Meilenstein in der Verwendung von RDCs für die Strukturaufklärung biomolekularer Systeme.

Die wichtigste Quelle für Strukturinformationen aus NMR-Experimenten bilden die dipolaren Wechselwirkungen zwischen Protonenspins, die von den Kernabständen und der Orientierung der internuclearen Vektoren zum statischen Magnetfeld abhängen. In der Festkörper-NMR-Spektroskopie wird die Abstands- und die Winkelabhängigkeit der dipolaren Wechselwirkungen schon lange für die Charakterisierung von Molekülstrukturen herangezogen.^[2] In Lösung dagegen heben sich die dipolaren Wechselwirkungen aufgrund der schnellen wahllosen Reorientierung eines Moleküls im Mittel gegenseitig auf. Allerdings führt das fluktuierende magnetische Dipolfeld eines Kerns zur Relaxation der Magnetisierung benachbarter Kerne. Diese Kreuzrelaxation ist die Grundlage des NOE und hängt vom Abstand r zwischen den Kernen ab ($\propto r^{-6}$). In Proteinen lassen sich darüber zahlreiche Proton-Proton-Abstände messen (mit einer Obergrenze von 5–6 Å), die zusammen die dreidimensionale Struktur festlegen. Schwierigkeiten treten dann auf,

wenn die Zahl der experimentell zugänglichen Abstände gering ist, z.B. zwischen strukturell unabhängigen Untereinheiten (Domänen) eines Proteins oder, aufgrund ihrer niedrigen Protonendichte, in Nucleinsäuren. Auch ist die relative Orientierung von räumlich weit entfernten Regionen eines Moleküls wegen der kurzen Reichweite der dipolaren Wechselwirkung nicht sehr präzise bestimmbar.

Dipolare Restkopplungen

Die Schwierigkeiten bei der Strukturbestimmung durch NOE-Experimente können durch RDC-Daten ausgeglichen werden. Ein Molekül in einem verdünnten flüssigkristallinen Medium ist teilweise ausgerichtet. Infolgedessen mitteln sich die dipolaren Wechselwirkungen nicht mehr vollständig aus, und es resultiert eine kleine, zusätzliche Aufspaltung der NMR-Signale. Tjandra und Bax demonstrierten, dass die RDCs Strukturinformationen enthalten, da sie Projektionswinkel von Bindungen (beispielweise von N-H-Peptidbindungen) in einem molekularen Koordinatensystem liefern (Abbildung 1).^[3] Da man den Grad der Orientierung klein halten kann, findet nach wie vor eine schnelle Reorientierung des Moleküls statt, sodass die Linienbrei-

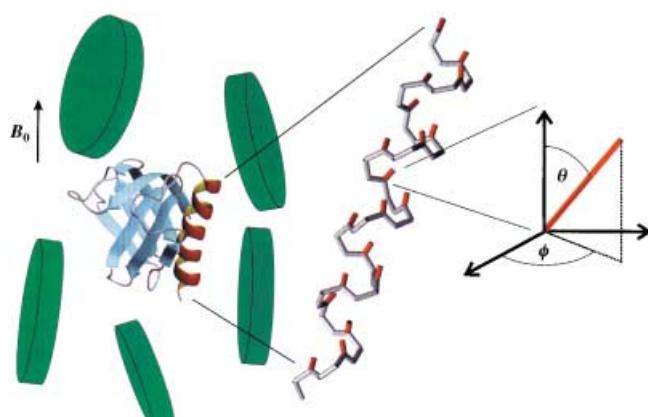


Abbildung 1. Teilweise Ausrichtung eines Proteins in einem verdünnten flüssigkristallinen Medium. Oberhalb einer bestimmten Temperatur richten sich die scheibenförmigen Bicellen (in grün) im statischen Magnetfeld B_0 aus. Dies führt zu dipolaren Restkopplungen, die von den Winkeln θ und ϕ der Vektoren zwischen den dipolar gekoppelten Kernen (beispielsweise der N-H-Bindung) abhängen.

[*] Dr. M. Sattler, Dr. B. Simon

Structural & Computational Biology
European Molecular Biology Laboratory
Meyerhofstraße 1, 69117 Heidelberg (Deutschland)
Fax: (+49) 6221-387-306
E-mail: sattler@EMBL-Heidelberg.de

[**] Wir danken Martin Blackledge für die Strukturkoordinaten von Ubiquitin und für hilfreiche Kommentare sowie Peer Bork und der DFG für ihre Unterstützung.

ten im NMR-Spektrum weitgehend unbeeinträchtigt bleiben. Die in den RDCs enthaltene Strukturinformation ist damit zugänglich, ohne dass man die in Festkörperspektren beobachteten großen Linienbreiten in Kauf nehmen müsste. Der geringe Grad der Ausrichtung (ca. 10^{-3}) wird durch sterische und/oder elektrostatische Wechselwirkung des Moleküls in einer verdünnten flüssigkristallinen Phase verursacht, kann aber auch aus der Wechselwirkung des statischen Magnetfelds mit der anisotropen magnetischen Suszeptibilität des Moleküls resultieren. Heutzutage stehen diverse flüssigkristalline Medien wie Lipid-Bicellen oder Phagen zur Verfügung, mit denen sich an nahezu jedem Molekül RDC-Messungen durchführen lassen.^[4]

Die unvollständige Ausmittelung der dipolaren Kopplungen in einer anisotropen Phase wird durch einen Ordnungstensor beschrieben. Diese „Saupe-Ordnungsmatrix“ ist ein symmetrischer Tensor zweiter Ordnung, dessen Spur verschwindet und der damit durch fünf unabhängige Parameter beschrieben werden kann. Im Hauptachsensystem des Ordnungstensors genügen zwei Tensorparameter (die axiale (A_a) und rhombische (A_r) Komponente) und die Polarkoordinaten (r_{ij} , θ , ϕ) des Verbindungsvektors der Spins (i , j) zur Berechnung der RDC-Werte D_{ij} (unter Vernachlässigung von interner Dynamik) [Gl. (1)]:

$$D_{ij} = k_{ij}[A_a(3\cos^2\theta - 1) + \frac{1}{2}A_r\sin^2\theta\cos(2\phi)] \quad (1)$$

Dabei ist $k_{ij} = -\gamma_i\gamma_j\mu_0h/(16\pi^3r_{ij}^3)$; γ ist das gyromagnetische Verhältnis, r_{ij} der als konstant angenommenen Abstand zwischen den dipolar koppelnden Kernen.

Strukturelle Parameter aus RDC-Daten

Um die Strukturinformationen aus den RDC-Messungen zu erhalten, müssen die Projektionswinkel der internuclearen Vektoren ermittelt werden. Sofern für ein Molekülfragment fünf oder mehr unabhängige RDCs gemessen werden, lässt sich der Ordnungstensor bestimmen. Zwei Ansätze zur Verwendung der Strukturinformationen aus RDCs wurden vorgeschlagen:

- Bei der Strukturbestimmung wird für die RDCs ein Pseudoenergiopotential in die Moleküldynamikrechnung/simulierte Tempern implementiert.^[5] Die Orientierung des Ordnungstensors wird durch Minimierung der Differenz zwischen den experimentellen und berechneten RDCs optimiert. Der Betrag des Ordnungstensors (A_a und A_r) kann zuvor aus der Verteilung der gemessenen RDCs abgeleitet werden;^[6] dieser wird bei der Strukturrechnung entweder konstant gehalten oder optimiert.^[7] Wie kürzlich gezeigt wurde, ist eine Strukturverfeinerung mit RDC-Daten auch ohne vorherige Kenntnis des Ordnungstensors möglich.^[8]

Die Abweichungen zwischen berechneten und gemessenen RDCs legen einen Qualitätsfaktor Q fest, der damit ein Maß ist für die Güte der Strukturbestimmung [Gl. (2)]:^[9, 10]

$$Q = \sqrt{\frac{\sum(D_{ij}^{\text{exp.}} - D_{ij}^{\text{ber.}})^2}{\sum(D_{ij}^{\text{exp.}})^2}} \quad (2)$$

- Der Ordnungstensor und die Bindungsprojektionswinkel können durch Angleich einer vorgegebenen Struktur an

die experimentellen RDCs simultan bestimmt werden. In diesem Fall lässt sich das System linearer Gleichungen, das die gemessenen RDCs mit dem Ordnungstensor und den Projektionswinkeln verbindet, mit einer Singulärwertzerlegung genannten Methode numerisch lösen.^[11] Ein Vergleich der Ordnungstensoren verschiedener Domänen eines Proteins liefert so Informationen über deren relative Orientierung und Dynamik.

Für beide Verfahren konnte gezeigt werden, dass die Güte von Strukturbestimmungen aus NMR-Messungen durch Verwendung von RDC-Daten enorm gesteigert werden kann. Insbesondere die relative Orientierung und die Dynamik einzelner Domänen in Biomakromolekülen lassen sich auf diese Weise wesentlich besser charakterisieren.^[4, 5, 12]

Der praktische Nutzen von RDC-Messungen konnte auch in Anwendungen demonstriert werden, bei denen nur wenige NOE-Daten zur Verfügung stehen oder bei denen strukturelle Informationen vorab bekannt sind.^[4] Beispiele sind die Verfeinerung von Homologiemodellen, das Docking zweier Proteine bekannter Struktur oder die Bestimmung von Tertiärstrukturen in Verbindung mit einer Fragmentsuche in der Proteinstrukturdatenbank (PDB). Durch Kombination mit der zu erwartenden vollständigen Charakterisierung aller möglichen räumlichen Tertiärstrukturmotive aus systematischen Proteinstrukturprojekten (structural genomics), kann die dreidimensionale Anordnung eines Proteins auf diese Weise schnell und effizient von einer bekannten homologen Struktur abgeleitet werden. Diese Methode auf RDC-Basis kann man als das „NMR-Äquivalent“ zum molekularen Ersatz auffassen, einer Methode, die mit großem Erfolg in der Kristallstrukturanalytik eingesetzt wird.

De-novo-Strukturbestimmung aus RDC-Daten

Mit den bisher beschriebenen Ansätzen ist es nicht möglich, die Struktur eines Proteins ausschließlich anhand von RDC-Daten zu bestimmen, d. h. ohne zusätzliche Strukturinformation aus beispielsweise NOE-Experimenten. Dies ist eine Folge der Entartung der gemessenen RDCs bezüglich der Orientierung der internuclearen Vektoren. Wie man anhand von Gleichung (1) sieht, gehören zu einer gegebenen RDC unendlich viele mögliche Kombinationen der Winkel θ und ϕ . Die infinite Menge von Lösungen für die Projektionswinkel (θ , ϕ) für eine gegebene RDC kann auf zweierlei Weise reduziert werden: 1) durch Messung mehrerer RDCs zwischen Kernen eines festen Molekülfragments wie der Peptidebene; 2) durch Kombination von RDC-Daten aus Messungen in mindestens zwei anisotropen Medien mit unterschiedlichen Ordnungstensoren.^[13, 14] Durch die Messung von fünf RDCs in einem planaren oder chiralen Molekülfragment wird dessen Orientierung bis auf eine acht- bzw. vierfache Entartung definiert. Selbst bei einem chiralen Fragment verbleiben vier mögliche Orientierungen, da die Vorzeichen von θ und ϕ in Gleichung (1) unabhängig voneinander invertiert werden können, ohne dass sich der Wert der RDCs ändert. Die Entartung lässt sich nun weiter aufheben, indem man RDCs in einem zweiten Orientierungsmedium misst. Dadurch wird die Entartung einer planaren Struktureinheit wie der Peptidbindung reduziert auf die tatsächliche und die

spiegelbildliche Orientierung, und ein chirales Motiv lässt sich eindeutig bestimmen.

Blackledge und Mitarbeiter führten erfolgreich die De-novo-Bestimmung der Faltung eines kleinen Proteins unter ausschließlicher Verwendung von RDC-Daten durch (Abbildung 2).^[1] Ihr Ansatz enthielt einen Algorithmus auf der

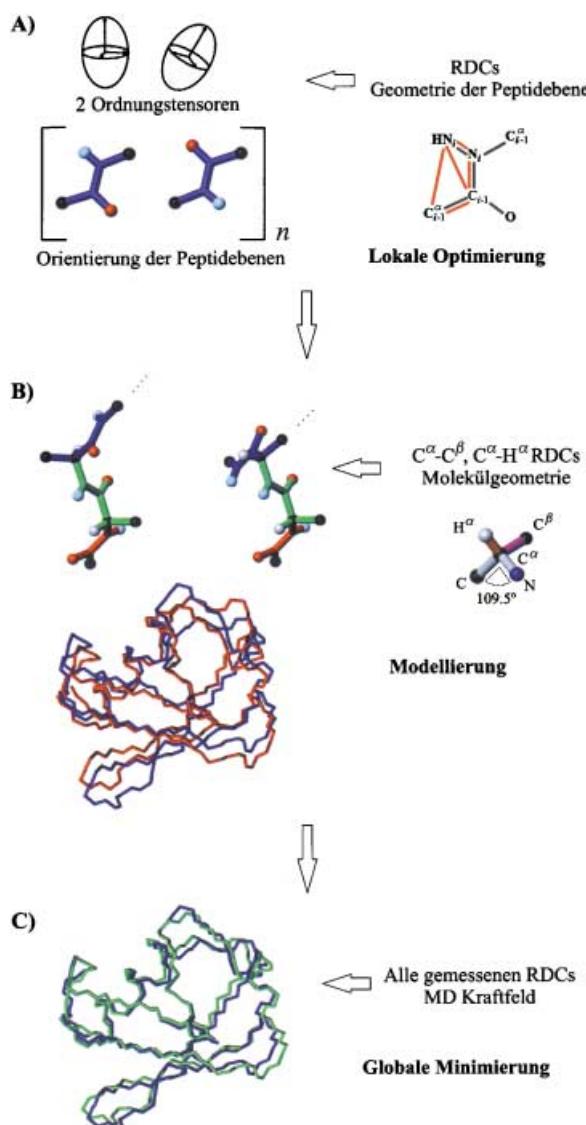


Abbildung 2. Flussdiagramm für die De-novo-Strukturbestimmung einer Proteinkette auf der Basis von RDC-Daten.^[1] A) Zwei Ordnungstensoren und die Orientierung aller Peptidebenen des Proteins werden aus RDCs (rote Linien) ermittelt, die zwischen den Kernen in der Peptidebene gemessen werden. Die lokale Optimierung definiert die Orientierung aller n Peptidebenen im Protein. Die tatsächliche und die spiegelbildliche Orientierung können auf dieser Stufe noch nicht unterschieden werden. B) Die Information über die lokale Orientierung jeder Peptidebene wird für die schrittweise Konstruktion der Peptidkette verwendet. Die zweifache Entartung der Orientierung der Peptidebenen wird unter Verwendung der RDCs der C^α-Atome aufgehoben. Überlagert man die so erhaltene Peptidkette (rot) mit der hochauflösten NMR-Struktur von Ubiquitin (PDB Strukturdatenbankeintrag 1d3z,^[10] blau), so zeigt sich, dass die Tertiärstruktur des Proteins bis auf wenige lokale Abweichungen eindeutig definiert ist. C) Zuletzt wird die Struktur mit einer Moleküldynamikrechnung optimiert, in die alle gemessenen RDCs eingehen. Die daraus resultierende Rückgratstruktur des Ubiquitins (grün) ist praktisch identisch mit einer konventionell bestimmten Struktur (blau).

Basis der Methode der kleinsten Quadrate, mit dem die Orientierung der Peptidebene anhand von RDCs in zwei verschiedenen Ausrichtungsmedien bestimmt werden kann. Die verschiedenen Peptidebenen werden sequentiell aneinandergesetzt und auf der Basis von RDCs am chiralen Verknüpfungspunkt (C^α) zweier benachbarter Peptidebenen (Abbildung 2 A) eindeutig ausgerichtet. Dadurch ist in den meisten Fällen die Orientierung der Peptidebene ($i+1$) relativ zur Peptidebene des Restes (i) eindeutig festgelegt. Wenn RDC-Daten fehlen (z.B. für Prolinreste), kann die Orientierung unter Einbeziehung von Informationen für benachbarte Reste bestimmt werden. Nach diesem ersten „tracing“ der Peptidkette wird die Konformation unter Verwendung aller gemessener RDCs in einer Moleküldynamikrechnung/simuliertes Tempern wie oben beschrieben optimiert (Abbildung 2B). Blackledge und Mitarbeiter haben dieses Verfahren an dem Protein Ubiquitin angewendet. Im schwierigsten Fall konnte dabei eine Sequenz von zwei Prolinresten, für die nur sehr wenige experimentelle Information zur Verfügung stand, erfolgreich bestimmt werden. Die daraus resultierende, allein durch RDC-Daten bestimmte Peptidrückgratkonformation von Ubiquitin ist praktisch identisch mit der Struktur, die auf einem vollständigen NOE-Datensatz basiert (Abbildung 2C).

Blackledge und Mitarbeitern ist damit ein entscheidender Durchbruch für die Anwendung von NMR-Methoden in der Strukturanalytik in flüssiger Phase gelungen. Die De-novo-Strukturbestimmung durch RDCs ist schnell und effizient. Dipolare Restkopplungen sind als Effekt erster Ordnung um eine Größenordnung stärker als NOEs, die das Ergebnis einer Störung zweiter Ordnung durch dipolare Kopplungen sind. Aus diesem Grund sind NMR-Experimente zur Messung von RDCs sehr empfindlich. RDCs zwischen heteronuclearen Spins lassen sich auch bei geringer Probenkonzentration einfach ermitteln. Zusammen mit jüngsten Entwicklungen in der NMR-Methodik (TROSY)^[15] und der Messgeräte-technik (supraleitende Empfängerspulen) sowie mit Verfahren zur Isotopenmarkierung werden NMR-Untersuchungen von Systemen mit hohem Molekulargewicht erleichtert. Die Verwendung von RDCs wird damit in Zukunft die Bestimmung der Tertiärstruktur größerer Proteine und ihrer Komplexe ermöglichen.

In der bisherigen Diskussion wurde der Einfluss der internen Dynamik auf die gemessenen RDCs vernachlässigt. Schon frühere Untersuchungen zeigten allerdings, dass sich durch RDC-Messungen Einblick in die Beweglichkeit von Biomolekülen gewinnen lässt, insbesondere in die Dynamik im Millisekundenbereich – einer Zeitskala, die über andere NMR-Parameter nur schwer zugänglich ist.^[16] Die strukturelle und dynamische Beschreibung von Biomolekülen mit RDC-Daten ist komplex, aber erste Studien sind viel versprechend und lassen auf weitere Anwendungsmöglichkeiten von RDC-Experimenten zur Charakterisierung der internen Bewegungen von Proteinen hoffen.^[17, 18]

[1] J.-C. Hus, D. Marion, M. Blackledge, *J. Am. Chem. Soc.* **2001**, *123*, 1541–1542.

[2] R. Tycko, *Annu. Rev. Phys. Chem.* **2001**, *52*, 575–606.

[3] N. Tjandra, A. Bax, *Science* **1997**, *278*, 1111–1114.

- [4] J. H. Prestegard, H. M. Al-Hashimi, J. R. Tolman, *Q. Rev. Biophys.* **2000**, *33*, 371–424.
- [5] N. Tjandra, J. G. Omichinski, A. M. Gronenborn, G. M. Clore, A. Bax, *Nat. Struct. Biol.* **1997**, *4*, 732–738.
- [6] G. M. Clore, A. M. Gronenborn, A. Bax, *J. Magn. Reson.* **1998**, *133*, 216–221.
- [7] J.-C. Hus, D. Marion, M. Blackledge, *J. Mol. Biol.* **2000**, *298*, 927–936.
- [8] S. Moltke, S. Grzesiek, *J. Biomol. NMR* **1999**, *15*, 77–82.
- [9] G. M. Clore, D. S. Garrett, *J. Am. Chem. Soc.* **1999**, *121*, 9008–9012.
- [10] G. Cornilescu, J. L. Marquardt, M. Ottiger, A. Bax, *J. Am. Chem. Soc.* **1998**, *120*, 6836–6837.
- [11] J. A. Losonczi, M. Andrec, M. W. F. Fischer, J. H. Prestegard, *J. Magn. Reson.* **1999**, *138*, 334–342.
- [12] N. R. Skrynnikov, N. K. Goto, D. Yang, W.-Y. Choy, J. R. Tolman, G. A. Mueller, L. E. Kay, *J. Mol. Biol.* **2000**, *295*, 1265–1273.
- [13] B. E. Ramirez, A. Bax, *J. Am. Chem. Soc.* **1998**, *120*, 9106–9107.
- [14] H. M. Al-Hashimi, H. Valafar, M. Terrell, E. R. Zartler, M. K. Eidsness, J. H. Prestegard, *J. Magn. Reson.* **2000**, *143*, 402–406.
- [15] K. Pervushin, *Q. Rev. Biophys.* **2000**, *33*, 161–197.
- [16] J. R. Tolman, J. M. Flanagan, M. A. Kennedy, J. H. Prestegard, *Nat. Struct. Biol.* **1997**, *4*, 292–297.
- [17] J. R. Tolman, H. M. Al-Hashimi, L. E. Kay, J. H. Prestegard, *J. Am. Chem. Soc.* **2001**, *123*, 1416–1424.
- [18] J. Meiler, J. J. Prompers, W. Peti, C. Griesinger, R. Bruschweiler, *J. Am. Chem. Soc.* **2001**, *123*, 6098–6107.