

Anpassung von Kleinwinkelröntgenspektren durch eine nicht-lineare Regressionsanalyse

Fitting of SAXS-Curves by Nonlinear Regression

V. Schilling, R. Jonak, Th. Nemetschek und H. Riedl

Pathologisches Institut der Universität Heidelberg, Abteilung für Ultrastrukturforschung

Chr. Pöppe und E. Schwander

Institut für Angewandte Mathematik der Universität Heidelberg

Z. Naturforsch. **36 c**, 333–335 (1981); received December 4, 1980

Small Angle X-Ray Diffraction, Synchrotron X-Ray, Collagen, Long Spacing

A method for fitting of background-based curves is described whereas, with its aid, the peak-positions and areas in fitted spectrum sought, are easily located.

Wie bereits gezeigt werden konnte, gelingt unter Einsatz der Synchrotronstrahlung die direkte Registrierung schnell ablaufender Umordnungen am Kollagenmolekül. Die Anzeige erfolgt durch Kurzzeitbeugungsspektren, die Änderungen der 67 nm Langperiode wiedergeben [1]. Hierdurch gelingt sowohl die Simulierung und röntgenographische Erfassung plötzlich einsetzender klinisch relevanter Faserstörungen als auch das Studium schnell ablaufender mechanochemischer Prozesse am Kollagenmolekül.

Zur genauen Auswertung der Peakorte und -flächen der ersten 12 Ordnungen dieses Kleinwinkelreflexes ist eine Glättung und Anpassung der untergrundbehafteten Kurven erforderlich. Eine Darstellung des Untergrundes durch ein Polynom bzw. eine e-Funktion mit einem Polynom im Exponenten lieferte für dieses Problem nur bedingt brauchbare Ergebnisse. Insbesondere konnte der steile Anstieg der Intensität der Untergrundstreuung zum Primärstrahl hin nicht befriedigend dargestellt werden. Ferner erwies sich der Versuch, alle eine Meßwertbildung beeinflussenden Störgrößen zu analysieren und mathematisch nachzubilden, als undurchführbar. Deswegen mußte empirisch eine Untergrundkurve gesucht werden, die superponiert mit den als Gaußkurven angenommenen Peaks die experimentell erhaltenen Werte am besten approximiert.

Die große Zahl gleichzeitig zu variierender Parameter machte es erforderlich, eine Großrechen-

anlage einzusetzen. Hierzu wurde das Programm NLSQA von Deuffhard und Apostolescu [2] mit folgendem Ansatz benutzt:

$$f_1(t) = \frac{\alpha_0}{1 + \frac{(t - x_0)^2}{\sigma_0^2}} + \beta + \alpha_s \cdot \exp \left\{ - \frac{(t - x_s)^2}{2 \sigma_s^2} \right\}$$

mit den Symbolen:

α_0 = Amplitude	} des Primärstrahls
σ_0 = Halbwertsbreite	
x_0 = Ort des Maximums	
α_s = Amplitude	} einer apparaturbedingten Störung
σ_s = Halbwertsbreite	
x_s = Ort des Maximums	
β = Konstante	
t = Kanalnummer (Abszisse des Spektrums).	

Dabei beschreibt der erste Term den prinzipiellen Verlauf des Untergrundes in Form einer Lorentz-Kurve. Die beiden anderen Terme berücksichtigen Störungen unbekannter Natur, die z. B. verursacht sein können durch Streuvorgänge am Kollimatorsystem und an Inhomogenitäten bzw. durch Abweichungen der Response-Funktion des Zählers von einer Konstanten.

Die einzelnen Peaks des Spektrums (Abb. 1) werden zunächst noch als Gauß-Kurven dargestellt, deren Maxima zuerst in einem festen Abstand Δx voneinander angenommen wurden.

Damit evtl. kleine Nichtlinearitäten, die an den Enden des Zähl drahtes des verwendeten ortsemp-

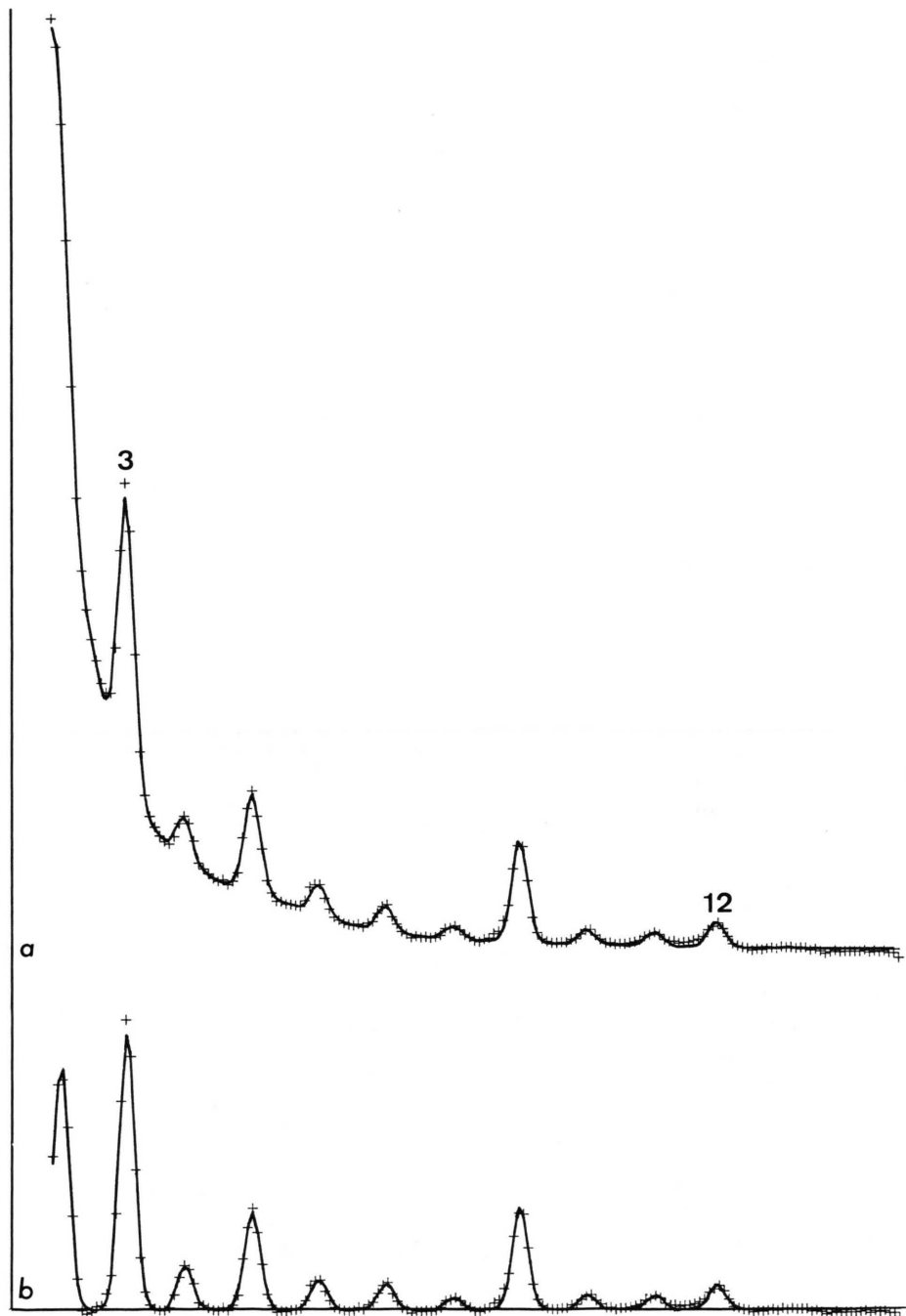


Abb. 1. a) Beispiel eines Röntgenspektrums an einer nativen Kollagenfaser aus menschlicher Beugesehne. b) Dasselbe Spektrum nach Untergrundsubtraktion. Synchrotronstrahlung: 3,3 GeV, 50 mA; Expositionszeit: 30 sec. Kleinwinkelkamera X 13 am Speicherring DORIS (DESY-Hamburg). (+: Meßdaten; —: angepaßte Kurve)

findlichen Detektors auftreten können, nicht die Anpassung global verschlechtern, wurden die Positionen der Peaks am Rande freigegeben. Für das untergrundfreie Kleinwinkelspektrum erhält man damit (für das Beispiel in Abb. 1) die Darstellung:

$$f_2(t) = \alpha_2 \cdot \exp \left\{ -\frac{(t - (\bar{x}_0 + x_2))^2}{2 \sigma_2^2} \right\} \\ + \alpha_3 \cdot \exp \left\{ -\frac{(t - (\bar{x}_0 + x_3))^2}{2 \sigma_3^2} \right\} \\ + \alpha_{12} \cdot \exp \left\{ -\frac{(t - (\bar{x}_0 + x_{12}))^2}{2 \sigma_{12}^2} \right\} \\ + \sum_{i=4}^{11} \alpha_i \cdot \exp \left\{ -\frac{(t - (\bar{x}_0 + i \cdot \Delta x))^2}{2 \sigma_i^2} \right\}.$$

Dabei bedeuten:

α_i = Amplitude des i -ten Peaks;
 x_i = Abstand des i -ten Peaks von \bar{x}_0 ;
 σ_i = Halbwertsbreite des i -ten Peaks;
 Δx = fester Abstand der Peaks untereinander;
 \bar{x}_0 = Ort des Reflexes 0-ter Ordnung.

Die registrierte Gesamtkurve läßt sich dann darstellen als:

$$f(t) = f_1(t) + f_2(t) \cong f(x, t).$$

Zur Anpassung wird die spezielle Zielfunktion $T(x)$ mit Hilfe des Gauß-Newton-Verfahrens [2] minimiert:

$$T(x) := \sum_{i=1}^m g_i^2(x),$$

wobei $g_i(x) := f(x, t_i) - y_i$; $i = 1, \dots, m$, dabei sind y_i die Meßwerte an den Orten t_i und x der Vektor der Kurvenparameter $(x_0, \dots, \alpha_0, \dots, \alpha_2, \dots, \alpha_{12}, \dots, \bar{x}_0)$.

Methodik

Das verwendete Programm NLSQA ist eine Realisierung des iterativen Gauß-Newton-Algorithmus zur nichtlinearen Regression, die sich durch besondere Eigenschaften auszeichnet: Das Verfahren gelangt auch für vergleichsweise schlechte Startdaten noch zur Konvergenz durch eine den Konvergenzbe- reich erweiternde Dämpfungsstrategie. – Auch wenn einzelne der Parameter aus den Daten nicht mehr zuverlässig bestimmt werden können, liefert das Verfahren für die übrigen korrekte Ergebnisse („numerische Rangentscheidung“).

Trotz der Größenordnung des Problems (ca. 200 Meßpunkte und ca. 25 Parameter) benötigt das Verfahren für die Anpassung der ersten Kurve aus einer Reihe zeitlich aufeinander folgender Messungen $\sim 7-10$ sec, für jede folgende < 1 sec Rechenzeit.

Wegen der großen Menge der anfallenden Daten wurde ein Programm zur interaktiven Startdatenerzeugung erarbeitet, womit die Meßdatenauswertung nahezu vollständig automatisiert ist.

Die Rechnungen wurden auf der Anlage IBM 370/168 des Rechenzentrums der Universität durchgeführt.

Ausblick

Ein neuer Ansatzpunkt für die Analyse ergibt sich, wenn man davon ausgeht, daß die Kollagenstruktur als Realisation eines „schwach stationären Systems“ [3] angesehen werden kann, dessen Fouriertransformierte das Röntgenbeugungsspektrum ist.

Dies bedeutet, daß innerhalb des Volumens, das vom Röntgenstrahl erfaßt wird und damit zur Entstehung des Beugungsdiagramms beiträgt, der statistische Zustand in jedem Volumenelement derselbe ist. Man kann nun zeigen, daß die Maxima der Fouriertransformierten einer solchen Struktur die Form von Cauchy-Kurven (allgemeine Form $y = \frac{ax+b}{c+x^2}$) besitzen. Für $a=0$ erhält man die symmetrische Lorentz-Kurve, während es bei geeigneter Wahl des Parameters möglich sein sollte, auch die in manchen Fällen beobachteten asymmetrischen Peakformen zu erklären.

Acknowledgements

Herrn Prof. Dr. P. Deuffhard danken wir für wertvolle Diskussion, Herrn W. Härdle für Hinweise.

Mit Unterstützung durch das BMFT (05113 NE P) und die DFG (SFB 123).

- [1] Th. Nemetschek, R. Jonak, H. Nemetschek-Gansler, H. Riedl und G. Rosenbaum, Z. Naturforsch. **33 c**, 928 (1978).
- [2] P. Deuffhard und V. Apostolescu, A study of the Gauß-Newton algorithm for the solution of nonlinear least squares problems, in: Special topics of applied mathematics (J. Frehse, D. Pallaschke, and U. Trottenberg, eds.), North-Holland, Amsterdam 1980.
- [3] U. Grenander und M. Rosenblatt, Statistical Analysis of Stationary Time Series, Chapter 1. John Wiley and Sons Ltd., New York-Stockholm 1957.