

## Ein Programmpaket zur rechnerischen Konformationsanalyse von Peptiden auf dem TR440

A Program Package for the Analysis of Peptide Conformations on the TR 440

I. Knack und K.-H. Röhm

Institut für Physiologische Chemie der Phillips-Universität, Lahnberge, D-3550 Marburg/Lahn

Z. Naturforsch. 36 c, 347–349 (1981);  
eingegangen am 29. September 1980

Peptide Conformations, Energy Calculations, Computer Programs

The "Empirical Conformational Energy Program for Peptides (ECEPP)" by Scheraga and coworkers which calculates the conformational energies of given peptide conformations was adapted for use with the CGK TR 440 computer system. New programs for energy minimization, statistical data analysis and graphical display of peptide conformations were added. In this way the range of application of the original ECEPP program was considerably extended. At the same time, its handling was greatly simplified by an improved data input/output.

### Einleitung

Konformationsuntersuchungen können wesentliche Beiträge zum Verständnis des Zusammenhangs zwischen Struktur und Funktion von Peptiden und Proteinen liefern. Die hohe biologische Aktivität von Enzymen und Peptidhormonen setzt eine sehr spezifische Wechselwirkung mit einem Substrat bzw. einem Rezeptor voraus, was eine räumlich wohldefinierte Molekülstruktur erfordert. Die Untersuchung solcher Strukturen sowie der konformationsbestimmenden Einflüsse ist daher für die biochemische Forschung von großem Interesse. Neben den bekannten experimentellen Methoden wie Röntgenstrukturanalyse, NMR, ORD, CD, IR usw. gewinnen hierbei empirische Energierchnungen zunehmend an Bedeutung. In den letzten Jahren wurden in der theoretischen Konformationsanalyse von Peptiden mit Hilfe von Computerrechnungen beachtliche Fortschritte erzielt. Empirische Energierchnungen wurden u. a. zu Strukturuntersuchungen von Enzym-Substratkplexen [1, 2] und von Peptidhormonen [3, 4] sowie zur Verfeinerung von Röntgenstrukturdaten eingesetzt [5, 6]. Große Probleme wirft nach wie vor die theoretische Behandlung der Faltung der Proteine

auf. In diesem Zusammenhang sind Untersuchungen der konformationsbestimmenden Wechselwirkungen an kleinen Modellpeptiden Gegenstand weiterer Arbeiten [7–11].

### Beschreibung der Programme

Ein sehr leistungsfähiges FORTRAN-Programm zur Berechnung des Energieinhalts von Peptidkonformeren wurde von Scheraga und Mitarbeitern entwickelt [12]. Das Programm ermittelt für einen gegebenen Satz von Diederwinkeln die Atomkoordinaten und die relative Konformationsenergie einer Peptidkette. Die Energie- und Geometrieparameter der im Molekül enthaltenen Aminosäuren und Endgruppen werden einem festgelegten Datensatz entnommen [13]. Dabei können insgesamt 36 verschiedene Aminosäuren und 20 verschiedene Endgruppen berücksichtigt werden. Aus empirischen Energiepotentialfunktionen [13] wird die Gesamtenergie  $E_{\text{tot}}$  als Summe aller intramolekularen Wechselwirkungsenergien berechnet:

$$E_{\text{tot}} = E_{\text{es}} + E_{\text{nb}} + E_{\text{tor}} + E_{\text{cystr}} + E_{\text{ring}}$$

mit:

$E_{\text{es}}$  = elektrostatische Energie;

$E_{\text{nb}}$  = nichtbindende Energie einschließlich H-Brückenenergie;

$E_{\text{tor}}$  = Torsionsenergie;

$E_{\text{cystr}}$  = Torsionsenergie bei S-S-Brücken;

$E_{\text{ring}}$  = Ringspannungsenergie bei S-S-Brücken.

Die Originalversion dieses Programms wurde an das Rechnersystem TR 440 adaptiert und der Anwendungsbereich durch verschiedene Unterprogramme sowie zusätzliche Möglichkeiten der Daten-Ein- und -Ausgabe erweitert. Dies war erforderlich, weil das Originalprogramm für systematische Konformationsstudien unzureichend ausgelegt ist. Für bestimmte Aufgabenstellungen wurden zwei in ihrem Leistungsumfang unterschiedliche Versionen geschaffen.

#### 1. ECEPP/G

Diese Programmversion ermöglicht einen sehr komfortablen interaktiven Dialogbetrieb am Terminal. So können für ein gegebenes Peptidmolekül beliebige Diederwinkel variiert und neben der Konformationsenergie ausgewählte Atomabstände berechnet

Sonderdruckanforderungen an Dr. K.-H. Röhm.  
0341-0382/81/0300-0347 \$ 01.00/0



Dieses Werk wurde im Jahr 2013 vom Verlag Zeitschrift für Naturforschung in Zusammenarbeit mit der Max-Planck-Gesellschaft zur Förderung der Wissenschaften e.V. digitalisiert und unter folgender Lizenz veröffentlicht: Creative Commons Namensnennung-Keine Bearbeitung 3.0 Deutschland Lizenz.

Zum 01.01.2015 ist eine Anpassung der Lizenzbedingungen (Entfall der Creative Commons Lizenzbedingung „Keine Bearbeitung“) beabsichtigt, um eine Nachnutzung auch im Rahmen zukünftiger wissenschaftlicher Nutzungsformen zu ermöglichen.

This work has been digitized and published in 2013 by Verlag Zeitschrift für Naturforschung in cooperation with the Max Planck Society for the Advancement of Science under a Creative Commons Attribution-NoDerivs 3.0 Germany License.

On 01.01.2015 it is planned to change the License Conditions (the removal of the Creative Commons License condition "no derivative works"). This is to allow reuse in the area of future scientific usage.

werden. Zusätzlich wurde ein Unterprogramm zur perspektivischen Darstellung von Molekülstrukturen eingefügt, mit dessen Hilfe sofort jede neue Konformation auf einem graphischen Sichtgerät dargestellt werden kann. Der Beobachtungspunkt, von dem aus das Molekül im Raum betrachtet wird, lässt sich dabei beliebig variieren. Die Größe der Atombole wird als Funktion des Abstandes vom Betrachter dargestellt; dadurch wird die perspektivische Wirkung noch erhöht (siehe Abb. 1). Die zusätzlich geschaffene Möglichkeit, Diederwinkel innerhalb eines beliebigen Bereiches mit frei wählbarer Schrittweite zu variieren ist besonders dann wichtig, wenn durch Modellbetrachtungen Startkonformationen für eine nachfolgende Energieminimierung gefunden werden sollen.

Zur allgemeinen graphischen Darstellung von Molekülstrukturen auf einem Plotter steht zusätzlich eine erweiterte Version des oben erwähnten Unterprogramms zur Verfügung.

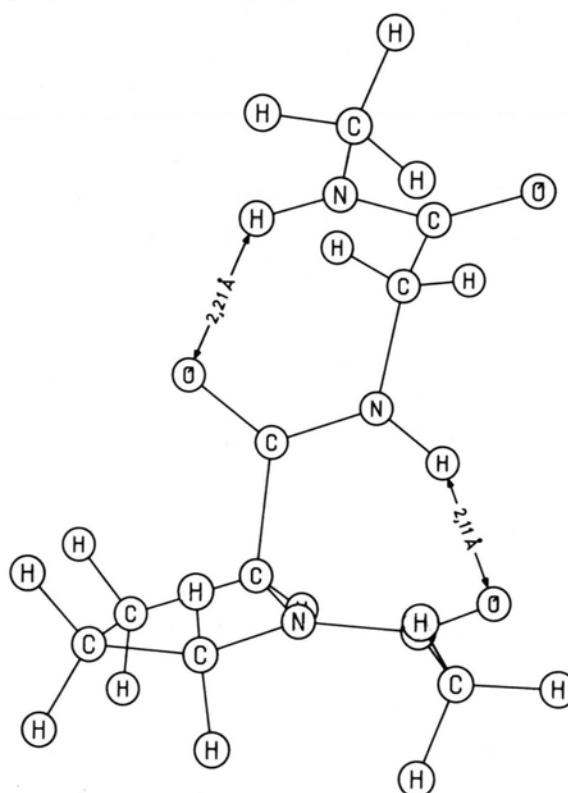


Abb. 1. Raumstruktur von N-Acetyl-Pro-Gly-methylamid am globalen Energieminimum. Die Konformation wird u. a. durch die eingezeichneten H-Brückenbindungen stabilisiert. Es wurden folgende Energieterme gefunden:  $E_{\text{es}} = -12.79$ ;  $E_{\text{nb}} = -3.09$ ;  $E_{\text{tot}} = -15.88$ ; (alle Angaben in kcal/mol).

Tab. I. Konformationsenergien und Besetzungswahrscheinlichkeiten  $w$  nach der Boltzmannverteilung für N-Acetyl-Pro-Gly-methylamid (alle Energieminima mit  $\Delta E \leq 3$  kcal/mol sind aufgeführt).

$\Delta E$	w	$h^b$ [Å]	$R^c$ [Å]	Diederwinkel <sup>a</sup>		
				$\psi_{\text{Pro}}$	$\phi_{\text{Gly}}$	$\psi_{\text{Gly}}$
0,0	0,208	4,88	5,98	78,9	83,7	-74,1
0,1	0,177	3,70	4,97	82,1	77,1	44,9
0,29	0,128	5,84	8,85	79,5	-82,5	76,1
0,34	0,117	3,51	5,86	79,5	169,1	-50,9
0,73	0,061	4,30	6,38	158,9	83,8	-73,6
0,88	0,047	6,25	8,24	161,0	-82,4	76,3
0,94	0,042	4,18	7,06	76,2	174,0	60,3
0,98	0,040	4,26	5,92	-49,1	-84,6	73,5
1,10	0,033	5,21	8,64	77,8	-73,8	-55,2
1,26	0,025	5,75	7,30	77,2	74,1	-154,1
1,37	0,021	6,30	8,85	-47,3	82,8	-76,7
1,41	0,019	5,47	8,08	80,1	180,0	180,0
1,45	0,018	6,33	9,28	80,0	-73,2	147,9
1,83	0,010	7,12	10,26	160,6	180,0	180,0
1,90	0,009	4,31	5,06	-22,4	-159,9	54,9
1,90	0,008	6,89	8,00	-47,3	-179,5	178,0
1,90	0,008	5,40	8,81	153,9	167,0	-62,5
1,92	0,008	3,57	5,44	-24,2	-71,3	-42,3
2,10	0,006	6,69	8,84	159,8	-71,7	-51,9
2,24	0,005	6,87	9,87	160,6	-169,4	61,1
2,42	0,004	6,63	7,58	-47,5	171,3	-59,2
2,69	0,003	6,61	8,60	-46,1	73,3	50,3
2,64	0,002	6,36	8,11	-26,8	73,2	50,1

<sup>a</sup> Nur die variablen Diederwinkel sind aufgeführt; alle anderen wurden konstant gehalten und betragen  $180^\circ$  bzw.  $-75^\circ$  für  $\phi_{\text{Pro}}$ .

<sup>b</sup> Abstand zwischen dem O-Atom im Acetylrest und dem N-Atom des Amidrestes in Å.

<sup>c</sup> Abstand zwischen den beiden C-Atomen der endständigen Methylgruppen in Å.

## 2. ECEPP/A

Diese Version wurde speziell für die Lösung von Problemen entwickelt, die einen hohen Aufwand an Rechenzeit erfordern. Das Programm berechnet neben Konformationen minimaler Energie auch die für die Darstellung von  $\varphi$ - $\psi$ -Winkeldiagrammen nach Ramachandran [14] erforderlichen Daten. Zur Energieminimierung wurde ein sehr leistungsfähiges Unterprogramm aus der TR 440-Programmbibliothek [15] eingefügt. Durch Variation von Diederwinkel wird – ausgehend von einer Anfangsnäherung –, nach einem Gradientenverfahren, eine iterativ verbesserte Näherung für ein lokales Energieminimum berechnet. Zur Iteration dient ein Quasi-Newton Verfahren [16], bei dem die Gradienten numerisch berechnet werden. Zu dem besonderen Problem der Auswahl günstiger Startkonformationen sei auf die Arbeiten von Scheraga und Mitarbeitern verwiesen [9, 17, 18, 22].

Zur weiteren Auswertung der bei der Energieminimierung erhaltenen Ergebnisse wurde ein Statistik-Programm geschrieben, das nach der Boltzmannverteilung die Besetzungswahrscheinlichkeiten für die einzelnen Energieminima berechnet. Außerdem können die mittl. Energie sowie mittl. Atomabstände berechnet werden. Eine Erweiterung auf die Berechnung anderer Mittelwerte, wie z. B. mittl. NMR-Kopplungskonstanten, ist ohne weiteres möglich. Da das genannte Minimierungs-Unterprogramm am Ende des Iterationsprozesses zusätzlich Werte für die zweiten partiellen Ableitungen (Hesse-Matrix) am jeweiligen Energieminimum liefert, kann bei der statistischen Analyse prinzipiell auch der Anteil der Schwingungsentropie berücksichtigt werden [11, 19–21]. Brauchbare Resultate sind hier allerdings nur bei extrem großer Rechengenauigkeit und damit hohen Rechenzeiten zu erwarten.

Die Ergebnisse einer Energieminimierung durch das Programm ECEPP/A sollen am Beispiel des N-Acetyl-Pro-Gly-methylamids vorgestellt werden. Scheraga und Mitarbeiter haben gezeigt, daß bei kleinen Peptiden die intramolekularen Wechselwirkungen innerhalb eines Aminosäurerestes dominie-

rend sind [8, 9]. Wenn dies zutrifft, stellen die meisten Energieminima eines Peptids Kombinationen der Minima der beteiligten Aminosäurereste dar; ein Befund, der für die Auswahl günstiger Startkonformationen sehr hilfreich ist. Durch Kombination der Diederwinkel der Energieminima der N-Acetyl-methylamide von Pro und Gly [7, 18] wurde zunächst ein Satz von 45 Startkonformationen ermittelt. Die anschließende Minimierung der Konformationsenergie führte dann zu dem in Tab. I zusammengefaßten Ergebnis. Die Konformationsenergie ist dabei in der Form  $\Delta E = E - E_0$  dargestellt, wobei  $E_0$  die Energie  $E$  am globalen Minimum ist;  $E_0$  beträgt hier – 15.88 kcal/mol. Eine perspektivische Darstellung der Molekülstruktur am globalen Energieminimum zeigt die Abb. 1. Dieselben Ergebnisse wurden auch von Scheraga und Mitarbeitern gefunden [7].

Über weitere Ergebnisse dieser Arbeiten, die im Rahmen einer Dissertation zum Studium der Konformationseigenschaften katalytisch aktiver Modellpeptide durchgeführt wurden, soll zu einem späteren Zeitpunkt berichtet werden. Die hier vorgestellten Programme stehen jedoch schon jetzt Interessenten zur Verfügung.

- [1] A. Warshel und M. Levitt, *J. Mol. Biol.* **103**, 227–149 (1976).
- [2] M. R. Pincus, S. S. Zimmerman und H. A. Scheraga, *Proc. Nat. Acad. Sci.* **73**, 4261–4265 (1976).
- [3] D. Kotelchuk, H. A. Scheraga und R. Walter, *Proc. Nat. Acad. Sci.* **69**, 3629–3633 (1972).
- [4] J. Vicar, E. Abillon, F. Toma, F. Piriou, K. Lintner, K. Bláha, P. Fromageot und S. Fermandjian, *FEBS-Letters* **97**, 275–278 (1979).
- [5] P. K. Warmer und H. A. Scheraga, *Biochemistry* **13**, 757–767 (1974).
- [6] M. Levitt, *J. Mol. Biol.* **82**, 393–420 (1974).
- [7] S. S. Zimmerman und H. A. Scheraga, *Biopolymers* **16**, 811–843 (1977).
- [8] S. S. Zimmerman und H. A. Scheraga, *Biopolymers* **17**, 1849–1869 (1978).
- [9] S. S. Zimmerman und H. A. Scheraga, *Biopolymers* **17**, 1871–1890 (1978).
- [10] S. Premilat und B. Maigret, *J. Chem. Phys.* **66**, 3418–3425 (1977).
- [11] S. S. Zimmerman, L. L. Shipman und H. A. Scheraga, *J. Phys. Chem.* **81**, 614–622 (1977).
- [12] M. J. Browman *et al.*, ECEPP – Empirical Conformational Energy Program for Peptides; das Programm ist erhältlich beim Quantum Chemistry Program Exchange, Indiana University, Bloomington Ind. 47401, USA (Bestellnummer QCPE 286) 1975.
- [13] F. A. Momany, R. F. McGuire, A. W. Burgess und H. A. Scheraga, *J. Phys. Chem.* **79**, 2361–2381 (1975).
- [14] G. N. Ramachandran und V. Sasisekharan, *Adv. Protein Chem.* **23**, 283–438 (1968).
- [15] FORTRAN-Unterprogramm FUMIOG der TR 440-Software Liste (Telefunken-Index-Nr.: F2.HH.01.24); Rechenzentrum der Universität Marburg, 1980.
- [16] P. E. Gill und W. Murray, *J. Inst. Maths. Applies.* **9**, 91–108 (1972).
- [17] P. N. Lewis, F. A. Momany und H. A. Scheraga, *Isr. J. Chem.* **11**, 121–152 (1973).
- [18] S. S. Zimmerman, M. S. Pottle, G. Némethy und H. A. Scheraga, *Macromolecules* **10**, 1–9 (1977).
- [19] P. J. Flory, *Macromolecules* **7**, 381 (1974).
- [20] N. Gö und H. A. Scheraga, *J. Chem. Phys.* **51**, 4751–4767 (1969).
- [21] N. Gö und H. A. Scheraga, *Macromolecules* **9**, 535–542 (1976).
- [22] I. Simon, G. Némethy und H. A. Scheraga, *Macromolecules* **11**, 797–804 (1978).