REFINEMENT OF THE FUNCTION FOR THE EVALUATION OF THE N-H...O HYDROGEN BOND POTENTIAL ENERGY*

BY R. BALASUBRAMANIAN AND C. RAMAKRISHNAN‡

(Centre of Advanced Study in Physics, University of Madras, Guindy Campus, Madras-25, India)

Received May 12, 1972

(Communicated by Prof. G. N. Ramachandran, F.A.SC.)

ABSTRACT

The parameters of the hydrogen bond potential energy function proposed earlier have been refined using the increased data available on N–H···O hydrogen bonds. The full-matrix least-square refinement procedure has been used and the discrepancy index has dropped down to a final value of 0·176. A linear interpolation has been proposed for the border regions where the hydrogen bonds are weak and the results are discussed.

THE original function proposed by Lippincott and Schroeder (1955) and Schroeder and Lippincott (1957) for the evaluation of the contribution to the energy of a conformation due to N–H...O hydrogen bonding has been modified by Chidambaram, Balasubramanian and Ramachandran (1970). Later an empirical form for this function has been proposed (Balasubramanian, Chidambaram and Ramachandran, 1970) and the parameters used were varied to get a satisfactory agreement between theory and experimental data taken from Ramachandran and Ramakrishnan (1968). In order to assess the agreement quantitatively a discrepancy index D (R, θ) was proposed based upon probability considerations and a minimum value of 0·405 had been obtained for this index (Balasubramanian, Chidambaram and Ramachandran, 1970).

Recently an analysis of the N–H...O hydrogen bonds occurring in amino acids and peptides has been done by Ramakrishnan and Nageshwar

---

[*] Contribution No. 346 from the Centre of Advanced Study in Physics, University of Madras, Madras, India.

[‡] Present address ; Molecular Biophysics Unit, Indian Institute of Science, Bangalore-12.

Prasad (1971) and thus more data has been made available. Hence a better fit of the theory with experiment has been attempted by using a full-matrix least-square refinement of the parameters involved in the earlier proposed function. The results are reported in this paper along with a possible simplification in the extrapolation procedure in the border regions of the (R, θ) plane (the regions where hydrogen bonds are weak and slowly goes over into the non-bonded interaction).

The form of the empirical function is given by (*see* Balasubramanian, Chidambaram and Chandrasekharan, 1970)

$$V_{hb}(R, \theta) = V_{\text{min.}} + p_1 (R - R_{\text{min.}})^2 + p_2 (R - R_{\text{min.}})^3 + \exp.\{p_3 (R - R_{\text{min.}})\} (q_1 \theta^2 + q_2 \theta^3) \qquad (1)$$

where R and θ are the hydrogen bond length and angle. The variable parameters are $R_{\text{min.}}$ p_1, p_2, p_3, q_1 and q_2.

As discussed in Balasubramanian, Chidambaram and Ramachandran (1970), the normalized probability distribution function is given by

$$P(R, \theta) = \frac{R^2 \sin\theta \exp.\{-V(R, \theta)/a\}}{\iint R^2 \sin\theta \exp.\{-V(R, \theta)/a\} dR d\theta}. \qquad (2)$$

For computational purposes a quantity $N'(R, \theta)$ proportional to $P(R, \theta)$ given by

$$N'(R, \theta) = \frac{R^2 \sin\theta \exp.\{-V(R, \theta)/a\}}{\Sigma\Sigma R^2 \sin\theta \exp.\{-V(R, \theta)/a\}} \qquad (3)$$

has been used, where the summation is done over the various grid points in the (R, θ) plane. The data given in Ramakrishnan and Nageshwar Prasad (1971) has been used for determining the experimental (or observed) values of $N'(R, \theta)$ using the method described in Balasubramanian, Chidambaram and Ramachandran (1970).

The parameters of equation (1) can be collectively referred to as s_k' $(k = 1, 6)$. Then the "normal equations" for the least square refinement can be written in the matrix form as:

$$\begin{bmatrix} \sum_i s'_{1i}{}^2 & \sum_i s'_{1i}s_{2i}' & \cdots & \sum_i s'_{1i}s_{6i}' \\ \sum_i s'_{2i}s_{1i}' & & \cdots & \sum_i s'_{2i}s_{6i}' \\ \cdots & & \cdots & \cdots \\ \cdots & & \cdots & \cdots \\ \sum_i s'_{6i}s_{1i}' & & \cdots & \sum_i s'_{6i}{}^2 \end{bmatrix} \begin{bmatrix} \triangle s_1 \\ \triangle s_2 \\ \cdots \\ \cdots \\ \triangle s_6 \end{bmatrix} = \begin{bmatrix} \sum_i s'_{1i}\triangle N_i' \\ \sum_i s'_{2i}\triangle N_i' \\ \cdots \\ \cdots \\ \sum_i s'_{6i}\triangle N_i' \end{bmatrix} \tag{4}$$

where

(*a*) $\triangle N_i'$ is the arithmetic difference between the observed and calculated probability densities of the *i*-th grid point of the (R, θ) mesh.

(*b*) $\triangle s_k$ ($k = 1$ to 6) are the shifts obtainable from refinement for the parameters, and

(*c*) $$s_k' = \frac{\partial N_i'}{\partial s_k}\,. \tag{5}$$

Further we have,

$$\frac{\partial N_i'}{\partial s_k} = \frac{\partial N_i'}{\partial V_i}\cdot\frac{\partial V_i}{\partial s_k} \tag{6}$$

$$\frac{\partial N_i'}{\partial V_i} = -\frac{1}{\alpha}\left\{\left(\frac{R_i^2 \sin\theta_i e^{-V_i/\alpha}}{\Sigma\Sigma R_i^2 \sin\theta_i e^{-V_i/\alpha}}\right) - \left(\frac{R_i^2 \sin\theta\, e^{-V_i/\alpha}}{\Sigma\Sigma R_i^2 \sin\theta_i e^{-V_i/\alpha}}\right)^2\right\} \tag{7}$$

and the derivatives $\partial V_i/\partial s_k$ can be obtained by differentiating (1) with respect to each parameter. From previous experience (Balasubramanian, Chidambaram and Ramachandran, 1970), $V_{min.}$ has been chosen to be $-5{\cdot}0$ kcal/mole. The refinement has been done by programming it for the IBM 1130 computer. The converged values of the parameters are,

$R_{min.} = 2{\cdot}8$ A
$p_1 = 27$
$p_2 = -25$
$p_3 = -1{\cdot}4$
$q_1 = 0{\cdot}0085$
$q_2 = -0{\cdot}0002$.

The final D index $(D(R, \theta))$ value reached is 0·176. The discrepancy indices D(R) and D(θ) for the two independent variables R and θ were found to be 0·024 and 0·095. It can be seen that these values have dropped down considerably from those given in Balasubramanian, Chidambaram and Ramachandran, 1970). The old and new values are:

D(R) = 0·196, 0·024
D(θ) = 0·170, 0·095
D(R, θ) = 0·405, 0·176.

For a small region in the (R, θ) plane which lies between the region where V_{hb} is expected to be valid and the region where the non-bonded energy is to be reckoned, a smoothing function using a cubic interpolation has been proposed by Balasubramanian, Chidambaram and Ramachandran (1970). The flow chart for such a procedure is given by Ramachandran, Chandrasekharan and Chidambaram (1971). Since this requires special computational efforts, a linear interpolation has been attempted here. For this purpose, the region where the value of R lies between 2·95 A and 3·60 A and/or the value of θ lies between 25° and 45°, has been considered where the energy of the conformation is expected to be in between that of a complete hydrogen bond and pure non-bonded interactions. With this linear interpolation, the energy values obtained in the (R, θ) plane are given in Table I.

TABLE I

The value of the potential energy (in kcal/mole) due to N–H...O hydrogen bonding obtained using refined parameters

θ in (°) \ R in (A)	2·60	2·70	2·80	2·90	3·00	3·10	3·20	3·30	3·40	3·50	3·60
0	−3·65	−4·69	−5·00	−4·76	−4·26	−3·78	−3·31	−2·83	−2·35	−1·88	−1·40
5	−3·38	−4·46	−4·81	−4·60	−4·13	−3·67	−3·22	−2·76	−2·31	−1·85	−1·39
10	−2·70	−3·90	−4·34	−4·22	−3·81	−3·40	−3·00	−2·59	−2·19	−1·78	−1·37
15	−1·84	−3·20	−3·76	−3·73	−3·40	−3·05	−2·71	−2·37	−2·03	−1·68	−1·34
20	−1·03	−2·52	−3·20	−3·27	−3·01	−2·72	−2·43	−2·15	−1·86	−1·58	−1·29
25	−0·47	−2·06	−2·82	−2·96	−2·74	−2·49	−2·24	−1·99	−1·73	−1·48	−1·23
30	−0·09	−1·57	−2·29	−2·48	−2·33	−2·14	−1·94	−1·74	−1·55	−1·35	−1·16
35	0·27	−1·07	−1·77	−2·00	−1·92	−1·78	−1·64	−1·50	−1·36	−1·22	−1·08
40	0·65	−0·58	−1·24	−1·51	−1·51	−1·42	−1·34	−1·25	−1·16	−1·08	−0·99
45	1·03	−0·08	−0·71	−1·03	−1·10	−1·07	−1·03	−1·00	−0·96	−0·93	−0·89

To compare these energies with those got from cubic interpolation, the difference between the values of the linear and the cubic methods, is given in Table II. From this table it is seen that in most of the places the difference is only 0·2 to 0·4 kcal/mole. When R is around 2·8 A and θ is around 30°, the present study gives an appreciably lower value for the energy, which seems to be justified by the occurrence of observed hydrogen bonds having values of R and θ around this value (Ramakrishnan and Nageshwar Prasad, 1971). A few cases that can be quoted from recent crystal structures are (*a*) 1, 4-cyclohexadiene-1-gly (Jandecek and Simonsen, 1969) ($R = 2{\cdot}75$ A and $\theta = 30°$); (*b*) N-methyl-DL-leu-gly HBr (Chandrasekharan and Subramanian, 1969) ($R = 2{\cdot}81$ A, $\theta = 30°$); (*c*) α-gly. gly (Biswas, Hughes and Wilson, 1968) ($R = 2{\cdot}71$ A, $\theta = 20°$); (*d*) L-citrulline HCl (Naganathan and Venkatesan, 1971) ($R = 2{\cdot}90$ A, $\theta = 30°$).

Thus the present approach has shown that the hydrogen bond parameters obtained from refinement, coupled with a linear interpolation in the border regions can be taken to be fairly satisfactory for practical purposes.

TABLE II

The difference in energy values obtained by cubic and linear interpolation. The blanks represent regions where interpolation is not necessary. Note that the difference is always positive showing that linear approximation gives a lower energy value

R in (A) / θ in (°)	2·60	2·70	2·80	2·90	3·00	3·10	3·20	3·30	3·40	3·50
0	..	..	..	..	0·08	0·19	0·30	0·36	0·34	0·24
5	..	..	..	..	0·08	0·19	0·29	0·34	0·34	0·23
10	..	..	..	..	0·07	0·16	0·25	0·30	0·29	0·20
15	..	..	..	..	0·06	0·13	0·20	0·24	0·23	0·15
20	1·70	0·43	0·02	0·02	0·05	0·10	0·14	0·18	0·17	0·12
25	3·82	1·79	0·76	0·23	0·03	0·07	0·11	0·14	0·13	0·09
30	4·75	3·03	1·64	0·62	0·15	0·04	0·05	0·06	0·01	0·03
35	2·78	2·10	1·62	0·88	0·32	0·01	0·14	0·01	0·11	0·17
40	1·21	0·93	0·76	0·58	0·32	0·04	0·11	0·05	0·01	0·05

ACKNOWLEDGEMENT

One of the authors (R. B.) thanks the University Grants Commission, India, for financial support.

REFERENCES

Balasubramanian, R., Chidambaram, R. and Ramachandran, G. N.	*Biochim. Biophys. Acta*, 1970, **221**, 197.
Biswas, A. B., Hughes, E. W. and Wilson, J. N.	*Acta Cryst.*, 1968, **24 B**, 40.
Chandrasekharan, R. and Subramanian, E.	*Ibid.*, 1968, **24 B**, 40.
Chidambaram, R., Balasubramanian, R. and Ramachandran, G. N.	*Biochim. Biophys. Acta*, 1970, **221**, 182.
Jandecek, R. J. and Simonsen, S. H.	*J. Amer. Chem. Soc.*, 1969, **91**, 6663.
Lippincott, E. R. and Schroeder, R.	*J. Chem. Phys.*, 1955, **23**, 1099.
Naganathan, P. S. and Venkatesan, K.	*Acta Cryst.*, 1971, **27 B**, 1079.
Ramachandran, G. N., Chandrasekharan, R. and Chidambaram, R.	*Proc. Indian Acad. Sci.*, 1971, **74 A**, 270.
——— and Ramakrishnan, C.	In W. G. Gewther, *Symp. on Fibrous Protein, Australia* 1967, Butterworths, Australia, 1968, p. 71.
Ramakrishnan, C. and Nageshwar Prasad	*Int. J. Protein Res.*, 1971, **3**, 209.
Schroeder, R. and Lippincott, E. R.	*J. Chem. Phys.*, 1957, **61**, 921.