

# POTENTIAL FUNCTIONS FOR HYDROGEN BOND INTERACTIONS

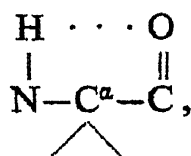
## III. Empirical Potential Function for the Peptide N-H...O=C Hydrogen Bond

BY G. N. RAMACHANDRAN,<sup>a,b</sup> R. CHANDRASEKHARAN<sup>a</sup> AND  
R. CHIDAMBARAM<sup>a,c</sup>

Received September 3, 1971

### ABSTRACT

A careful comparison of the distribution in the  $(R, \theta)$ -plane of all  $\text{NH} \cdots \text{O}$  hydrogen bonds with that for bonds between neutral  $\text{NH}$  and neutral  $\text{C}=\text{O}$  groups indicated that the latter has a larger mean  $R$  and a wider range of  $\theta$  and that the distribution was also broader than for the average case. Therefore, the potential function developed earlier for an average  $\text{NH} \cdots \text{O}$  hydrogen bond was modified to suit the peptide case. A three-parameter expression of the form  $V_{hb} = V_{\min} + p_1 \Delta^2 + q_1 e^{p_2 \Delta} \theta^2$ , with  $\Delta = R - R_{\min}$ , was found to be satisfactory. By comparing the theoretically expected distribution in  $R$  and  $\theta$  with observed data (although limited), the best values were found to be  $p_1 = 25$ ,  $p_2 = -2$  and  $q_1 = 1 \times 10^{-3}$ , with  $R_{\min} = 2.95 \text{ \AA}$  and  $V_{\min} = -4.5 \text{ kcal/mole}$ . The procedure for obtaining a smooth transition from  $V_{hb}$  to the non-bonded potential  $V_{nb}$  for large  $R$  and  $\theta$  is described, along with a flow chart useful for programming the formulae. Calculated values of  $\Delta H$ , the enthalpy of formation of the hydrogen bond, using this function are in reasonable agreement with observation. When the atoms involved in the hydrogen bond occur in a five-membered ring as in the sequence



a different formula for the potential function is needed, which is of the form  $V_{hb} = V_{\min} + p_1 \Delta^2 + q_1 x^2$  where  $x = \theta - 50^\circ$  for  $\theta \geq 50^\circ$ , with  $p_1 = 15$ ,  $q_1 = 0.002$ ,  $R_{\min} = 2.8 \text{ \AA}$  and  $V_{\min} = -2.5 \text{ kcal/mole}$ .

<sup>a</sup> Department of Biophysics, University of Chicago, Chicago, Illinois 60637, U.S.A.

<sup>b</sup> Indian Institute of Science, Bangalore-12, India (address to which reprint requests should be sent).

<sup>c</sup> Nuclear Physics Division, Bhabha Atomic Research Centre, Trombay, Bombay-85, India.

## INTRODUCTION

POTENTIAL functions giving the variation of the potential energy with the length  $R$  and the angle  $\theta$  of an  $\text{NH} \cdots \text{O}$  hydrogen bond for different values of  $R$  and  $\theta$  have been worked out previously in the authors' laboratory.<sup>1</sup> In this paper, the same procedure adopted in Ref. 1 is used for working out the potential function for the  $\text{NH} \cdots \text{O}$  hydrogen bond which occurs specifically in the backbone of peptide chains, namely between  $\text{NH}$  groups and  $\text{CO}$  groups. The study of the observed values of  $R$  and  $\theta$  in a large number of amino-acid and peptide structures made by Ramachandran and Ramakrishnan some time ago<sup>2</sup> showed that the average length of an  $\text{NH} \cdots \text{O}$  hydrogen bond between neutral  $\text{NH}$  and neutral  $\text{CO}$  groups was larger than the average length of all the hydrogen bonds of the type  $\text{NH} \cdots \text{O}$  put together. This indicates that the potential function for an average  $\text{NH} \cdots \text{O}$  hydrogen bond cannot be used for calculating the contribution from the hydrogen bonds occurring in peptide structures. The analysis was therefore made by taking into account those examples in the study of Ramachandran and Ramakrishnan which corresponded to the situation similar to that which occurs in a peptide unit. In fact, most of the examples studied were hydrogen bonds between  $\text{NH}$  and  $\text{CO}$  groups in peptide units in simple peptides. They were plotted in the  $(R, \theta)$ -plane and their distribution, normalised with respect to the total number of samples, was analysed. The data are summarised in Table I. This also contains for comparison the values calculated from the latest potential energy function. The notation follows that of Ref. 1.

## FORMULATION OF THE FUNCTION

The empirical form of the potential function discussed in Ref. 1 was considered for the purpose of this paper. This function may be represented by

$$V_{hb} = V_{\min} + P_1 \Delta^2 + P_2 \Delta^3 + e^{p_3 \Delta} (q_1 \theta^2 + q_2 \theta^3). \quad (1)$$

The second and third terms in this expression take care of the variation of the hydrogen bond energy with distance  $R$ . The last term accounts for the effect of nonlinearity of the bond. For a straight bond, *i.e.*,  $\theta = 0^\circ$ , the expression simplifies to

$$V_{hb} = V_{\min} + p_1 \Delta^2 + p_2 \Delta^3 \quad (2)$$

where  $\Delta = R - R_{\min}$ . The distribution of peptide hydrogen bonds in the  $(R, \theta)$ -plane (*see* Table I) clearly showed a maximum density between

$R = 2.95 \text{ \AA}$  and  $3.00 \text{ \AA}$ . However,  $R_{\min}$  was chosen to be  $2.95 \text{ \AA}$ , as the resultant theoretical distribution agreed well with observation. From the analogy of the previous study, the empirical function was expected to be valid in the range of  $\Delta$  between  $-0.2 \text{ \AA}$  and  $+0.2 \text{ \AA}$  for  $\theta = 0^\circ$ , or, in other words, for  $2.75 \leq R \leq 3.15 \text{ \AA}$ . Although very few examples were observed with  $R$  less than  $2.75 \text{ \AA}$ , the lower limit of  $R$  was taken to be  $2.6 \text{ \AA}$  in the present study so that the function would be valid in the range of  $R$  from  $2.6 \text{ \AA}$  to  $3.15 \text{ \AA}$ .

TABLE I

Experimental and theoretical probability distribution giving values of  $N'(R, \theta)$  in per cent.

$N'(\theta)^a$		R (Å)									
		2.75	2.80	2.85	2.90	2.95	3.00	3.05	3.10	3.15	
9.09 (19.49)	25	0.00 (0.00)	0.00 (0.00)	1.14 (1.90)	3.41 (3.03)	2.27 (3.88)	1.14 (4.00)	1.14 (3.30)	0.00 (2.20)	0.00 (1.18)	$\mathcal{D}(R, \theta)^e$ = 62.6
28.41 (24.28)	20	0.00 (0.64)	2.27 (1.28)	4.45 (2.43)	3.41 (3.71)	7.95 (4.57)	7.95 (4.54)	1.14 (3.64)	0.00 (2.35)	1.14 (1.23)	
34.09 (24.50)	15	3.41 (0.63)	2.27 (1.43)	3.41 (2.62)	3.41 (3.88)	9.09 (4.63)	7.95 (4.47)	2.27 (3.49)	1.14 (2.21)	1.14 (1.13)	
15.91 (20.23)	10	4.55 (0.58)	0.00 (1.27)	0.00 (2.27)	1.14 (3.27)	1.14 (3.83)	2.27 (3.62)	3.41 (2.78)	2.27 (1.73)	1.14 (0.87)	
12.50 (11.51)	5	1.14 (0.35)	0.00 (0.76)	1.14 (1.33)	3.41 (1.89)	2.27 (2.18)	1.14 (2.04)	1.14 (1.55)	1.14 (0.95)	1.14 (0.48)	
$\mathcal{D}(\theta)^c = 29.4$	$N'(R)^b$	9.09 (2.10)	4.55 (4.74)	10.23 (10.54)	14.77 (15.78)	22.73 (19.09)	20.45 (18.67)	9.09 (14.76)	4.55 (9.44)	4.55 (4.88)	$\mathcal{D}(R)^d = 24.8$

\* The experimental distribution  $N'_{\text{exp}}(R, \theta)$  is the number of examples normalized with respect to the total number. The theoretical distribution is

$$N'_{\text{th}}(R, \theta) = \frac{e^{-V(R, \theta)/RT}}{\sum_R \sum_\theta e^{-V(R, \theta)/RT}}$$

All experimental values are unbracketed and the theoretical values are within parentheses.

$$^a N'(\theta) = \sum_R N'(R, \theta)$$

$$^b N'(R) = \sum_\theta N'(R, \theta)$$

$$^c \mathcal{D}(\theta) = \sum_\theta |N'_{\text{exp}}(\theta) - N'_{\text{th}}(\theta)|$$

$$^d \mathcal{D}(R) = \sum_R |N'_{\text{exp}}(R) - N'_{\text{th}}(R)|$$

$$^e \mathcal{D}(R, \theta) = \sum_R \sum_\theta |N'_{\text{exp}}(R, \theta) - N'_{\text{th}}(R, \theta)|$$

For nonlinear hydrogen bonds (*i.e.*,  $\theta \neq 0^\circ$ ) with  $R$  in the above region, an upper limit for  $\theta$  was set up below which Eqn. (2) could be used straightaway. This has been formulated on an empirical basis (from an examination of the observed distribution) as follows:  $\theta_{\text{lim}} = 20^\circ$  for  $R = 2.6 \text{ \AA}$ ,  $25^\circ$  for  $R = 2.8 \text{ \AA}$  and  $30^\circ$  for  $R = 3.0 \text{ \AA}$  to  $3.15 \text{ \AA}$ . The region of validity of Eqn. (2) for  $V$  (denoted by  $V_{nb}$ ) is marked by the symbol HB in Fig. 1, and it is enclosed by solid lines.

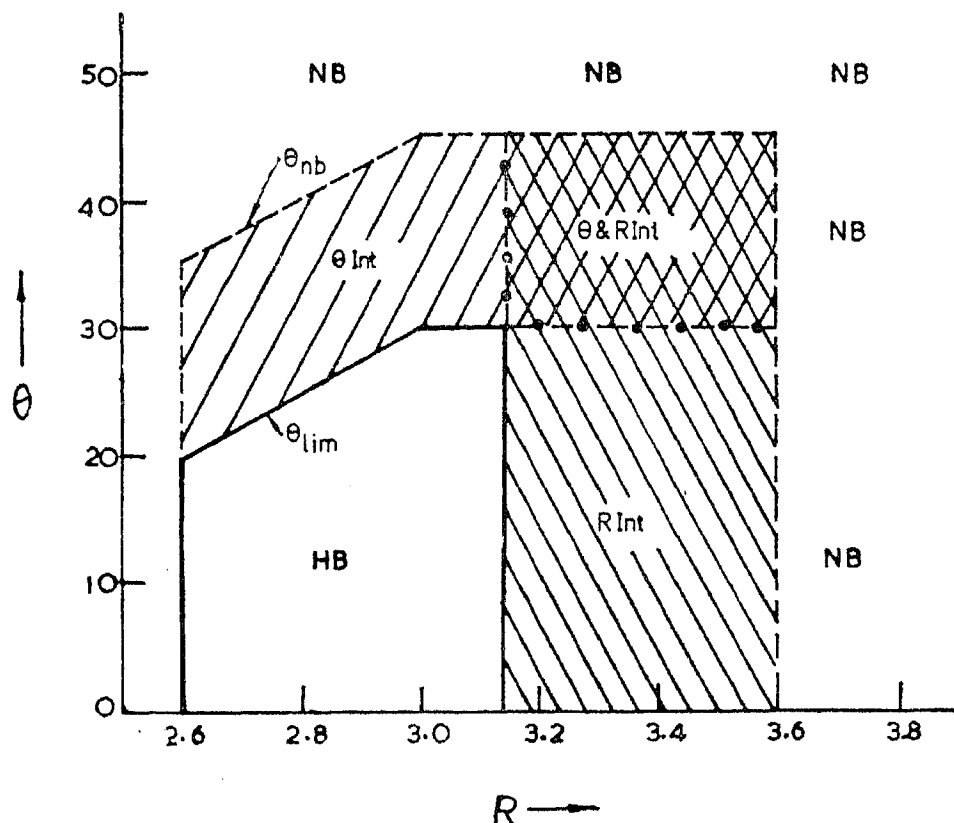


FIG. 1. Ranges of  $R$  and  $\theta$  in the  $(R, \theta)$ -plane showing the procedure for using the empirical hydrogen bonding function for the general peptide  $\text{NH}\cdots\text{O}$  bonds. The symbols are HB = hydrogen bonding function, NB = non-bonded function,  $\theta_{Int}$  = Theta-interpolation,  $R_{Int}$  = R-interpolation and  $\theta \& R_{Int}$  = First Theta interpolation, and then R-interpolation.

In the region of the  $(R, \theta)$ -plane outside that enclosed by the broken lines, marked by the symbol NB in Fig. 1, the non-bonded potential function  $V_{nb}$  is used. In the region in between these, which is shaded in Fig. 1, some type of interpolation is used, the nature of which is marked in the three subdivisions of this region in Fig. 1.

#### *Smoothing Between Hydrogen-Bonded and Non-bonded Potential Functions*

In the region in which R-interpolation alone is used, namely with  $\theta \leq 30^\circ$  and  $3.15 \text{ \AA} \leq R \leq 3.6 \text{ \AA}$ , the interpolation is made by fitting a cubic equa-

tion of the type  $V = A + BR + CR^2 + DR^3$  to the values of  $V$  at  $3.05 \text{ \AA}$  and  $3.15 \text{ \AA}$  (both hydrogen-bonded) and at  $3.6 \text{ \AA}$  and  $3.7 \text{ \AA}$  (both non-bonded). The details are given in the flow chart (Fig. 2) for programming the function for computational purposes. For the region in which  $\theta$ -interpolation is used, a similar cubic equation in  $\theta$ , namely  $V = A + B\theta + C\theta^2 + D\theta^3$  is employed, using values of  $\theta = \theta_{lim} - 5^\circ$ ,  $\theta_{lim}$ ,  $\theta_{nb}$  and  $\theta_{nb} + 5^\circ$  for obtaining the relevant constants  $A$ ,  $B$ ,  $C$  and  $D$  (where  $\theta_{nb} = \theta_{lim} + 15^\circ$ ). In the third interpolation region, denoted by  $\theta$  and  $R$  Int in Fig. 1, the procedure adopted is to evaluate  $V$  first by  $\theta$ -interpolation corresponding to values of  $R = 3.05 \text{ \AA}$  and  $3.15 \text{ \AA}$ , then to compute  $V$  at  $R = 3.6 \text{ \AA}$  and  $3.7 \text{ \AA}$  using the non-bonded function and finally to perform an  $R$ -interpolation using these (see flow chart).

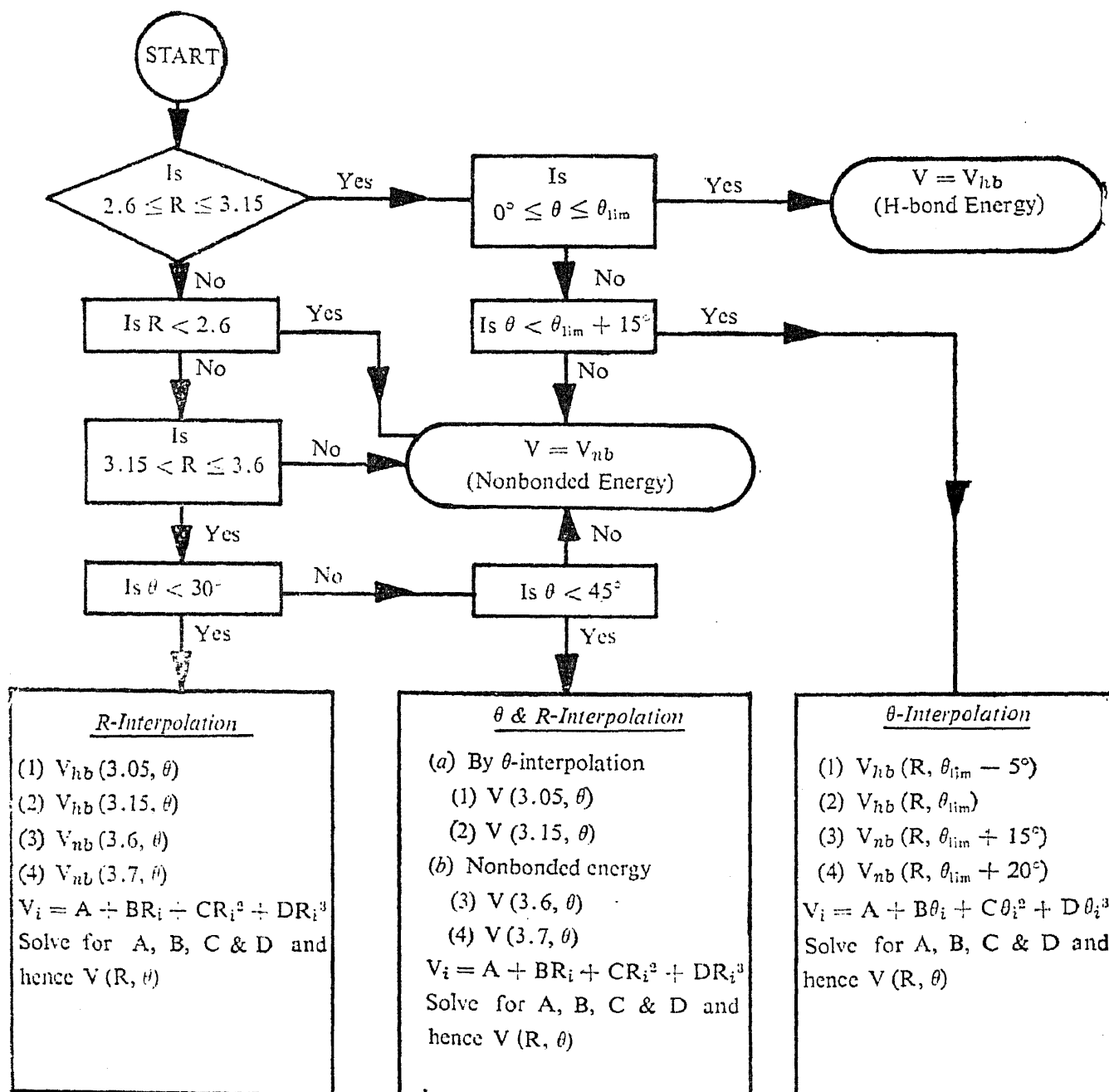


FIG. 2. Flow chart for using the hydrogen bond function.

## CHOICE OF THE PARAMETERS

The theoretical distribution was worked out for different sets of parameters  $p_1, p_2, p_3, q_1,$  and  $q_2$  in the empirical function given by Eqn. 1. It was found that as good an agreement with the observed distribution could be obtained with  $p_2$  and  $q_2$  made equal to zero, as when all the five constants were used. Hence only three parameters  $p_1, q_1$  and  $p_3$  were adjusted to get the best fit with the observed distribution. The following are the values finally chosen:

$$p_1 = 25, p_3 = -2 \text{ and } q_1 = 0.001; p_2 = q_2 = 0;$$

$$R_{\min} = 2.95 \text{ \AA}; V_{\min} = -4.5 \text{ kcal/mole.} \quad (3)$$

The theoretical distribution shown in Table I was calculated with this set of parameters. The variation of  $V$  with  $R$ , for different values of  $\theta$ , is shown in Fig. 3.

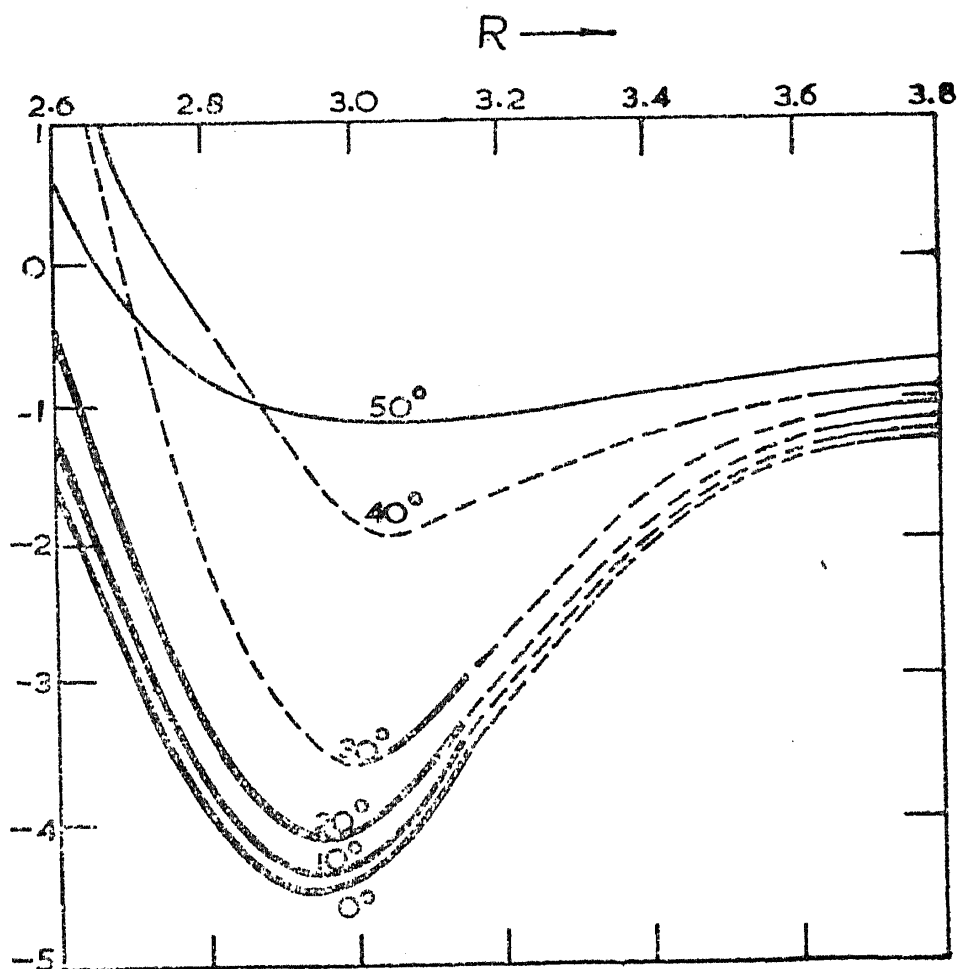


FIG. 3. Variation of hydrogen bonded potential energy  $V$  with  $R$  for different values of  $\theta$ . The thick line (—) represents  $V_{hb}$ , the thin line (—) represents  $V_{nb}$  and the broken line (....) shows the smooth transition from the HB region to the NB region.

Table II gives a few examples showing the effect of taking different values of the parameters  $p_1$ ,  $p_3$  and  $q_1$  in Eqn. (1) on the comparison between the observed distribution of hydrogen bonds in the  $(R, \theta)$ -plane and the theoretically calculated distribution. It may be seen that the values of the discrepancy indices  $\mathcal{D}$  are not very sensitive to small variations in the parameters  $p_1$ ,  $p_3$ , and  $q_1$ . This suggests that any one of the sets from Table II will prove to be a satisfactory choice for use in the empirical function. It should be mentioned that the total number of examples available for inclusion in the study was only 27, which explains the large magnitude of the best value of  $\mathcal{D}(R, \theta)$  obtained in this analysis, namely, about 60%.

TABLE II

*Variation of  $\mathcal{D}(\theta)$ ,  $\mathcal{D}(R)$  and  $\mathcal{D}(R, \theta)$  in per cent for typical values of  $p_1$ ,  $p_3$  and  $q_1$ .*

$p_1$	$p_3$	$q_1$	$\mathcal{D}(\theta)$	$\mathcal{D}(R)$	$\mathcal{D}(R, \theta)$
25	-3.0	0.0010	30.0	25.6	64.0
25	-3.5	0.0011	29.7	27.0	64.8
20	-2.0	0.0010	29.2	28.9	63.2
25	-2.0	0.0010	29.4	24.8	62.6
30	-2.0	0.0010	29.6	23.0	63.0
25	-2.5	0.0010	29.7	24.8	63.3
25	-1.5	0.0010	29.2	24.9	61.9
25	-2.0	0.0008	30.9	24.8	63.2
25	-2.0	0.0012	28.2	24.9	62.5
30	-3.5	0.0011	29.8	24.4	64.7
35	-3.0	0.0007	32.6	22.8	65.4

It may be seen from Table 1 that the experimental and the theoretical distributions of  $N'(R, \theta)$  significantly differ from each other at several grid points. Consequently, the discrepancy index  $\mathcal{D}(R, \theta)$ , which gives the sum of the magnitude of the individual deviations over all the grid points, is larger than 60%. However, if we compare the distributions of  $N'(R)$ , or  $N'(\theta)$ , along the row lines, the agreement between theory and experiment at each point is much better. The corresponding discrepancy indices  $\mathcal{D}(R$

and  $\mathcal{D}(\theta)$  are less than 30%, which should be considered satisfactory, since the total number of data points used is less than 30.

#### *Variation of the Energy $V$ with $\theta'$*

In Part I of this series,<sup>3</sup> the angle  $\theta'$  was defined to be that between the direction of a lone pair orbital at the acceptor atom and the line joining the proton to the acceptor. In general, there are two lone pair orbitals and two values of  $\theta'$ , namely  $\theta'_1$  and  $\theta'_2$ . A preliminary analysis of the distribution of  $\theta'$  for  $\text{NH} \cdots \text{O}$  bonds in general indicated that it has a maximum corresponding to values of either  $\theta'_1$  or  $\theta'_2$  close to  $0^\circ$ . This would indicate a dependence of  $V$  on  $\theta'$  of a lone pair orbital,  $V$  decreasing with increasing  $\theta'$ . However, it was learnt from Professor B. Pullman (personal communication) that quantum mechanical calculations indicated little or no variation of  $V$  with  $\theta'$  for the case of a single NH donor making a hydrogen bond with a  $\text{C}=\text{O}$  acceptor. Recently, we have learnt from Dr. C. Ramakrishnan (personal communication) that, if those examples in which a  $\text{C}=\text{O}$  group accepts a *single* hydrogen bond are examined, then the curve showing the variation of the number of hydrogen bonds with  $\theta'$  is flat near  $\theta' = 0^\circ$ , and has no maximum at  $\theta' = 0^\circ$ . This, together with the correlation between the variation of the energy and the distribution of the number of hydrogen bonds which was used in Ref. 1, indicates that experimental data support the quantum mechanical conclusion, namely that the change of  $V$  with  $\theta'$  is negligible.

On the other hand, Ramakrishnan finds that, when there are *two* hydrogen bonds directed towards the same oxygen atom, then the two  $\text{H} \cdots \text{O}$  directions are not far away from the directions of the two lone-pair orbitals. The effect in this case must be attributed to a number of causes including the repulsion between the atoms contained in the two donor groups. It was probably the occurrence of a large number of examples of such multiple bonds going to the same receptor that led to the histogram shown in Fig. 4 of Ref. 1, in which there is a maximum at  $\theta'' \approx 55^\circ$  (or  $\theta' \approx 0^\circ$ ). For reasons mentioned above, this is not significant in so far as the variation of energy with  $\theta'$  of a *single*  $\text{NH} \cdots \text{O}$  bond is concerned.

We therefore conclude that no term need be added in our formula for a possible variation of  $V_{hb}$  with  $\theta'$ .

#### *Calculation of $\Delta H$ , Enthalpy of Formation of the General Peptide $\text{NH} \cdots \text{O}$ Hydrogen Bond*

The formula for the variation of the energy  $V$  with the parameters  $R$  and  $\theta$  which has been obtained in this work may be used for finding the value



of  $\Delta H$  due to hydrogen bonding. Data of this type have been presented for various kinds of hydrogen bonds by Murthy and Rao<sup>4</sup> and for the case of N-methyl acetamide, which is a very good analogue of a peptide unit, two values are given namely,  $-3.3$  kcal/mole at  $36^\circ\text{C}$  obtained from NMR data and  $-4.7$  kcal/mole at  $26^\circ\text{C}$  obtained from IR data. The measurements in both cases were made in  $\text{CCl}_4$  solution.

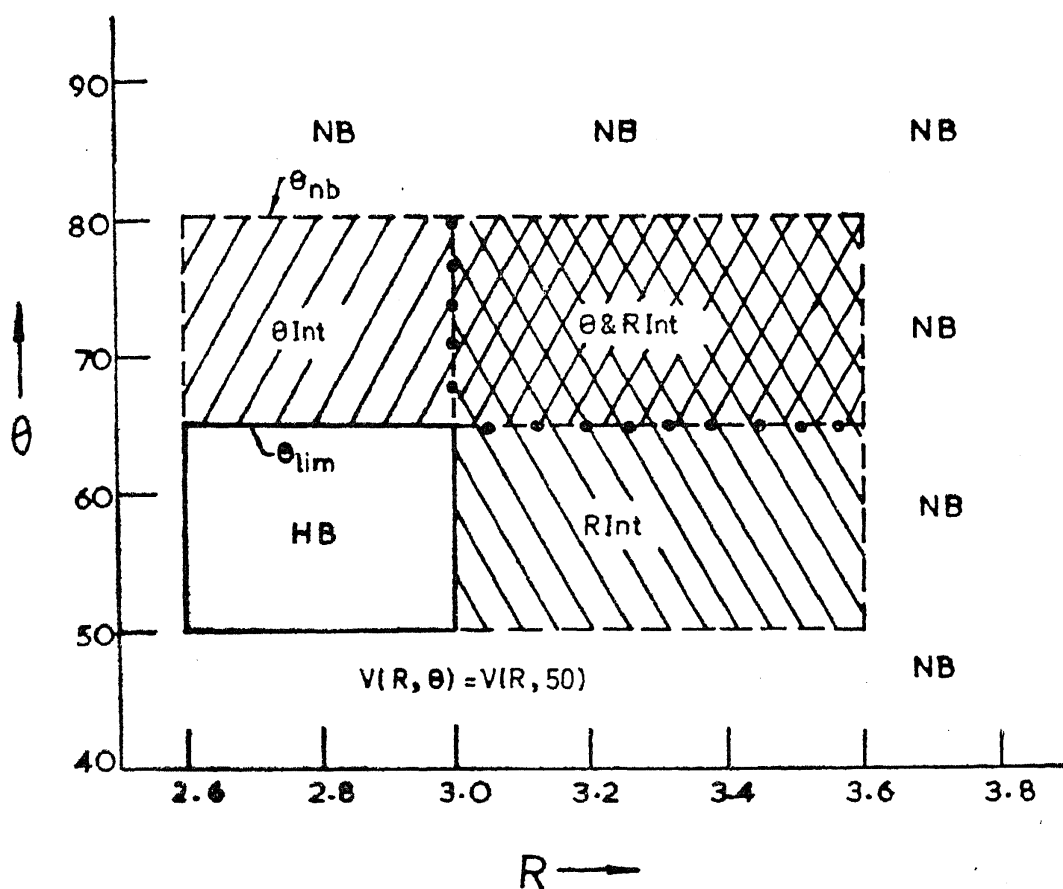


FIG. 4. Diagram similar to Fig. 1 indicating in outline the procedure for calculating  $V(R, \theta)$  for the five-membered ring peptide  $\text{NH}\cdots\text{O}$  hydrogen bond. Symbols as in Fig. 1.

For evaluating  $\Delta H$  from theory, it is necessary to define the range of values of  $R$  and  $\theta$  which have to be included in the hydrogen-bonded state. It is usual to consider the  $\text{NH}\cdots\text{O}$  hydrogen bond to have a length  $R$  less than  $3.2 \text{ \AA}$  and value of  $\theta$  less than  $35^\circ$ . Since  $V$  becomes very large for  $R < 2.6 \text{ \AA}$ , the ranges  $2.6 \text{ \AA} \leq R \leq 3.2 \text{ \AA}$  and  $0^\circ \leq \theta \leq 35^\circ$  were used for the definition of the hydrogen-bonded state. The value of  $\langle V_{hb} \rangle$  was obtained by using a formula analogous to that given in an earlier paper by two of us<sup>5</sup>, namely,

$$\langle V_{hb} \rangle = \frac{\iint V(R, \theta) \exp [-V(R, \theta)/RT] dRd\theta}{\iint \exp [-V(R, \theta)/RT] dRd\theta}, \quad (4)$$

where the term  $RT$  is the product of the gas constant and temperature and is equal to  $0.6$  kcal/mole for  $T = 300^\circ \text{K}$ ; the integration is over the ranges mentioned above and is approximated by a summation over a set of values  $(R_i, \theta_j)$  at intervals of  $0.02 \text{ \AA}$  for  $R$  and  $2.5^\circ$  for  $\theta$ .

In order to see the effect of the variation of the ranges included under the hydrogen-bonded state on the value of  $\langle V_{hb} \rangle$ , calculations were made taking the upper limit for  $R$  to be  $3.2 \text{ \AA}$ ,  $3.3 \text{ \AA}$  and  $3.4 \text{ \AA}$  and that for  $\theta$  to be  $35^\circ$ ,  $40^\circ$  and  $50^\circ$ , in different combinations. The results thus obtained are summarised in Table III (a).

TABLE III

Values of  $\langle V_{hb} \rangle$  in kcal/mole for the  $\text{NH} \cdots \text{O}$  hydrogen bond

Range of R ( $\text{\AA}$ )	Range of $\theta$		
	$0^\circ$ to $35^\circ$	$0^\circ$ to $40^\circ$	$0^\circ$ to $50^\circ$
	(a) $V_{\min} = -4.5$ kcal/mole		
2.6 to 3.2	-4.01	-4.00	-3.99
2.6 to 3.3	-3.97	-3.96	-3.95
2.6 to 3.4	-3.95	-3.94	-3.93
	(b) $V_{\min} = -4.0$ kcal/mole		
2.6 to 3.2	-3.50	-3.49	-3.48
2.6 to 3.3	-3.47	-3.45	-3.43
2.6 to 3.4	-3.44	-3.42	-3.40
	(c) $V_{\min} = -5.0$ kcal/mole		
2.6 to 3.2	-4.51	-4.50	-4.50
2.6 to 3.3	-4.48	-4.47	-4.47
2.6 to 3.4	-4.46	-4.46	-4.45
	(d) $V_{\min} = -6.0$ kcal/mole		
2.6 to 3.2	-5.52	-5.51	-5.51
2.6 to 3.3	-5.49	-5.49	-5.49
2.6 to 3.4	-5.48	-5.48	-5.48
	(e) $V_{\min} = -7.0$ kcal/mole		
2.6 to 3.2	-6.52	-6.52	-6.52
2.6 to 3.3	-6.50	-6.50	-6.50
2.6 to 3.4	-6.49	-6.49	-6.49

It will be noticed that  $\langle V_{hb} \rangle$  is insensitive to the upper limits of  $R$  and  $\theta$ . Since in a dilute solution, the solute molecules will be at fairly large distances from each other most of the time, it is reasonable to assume that

in the non-hydrogen-bonded state,  $\langle V_{nb} \rangle$  is practically zero. Then  $\Delta H$  becomes equal to  $\langle V_{hb} \rangle$  which is listed in Table III (a). The theoretical value of  $-4.0$  kcal/mole calculated from Eqn. (3) comes out to be intermediate between the NMR and IR experimental data mentioned above.

Since the experimental value is rather uncertain, the theoretical calculations were repeated using  $V_{\min} = -4, -5, -6,$  and  $-7$  kcal/mole but without making any other changes in the formula for  $V$  and the values of the other constants in Eqn. (3). It may be mentioned that in the method of adjustment used here, only the values of the constants  $p_1, p_3, q_1$  and  $R_{\min}$  are obtained from the distribution of the hydrogen bonds as observed in crystal structures, while the distribution is practically insensitive to the value of  $V_{\min}$ . However,  $\langle V_{nb} \rangle$ , and hence  $\Delta H$ , do depend on  $V_{\min}$ , as may be seen from Table III (b) to (e). It will be noticed that there is a corresponding increase in the numerical value of  $|\Delta H|$  by about the same order as the increase in  $|V_{\min}|$ .

One may thus express the relation between  $\Delta H$  and  $V_{\min}$  by the equation

$$-\Delta H = -V_{\min} - 0.5 \text{ kcal/mole}, \quad (5)$$

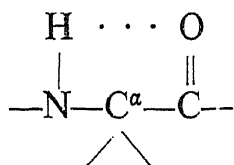
to a good degree of approximation, for the distribution we have deduced in this work. Therefore, if more accurate values of  $\Delta H$  are obtained from experiment in specific examples, then the value of  $V_{\min}$  for that example may be evaluated from this equation.

The function given by Eqn. (1), with the set of constants given in Eqn. (3), has been adopted for the study of various types of polypeptide chains and for hydrogen-bonded conformations in peptide units, and these studies are being reported elsewhere. In the succeeding part, the formula is applied to calculate the minimum energy conformation of the  $\alpha$ -helix of poly-L-alanine and it is found to give excellent agreement with the observed helical parameters and dihedral angles.

#### *Empirical Potential Function for the Five-Membered Ring NH ... O Bond*

In the course of our recent studies on the conformation of dipeptides in relation to NMR and IR data, our attention was drawn to an interesting conformation near  $\phi = 180^\circ$  and  $\psi = 180^\circ$ . (We follow here the latest nomenclature<sup>6</sup> for the description of the peptide chain.) This corresponds

to the fully extended conformation of the dipeptide unit with an NH  $\cdots$  O bond



involving atoms in a five-membered ring having a length  $R$  of 2.69 Å and a value of  $\theta$  equal to 51°. When the angles  $\phi$  and  $\psi$  are varied, both the length and the nonlinearity increase so that this particular bond cannot have values of  $R$  and  $\theta$  less than the above.

Both IR studies<sup>7</sup> and NMR data (Professor K. D. Kopple, personal communication) lead to evidence in support of this type of hydrogen-bonded conformation for compounds like acetyl-glycyl-N-methylamide in suitable solvents, *e.g.*, CCl<sub>4</sub>, CDCl<sub>3</sub>. In order to make conformational energy calculations for such model systems and to compare them with observation, another empirical potential function has been devised specifically for the five-membered ring NH  $\cdots$  O bond. This type of NH  $\cdots$  O bond would be expected to be different in energy from the usual type for the following reason. Since the best values of  $R$  and  $\theta$  for this bond are 2.69 Å and 51° respectively, one would obtain only the non-bonded energy  $V_{nb}$  if Eqn. (1) with the set of constants in Eqn. (3) is used. However, it is reasonable to expect that the energy for the five-membered system is lower than  $V_{nb}$  in order to explain the observed preference of the conformation in which such a 'bond' is found to occur. In the absence of any other information, it was assumed that the energy variation with  $R$  and  $\theta$  would not be very different in form from Eqn. (1). The following formula appears to be reasonable:

$$V = V_{\min} + p_1 \Delta^2 + q_1 x^2$$

where

$$x = \theta - 50^\circ \quad \text{for } \theta \geq 50^\circ. \quad (6)$$

The various constants were chosen to be

$$R_{\min} = 2.8 \text{ \AA}, \quad V_{\min} = -2.5 \text{ kcal/mole}$$

with

$$p_1 = 15 \quad \text{and} \quad q_1 = 0.002, \quad (7)$$

The value of  $\theta$  can never be lower than 50° for this type of bond occurring for standard peptide units<sup>8</sup> when the number of atoms involved in the ring

closure through hydrogen bond is only five, and hence the function has not been defined for  $\theta \leq 50^\circ$ . From  $50^\circ$  to  $65^\circ$ , the variation of energy with  $\theta$  is taken into account by the term  $q_1 x^2$ . If, for any reason,  $\theta$  becomes slightly less than  $50^\circ$ , we suggest the formula

$$V(R, \theta) = V(R, 50^\circ), \text{ for } \theta \leq 50^\circ. \quad (8)$$

Analogous to Fig. 1, the various regions in the  $(R, \theta)$ -plane, in which different formulae are used for calculating  $V$ , are shown in Fig. 4. The procedures for making the different interpolations are exactly similar to those mentioned earlier. The variation of energy  $V$  with  $R$ , for different values of  $\theta$  (starting from  $50^\circ$ ) is shown in Fig. 5.

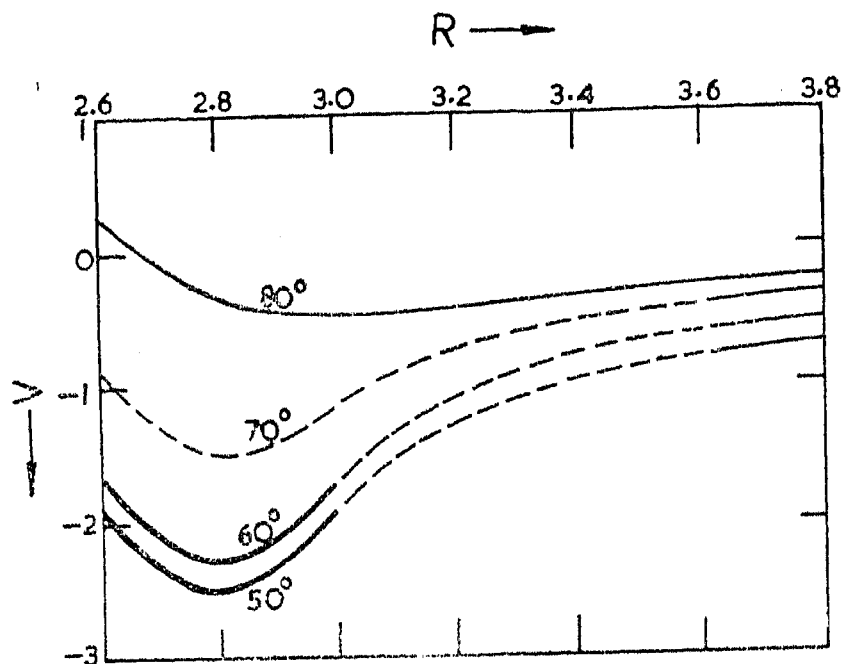
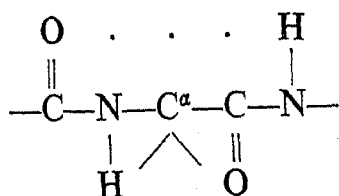


FIG. 5. Variation of  $V$  with  $R$  for different  $\theta$ 's for the five-membered ring peptide  $\text{NH} \cdots \text{O}$  hydrogen bonds. Symbols as in Fig. 3.

The five-membered ring hydrogen bond potential function [Eqns. (6), (7) and (8)] was found to give results in reasonable agreement with NMR and IR data for some of the dipeptides which were examined in continuation of the work reported in Ref. 5. The function given by Eqns. (1) and (3) for the general type of hydrogen bond is, however, expected to be valid for those involved in rings having 6, 7 and larger numbers of atoms. In particular, it has been used for the seven-membered ring hydrogen bond



in a dipeptide as part of this analysis and was found to fit the NMR and IR data reasonably well. These will be published as part of a separate study in due course.

#### ACKNOWLEDGMENT

We wish to acknowledge that the earlier work (Ref. 1) was carried out in collaboration with Mr. R. Balasubramanian. We thank Dr. C. Ramakrishnan for his assistance in obtaining the data which were used for this study. This work was supported by N.I.H. Grants AM11493 given to G. N. R. in Chicago and AM 15964 given to him in Bangalore.

#### REFERENCES

1. Balasubramanian, R.,  
Chidambaram, R. and  
Ramachandran, G. N. *Biochim. Biophys. Acta*, 1970, **221**, 196.
2. Ramachandran, G. N.  
and Ramakrishnan, C. In *Symposium on Fibrous Proteins, Australia, 1967*, W. G. Crewther, Ed., Butterworths, Australia, 1968, p. 71.
3. Chidambaram, R.  
Balasubramanian, R.  
and Ramachandran, G. N. *Biochim. Biophys. Acta*, 1970, **221**, 182.
4. Murthy, A. S. N. and Rao,  
C. N. R. In *Applied Spectroscopy Reviews*, E. G. Brame Ed., Marcel Dekker, New York, 1969, p. 69.
5. Ramachandran, G. N. and  
Chandrasekaran, R. *Biopolymers* (in press).
6. IUPAC-IUB Commission on Biochemical Nomenclature, *Biochemistry*, 1970, **9**, 3471.
7. Avignon, M., Huong, P. V.,  
Lascombe, J., Murraud,  
M., and Néel, J. *Biopolymers*, 1969, **8**, 69.
8. Ramachandran, G. N. and  
Sasisekharan, V. *Adv. Protein Chem.*, 1968, **23**, 283.