Acta Cryst. (1991). A47, 314-317

Repulsive Restraints for Hydrogen Bonding in Least-Squares Refinement of Protein Crystals. A Neutron Diffraction Study of Myoglobin Crystals

BY XIAODONG CHENG*

Center for Structural Biology, Department of Biology, Brookhaven National Laboratory, Upton, New York 11973, *USA, and Department of Physics, State University of New York at Stony Brook, Stony Brook, New York* 11794, *USA*

AND BENNO P. SCHOENBORN

Center for Structural Biology, Department of Biology, Brookhaven National Laboratory, Upton, New York 11973, *USA*

(*Received* 4 *June* 1990; *accepted* 21 *December* 1990)

Abstract

The purpose of this article is to describe stereochemical restraints on hydrogen bonding within proteins and their associated solvent which can be included in the refinement (*PROLSQ*) of X-ray or neutron structures of protein crystals. The parameters which define the geometry of hydrogen bonding, *i.e.* the correlation between distances and angles, are based on the results of an analysis of hydrogen bonding in crystal structures of myoglobin derivatives analyzed by neutron diffraction.

Nearly half of the atoms in a protein are H atoms. The locations and orientations of H atoms are important for understanding the geometry and function of the H atoms and hydrogen bonding. These atoms cannot be directly visualized by X-ray crystallography and in general their positions are inferred from the location and orientation of the various chemical groups to which they are bound. Geometric data based on stereochemical geometry are essential for this procedure. Neutron crystallography provides, however, a direct method to locate H atoms in proteins.

Neutron data from three different crystals (one carbonmonoxymyoglobin and two metmyoglobins) were analyzed in detail by including the effect of solvent. The analysis consisted of an iterative procedure of solvent evaluation in real space and protein refinement in reciprocal space. The analysis yielded detailed information on hydrogen bonds within the protein and between the protein and solvent and formed a suitable database for the assessment of the hydrogen-bonding properties in this helical protein.

The refinement procedure required the observed structure factors (F_{obs}) be resolved into two components: protein (F_{pro}) and solvent (F_{sol}) (Schoenborn, 1988; Cheng & Schoenborn, 1990):

$$F_{obs} = |F_{pro} + F_{sol}|.$$

The contribution of solvent (F_{sol}) to the low-order F_{obs} was evaluated by dividing the solvent volume into n shells extending outward from the surface of the protein. Each shell was given two parameters (one for solvent scattering density and one for disorder) and treated independently in the solvent evaluation. The two solvent parameters for each of the n shells were then adjusted to minimize

$$\delta F = (F_{obs} - F_{pro}) - F_{sol}.$$

Based on the solvent evaluation, the modified structure-factor terms (F_{om}) were used to refine the protein structure in a reciprocal-space least-squares refinement (*PROLSQ*) by Konnert (1976), Hendrickson & Konnert (1981) and Hendrickson (1985):

$$F_{om} = |F_{obs} - F_{sol}|.$$

The use of all data including the low-order diffraction terms (in which the contributions from solvent and protein are of the same order of magnitude) in refinement is advantageous, since the low-order data are generally the most intense with the best counting statistics and help condition the least-squares refinement.

After completion of several cycles of protein refinement, the solvent evaluation procedure was repeated. This was followed by a Fourier synthesis and calculation of another set of modified F_{om} which was used in the next round of protein refinement. During the process, about 90 water molecules, as well as a few ions, were added to the protein model. All the water molecules were found to be bound to polar or charged groups (Cheng & Schoenborn, 1990).

After eight rounds (101 cycles of refinement), the geometry of *intra-protein* hydrogen bonding was surveyed based on the refined coordinates of CO-

* Present address: Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA.

Table 1. *Summary of observed hydrogen bonding in three refined structures*

	N···O (Å)	H···O (Å)	N–H···O (°)	N···O=C (°)
Main chain	3·12±0·17	2·25±0·18	151±12	146±15
Side chain	3·02±0·31	2·18±0·35	146±16	132±19

myoglobin. A similar procedure was followed for two metmyoglobin structures with different D_2O concentrations.

The intra-protein hydrogen bonding derived from these analyses is classified into main-chain bonds and side-chain-related bonds. The geometry of each hydrogen bond is described by bond distances and bond angles. The N–H···O=C hydrogen bonds formed by the peptide group of the main chains are described by N···O distance, H···O distance and N–H···O angle. The criterion used to define main-chain hydrogen bonds was based on helix assignment with the N···O distance less than 3·5 Å; for side-chain-related hydrogen bonds and main-chain hydrogen bonds in nonhelical regions, the criterion was that the distance between non-H atoms (N···O or O···O) was less than 3·5 Å, and the angle at the H (D) atom (N–H···O or O–H···O) was greater than 100°. Table 1 gives the mean values for all intra-protein hydrogen bonds found in three myoglobin derivatives and the root-mean-square deviations in the bond distance and angle. Figs. 1 and 2 show the dependence of the angles at H atoms on bond lengths for all uncharged N–H···O and O–H···O bonds. The three filled symbols indicate the hydrogen bonds formed between main-chain atoms from the three independent structures. The open circles indicate the wider distribution of side-chain-related hydrogen bonds in three structures. There is a significant scatter in the hydrogen-bond geometric data, but a limiting line can be drawn representing minimum allowed values. Below this line configurations are not observed. In the absence of a distinct algorithm for the nature of hydrogen-bonding interactions, experience will show the best placement of these lines that describe the minimum values of possible hydrogen bonding. The line immediately suggests an excluded region. These observations of the hydrogen-bonding configurations between the protein atoms are a result of refinement against the F_{obs} including the solvent effect and do not include any stereochemical hydrogen-bonding restraints. They are only subject to the stereochemical restraints described by Konnert (1976), Hendrickson & Konnert (1981) and Hendrickson (1985). Figs. 3 to 5 plot N···O distance, H···O distance and N–H···O angle against N···O=C angle, respectively, and show limiting lines that represent the extreme limits of acceptable hydrogen-bond geometry. This can be interpreted in terms of the limits of the repulsions of the

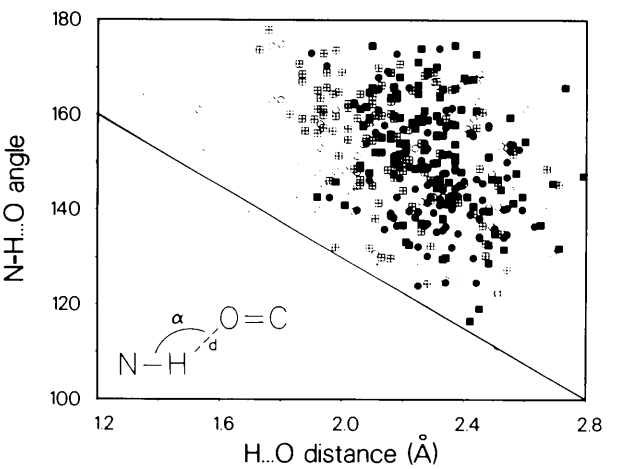

Fig. 2. Plot of the dependence of hydrogen-bonded angle at the H atom (N–H···O)(°) on the H···O distance.

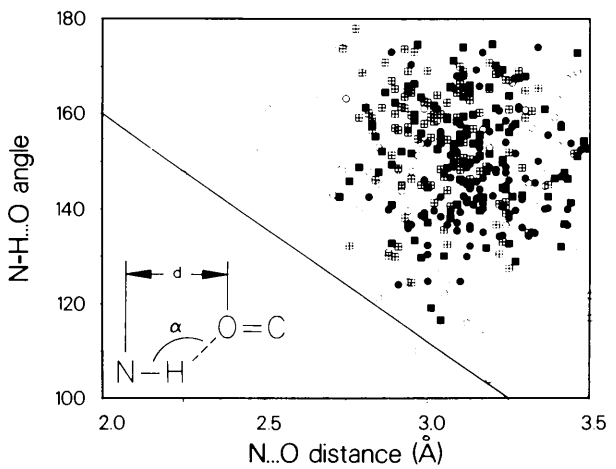

Fig. 1. Plot of the dependence of the hydrogen-bonded angle at the H atom (N–H···O)(°) on the N···O distance.

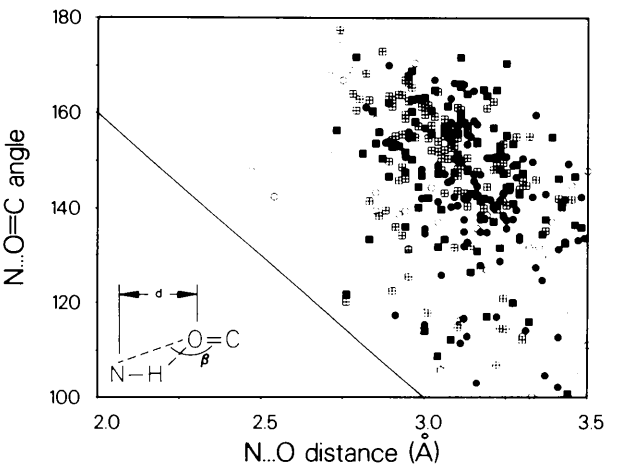

Fig. 3. Plot of the dependence of the N···O=C angle at the O atom (°) on the N···O distance.

hydrogen-bonded neighbor atoms. These 'repulsive restraints' are important and useful additional stereochemical conditions in restrained least-squares refinement of both X-ray and neutron structures.

These observations can be introduced as further restraints into the refinement procedures using the *PROLSQ* algorithm. The *PROLSQ* algorithm (Hendrickson, 1985) treats the repulsive interactions as simple interatomic potential-energy functions $U(d)$ which can be approximated to the formulation

$$U(d) - U(d_{\min})|_{d<d_{\min}} \simeq (d - d_{\min})^4/\sigma^4$$

for repulsive contacts ($d < d_{\min}$) only. $d_{\min}$ locates the minimum energy and σ is the associated standard deviation. In principle, σ varies with the type of contacts and relies on the accuracy of the coordinates. Experience shows that the values are quite uniform ($\sigma = 0{\cdot}65$) in restraints on water-related hydrogen bonds (Cheng & Schoenborn, 1990). In practice, $\sigma = 0{\cdot}60$–$0{\cdot}80$ is a good starting point. The value of $d_{\min}$ depends on the atoms in contact and the type of contact involved. The observed hydrogen-bonding pattern can be expressed in a similar fashion with $d_{\min}$ representing the 'exclusion line'. Hydrogen-bonding configurations above these lines are not restrained. Five distinct groups are identified and the following five formulas are based on the limiting lines provided in Figs. 1 to 5. Because of the uncertainty in the coordinates for the atoms in the three structures surveyed, each point in Figs. 1 to 5 has a maximum uncertainty of 0·2 Å in bond distance for the current *PROLSQ* refinement, with a resulting uncertainty of 10° in bond angle. For Figs. 1–5, the three filled symbols indicate the hydrogen bonds formed between main-chain atoms and the open circles those of side-chain-related hydrogen bonds from three independent structures.

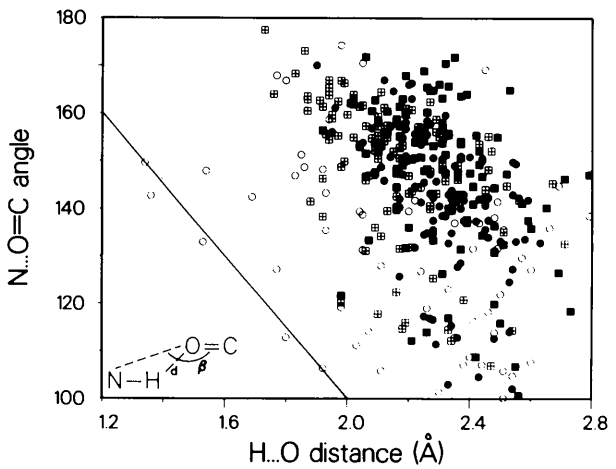

Fig. 4. Plot of the dependence of the N···O=C angle at the O atom (°) on the H···O distance.

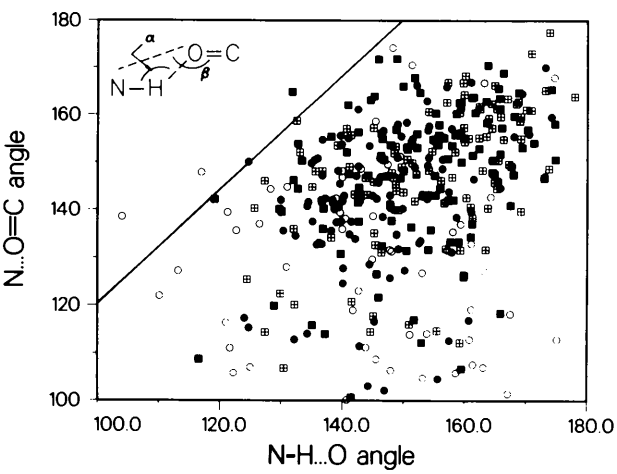

Fig. 5. Plot of the dependence of the N···O=C angle at the O atom (°) on the N–H···O angle at the H atom (°).

1. N···O repulsive of hydrogen-bonded distance against N–H···O angle (Fig. 1):

$$d_{\min}(\text{Å})|_{100\le\alpha\le180°} \simeq -\alpha/48 + 5{\cdot}3,$$

where α is N–H···O angle (°).

2. H···O repulsive of hydrogen-bonded distance against N–H···O angle (Fig. 2):

$$d_{\min}(\text{Å})|_{100\le\alpha\le180°} \simeq -2\alpha/75 + 5{\cdot}5,$$

where α is N–H···O angle (°).

3. N···O repulsive of hydrogen-bonded distance against N···O=C angle (Fig. 3):

$$d_{\min}(\text{Å})|_{100\le\beta\le180°} \simeq -\beta/60 + 4{\cdot}7,$$

where β is N···O=C angle (°).

4. H···O repulsive of hydrogen-bonded distance against N···O=C angle (Fig. 4):

$$d_{\min}(\text{Å})|_{100\le\beta\le180°} \simeq -\beta/75 + 3{\cdot}3,$$

where β is N···O=C angle (°).

5. The relationship between α angle (N–H···O) and β angle (N···O=C) (Fig. 5) can be represented by

$$\beta_{\max}(°)|_{100\le\alpha\le180°,\ 100\le\beta\le180°} \simeq 1{\cdot}2\alpha,$$

where α is N–H···O angle (°) and β is N···O=C angle (°).

While these observations are based on data derived from three derivatives of only one helical protein, the results show that the dependence of hydrogen-bond length on angle parallels the observations for water structures in ices and small crystalline hydrates (Savage & Finney, 1986). Therefore, the placement of the limiting lines that define the restraint algorithm should be chosen to reflect these geometrical characteristics shared by proteins, ices and small hydrates.

The five restraints discussed above are easily introduced into refinement procedures and suitable additions to the *PROLSQ* algorithm to prevent unacceptable hydrogen bonding and apply to neutron as well as X-ray data.

The original CO-myoglobin data was collected in collaboration with A. C. Nunes and J. C. Norvell. The original coordinates of myoglobin were provided by J. C. Kendrew and H. C. Watson. We are grateful to Drs A. J. Wonacott and R. Knott for helpful discussions and critically reading the manuscript. The work was supported by the Office of Health and Environmental Research and the calculations were performed under the Supercomputing Program of the United States Department of Energy.

References

CHENG, X. & SCHOENBORN, B. P. (1990). *Acta Cryst.* B**46**, 195-208.
HENDRICKSON, W. A. (1985). *Meth. Enzymol.* **115**, 252-270.
HENDRICKSON, W. A. & KONNERT, J. H. (1981). In *Biomolecular Structure, Function, Conformation and Evolution*, Vol. 1, edited by R. SRINIVASAN, pp. 43-57. Oxford: Pergamon.
KONNERT, J. H. (1976). *Acta Cryst.* A**32**, 614-617.
SAVAGE, H. F. J. & FINNEY, J. L. (1986). *Nature (London)*, **322**, 717-720.
SCHOENBORN, B. P. (1988). *J. Mol. Biol.* **201**, 741-749.

Acta Cryst. (1991). A**47**, 317-327

Theoretical and Experimental Studies of Electron Resonance Effects in Reflection High-Energy Electron Diffraction

BY P. LU,* J. LIU AND J. M. COWLEY

Department of Physics, Arizona State University, Tempe, AZ 85287, *USA*

(*Received* 3 *April* 1990; *accepted* 11 *September* 1990)

Abstract

The electron resonance effect for 100 keV electrons incident on the surface of GaAs (110) in RHEED is studied by theoretical simulations using multislice theory and experimental observations. The exact resonance conditions, effective resonance region, effective penetration depth of electrons at or near the resonance condition and scattering processes involved for the resonance effect are investigated. It is found that the intensity of the 440 specularly reflected beam is mainly due to direct reflection from the surface atomic layer and beam enhancement due to surface channeling effects under the resonance conditions seems to be insignificant. For the 880 specularly reflected beam most of the electron intensity penetrates the surface and is diffracted by the crystal. The resonance condition for the 880 specular beam is satisfied when the transmitted beam excites a strong surface wave which propagates in the direction parallel or nearly parallel to the surface and is localized in the surface region by the surface potential barrier; double diffraction from the surface beam to the specular beam then enhances the total intensity in the specular beam. The exact resonance condition for the 880 beam is found to be at the glancing angle of 35·7 mrad and the azimuth angle of 29·7 mrad. A strong wave field is localized in the surface region at the resonance condition with an effective electron penetration depth of ~5 Å, which increases to ~35 Å on going to the nonresonance conditions. The effective resonance region for the 880 spot is ~2 mrad about the azimuth angle and ~1 mrad about the glancing angle.

1. Introduction

The technique of reflection electron microscopy (REM), in which a high-energy electron beam (20-200 keV) is incident on a surface at a glancing angle of 1-2° and forms a reflection high-energy electron diffraction (RHEED) pattern, is now becoming one of the most important microscope methods for investigating surface structures (Yagi, 1987). The REM image which is formed by selecting one of the reflected beams provides structural information about the surface (Yagi, 1987) and electron energy-loss spectroscopy (EELS) with the reflected beams can provide complementary chemical information (Krivanek, Tanishiro, Takayanagi & Yagi, 1983; Howie, 1983; Wang & Cowley, 1988). It has been found that the

* Present address: Department of Mechanics & Materials Science, Rutgers, Piscataway, NJ 08855-0909, USA.

0108-7673/91/040317-11$03.00