pl5 The Probabilistic Theory of the Structure Invariants Assuming a Substructure to be Known.

H. Hauptman and H. Xu, *Hauptman-Woodward Medical Research Institute, Inc., 73 High St., Buffalo,NY 14203 USA.*

Keywords: protein direct methods, SIR direct methods, SIRAS direct methods.

The mathematical machinery needed to develop the probabilistic theory of the structure invariants (triplets) for an arbitrary pair of isomorphous structures (F, G) in the case that only normal diffraction intensities are available was initiated by Hauptman[1]. A similar theory was developed when single wavelength anomalous data are available for a single structure[2]. In this paper the earlier techniques are combined to yield the probabilistic theory of the two phase structure invariants (doublets) in the case that single wavelength anomalous scattering data are available for each member of the isomorphous pair (F, G); and the complex-valued atomic scattering factors f and g for the structures F and G, respectively, are assumed to be known.

For fixed reciprocal lattice vector H denote by $|E_H| = R$, $|E_{\bar{H}}| = \bar{R}$, $|G_H| = S$, $|G_{\bar{H}}| = \bar{S}$ the magnitudes, and by $\varphi_H = \varphi$, $\varphi_{\bar{H}} = \bar{\varphi}$, $\psi_H = \psi$, $\psi_{\bar{H}} = \bar{\psi}$ the phases, of the normalized structure factors $E = |E|\exp(i\varphi)$, $G = |G|\exp(i\psi)$ for the structures F and G, respectively. Note that, owing to the breakdown of Friedel's Law , in general, $R \,?\, \bar{R}$, $S \,?\, \bar{S}$; $\bar{\varphi} \,?\, -\varphi$, $\bar{\psi} \,?\, -\psi$.

Making the usual assumption that the atomic coordinates are the primitive random variables, assumed to be uniformly and independently distributed, one uses the machinery developed in 1982 to derive first the conditional joint probability distribution of the six doublets $\varphi + \bar{\varphi}$, $\varphi - \psi$, $\bar{\varphi} - \bar{\psi}$, $\varphi + \bar{\psi}$, $\bar{\varphi} + \psi$; $\psi + \bar{\psi}$, assuming the four magnitudes $R, \bar{R}, S, \bar{S}$ to be known. This distribution leads in the usual way (See the 1982 papers) to the conditional probability distribution of the single phase ψ, given the four magnitudes $R, \bar{R}, S, \bar{S}$, and the two phases $\varphi, \bar{\varphi}$:

$$P\left(\psi / R; \bar{R}; S; \bar{S}; \varphi, \bar{\varphi}\right) = \frac{1}{K}\exp\left\{B\cos(\psi - \alpha)\right\} \qquad (1)$$

where the parameters K, B(>0), and α are expressible in terms of the four magnitudes $R, \bar{R}, S, \bar{S}$ and the two phases $\varphi, \bar{\varphi}$, as well as the complex-valued atomic scattering factors f and g, presumed to be known.

In the special case that the structure F is a known substructure of G then F and G are isomorphous (Some of the f's will then be zero), R, $\bar{R}$, φ, and $\bar{\varphi}$ can be calculated for any wavelength from the known substructure F, and S, $\bar{S}$ and are determined from experiment at a single wavelength. Eq. (1) then yields the estimate α for the phase ψ with reliability depending on the magnitude of B. Thus this theory leads to a method for going from a substructure (F), presumed to be known, to the full structure (G), when single wavelength anomalous diffraction data are available for G.

This research is supported by a grant from the National Institutes of Health, No. GM46733.

[1] Hauptman HA. "On Integrating the Techniques of Direct Methods and Isomorphous Replacement. I. The Theoretical Basis", Acta Cryst. (1982), A38:289-294.

[2] Hauptman HA. "On Integrating the Techniques of Direct Methods with Anomalous Dispersion. I. The Theoretical Basis", Acta Cryst. (1982), A38: 632-641.

Notes