Chapter 6. How reliable are B-factor data?

Since they are rarely discussed at the same level of significance given to coordinates, B-factors seem to viewed with considerable suspicion by protein crystallographers. To the non-crystallographer this is puzzling and we have not found any convincing explanations nor, in view of the quantitative success of intra-palindrome comparisons, do we expect to. The basis for suspicion almost always offered is lattice disorder but the B factors show that disorder if negligible A more legitimate explanation is variations among substates and particularly the matrix contraction transition. The degree of the latter varies with the protein activity coefficient which in turn reflects the solvent or solvent composition hydration, temperature and ligation state. Residue exchange and chemical modification can produce significant changes themselves often of much interest. In some cases the mean B values for two otherwise identical proteins in the same unit cell are quite different. Competitive inhibitor binding and temperature variation can produce reductions in matrix B factors by as much as 60%. That is close to the minimum available free volume of matrices estimated from low-temperature diffraction data in which vibrational excitation is largely frozen out

The accuracy of protein x-ray diffraction data is not statistically established because there are no reference sets of absolute coordinates but despite good precision the errors in coordinates are much larger than those in the B factors. Only with Fourier–difference methods do the values of coordinate differences approach those that can be estimated from B factors. This is because the errors tend to be percentages of the ranges and the B factors have a much smaller range. The B range obtained in modern high-resolution x-ray studies is just that required to measure the relative differences in comparing sets of B factors. Despite the absence of residue similarity in their sequences evolution has been very successful in matching the two branches of the low-B palindromic

patterns so that the best comparisons to estimate both precision and accuracy in protein construction are those between the two branches of the palindromes in single-protein studies. Using the isotropic approximation B values in the normal range of 5 to 35 cover a range of displacement of the mean center of an atom from its ideal lattice position from 0.07 $Å^2$ to 0.5 $Å^2$ equivalent to mean displacement-radius values of 0.27 - 0.71 Å. The corresponding mean free volumes are 0.085 to 1.55 $Å^3$ giving a total molecule free volume in the expanded state of about 4%. On complete contraction to the glasslike state the atom free volumes decrease nearly in proportion to their B values in the liquidlike state producing hard matrices with uniform B values not much larger than those of knot atoms. (Fig. $).

For the streptococcus G protein the palindromic pairs of knot atoms in this high-precision study are matched in radius to about 0.05Å. Any constraints applied to B factors in data analysis will not change matching although the coordinate scale errors may be changed. Nevertheless 0.05Å computed from the mean-square displacement values is an important number as an estimate from the knot B factors of the streptococcus G protein. It can be too low for a general estimate insofar as it reflects some cooperative stiffening of the fused knots. However, so far as we now know, mutations in the two arms evolved independently. That being so, the evolutionary process for enzymes converges toward C-2 symmetry in knot assemblies not because of any intrinsic feature of DNA construction but rather because of natural selection among possible products. The estimate sets an upper limit for the accuracy and a lower limit for the precision in protein construction,. It also demonstrates the precision and accuracy in B determinations possible with modern x-ray-diffraction methods. Since there is rarely much palindromy in the residue sequences, each atom of a palindromic pair is likely to be from a different kind of residue. This may not be surprising since the domains are constructed back-to-back and, as noted above,

none of the know ways to construct and express DNA appear have the intrinsic ability to force palindromy. That important observation establishes that the palindromes are neither an intrinsic feature of DNA transcription nor a result of post-translational modifications.

In general the PDB data indicate that matrices rarely have as well developed B-factor palindrome patterns as their knots. This may be somewhat misleading since the branches of a palindrome pattern appear to improve in accuracy and precision with increased x-ray resolution and refinement. The quality of x-ray data is rapidly improving so that the assumption of a mean square displacement parameter still necessary in most analysis, can be replaced by accurate determination of the three axes of the scattering ellipsoids but we have not found it easy to extract useful molecular information from such data.

In Fig.@ H@$ the slowest exchanging atoms, a total of 64, 16 per group, have B values less than about 6.3. The few proteins for which 1Å resolution have been reported set the current goal. It is also of considerable utility that as shown by PDB data for the G protein, the free-volume map for the entire protein can be displayed with some accuracy using atom B factors. Such maps give detailed descriptions of the construction and the conformational adjustments which support physiological function.

At this time the use of B factors for comparisons among x-ray-diffraction studies is limited by uncertainties in just what experimental B values actually measure in any specific applicaition. How much is actual free-volume, how much is experimental artifact Better ways to estimate these quantities are required since the future of the use of protein diffraction data depends to a large degree of the use of B factors. Coordinates even from high-precision studies are not likely to provide more than pictorial information inadequate to quantitative applications such as accurate potential-energy functions. The B factors, already extensively

tabulated , open the door to quantitative study of proteins but this use is not reliable until they can be related in an unambiguous way to molecular changes. This weakness is particularly important for better understanding of the matrix expansion-contraction process. Singh, Huber and Bone, in one of the few x-ray studies constructed to test the use of B values in this connection, found that the process could be followed over temperature range by its phase behavior. The standard enthalpy change was only weakly dependent on the amount of methanol used to prevent freezing but the B-value scale was shifted by methanol concentration and temperature. The expanded and contracted species are very different, one dynamic the other static, but few x-ray-diffraction experiments have been constructed to attempt to detect these differences essential in any attempt to understand the conformational basis of physiological function. The pepsini+pepstatin study (Fig. @) is a notable exception but not in itself sufficient to provide essential molecular detail.