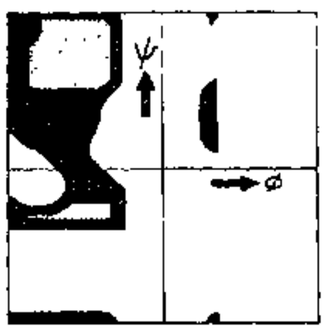

Special issue:
The Ramachandran map and
protein and peptide structures

Some uses of the Ramachandran (ϕ,ψ) diagram in the structural analysis of lysozymes

L. H. Weaver, D. E. Tronrud, H. Nicholson and B. W. Matthews*

Institute of Molecular Biology, Howard Hughes Medical Institute and Department of Physics, University of Oregon, Eugene, Oregon 97403, USA

Two illustrations of the utility of the Ramachandran diagram in the analysis of protein structures are provided. In the first example, two mutants of T4 lysozyme are examined, in which two non-glycine residues, Asn 55 and Lys 124, that have left-handed helical conformations in the wild-type protein are replaced with glycines. In the Asn 55 → Gly case there is a change of about 20° in both ϕ and ψ in the mutant. In the Lys 124 → Gly case there is essentially no change in backbone conformation. In both cases, substitution had little effect on protein stability. A second example describes the use of the Ramachandran diagram in structure evaluation. An abortive model for goose egg-white lysozyme with a crystallographic residual of 28% at 2.8 Å resolution shows a significantly greater proportion of residues in Ramachandran disallowed regions than a more acceptable model with R = 15% at 1.6 Å resolution.

THE use of the 'Ramachandran diagram'[1] has become commonplace in the analysis of the structures of proteins and peptides.

In this article we provide two illustrations of the usefulness of these diagrams. The first example is taken from studies of the structure and stability of mutant T4 lysozymes[2,3]. The second example shows how the Ramachandran diagram can be used to check the reliability of a proposed X-ray (or NMR) protein structure.

Analysis of left-handed helical residues in T4 lysozyme

As is well known, non-glycine residues in folded proteins usually have (ϕ,ψ) values that correspond to the 'right-handed-helical' or to the 'extended' regions of the Ramachandran diagram. The 'left-handed-helical' configuration is 'allowed', but is observed only rarely. Glycines, on the other hand, occur with approximately equal frequencies on the left-hand and right-hand sides of the Ramachandran diagram. The difference, of course, is due to the fact that glycine lacks a β-carbon.

Notwithstanding these well-known facts, the exact difference between the backbone configurational energy of a glycine and a non-glycine residue in the left-handed helical conformation is unclear.

If the difference in energy between a left-handed glycine and non-glycine is large, then a significant increase in the free energy of stabilization of a protein could be realized by judiciously replacing any 'left-handed' residues with glycines. The anticipated stabilization would be offset if removal of the given residue resulted in the loss of favorable interactions from the folded structure of the protein.

The structure of phage T4 lysozyme contains two non-glycine residues, Asn 55 and Lys 124, that have left-handed helical conformations[4]. The energetics of each of these residues was investigated by replacement with a glycine[5].

Asparagine 55 → glycine. Asn 55 is located within an exposed, irregular, relatively mobile segment that connects α-helix 39–50 with α-helix 60–80. The polypeptide backbone at and adjacent to Asn 55 forms a network of solvent-mediated interactions as well as an H-bond to the side chain of Asp 47. Other H-bonding interactions are shown in Figure 1,*a*.

The (ϕ,ψ) values of Asn 55 in wild-type lysozyme are shown in Figure 2. Replacement of Asn 55 with glycine results in a change of about 20° in both ϕ and ψ. Because the side chain of Asn 55 is not obviously constrained it suggests that the change in (ϕ, ψ) occurs because the preferred energy minimum for glycine is not the same as that for asparagine.

A survey of accurately determined protein struc-

*For correspondence.

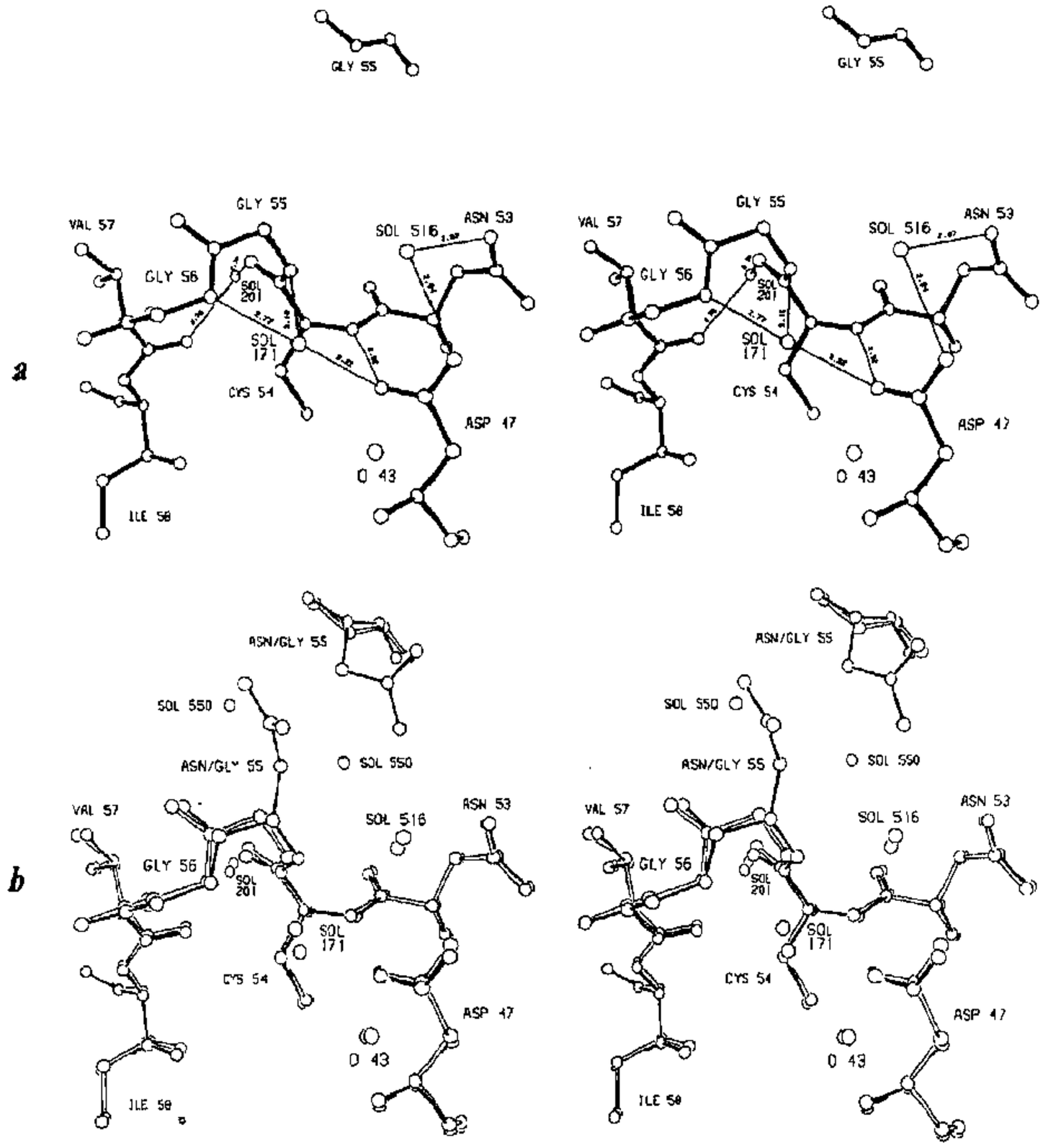

Figure 1. *a*, Structure of the mutant lysozyme N55G in the vicinity of residue 55. The isolated chain segment (top) labelled Gly 55 is part of a neighbouring molecule in the crystal. *b*, Superposition of N55G mutant lysozyme (open bonds) on wild-type lysozyme. The two structures are very similar, but rotation of the residue 54–55 and 55–56 peptide groups can be seen in the figure. (From ref. 5.)

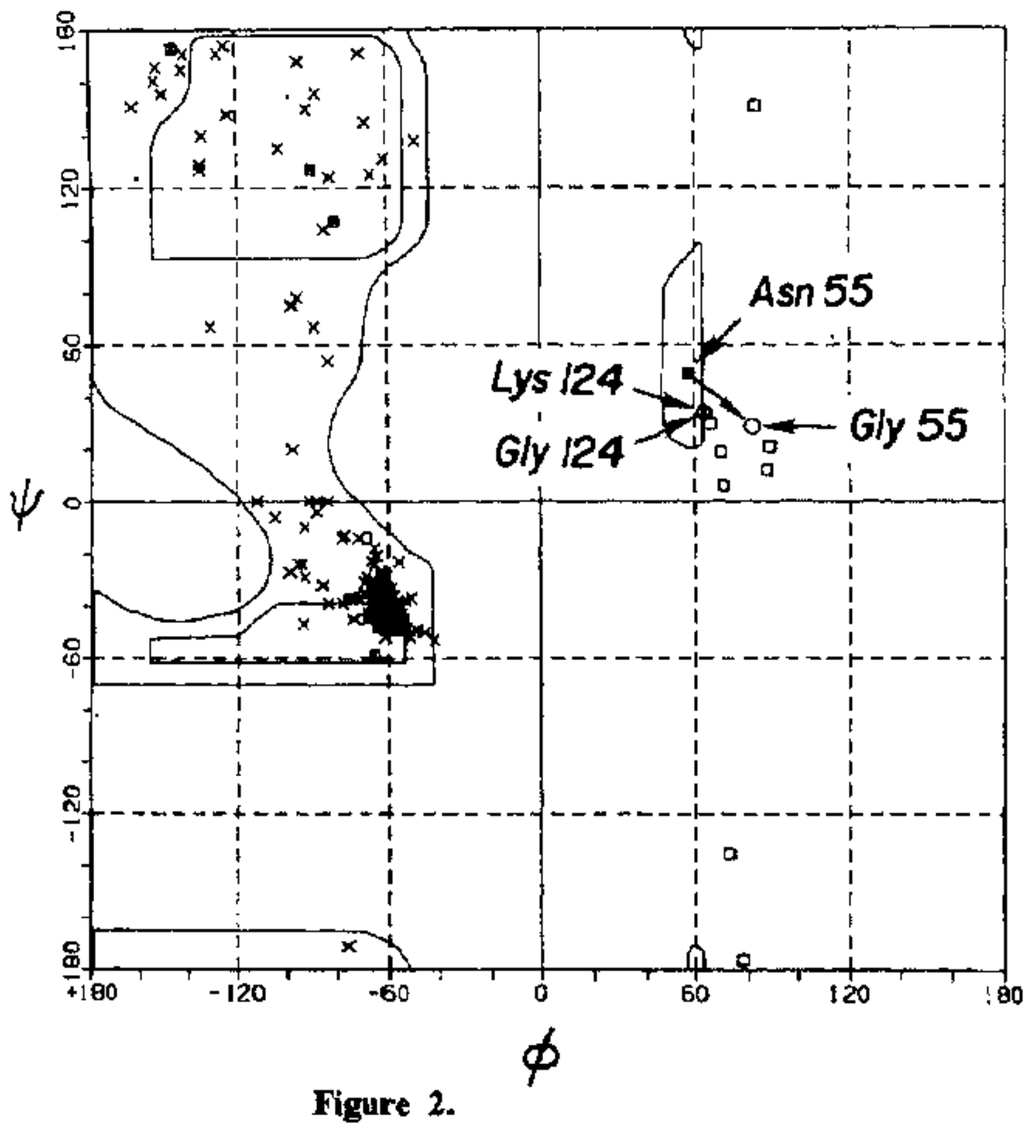

Figure 2.

tures suggests that the changes in (ϕ,ψ) associated with the Asn 55 → Gly replacement may be just one specific example of a more general phenomenon. Figure 3,*a* shows the (ϕ,ψ) values observed for non-glycine residues in 37 refined high-resolution protein crystal structures taken from the Brookhaven Data Bank. (Parenthetically, the correspondence between this observed distribution and the 'allowed' (ϕ,ψ) values anticipated in Ramachandran's original paper is quite remarkable.) Figure 3,*b* shows just the asparagines. In both cases the left-handed helical residues have conformations that cluster around $\phi = 60°$, $\psi = 30°$. [The conformation of Asn 55 in wild-type lysozyme ($\phi = 58°$, $\psi = 49°$) is within this

Figure 2. Ramachandran diagram showing the backbone conformational angles for wild-type lysozyme. Glycines (□); non-glycines (×). Also shown are the conformational angles for Gly 55 and Gly 124 in the respective mutant structures (○). The Asn 55 → Gly 55 substitution is associated with changes of about 20° in ϕ and ψ (shown by the arrow) whereas the Lys → Gly substitution at residue 124 leaves the conformational angles invariant. (From ref. 5.)

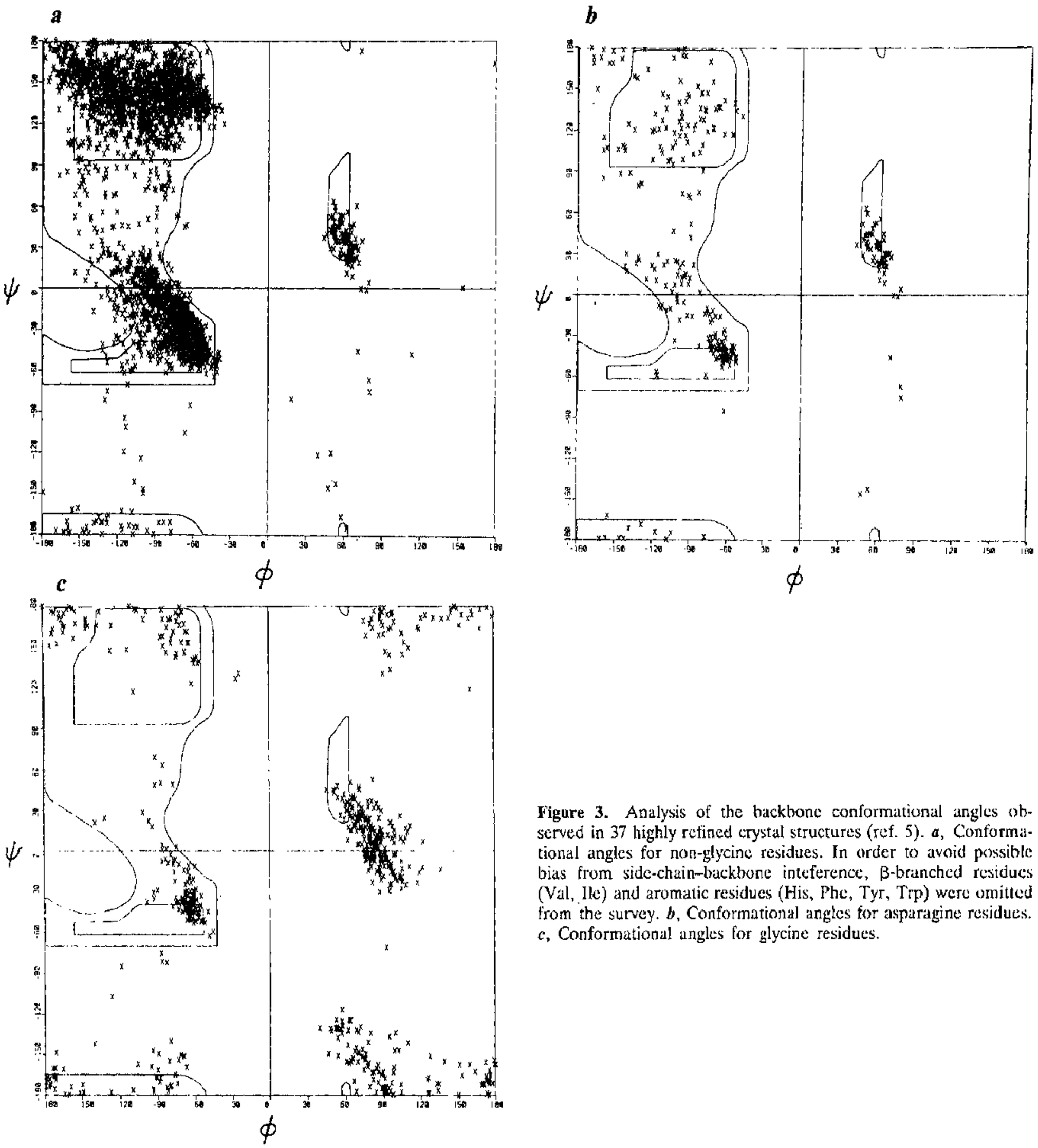

Figure 3. Analysis of the backbone conformational angles observed in 37 highly refined crystal structures (ref. 5). *a*, Conformational angles for non-glycine residues. In order to avoid possible bias from side-chain–backbone inteference, β-branched residues (Val, Ile) and aromatic residues (His, Phe, Tyr, Trp) were omitted from the survey. *b*, Conformational angles for asparagine residues. *c*, Conformational angles for glycine residues.

cluster.] The observed conformations of left-handed glycines, however, are different (Figure 3,*c*). The range of (ϕ,ψ) values is broader and the center of the distribution moves to $\phi = 90°$, $\psi = 0°$. The statistical distributions (Figures 3, *b* and *c*) taken together with the changes in (ϕ,ψ) associated with the specific Asn → Gly substitution (Figure 2) strongly suggest that the minimum energy conformation for left-handed glycines is close to ($\phi = 90°$, $\psi = 0°$), while that for left-handed helical non-glycines is close to ($\phi = 60°$, $\psi = 30°$). This result is, however, at variance with theoretical (ϕ,ψ) energy contour maps since none of these predict that glycine has an energy minimum at $\phi = 90°$, $\psi = 0°$ (refs. 6 – 11). It should be noted that the (ϕ,ψ) values shown in Figure 3,*c* are for glycines in folded protein structures. It could be argued, therefore, that the local secondary structure influences the (ϕ,ψ) values and that Figure 3,*c*, does not constitute evidence that an isolated glycine has an energy minimum at ($\phi = 90°$, $\psi = 0°$). On the other hand, it is not as if all the 'left-handed helical' glycine residues shown in Figure 3,*c* occur within the same type of local secondary structure. Rather, the survey includes all glycines, no matter what their environment.

Lysine 124 → glycine. Lysine 124 is part of an extended segment on the surface of the molecule that

connects the antiparallels helices 115–123 and 126–134. As required by the left-handed helical conformation of Lys 124, the amides of the 123–124 and 124–125 peptide linkages are essentially parallel to each other. These two amides both make H-bonds to the carbonyl oxygens in the last turn of the 126–134 helix (Figure 4,*b*).

Comparison of the refined coordinates of the mutant and wild-type structures (Figure 4,*b*) shows that the Lys 124 → Gly substitution does not result in any significant change in the protein backbone at or adjacent to residue 124. The (ϕ,ψ) values in the mutant structure are, within experimental error, the same as wild-type (Figure 2). Because the backbone amides of Lys 124 and Arg 125 are both involved in a tight network of intramolecular H-bonds, the (ϕ, ψ) values of residue 124 may be essentially 'locked', and unable to relax as was observed in the Asn 55 → Gly substitution. The conformation ($\phi = 66°$, $\psi = 33°$) maintained in both the wild-type and mutant structures is well within the observed distributions for both glycines (Figure 3,*c*) and non-glycines (Figure 3,*a*).

The replacement of Lys 124 with glycine causes virtually no change in the melting temperature of the protein. The simplest and most plausible interpretation of this result is that the backbone conformational energy of the 'left-handed' lysine is essentially the same as that of the glycine. On the other hand, the results do not rule out a more complicated scenario in which the introduction of the left-handed glycine contributes backbone conformational stability, but this is offset by the loss of favorable interactions consequent on the removal of the lysine side chain.

One of the reasons for carrying out the Asn 55 → Gly and Lys 124 → Gly substitutions was to see if such replacements would increase the thermostability of the protein. This was not the case. In one case the stability of the mutant protein was unaffected and in the other case it was marginally reduced relative to wild-type. Contrary to some estimates, the present results suggest that a glycine and a non-glycine with left-handed helical conformations have very similar backbone conformational energies, probably within 0.5 kcal/mol (ref. 5).

Use of the Ramachandran diagram in structure evaluation

Recently several proposed models of protein structures have subsequently been found to be in error[12].

Traditionally the crystallographic '*R*-value' has been used as a diagnostic for the reliability of a given structure determination but experience has shown that this need not be a reliable measure. To take an example from our own laboratory, an initial, abortive attempt to solve the structure of the lysozyme from goose egg-white led to an (incorrect) model with a crystallographic residual of 28% at 2.8 Å resolution. This relatively low *R*-value, together with the fact that the model had good stereochemistry and did not include any solvent, could have been taken as evidence that the model was substantially correct. On the other hand, the model included relatively few

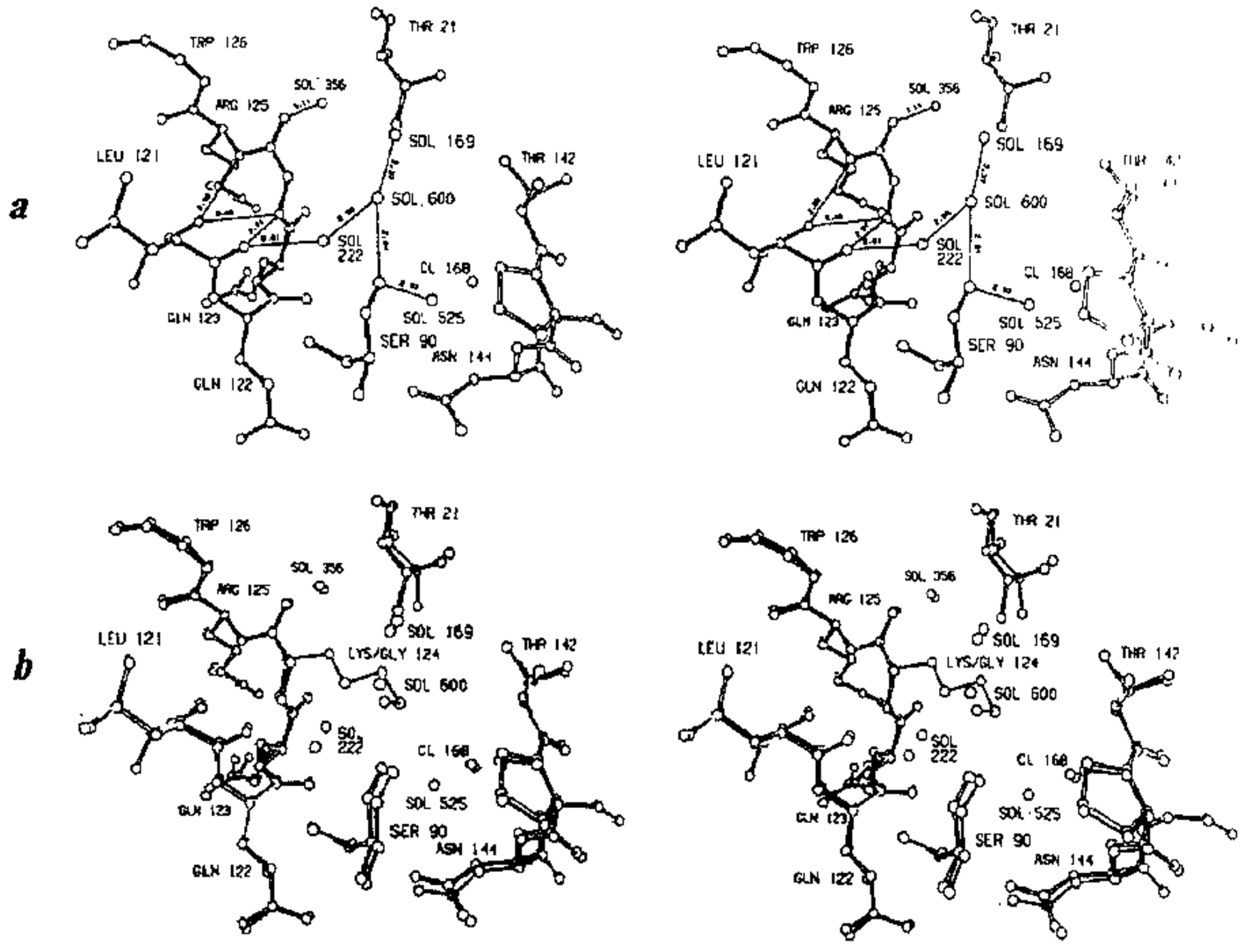

Figure 4. *a*, Structure of the mutant lysozyme K124G in the vicinity of residue 124. CL 168 is thought to be a solvent molecule that replaces a chloride ion bound in the wild-type crystal structure. *b*, Superposition of K124G (open bonds) and wild-type lysozyme (solid bonds). The residues to the right and top of the figure (Thr 21, Thr 142–Asn 144) correspond to an adjacent lysozyme molecule in the crystal. (From ref. 5.)

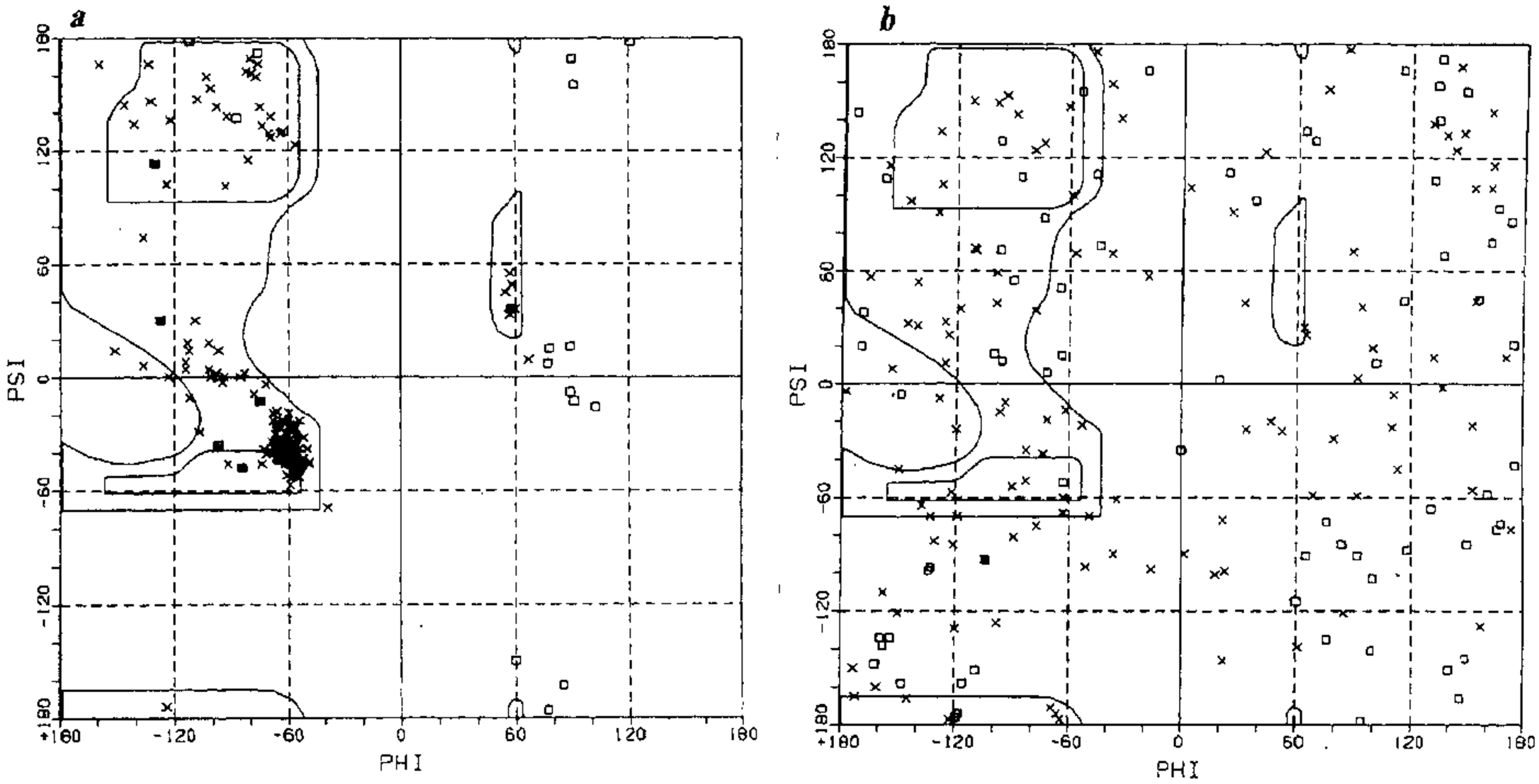

Figure 5. Ramachandran diagrams for two models of the structure of the lysozyme from goose egg-white. Squares indicate glycine residues and crosses non-glycines. *a*, Structure after refinement to a crystallographic residual $R = 15\%$ at 1.6 Å resolution. *b*, An incorrect, abortive model, that refined to a residual $R = 28\%$ at 2.8 Å resolution (unpublished results of L.H.W. and B.W.M.).

side chains (beyond the β-carbon) and consisted of 15 disconnected polypeptide segments. These limitations caused us to abandon the model, and, in due course to solve the structure via the inclusion of a new isomorphous heavy-atom derivative[13,14]. The structure has subsequently been refined to a residual of 15% at 1.6 Å resolution (Weaver and Matthews unpublished results).

Having these two models for goose egg-white lysozyme (the abortive, incorrect one and the final refined structure) provides an opportunity to test the use of the Ramachandran diagram as a tool for structure diagnosis. The calculated (ϕ,ψ) distributions for the two models of goose lysozyme are shown in Figures 5, *a* and *b*. The difference is self-evident. Because the 'allowed' regions of the Ramachandran map are substantially smaller than the 'disallowed' regions, it is likely that a region of a protein structure that is incorrect will likely include one or more residues with (ϕ,ψ) values that fall into the 'disallowed' region. Of course one must recognize that unusual (ϕ,ψ) values do occur, from time to time, in well-determined structures, so that such observations must not be discounted as 'errors'. On the other hand, if a proposed protein structure has a high proportion of residues with unusual (ϕ,ψ) values, then one is entitled to regard such a structure with skepticism. The expected frequency of occurrence of unusual (ϕ,ψ) values in a well-determined protein structure can be estimated by inspection of Figure 3,*a*.

1. Ramachandran, G. N., Ramakrishnan, C. and Sasisekharan, V., *J. Mol. Biol.*, 1963, 7, 95.
2. Matthews, B. W., *Biochemistry*, 1987, **26**, 6885.
3. Alber, T. and Matthews, B. W., *Methods Enzymol.*, 1987, **154**, 511.
4. Weaver, L. H. and Matthews, B. W., *J. Mol. Biol.*, 1987, **193**, 189.
5. Nicholson, H., Solderlind, E., Tronrud, D. E. and Matthews, B. W., *J. Mol. Biol.*, 1989, **210**, 181.
6. Brandt, D. A. and Flory, P. J., *J. Mol. Biol.*, 1967, **23**, 47.
7. Hoffmann, R. and Imamura, A., *Biopolymers*, 1969, **7**, 207.
8. Pullman, B., *Aspects de la Chimie Quantique Contemporaire* (eds. Daudel, R. and Pullman, A.), Colloque International du CNRS, Paris, 1971, pp. 261–300.
9. Dunfield, L. G., Burgess, A. W. and Scheraga, H. A., *J. Phys. Chem.*, 1978, **82**, 2609.
10. Peters, D. and Peters, J., *J. Mol. Struct.*, 1981, **85**, 107.
11. Weiner, S. J., Kollmann, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta, S. Jr. and Weiner, P., *J. Am. Chem. Soc.*, 1984, **106**, 765.
12. Branden, C-I. and Jones, T. A., *Nature*, 1990, **343**, 687.
13. Grütter, M. G., Weaver, L. H. and Matthews, B. W., *Nature*, 1983, **303**, 828.
14. Weaver, L. H., Grütter, M. G., Remington, S. J., Gray, T. M., Isaacs, N. W. and Matthews, B. W., *J. Mol. Evol.*, 1985, **21**, 97.

ACKNOWLEDGEMENTS. This work was supported in part by grants to B.W.M. from the National Institutes of Health (GM21967), the National Science Foundation (DMB8611084) and also by a grant from the Markey Charitable Trust.