diffracting planes, the 'arithmetical' adding of 'dynamical' peaks from different parts of the illuminated spot would seem to be more appropriate.

If the layer consists of different sublayers (crystallites) of thicknesses greater than a few hundred ångströms, the diffraction peak becomes broader than in a perfect crystal (or is split into a series of peaks). In such a case, establishing a correlation between the peak maximum and the lattice constant is extremely difficult. The peak position will also be dependent on block disorientations, their size, sample bending, the presence of precipitates, which give rise to diffuse scattering (*e.g.* Holy & Hartwig, 1988) *etc.*

This work was sponsored by the ICTP Programme for Training and Research in Italian Laboratories and a KBN (Poland) grant (201609101).

References

Bartels, W. J. (1983). *J. Vac. Sci. Technol.* **B1**, 338-345.
Bond, W. L. (1960). *Acta Cryst.* **13**, 814-816.
Cargill, G. S. III, Angilello, J. & Kavanagh, K. L. (1988). *Phys. Rev. Lett.* **61**, 11748-11751.
Holy, V. & Hartwig, J. (1988). *Phys. Status Solidi B*, **145**, 363-372.
Mikkelsen, J. C. & Boyce, J. B. (1983). *Phys. Rev. B*, **28**, 7130-7140.
Nakajima, M., Sato, T., Inada, T., Fukuda, T. & Ishida, K. (1986). *Appl. Phys. Lett.* **49**, 1251-1253.
Ohmori, H., Tomita, M., Tsuchiya, N. & Matsushita, Y. (1991). *Adv. X-ray. Anal.* **34**, 577-586.
Tanner, B. K., Turnbull, A. G., Stanley, C. R., Kean, A. H. & McElhinney, M. (1991). *Appl. Phys. Lett.* **59**, 2272-2274.
Wie, C. R., Tombrello, T. A. & Vreeland, T. Jr (1986). *J. Appl. Phys.* **59**, 3743-3746.

Computer Programs

J. Appl. Cryst. (1993). **26**, 283-291

PROCHECK: a program to check the stereochemical quality of protein structures.

By Roman A. Laskowski, *Biomolecular Structure and Modelling Unit, Department of Biochemistry and Molecular Biology, University College, Gower Street, London WC1E 6BT, England*, Malcolm W. MacArthur, *Biomolecular Structure and Modelling Unit, Department of Biochemistry and Molecular Biology, University College, Gower Street, London WC1E 6BT, England, and Crystallography Department, Birkbeck College, Malet Street, London WC1E 7HX, England*, David S. Moss, *Crystallography Department, Birkbeck College, Malet Street, London WC1E 7HX, England* and Janet M. Thornton, *Biomolecular Structure and Modelling Unit, Department of Biochemistry and Molecular Biology, University College, Gower Street, London WC1E 6BT, England*

(*Received* 21 *August* 1992; *accepted* 23 *September* 1992)

Abstract

The *PROCHECK* suite of programs provides a detailed check on the stereochemistry of a protein structure. Its outputs comprise a number of plots in PostScript format and a comprehensive residue-by-residue listing. These give an assessment of the overall quality of the structure as compared with well refined structures of the same resolution and also highlight regions that may need further investigation. The *PROCHECK* programs are useful for assessing the quality not only of protein structures in the process of being solved but also of existing structures and of those being modelled on known structures.

1. Introduction

A protein structure derived from experimental data will be subject to many sources of error, both experimental and in the interpretation of results (Brändén & Jones, 1990). It is important, therefore, to have an assessment of a structure's overall 'quality' and to be able to identify regions which may need careful investigation. This applies not only at the structure-determination stage but also when one is 'model building' a sequence onto a known structure (Blundell *et al.*, 1988), or using the existing database of known structures for analysis and prediction.

For determining overall quality, the two most widely used measures are the resolution and the *R* factor. The higher the resolution, the greater the number of independent experimental observations obtained from the diffraction data and hence the greater the accuracy of the molecular structure. This was illustrated by Hubbard & Blundell (1987), who showed how the r.m.s. distance between corresponding C^{α} atoms of independently refined structures of identical amino-acid sequences decreases as the resolution improves. In other words, the structures approach one another as the resolution gets better, which suggests that the errors in each are decreasing, as might be expected. The *R* factor, on the other hand, is a less certain guide; it can be artificially reduced in a number of ways and so can sometimes be misleading (Brändén & Jones, 1990). Nevertheless, it is common to take reliable structures as being those with a resolution of 2.0 Å or better and an *R* factor no worse than 20%.

Other measures exist and fall into two categories. The first covers structures that are calculated from the experimental data. One such measure is the estimated *average* positional error, Δr, in the structure's atomic coordinates. This can be obtained from

0021-8898/93/020283-09$06.00 © 1993 International Union of Crystallography

Table 1. *Stereochemical parameters of Morris, MacArthur, Hutchinson & Thornton* (1992), *derived from high-resolution protein structures, against which the structure is compared on a residue-by-residue basis*

Stereochemical parameter	Mean value	Standard deviation
φ–ψ in most favoured regions of Ramachandran plot (%)	>90	
χ_1 dihedral angle (°)		
gauche minus	64.1	15.7
trans	183.6	16.8
gauche plus	−66.7	15.0
χ_2 dihedral angle (°)	177.4	18.5
Proline φ torsion angle (°)	−65.4	11.2
Helix φ torsion angle (°)	−65.3	11.9
Helix ψ torsion angle (°)	−39.4	11.3
χ_3 (S−S bridge) (°)		
Right-handed	96.8	14.8
Left-handed	−85.8	10.7
Disulfide bond separation (Å)	2.0	0.1
ω dihedral angle (°)	180.0	5.8
Main-chain hydrogen-bond energy (kJ mol^{-1})*	−8.50	3.14
C^{α} chirality: ζ 'virtual' torsion angle ($C^{\alpha}-N-C-C^{\beta}$) (°)	33.9	3.5

* Evaluated using the Kabsch & Sander (1983) method.

a Luzzati plot (Luzzati, 1952), a σ_A plot (Read, 1986), or the R indices and tables of Elango & Parthasarathy (1990). Another measure is the 'free R value', or $R_T^{\rm free}$, of Brünger (1992), which is a means of calculating an *unbiased R* factor. For structures determined using nuclear magnetic resonance methods, for which neither the resolution nor the R factor have any meaning, quality is usually assessed by the variability between the many trial structures that are consistent with the experimental data.

The second category of measures are calculated directly from the coordinates of the structure and so do not require the experimental data. These include: the 3D profiles of Lüthy, Bowie & Eisenberg (1992), which can verify whether the given structure is compatible with the sequence; the calculation of the numbers of bad contacts between nonbonded atoms in the protein (Islam, Sternberg & Weaver, 1990); and the use of stereochemical parameters to classify a structure according to its relative reliability (Morris, MacArthur, Hutchinson & Thornton, 1992).

All the above are measures of the overall quality and so do not provide information on the different parts of the structure. Some regions are likely to be more reliable than others. For example, regions in the core of the protein will tend to have more clearly defined electron densities and so will have been easier to interpret during the determination of the structure. Conversely, the loop regions on the surface will tend to exhibit greater static and dynamic disorder and thus have weaker (or, rather, more 'smeared out') electron densities, which makes them more difficult to interpret reliably.

During the determination and refinement of a structure, regions in error are usually identified by stereochemical considerations. Standard checks include the identification of bad contacts, the careful inspection of the structure on a graphics terminal and the use of a Ramachandran plot (Ramachandran, Ramakrishnan & Sasisekharan, 1963) to see which residues lie in the 'disallowed' regions. A further guide is provided by the atomic B values, which will be large for erroneously placed atoms; however, when looking at existing structures, one finds that not all published structures show the B values – about 16% of the coordinate files in the Brookhaven databank (Bernstein *et al.*, 1977) do not.

Three methods have been devised fairly recently to test the quality of a structure's electron-density map on a residue-by-residue basis (Wierenga, Kalk & Hol, 1987; Jones, Zou, Cowan & Kjeldgaard, 1991; Engh & Huber, 1991). These can give an additional measure of the local reliability, but require the experimental data.

Here we describe a suite of programs, *PROCHECK*, that uses stereochemical considerations alone, both to provide an overall assessment of the stereochemistry of a given structure and to highlight regions that may need further investigation. The programs can therefore be used independently of the experimental data and so may be applied to structures that have already been published, to those in the process of being solved, or to model-built structures.

One of the by-products of running *PROCHECK* is that the coordinates file will be 'cleaned up', in that any mislabelled atoms will be relabelled in accordance with the IUPAC naming conventions (IUPAC–IUB Commission on Biochemical Nomenclature, 1970).

2. The programs

The *PROCHECK* suite comprises five programs, four written in Fortran77 and one written in C, that are run in succession: *CLEAN.F*, *SECSTR.F*, *NB.C*, *ANGLEN.F* and *PPLOT.F*. The main input to the suite is a file containing the structure's coordinates in Brookhaven format (Bernstein *et al.*, 1977). An additional input is a parameter file containing various program options that govern the plots and printouts produced. This is a text file which can be modified using any text editor.

The stereochemical parameters used for assessing the structure are those derived by Morris, MacArthur, Hutchinson & Thornton (1992), together with the bond lengths and bond angles derived from a recent and comprehensive analysis (Engh & Huber, 1991) of small-molecule structures in the Cambridge Structural Database (Allen *et al.*, 1979), which now numbers over 80 000 structures. These sets of parameters are listed in Tables 1 and 2, respectively.

The output produced by *PROCHECK* comprises a number of plots, together with a detailed residue-by-residue listing. The plots for γII-crystallin (Brookhaven code 1gcr; Wistow *et al.*, 1983) (see Figs. 1 to 5) show clearly how the structure compares with well refined structures of the same resolution and thus give a quick visual picture of its overall stereochemical quality. The last of the plots (Fig. 5) also gives a view of how good the stereochemistry is on a residue-by-residue basis. All the plots are output in PostScript format (Adobe Systems Inc., 1985) and so can be printed on a PostScript laser printer or displayed on a graphics screen using the appropriate software (*e.g. GHOSTSCRIPT* on Sun workstations or *PSVIEW* on Silicon Graphics IRIS-4D systems).

The residue-by-residue listing gives more detailed information and highlights regions of the structure that may be in error, or just need further examination. For each residue, the listing shows the calculated value of each of the stereochemical parameters and highlights any values that deviate from the 'ideal'. Summary pages provide the information in a concise form and various user-definable options make it possible to select how much of the information is printed.

The five programs in the suite are as follows. The first, *CLEAN.F*, produces a 'cleaned-up' version of the input structure; it ensures that the atoms have been correctly labelled in accordance with the IUPAC naming conventions (IUPAC–IUB Commission on Biochemical Nomenclature, 1970). A typical error is that the $N^{\eta 1}$ and $N^{\eta 2}$ atoms of arginine are labelled the

Table 2. *Main-chain bond lengths and bond angles and their standard deviations, as observed in small molecules (Engh & Huber,* 1991)

Atom labelling follows that used in the *X-PLOR* dictionary (Brünger, 1988), with some additional atom types (marked with asterisks) as defined by Engh & Huber (1991).

Bond lengths

Bond	*X-PLOR* labelling		Value (Å)	σ (Å)
C−N	C−NH1	(except Pro)	1.329	0.014
	C−N	(Pro)	1.341	0.016
C−O	C−O		1.231	0.020
C^{α}−C	CH1*E*−C	(except Gly)	1.525	0.021
	CH2*G**−C	(Gly)	1.516	0.018
C^{α}−C^{β}	CH1*E*−CH3*E*	(Ala)	1.521	0.033
	CH1*E*−CH1*E*	(Ile, Thr, Val)	1.540	0.027
	CH1*E*−CH2*E*	(the rest)	1.530	0.020
N−C^{α}	NH1−CH1*E*	(except Gly, Pro)	1.458	0.019
	NH1−CH2*G**	(Gly)	1.451	0.016
	N−CH1*E*	(Pro)	1.466	0.015

Bond angles

Angle	*X-PLOR* labelling		Value (°)	σ (°)
C−N−C^{α}	C−NH1−CH1*E*	(except Gly, Pro)	121.7	1.8
	C−NH1−CH2*G**	(Gly)	120.6	1.7
	C−N−CH1*E*	(Pro)	122.6	5.0
C^{α}−C−N	CH1*E*−C−NH1	(except Gly, Pro)	116.2	2.0
	CH2*G**−C−NH1	(Gly)	116.4	2.1
	CH1*E*−C−N	(Pro)	116.9	1.5
C^{α}−C−O	CH1*E*−C−O	(except Gly)	120.8	1.7
	CH2*G**−C−O	(Gly)	120.8	2.1
C^{β}−C^{α}−C	CH3*E*−CH1*E*−C	(Ala)	110.5	1.5
	CH1*E*−CH1*E*−C	(Ile, Thr, Val)	109.1	2.2
	CH2*E*−CH1*E*−C	(the rest)	110.1	1.9
N−C^{α}−C	NH1−CH1*E*−C	(except Gly, Pro)	111.2	2.8
	NH1−CH2*G**−C	(Gly)	112.5	2.9
	N−CH1*E*−C	(Pro)	111.8	2.5
N−C^{α}−C^{β}	NH1−CH1*E*−CH3*E*	(Ala)	110.4	1.5
	NH1−CH1*E*−CH1*E*	(Ile, Thr, Val)	111.5	1.7
	N−CH1*E*−CH2*E*	(Pro)	103.0	1.1
	NH1−CH1*E*−CH2*E*	(the rest)	110.5	1.7
O−C−N	O−C−NH1	(except Pro)	123.0	1.6
	O−C−N	(Pro)	122.0	1.4

wrong way round. Similarly, atom labels for Phe, Tyr, Asp and Glu residues are corrected where necessary. The program also checks that the correct L/D stereochemical labels have been assigned to individual residues and that chain breaks have been correctly identified. (However, no check is made of H-atom nomenclature.)

The second program, *SECSTR.F*, is responsible for making secondary-structure assignments, on a residue-by-residue basis, in accordance with the modified method of Kabsch & Sander (1983).

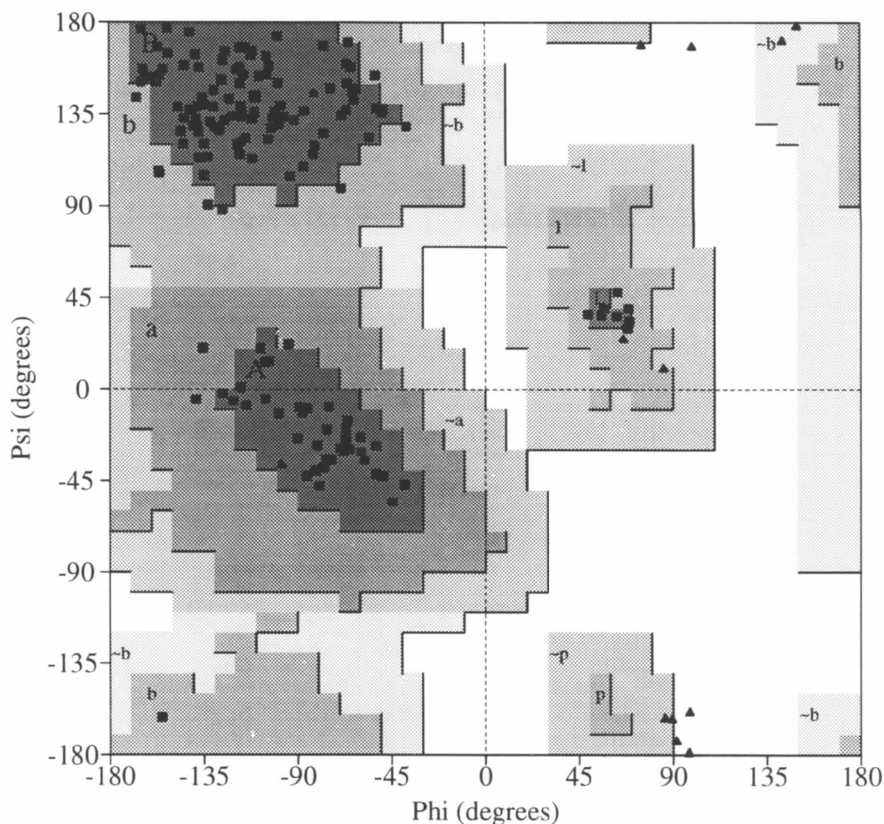
The third program, *NB.C*, identifies all nonbonded interactions between different pairs of residues. Nonbonded interactions are defined as those where the closest atom–atom contact between two residues is less than 4.0 Å and the atoms concerned are four or more bonds apart.

The fourth program, *ANGLEN.F*, calculates all main-chain bond lengths and bond angles in the structure, while the last, *PPLOT.F*, produces all the plots and the detailed residue-by-residue listing.

The *PROCHECK* suite is easy to use and should prove useful for the solution of new structures, assessment of existing structures and model building of unknown structures. In the latter case, it also provides a means of checking the quality of the structure(s) on which the model is to be based before model building is started.

3. Availability

The *PROCHECK* source code and operating instructions are available to academic institutions, free of charge, from the authors. The programs are supplied with script files for running on UNIX operating systems and command files for running under VAX/VMS. Requests for the programs can be made by post, or by email to roman@uk.ac.ucl.bioc.bsm.



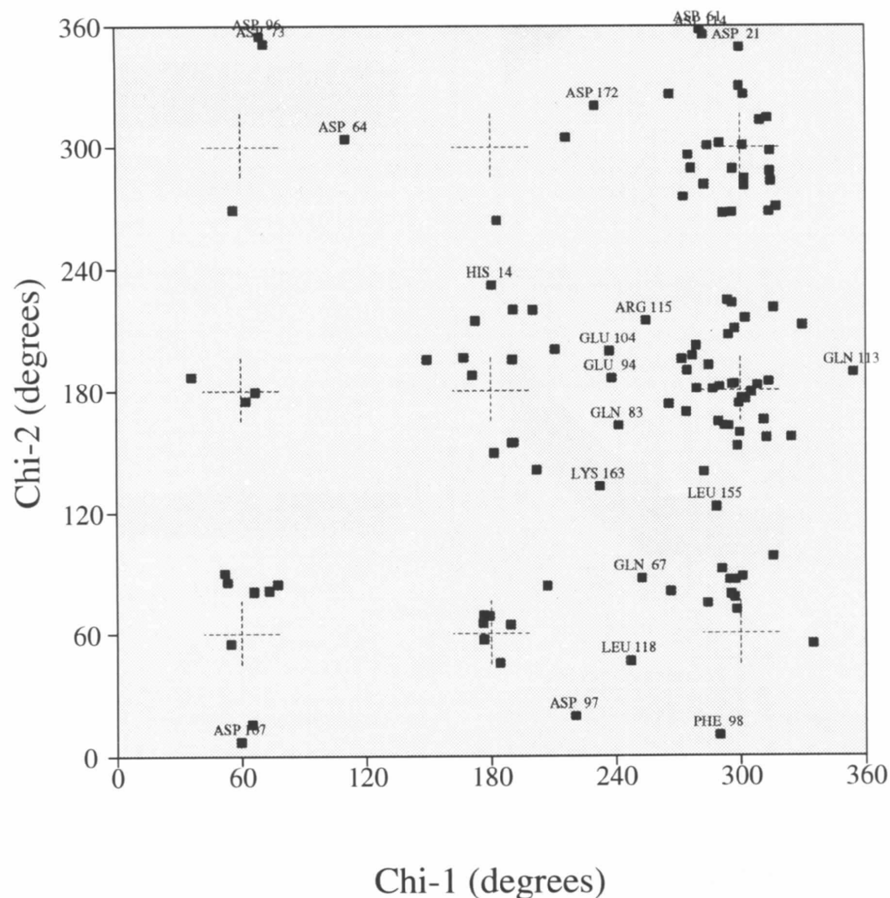
Plot statistics

Residues in most favoured regions [A,B,L]	136	90.1%
Residues in additional allowed regions [a,b,l,p]	15	9.9%
Residues in generously allowed regions [~a,~b,~l,~p]	0	0.0%
Residues in disallowed regions	0	0.0%
	----	------
Number of non-glycine and non-proline residues	151	100.0%
Number of end-residues	1	
Number of glycine residues (shown as triangles)	14	
Number of proline residues	8	

Total number of residues	174	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

Fig. 1. The first of the plots produced by *PROCHECK* is a Ramachandran plot. The example here is for γII-crystallin, Brookhaven code 1gcr (Wistow *et al.*, 1983). Glycine residues are separately identified by triangles. The shading represents the different regions of the plot; the darker the area the more favourable the φ–ψ combination. The different regions are those described in Morris, MacArthur, Hutchinson & Thornton (1992). Ideally, the structure should have over 90% of the residues in the darkest 'core' regions (marked A, B and L) as is the case in this example. The appearance of the plot itself can be modified to some extent by amending the program parameters. Thus, the shading and/or lettering of the different regions can be switched on or off, the region borders can be drawn in or not drawn in and the individual residues can be labelled.



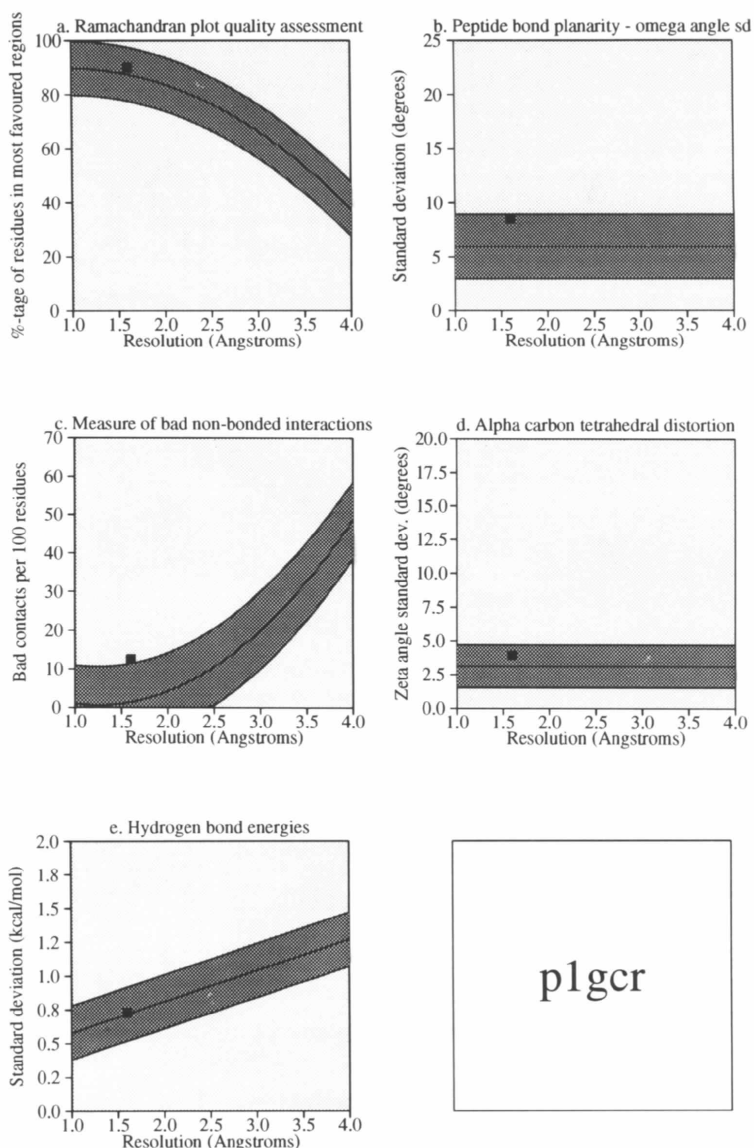
Total number of residues = 174

Number of residues plotted = 119

Number of labelled residues = 20

The dashed crosses represent the gauche minus, trans, and gauche plus regions for the Chi-1 and Chi-2 dihedral angles. The width of each cross is approximately one standard deviation. The points should therefore cluster around these crosses. Points more than 2.5 standard deviations from the ideal are labelled

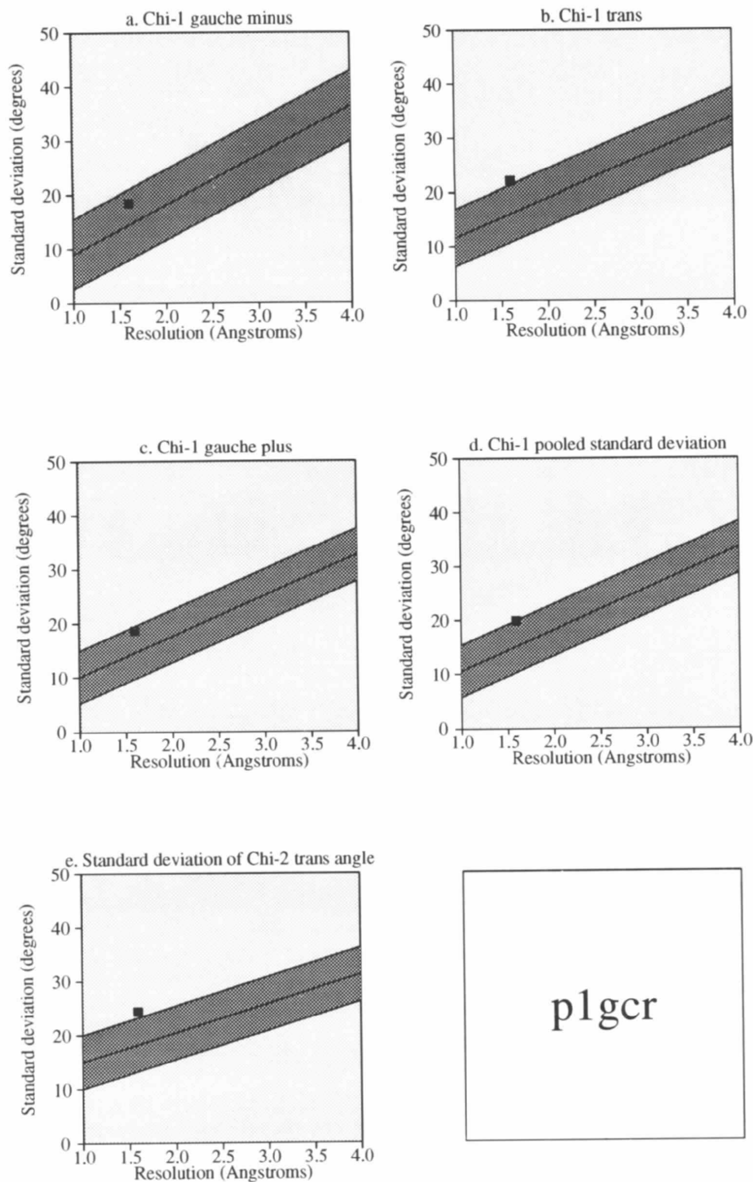
Fig. 2. The second of the plots shows a graph of the χ_1 *versus* the χ_2 torsion angles for each residue, where applicable. As each χ_1 and χ_2 can be in one of three preferred configurations (*gauche minus*, *trans* and *gauche plus*), there are $3\times3=9$ combinations for the two angles. The nine 'ideal' positions are marked by crosses on the graph. Residues lying more than 2.5 standard deviations away from these positions are labelled. (The number 2.5 can be altered in the parameter file if required.)



Plot statistics

Stereochemical parameter	No. of data pts	Parameter value	Comparison values: Typical value	Band width	No. of band widths from mean	
a. %-tage residues in A, B, L	151	90.1	87.5	10.0	0.3	Better
b. Omega angle st dev	173	8.6	6.0	3.0	0.9	Worse
c. Bad contacts / 100 residues	22	12.6	1.4	10.0	1.1	WORSE
d. Zeta angle st dev	160	3.9	3.1	1.6	0.5	Worse
e. H-bond energy st dev	99	0.7	0.7	0.2	0.1	Worse

Fig. 3. The third plot shows the graphs of five main-chain properties of the structure (values marked by solid squares) and how these properties compare with well refined structures at a similar resolution. The dark band in each graph represents the results from the well refined structures; the central line is a least-squares fit to the mean trend as a function of resolution, while the width of the band either side of it corresponds to a variation of one standard deviation about the mean. In some cases, the trend is dependent on the resolution, and in other cases is independent of it. Graph (*a*) shows the Ramachandran-plot quality, as measured by the percentage of the protein's residues that are in its most favoured, or 'core', regions; (*b*) shows the planarity of the peptide bond as measured by the standard deviation of the ω torsion angles; (*c*) shows the number of bad contacts per 100 residues; (*d*) shows the C^{α} tetrahedral distortion, measured by the standard deviation of the ζ 'torsion' angle (a notional torsion angle in that it is not defined about any actual bond in the structure but by the residue's C^{α}, N, C and C^{β} atoms); and (*e*) shows the standard deviation of the hydrogen-bond energies for main-chain hydrogen bonds, calculated using the method of Kabsch & Sander (1983).



Plot statistics

Stereochemical parameter	No. of data pts	Parameter value	Comparison values: Typical value	Comparison values: Band width	No. of band widths from mean	
a. Chi-1 gauche minus st dev	27	18.4	14.5	6.5	0.6	Worse
b. Chi-1 trans st dev	39	22.0	16.1	5.3	1.1	WORSE
c. Chi-1 gauche plus st dev	84	18.6	14.5	4.9	0.8	Worse
d. Chi-1 pooled st dev	150	19.9	15.1	4.8	1.0	Worse
e. Chi-2 trans st dev	54	24.4	18.2	5.0	1.2	WORSE

Fig. 4. The fourth plot shows five different side-chain properties. Like the graphs in Fig. 3, these show how the structure (values marked by solid squares) compares with well refined structures at a similar resolution. Again, the dark band in each graph represents the results from the well refined structures, giving one standard deviation about a mean trend. The graphs show the standard deviations of the side-chain torsion angles: (*a*) χ_1 *gauche minus*; (*b*) χ_1 *trans*; (*c*) χ_1 *gauche plus*; (*d*) pooled χ_1 torsion angles; and (*e*) χ_2 *trans*.

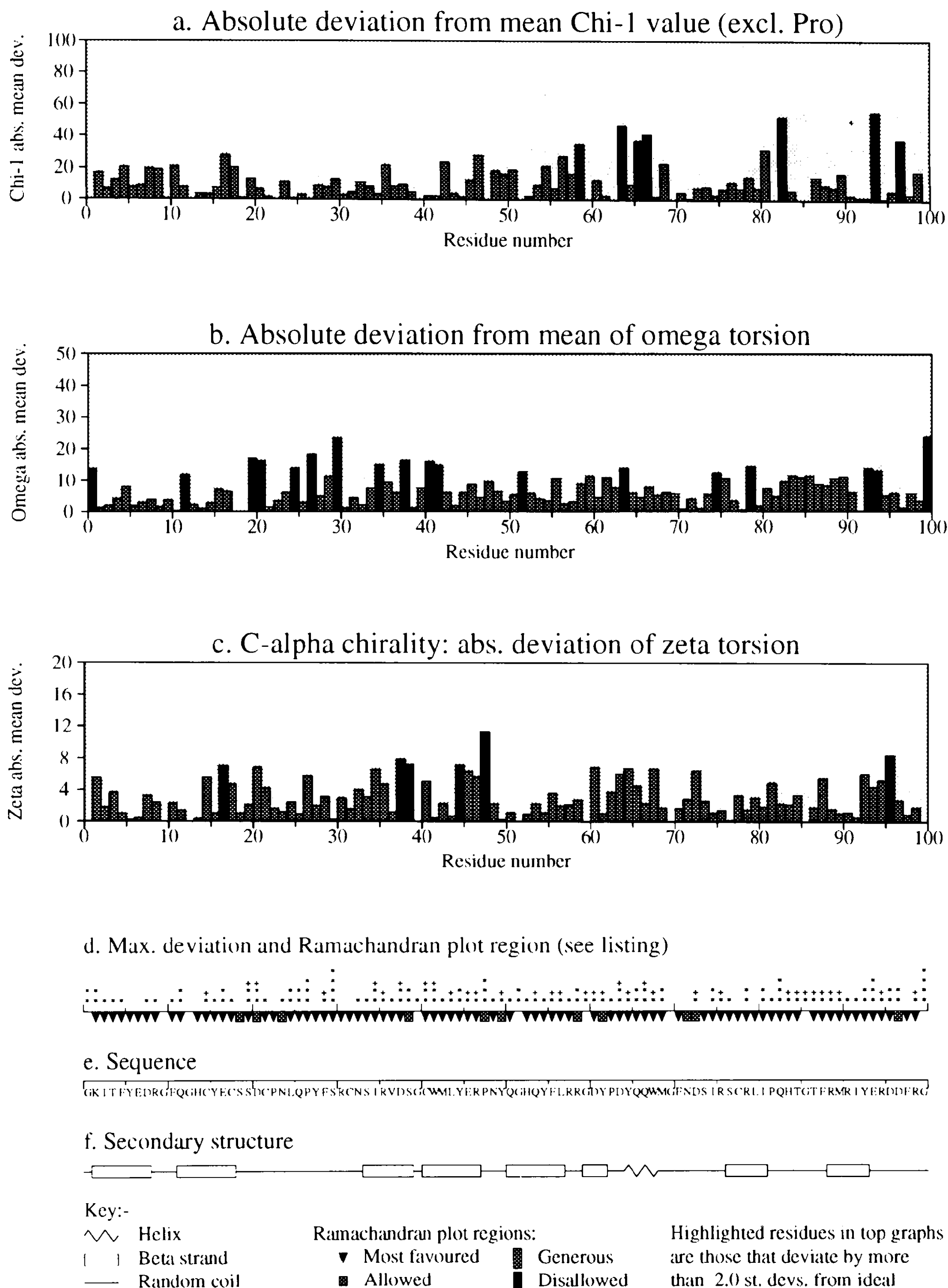


Fig. 5. The fifth plot shows a number of residue properties. In this example, only the first 'page' for 1gcr is shown, representing only the first 100 residues. The top three graphs, (*a*), (*b*) and (*c*), can be selected from seven possibles that comprise: the three shown, together with the absolute deviation of main-chain hydrogen-bond energy from the 'ideal' value; the *B* value of the γ atom (O, C or S, whichever is used in the definition of the χ_1 torsion angle); the average *B* value of main-chain atoms; and the average *B* value of side-chain atoms. The bottom part of the diagram illustrates information given in greater detail in the residue-by-residue listing. Graph (*d*) shows each residue's 'maximum deviation' from ideal values, as shown on the listing, and in which region of the Ramachandran plot the residue is located. The key at the bottom of the page explains the markers used for the latter. Graph (*e*) gives the protein's sequence and graph (*f*) shows a schematic picture of its secondary structure, as defined using the Kabsch & Sander (1983) assignments. The key at the bottom left of the plot shows which structure is which.

MWM is supported by a SERC studentship. Major contributors to the programs in the suite are Gail Hutchinson, David Jones, Louise Morris, Dorica Naylor and David Smith. We thank Oxford Molecular Ltd for providing financial assistance during the final stages of the project to compile the *PROCHECK* suite. Thanks are also due to Paula Kuser, Helen Stirk, V. Dhanaraj and Geoff Barton for helpful comments and suggestions. Finally, the PostScript output of the *MOLSCRIPT* program (Kraulis, 1991) proved an invaluable guide to the generation of PostScript files.

References

Adobe Systems Inc. (1985). *PostScript Language Reference Manual.* Reading, MA: Addison–Wesley.
ALLEN, F. H., BELLARD, S., BRICE, M. D., CARTWRIGHT, B. A., DOUBLEDAY, A., HIGGS, H., HUMMELINK, T., HUMMELINK-PETERS, B. G., KENNARD, O., MOTHERWELL, W. D. S., RODGERS, J. R. & WATSON, D. G. (1979). *Acta Cryst.* **B35**, 2331–2339.
BERNSTEIN, F. C., KOETZLE, T. F., WILLIAMS, G. J. B., MEYER, E. F. JR, BRICE, M. D., RODGERS, J. R., KENNARD, O., SHIMANOUCHI, T. & TASUMI, M. (1977). *J. Mol. Biol.* **112**, 535–542.
BLUNDELL, T., CARNEY, D., GARDNER, S., HAYES, F., HOWLIN, B., HUBBARD, T., OVERINGTON, J., SINGH, D. A., SIBANDA, B. L. & SUTCLIFFE, M. (1988). *Eur. J. Biochem.* **172**, 513–520.
BRÄNDÉN, C. -I. & JONES, T. A. (1990). *Nature (London)*, **343**, 687–689.
BRÜNGER, A. T. (1988). *X-PLOR* Version 1.3. *Crystallographic Computing* 4, edited by N. W. ISAACS & M. R. TAYLOR, pp. 126–140. Oxford Univ. Press.
BRÜNGER, A. T. (1992). *Nature (London)*, **355**, 472–475.
ELANGO, N. & PARTHASARATHY, S. (1990). *Acta Cryst.* **A46**, 495–502.
ENGH, R. A. & HUBER, R. (1991). *Acta Cryst.* **A47**, 392–400.
HUBBARD, T. J. P. & BLUNDELL, T. L. (1987). *Protein Eng.* **1**, 159–171.
ISLAM, S. A., STERNBERG, M. J. E. & WEAVER, D. L. (1990). *Accuracy and Reliability of Macromolecular Crystal Structures*, edited by K. HENRICK, D. S. MOSS & I. J. TICKLE, pp. 53–62. Proceedings of the CCP4 Study Weekend, 26–27 January 1990, Daresbury Laboratory, Warrington, England.
IUPAC–IUB COMMISSION ON BIOCHEMICAL NOMENCLATURE (1970). *J. Mol. Biol.* **52**, 1–17.
JONES, T. A., ZOU, J. -Y., COWAN, S. W. & KJELDGAARD, M. (1991). *Acta Cryst.* **A47**, 110–119.
KABSCH, W. & SANDER, C. (1983). *Biopolymers*, **22**, 2577–2637.
KRAULIS, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.
LÜTHY, R., BOWIE, J. U. & EISENBERG, D. (1992). *Nature (London)*, **356**, 83–85.
LUZZATI, P. V. (1952). *Acta Cryst.* **5**, 802–810.
MORRIS, A. L., MACARTHUR, M. W., HUTCHINSON, E. G. & THORNTON, J. M. (1992). *Proteins*, **12**, 345–364.
RAMACHANDRAN, G. N., RAMAKRISHNAN, C. & SASISEKHARAN, V. (1963). *J. Mol. Biol.* **7**, 95–99.
READ, R. J. (1986). *Acta Cryst.* **A42**, 140–149.
WIERENGA, R. K., KALK, K. H. & HOL, W. G. J. (1987). *J. Mol. Biol.* **198**, 109–121.
WISTOW, G., TURNELL, B., SUMMERS, L., SLINGSBY, C., MOSS, D., MILLER, L., LINDLEY, P. & BLUNDELL, T. (1983). *J. Mol. Biol.* **170**, 175–202.

J. Appl. Cryst. (1993). **26**, 291–294

The FROG PC series: programs for electron-density and model investigations for proteins. By E. A. VERNOSLOVA and V. YU. LUNIN, *Institute of Mathematical Problems of Biology, Russian Academy of Sciences, Pushchino, Moscow Region* 142292, *Russia*

(*Received* 2 *April* 1992; *accepted* 2 *September* 1992)

Abstract

A set of computer programs, developed for IBM-compatible personal computers and aimed at crystallographic use, is described. The programs have user-friendly interfaces and allow the calculation of various Fourier syntheses, which can be vizualized and compared. The possibility of obtaining a synthesis and an atomic model together and performing the model image rotations and translations with respect to the synthesis also exists.

Introduction

Computer calculations are involved in all stages of X-ray structure determination. The present tendency is towards the use of efficient computers and supercomputers to perform laborious calculations and graphic stations to analyze and interpret results. At the same time, the constantly growing possibilities offered by personal computers (PCs) allow the performance of some investigative steps by means of relatively cheap and widely available equipment. This arouses interest in the development of crystallographic programs aimed at the facilities of personal computers.

0021-8898/93/020291-04$06.00

This paper briefly describes some of the PC programs developed at the Institute of Mathematical Problems of Biology (Research Computing Center) of the Russian Academy of Sciences. A full description of the programs is available from the authors on request. The programs use IBM-compatible personal computers with VGA or EGA graphic-display adapters running MS-DOS version 3.3 or higher.

The program organizations have similar features. Each has a built-in window editor to set the parameters managing a program session. The current parameter values can be saved in a special file created by a program at the end of a session and used when starting a new program session. The programs have the means to control the external files, which allows the detection and correction of errors introduced when entering file names. A user can obtain the short prompt at any time.

1. *FAN* (scalar field analysis)

The *FAN* program is designed for the visual investigation of functions that depend on three variables (in crystallography, as a rule, electron-density Fourier synthesis or

© 1993 International Union of Crystallography