Fourier Analysis and the Structure of DNA

Three of the following papers are comments on an article on Fourier analysis and the structure of DNA [Science **165**, *1091 (1969)*]. *The fourth paper is a reply.*—EDITOR

Some Misconceptions on Fourier Analysis and Watson-Crick Base Pairing

Donohue (*1*) criticizes our x-ray diffraction studies of DNA structure and, especially, our use of Fourier syntheses (*2, 3*). He asserts that we are wrong to conclude that the Watson-Crick base-pairing scheme is clearly preferable to alternative schemes proposed by him (*4*) or by Hoogsteen (*5*). The structure of DNA has unique importance in biology; therefore we agree that it would be most desirable that, if there are weaknesses in the evidence for the generally accepted structure, these be brought to light. However, in this connection Donohue's article is not useful. It misleads, not only by direct statement, but by implication and by omission.

The article is concerned with the very real limitations in the x-ray data from DNA. What is misleading is Donohue's implication, despite our published statements (*2, 3*), that we have not adequately taken these limitations into account in our studies. He claims that Fourier analysis in the form of $|F_0|$ syntheses would not form a satisfactory basis for our conclusions. He implies that we based our results only on $|F_0|$ syntheses and claims therefore that our conclusions are unjustified. But we did not base our conclusions on $|F_0|$ syntheses; we based them on difference syntheses. Donohue makes no mention of these.

We give here some brief comments on Donohue's article. A more technical criticism is presented by one of us (*6*).

Donohue presents, as electron density maps, $|F_0|$ syntheses of various base pairs, pointing out that such syntheses do not enable one easily to discriminate between alternative models. His examples show this clearly. Examples as good have already been provided by us (*7*). The well-known principle reiterated by Donohue—that a synthesis involving observed intensities and phases derived from a model is weighted toward the model—has also been stated explicitly by us (*3*). The important consideration is that the difficulty of discriminating between alternative models is greatly reduced by using difference syntheses, a technique widely used where sensitivity is required in Fourier studies (*8*). Although a difference synthesis contains no more information than the corresponding $|F_0|$ and model syntheses, it greatly aids recognition of their differences. As a result one can see clearly that Watson-Crick pairs agree better with DNA diffraction data than do Hoogsteen pairs (*3, 6*).

Donohue, by confining his attention to $|F_0|$ syntheses, has also erred in previous work. In his study (*9*) of the structure of 2-phenylazulene, he concluded that the x-ray data cannot be used to discriminate between the two possible structures. It is shown, however (*6*), that in this case difference syntheses make discrimination possible when $|F_0|$ syntheses do not.

Donohue discusses the general question of proof of the DNA structure but says nothing new. He says, as we have often said before (*10*) that, because DNA is only available in microcrystalline fibers, conclusive proof of the correctness of the DNA structure has not been obtained as directly as is generally the case with single crystals. But, as we also emphasized, the disadvantage of having only fiber data for DNA is to a fair extent overcome by DNA having several conformations—all providing data in good agreement with the Watson-Crick scheme. Moreover, we have shown (*3*) that, in spite of their low resolution, the data are adequate to discriminate between the electron density distributions of different base-pairing schemes. The Hoogsteen scheme shows discrepancies, and these are of the kind expected if the Watson-Crick scheme is correct. As Donohue says, it is undesirable to use Fourier syntheses to prove directly that a structure is correct. The important thing he should explain is that they can readily show that the alternative structures are incorrect, thereby providing proof by elimination.

We might also add that we have frequently emphasized the limitations of the x-ray data from DNA, in particular that the resolution is insufficient to resolve single atoms. To make this clear we have published (*2, 3*) $|F_0|$ synthesis maps to show the resolution directly. These are, we think, more useful than Donohue's attempts to give a numerical value to the resolution. It is unfortunate that Donohue, in publishing such an $|F_0|$ map as the sole example of our work, gives the impression that we used only such maps to derive the structure. Also, Donohue's comparison (*1*) of Fourier maps of a Watson-Crick and of a particular Donohue pair could be very misleading. The electron density distributions in both pairs are similar, but he does not draw attention to the fact that the glycosidic bonds are in different positions; as a result the radical difference of the pairs as they would appear in DNA is concealed. We further note that the caption of his Fig. 6 appears to be wrong, the phases being supplied for pairing H and not D.

In conclusion, we briefly restate our position regarding base pairs in DNA. We, and presumably others, have had no success in building stereochemically reasonable DNA models with Donohue pairs: the stage has never been reached where the addition of Fourier studies was thought necessary. This, of course, in no way detracts from the contribution, widely acknowledged [for example, in J. D. Watson, *The Double Helix* (Atheneum, New York, 1968)], made by Jerry Donohue in the original formulation of base-pairing in DNA. The Hoogsteen scheme had more promise of being an alternative to Watson-Crick but did not survive subsequent tests. As well as the Hoogsteen Fourier studies (*3*), it has not been possible to build

Hoogsteen models (*3*) with good stereochemistry, and there is spectroscopic evidence (*11*) against the existence of the tautomeric form of cytosine required for this pairing. There appears therefore to be no alternative to accepting the Watson-Crick scheme.

M. H. F. WILKINS
S. ARNOTT
Medical Research Council Biophysics Unit, London, W.C.2, England
D. A. MARVIN
Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06520
L. D. HAMILTON
Medical Research Center, Brookhaven National Laboratory, Upton, New York 11973

References and Notes

1. J. Donohue, *Science* **165**, 1091 (1969).
2. S. Arnott, *Bull. Inst. Phys. Soc. London* **15**, 169 (1964); D. A. Marvin, M. H. F. Wilkins, L. D. Hamilton, *Acta Cryst.* **20**, 663 (1966).
3. S. Arnott, M. H. F. Wilkins, L. D. Hamilton, R. Langridge, *J. Mol. Biol.* **11**, 391 (1965).
4. J. Donohue, *Proc. Nat. Acad. Sci. U.S.* **42**, 60 (1956).
5. K. Hoogsteen, *Acta Cryst.* **12**, 822 (1959).
6. S. Arnott, *Science*, this issue.
7. Figures 9 and 10 in Arnott (*2*) and figures 2a, 4a, and 5a in reference (*3*).
8. Textbook references are H. Lipson and W. Cochrane, *The Determination of Crystal Structures* (Bell, London, 1957), pp. 202–204 and 300–305; G. H. Stout and L. H. Jensen, *X-ray Structure Determination, a Practical Guide* (Macmillan, London, 1969), pp. 354 and 356–362.
9. B. D. Sharma and J. Donohue, *Nature* **198**, 878 (1963).
10. R. Langridge, D. A. Marvin, W. E. Seeds, H. R. Wilson, C. W. Hooper, M. H. F. Wilkins, L. D. Hamilton, *J. Mol. Biol.* **2**, 38 (1960).
11. H. T. Miles, *Proc. Nat. Acad. Sci. U.S.* **47**, 791 (1961).

DNA: Test of Structure?

Donohue (*1*) has argued that the DNA x-ray data of Wilkins and his colleagues might fit just as well a model with alternative base-pairing. Wilkins, Arnott, Marvin, and Hamilton have replied to this (*2*) by pointing out that Donohue has not appreciated the power of difference syntheses. In particular, they state that the Fourier obtained with the use of phases calculated from a Hoogsteen scheme show discrepancies of the kind expected if the Watson-Crick scheme is correct. I am, of course, a biased witness but it seems to me that Donohue has not yet made his case. Moreover, he has not allowed for the fact that (owing to the numerous dyads in the structure) many of the reflections of the pattern are effectively centric in character. In centric structures his arguments have much less force.

It seems to me that a long, involved, and possibly acrimonious theoretical argument on these issues would be fruitless. If Donohue thinks that an equally effective model for DNA could be produced with some alternative base-pairing, let him build such a model and publish the coordinates. The fit of this model with the observed x-ray data could be compared with that of the models described by Wilkins and his colleagues. We would then all see which model fits the data better, or whether there is nothing to choose between them. The King's College group have already tried to do this and failed, but it could be argued that they may be subconsciously prejudiced against such models and may have overlooked a possible solution. Such an argument could hardly be applied to Donohue, and in fact I can think of no one better qualified to make the attempt. I agree that it would involve him in a fair amount of work, but I see no other way of deciding the matter.

F. H. C. CRICK
Medical Research Council, Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, England

References

1. J. Donohue, *Science* **165**, 1091 (1969).
2. M. A. F. Wilkins, S. Arnott, D. A. Marvin, L. D. Hamilton, *ibid.*, this issue.

Crystallography of DNA: Difference Synthesis Supports Watson-Crick Base Pairing

The inability of Donohue to demonstrate rejection of an incorrect structural hypothesis, using Fourier syntheses of electron density in a number of cases (*1, 2*), has prompted him to make far-reaching conclusions implying the futility of using this method in x-ray analysis of structure. His pessimism is not confined to situations where, as with DNA, only low-resolution data are available ". . . the Fourier method of structure refinement has, in fact, contributed nothing toward either the proof of DNA structure, nor toward the elucidation of its details . . ." (*2*), but extends more widely. "It might be thought that this failure of the Fourier method to reject an incorrect hypothesis resulted from the use of low-resolution data. This conjecture is in fact untrue . . ." (*2*). Since in the case of DNA many experiments have been described (*3, 4*) that contradict Donohue's opinion but were not discussed by him, and since it will be shown below that the examples presented by him (*1, 2*) do not support his contention when subjected to thorough analysis, I feel it important to refute his conclusion lest others should be discouraged from using the Fourier method as one of a number of important weapons in the armory of structure analyzers.

Fourier Method in X-ray Analysis

The main difficulty in using the Fourier method, correctly highlighted by Donohue (*2*), is that only the amplitudes, $|F_{\mathbf{H}}|$, and not the phases, $\alpha_{\mathbf{H}}$, may be available experimentally to be used in calculating electron density ρ from

$$\rho(\mathbf{r}) = (2/V) \sum_{\mathbf{H}} | F_{\mathbf{H}} | \cos (2\pi \mathbf{H} \cdot \mathbf{r} - \alpha_{\mathbf{H}})$$

Moreover when phases calculated from an only partially correct, provisional structure are combined with experimental amplitudes, the resulting electron density map usually more closely resembles that of the provisional structure rather than that of the structure that would provide the amplitudes. This is a crystallographic commonplace, well known to us and excellently demonstrated by our Fourier experiments with different DNA models (*3, 4*).

If, however, this procedure of using "true" amplitudes with phases from a provisional structure always merely resulted in the same density as that of the provisional model crystallographers would be out of business. The science of crystallographic structure determination, particularly in the 30 years just past, has flourished mainly because this, in fact, is fortunately not true. Perhaps the most familiar instance of this occurs when the provisional phases have been calculated from the previ-

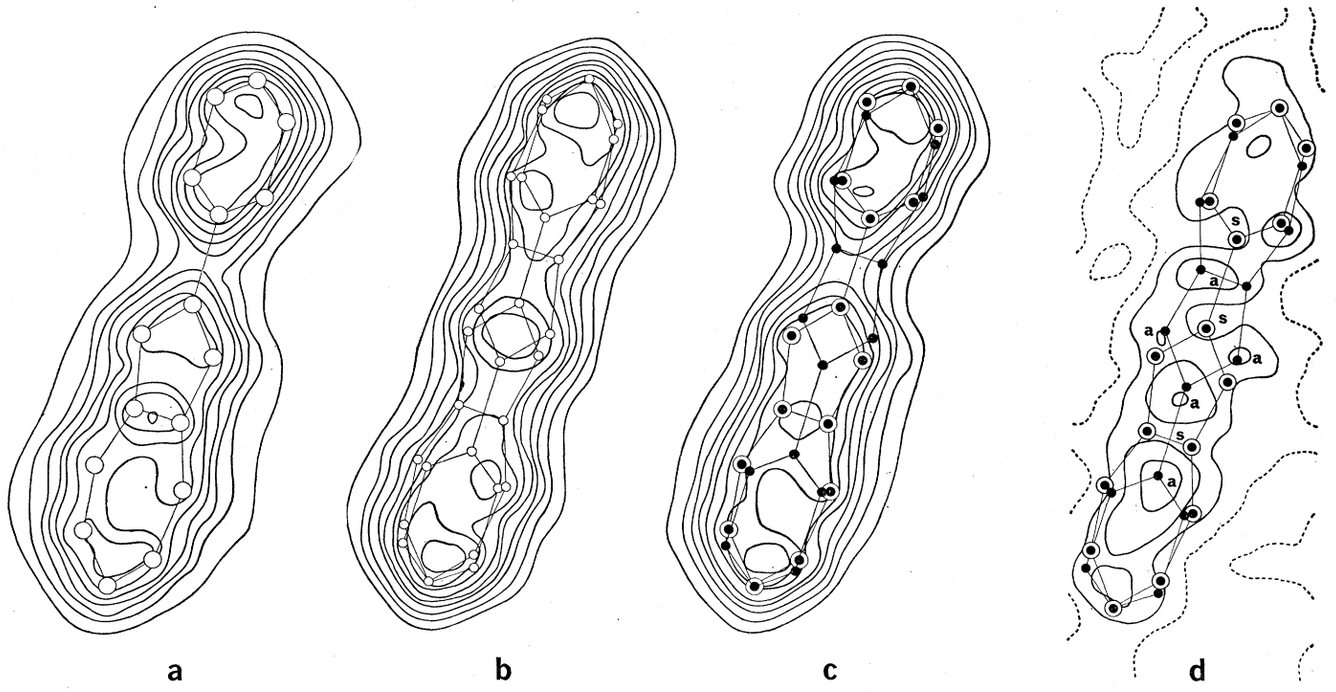


Fig. 1. Electron density syntheses for (a) 2-phenylazulene; (b) an alternative crystal structure. In preparing these syntheses phases *and* amplitudes were provided by the superposed (open-circle) models; (c) the ρ_{obs} synthesis where the phases were derived from the model in part a, but the amplitudes from the model in part b (closed circles); (d) the difference between parts c and a. In parts a, b, and c the lowest contour is 1 electron/Å^2 and the contour interval 0.5 electron/Å^2. In part d the negative contours are shown by broken lines; the contour interval is everywhere 0.25 electron/Å^2; the zero contour is not shown.

ously determined positions of a few "heavy" atoms in the structure. It is usually found that combination of these phases with the true amplitudes results in an electron density map portraying not only the atoms providing the phases but also minor maxima indicating the positions of some, if not all, of the remaining atoms in the structure. The provisional model can then be augmented by the new atoms, better phases can be calculated, and the process can be repeated as often as deemed necessary. Alternatively the provisional model may be defective not in the number of atoms it contains but in the positions and shapes of its atoms. Once again it is often possible to determine from the provisional density map the manner in which the model should be amended.

Such progress depends on the Fourier synthesis ρ_{obs}, prepared from the true amplitudes and provisional phases, not being identical with the electron density of the provisional model ρ_{model}. It has been noted above and elsewhere that ρ_{obs} and ρ_{model} are generally similar; therefore the obvious procedure to adopt is to examine the difference synthesis (*5*), $\Delta\rho = (\rho_{obs} - \rho_{model})$ to concentrate attention not on the similarities but on the differences between ρ_{obs} and ρ_{model}. The difference synthesis also has the advantage that the series termination errors common to both ρ_{obs} and ρ_{model} are diminished, a helpful property in DNA studies where, for a variety of experimental reasons, the Fourier series are more than usually prematurely terminated.

When the provisional model is correct, the $\Delta\rho$ map is relatively featureless apart from minor fluctuations occasioned by errors in the experimental amplitudes, but when it is only partially correct the $\Delta\rho$ synthesis shows a pattern of positive and negative electron density with the less correct features of the provisional model in hollows or on steep slopes. An improved model is one that provides more electron density on or toward the peaks in $\Delta\rho$ and away from the hollows.

2-Phenylazulene

Use of the difference synthesis can be illustrated best by an example, and for this I have chosen a molecule where alternative crystal structures were possible and for which Donohue found his version of Fourier analysis inadequate to distinguish between the alternatives (*1*).

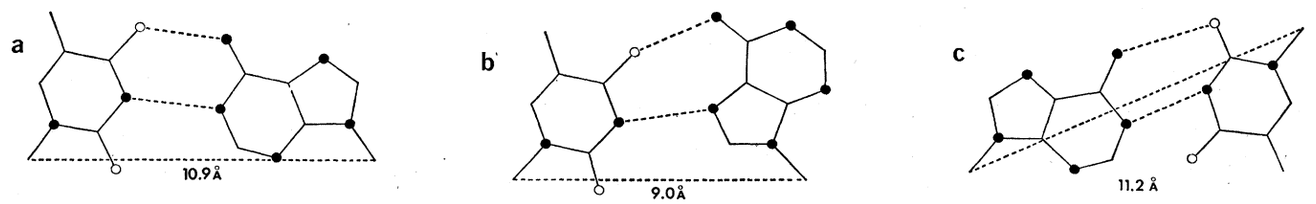


Fig. 2. The three base-pairing schemes designated (a) W, (b) H, (c) D. The hydrogen-bonds are shown and also the line (and its length) between the carbon atoms of the two glycosidic bonds in each pair. For the electron density in a D pair to most closely resemble that of a W pair the interglycosidic line has to be oriented as shown.

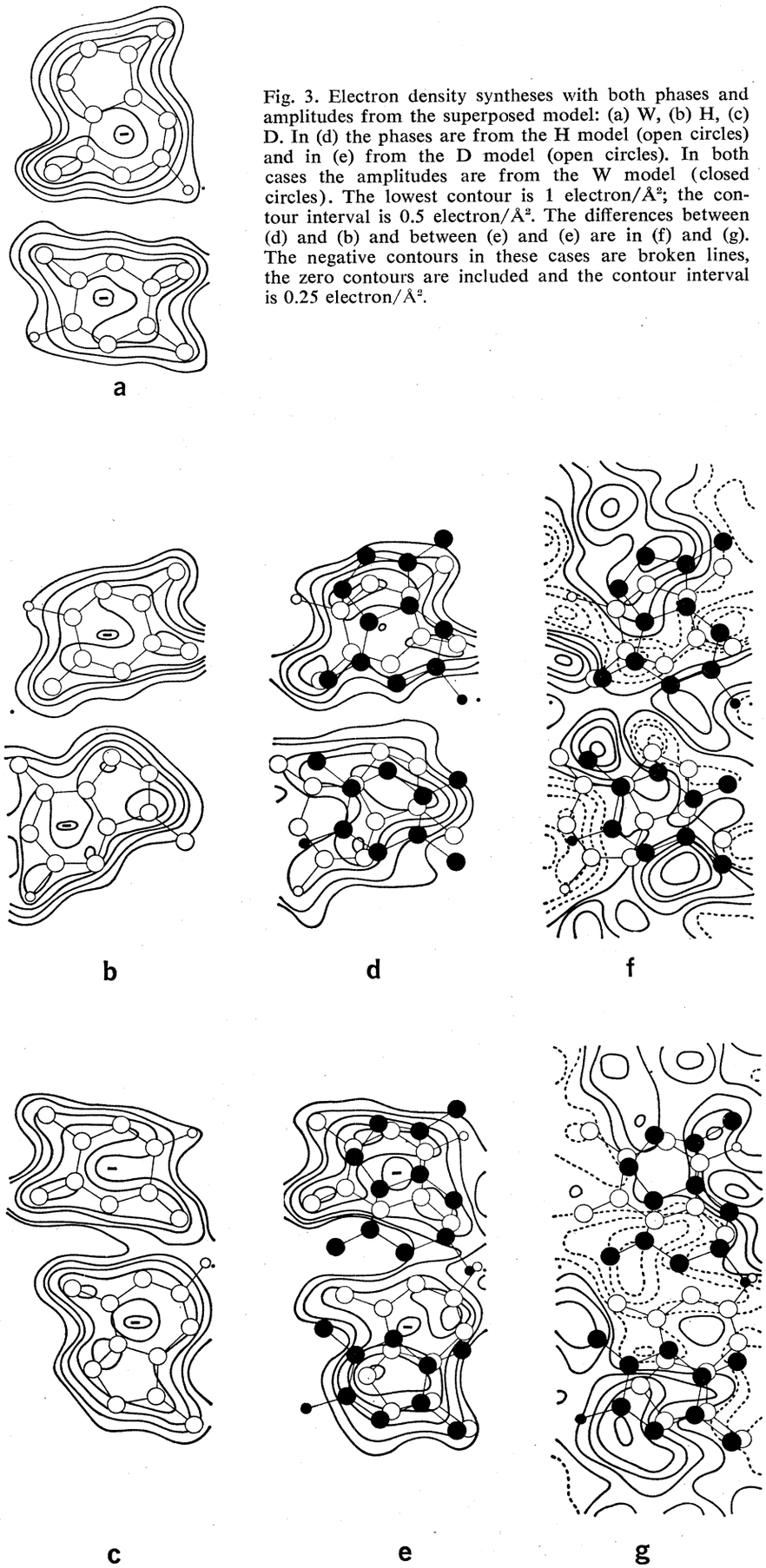


Fig. 3. Electron density syntheses with both phases and amplitudes from the superposed model: (a) W, (b) H, (c) D. In (d) the phases are from the H model (open circles) and in (e) from the D model (open circles). In both cases the amplitudes are from the W model (closed circles). The lowest contour is 1 electron/Å^2; the contour interval is 0.5 electron/Å^2. The differences between (d) and (b) and between (e) and (e) are in (f) and (g). The negative contours in these cases are broken lines, the zero contours are included and the contour interval is 0.25 electron/Å^2.

2-Phenylazulene was conjectured (*1*) either to have the structure shown in projection in Fig. 1a or the statistical structure shown in Fig. 1b. For the purposes of this test I have assumed that the statistical structure is indeed the true structure and will therefore use, as the "experimental data," amplitudes calculated from it. If the provisional model is assumed to be the nonstatistical structure, phases and thence ρ_{obs} (Fig. 1c) can be calculated (*6*). We note that this ρ_{obs} more closely resembles Fig. 1a (the map of the provisional model) than Fig. 1b (the map of the true structure). Like Donohue (*2*), I would agree that this similarity was no proof that the phasing model was correct, but I would not conclude that I could not reject it as wrong before examining $\Delta\rho$ (Fig. 1d), the difference between the distributions in Figs. 1, c and a. This difference map has pronounced features indicating that a new model is required that would have more electron density in the regions marked add and less in the regions marked subtract. The former regions correspond closely with features of the true structure (the statistical model) not represented in the provisional model, and the latter with features overrepresented in the model providing the phases. Even in those parts of the structure where the provisional model has provided a full atom, rather than the two half-atoms of the true structure, the minor adjustments necessary are indicated by the fact that the half-atoms not coincident with the atomic positions of the provisional model lie uphill from these positions.

Donohue (*1*) was therefore correct in concluding that his version of Fourier analysis (that excludes difference syntheses) was inadequate to distinguish alternate structural hypotheses, but wrong in concluding that no other version of the method might do so.

Purine-Pyrimidine Pairs at Low Resolution

Donohue has also explored (*2*) the possibility of distinguishing among three purine-pyrimidine pairs, W, H, and D (Fig. 2) that have been advanced as base-pairing schemes for DNA and other duplex nucleic acids with complementary helices (*7*). Figure 3, a to c, shows the true electron density distributions for these when terms cor-

responding to periodicities less than 2 Å have been omitted. There it can be seen that, with only insignificant changes (*8*), the situations discussed by Donohue (*2*) are being reproduced. Like him, I have calculated ρ_{obs} for the H (Fig. 3d) and D (Fig. 3e) cases, assuming that the true structure is W (that is, in both cases the amplitudes, but not the phases, derive from W). Like him I am struck, but not surprised, by the similarity of these to the electron density distributions of the models providing the phases (Figs. 3, b and c). Unlike him, I would not deduce that therefore the hypotheses of H or D pairing could not be rejected. The lack of identity between the ρ_{obs} and the electron density distributions of the provisional models is clearly revealed in the $\Delta\rho$ maps (Fig. 3, f and g) ignored by Donohue.

The Donohue doctrine (of the inability of Fourier methods to reject an incorrect hypothesis) rests on the indistinguishability between Fig. 3, b and d, and between Fig. 3, c and e. Certainly it is difficult visually to detect the changes in gradient, and shifts in peak positions that show that these pairs of maps are far from identical. But the difference syntheses in Fig. 3, f and g, not exploited by Donohue (*2*), reveal that the ρ_{obs} maps have significantly failed perfectly to reproduce the electron density of the models providing the phases. Nor are the features in these $\Delta\rho$ maps of trivial magnitudes: the maximum heights in the ρ_{obs} maps are about 3.3 electron/Å^2; in the $\Delta\rho$ maps the error-indicating features are not only extensive in area but achieve amplitudes of $\pm$ 1.3 electron/Å^2.

It is noticeable that in these $\Delta\rho$ maps the most prominent changes of gradient occur where the atoms of the phasing models (open circles) should be redistributed to positions corresponding to those of the W model (filled circles). (Since these are low-resolution situations where individual atoms are not resolved, for workers familiar only with single-crystal analyses, a good analogy is to consider the bases as large "atoms" of complex—but known—shape, and to consider the $\Delta\rho$ maps as providing evidence of how these "atoms" should be reoriented in much the same way that one would determine the orientation of the axes of an atom subject to anisotropic thermal vibrations in high-resolution studies.)

Base Pairing in DNA

So far I was concerned to demonstrate that essentially the same model schemes that Donohue discusses (*2*) are indeed distinguishable by Fourier methods. I now turn to versions of these pairing schemes more likely to be relevant for DNA and more closely corresponding to the models we considered when analyzing the experimental data from DNA itself.

All the indications are that DNA has a statistical crystal structure in the sense that the diffraction corresponds to a situation where either purine or pyrimidine may occupy any base site and where each molecular site is occupied by a constellation of atoms that looks the same when turned upside down (*9*). This situation is neatly achieved by the identical, antiparallel, sugar phosphate chains that are possible with W and H pairs, where the glycosylic links are related by a dyad axis perpendicular to the helix axis. Statistical base pairs for such DNA molecular models are shown in Fig. 4, a and b. For D pairs the situation in DNA would be more complex and the necessary statisticality would be attained only at the cost of complex-looking average base structures like that in Fig. 4c.

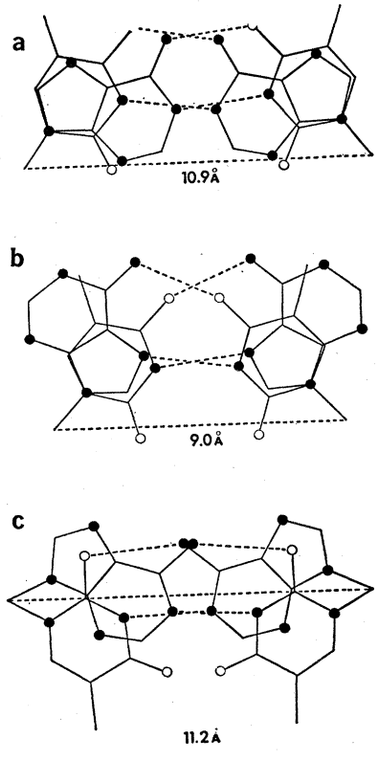


Fig. 4. Statistical base-pairing schemes for DNA involving (a) W pairs, (b) H pairs, (c) D pairs.

With the base-pairing schemes shown in Fig. 4, I conducted experiments similar to the previous ones. The model electron density distributions are shown in Fig. 5, a–c. The values for ρ_{obs}—W amplitudes but phases from statistical H and D models were used—are in Fig. 5, d and e. The nature and extent of the differences between W and H and between W and D pairs, which can be obtained in spite of the bias toward the phasing models, is shown in Fig. 5, f and g. In both cases, these differences are substantial. Further, the difference between W and H pairs in these model systems shows features (for example, the central ridge and flanking troughs) similar to those seen when we used (*4*) the experimental amplitudes from DNA itself and a DNA model with H pairs as the source of phases (Fig. 6, a and b). Also, there are no features in the analogous $\Delta\rho$ map, where a DNA model is used with W pairs for phasing (Fig. 6c), which would suggest that the DNA data support either H or D pairs. The minor residual features there can all be explained by the need to develop a more adequate representation of the large number of water molecules and counterions known to be present between the DNA molecules.

Potential of DNA Data in Distinguishing Models

Donohue has commented pejoratively (*2*) on the values for DNA of the average discrepancy (R) between the observed structure amplitudes, $|F_{obs}|$ and those calculated on the basis of a proposed structure, $|F_{calc}|$.

$$R = \sum \big| |F_{obs}| - |F_{calc}| \big| \Big/ \sum |F_{obs}|$$

These values (0.3 to 0.4) are high compared with those (0.05 to 0.2) obtained in high resolution single-crystal studies. The accuracy of the intensities (proportional to $|F_{obs}|^2$) measured by the densitometer from fiber diagrams of quality is certainly not much inferior to visually estimated intensities from single-crystal diffractograms, and therefore one would expect to find that the best values of R obtainable were about 0.2.

That the best values attained for nu-

cleic acids are higher than this is mainly due to errors in $|F_{calc}|$! All nucleic acid crystals contain water and counterions amounting to some 40 percent by weight of the whole structure. These additional scattering features do not appear to be well-ordered (*10*), and it is computationally convenient to derive the $|F_{calc}|$ assuming that each atom in the DNA molecule is surrounded by a spherically symmetric cloud of additional scattering material to an extent determined by the hydrophilicity of each atom. Clearly a better model would not place this additional density around every atom but only on the exterior of the DNA molecule. Indeed, in a model where this was done (*10*) R was reduced to 0.24. It is relevant in this connection to mention that after the structure refinements of α- and β-poly-L-alanine and poly-L-proline II, crystalline fiber systems with no solvent problems, the values of R were respectively 0.21, 0.15, and 0.08 (*11*).

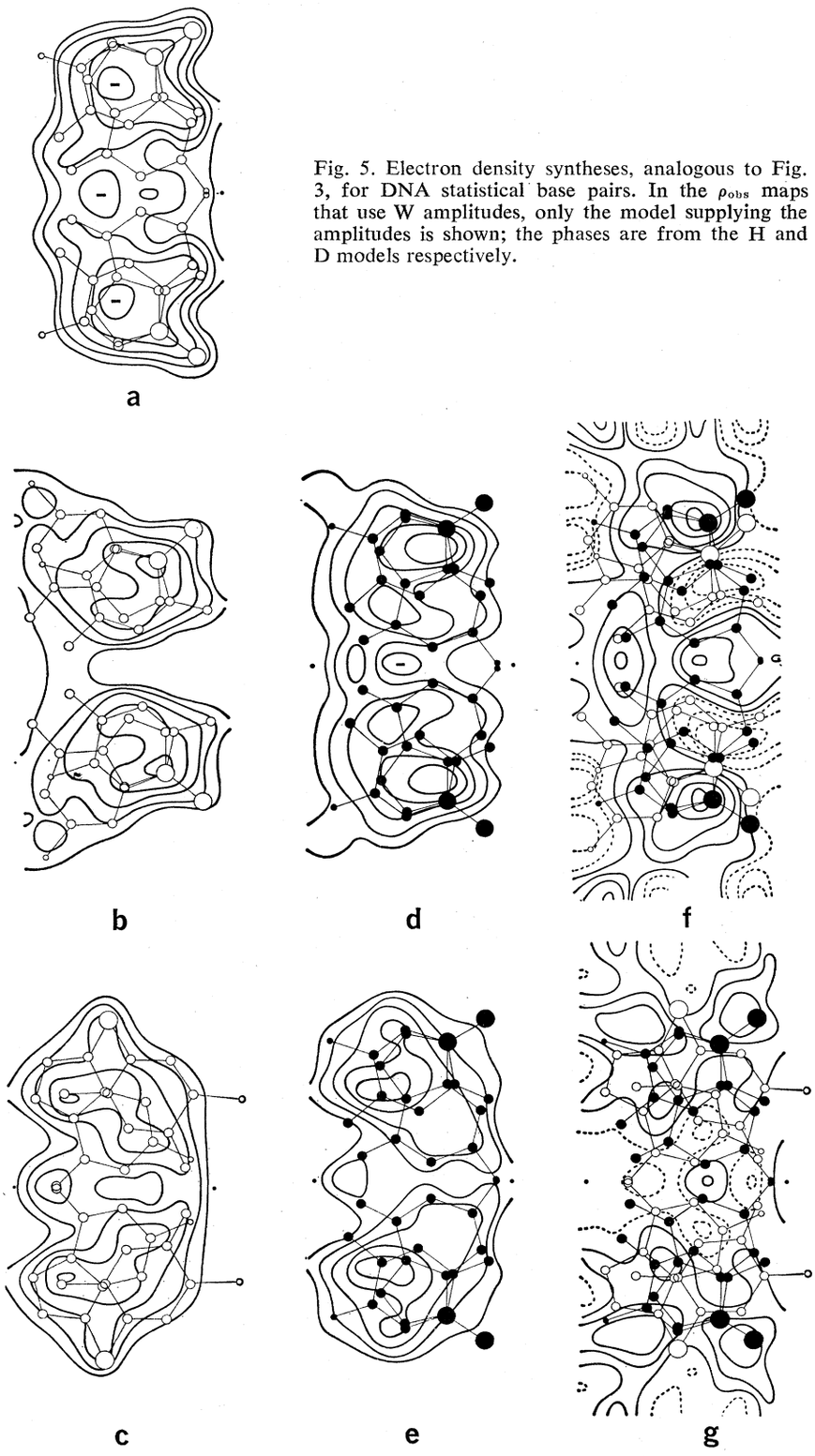


Fig. 5. Electron density syntheses, analogous to Fig. 3, for DNA statistical base pairs. In the ρ_{obs} maps that use W amplitudes, only the model supplying the amplitudes is shown; the phases are from the H and D models respectively.

Another factor tending to inflate R values for nucleic acids is the high symmetry assumed for these molecules. In B-DNA, for example, the atoms of one average nucleotide are assumed to be precisely related to those in other nucleotides either by a twofold axis or a tenfold screw axis, or both. Any minor errors in the atomic positions of the standard residue are therefore systematically repeated throughout the molecule and will give rise to higher R values for the same reason that centrosymmetric structures have characteristically higher R values than similar structures with no symmetry (*12*).

It should not be forgotten that R tends to increase with the number of terms included (M) and to decrease with increasing number of degrees of structural freedom (N). Contemporary refinements of nucleic acids are made with the use of linked-atom systems (*13*) that have standard bond lengths and angles maintained, and that are further constrained to have the correct molecular symmetry and any given base-pairing scheme, with the result that the number of degrees of freedom allowed for the crystal structure is most often 5 ± 1. The number of x-ray reflections is usually 100 to 200 so that the ratio of data to parameter, M/N, can be as high as 50 and is never less than 16. The values of $R \sim 0.05$ achieved in single-crystal structures are obtained with, say, 1000 reflections but 200 parameters (individual atom positions and anisotropic thermal vibration parameters), a data-parameter ratio of only 5.

All these factors that tend to inflate nucleic acid R values in contrast to those obtained in more conventional crystal systems are, of course, similar for alternative DNA models, and therefore it is more useful to compare the best values obtained for each DNA system rather than to appeal to an absolute standard. In this connection it is relevant that I have developed

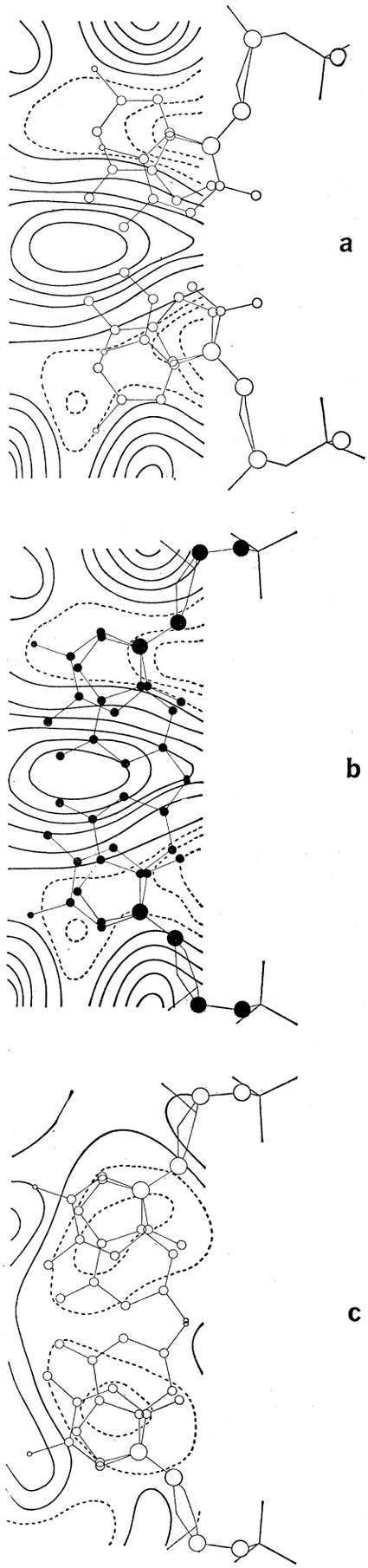


Fig. 6. (a and b) A section of a three-dimensional $\Delta\rho$ map (*3*, *4*) where ρ_{obs} was computed using DNA experimental amplitudes and phases from a DNA model with H pairs. The atoms of the phasing model in or near this section are shown (open circles) in (a). A DNA model with W pairs that would remedy the deficiencies revealed is shown (closed circles) in (b). (c) A similar $\Delta\rho$ map where the DNA model (superposed) has the W pairs. In (a), (b), and (c) only the area corresponding to the model maps in Fig. 5 is contoured.

an "even-handed" method for comparing different structural hypotheses for nucleic acids. This linked-atom refinement (*14*) scheme produces that model of any particular type that provides a least-squares best fit between the observed and calculated amplitudes. The extent to which different best models with the same number of parameters fit the experimental data can be readily and objectively determined. For B-DNA the best W-containing model (*13*) has $R = 0.35$, but the best H- and D-containing models (*15*) have $R > 0.50$. For double-helical RNA the best W model (*13*) has $R = 0.33$, but the alternative schemes give $R > 0.40$ (*15*).

The gross differences in R between different models imply that these nucleic acid x-ray data are more than adequate to distinguish between models of different types and that both for DNA and for RNA the type favored possesses W base pairs.

Conclusion

The types of models that need to be considered for DNA are few in number, and each has no more than half a dozen degrees of freedom. There are more than 200 experimental amplitudes for DNA, of a quality comparable with visually estimated single crystal amplitudes, so that the problem is well overdetermined, if the different types of models are sufficiently different from one another. Donohue has ignored the one component of the Fourier method (difference syntheses) which clearly shows that the alternative base-pairing schemes are sufficiently different from one another for the DNA data successfully to arbitrate between them in spite of phasing bias. Although the experiments presented above support many of Donohue's general, cautionary statements on the use of the Fourier syntheses approach (*2*), they contradict his conclusions about DNA. These experiments, combined with our previous work on DNA with the difference syntheses method, show that a decision can be made in favor of a DNA model of the Watson-Crick type. This is not only because Fourier analysis with phases from such models do not lead to inconsistencies but also because, in contrast, Fourier syntheses with phases from other models invariably lead one to conclude that they are grossly inadequate.

Summary

Examination of the Fourier method of crystal structure analysis, in which the distribution of electron density is calculated with the observed structure amplitudes combined with phase angles obtained from an assumed model and in which the differences between this distribution and that of the phasing model are thoroughly examined can result in the rejection of incorrect structural hypotheses even when only low resolution data are available.

STRUTHER ARNOTT
Medical Research Council Biophysics Unit, London, W.C.2, England

References and Notes

1. B. D. Sharma, J. Donohue, *Nature* **198**, 878 (1963).
2. J. Donohue, *Science* **165**, 1091 (1969).
3. S. Arnott, *Bull. Inst. Phys. Soc. (London)* **15**, 169 (1964).
4. S. Arnott, M. H. F. Wilkins, L. D. Hamilton, R. Langridge, *J. Mol. Biol.* **11**, 391 (1965).
5. Difference syntheses were suggested by A. D. Booth, *Nature* **161**, 765 (1948) and soon afterward exploited by C. W. Bunn, in the form of "error syntheses" [*The X-ray Crystallographic Investigation of the Structure of Penicillin* (Princeton University Press, Princeton, N.J., 1949), pp. 310–367]; in this work he used them to eliminate an incorrect structural hypothesis. More conventional use of the syntheses is described in elementary textbooks, for example M. J. Buerger's *Crystal-Structural Analysis* (Wiley, New York, 1960), pp. 697–8.
6. In this experiment the unit rectangle has dimensions $a = 12$ Å and $b = 4$ Å. The atomic form factors for carbon are from *International Tables for X-ray Crystallography*. Hydrogens are ignored. In the statistical structure half-weight carbon atoms are used. All $|F_H|$ are included corresponding to periodicities > 1 Å. The temperature factor has $B = 5$ Å^2.
7. The W pairs are a vital feature of the DNA structural hypothesis → J. D. Watson and F. H. C. Crick, *Nature* **171**, 737 (1953); the H pairs were first observed by K. Hoogsteen, *Acta Cryst.* **12**, 282 (1959), and are the kind most commonly observed when adenine and thymine derivatives cocrystallize; the D pairs were discussed → J. Donohue, *Proc. Nat. Acad. Sci. U.S.* **42**, 60 (1956), and by J. Donohue and K. N. Trueblood, *J. Mol. Biol.* **2**, 363 (1960).
8. In the experiments with base pairs, the unit rectangles have the same dimensions, $a = 16$ Å, $b = 8$ Å. (The areas shown in the diagrams do not correspond to the complete unit rectangle.) Instead of adenine-thymine pairs, I use guanine-cytosine pairs with a half-weight guanine amino-group and with a

half-weight methyl added to cytosine to produce an average purine-pyrimidine pair for DNA. The atomic form factors are from *International Tables for Crystallography*; they are multiplied by weighting factors of ½ or ¼ where appropriate. The temperature factor has $B = 6$ Å^2 since this is more appropriate for DNA. (Those familiar with the diffraction from DNA models are aware that much of the sparseness of DNA data with periodicities less than 3 Å is not due only to attenuation factors that increase continuously with diffraction angle but to the fact that the molecular transform itself is small between 3 and 2 Å.) However we have repeated the experiments with $B = 15$ Å and 30 Å and reach the same conclusions as when $B = 6$ Å^2.
9. M. H. F. Wilkins, in *Biological Structure and Function* (Academic Press, New York, 1960), vol. 1, pp. 13–32.
10. D. A. Marvin, M. H. F. Wilkins, L. D. Hamilton, *Acta Cryst.* **20**, 663 (1966).
11. S. Arnott and A. J. Wonacott, *J. Mol. Biol.* **21**, 371 (1966); S. Arnott and S. D. Dover, *ibid*, **30**, 209 (1967); ———, A. Elliott, *ibid.*, p. 201; S. Arnott and S. D. Dover, *Acta Cryst.* **B24**, 599 (1968).
12. V. Luzzati, *Acta Cryst.* **5**, 802 (1952).
13. S. Arnott, S. D. Dover, A. J. Wonacott, *ibid.* **B25**, 2192 (1969).
14. S. Arnott and A. J. Wonacott, *Polymer* **7**, 157 (1966); S. Arnott, in *Symposium Fibrous Proteins, Australia, 1967* (Butterworths, New York and Sidney, 1968), pp. 25–41; also reference *11*.
15. Unpublished experiments of S. Arnott and D. W. L. Hukins.
16. From all $\Delta\rho$, terms with $|F_{obs}|/|F_{calc}| > 3$ were excluded.
17. I thank Professor Sir John Randall for facilities, Professor M. H. F. Wilkins for discussion, Mrs. Janet Rickard for assistance in preparing this article, and Z. Gabor for photography.

Fourier Series and Difference Maps as Lack of Structure Proof: DNA Is an Example

Wilkins, Arnott, Marvin, and Hamilton (*1*) have taken issue with some of the observations in my article on Fourier analysis (*2*). They refer to the DNA structure as being "generally accepted," as if such an important matter could be decided by an opinion poll. May I point out that enol formulation of the bases in nucleic acid was "generally accepted" save for a very small group of skeptical specialists. One of the points in my article is that proof of a proposed structure cannot be furnished by calculation of electron density distributions with a set of observed structure amplitudes and a set of phase angles based on the proposed structure. This feature of $|F_0|$ syntheses is well known to practicing crystallographers. One of the "omissions" in my article decried by Wilkins *et al.* is that there is no mention of difference syntheses, on which they say they based their conclusions. They then state "although a difference synthesis contains no more information than the corresponding $|F_0|$ and model synthesis, it greatly aids recognition of their differences." I wholly agree with the first part of this statement and point out that Wilkins *et al.* chose not to mention the fact that I did present both $|F_0|$ and model syntheses. An aid in the recognition of such differences for the uninitiated is found in a difference map, Fig. 1, which was, ironically, removed from an early draft of my manuscript on the advice of referees. There is very little, if any, indication in this map that the assumed structure, indicated by the crosses, is the "correct" one. The view of Wilkins *et al.* that the difference synthesis technique should be emphasized at the expense of $|F_0|$ and model syntheses is thus inappropriate.

Wilkins *et al.* refer to my "error" with regard to the structure of 2-phenylazulene, citing a more technical criticism by Arnott (*3*); my comments on his discussion form a later part of this paper. But I am led to wonder why Wilkins *et al.* believe that the 2-phenylazulene problem is germane to the present issue. They appear to imply that I am unaware of the properties of difference syntheses, a conclusion which displays a degree of unfamiliarity with certain portions of the literature (*4*). In any case, even if it is assumed that I made a previous error of omission with 2-phenylazulene, this in no way affects any of the arguments presented by me in (*2*).

The fact remains that Fourier syntheses have a distressing tendency always to give back what was put in. Thus, in addition to the sources cited (*2*), in the two sources on difference syntheses cited by Wilkins *et al.* we find: ". . . slight errors in any postulated atomic position will lead to slight errors in the phase angles as well as in the structure amplitudes, and Fourier syntheses computed with these values of phase angles will tend to reproduce the atomic positions from which they were derived" (*5*, p. 194); "Unfortunately, the presence of an atom in the model tends to bias the phasing to favor it . . ." (*6*). Wilkins *et al.* are aware that "Fourier syntheses are inevitably biased toward the structural model from which the phases are derived" (*7*). Figure 1 shows just how extreme that bias can be.

In any case, discussions of the Fourier method as applied to crystal structure analysis would carry more force were they made by those with considerable experience in the use of that method on a large number of different structures.

Basis of Structure Proof

Some of the questions raised by Arnott are not relative to the question of proof of structure based on Fourier analysis, which was the subject of my article (*2*). A basic difference in the philosophy of what constitutes proof of structure with x-ray data has been laid bare. One view held, for example, by Arnott (*3*) and by Wilkins, Arnott, Marvin, and Hamiliton (*1*), is that it is sufficient to consider various models, and then to choose as correct (after adjustment) one that gives satisfactory best agreement with experiment. A different view, held by me, among others, is that model-building used in this way is insufficient for proof of structure; the correct structure should be arrived at by considerations based on, for example, heavy atoms, isomorphous replacement, intensity statistics ("direct" methods), Patterson (vector) maps, and gradual introduction of atoms into electron density functions based on partial structures. A flaw in the first method is that one can never be certain that a model sufficiently close to the true structure has been constructed; thus I must point out that this flaw is considerably more serious when only low-resolution data are available, so that the atoms are not resolved in the various refinement procedures.

2-Phenylazulene

Arnott cites the problem of 2-phenylazulene as an example of my inability to decide between two alternative structures when I used my "version of Fourier analysis." He then attempts to show that when *his* version of Fourier analysis, that is, difference synthesis, is applied, the problem can be solved. Although it is interesting that he chose an example of my own work to bolster his case on the power of difference syntheses, it is unfortunate that he chose to treat this problem in the way in which he did. Briefly, the problem is this: is 2-phenylazulene ordered (Arnott's figure 1a) or disordered (figure 1b)? Sharma and I concluded that, on the

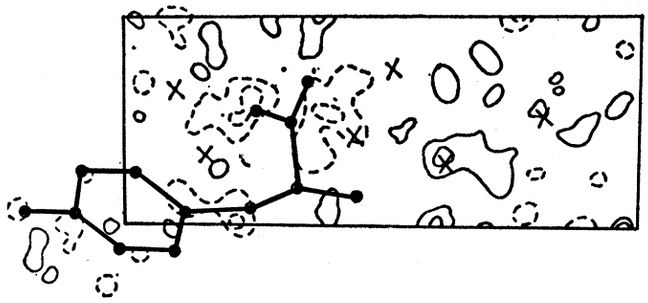
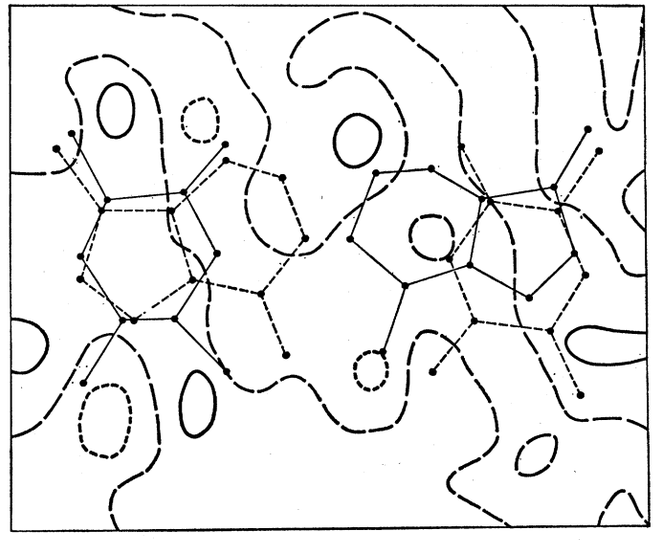


Fig. 1 (above). Difference synthesis calculated with the difference between the amplitudes of a six-atom random structure (crosses) and the tyrosine structure (dots), and the phases of the tyrosine structure. Contours are at intervals of 1 electron/Å^2. Positive contours are shown by continuous lines and negative contours by broken lines. The zero contour is omitted. According to the theory large positive peaks should occur at the X positions, and large negative peaks at the tyrosine positions. [After reference *16* in (*2*)] Fig. 2 (right). Difference map prepared with the differences between the amplitudes of pairings D (solid lines) and W (dashed lines) and the phases of pairing D.



basis of the observed $|F|$ values, from which we calculated two $|F_0|$ syntheses, it was not possible to decide (*8*). Arnott, on the other hand, uses calculated $|F|$ values and concludes that the difference synthesis method could (should?) have been used to make the decision. This conclusion is based on the fact that in his difference synthesis (Arnott's figure 1d) maxima occur near some of the positions corresponding to the true structure, and indications are present that less electron density is required at some of the positions corresponding to the provisional structure. In an actual crystal structure determination, such maxima and indications indeed might be used to decide between the two possibilities. However, what steps, in an actual determination, must precede the production of such a map? These are (i) removal, since this map was prepared with calculated $|F|$ values, of all experimental errors in the observed intensities, and (ii) determination of very precise values for all of the positional and thermal parameters. Achievement of (i) is, of course, impossible and it should be noted that spurious peaks of the order of about 0.3 to 0.8 electron/Å^3 have been observed in difference syntheses and attributed to experimental errors in the intensity data (*4*). Achievement of (ii) is necessary before a difference synthesis can be interpreted in the way that Arnott does, because peaks as high as 1.5 electron/Å^2 or more may appear in preliminary difference syntheses prepared with an incompletely refined structure (*5*, p. 334). In the case of 2-phenylazulene, for complete refinement of the ordered structure, if allowance is made for thermally anisotropic carbon atoms and isotropic hydrogen atoms, a total of 117 variables would have to be determined from the 72 observed values of $|F_0|$. Our original conclusion that this structure cannot be established on the basis of the available data is correct, whereas Arnott's conclusion that it could have been established by difference synthesis is not correct.

Determination of Structures by Difference Syntheses

Arnott refers to "$\Delta\rho$ maps that Donohue either did not consider or chose not to present." As a matter of fact, neither of these alternatives is correct. Several difference syntheses were included in an early draft of my manuscript, but these were removed on the advice of referees on the grounds that nothing was to be gained by including them. The final draft contains both $|F_0|$ and model syntheses (*9*), and, as pointed out by Wilkins *et al.* (*1*), a difference synthesis contains no more information than these. Thus, simultaneous examination of the illustration in my article (*2*, figure 3) ($|F_D|$ and α_D) and (*2*, figure *4*) ($|F_W|$ and α_D) gives no more clue that the "correct" structure is W than the corresponding difference synthesis does; the same is true of examination of figure 5 (*2*) ($|F_H|$ and α_H) and figure 6 (*2*) ($|F_W|$ and α_H). In case there is any doubt about this point, one of the difference syntheses removed from the original manuscript is presented above as Fig. 2. The general appearance of this map is that described by Arnott for the case when the provisional model is correct—"the $\Delta\rho$ map is relatively featureless apart from minor fluctuations. . . ." In the present example, the provisional model is not correct, and there are no clear indications from Fig. 2 that the structure is W rather than D.

Discrepancy Index, R

Values of R of 24 percent and less have been achieved for incorrect structures, some of which are more incorrect than others. Values this high, and higher, have been reported for DNA. It is immaterial whether the source of such values is inadequate treatment of the solvent, inaccuracies in the measured intensities, incorrectly assumed high symmetry, and the like. If the R value for a DNA structure is reduced, say, to the value 8 percent, which Arnott cites for polyproline II, then confidence in the correctness of the DNA structure would be substantially increased. In any case, it would be helpful if tables of $|F_0|$ and F_{calc}, such as crystallographers conventionally include in their reports of structure determinations, would be published; not only is the overall agreement as reflected in the R value of importance, but also such discrepancies as may occur with the individual F_{hkl}.

Crick's Remarks

It appears to me that Crick (*10*) has misinterpreted part of my article on the Fourier method. I did not mean to im-

ply that the x-ray data from DNA could be fitted just as well by a model with alternative base-pairing. What I said was "the Fourier method of structure refinement has, in fact, contributed nothing toward either the proof of that structure [that is, the Watson-Crick model], nor toward the elucidation of its details. . . ." This is a negative conclusion affirming that one particular model has not been established on the basis of evidence furnished by a particular method. Such models have not yet been tested rigorously.

Crick says that I did not allow for the numerous dyads in the structure which would cause many of the reflections to be effectively centric. These numerous dyads are, in fact, part of a model of a single molecule, and not of any of the proposed crystal structures, which are assemblages of molecules, and do not contain numerous dyads which must be "allowed for."

I agree with Crick that this matter needs deciding.

JERRY DONOHUE
Department of Chemistry and
Laboratory for Research on the Structure of Matter,
University of Pennsylvania,
Philadelphia 19104

References and Notes

1. M. F. H. Wilkins, S. Arnott, D. A. Marvin, L. D. Hamilton, *Science*, this issue.
2. J. Donohue, *ibid.* **165**, 1091 (1969).
3. S. Arnott, *ibid.*, this issue.
4. For example, R. L. Sass and J. Donohue, *Acta Cryst.* **11**, 497 (1958); K. N. Trueblood, E. Goldish, J. Donohue, *ibid.* **14**, 1009 (1961).
5. H. Lipson and W. Cochran, *The Determination of Crystal Structures* (Bell, London, 1957).
6. G. H. Stout and L. H. Jensen, *X-ray Structure Determination, a Practical Guide* (Macmillan, London, 1969), p. 361.
7. S. Arnott, M. H. F. Wilkins, L. D. Hamilton, R. Langridge, *J. Mol. Biol.* **11**, 391 (1965).
8. B. D. Sharma and J. Donohue, *Nature* **198**, 878 (1963).
9. Wilkins *et al.* have correctly detected a typographical error in the legend to figure 6 (*2*); the phases used in the preparation of this figure were those of pairing H.
10. F. H. C. Crick, *Science*, this issue.
11. Supported by the ARPA, Office of Secretary of Defense.

Animal Remains from Lepenski vir

The vertebrate fauna of this early center of domestication represent an atypical animal husbandry.

Sándor Bökönyi

In the Iron Gate gorge, where the Danube traverses the southern chain of the Carpathians, Yugoslavia and Rumania are planning a giant hydroelectric power system. Large areas of land will be inundated by the artificial lake, which will be formed by the proposed dam. Therefore, in 1965 large-scale salvage work was begun to save the most important archeological monuments of the area. Most of this work has been done by the Archeological Institute and the National Museum, Belgrade.

The most interesting site of the area is Lepenski vir. The excavation of this site has been one of the most rewarding of any in the field of European prehistory in recent years.

Lepenski vir lies on the right bank of the Danube, about 100 miles (160 kilometers) downstream from Belgrade, near the town of Donji Milanovac, not far from the mouth of the small Boljetinska River. At the site location the valley of the Danube widens from its narrow course in the Iron Gate gorge, making room for a small but almost inaccessible settlement.

Excavation has been going on since 1965, under the direction of D. Srejović and Z. Letica, of the University of Belgrade. By 1967, 1643 square meters had been explored; since then, more of the site has been unearthed.

According to Srejović (*1*) the site has three main occupation phases, Lepenski vir I, II, and III. Archeologically, the age of the first two is not yet clear, though it is certain that they represent a period before the earliest pottery-Neolithic culture of the Balkans. The first phase is characterized by trapeziform houses with hard, red-lime plaster floors and by stone sculptures. These finds are unique in Europe, particularly the figural and abstract sculptures, and they indicate that the site will have great significance for both archeologists and art historians.

The houses of the second phase did not have hard floors, but they did have sculptures, which are more monumental, though less finely worked, than those of the first phase.

The third phase belongs to the Starčevo-Körös culture complex that heralds the beginning of the pottery-Neolithic, found in the northern Balkans and in the southeastern half of the Carpathian Basin. Radiocarbon data place the early period of this complex between 5410 ± 100 and 4449 ± 75 B.C. (*2*).

Along with the archeological features and artifacts, a rich collection of animal remains was uncovered at the excavations. The bones were fragmentary, the sample yielding only one whole skull, some larger skull fragments, and a few whole long bones. However, despite their fragmentary state, they were well preserved, so a high percentage of them could be identified.

Unfortunately, as Table 1 clearly shows, the early phases of the site are the poorest in bone material. Despite this, one can obtain much valuable information about animal husbandry, hunting, and fishing within those phases.

The domestic faunas of the two early phases are very similar, and they differ sharply from the fauna of the third phase. Their most specific characteristic is that they have only one domestic species, the dog. This is quite an unusual situation, since we have not yet found in the temperate zone of Europe any Neolithic sites where the dog was the only domestic animal. Therefore, if the two early phases do not belong to the Mesolithic period, the first of the two phases may represent an independent, local evolution, and the second may be a survival of the first.

The wild faunas of the two early phases resemble each other closely. Common to the two phases is the high ratio of fishes, proving the great im-

The author is curator of the Hungarian National Museum, Budapest, Hungary.