Acta Cryst. (1993). D**49**, 66–74

Experience with Phase Extension and *Ab Initio* Phase Determination in Macromolecular Crystallography Using Maximum-Entropy Methods

BY LENNART SJÖLIN

Institute for Inorganic Chemistry, Chalmers University of Technology and University of Göteborg, S-412 96 Göteborg, Sweden

AND L. ANDERS SVENSSON

Department of Molecular Biophysics, University of Lund, PO Box 124, *S-221 00 Lund, Sweden*

(*Received* 22 *June* 1992; *accepted* 18 *September* 1992)

Abstract

Three procedures, or 'tools', have been developed and tested for applying maximum-entropy methods to phase extension and to *ab initio* phase determination. The phase expander tool has been used in connection with the solution of two previously unknown macromolecular structures. An efficient algorithm for the determination of an electron-density distribution that is everywhere positive and that agrees with observed structure amplitudes (tools II and III) has been used to determine the phases of X-ray diffraction data from recombinant bovine chymosin, a protein with 323 amino-acid residues in the molecular chain, the structure of which was recently determined using replacement methods. By use of the same maximum-entropy methods, the structure amplitudes from the unknown structure of bovine heart creatine kinase, a protein with 381 amino-acid residues, have been phased *ab initio* to 2.7 Å resolution. The phases of the centric reflections have also been confirmed by a satisfactory solution of the Patterson map of a mercury derivative. The current status of the structure interpretation is presented. This technique has also been applied to a test case where 48 centric reflections from bovine prothrombin fragment 1 data were phased *ab initio* and subsequently used in the determination of Patterson solutions for a heavy-atom derivative data set.

Introduction

In many branches of physics and chemistry such as spectroscopy, crystallography and astronomy, the primary data are the Fourier transforms of the physical quantities that are of interest to the scientist. One of the more crucial steps in protein crystallography, for example, is to interpret Fourier transforms that describe the pertinent electron-density distribution for the macromolecule under investigation. These electron-density maps are frequently difficult to interpret unambiguously, generally because of ill defined phases. This problem can usually be overcome if a better phase set can be generated by using, for example, phase information obtained from heavy-atom derivatives. However, there are other options available for improving phases; these options include various forms of real-space density modification.

The original application of density modification was to a small organic compound (Hoppe & Gassman, 1968), but currently these methods are widely, if not routinely, applied to macromolecules. In one of the earliest applications to protein structure, a low-resolution map of myohemerythrin (Hendrickson, Klippenstein & Ward, 1975) was improved by solvent flattening. Schevitz, Podjarny, Zwick, Hughes & Siegler (1981) subsequently developed a protocol that utilizes both solvent flattening and attenuation of negative density after manual determination of the molecular envelope. Another commonly used method is the one developed by Wang, in which a molecular envelope is automatically determined (Wang, 1985*a*,*b*). Like the Schevitz method, the Wang algorithm imposes non-negativity and smooth solvent restraints on the real-space electron-density distribution.

Another technique for density modification is described by the term 'Wiener filters'. Different formalisms for using Wiener filters for determining high-resolution positive-definite density estimates were developed by Wiener (1949), Bode & Shannon (1950), Robinson (1967), Lacoss (1971), Collins (1978) and Collins & Mahar (1983). These filters make statistical predictions of the values of unobserved Fourier coefficients on the basis of the observed ones.

Shannon (1948) explored the possibilities of using maximum-entropy statistics combined with various constraints, formulated on the basis of some prior knowledge of the system under investigation, in order to extract more information from the system. The immediate use of the analysis was in the field of information theory. Later Jaynes (1979) showed that the expression for maximum entropy, as derived by Shannon, was applicable to any scientific field where the prior knowledge and the experimental data could both be formalized in a statistical way. Jaynes also showed that a maximization of the entropy system subjected to these constraints would give rise to the least biased inference possible with respect to inherent noise and limitations in the data set.

0907-4449/93/010066-09$06.00

The idea of using the maximum-entropy method in order to obtain the least biased statistical measure of inverted information has been successfully used in a variety of applications. The Cambridge maximum-entropy group has developed a frequently used maximum-entropy algorithm (Bryan, 1980; Skilling, 1981) principally for linear data such as phased structure factors. It has also been used with Fourier data in astronomical interferometry (Scott, 1981), with convolution data in optics (Daniell & Gull, 1980; Burch, Gull & Skilling, 1983) and with nuclear magnetic resonance spectroscopy (Sibisi, 1983; Sibisi, Skilling, Brereton, Laue & Staunton, 1984), as well as in high-energy astronomy, with tomographic data and with extended X-ray absorption fine structure data (Kemp, 1980; Skilling, Strong & Bennett, 1979; Livesey, 1984).

Maximum-entropy methods have also been applied to the crystallographic problem by Collins (1982), Wilkins, Varghese & Lehmann (1983), Navaza (1985) and Bricogne (1984). The latter manuscript presents the most comprehensive work in this field. Gull, Livesey & Sivia (1987) have also discussed the theoretical concept of finding a maximum-entropy probability distribution with the constraints of satisfying observed Fourier amplitudes. Recently, Bricogne & Gilmore (1990) and Gilmore, Bricogne & Bannister (1990) have described a procedure in which likelihood is combined with maximum entropy that has shown considerable promise for application to problems involving several hundred atoms.

The numerical problem of finding an exponential model that satisfies experimental constraints was also addressed by Prince, Sjölin & Alenljung (1988), Prince (1989) and Sjölin, Prince, Svensson & Gilliland (1991), who showed that logarithmic space coefficients could be determined to any arbitrary precision using an iterative procedure similar to non-linear least squares. This procedure is equivalent to finding the unconstrained minimum of a dual function (Luenberger, 1984), which is equal to the constrained maximum of the entropy function. This suggests the use of the entropy of the map as a figure of merit for comparing different maps that satisfy the non-negativity condition at the same time as their Fourier transforms satisfy a specified number of Fourier amplitudes. Prince, Sjölin & Alenljung (1988) describe a procedure for finding a maximum-entropy distribution under conditions of an overall 'soft' constraint. This procedure was applied to phase extension in the solution of the structure of fragment TR_2C of bull testis calmodulin (Sjölin, Svensson, Prince & Sundell, 1990).

This manuscript summarizes previous work on phase extension and *ab initio* phase determination and presents new results with the solution of Patterson functions using methods based on the maximum-entropy concept. Finally, a trial of an *ab initio* phase determination on the unknown structure of bovine heart creatine kinase (381 amino-acid residues) and estimations of the pertinent phase errors are presented.

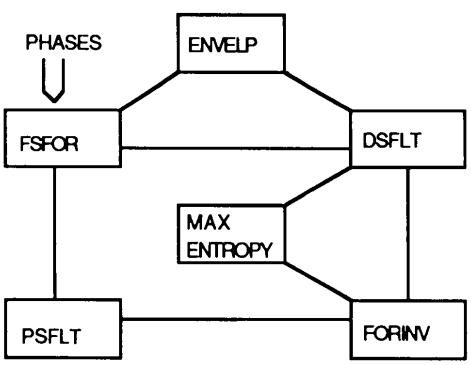


Fig. 1. A schematic diagram of the program system for density modification according to Wang (1985*a*,*b*) into which the maximum-entropy routine has been inserted. *SIRSAS* loads phases into the general files; *FSFOR* performs direct-space Fourier synthesis; *ENVELP* calculates the molecular envelope; *DSFLT* performs direct-space filtering; *MAX ENTROPY* maximizes the entropy of the map; *FORINV* performs Fourier inversion, and *PSFLT* is a reciprocal space filtering program.

Use of the maximum-entropy method

We have developed three 'tools' for applying maximum-entropy methods in phase extension and in *ab initio* phase determination. Each tool consists of a small program system for a specific purpose. Tool I is a phase expander tool. Tool II is used to determine the signs for 16 'new' centric reflections. This tool has been used for finding Patterson solutions for heavy-atom derivative data. Finally, tool III determines the phase for one acentric reflection at a time. It has been used in combination with the other tools in *ab initio* phase calculations.

Phase extension procedures

The maximum-entropy procedure described by Prince, Sjölin & Alenljung (1988) was coded as a subroutine in such a way that it could be easily incorporated in the commonly used density-modification system described by Wang (1985*a*,*b*). Ribonuclease A was chosen as a test structure because it had been refined to a high degree of precision (Wlodawer & Sjölin, 1983; Svensson, Sjölin, Gilliland, Finzel & Wlodawer, 1987; Wlodawer, Svensson, Sjölin & Gilliland, 1988). The procedure is summarized in Fig. 1. Since the method uses a definition of the molecular envelope, this was first calculated from the ribonuclease A data to a nominal resolution of 2.0 Å. Then two Fourier maps, using data to 2.0 and 5.0 Å resolution, respectively, were calculated and contoured at the 1σ level; these maps served as reference maps. Phases to resolution higher than 5.0 Å were then stripped from the original data set and thereafter treated as unknown. A series of calculations was undertaken according to Fig. 1, and the resolution was extended stepwise. During this process the original envelope was never altered. It was found after each cycle that some unphased reflections had strong

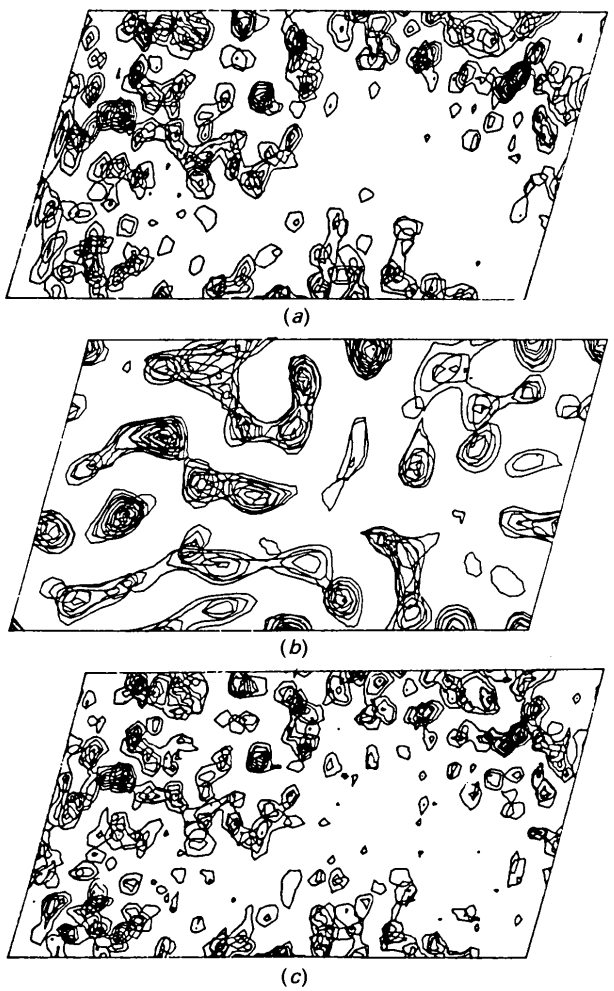
indications for certain phases, while others had only weak indications. Only reflections with strong indications were accepted and included in the next cycle. After the completed phase extension out to 2.0 Å resolution a Fourier map was calculated. Portions of this map and the reference maps are presented in Fig. 2. A calculation of the differences between the original phases and the extended phases revealed an average shift of roughly 35°.

The procedure was also tested on the structure of fragment 1 from bovine prothrombin which was unknown at that time. The initial phases were obtained from single isomorphous replacement data and were subjected to tool I. Fig. 3 shows the results from these calculations.

Tool I was also used in the structure determination of the calmodulin fragment TR_2C. The phased model of this structure was obtained by a combination of the molecular replacement method (Lattman, 1985) and maximum-entropy calculations. As was reported previously (Sjölin, Svensson, Prince & Sundell, 1990), there were two molecules in the asymmetric unit, but the rotation/translation function calculations could determine only one. However, the phases obtained from this solution were used as a starting set for phase improvement calculations using the maximum-entropy method with the soft constraints. After a new Fourier map had been calculated and displayed on the Evans & Sutherland screen, it was found that the solution for the missing molecule could be obtained from a translation of a previously determined rotation solution, a solution that had not produced any significant translation peak in the previous calculations.

Test of *ab initio* phase determination for centric and acentric reflections

In order to investigate the possibility of phasing centric reflections *ab initio*, we started to use tool II to produce electron-density maps of the previously solved structure of recombinant bovine chymosin (Gilliland, Winborne, Nachman & Wlodawer, 1990). This protein contains 323 amino-acid residues and crystallizes in the space group *I*222 with unit-cell dimensions a = 72.7, b = 80.2 and c = 114.8 Å. There is one molecule in the asymmetric unit. In this space group the structure factors for all *hk*0, *h*0*l* and

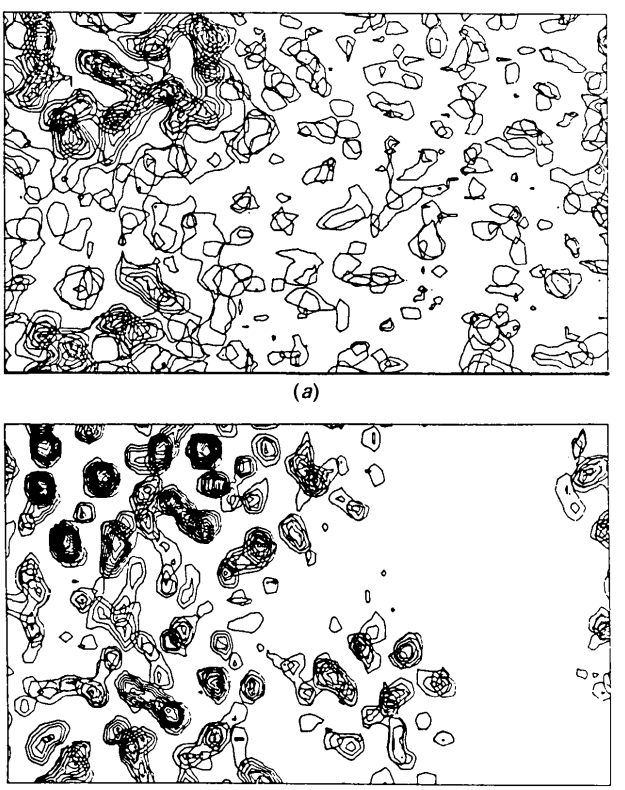


(*a*)

(*b*)

(*c*)

Fig. 2. Overlaid sections of Fourier maps from native ribonuclease A. (*a*) Map with phases to 2.0 Å resolution. (*b*) Map with phases to 5.0 Å resolution. (*c*) Map with phases extended from 5.0 to 2.0 Å.



(*a*)

(*b*)

Fig. 3. Overlaid Fourier sections calculated from SIR phases on fragment 1 from bovine prothrombin. (*a*) Regular F_{obs} map. (*b*) Map after maximum-entropy calculations.

0*kl* reflections must be real. The phase of one reflection from each set may be chosen arbitrarily to define the origin. A strong reflection was therefore selected from each set and assigned the phase that it had in the refined structure. Using methods described previously (Prince, 1989; Sjölin, Prince, Svensson & Gilliland, 1991), an initial map was calculated that maximized the entropy and matched the amplitudes of these three reflections. In the next step 16 strong reflections were chosen from the three centric groups, and 32 of all the possible sign combinations were chosen according to a 16-point fractional factorial design (Box, Hunter & Hunter, 1978). (A 16-point fractional factorial design gives 32 sign combinations that are the positive and negative senses of 16 mutually orthogonal vectors, so that there are 16 entropy differences, one for each pair. A procedure known as Yates's algorithm, actually a two-point fast Fourier transform in four dimensions, is used to extract estimates of the effects of the individual reflections.) The procedure that was adopted for this work was to add reflections in blocks of 16 reflections, using the 32 sign combinations in the fractional factorial design, the one additional one indicated by Yates's algorithm, and 16 more generated by single changes from the previously highest entropy, for a total of 49 trial sets. Maximum-entropy maps were computed with each sign combination and the total resulting entropy for each map was determined. From this it could be found, using the method described above, which set of signs produced the highest entropy. Following this procedure, 13 out of 16 reflections had the same signs as they had in the refined structure. (It was found subsequently that the correct set of signs had a still higher entropy, but the procedure had failed to find it.) After the first 16, 48 additional reflections were added, again in sets of 16, by the same procedure.

Continuing to use the chymosin data, we then extended the technique to handle acentric reflections as well. When the first 67 centric phases had been calculated, the strongest acentric reflection, 7 13 14, was selected and its amplitude fitted by maximum entropy with its phase set at each multiple of 20° from 0 to 340°. The result is presented in Fig. 4. It is apparent that, because the prior map contains no enantiomorph-determining information, the curve is symmetrical around 180°. The phase in the refined structure of chymosin for this reflection was 64°. We therefore selected the 60° maximum as our calculated phase, to define the enantiomorph. This reflection was now added to our map, and the phase of the next strongest acentric reflection, 9 1 20, was determined using the same procedure. The result is also plotted in Fig. 4 and, as can be seen, the plot is no longer symmetrical, but has a unique maximum at 316°. (The refined phase value was 336°.) After we had determined the phases of another 27 reflections manually using this technique, the whole procedure was automated by writing a computer program to do the necessary calculations. After tedious calculations using ten MicroVAX II computers, 159 centric and 1811 acentric reflections out of a total of 12 436 reflections to 1.6 Å resolution had been phased. Of these 1970 reflections, 1809 were in the resolution range 2.0–8.0 Å and 36 were of higher resolution than 2.0 Å.

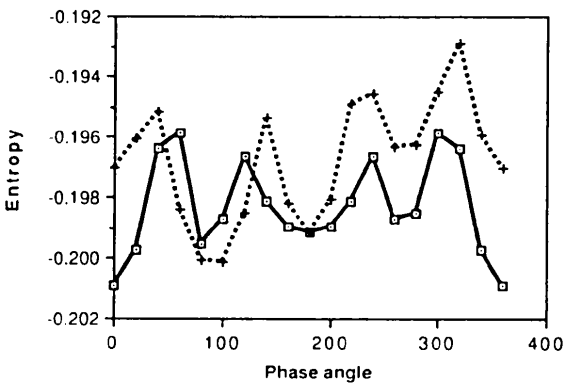


Fig. 4. The maximum entropy of a fitted electron-density map, plotted as a function of the initial phases of the two strongest acentric reflections. For the solid line the prior map contains only centric reflections, and the 7 13 14 reflection has been added. The plot is symmetrical, reflecting the two possible enantiomorphs. For the dotted line the prior map contains the 7 13 14 reflection, and the 9 1 20 reflection has been added. The plot is no longer symmetrical.

As reported earlier, a comparison of the phases computed by inversion of the final map with those from the refined structure showed that 138 out of 156 centric phases (88%) were the same as in the refined structure. Earlier tests on selected known simpler structures had suggested that the phasing accuracy could be in the range 80–93%. The median absolute phase difference for the acentric reflections was 32°, and 90% of them had differences less than 77°. Fig. 5 shows portions of density maps containing a four amino-acid residue segment of the chymosin molecule. There is a remarkable agreement between the map calculated from *ab initio* phases and the one calculated with phases from the previously refined structure. However, the *ab initio* map is based on only 1970 reflections and, consequently, there are gaps in the chain densities. In the whole 323-residue chain there are about 20 more gaps than in the map of the refined structure. Most of these gaps are about 3.5–4.0 Å wide. They are probably a result of the use of only 15% of the X-ray data in the calculations. Even if the map in most parts is very accurate, it is uncertain whether these gaps would prevent a real structure determination or not. One can conclude nevertheless that the total entropy of the map is an interesting candidate for a figure of merit for the choice of phases in macromolecular structures.

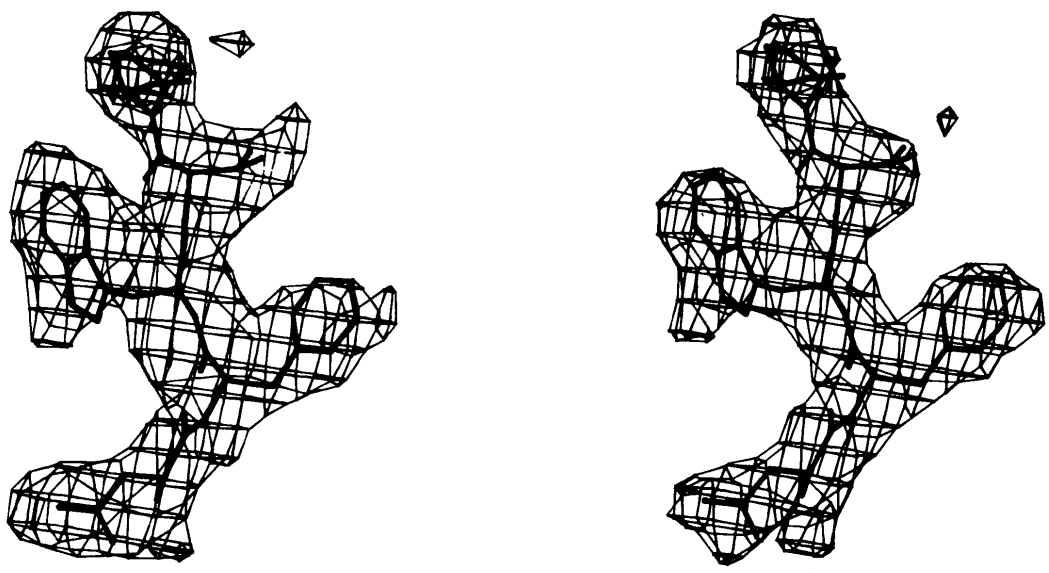
Solution of Patterson problems

Another test of the usefulness of the method was an attempt to find a heavy-atom site in a derivative. Normally, these positions can be found by regular Patterson methods. However, it should also be possible to find these sites, in many space groups, if a subset of centric reflections from the native data set could be phased properly, so that both

the magnitudes of the difference vectors between the native and the derivative structure factors and their phases could be calculated. The heavy-atom structure would then, in theory, be easily found from regular Fourier maps based on this information. This was tested by phasing the 48 strongest native centric reflections from bovine prothrombin fragment 1. The fragment 1 molecule crystallizes in space group $P2_12_12_1$ with the cell constants a = 39, b = 54.0 and c = 129.0 Å. This system was selected because we had earlier collected two different heavy-atom derivative data sets to 4.0 Å resolution, and the solution to the Patterson problem using standard methods in space group $P2_12_12_1$ (three Harker sections) showed one major site and a few minor sites for each derivative. Multiple isomorphous replacement maps based on these solutions had earlier been used to deduce a 4.0 Å resolution structure for this molecule (Olsson, Andersen, Lindqvist, Sjölin, Magnusson, Petersen & Sotterup-Jensen, 1982). The three strongest reflections, one from each centric zone, were selected to be the origin-defining reflections. Then a set of 16 reflections, evenly selected from the three centric groups, were analysed according to the method described above. Before these phases were accepted and included in the new map, another set of 16 reflections was analysed. This time groups of four reflections from these first two calculations were combined with each other and analysed once more. In this way ten reflections that showed stability through these primary tests were included in the new map, which at this point consisted of the three origin-defining reflections and the ten new reflections. Consistently, at least ten of the 'old' reflections were used in the trial set, and each reflection was phased at least twice before its phase was accepted and included in the map. This procedure continued until 48 reflections plus the three origin-defining reflections had been phased. The derivative data set was then scaled against the native data set using standard methods, and the differences between the structure factors were calculated. This revealed 41 common reflections with accepted structure-factor values. From the magnitudes and signs of the differences, these reflections could then be phased from the previously determined phase information for the native reflections. A regular Fourier map was calculated, based on the 41 phased differences. The major peaks in this map should, in theory, indicate the positions of plausible heavy-atom sites. The Fourier map was subsequently displayed on an Evans & Sutherland PS390 graphics system and, although the map was somewhat noisy due to the small number of reflections used in the calculations, the major site appeared as one of the highest peaks in this Fourier map.

Test of *ab initio* phase calculations for the unknown structure of bovine heart creatine kinase

Another test of maximum-entropy methods was the determination of a set of phases of the unknown structure of bovine heart creatine kinase, which is a protein with 381 amino-acid residues (Gilliland, Sjölin & Olsson, 1983). The data were collected from crystals of the native protein that belonged to space group $P4_22_12$ with a = b = 132.0, c = 75.0 Å, and one molecular subunit in the asymmetric unit. In this space group the structure factors of all $hk0$ and hhl reflections must be real, while those of $h0l$ reflections may be real or imaginary, but may not be complex. The



(*a*) (*b*)

Fig. 5. Comparison of portions of the two density maps for a four amino-acid residue segment of the molecule of recombinant bovine chymosin. (*a*) The map with phases extended by maximum entropy to 1970 reflections. (*b*) The corresponding map computed using all data with phases calculated from the refined structure.

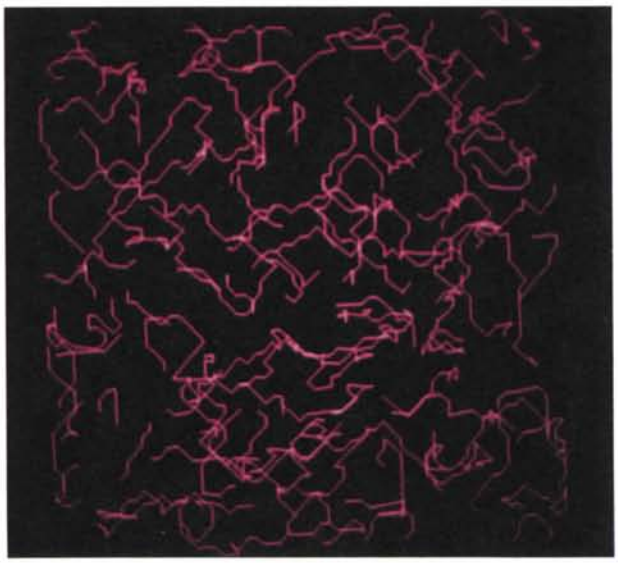


Fig. 6. A portion of the unit cell of bovine heart creatine kinase, with density ridges as traced by the *BONES* feature of the *FRODO* computer-graphics system. The border between two molecules is clearly visible.

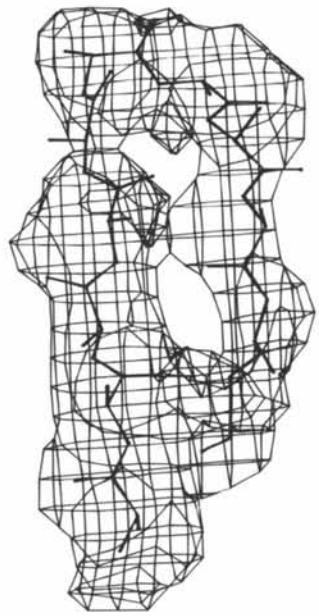
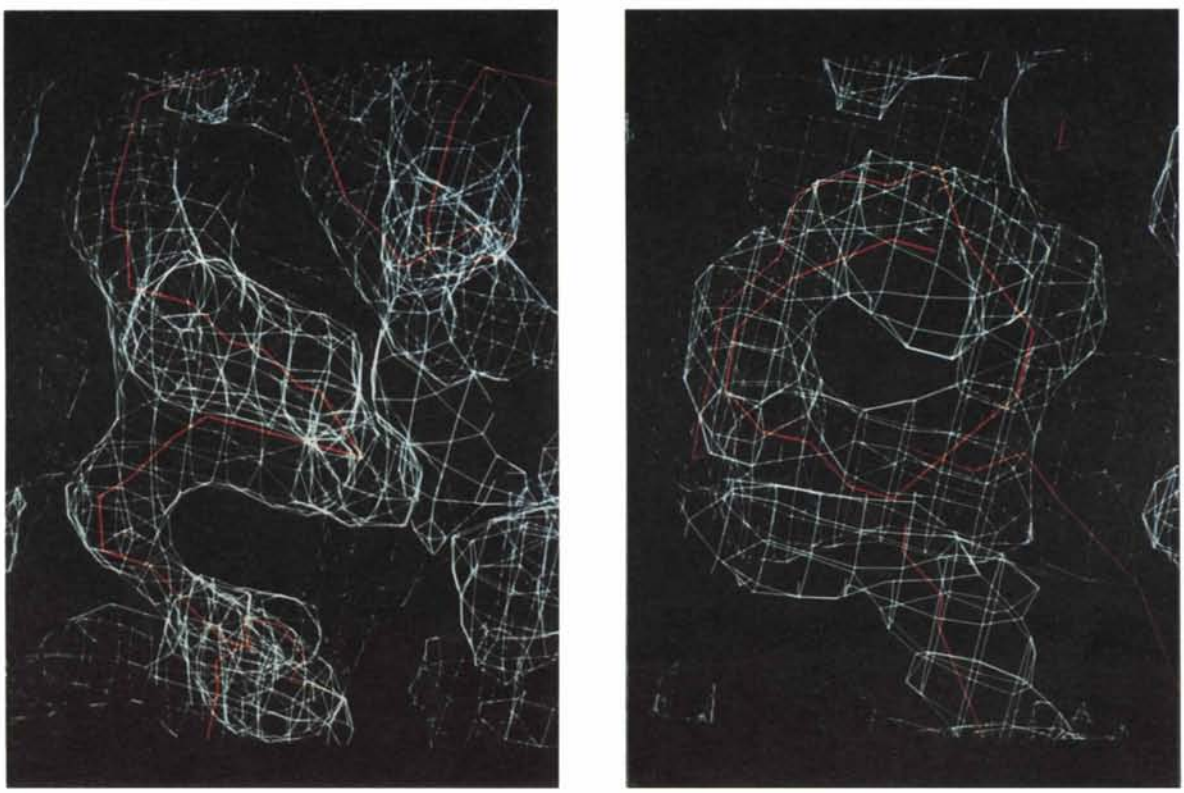


Fig. 7. A section of the electron-density map of creatine kinase into which two antiparallel segments of β-strands with the connecting hairpin loop have been modeled.



(*a*) (*b*)

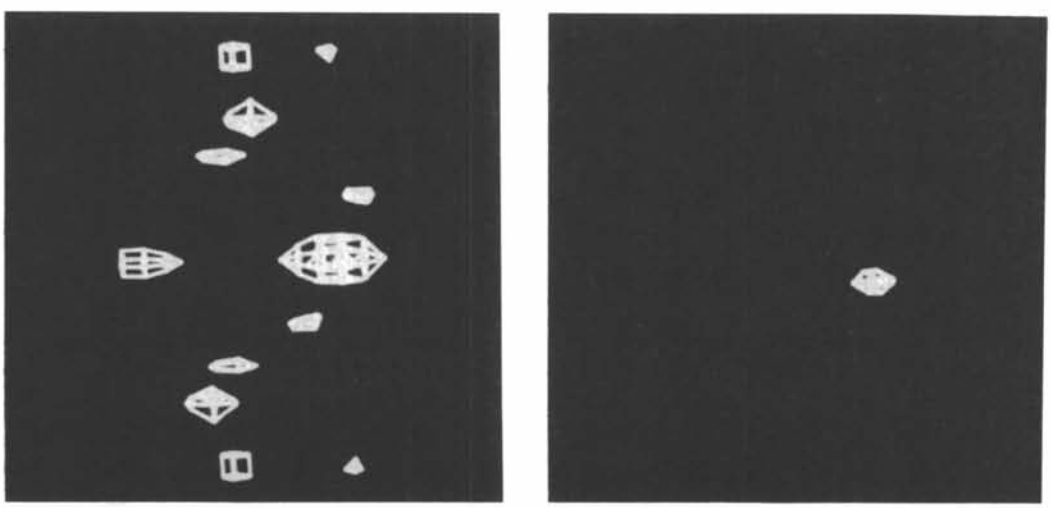
Fig. 8. (*a*), (*b*) One region of the map from creatine kinase is viewed from two directions, showing the skeleton structure obtained by *BONES*. The diameter and pitch of the helix are close to those of a typical α-helix.

phases of one *hhl* reflection with *l* odd and one *hk*0 reflection with $h + l$ odd may be chosen arbitrarily, to fix the origin, and the phase of one *h*0*l* reflection with $h + l$ odd may be chosen to fix the enantiomorph. An initial map was calculated that maximized entropy and fitted the amplitudes of three reflections chosen from these sets. Then, similar to the analysis of the centric phases for chymosin, a set of 16 reflections was chosen from the three centric groups and the calculations continued, adding reflections in blocks of 16, until 240 centric reflections had been phased out to a resolution of 4.0 Å. Next, acentric reflections were added using a procedure similar to that described for chymosin, except that, because the enantiomorph is defined by acentric reflections in this space group, no preliminary choice between symmetrical peaks was necessary. The procedure was continued, adding centric reflections in blocks of 16 and acentric reflections one at a time, until all reflections up to 4.0 Å had been phased. Finally, continuing use of the maximum-entropy procedure extended phases to 2.7 Å resolution, first for reflections from the three special sets and then for all reflections. This final map was used as input to the *FRODO* computer-graphics system (Jones, 1978), in order to fit a chain backbone into the region of contiguous density.

Results from the *FRODO* computer-graphics system using the *BONES* feature are presented in Fig. 6. Although the connectivity of these skeleton sections is very likely spurious, the division between molecules is clearly visible, and the molecule is situated in a favourable position in the cell if space-group considerations are taken into account. When the electron density is further examined in more detail, secondary structure elements can be found. Fig. 7 shows the electron-density map containing a region into which a segment of an antiparallel β-strand and the connecting hairpin loop found in avian pancreatic polypeptide fits remarkably well. Fig. 8 shows another region of the map where an α-helix from crambin has been modeled into the density. The probability that a map displaying such common protein structure features could be produced by a random choice of phases would appear to be rather small. However, despite a rather substantial effort, the fitting of a complete chain backbone into the electron-density map has not yet been successful. The quality of the map in general is probably not good enough to allow this in a straightforward manner. We must therefore conclude that the incorporation of other information into our maximum-entropy procedure will probably be necessary.

Since we were not readily able to interpret the electron-density map, we continued the investigation by using the mercury derivative X-ray data set that had previously been collected at the synchrotron DESY in Hamburg. Using the same technique as discussed above for the Patterson solution of the fragment 1 test case, we now rephased 243 of the centric reflections from creatine kinase, using the same origin-defining reflections as before, but selecting the prior map according to the test case. We also used faster versions of the programs *INIMAP* and *MESIGNS*, in which Fourier transforms are performed by FFT algorithms, for the calculations.

When these 243 centric reflections had been phased, the two data sets were scaled together, the differences determined and subsequently phased. The resulting Fourier map was again displayed on the computer-graphics system, and a representation of the map on the screen using the program *FRODO* is shown in Fig. 9. As can be seen from this picture, there is one major peak and a number of smaller peaks. The coordinates from the major peak and the minor peaks were then obtained from the screen and subjected to phase refinement using *MIRPH*, a program in the *PROTEIN* package (Steigemann, 1974). A regular



(*a*) (*b*)

Fig. 9. Results from a Fourier calculation of phased structure-amplitude differences from centric creatine kinase reflections. (*a*) The map peaks at 3σ cut-off level. (*b*) The map peaks at 5σ cut-off level.

Table 1. *A summary of the different maximum-entropy experiment tools that were used*

The tools are indicated by roman numerals.

Structure	Unknown/ known	Phase extension	Heavy-atom location	*Ab initio* phasing	Results
Ribonuclease A	K	I			Satisfactory
Calmodulin TR_2C	UK	I			Satisfactory
Prothrombin fragment 1	UK	I	II		Satisfactory
Chymosin	K			II, III	Good indications
Creatine kinase	UK		II	II, III	Not yet

Patterson map based on the squared amplitude differences was calculated at this point. The reliability of the extracted coordinates could then be examined by a visual inspection of the Patterson map. It was now found that the coordinates from the three strongest solutions were supported satisfactorily by related peaks in the Patterson map.

In addition, these 243 new native phases, obtained by using the new version of the program system and with the somewhat modified selection technique, were compared with the old phases determined as described earlier. About 75% of the phases were of equal sign.

Results and concluding remarks

Phase extension, as we have shown previously and as shown by Dong *et al.* (1992) and others, is now a relatively established technique. As presented in Fig. 2, the results from the phase extension calculations are in good agreement with the original map. The procedure was also tested on two structures unknown at that time, fragment 1 from bovine prothrombin (*cf.* Fig. 3), in which the initial phases were obtained from isomorphous replacement data, and fragment TR_2C from bull testis calmodulin, where the starting phases were determined by using molecular replacement methods. Each phase set was subjected to entropy maximization according to the scheme in Fig. 1. However, in these cases the molecular envelopes were recalculated several times during the course of the study. In the case of fragment 1 the quality of the map improved and the structure determination could then proceed in a straightforward manner. A more detailed description of the solution of the TR_2C structure has previously been published (Sjölin, Svensson, Prince & Sundell, 1990; Svensson, 1989).

These results imply that the maximum-entropy procedure can extend phases rather accurately. If low-resolution phases of good quality could be obtained by either isomorphous replacement techniques or by molecular replacement methods then phase extension is an interesting alternative, particularly if the system under investigation shows a lack of isomorphism for higher resolution data. It is also likely that the phasing power of the maximum-entropy fitting increases as the amount of information incorporated in the prior map increases, making the phases potentially more and more accurate.

In summary, three tools have been used in three different phasing experiments and tested on five different structures, four of them known and one unknown. These experiments are summarized in Table 1. Our first attempt to phase *ab initio*, and consequently solve, the unknown structure of bovine heart creatine kinase, using the total entropy as a figure of merit, has not yet been successful, in spite of the good indications obtained when using the method on the test structure of recombinant bovine chymosin. Therefore, the conclusion must be that more information, or a better figure of merit, is needed. We will continue to investigate the effects of various parameters that appear to be important for a general solution of the problem and that at the same time can feasibly be incorporated in an efficient numerical procedure.

The authors wish to thank the National Institute of Standards and Technology and the Swedish Natural Science Research Council (NFR) for financial support of this project.

References

Bode, H. W. & Shannon, C. E. (1950). *Proc. IRE*, **38**, 417-425.

Box, G. E. P., Hunter, W. G. & Hunter, J. S. (1978). *Statistics for Experimenters*. New York: John Wiley.

Bricogne, G. (1984). *Acta Cryst.* A**40**, 410-445.

Bricogne, G. & Gilmore, C. J. (1990). *Acta Cryst.* A**46**, 284-297.

Bryan, R. K. (1980). PhD thesis, Univ. of Cambridge, England.

Burch, S. F., Gull, S. F. & Skilling, J. (1983). *Comput. Graph. Image Process.* **23**, 113-128.

Collins, D. M. (1978). *Acta Cryst.* A**34**, 533-541.

Collins, D. M. (1982). *Nature (London)*, **298**, 49-51.

Collins, D. M. & Mahar, M. C. (1983). *Acta Cryst.* A**39**, 777-783.

Daniell, G. J. & Gull, S. F. (1980). *Proc. Inst. Electr. Eng. Part E*, **127**, 170-172.

Dong, W., Baird, T., Fryer, J. R., Gilmore, C. J., Macnicol, D. D., Bricogne, G., Smith, D. J., O'Keefe, M. A. & Hovmöller, S. (1992). *Nature (London)*, **355**, 605-609.

Gilliland, G. L., Sjölin, L. & Olsson, G. (1983). *J. Mol. Biol.* **170**, 791-793.

Gilliland, G. L., Winborne, E. L., Nachman, J. & Wlodawer, A. (1990). *Proteins Struct. Funct. Genet.* **8**, 82-101.

Gilmore, C. J., Bricogne, G. & Bannister, C. (1990). *Acta Cryst.* A**46**, 297-308.

Gull, S. F., Livesey, A. K. & Sivia, D. S. (1987). *Acta Cryst.* A**43**, 112-117.

Hendrickson, W. A., Klippenstein, G. L. & Ward, K. B. (1975). *Proc. Natl Acad. Sci. USA*, **72**, 2160-2164.

Hoppe, W. & Gassman, J. (1968). *Acta Cryst.* B**24**, 97-107.

Jaynes, E. T. (1979). *The Maximum-Entropy Formalism*, edited by R. D. Levine & M. Tribus, pp. 15-118. Cambridge, MA: MIT Press.

Jones, T. A. (1978). *J. Appl. Cryst.* **11**, 268-272.

Kemp, M. C. (1980). *Med. Radionucl. Imag. IAEA (Vienna)*, **1**, 313-323.

Lacoss, R. T. (1971). *Geophysics*, **36**, 661-675.

LATTMAN, E. (1985). *Methods Enzymol.* **115**, 55-77

LIVESEY, A. K. (1984). PhD thesis, Univ. of Cambridge, England.

LUENBERGER, D. G. (1984). *Linear and Non-Linear Programming*, pp. 396-401. Cambridge, MA: Addison-Wesley.

NAVAZA, J. (1985). *Acta Cryst.* A**41**, 232-244.

OLSSON, G., ANDERSEN, L., LINDQUIST, O., SJÖLIN, L., MAGNUSSON, S., PETERSEN, T. E. & SOTTERUP-JENSEN, L. (1982). *FEBS Lett.* **145**, 317-322.

PRINCE, E. (1989). *Acta Cryst.* A**45**, 200-203.

PRINCE, E., SJÖLIN, L. & ALENLJUNG, R. (1988). *Acta Cryst.* A**44**, 216-222.

ROBINSON, E. A. (1967). *Statistical Communication and Detection*. New York: Hafner.

SCHEVITZ, R. W, PODJARNY, A. D., ZWICK, M., HUGHES, J. J. & SIEGLER, P. B. (1981). *Acta Cryst.* A**37**, 669-677.

SCOTT, P. F. (1981). *Mon. Not. R. Astron. Soc.* **194**, 23P-29P.

SHANNON, C. E. (1948). *Bell Syst. Tech. J.* pp. 379-423, 623-656.

SIBISI, S. (1983). *Nature (London)*, **301**, 134-136.

SIBISI, S., SKILLING, J., BRERETON, R. G., LAUE, E. D. & STAUNTON, J. (1984). *Nature (London)*, **311**, 446-447.

SJÖLIN, L., PRINCE, E., SVENSSON, L. A. & GILLILAND, G. L. (1991). *Acta Cryst.* A**47**, 216-223.

SJÖLIN, L., SVENSSON, L. A., PRINCE, E. & SUNDELL, S. (1990). *Acta Cryst.* B**46**, 209-215.

SKILLING, J. (1981). 1st Workshop on Maximum Entropy Estimation and Data Analysis, Univ. of Wyoming, USA.

SKILLING, J., STRONG, A. W. & BENNETT, K. (1979). *Mon. Not. R. Astron. Soc.* **187**, 145-152.

STEIGEMANN, W. (1974). PhD thesis, Technische Univ., München, Germany.

SVENSSON, L. A. (1989). PhD thesis, Univ. of Göteborg, Sweden.

SVENSSON, L. A., SJÖLIN, L., GILLILAND, G. L., FINZEL, B. C. & WLODAWER, A. (1987). *Proteins Struct. Funct. Genet.* **1**, 370-375.

WANG, B. C. (1985*a*). *Methods Enzymol.* **115**, 90-112.

WANG, B. C. (1985*b*). *Diffraction Methods in Biological Macromolecules*, edited by H. W. WYCOFF, C. H. W. HIRS & S. N. TIMASHEFF, pp. 114-167. New York: Academic Press.

WIENER, N. (1949). *Extrapolation, Interpolation and Smoothing of Stationary Time Series*. New York: John Wiley.

WILKINS, S. W., VARGHESE, J. N. & LEHMANN, M. S. (1983). *Acta Cryst.* A**39**, 47-60.

WLODAWER, A. & SJÖLIN, L. (1983). *Biochemistry*, **22**, 2720-2728.

WLODAWER, A., SVENSSON, L. A., SJÖLIN, L. & GILLILAND, G. L. (1988). *Biochemistry*, **27**, 2705-2717.