

АКАДЕМИЯ НАУК СССР  
НАУЧНЫЙ ЦЕНТР БИОЛОГИЧЕСКИХ ИССЛЕДОВАНИЙ

---

В. Ю. ЛУНИН  
ФАЗОВАЯ ПРОГРАММА ТЕН ЭЙКА

# ФОРТРАН

МАТЕРИАЛЫ ПО  
МАТЕМАТИЧЕСКОМУ  
ОБЕСПЕЧЕНИЮ ЭВМ

---

4

ПУЩИНО-1979

АКАДЕМИЯ НАУК СССР  
НАУЧНЫЙ ЦЕНТР БИОЛОГИЧЕСКИХ ИССЛЕДОВАНИЙ  
НАУЧНО-ИССЛЕДОВАТЕЛЬСКИЙ  
ВЫЧИСЛИТЕЛЬНЫЙ ЦЕНТР

В.Ю.ЛУНИН

ФАЗОВАЯ ПРОГРАММА ТЕН ЭЙКА

Программное обеспечение структурных  
исследований

ПУЩИНО • 1979

УДК 548.73:681.3.06

Описываемая в данном выпуске программа предназначена для решения двух задач рентгеноструктурного анализа:

- расчета фаз структурных факторов нативного белка по данным изоморфного замещения;
- уточнения параметров тяжелоатомных производных.

Программа работает в любой пространственной группе и с произвольным числом рефлексов. При наличии данных по аномальному рассеянию программа позволяет включить их в расчет.

The program to be described in this issue aims to solving of two X-ray structure analysis tasks:

- a definition of structure factors phases from the isomorphous replacement data;
- a refinement of the heavy atoms parameters of isomorphous derivatives.

The anomalous scattering data may be added in the calculations.

Научный редактор д-р биол. наук Ю.Н.Чиргадзе

© Научный центр биологических исследований АН СССР  
в Пушкине, 1979

## ВВЕДЕНИЕ

Описываемая в данном выпуске программа предназначена для решения двух задач, возникающих при рентгеноструктурном исследовании белков: расчета фаз структурных факторов нативного белка по данным изоморфного замещения; уточнения параметров тяжелых атомов.

В зависимости от желания пользователя программа может выполнять одну из указанных функций или их комбинацию. При наличии данных по аномальному рассеянию программа позволяет включить их в работу. Программа может работать в любой пространственной группе и с любым числом рефлексов во входном файле. При обработке рефлексов из centrosymmetric зон учитывается, что для них возможны только два значения фазы.

Настоящая программа представляет собой незначительно модифицированный вариант фазовой программы доктора Л.Тен Эйка, любезно переданной им заведующему лабораторией структурного анализа Института белка АН СССР Ю.Н.Чиргадзе. Алгоритм решения фазовой задачи в этой программе основан на использовании вероятностной схемы оценки фазовых ошибок согласно Блоу и Крику /1/. Принципиальным дополнением является предложенный Тен Эйком способ учета экспериментальных ошибок структурных амплитуд, позволяющий рассматривать их влияние отдельно от ошибок в определении параметров тяжелых атомов, которые возникают в процессе решения фазовой задачи /2/. Процедура уточнения, используемая во второй части программы, основана на минимизации разности модулей экспериментальной и вычисленной структурных амплитуд производной. Модификации программы были сделаны с целью ее использования на ЭВМ серии ЕС.

Указанная адаптация программ к ЕС ЭВМ была осуществлена в лаборатории вычислительной математики Научно-исследовательского вычислительного центра АН СССР. В 1977-1979 гг. эта программа использовалась сотрудниками Института белка АН СССР Ю.Н.Чиргадзе и Ю.В.Сергеевым в практической работе, что помогло устранить ряд мелких недоработок.

Программа состоит из главной программы и восьми подпрограмм. Все программы написаны на алгоритмическом языке Фортран IV. Для работы программы требуется около 125 К байт оперативной памяти.

## ГЛАВА I. ПРИНЦИПЫ ОРГАНИЗАЦИИ ПРОГРАММЫ

### §1. Логическая структура программы

Выполнение программы состоит из заданного числа циклов расчета фаз структурных факторов нативного белка и уточнения параметров тяжелоатомных производных. Число необходимых циклов задается в управляющих данных величиной `NCYCLE`.

На каждом цикле фазы структурных факторов могут быть вычислены заново или оставлены такими же, как на предыдущем цикле. Кроме того, для указанных производных может быть уточнен ряд параметров. Пересчет фаз на  $J$ -м цикле работы задается ненулевым флагом `KCYCLE(J)` в управляющих данных. Исключение составляет первый цикл. Здесь для рефлексов, у которых значение фазы равно 1000.0 (признак неопределенности фазы), фазы будут вычисляться заново независимо от значения флага `KCYCLE(1)`.

В работе программа может использовать до шести производных. Конкретное число используемых производных задается величиной `NCPDS` в управляющих данных. При работе программы производные нумеруются в том порядке, в котором поступает о них информация с управляющих перфокарт (см. гл. II, §1). Во входном файле со структурными амплитудами (см. гл. II, §4) информация по этим производным может располагаться в ином порядке. Для установления соответствия для каждой из используемых в работе программы производных в управляющих данных задается величина `MAP(J)` - порядковый номер во входном файле данных по  $J$ -й производной.

Использование данных по  $J$ -й производной при расчете фаз структурных факторов определяется ненулевым флагом `NPHASE(J)` в управляющих данных.

При желании уточнить для J-й производной какие-либо параметры следует задать в управляющих картах ненулевой флаг NREFIN (J). В этом случае набор уточняемых для данной производной параметров определяется значениями флагов IFBOV и LABEL1÷LABELA (гл. I, §5).

Включение для J-й производной данных по аномальному рассеянию осуществляется путем задания в управляющих данных ненулевого флага KAY(J).

Программа может работать с любой пространственной группой. Информация о пространственной группе задается числом NSYM эквивалентных положений и матрицами симметрии. Для каждого из NSYM эквивалентных положений соответствующее преобразование симметрии, включая тождественное, представляется в виде  $\vec{x} = R\vec{x} + \vec{t}$  и для него задается матрица вращения и вектор трансляции:

$$R = \begin{pmatrix} T1 & T2 & T3 \\ T5 & T6 & T7 \\ T9 & T10 & T11 \end{pmatrix}, \vec{t} = (T4, T8, T12).$$

При определении фаз программа учитывает, что для рефлексов из centrosимметричных зон фазовый угол может принимать только одно из двух возможных значений, отличающихся по фазе на 180°. Значение одной из возможных фаз определяется при этом выражением:

$$\varphi = \text{CENT1} * h + \text{CENT2} * k + \text{CENT3} * l.$$

Условие принадлежности рефлекса зоне можно записать в виде:

$$CT1 * h + CT2 * k + CT3 * l = 0.$$

В управляющих картах задается число NCEN centrosимметричных зон во входном файле (а не вообще в группе) и для каждой из зон задаются величины CT1, CT2, CT3, CENT1, CENT2, CENT3.

Например, для группы  $P4_12_1$  коэффициенты уравнения centrosимметричной плоскости (h0l) в обратном пространстве запишутся в виде 0, 1, 0, а коэффициенты в формуле определения фаз соответственно в виде 90, 0, 45 (размерность угла - градусы).

На число рефлексов, информация о которых дана во входном файле, не накладывается никаких ограничений. Однако

при работе программа рассматривает не все рефлексy, а только те из них, которые относятся к заданной зоне разрешения  $[D_{MIN}, D_{MAX}]$ , (то есть  $D_{MIN} \leq d = \frac{\lambda}{2 \sin \theta} \leq D_{MAX}$ ).

Для ведения статистики интервал  $(\frac{1}{D_{MIN}}, \frac{1}{D_{MAX}})$  разбивается по  $\frac{\sin \theta}{\lambda}$  на восемь равных подинтервалов, и статистика ведется по всем включенным в работу рефлексам и отдельно - по рефлексам, попавшим в каждый из этих восьми подинтервалов.

## 82. Шкалирование данных в программе

На каждом цикле, начиная обработку очередного рефлекса, программа считывает для него из файла со структурными амплитудами значения:

$F_P^{tape}$  - амплитуда структурного фактора нативного белка;  
 $\sigma_P^{tape}$  - стандартное отклонение для  $F_P^{tape}$  ;  
 $F_{PHj}^{tape}$  - амплитуды структурных факторов тяжелоатомных производных, участвующих в работе ( $j = 1, 2, \dots, NCPDS$ ;  
 где  $NCPDS$  - число производных, участвующих в работе);  
 $\sigma_{PHj}^{tape}$  - стандартное отклонение для  $F_{PHj}^{tape}$ .

При наличии данных по аномальному рассеянию берется также значение бейфтовой разности  $\Delta_j^{tape} = F_{PHj}^{+tape} - F_{PHj}^{-tape}$ .

В процессе расчета фаз структурных факторов данные шкалируются следующим образом (шкала нативного белка):

$$F_P = SCALE * \{ F_P^{tape} \};$$

$$\sigma_P = SCALE * \{ SENAT * \sigma_P^{tape} \};$$

$$F_{PHj}^{obs} = SCALE * \left\{ \frac{1}{Q(J)} * e^{-BOVER(J)} * \left( \frac{\sin \theta}{\lambda} \right)^2 * F_{PHj}^{tape} \right\};$$

$$\sigma_{PHj}^{obs} = SCALE * \left\{ \frac{SCESD(J)}{Q(J)} * e^{-BOVER(J)} * \left( \frac{\sin \theta}{\lambda} \right)^2 * \sigma_{PHj}^{tape} \right\};$$

$$\Delta_j^{obs} = SCALE * \left\{ \frac{1}{Q(J)} * e^{-BOVER(J)} * \left( \frac{\sin \theta}{\lambda} \right)^2 * \Delta_j^{tape} \right\}.$$

Здесь общие шкальные коэффициенты  $SCALE$ ,  $SENAT$  и шкальные коэффициенты  $J$ -й производной -  $BOVER(J)$ ,

$Q(J)$ ,  $SCESD(J)$  задаются в управляющих данных;  $F_P$ ,  $G_P$ ,  $F_{PHj}^{obs}$ ,  $G_{PHj}^{obs}$ ,  $\Delta_j^{obs}$  - используемые при расчете фаз амплитуда и стандартное отклонение амплитуды структурного фактора нативного белка; амплитуды, аномальные разности и стандартные отклонения амплитуд структурных факторов производных.

В процессе уточнения параметров  $J$ -й производной данные шкалируются иначе (шкала  $J$ -й производной):

$$F_P = SCALE * \left\{ Q(J) * e^{BOVER(J) * \left(\frac{\sin \theta}{\lambda}\right)^2} \right\} * F_P^{tape};$$

$$G_P = SCALE * \left\{ SENAT * Q(J) * e^{BOVER(J) * \left(\frac{\sin \theta}{\lambda}\right)^2} \right\} * G_P^{tape};$$

$$F_{PHj}^{obs} = SCALE * F_{PHj}^{tape};$$

$$G_{PHj}^{obs} = SCALE * SCESD(J) * G_{PHj}^{tape};$$

$$\Delta_j^{obs} = SCALE * \Delta_j^{tape}.$$

### 83. Расчет вклада тяжелых атомов в структурный фактор

Информация, необходимая для расчета вклада тяжелых атомов в структурный фактор тяжелоатомной производной, поставляется управляющими данными на перфокартах (см. §1, гл. II).

При расчете отдельно вычисляются вещественные и мнимые части требуемых комплексных величин, причем функции  $\sin x$  и  $\cos x$  определяются из таблицы с шагом в  $1^\circ$ .

Остановимся на особенности задания в программе факторов тяжелых атомов и коэффициентов заполнения.

Программа использует для работы только те рефлексы, для которых межплоскостное расстояние в обратном пространстве  $d^* = \frac{2 \sin \theta}{\lambda}$  находится в заданных пределах:

$$\frac{1}{D_{MAX}} \leq d^* \leq \frac{1}{D_{MIN}}. \text{ Интервал } \left( \frac{1}{2 * D_{MAX}}, \right.$$

$\frac{1}{2 * DMIN}$  ) разбивается на восемь равных частей, и для значений  $\frac{\sin \theta}{\lambda}$ , лежащих внутри каждого такого подинтервала, значение формфактора считается постоянным. То есть формфактор (аналогично и аномальная рассеивающая способность) задается восьмью значениями.

Для каждого структурного фактора  $\vec{F}_{PHj}$  j-й производной во входном файле данных хранится условный номер JXL(J). Этот номер должен быть в пределах от 1 до 100 и может означать, например, номер рентгенограммы, с которой снят этот структурный фактор.

При обработке очередной производной все ее структурные факторы из входного файла разбиваются на заданное число NXL "рабочих групп" (от 1 до 8). Для каждого условного номера JXL указывается в управляющих данных номер JCRIN(JXL) "рабочей группы", к которой будет отнесен структурный фактор с таким номером. Например, в одну "рабочую группу" могут быть отнесены все структурные факторы, снятые с одного и того же кристалла. При расчете вклада атомов коэффициент заполнения задается и уточняется отдельно для каждой "рабочей группы" рефлексов. Это позволяет, например, уточнять коэффициенты межплоскостного шкалирования.

В простейшем случае можно положить JXL=1 для всех рефлексов, JCRIN(1)=1. Тогда коэффициент заполнения будет одинаков для всех рефлексов.

При расчете вклада тяжелых атомов температурный фактор тяжелого атома может считаться как изотропным (значение величины NTROP в управляющих данных равно единице), так и анизотропным (NTROP =6). В первом случае в управляющих картах задается одно значение  $V_1$  температурного фактора. Во втором - шесть значений  $V_1, V_2, V_3, V_4, V_5, V_6$ , определяющих матрицу температурного фактора. Обратим внимание на то, что вычисления температурного фактора в изотропном и анизотропном случаях проводятся в программе в шкалах разной размерности.

Нормальная и аномальная (только при учете аномального рассеяния) компоненты вклада тяжелых атомов в структурный фактор тяжелоатомной производной рассчитываются по формулам:

$$\vec{F}_{Hj}^n = \sum_{q=1}^n RK_q(\vec{h}) * RZ_q(\vec{h}) * \sum_{j=1}^{NSYM} T_j(\vec{h}) e^{2\pi i(\vec{h}, R_j \vec{x}_q + \vec{t}_j)},$$

$$\vec{F}_{Hj}^{an} = \sum_{q=1}^n AK_q(\vec{h}) * AZ_q(\vec{h}) * \sum_{j=1}^{NSYM} T_j(\vec{h}) e^{2\pi i(\vec{h}, R_j \vec{x}_q + \vec{t}_j)},$$

где

$$T_j(\vec{h}) = \begin{cases} e^{-B_1 \left(\frac{\sin \theta}{\lambda}\right)^2} & \text{- в случае изотропного температурного фактора;} \\ e^{-(\vec{h}, R_j B R_j^T \vec{h})} & \text{- в случае анизотропного температурного фактора.} \end{cases}$$

Здесь  $n$  - число независимых (то есть не связанных симметрией группы) мест присоединения тяжелых атомов;

$RZ_q(\vec{h})$  - атомный фактор  $q$ -го тяжелого атома; берется постоянным внутри каждой из восьми равных частей заданного интервала  $\left(\frac{1}{2 * DMAX}, \frac{1}{2 * DMIN}\right)$  по  $\frac{\sin \theta}{\lambda}$ ;

$RK_q(\vec{h})$  - коэффициент заполнения для  $q$ -го тяжелого атома для той "рабочей группы" рефлексов, к которой относится данный рефлекс  $\vec{h} = (h, k, l)$ ;

$NSYM$  - число эквивалентных позиций в рассматриваемой пространственной группе;

$R_j, \vec{t}_j$  - матрица поворота и вектор трансляции для  $j$ -го преобразования симметрии;

$\vec{h} = (h, k, l)$  - вектор миллеровских индексов;

$\vec{x}_q$  - вектор координат  $(x_q, y_q, z_q)$   $q$ -го тяжелого атома (в долях периодов элементарной ячейки);

$B_1$  - изотропный температурный фактор;

$B$  - матрица анизотропного температурного фактора:

$$B = \begin{pmatrix} B_1 & \frac{1}{2}B_2 & \frac{1}{2}B_3 \\ \frac{1}{2}B_2 & B_4 & \frac{1}{2}B_5 \\ \frac{1}{2}B_3 & \frac{1}{2}B_5 & B_6 \end{pmatrix};$$

$AK_q(\vec{h})$  - коэффициент аномального замещения  $q$ -го тяжелого атома для "рабочей группы" рефлексов, к которой относится рефлекс  $\vec{h} = (h, k, l)$ ;

$AZ_q(\vec{h})$  - аномальная рассеивающая способность  $q$ -го атома.

#### §4. Расчет фаз структурных факторов нативного белка

Расчет фаз в программе осуществляется по методике,

восходящей к работе /1/ и описанной подробно в статье /2/.

Расчет фазы осуществляется для каждого рефлекса, принадлежащего заданной зоне разрешения [DMIN, DMAX] и имеющего положительную амплитуду структурного фактора нативного белка.

Для вычисления фазы используются тяжелоатомные производные, отмеченные ненулевым флагом NPHASE (J), причем при работе с очередным рефлексом во внимание принимаются только те из них, для которых значения модулей структурных факторов после шкалирования (гл. I, § 2) оказались положительными.

Для каждого обрабатываемого рефлекса программа вычисляет "плотность распределения вероятностей" значений фазового угла  $\varphi$  по формулам:

$$P(\varphi) = \prod_{j=1}^k p_j(\varphi),$$

$$p_j(\varphi) = \begin{cases} n_j \exp(-\varepsilon_j^2(\varphi)/2E_j) & \text{- при отсутствии аномального рассеяния;} \\ n'_j \exp(-\varepsilon_j^2(\varphi)/2E_j^2) \exp(-\delta_j^2(\varphi)/2A_j^2) & \text{- при учете аномального рассеяния;} \end{cases}$$

где  $k$  - число производных, участвующих в расчете фазы данного рефлекса;

$n_j, n'_j$  - нормирующие множители, определяемые из условия:

$$\int_0^{2\pi} p_j(\varphi) d\varphi = 1;$$

$\varepsilon_j(\varphi)$  - ошибка замкнутости фазового треугольника для  $j$ -й производной при условии, что фаза структурного фактора нативного белка есть  $\varphi$ ;

$\delta_j(\varphi)$  - ошибка замкнутости для бейфтовой пары рефлексов  $j$ -й производной при условии, что фаза структурного фактора нативного белка есть  $\varphi$ ;

$E_j, A_j$  - величины, характеризующие разброс "плотности распределения вероятностей" (об их определении см. ниже).

После определения функции  $P(\varphi)$  для рефлекса определяется "наиболее вероятная" (most probable) фаза  $\varphi_{m.p.}$ , "наилучшая" (best) фаза  $\varphi_{best}$  и показатель достоверности  $m$  (figure of merit).

Наиболее вероятная фаза  $\varphi_{m.p.}$  определяется по максимальному значению функции правдоподобия:

$$P(\varphi_{m.p.}) = \max P(\varphi), \quad 0 \leq \varphi \leq 2\pi.$$

Наилучшая фаза  $\varphi_{best}$  и показатель достоверности  $m$  определяются из соотношений:

$$m e^{i\varphi_{best}} = \int_0^{2\pi} P(\varphi) e^{i\varphi} d\varphi$$

или, что то же,

$$\begin{cases} m \cos \varphi = \int_0^{2\pi} P(\varphi) \cos \varphi d\varphi, \\ m \sin \varphi = \int_0^{2\pi} P(\varphi) \sin \varphi d\varphi. \end{cases}$$

Значения функции  $P(\varphi)$  вычисляются в программе в отдельных точках с шагом `ANGLE`, задаваемым в управляющих данных. Исключения составляют рефлексy, принадлежащие центросимметричным зонам, для которых функция  $P(\varphi)$  вычисляется только в двух точках, соответствующих двум возможным для этого рефлекса значениям фазового угла.

Фаза  $\varphi_{m.p.}$  находится просмотром значений  $\varphi$  в интервале  $(0, 2\pi)$ , для которых вычислена величина  $P(\varphi)$ . Если максимум  $P(\varphi)$  достигается в двух точках, то в качестве  $\varphi_{m.p.}$  берется первая из них. Исключение составляет случай, когда фаза рассчитывается по одной производной. Здесь  $\varphi_{m.p.}$  принимается равной фазе  $\varphi_{best}$ .

При определении  $\varphi_{best}$  и  $m$  интегралы рассчитываются по формуле прямоугольников с шагом `ANGLE`.

Значения функций  $\sin x$  и  $\cos x$  определяются по таблице с шагом  $1^\circ$ , значения функции  $\exp x$  по таблице с шагом  $0.02$ .

У рефлексов, фазы которых определились с нулевым показателем достоверности  $m$ , фазам  $\varphi_{m.p.}$  и  $\varphi_{best}$  присваиваются значения  $1000.0$ , что служит признаком неопределенности фазы.

Ошибки замкнутости фазового треугольника  $\varepsilon_j(\varphi)$  и  $\delta_j(\varphi)$  определяются по формулам:

$$\begin{aligned} \varepsilon_j(\varphi) &= F_{PHj}^{obs} - F_{PHj}^{calc}, \\ \delta_j(\varphi) &= \Delta_j^{obs} - \Delta_j^{calc}, \end{aligned}$$

где  $F_{PHj}^{obs}$ ,  $\Delta_j^{obs}$  - амплитуда и аномальная разность для  $j$ -й производной, полученная шкалированием входных данных (см. гл. I, § 2);

$$F_{\text{PHj}}^{\text{calc}} = \begin{cases} |F_{\text{P}} e^{i\varphi} + f_{\text{Hj}}^{\text{n}}| & \text{— при отсутствии аномального} \\ & \text{рассеяния;} \\ \frac{1}{2} \left[ |F_{\text{P}} e^{i\varphi} + f_{\text{Hj}}^{\text{n}} + f_{\text{Hj}}^{\text{an}}| + |F_{\text{P}} e^{i\varphi} + f_{\text{Hj}}^{\text{n}} - f_{\text{Hj}}^{\text{an}}| \right] & \text{—} \\ & \text{при учете аномального рассеяния;} \end{cases}$$

$$\Delta_j^{\text{calc}} = |F_{\text{P}} e^{i\varphi} + f_{\text{Hj}}^{\text{n}} + f_{\text{Hj}}^{\text{an}}| - |F_{\text{P}} e^{i\varphi} + f_{\text{Hj}}^{\text{n}} - f_{\text{Hj}}^{\text{an}}| ;$$

$F_{\text{P}}$  — значение амплитуды нативного белка;  
 $f_{\text{Hj}}^{\text{n}}, f_{\text{Hj}}^{\text{an}}$  — нормальная и аномальная компоненты вклада в структурный фактор тяжелых атомов  $j$ -й производной, вычисленные согласно §3, гл. I.

Величины  $E_j$  и  $A_j$ , характеризующие разброс "распределения вероятностей", берутся равными:

$$E_j^2 = E_{0j}^2 + (\zeta_{\text{PHj}}^{\text{calc}})^2 + (\zeta_{\text{PHj}}^{\text{obs}})^2,$$

$$A_j^2 = A_{0j}^2 + 4(\zeta_{\text{PHj}}^{\text{obs}})^2,$$

где  $\zeta_{\text{PHj}}^{\text{obs}}$  — стандартное отклонение для  $F_{\text{PHj}}^{\text{obs}}$ , шкалированное согласно §2, гл. I;  
 $\zeta_{\text{PHj}}^{\text{calc}}$  — стандартное отклонение для  $F_{\text{PHj}}^{\text{calc}}$ , принимаемое равным  $\sqrt{\left(\frac{\partial F_{\text{PHj}}}{\partial F_{\text{P}}}\right)^2} \zeta_{\text{P}}$ ;

$\zeta_{\text{P}}$  — стандартное отклонение для  $F_{\text{P}}$ ;

$E_{0j}, A_{0j}$  — величины, характеризующие ошибку изоморфизма и ошибку в определении параметров тяжелых атомов.

Величины  $E_{0j}$  и  $A_{0j}$ , зависящие от индексов  $h, k, l$ , задаются постоянными внутри каждой из восьми равных частей заданного для работы интервала по  $\frac{2\pi l \theta}{\lambda}$  и пересчитываются после каждого цикла работы программы.

На входе в программу величины  $E_{0j}$  и  $A_{0j}$  задаются в управляющих картах. Если при этом начальные значения величин сильно занижены, то на первых циклах работы программы показатели достоверности и значение R-фактора будут неоправданно хороши. Если же на входе в программу  $E_{0j}$  и  $A_{0j}$  заданы слишком большими, то на нескольких первых циклах значения показателей достоверности и R-фактора будут занижены.

Величины  $E_{0j}$  и  $A_{0j}$  пересчитываются на очередном

цикле для каждого подинтервала в пределах заданных границ разрешения по формулам:

$$(E_{0j}^{new})^2 = \frac{1}{M} \sum_1 (\varepsilon_j^2(\varphi_{best}) - (G_{PHj}^{calc})^2 - (G_{PHj}^{obs})^2),$$

$$(A_{0j}^{new})^2 = \frac{1}{N} \sum_2 (\delta_j^2(\varphi_{best}) - 4(G_{PHj}^{obs})^2).$$

Здесь  $G_{PHj}^{obs}$ ,  $G_{PHj}^{calc}$ ,  $\varepsilon_j$ ,  $\delta_j$  - определены выше;  
 $E_{0j}^{new}$ ,  $A_{0j}^{new}$  - новые значения величин  $E_{0j}$  и  $A_{0j}$ ,  
 которые будут использоваться на следующем цикле работы программы;

$M$  - число обработанных рефлексов из центросимметричных зон, для которых  $\frac{\sin \theta}{\lambda}$  лежит в рассматриваемом подинтервале, и таких, что для них фаза определена с ненулевым показателем достоверности и  $\varepsilon_j^2(\varphi_{best}) \leq 16E_j^2$ ;

$N$  - число обработанных рефлексов, для которых  $\frac{\sin \theta}{\lambda}$  лежит в рассматриваемом подинтервале, и таких, что фаза для них определена с ненулевым показателем достоверности и  $\delta_j^2(\varphi_{best}) \leq 16A_j^2$ ;

$\sum_1$  - суммирование ведется по рефлексам, обладающим указанными в определении величины  $M$  свойствами и таким, что

$$\varepsilon_j^2(\varphi_{best}) - (G_{PHj}^{calc})^2 - (G_{PHj}^{obs})^2 > 0;$$

$\sum_2$  - суммирование ведется по рефлексам, обладающим указанными в определении величины  $N$  свойствами и таким, что  $\delta_j^2(\varphi_{best}) - 4(G_{PHj}^{obs})^2 > 0$ .

Подчеркнем, что при вычислении  $E_{0j}$  учитываются лишь рефлексы центросимметричных зон, а при вычислении  $A_{0j}$  - все рефлексы. Исключением является случай, когда во входных данных не указано наличие центросимметричных зон ( $NCEN=0$ ). Тогда  $E_{0j}$  определяется по всем рефлексам.

Отметим еще, что если при вычислении величины  $E_j$  (либо  $A_j$ ) мы получим нулевое значение, то при расчёте фазы будет считаться, что  $E_j=1$  (соответственно  $A_j=1$ ).

#### §5. Уточнение параметров тяжелоатомных производных

Для каждой производной в процессе работы программы может уточняться ряд параметров. При этом, если задано

уточнение какого-либо параметра, то его уточнение будет производиться на каждом цикле работы программы.

В процессе уточнения параметры каждой производной уточняются независимо от других производных. Для каждой производной могут уточняться шкальные коэффициенты  $Q(J)$  и  $BOVER(J)$  (см. гл. I, §2), коэффициенты заполнения  $RK_q$  и  $AK_q$  (см. гл. I, §3), координаты и компоненты температурного фактора для каждого из тяжелых атомов.

В случае наличия у  $J$ -й производной параметров, которые требуется уточнить, необходимо задать ненулевой флаг  $NREFIN(J)$  в управляющих данных. В этом случае всегда будет уточняться шкальный коэффициент  $Q(J)$ . Остальные параметры будут уточняться, если для них будут заданы ненулевые соответствующие флаги, управляющие уточнением параметров. При этом флаги, отвечающие за уточнение параметров тяжелых атомов задаются индивидуально для каждого атома. Соответствие флагов уточняемым параметрам следующее:

- IFBOV - уточнение шкального коэффициента BOVER;
- LABEL1 - уточнение коэффициентов заполнения RK (для всех "рабочих групп" рефлексов отдельно);
- LABEL2 - уточнение компонент температурного фактора (в случае анизотропного температурного фактора уточняются все шесть компонент);
- LABEL3 } - уточнение координат x, y, z по данным
- LABEL4 } - изоморфных разностей;
- LABEL5 }
- LABEL6 - уточнение аномальной рассеивающей способности (для всех "рабочих групп" рефлексов);
- LABEL7 - уточнение компонент температурного фактора по данным аномального рассеяния;
- LABEL8 } - уточнение координат x, y, z по данным
- LABEL9 } - аномальных разностей.
- LABELA }

Флаг IFBOV задается отдельно для каждой производной, а флаги LABEL отдельно для каждого атома каждой производной.

Уточнение каких-либо параметров по данным аномального рассеяния возможно лишь при ненулевом флаге  $KAY(J)$  для  $J$ -й производной. При этом уточнение какого-либо параметра будет производиться по данным аномального рассея-

ния только в том случае, если для этого параметра не было ранее указано уточнение по данным нормального рассеяния.

В основе уточнения параметров  $J$ -й производной лежит минимизация функций

$$R = \sum_{\vec{h}} \tau_{\vec{h}}^2,$$

где

$$\tau_{\vec{h}}^2 = \begin{cases} \varepsilon_j^2(\varphi_{best}) / [(\sigma_{\rho_{nj}}^{obs})^2 + (\sigma_{\rho_{nj}}^{calc})^2] & \text{при } (\sigma_{\rho_{nj}}^{obs})^2 + (\sigma_{\rho_{nj}}^{calc})^2 \neq 0, \\ \varepsilon_j^2(\varphi_{best}) & \text{при } (\sigma_{\rho_{nj}}^{obs})^2 + (\sigma_{\rho_{nj}}^{calc})^2 = 0; \end{cases}$$

$$R^{an} = \sum_{\vec{h}} (\tau_{\vec{h}}^{an})^2,$$

где

$$\left(\tau_{\vec{h}}^{an}\right)^2 = \begin{cases} \delta_j^2(\varphi_{best}) / (\sigma_{\rho_{nj}}^{obs})^2 & \text{при } \sigma_{\rho_{nj}}^{obs} \neq 0, \\ \delta_j^2(\varphi_{best}) & \text{при } \sigma_{\rho_{nj}}^{obs} = 0 \end{cases}$$

(функция  $R$  используется для уточнения параметров на базе нормального,  $R^{an}$  - аномального рассеяния)

$\varepsilon_j(\varphi_{best})$ ,  $\delta_j(\varphi_{best})$  - ошибки замкнутости фазового треугольника при фазе  $\varphi_{best}$  структурного фактора нативного белка (см. гл. I, 84);

$\sigma_{\rho_{nj}}^{obs}$  - стандартное отклонение для  $F_{\rho_{nj}}^{obs}$ ;

$\sigma_{\rho_{nj}}^{calc}$  - стандартное отклонение для вычисленной амплитуды  $F_{\rho_{nj}}^{calc}$ , принимаемое равным  $\sqrt{\left(\frac{\partial F_{\rho_{nj}}^{calc}}{\partial F_{\rho}}\right)^2 \sigma_{\rho}^2}$ ;

$\sigma_{\rho}$  - стандартное отклонение для  $F_{\rho}$ ;

$\sum$  - суммирование ведется по всем рефлексам таким, что фаза для них определена с ненулевым показателем достоверности и

$$\varepsilon_j^2(\varphi_{best}) < 16[(E_{0j})^2 + (\sigma_{\rho_{nj}}^{obs})^2 + (\sigma_{\rho_{nj}}^{calc})^2] \quad \text{для } R,$$

$$\delta_j^2(\varphi_{best}) < 16[(A_{0j})^2 + 4(\sigma_{\rho_{nj}}^{obs})^2] \quad \text{для } R^{an}$$

(ошибки изоморфизма  $E_{0j}$  и  $A_{0j}$  определены в §4 гл. I).

Для минимизации описанных выше функций  $R$  и  $R^{an}$  применяется стандартная разновидность "линеаризованного метода наименьших квадратов". Именно, пусть нам необходимо минимизировать функцию вида:

$$R(d_1, d_2, \dots, d_n) = \sum_{q=1}^N r_q^2(d_1, \dots, d_n),$$

зависящую от параметров  $d_1, d_2, \dots, d_n$ . Предлагается определить смещения параметров  $\delta_1, \dots, \delta_n$  из системы линейных уравнений:

$$A \vec{\delta} = \vec{b},$$

где  $A = ((a_{ij}))$  - матрица с элементами

$$a_{ij} = \sum_{q=1}^N \frac{\partial r_q}{\partial d_i} \cdot \frac{\partial r_q}{\partial d_j},$$

$\vec{b} = (b_1, \dots, b_n)$  - вектор правых частей системы:

$$b_i = - \sum_{q=1}^N r_q \frac{\partial r_q}{\partial d_i}$$

и затем взять в качестве уточненных значений параметров величины:  $d_1 + \delta_1, d_2 + \delta_2, \dots, d_n + \delta_n$ .

В данной программе реализуется описанная выше методика, причем производные по уточняемым параметрам вычисляются аналитически в предположении, что фазы  $\varphi_{best}$  нативного белка от этих параметров не зависят.

Для решения системы линейных уравнений применяется метод разложения матрицы системы в произведение треугольных матриц, описанный в работе /3/.

Для контроля процесса уточнения вычисляются параметр уточнения  $S$ , стандартные отклонения  $G_i$  уточняемых параметров и корреляционная матрица  $K = ((k_{ij}))$  по формулам:

$$S = \frac{\sum_{q=1}^N r_q^2(d_1, \dots, d_n)}{N-n} = \frac{R}{N-n}; \quad G_i = \sqrt{C_{ii} * S}, \quad k_{ij} = \frac{C_{ij}}{\sqrt{C_{ii} * C_{jj}}}$$

( $S = \frac{R^{an}}{N-n}$ , если уточнение по данным аномального рассеяния),

где  $n$  - число уточняемых параметров;

$N$  - число рефлексов, участвующих в уточнении;

$R$  - минимизируемая функция;

$C_{ij}$  - элемент матрицы  $C = ((c_{ij}))$ , обратной к матрице системы уравнений для определения смещений параметров.

Программа допускает возможность регулировать процесс уточнения параметров изменением двух величин THR и ACL, задаваемых в управляющих данных на перфокартах. Регулировка осуществляется принудительным уменьшением тех смещений, которые велики по сравнению со стандартными отклонениями соответствующих параметров. Если вычисленное значение приращения  $\delta_i$  для параметра  $\mathcal{L}_i$  таково, что  $|\delta_i| > THR * \sigma_i$  ( $\sigma_i$  - стандартное отклонение для  $\mathcal{L}_i$ ), то в качестве нового значения параметра берется  $\mathcal{L}_i + ACL * \delta_i$ .

Отметим, что минимизируемые функции R и  $R^{an}$  могут меняться в процессе работы - число слагаемых, участвующих в их определении, зависит от текущих значений величин  $E_{0j}$  и  $A_{0j}$ . Это может приводить к тому, что на некоторых этапах уточнения значения величин R и  $R^{an}$  будут несколько возрастать.

#### 86. Организация программы

Фазовая программа состоит из вызывающей программы и восьми подпрограмм (см. приложение 1). Все программные единицы написаны на языке Фортран IV за исключением подпрограммы CLOKIT. Назначение этой подпрограммы - вывести при обращении к ней

```
CALL CLOKIT
```

на печать время суток. Эта подпрограмма может быть либо написана для конкретной операционной системы на автокоде, либо заменена фиктивной:

```
SUBROUTINE CLOKIT
RETURN
END.
```

Главная программа вводит управляющие данные с перфокарт и организует циклы расчета фаз и уточнения параметров. Для ввода и вывода структурных факторов она обращается к подпрограммам INPUT и OUTPUT. Подпрограмма PHASE для данного рефлекса определяет фазы  $\varphi_{m.p.}$  и  $\varphi_{best}$  и выводит (если требуется) данные для Фурье-синтезов. Подпрограмма FINAL обрабатывает и выводит статистику по процессу уточнения параметров тяжелоатомных

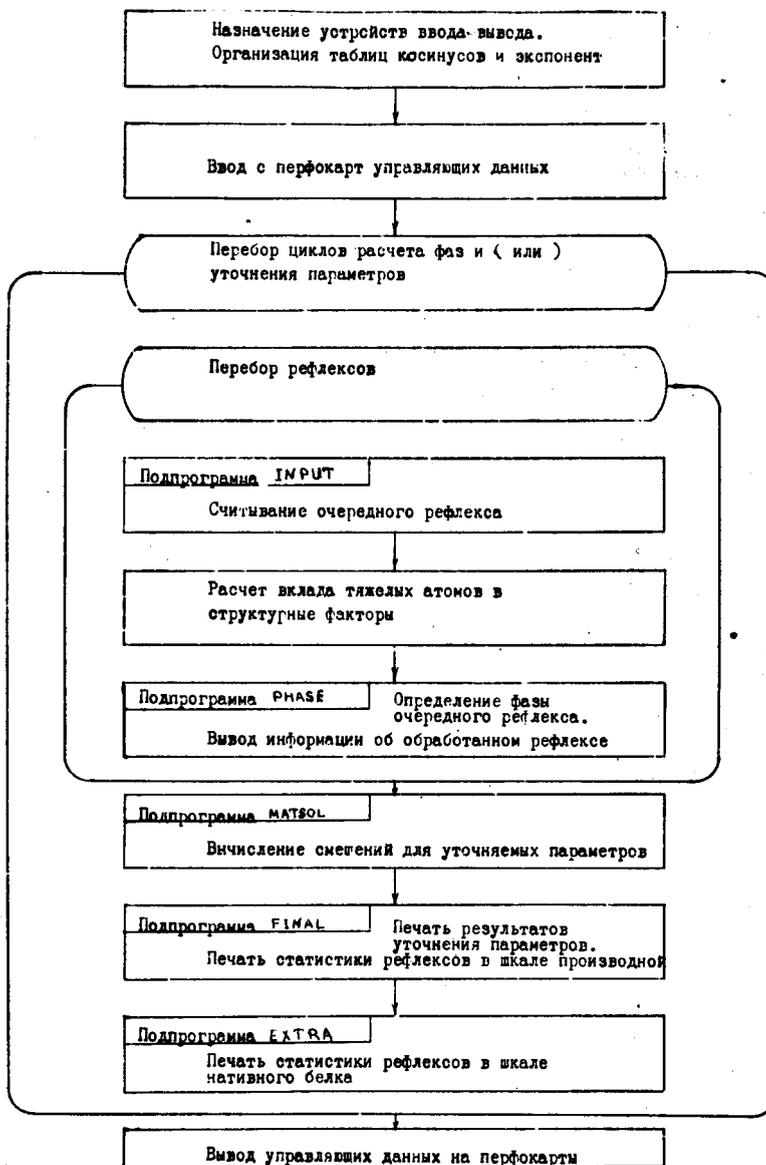


Рис. 1. Схема работы фазовой программы

производных, а подпрограмма EXTRA - статистику по расчету фаз. Подпрограмма MATSOL решает систему линейных уравнений в процессе уточнения параметров. Подпрограмма HIDE организует запись в рабочий файл и чтение из него промежуточных результатов. Блок-схема программы приведена на рис. 1.

## ГЛАВА II. ВХОДНЫЕ И ВЫХОДНЫЕ ДАННЫЕ ПРОГРАММЫ

В зависимости от указанного характера работы программа использует от четырех до семи внешних устройств. Сведения об использовании внешних устройств даны в табл. 1.

Обычно для файла 1 используется ввод перфокарточный, файла 2 - вывод перфокарточный, файла 3 - алфавитно-цифровое печатающее устройство, файлов 4, 5, 6, 7 - магнитные ленты или диски.

Ссылочные номера используемых файлов присваиваются идентификаторам этих файлов в самом начале работы вызывающей программы. Поэтому в случае необходимости они могут быть легко изменены.

### §1. Управляющие данные на перфокартах

Информация о параметрах элементарной ячейки, о пространственной группе, о параметрах тяжелых атомов и управляющие работой программы флаги задаются на входе с перфокарт. При работе программы предполагается, что все рассмотрения ведутся в единой системе координат. В частности, координаты тяжелых атомов в различных производных перед работой с фазовой программой должны быть приведены к общему началу. Относительно этого же начала задается информация о центросимметричных зонах (если они есть) и рассчитываются фазы рефлексов.

Данные на перфокартах вводятся в программу из файла с идентификатором MONINP, причем в начале работы программы этому идентификатору присваивается значение 1. При необходимости заменить в программе номер этого файла достаточно заменить одну карту: MONINP=1 в тексте программы.

Таблица 1

Идентификатор файла в программе	Ссылочный номер	Назначение файла	Необходимость присутствия
MONINP	1	Ввод управляющих данных с перфокарт	Всегда
MONCDP	2	Вывод управляющих данных (на перфокарты)	Всегда
MONPRT	3	Вывод информации на печать	Всегда
BININP	4	Ввод структурных амплитуд	Всегда
BINOUT	5	Вывод структурных факторов после расчета фаз	Если указан такой вывод
BINFOU	6	Вывод данных для Фурье-синтезов	Если указан такой вывод
BINSCR	7	Рабочий файл	Используется, если на некоторых циклах пересчет фаз не производится

Примечание. Все используемые файлы рассматриваются как файлы последовательного доступа.

Ниже приведено описание последовательности перфокарт с данными, требуемыми для работы программы. В круглых скобках приведены описания форматов, согласно которым считываются эти величины.

Последовательность данных следующая:

**A** (TITLE(J), J=1, 20) (20A4) - произвольный текст, выводимый периодически на печать в ходе работы программы.

**B** A, B, C, ALPHA, BETA, GAMMA (6F10.0) - параметры элементарной ячейки. Стороны A, B, C, задаются в ангстремах, углы ALPHA, BETA, GAMMA - в градусах.

**C** Информация о пространственной группе:

а) (GROUP(J), J=1, 3), NSYM (3A4, 13) - название пространственной группы GROUP (несущественно для работы - используется только для вывода на печать), число эквивалентных позиций NSYM (преобразований симметрии) в группе, включая тождественное.

б) (T(J), J=1, 12) (12 F 6.0) (NSYM карт) - для каждой из NSYM эквивалентных позиций соответствующее преобразование симметрии представляется в виде  $x = Rx + t$  и для него задается матрица поворота и вектор трансляции:

$$R = \begin{pmatrix} T(1) & T(2) & T(3) \\ T(5) & T(6) & T(7) \\ T(9) & T(10) & T(11) \end{pmatrix}, \quad \vec{t} = \begin{pmatrix} T(4) \\ T(8) \\ T(12) \end{pmatrix}.$$

Порядок, в котором следуют карты различных преобразований симметрии, не существует.

с) NCEN (15) - число centrosимметричных зон во входных данных. Это число может быть равным нулю. Если во входных данных нет рефлексов из centrosимметричных зон, то NCEN необходимо задать равным нулю, даже если в пространственной группе такие зоны есть.

д) CT1, CT2, CT3, CENT1, CENT2, CENT3 (6 F 10.0) (NCEN карт, при NCEN ≠ 0).

Условие принадлежности рефлекса данной centrosимметричной зоне есть:  $CT1 \cdot h + CT2 \cdot k + CT3 \cdot l = 0$  и один из двух возможных для этого рефлекса фазовых углов

$$\varphi = CENT1 \cdot h + CENT2 \cdot k + CENT3 \cdot l.$$

При NCEN=0 группа карт d отсутствует. Порядок, в котором следуют карты для различных зон, не существует.

**D** DMIN, DMAX, SCALE, SENAT, THR, ACL (6 F 10.0) DMIN, DMAX - нижняя и верхняя границы

зоны разрешения в прямом пространстве в ангстремах; обрабатываются только те рефлексы, для которых  $DMIN \leq d \leq DMAX$ , где  $d = \frac{\lambda}{2 \sin \theta}$  - межплоскостное расстояние в прямом пространстве. SCALE, SENAT - шкальные коэффициенты (см. гл. I, §2.). Все структурные амплитуды после ввода умножаются на SCALE, стандартные отклонения для амплитуд структурных факторов нативного белка умножаются на SENAT.

THR, ACL - параметры, используемые для управления процессом уточнения параметров (см. гл. I, §5). Если получающиеся при уточнении сдвиги параметров больше, чем  $THR * \sigma$  (здесь  $\sigma$  - стандартное отклонение для соответствующего параметра), то сдвиги перед добавлением умножаются на ACL. Если задано  $THR \leq 0$ , то программа устанавливает  $THR = 10^6$ . В случае  $ACL \leq 0$  программа устанавливает  $ACL = 1,0$ .

**E** IANGLE, IOUT, KOUT (315) - при расчете фаз "плотность распределения вероятностей" значений фазовых углов для рефлексов, не принадлежащих centrosymmetric зонам, рассчитывается с шагом IANGLE градусов. При значении флага IOUT  $\neq 0$  производится вывод результата работы программы в файл BINOUT (см. гл. II, §4). Значение флага KOUT контролирует вывод в файл BINFOU различного рода информации для синтезов Фурье согласно таблице, приведенной в §5.

**F** NCYCLE, (KCYCLE(J), J=1, NCYCLE) (1615). Работа программы состоит из NCYCLE циклов пересчета фаз и (или) уточнения параметров. Если KCYCLE(J)  $\neq 0$ , то на J-м цикле работы программы фазы будут вычисляться заново. В противном случае для уточнения параметров на этом цикле будут использоваться фазы, полученные на предыдущих циклах.

**G** NCPDS (15) - число тяжелоатомных производных, данные по которым будут использоваться программой в работе.

Далее для каждой из производных следует ввести данные по этой производной. В программе производные нумеруются в том порядке, в каком поступает о них эта информация.

I. (CTITLE(J), J=1, 20) (20A4) - текст, идентифицирующий производную, который периодически выводится на печать.

II. NATOMS, NPHASE, NREFIN, KAY, MAP (5I5). Здесь NATOMS - число независимых (то есть не связанных симметрией группы) мест присоединения тяжелых атомов в текущей производной. Если флаг NPHASE $\neq$ 0, то эта производная используется при определении фаз структурных факторов нативного белка. В противном случае параметры производной могут уточняться, но при расчете фаз эта производная не используется. Если флаг NREFIN $\neq$  0, то для текущей производной будут уточняться какие-либо параметры. Флаг KAY $\neq$ 0, если для данной производной предполагается использовать данные по аномальному рассеянию. Номер MAP указывает порядковый номер данных по этой производной во входном файле BININP (см. гл. II, §3).

III. MXL, NXL, (JCRIN(J), J=1, MXL) (2I5, 35I2/40I2) - информация о группировке данных в "рабочие группы" рефлексов, как это описано в гл. I, §3. Здесь MXL - число "входных групп" рефлексов (максимальное значение условного номера для рефлексов данной производной); NXL - число "рабочих групп" рефлексов; JCRIN(J) - номер "рабочей группы", к которой будет отнесен рефлекс с условным номером J.

IV. SCESD, Q, BOVER, IFBOV (3F10.0, I5). SCESD, Q, BOVER - шкальные коэффициенты для текущей производной (см. гл. I, §2). Если флаг IFBOV $\neq$ 0, то шкальный коэффициент BOVER будет уточняться.

V. Далее для каждого из NATOMS тяжелых атомов следует группа карт a-g, содержащих информацию по этому атому. В программе далее атомы нумеруются в том порядке, в котором поступает о них эта информация.

a) (RK(J), J=1, NXL) (8F10.0). Коэффициент заполнения как "функция номера "рабочей группы" (см. гл. I, §3).

b) (RZ(J), J=1, 8) (8F10.0). Атомный формфактор как функция межплоскостного расстояния  $d^*$  в обратном пространстве. Заданная зона разрешения в обратном пространстве  $\frac{1}{D_{MAX}} \leq d^* \leq \frac{1}{D_{MIN}}$  разбивается на восемь равных частей, и атомный фактор задается постоянным для каждой из таких частей.

c) Эта карта присутствует только в том случае, если флаг KAY для текущей производной не нулевой.

(AK(J), J=1, NXL) (8F10.0).

Аномальный коэффициент заполнения как функция номера "рабочей группы".

d) Эта карта присутствует только в том случае, когда для текущей производной флаг KAY $\neq$ 0.

(AZ(J), J=1, 8) (8F10.0).

Аномальная рассеивающая способность атома как функция межплоскостного расстояния в обратном пространстве.

e) NTROP, B1, B2, B3, B4, B5, B6, (15, 6F10.0).  
где NTROP - размерность температурного фактора атома. Если NTROP $\neq$ 6, то программа устанавливает NTROP=1, и в этом случае для работы используется только величина B1 температурного фактора. Если NTROP=6, то температурный фактор анизотропен с матрицей

$$\begin{pmatrix} B1 & \frac{1}{2}B2 & \frac{1}{2}B3 \\ \frac{1}{2}B2 & B4 & \frac{1}{2}B5 \\ \frac{1}{2}B3 & \frac{1}{2}B5 & B6 \end{pmatrix}.$$

f) X, Y, Z (3F10.0)

Координаты атома в долях элементарной ячейки.

g) Эта карта присутствует только в том случае, когда для текущей производной флаг NREFIN $\neq$ 0.

LABEL1÷LABEL9, LABELA (8F10.0).

Флаги, определяющие набор уточняемых параметров согласно гл. I §5.

Далее группа карт a-g повторяется для второго атома текущей производной, затем для третьего и т.д. - для каждого из NATOMS атомов.

VI. (ESQ(J), J=1, 8) (8F10.0).

Среднеквадратичная величина ошибки изоморфизма как функция межплоскостного расстояния в обратном пространстве. Заданная зона разрешения в обратном пространстве разбивается на восемь равных частей, и среднеквадратичная величина ошибки изоморфизма считается постоянной для каждой из таких частей. Эта величина служит начальным значением для величины E<sub>0</sub> (см. гл. I, §4).

VII. Эта карта присутствует только в том случае, когда

для текущей производной флаг  $KAY \neq 0$ .

(ASQ(J), J=1, 8) (8F10.0).

Среднеквадратичная ошибка изоморфизма для аномального рассеяния как функция межплоскостного расстояния в обратном пространстве. Служит начальным значением для величины  $A_0$  (см. гл. I, §4).

Далее группа карт I-VII повторяется для второй производной, затем для третьей и т.д. - для каждой из NCPDS производных.

#### Ограничения на управляющие данные

Размеры массивов, используемых в программе, определены в предположении, что выполнены следующие условия:

- а) число эквивалентных позиций в рассматриваемой группе не превосходит 12;
- б) число центросимметричных зон во входных данных не превосходит 6;
- в) число циклов работы программ не превосходит 10;
- г) число производных, используемых в работе, не превосходит 6;
- д) общее число тяжелых атомов во всех производных не превосходит 30 (считая для каждой производной лишь атомы в независимой части ячейки);
- е) число "входных групп" рефлексов не превосходит 100;
- ж) число "рабочих групп" рефлексов не превосходит 8;
- з) выполнены условия:

$$\sum_{j=1}^k \frac{n_j(n_j+1)}{2} + \sum_{j=1}^k \frac{an_j(an_j+1)}{2} \leq 3000, \quad \sum_{j=1}^k n_j + \sum_{j=1}^k an_j \leq 150,$$

где  $k$  - число производных, для которых уточняются параметры;

$n_j$  - число параметров  $j$ -й производной, уточняемых по данным нормального рассеяния;

$an_j$  - число параметров  $j$ -й производной, уточняемых по данным аномального рассеяния.

## §2. Вывод данных на перфокарты

По завершении работы программы число тяжелоатомных производных и информация по каждой из них выводится на перфокарты (или в образах перфокарт в файл последовательного доступа). По характеру информация о производных в точности повторяет вводную информацию (гл. II, §1, группа данных I-VII) за следующими исключениями:

а) все параметры, для которых задавалось уточнение, выводятся уточненными;

б) для среднеквадратичных ошибок изоморфизма (карты VI, VII в управляющих данных) выводятся их значения, определенные на последнем цикле работы программы;

в) флаги LABEL1; LABELA выводятся соответствующими проведенному уточнению.

Вывод указанных данных осуществляется в файл с идентификатором MONCDP, которому в начале работы программы присваивается значение 2. При необходимости изменить в программе номер этого файла достаточно заменить одну карту MONCDP=2 в главной программе.

## §3. Входные данные на магнитной ленте

Экспериментально определенные амплитуды структурных факторов нативного белка и тяжелоатомных производных считываются программой из файла последовательного доступа. Мы будем для конкретности говорить далее о магнитной ленте. Рефлексы обрабатываются в том порядке, в каком они записаны на магнитной ленте. Информация по каждому обрабатываемому рефлексу считывается перед началом его обработки и хранится в оперативной памяти в течение работы над этим рефлексом.

Конец данных на магнитной ленте определяется программой по значению  $-2.0$  величины  $f_p^{\text{tape}}$  (см. ниже). Перед началом очередного цикла работы лента перематывается программой к началу.

Информация по каждому рефлексу считывается отдельным блоком длиной в 41 слово. Это считывание осуществляется из файла с идентификатором BININP, причем в начале работы программы этому идентификатору присваивается значе-

ние 4. При необходимости изменить в программе номер файла достаточно заменить одну карту BININP=4 в начале текста программы.

Считывание информации по каждому рефлексу осуществляется в подпрограмме INPUT, поэтому при необходимости изменения формата входных данных следует заменить подпрограмму INPUT.

Блок данных длиной в 41 слово, считываемый для каждого рефлекса бесформатным оператором READ, содержит информацию по этому рефлексу согласно табл. 2.

Т а б л и ц а 2

№ слова в блоке	Обозначение величины	Назначение величины
1	2	3
1-3	$h, k, l$	Миллеровские индексы рефлекса
4	$\varphi_{best}$	"Наилучшая" фаза структурного фактора нативного белка
5	$\varphi_{m.p.}$	"Наиболее вероятная" фаза структурного фактора нативного белка
6	$m$	Показатель достоверности определения фазы
7	$F_p^{tape}$	Амплитуда структурного фактора нативного белка
8	$\sigma_p^{tape}$	Стандартное отклонение для $F_p^{tape}$ (считается, что экспериментальная ошибка в определении $F_p^{tape}$ распределена нормально с таким стандартным отклонением)
9	NXLPRO	Программой не используется
10-17	$F_{PH1}^{tape} \div F_{PH8}^{tape}$	Амплитуды структурных факторов тяжелоатомных производных
18-25	$\Delta_1^{tape} \div \Delta_8^{tape}$	Аномальные разности для тяжелоатомных производных

1	2	3
26-33	$\sigma_{pH1}^{tape} \div \sigma_{pH8}^{tape}$	Стандартные отклонения для амплитуд структурных факторов тяжелоатомных производных (считается, что экспериментальная ошибка в определении $F_{pH}^{tape}$ распределена нормально с таким стандартным отклонением)
34-41	JXL(1)÷JXL(8)	Условные номера рефлексов для тяжелоатомных производных (см. гл. I, §3)

Величины NXLPPO, JXL(1)÷JXL(8) представляются в ЭВМ в формате целых чисел, остальные величины - в формате вещественных чисел.

#### Примечания

- 4-6 Если значения величин  $\varphi_{best}$ ,  $\varphi_{m.p.}$ ,  $m$  еще не определены, рекомендуется присвоить значения  
 $\varphi_{best} = 1000.0$ ,  $\varphi_{m.p.} = 1000.0$ ,  $m = 0.0$
- 7 Значение  $F_p^{tape} = -2.0$  означает конец данных на магнитной ленте.
- 10-41 На ленте могут быть записаны данные по восьми производным, однако, в работе будут участвовать не все из них, а только те, которые будут указаны особо во входных данных. При наличии данных менее чем по восьми производным свободные места должны быть заполнены произвольной информацией (например, нулями).
- 10-17 В случае использования для  $j$ -й производной данных по аномальному рассеянию здесь  $f_{pHj}^{tape} = \frac{1}{2}(F_{pHj}^{+tape} + F_{pHj}^{-tape})$ , где  $F_{pHj}^{+tape}$ ,  $F_{pHj}^{-tape}$  - амплитуда бейфутовой пары рефлексов.

18-25 Здесь даны величины  $\Delta_j^{tape} = F_{PHj}^{+tape} - F_{PHj}^{-tape}$ .  
Если для некоторого рефлекса задано  $\Delta_j^{tape} = 999.0$ ,  
то для такого рефлекса при расчете фазы структурно-  
го фактора нативного белка данные по аномальному  
рассеянию использоваться не будут.

---

#### §4. Выходные данные на магнитной ленте

На последнем цикле работы программы входная информация, считанная из файла BININP (см. гл. II, §3), может быть выведена в другой файл последовательного доступа с измененными согласно проделанной работе значениями фаз. Такой вывод осуществляется, если в управляющих данных флаг ICUT  $\neq 0$ .

При наличии указанного вывода информация выводится в том же виде, в котором она вводилась из файла BININP, за следующими исключениями:

а) фазы  $\varphi_{best}$ ,  $\varphi_{m.p.}$  и показатель достоверности  $m$  выводятся такими, какими они оказались после последнего пересчета фаз;

б) если фаза  $\varphi_{best}$  для некоторого рефлекса определилась с нулевым показателем достоверности, то выводится  $\varphi_{best} - 1000.0$ ,  $m = 0.0$ ;

в) рефлексы, для которых с входной ленты была считана неположительная величина  $F_p^{tape}$ , в выводной файл не берутся;

г) рефлексы, выходящие за пределы заданной зоны разрешения, отбрасываются.

В конце выводимой информации записывается блок с  $F_p^{tape} = -2.0$ , сигнализирующий о конце данных в файле.

Указанный вывод осуществляется в файл с идентификатором BINOUT, которому в начале работы программы присваивается значение 5. Для замены номера файла в программе нужно заменить оператор BINOUT=5. Запись информации по каждому рефлексу осуществляется в подпрограмме OUTPUT, поэтому при необходимости изменения формата выходных данных следует заменить подпрограмму OUTPUT.

### §5. Вывод данных для синтезов Фурье

На последнем цикле работы программы кроме вывода, описанного в предыдущем параграфе, может быть осуществлен вывод в файл последовательного доступа информации для различного рода синтезов. Возможность и характер вывода задаются значением флага KOUT в управляющих данных.

KOUT	Характер выводимой информации
0	Информация не выводится
1	$F_{PHj}^{obs} e^{i\alpha_j}$ - синтез Фурье для производной
2	$(F_{PHj}^{obs} - F_P) e^{i\psi_{best}}$ - разностный синтез Фурье (производная - белок)
3	$-i(F_{PHj}^{+obs} - F_{PHj}^{-obs}) e^{i\psi_{best}}$ - аномальный синтез Фурье
4	$(F_{PHj}^{obs} - F_{PHj}^{calc}) e^{i\alpha_j}$ - синтез Фурье фазовых ошибок для производной
5	$(F_{PHj}^{obs} - F_P)^2$ - разностный синтез Патерсона для производной

Здесь  $F_P^{obs}$ ,  $F_{PHj}^{obs}$  - амплитуды структурных факторов нативного белка и тяжелоатомных производных, шкалированные в шкале нативного белка (см. гл. I, §2);  $F_{PHj}^{+obs}$ ,  $F_{PHj}^{-obs}$  - амплитуды бейфутовой пары рефлексов в той же шкале;  $\psi_{best}$  - "наилучшая" фаза структурного фактора нативного белка;  $\alpha_j$  - фаза структурного фактора j-й тяжелоатомной производной, рассчитанная в предположении, что для белка фаза есть  $\psi_{best}$  из равенства

$$F_{PHj}^{obs} e^{i\alpha_j} = F_P^{obs} e^{i\psi_{best}} + \vec{f}_{Hj},$$

где  $\vec{f}_{Hj}$  - вклад тяжелых атомов в структурный фактор j-й производной.

В случае, когда флагу KOUT присвоено значение меньше нуля или больше пяти, оно автоматически устанавливается равным нулю.

Информация по каждому рефлексу записывается отдельным блоком длиной в 19 слов. Запись производится в подпрограмме PHASE в файл с идентификатором BINFOU, которому в начале работы программы присваивается значение 6. При необходимости сменить номер устройства требуется заменить карту BINFOU=6 в начале главной программы. Данные для синтеза выводятся только для тех рефлексов, для которых фаза определена с ненулевым показателем достоверности. В конце выводимой информации записывается блок со значением показателя достоверности  $m = -1.0$ , сигнализирующий о конце данных.

Блок данных длиной 19 слов, записываемый для каждого рефлекса бесформатным оператором WRITE, содержит информацию по этому рефлексу согласно следующей таблице. При этом все величины записываются как вещественные числа.

№ слова в блоке	Обозначение величины	Назначение величины
1	2	3
1-3	$h, k, l$	Миллеровские индексы рефлекса
4	$m$	Показатель достоверности определения фазы структурного фактора нативного белка
5	$\psi_{m,p}$	"Наиболее вероятная" фаза структурного фактора нативного белка
6	$F_P^{obs}$	Амплитуда структурного фактора нативного белка, шкалированная в шкале нативного белка
7-12	$F_1 \div F_6$	$F_j$ - модуль коэффициента Фурье заданного синтеза для $j$ -й производной

1	2	3
13	$\varphi_{best}$	"Наилучшая" фаза структурного фактора нативного белка
14-19	$A_1 \div A_6$	$A_j$ - фаза коэффициента Фурье заданного синтеза для $j$ -й производной

#### Примечания

- 7-12 Значения модулей коэффициентов Фурье выводятся неотрицательными. При выводе данных производные просматриваются в том порядке, в котором информация о них поступила с управляющих карт в программу. Если число назначенных для работы в программе производных меньше шести, то оставшиеся неопределенными величины  $F_j$  полагаются равными нулю.
- 14-19 Значения фаз коэффициентов Фурье выводятся лежащими на отрезке  $[0, 2\pi]$ . Если число производных в программе меньше шести, то оставшиеся неопределенными величины  $A_j$  полагаются равными нулю.

#### §6. Сообщения программы

По ходу работы программы на печать выводится ряд сообщений, отражающих ход расчета фаз и уточнения параметров. Эти данные выводятся в программе в файл с идентификатором MONPRT, которому в начале работы присваивается значение 3 (устройство SYSLSY в операционной системе DOS EC). При необходимости изменить в программе номер этого файла достаточно заменить одну карту MONPRT=3 в тексте программы.

Первая часть сообщений дублирует вводимую программой с перфокарт информацию. В качестве примера в приложении 3 приведены протоколы 1 и 2, комментарии к которым даны ниже, где идентификаторы величины выбраны такими же, как в §1 гл.

Протокол 1.

1. Заголовок фазовой программы.
  2. Текст TITLE, идентифицирующий программу.
  3. Кристаллографические параметры: A, B, C - ребра (в ангстремах), ALPHA, BETA, GAMMA - углы (в градусах) элементарной ячейки в прямом пространстве.  
A\*, B\*, C\*, ALPHA\*, BETA\*, GAMMA\* - ребра и углы в обратном пространстве (в  $1/\text{Å}$  и градусах).
  4. Объем элементарной ячейки в  $\text{Å}^3$ .
  5. Название GROUP пространственной группы, число NSYM эквивалентных позиций.
  6. Для каждой из эквивалентных позиций матрица вращения R и вектор трансляции  $\vec{T}$  для соответствующего преобразования симметрии.
  7. Число NCEN центросимметричных зон.
  8. Уравнение каждой из центросимметричных зон и значение фазового угла с точностью до  $\pi$  для этой зоны.
  9. Границы зоны разрешения в прямом пространстве.
  10. Шкальный коэффициент SCALE.
  11. Шкальный коэффициент SENAT.
  12. Параметры THR и ACL, регулирующие процесс уточнения.
  13. Шаг, с которым рассчитывается распределение вероятностей значений фаз.
  - 14-15. Флаги IOUT и KOUT вывода данных в файлы BINOUT и BINFOU, соответственно.
  16. Число NCYCLE циклов расчета фаз и (или) уточнения параметров.
  17. Флаги KCYCLE(I), задающие наличие пересчета фаз на каждом из циклов работы программы.
- Далее отражена информация по одной из тяжелоатомных производных. При наличии нескольких производных такого рода информация повторяется для каждой из них.
18. Текст TITLE, идентифицирующий программу.
  19. Число производных, для которых дана информация.
  20. Порядковый номер производной в программе.
  21. Текст CTITLE, идентифицирующий производную.
  22. Число тяжелых атомов в производной, флаги NPHASE и NREFIN.

23. Флаг КАУ и порядковый номер данных текущей производной на магнитной ленте.

24. Число "входных групп" и "рабочих групп" рефлексов, карта соответствия для них.

25-27. Шкальные коэффициенты SCESD, Q, BOVER.

28. Флаг уточнения для коэффициента BOVER.

29. Порядковый номер тяжелого атома внутри производной.

30. Коэффициент RK нормального замещения.

31. Вещественный формфактор RZ в восьми точках интервала по  $\frac{\sin \theta}{\lambda}$ .

32. Размерность температурного фактора.

33. Элементы матрицы температурного фактора.

34. Координаты тяжелого атома.

35. Флаги уточнения LABEL1÷LABEL5 для коэффициента замещения, температурного фактора и координат.

36. Оценка ошибки изоморфизма ESQ в восьми точках заданного интервала по  $\frac{\sin \theta}{\lambda}$ .

При наличии в производной нескольких независимых (не связанных симметрией группы) тяжелых атомов информация, содержащаяся в строках 29-35 повторяется для каждого из мест посадки тяжелых атомов.

В случае, когда флаг КАУ ненулевой, после сообщения 31 выводятся значения аномального коэффициента замещения и формфактора, а после сообщения 35 - флаги уточнения по аномальному рассеянию.

На каждом цикле работы программы на печать выводится информация о рефлексах, не взятых в уточнение. Исключение рефлекса из уточнения означает, что при вычислении производных по параметрам и при определении статистики этот рефлекс использоваться не будет. Для j-й тяжелоатомной производной исключаются те рефлексы, для которых (см. гл. I, §5):

$$\epsilon_j^2(\psi_{est}) \geq 16 \left[ E_{0j}^2 + (\sigma_{RNj}^{calc})^2 + (\sigma_{RNj}^{obs})^2 \right].$$

Пример такого рода информации приведен в протоколе 2.

Протокол 2.

1. Текст TITLE, идентифицирующий программу.
2. Порядковый номер цикла работы.

3, 4. Для каждой из тяжелоатомных производных, включенных в работу, ее порядковый номер и значение величины  $E_{0j}^2$  в восьми точках заданного интервала по  $\frac{\sin \theta}{\lambda}$ . Значения  $E_{0j}^2$  выводятся такими, какими они были вычислены на предыдущем цикле работы (см. гл. I, §4).

5. Информация о рефлексах, исключенных из уточнения:

N - порядковый номер производной, для которой исключен данный рефлекс;

H, K, L - миллеровские индексы рефлекса;

RES - разрешение в ангстремах, которому соответствует этот рефлекс;

FNAT - амплитуда  $F_p$  нативного белка;

SIGMA - стандартное отклонение  $\sigma_p$  для  $F_p$ ;

FDER - амплитуда  $F_{pH}^{obs}$  для N-й производной;

SIGMA - стандартное отклонение для  $F_{pH}^{obs}$ ;

FPHC - амплитуда  $F_{pH}^{calc}$ ;

EPS - ошибка замкнутости  $\mathcal{E}(\varphi_{best}) = F_{pH}^{obs} - F_{pH}^{calc}$ ;

$\vec{F}_H, PHH$  - амплитуда и фаза нормальной компоненты структурного фактора тяжелых атомов N-й производной;

FOM - показатель достоверности определения фазы нативного белка.

6. Число рефлексов, исключенных из уточнения.

В случае присутствия данных по аномальному рассеянию после 4 выводятся значения величины  $A_{0j}^2$  (см. гл. I, §4) в восьми точках интервала разрешения.

Для каждой из тяжелоатомных производных на печать выводится информация, отражающая ход уточнения параметров и статистику рефлексов. Пример такого рода информации приведен далее в протоколе 2.

7, 8. Порядковый номер производной в программе и текст, ее идентифицирующий.

9. Шкальный коэффициент SCESD для текущей производной.

10. Для шкального коэффициента Q величины уточненного значения, произведенного изменения значения и стандартного отклонения.

11. Значение параметра BOVER (в данном случае не уточнялось).

12-15. Для каждого из независимых тяжелых атомов производной порядковый номер атома в производной и информация об уточнении коэффициента замещения, температурного

фактора и координат атома в форме: уточненное значение параметра, произведенный сдвиг параметра, стандартное отклонение для параметра.

16. Корреляционная матрица порядка  $n$  ( $n$  – число уточняемых для данной производной параметров). Представлена только верхняя треугольная часть матрицы – нижняя заполнена нулями.

17. Определитель матрицы системы линейных уравнений для сдвигов параметров.

При наличии в текущей производной нескольких тяжелых атомов информация 12–15 выводится для каждого из них.

При использовании для текущей производной данных по аномальному рассеянию выводится также информация об уточнении аномального коэффициента замещения.

Если регулирующий уточнение параметр  $ACL \neq 1.0$ , то сдвиги параметров, для которых необходимо умножение на  $ACL$ , будут выводиться после умножения.

Следующая часть протокола 2 посвящена анализу изоморфного замещения для одной из производных. Такого рода информация выводится для каждой из участвующих в работе производных. При проведении указанного анализа участвующие в нем величины  $F_p, G_p, F_{pnj}^{obs}, G_{pnj}^{obs}, \Delta_{pnj}^{obs}$  шкалируются, как это описано в §2 гл. I (шкала  $j$ -й производной).

Анализ проводится для заданной зоны разрешения  $\frac{1}{D_{MAX}} \leq d^* \leq \frac{1}{D_{MIN}}$  в обратном пространстве в целом и для каждой из восьми равных частей зоны разрешения в обратном пространстве. Подчеркнем, что на восемь равных частей разбивается интервал  $(1/D_{MAX}, 1/D_{MIN})$  в обратном пространстве, то есть длины соответствующих подинтервалов интервала  $(D_{MIN}, D_{MAX})$  в прямом пространстве различны.

В строках 18–30 протокола 2 (см. приложение 2) приводится таблица, смысл каждой строки которой будет разъяснен ниже. Первые восемь столбцов таблицы соответствуют восьми указанным подинтервалам зоны разрешения, девятый содержит статистику по зоне разрешения в целом.

Протокол 2.

18. Разрешение в ангстремах - минимальное значение межплоскостного расстояния  $d = \frac{\lambda}{2 \sin \theta}$  в прямом пространстве для рефлексов из соответствующего подинтервала.

19. Число рефлексов, попавших в соответствующую часть зоны разрешения - величина  $M$ .

20. Среднее значение амплитуды производной - величина  $\frac{\sum F_{\text{PH}_j}^{\text{obs}}}{M}$ .

21. Взвешенное среднее значение амплитуды для производной - величина:

$$\frac{\sum F_{\text{PH}_j}^{\text{obs}} / (\sigma_{\text{PH}_j}^{\text{obs}})^2}{\sum 1 / (\sigma_{\text{PH}_j}^{\text{obs}})^2}.$$

22. Среднее значение амплитуды для тяжелых атомов - величина:

$$\frac{1}{M} \sum |\vec{f}_{\text{H}_j}|.$$

23. Среднее значение абсолютной величины ошибки замкнутости - величина:

$$\frac{1}{M} \sum |\varepsilon_j(\varphi_{\text{best}})|.$$

24. Взвешенная средняя ошибка замкнутости - величина:

$$\frac{\sum \varepsilon_j(\varphi_{\text{best}}) / ((\sigma_{\text{PH}_j}^{\text{obs}})^2 + (\sigma_{\text{PH}_j}^{\text{calc}})^2)}{\sum 1 / ((\sigma_{\text{PH}_j}^{\text{obs}})^2 + (\sigma_{\text{PH}_j}^{\text{calc}})^2)}.$$

25. Взвешенное среднеквадратичное значение амплитуды производной - величина:

$$\sqrt{\frac{\sum (F_{\text{PH}_j}^{\text{obs}})^2 / (\sigma_{\text{PH}_j}^{\text{obs}})^2}{\sum 1 / (\sigma_{\text{PH}_j}^{\text{obs}})^2}}.$$

26. Среднеквадратичное значение амплитуды тяжелых атомов - величина:

$$\sqrt{\frac{1}{M} \sum |\vec{f}_{\text{H}_j}|^2}.$$

27. Среднеквадратичная ошибка замкнутости - величина:

$$\sqrt{\frac{1}{M} \sum \varepsilon_j(\varphi_{best})}$$

28. Взвешенная среднеквадратичная ошибка замкнутости - величина:

$$\frac{\sum \varepsilon_j(\varphi_{best}) / ((G_{\rho H_j}^{obs})^2 + (G_{\rho H_j}^{calc})^2)}{\sum 1 / ((G_{\rho H_j}^{obs})^2 + (G_{\rho H_j}^{calc})^2)}$$

29. R-фактор - величина:

$$100.0 * \frac{\sum \varepsilon_j(\varphi_{best})}{\sum F_{\rho H_j}^{obs}}$$

30. Среднеквадратичный R-фактор - величина:

$$100.0 * \frac{\sum \varepsilon_j(\varphi_{best}) / ((G_{\rho H_j}^{obs})^2 + (G_{\rho H_j}^{calc})^2)}{\sum (F_{\rho H_j}^{obs})^2 / (G_{\rho H_j}^{obs})^2} * \frac{\sum 1 / (G_{\rho H_j}^{obs})^2}{\sum 1 / ((G_{\rho H_j}^{obs})^2 + (G_{\rho H_j}^{calc})^2)}$$

31. Параметр уточнения S.

Здесь M - число рефлексов, попавших в соответствующую часть интервала разрешения;

$F_{\rho H_j}^{obs}$ ,  $G_{\rho H_j}^{obs}$  - экспериментальная амплитуда текущей производной и ее стандартное отклонение;

$\vec{F}_{H_j}$  - структурный фактор тяжелых атомов текущей производной;

$\varepsilon_j(\varphi_{best})$  - ошибка замкнутости фазового треугольника для нормальной компоненты текущей производной в пред-

положении, что фаза нативного белка есть  $\varphi_{best}$ ;  
 $G_{\rho H_j}^{calc}$  - стандартное отклонение для вычисленного в предположении, что фаза белка есть  $\varphi_{best}$  значения амплитуды производной;

$\sum$  - суммирование ведется по взятым в уточнение рефлексам, попавшим в соответствующий интервал зоны разрешения.

В случае, когда при расчете взвешенных величин в каком-либо слагаемом  $\sigma_{\text{PHj}}^{\text{obs}} = 0$  или  $\sigma_{\text{PHj}}^{\text{calc}} = 0$ , то полагается

$$1/(\sigma_{\text{PHj}}^{\text{obs}})^2 = 1, \quad 1/((\sigma_{\text{PHj}}^{\text{obs}})^2 + (\sigma_{\text{PHj}}^{\text{calc}})^2) = 1.$$

Если для производной используются данные по аномальному рассеянию, то выводятся также величины:

NUMBER OF MEASUREMENTS - число N таких рефлексов, попавших в соответствующую часть зоны разрешения, что для них  $\delta_j(\varphi_{\text{best}}) < 16 \text{ \AA}^2$ ;

WEIGHTED RMS DIFF - взвешенное среднеквадратичное значение аномальной разности

$$\sqrt{\frac{\sum (\Delta_j^{\text{obs}})^2 / (\sigma_{\text{PHj}}^{\text{obs}})^2}{\sum 1/(\sigma_{\text{PHj}}^{\text{obs}})^2}};$$

CALCULATED RMS DIFF - среднеквадратичное значение, вычисленное в предположении, что фаза белка есть  $\varphi_{\text{best}}$ , аномальной разности:

$$\sqrt{\frac{1}{N} \sum (\Delta_j^{\text{calc}})^2};$$

WEIGHTED RMS RESIDUAL - взвешенная среднеквадратичная ошибка

$$\sqrt{\frac{\sum (\delta_j(\varphi_{\text{best}}))^2 / (\sigma_{\text{PHj}}^{\text{obs}})^2}{\sum 1/(\sigma_{\text{PHj}}^{\text{obs}})^2}};$$

GENERALIZED R FACTOR - R-фактор по аномальному рассеянию

$$100.0 * \sqrt{\frac{\sum (\delta_j(\varphi_{\text{best}}))^2 / (\sigma_{\text{PHj}}^{\text{obs}})^2}{\sum (\Delta_j^{\text{obs}})^2 / (\sigma_{\text{PHj}}^{\text{obs}})^2}};$$

ANOMALOUS REFINEMENT PARAMETER - параметр уточнения по аномальному рассеянию. Здесь в дополнение к предыдущему обозначено:

$N$  - число рефлексов, попавших по данным аномального рассеяния в соответствующую часть интервала разрешения и взятых в уточнение;

$\delta_j(\varphi_{best})$  - ошибка замкнутости фазового треугольника для аномального рассеяния в предположении, что фаза нативного белка есть  $\varphi_{best}$ ;

$A_j$  - величина, характеризующая разброс распределения вероятностей;

$\Delta_j^{obs}$ ,  $\Delta_j^{calc}$  - экспериментальная и вычисленная (в предположении, что фаза белка есть  $\varphi_{best}$ ) аномальные разности.

Для рефлексов, фазы которых определились с ненулевыми показателями достоверности, проводится анализ по показателю достоверности. Пример такого анализа дан в начале протокола 3, приведенного в приложении 3.

### Протокол 3.

1. Текст TITLE, идентифицирующий программу.

2, 3. Статистика рефлексов по показателю достоверности.

В  $k$ -м столбце строки 3 приводится число рефлексов, для которых  $\frac{k-1}{10} \leq m < \frac{k}{10}$ ;

4-7. Статистика показателей достоверности для различных частей зоны разрешения.

5. Минимальное межплоскостное расстояние  $d$  в прямом пространстве для соответствующей части зоны разрешения.

6. Число рефлексов, попавших в эту часть зоны разрешения.

7. Среднее значение показателя достоверности.

Во второй части протокола 3 приведены таблицы, содержащие анализ тяжелоатомных производных. Такие таблицы выводятся на печать для каждой из производных, участвующих в работе.

В отличие от описанной выше таблицы из протокола 2 величины  $F_p$ ,  $G_p$ ,  $F_{pni}^{obs}$ ,  $G_{pni}^{obs}$ ,  $\Delta_j^{obs}$  приводятся к шкале нативного белка (см. гл. II, § 2).

### Протокол 3 (продолжение)

8, 9. Порядковый номер производной в программе и текст TITLE, идентифицирующий ее.

10. Разрешение в ангстремах - минимальное значение меж-плоскостного расстояния  $d$  в прямом пространстве для реф-лексов соответствующей зоны разрешения.

11. Число рефлексов, попавших в соответствующую часть зоны разрешения, - величина  $M$ .

12. Среднеквадратичное значение изоморфной разности - величина

$$\sqrt{\frac{\sum (F_{\rho h_j}^{obs} - F_{\rho})^2 / ((G_{\rho h_j}^{obs})^2 + (G_{\rho})^2)}{\sum 1 / ((G_{\rho h_j}^{obs})^2 + (G_{\rho})^2)}} .$$

13-23. Те же величины, что и в строках 20-30 протокола 2, но вычисленные в шкале нативного белка.

24. Взвешенный средний квадрат ошибки изоморфизма - величина:

$$\sqrt{\frac{\sum (\varepsilon_j(\varphi_{best}))^2 / ((G_{\rho h_j}^{obs})^2 + (G_{\rho h_j}^{calc})^2)}{\sum 1 / ((G_{\rho h_j}^{obs})^2 + (G_{\rho h_j}^{calc})^2)}} .$$

25. R-фактор Cullis - величина:

$$100.0 * \frac{\sum |\varepsilon_j(\varphi_{best})|}{\sum |F_{\rho h_j}^{obs} - F_{\rho}|} ;$$

26. R-фактор Kraut по центросимметричным рефлексам - величина

$$100.0 * \frac{\sum_{\frac{h}{c}} |\varepsilon_j(\varphi_{best})|}{\sum_{\frac{h}{c}} F_{\rho h_j}^{obs}} ,$$

где  $\sum_{\frac{h}{c}}$  берется по рефлексам, принадлежащим центросимметричным зонам.

27. R-фактор Cullis по центросимметричным рефлексам

$$100.0 * \frac{\sum_{\frac{h}{c}} |\varepsilon_j(\varphi_{best})|}{\sum_{\frac{h}{c}} |F_{\rho h_j}^{obs} - F_{\rho}|} .$$

28. Среднеквадратичное значение ошибки изоморфизма - величина  $E_0$ .

29. Сообщение об окончании работы.

Здесь приняты те же обозначения, что и при описании протокола 2, за исключением того, что величины  $F_p$ ,  $G_p$ ,  $F_{pH_j}$ ,  $G_{pH_j}$ ,  $\Delta_j$  приводятся к шкале нативного белка.

Если для производной используются данные по аномальному рассеянию, то выводятся также величины:

NUMBER OF MEASUREMENTS	} те же, что и выше,
WEIGHTED RMS DIFF	
CALCULATED RMS DIFF	
WEIGHTED RMS RESIDUAL	

а также RMS R FACTOR - среднеквадратичный R-фактор:

$$100.0 * \sqrt{\frac{\sum (\delta_j(\varphi_{best}))^2 / (G_{pH_j}^{obs})^2}{\sum 1 / (G_{pH_j}^{obs})^2}}$$

WEIGHTED MEAN SQUARE RESIDUAL - взвешенный средний квадрат ошибки изоморфизма по аномальному рассеянию:

$$\frac{\sum (\delta_j(\varphi_{best}))^2 / (G_{pH_j}^{obs})^2}{\sum 1 / (G_{pH_j}^{obs})^2}$$

RMS LACK OF ISOMORPHISM - среднеквадратичная ошибка изоморфизма.

Сообщения программы в случае ненормального завершения работы

При нарушении некоторых из условий на размеры массивов (см. гл. II, §2) программа прекращает работу и выдает на печать соответствующее сообщение. Ниже приводится перечень таких сообщений.

q COMPOUNDS, LIMIT IS 6 - указанное для работы число q производных превосходит 6;

MORE THAN 30 ATOMS - общее число тяжелых атомов во всех производных превосходит 30;

q VARIABLES, LIMIT IS 150

n LOCATIONS REQUIRED FOR MATRICES,  
LIMIT IS 3000 - общее число уточняемых параметров  
q > 150 либо размер, требуемый для уточнения матрицы n >  
3000.

Если при решении системы линейных уравнений для сдви-  
гов параметров определитель матрицы обратился в нуль, то  
для рассматриваемой производной уточнение не производится,  
работа продолжается, а на печать выводится сообщение:

A SINGULAR MATRIX DETECTED IN MATSOL,  
RETURNING A AS ANSWERS

#### 87. Использование рабочего файла

Если на каком-либо цикле работы программы пересчет  
фаз не производится (то есть проводится только уточнение  
параметров), то на предыдущем цикле фазовая информация  
выводится в рабочий файл. Информация выводится в файл по-  
следовательного доступа с идентификатором BINSCR, ко-  
торому в начале работы программы присваивается значение 7.  
При необходимости изменить номер этого файла нужно заме-  
нить карту: BINSCR=7 в начале главной программы. Рабо-  
чий файл должен допускать применение оператора REWIND.

ПРИЛОЖЕНИЕ 1

Список подпрограмм

Имя	Число операторов
Главная программа	999
PHASE	255
FINAL	502
EXTRA	250
MATSOL	184
HIDE	34
CLOCK	3
INPUT	14
OUTPUT	14

## ПРИЛОЖЕНИЕ 2

### Генерация тестовых данных

```
DIMENSION REC(41)
EQUIVALENCE (JXL1,REC(34)),(JXL2,REC(35))
DATA X1,Y1,Z1,X2,Y2,Z2,
* XA1,YA1,ZA1,XA2,YA2,ZA2,TP
* /0.1,0.1,0.1,0.3,0.3,0.3,
* 0.1,0.2,0.3,0.4,0.3,0.2,6.283185/
DO 10 J=1,41
10 REC(J)=0.0
JXL1=1
JXL2=1
DO 100 JH=1,11
DO 100 JK=1,11
DO 100 JL=1,6
AH=JH-6
AK=JK-6
AL=JL-1
REC(1)=AH
REC(2)=AK
REC(3)=AL
R1=TP*(AH*X1+AK*Y1+AL*Z1)
RS1=TP*(-AH*X1+AK*(Y1+0.5)-AL*Z1)
R2=TP*(AH*X2+AK*Y2+AL*Z2)
RS2=TP*(-AH*X2+AK*(Y2+0.5)-AL*Z2)
C1=COS(R1)+COS(RS1)+COS(R2)+COS(RS2)
S1=SIN(R1)+SIN(RS1)+SIN(R2)+SIN(RS2)
REC(7)=SQRT(C1*C1+S1*S1)
R1=TP*(AH*X1+AK*Y1+AL*ZA1)
RS1=TP*(-AH*X1+AK*(Y1+0.5)-AL*ZA1)
C2=C1+COS(R1)+COS(RS1)
S2=S1+SIN(R1)+SIN(RS1)
REC(10)=SQRT(C2*C2+S2*S2)
R1=TP*(AH*X2+AK*Y2+AL*ZA2)
RS1=TP*(-AH*X2+AK*(Y2+0.5)-AL*ZA2)
C2=C1+COS(R1)+COS(RS1)
S2=S1+SIN(R1)+SIN(RS1)
REC(11)=SQRT(C2*C2+S2*S2)
100 WRITE(9) REC
CONTINUE
REC(7)=-2.0
WRITE(9) REC
STOP
END
```

Рис. 2

На рис. 2 приведен текст программы, генерирующий тестовые данные. На рис. 3 приведен пример задания управляющих данных на перфокартах.

В приложении 3 приведены протоколы работы фазовой программы при использовании этих управляющих данных и со входными данными, сгенерированными приведенной на рис. 2 программой. Счет проводился на ЭВМ ЕС 1040 в рамках операционной системы ДОС ЕС.

TEST FOR PHASING PROGRAM										
	25.0	25.0	50.0	90.0	90.0	90.0	90.0	90.0	90.0	
P21										
1.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	1.0
-1.0	0.0	0.0	0.0	0.0	1.0	0.0	0.5	0.0	0.0	-1.0
1										0.0
0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
5.0	500.0	1.0	1.0	1.0	1.0	1000.0	1.0	1.0	1.0	
3	1	1	1							
2										
TEST DER. 1										
1	1	1	0	1						
1	1	1								
1.0	1.0	1.0	0.0	0						
0.9										
1.0	1.0	0.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
0.12	0.2	0.2	0.3							
1	0	1	0	0						
1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
TEST DER. 2										
1	1	1	0	2						
1	1	1								
1.0	1.0	1.0	0.0	0						
1.0										
1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
1	0.0	0.0								
0.4	0.29	0.201								
0	0	0	1	1						
1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

Рис. 3

### ПРИЛОЖЕНИЕ 3. Пример работы программы

```

1) HEAVY ATOM PHASING AND REFINEMENT PROGRAM
2) TEST FOR PHASING PROGRAM
3) CRYSTALLOGRAPHIC PARAMETERS
    A = 25.00   B = 25.00   C = 50.00   ALPHA = 90.00   BETA = 90.00   GAMMA = 90.00
    A* = 0.040000   B* = 0.040000   C* = 0.020000   ALPHA* = 90.00   BETA* = 90.00   GAMMA* = 90.00
4) VOLUME OF UNIT CELL = 3.125E 04 CUBIC ANGSTROMS
5) SPACE GROUP = P21   WITH 2 EQUIVALENT POSITIONS
6) POSITION 1
    1.0 0.0 0.0   0.0
    0.0 1.0 0.0   0.0
    0.0 0.0 1.0   0.0
POSITION 2
    -1.0 0.0 0.0   0.0
    0.0 1.0 0.0   0.5000000
    0.0 0.0 -1.0   0.0
7) THERE ARE 1 CENTROSYMMETRIC PROJECTIONS
8) IF H*( 0.0) + K*( 1.0) + L*( 0.0) = 0 THEN PHI = H*( 0.0) + K*( 0.0) + L*( 0.0)
9) RESOLUTION LIMITS ARE FROM 500.00 TO 5.00 ANGSTROMS
10) DATA WILL BE SCALED BY 1.000E 00
11) NATIVE COUNTING STATISTICS WILL BE SCALED BY 1.000E 00
12) PARAMETER SHIFTS GREATER THAN 1.000E 03 SIGMA WILL BE SCALED BY 1.000E 00
13) PHASE PROBABILITIES WILL BE CALCULATED EVERY 5 DEGREES
14) FLAG FOR NEW OUTPUT TAPE = 0
15) FLAG FOR FOURIER OUTPUT = 0
16) NUMBER OF PHASING AND REFINEMENT CYCLES = 3
17) PHASING FLAGS FOR EACH CYCLE ARE 1 1 1
18) TEST FOR PHASING PROGRAM
19) DERIVATIVE INFORMATION FOR 2 COMPOUNDS
20) COMPOUND 1   21) TEST DER.1
    22) SITES = 1   PHASING FLAG = 1   REFINEMENT FLAG = 1
    23) ANOMALOUS DISPERSION FLAG = 0   INPUT COLUMN NUMBER = 1
    24) 1 GROUPS OF INPUT DATA TO BE PLACED IN 1 GROUPS ON OUTPUT
    MAPPING = 1
    25) COUNTING STATISTICS WILL BE SCALED BY 1.000E 00
    26) DIVIDING SCALE FACTOR = 1.00000E 00
    27) OVERALL TEMPERATURE FACTOR = 0.0
    28) REFINEMENT FLAG FOR OVERALL TEMPERATURE FACTOR = 0
    29) SITE NUMBER 1
    30) REAL OCCUPANCY   0.50000
    31) REAL FORM FACTOR   1.00000   1.00000   1.00000   1.00000   1.00000   1.00000   1.00000   1.00000
    32) NUMBER OF B VALUES USED = 1
    33) B VALUES ARE   0.0   0.0   0.0   0.0   0.0   0.0   0.0
    34) COORDINATES   0.120000   0.200000   0.300000
    35) REFINEMENT FLAGS FOR OCCUPANCY, B AND COORDINATES ARE 1 0 1 0 0
    36) ESTIMATED ROOT MEAN SQUARE LACK OF CLOSURE
    1.000E 00   1.000E 00
COMPOUND 2   TEST DER.2
    SITES = 1   PHASING FLAG = 1   REFINEMENT FLAG = 1
    ANOMALOUS DISPERSION FLAG = 0   INPUT COLUMN NUMBER = 2
    1 GROUPS OF INPUT DATA TO BE PLACED IN 1 GROUPS ON OUTPUT
    MAPPING = 1
    COUNTING STATISTICS WILL BE SCALED BY 1.000E 00
    DIVIDING SCALE FACTOR = 1.00000E 00
    OVERALL TEMPERATURE FACTOR = 0.0
    REFINEMENT FLAG FOR OVERALL TEMPERATURE FACTOR = 0
    SITE NUMBER 1
    REAL OCCUPANCY   1.00000
    REAL FORM FACTOR   1.00000   1.00000   1.00000   1.00000   1.00000   1.00000   1.00000   1.00000
    NUMBER OF B VALUES USED = 1
    B VALUES ARE   0.0   0.0   0.0   0.0   0.0   0.0
    COORDINATES   0.400000   0.290000   0.201000
    REFINEMENT FLAGS FOR OCCUPANCY, B AND COORDINATES ARE 0 0 0 1 1
    ESTIMATED ROOT MEAN SQUARE LACK OF CLOSURE
    1.000E 00   1.000E 00

```

4) TEST FOR PHASING PROGRAM

2) CYCLE 3

3) COMPOUND 1  
 4) MEAN SQUARE ISOMORPHOUS ERROR  
 3.578E-03 4.947E-02 3.512E-02 2.970E-02 4.689E-02 4.831E-02 4.997E-02 0.110E-02

COMPOUND 2  
 MEAN SQUARE ISOMORPHOUS ERROR  
 0.0 3.416E-04 0.160E-04 2.075E-03 2.454E-03 4.436E-03 4.384E-03 2.511E-03

5) REJECTED REFLECTIONS

N	H	K	L	RES	FNIT	SIGMA	FDER	SIGMA	FPHC	EPS	FN	PHI	H	FOM
2	-2.	-4.	4.	5.1	1.54	0.0	0.04	0.0	0.40	0.35	2.00	302	0.93	
2	-2.	4.	4.	5.1	1.54	0.0	0.04	0.0	0.40	0.35	2.00	50	0.93	
2	-1.	0.	1.	22.4	4.30	0.0	4.70	0.0	4.62	0.00	0.62	0	1.00	
2	0.	0.	1.	50.0	1.30	0.0	1.65	0.0	1.61	0.03	0.61	0	1.00	

6) 4 REFLECTION REJECTED FOR REFINEMENT

7) COMPOUND 1 8) TEST DER.1

9) STANDARD DEVIATIONS SCALED BY 1.000E 00

10) SCALE FACTOR = 1.002E 00 SHIFT = 7.351E-03 S. D. OF SHIFT = 3.798E-03

11) OVERALL B (NOT REFINED) = 0.0

12) SITE 1

13) REAL OCCUPANCY 1.00176  
 OCCUPANCY SHIFT -0.00035  
 S. D. OF SHIFT 0.00035

14) ISOTROPIC TEMPERATURE FACTOR 0.0

15) ATOMIC COORDINATES 0.10014 0.20000 0.30000  
 COORDINATE SHIFT 0.00052  
 S. D. OF COORDINATE SHIFT 0.00040

16) CORRELATION MATRIX FOR REAL PARAMETERS OF ORDER 3

1.00 -0.73 0.12  
 0.0 1.00 -0.06  
 0.0 0.0 1.00

17) DETERMINANT OF NORMAL MATRIX WAS 6.4700 TIMES 10 TO THE 10

ANALYSIS OF ISOMORPHOUS DIFFERENCES

	37.30	19.42	13.11	9.90	7.95	6.64	5.71	5.00	TOTAL
18) RESOLUTION IN ANGSTROMS	1	9	21	54	40	96	04	100	413
19) NUMBER OF MEASUREMENTS	0.30	2.66	2.20	2.34	1.05	2.12	2.11	2.51	2.22
20) MEAN DERIVATIVE F	0.30	2.66	2.20	2.34	1.05	2.12	2.11	2.51	2.22
21) WEIGHTED M DERIV. F	0.52	1.22	1.25	1.34	1.13	1.33	1.27	1.20	1.27
22) MEAN HEAVY ATOM F	0.01	0.03	0.06	0.07	0.09	0.07	0.10	0.13	0.09
23) MEAN ABS. RESIDUAL	0.01	0.02	0.00	0.00	-0.02	0.01	0.02	0.02	0.01
24) WEIGHTED MEAN RESID.	0.30	3.14	2.41	2.60	2.14	2.36	2.34	2.74	2.48
25) WEIGHTED RMS DERIV. F	0.62	1.42	1.30	1.48	1.31	1.47	1.41	1.42	1.42
26) RMS HEAVY ATOM F	0.01	0.05	0.07	0.10	0.13	0.09	0.14	0.17	0.13
27) RMS RESIDUAL	0.01	0.05	0.07	0.10	0.13	0.09	0.14	0.17	0.13
28) WEIGHTED RMS RESIDUAL	2.20	1.01	2.55	2.06	4.04	3.20	4.65	5.16	4.05
29) R FACTOR	2.20	1.96	3.10	3.01	6.16	4.02	5.04	6.39	5.22

34) REFINEMENT PARAMETER = 1.6064920-02

COMPOUND 2 TEST DER.2

STANDARD DEVIATIONS SCALED BY 1.000E 00

SCALE FACTOR = 9.900E-01 SHIFT = 8.279E-03 S. D. OF SHIFT = 7.250E-04

OVERALL B (NOT REFINED) = 0.0

Протокол 2

4) TEST FOR PHASING PROGRAM

COMPARISON AND SUMMARY OF DERIVATIVES ON NATIVE SCALE

2) FIGURE OF MERIT LESS THAN 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1.0  
 3) NUMBER OF REFLECTIONS 56 8 2 0 0 0 0 0 0 0 347  
 4) FIGURE OF MERIT WITH RESOLUTION

5) RESOLUTION IN ANGSTROMS	37.38	19.42	13.11	9.90	7.95	6.64	5.71	5.00	TOTAL
6) NUMBER OF OBSERVATIONS	1	9	21	54	48	96	84	188	413
7) MEAN FIGURE OF MERIT	1.00	0.77	0.83	0.86	0.83	0.79	0.83	0.89	0.84
8) COMPOUND 1 9) TEST DER.1									
10) RESOLUTION IN ANGSTROMS	37.38	19.42	13.11	9.90	7.95	6.64	5.71	5.00	TOTAL
11) NUMBER OF REFLECTIONS	1	9	21	54	48	96	84	188	413
12) RMS ISOMORPHOUS DIFFERENCE	0.62	1.43	0.91	1.10	0.85	0.98	0.92	0.94	0.97
13) MEAN DERIVATIVE F	0.30	2.67	2.21	2.35	1.86	2.13	2.12	2.53	2.23
14) WEIGHTED MEAN DERIVATIVE F	0.30	2.67	2.21	2.35	1.86	2.13	2.12	2.53	2.23
15) MEAN HEAVY ATOM F	0.62	1.22	1.26	1.35	1.13	1.34	1.27	1.28	1.28
16) MEAN ABSOLUTE RESIDUAL	0.01	0.03	0.06	0.07	0.09	0.07	0.10	0.13	0.09
17) WEIGHTED MEAN RESIDUAL	0.01	0.02	0.00	0.00	-0.02	0.01	0.02	0.02	0.01
18) WEIGHTED RMS DERIVATIVE F	0.30	3.14	2.41	2.58	2.14	2.36	2.34	2.74	2.48
19) RMS HEAVY ATOM F	0.62	1.43	1.39	1.49	1.31	1.48	1.42	1.43	1.43
20) RMS RESIDUAL	0.01	0.05	0.07	0.10	0.13	0.09	0.14	0.17	0.13
21) WEIGHTED RMS RESIDUAL	0.01	0.05	0.07	0.10	0.13	0.09	0.14	0.17	0.13
22) R FACTOR	2.20	1.01	2.35	2.06	4.84	3.28	4.65	5.16	4.85
23) RMS R FACTOR	2.20	1.56	3.10	3.01	6.16	4.82	5.94	6.30	5.22
24) WEIGHTED MEAN SQUARE RESIDUAL	0.00	0.00	0.01	0.01	0.02	0.01	0.02	0.03	0.02
25) CULLIS R FACTOR	1.37	2.20	0.79	7.44	14.08	9.16	14.37	10.45	12.29
26) KRAUT CENTRIC R FACTOR	2.20	0.24	1.09	0.72	0.50	1.48	0.50	0.23	0.59
27) CULLIS CENTRIC R FACTOR	1.37	0.51	2.05	1.38	0.85	2.82	1.16	0.58	1.14
28) RMS LACK OF ISOMORPHISM	0.01	0.01	0.01	0.02	0.01	0.02	0.01	0.01	0.01
COMPOUND 2 TEST DER.2									
RESOLUTION IN ANGSTROMS	37.38	19.42	13.11	9.90	7.95	6.64	5.71	5.00	TOTAL
NUMBER OF REFLECTIONS	0	8	21	54	48	96	84	98	409
RMS ISOMORPHOUS DIFFERENCE	0.0	1.27	1.84	1.21	1.13	1.17	1.04	1.18	1.13
MEAN DERIVATIVE F	0.0	1.92	2.43	2.27	2.00	2.15	2.16	2.14	2.16
WEIGHTED MEAN DERIVATIVE F	0.0	1.92	2.43	2.27	2.00	2.15	2.16	2.14	2.16
MEAN HEAVY ATOM F	0.0	1.29	1.25	1.33	1.12	1.33	1.26	1.26	1.27
MEAN ABSOLUTE RESIDUAL	0.0	0.02	0.02	0.03	0.03	0.02	0.03	0.02	0.03
WEIGHTED MEAN RESIDUAL	0.0	0.01	0.01	0.02	0.03	0.02	0.02	0.02	0.02
WEIGHTED RMS DERIVATIVE F	0.0	2.59	2.60	2.59	2.31	2.47	2.42	2.45	2.46
RMS HEAVY ATOM F	0.0	1.40	1.37	1.47	1.30	1.47	1.40	1.40	1.42
RMS RESIDUAL	0.0	0.03	0.03	0.03	0.04	0.03	0.04	0.04	0.04
WEIGHTED RMS RESIDUAL	0.0	0.03	0.03	0.03	0.04	0.03	0.04	0.04	0.04
R FACTOR	0.0	1.13	0.90	1.16	1.50	1.12	1.32	1.16	1.21
RMS R FACTOR	0.0	1.00	0.96	1.30	1.77	1.25	1.63	1.44	1.43
WEIGHTED MEAN SQUARE RESIDUAL	0.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CULLIS R FACTOR	0.0	2.23	2.56	2.53	3.44	2.41	3.35	2.57	2.70
KRAUT CENTRIC R FACTOR	0.0	1.00	1.79	1.63	1.00	2.07	2.14	1.09	1.89
CULLIS CENTRIC R FACTOR	0.0	2.91	3.25	2.95	3.76	4.29	3.95	3.40	3.54
RMS LACK OF ISOMORPHISM	0.0	0.03	0.03	0.04	0.04	0.05	0.04	0.05	0.04

29) END OF JOB

Протокол 3

#### ЛИТЕРАТУРА

1. Blow D.M., Crick E.H.C. - Acta Cryst., 1959, v. 12, p. 794-802.
2. Ten Eyck L.F., Arnone A. - J. Mol. Biol., 1976, v. 100, p. 3-11.
3. Martin R.S., Peters G., Wilkinson J.H. - Num. Math., 1965, v. 7, p. 362-383.

## СОДЕРЖАНИЕ

ВВЕДЕНИЕ . . . . .	3
Глава I. Принципы организации программы . . . . .	5
§1. Логическая структура программы . . . . .	5
§2. Шкалирование данных в программе . . . . .	7
§3. Расчет вклада тяжелых атомов в структурный фактор . . . . .	8
§4. Расчет фаз структурных факторов нативного белка . . . . .	10
§5. Уточнение параметров тяжелоатомных производных . . . . .	14
§6. Организация программы . . . . .	18
Глава II. Входные и выходные данные программы . . .	21
§1. Управляющие данные на перфокартах . . . . .	21
§2. Вывод данных на перфокарты . . . . .	28
§3. Входные данные на магнитной ленте . . . . .	28
§4. Выходные данные на магнитной ленте . . . . .	31
§5. Вывод данных для синтезов Фурье . . . . .	32
§6. Сообщения программы . . . . .	34
§7. Использование рабочего файла . . . . .	45
Список подпрограмм . . . . .	46
Приложение 1 . . . . .	46
Приложение 2. Генерация тестовых данных . . . . .	47
Приложение 3. Пример работы программы . . . . .	49
Литература . . . . .	52

Владимир Юрьевич Луни

ФАЗОВАЯ ПРОГРАММА ТЕН ЭЙКА  
Программное обеспечение структурных  
исследований - ФОРТРАН

Отредактировано и подготовлено к печати в Отделе  
научно-технической информации НЦБИ

Редактор Р.Г.Цветникова  
Технический редактор С.М.Ткачук  
Корректоры Т.К.Крейцер, Л.М.Орлова

Подписано в печать 26/ХІІ-79г. Т-21369.  
Уч.-изд.л. 2,58. Формат 60х90/16. Тираж 200 экз.  
Заказ 12Р. Бесплатно. Изд. №163

Отпечатано на ротапринтере в Отделе научно-техни-  
ческой информации Научного центра биологических  
исследований АН СССР в Пушкине