*Acta Cryst.* (1998). D**54**, 822–827

# Detection and Use of Pseudo-Translation in Determination of Protein Structures

YUH MIN CHOOK,[a] WILLIAM N. LIPSCOMB[b] AND HENGMING KE[c]*

[a]*Laboratory of Cellular Biology, Howard Hughes Medical Institute, The Rockefeller University, 1230 York Avenue, New York, NY 10021, USA,* [b]*Gibbs Laboratory, Harvard University, Cambridge, MA 02138, USA, and* [c]*Department of Biochemistry and Biophysics, School of Medicine, The University of North Carolina, Chapel Hill, NC 27599, USA. E-mail: hke@med.unc.edu*

(*Received 18 August 1997; accepted 22 December 1997*)

## Abstract

Two types of pseudo-translation symmetry, pseudo-twofold translational symmetry and pseudo-body-centered symmetry, have been found in protein crystals of chorismate mutase and cyclophilin C. Statistics on diffraction intensity from these two crystals showed that the presence of pseudo-translations in atomic space yielded a distribution of systematically strong and weak reflections at low resolution. The diffraction pattern resulting from pseudo-translational symmetry was apparently similar to that from true crystallographic symmetry at 4 Å resolution, but was distinct at high resolution. Pseudo-translation can be detected by comparing the average magnitudes of certain parity groups of reflections in three-dimensional *hkl* space. Based on the structures of chorismate mutase and cyclophilin C, the ratio of >1.2 for the average magnitudes of parity groups is sufficient to indicate the existence of pseudo-translation. Although pseudo-translation often makes structure determination more difficult, the additional information of pseudo-translation has been used successfully in the structure determination of chorismate mutase by multiple isomorphous replacement and of cyclophilin C by molecular replacement. Thus, examination of pseudo-translation is recommended at an early stage of structure determination.

## 1. Introduction

Multiple molecules in a crystallographic asymmetric unit may be related by rotational and translational symmetries. These non-crystallographic symmetries, or molecular symmetries, have been widely used for phase improvement in structure determination of proteins. Pseudo-translation is a special type of non-crystallographic symmetry. It has been observed and reported in many crystal structures of small organic and inorganic molecules, but its presence in protein structures has not been specially addressed. Pseudo-translation often makes structure determination difficult (Fan *et al.*, 1990). For example, an ordinary direct-methods programs may fail to solve a structure with pseudo-translation symmetry because the program may select only strong reflections which represent a portion of the information in the unit cell. We report here two examples of pseudo translation in the protein structures of chorismate mutase (Chook *et al.*, 1993) and cyclophilin C complexed with the immunosuppressive drug cyclosporin A (CyPC–CsA, Ke *et al.*, 1993). The detection of pseudo-translation and its usage in protein structure determination will also be discussed.

## 2. Methods and materials

### 2.1. *Definition and detection of pseudo-translation*

Pseudo-translation refers to translational symmetry in atomic space, which relates multiple molecules in a crystallographic asymmetric unit and is close but not exactly equal to crystallographic translation. It can be illustrated from the definition of structure factor,

$$\mathbf{F}(hkl) = \sum_{j=1\rightarrow n} f_j \exp[2\pi i(hx + ky + lz)].$$

When $n$ atoms in the unit cell are related to one another by translation of half a unit cell along the $X$ direction, the structure factor can then be expressed as,

$$\begin{aligned}\mathbf{F}(hkl) = &\sum_{j=1\rightarrow n/2} f_j \exp[2\pi i(hx + ky + lz)] \\ &+ \sum_{j=n/2\rightarrow n} f_j \exp\{2\pi i[h(x + \tfrac{1}{2}) + ky + lz]\} \\ &+ \sum_{j=1\rightarrow n/2} f_j \exp[2\pi i(hx + ky + lz)] \\ &\times [1 + \exp 2\pi i(h/2)].\end{aligned}$$

Hence, $\mathbf{F}(hkl) = 0$ if $h$ = odd, or the reflections with odd $h$ indices will be systematically absent. In fact, the unit-cell dimension along the $X$ direction can be reduced to half the original in this special case. However, when atoms in a unit cell are related by a translation close but not exactly equal to half a unit along the $X$ direction, the reflections with odd $h$ indices will be systematically weak while the reflections with even $h$ indices will be systematically strong. This type of translation is thus

Table 1. *Statistics on magnitudes of reflections in parity groups for CyPC–CsA*

| Resolution (Å) | $H + K + L$ = odd $\langle F \rangle$† | # ref | $H + K + L$ = even $\langle F \rangle$ | # ref | $\langle F_e/F_o \rangle$‡ | $H + K$ = odd $\langle F \rangle$ | # ref | $H + K$ = even $\langle F \rangle$ | # ref | $\langle F_e/F_o \rangle$ |
|---|---|---|---|---|---|---|---|---|---|---|
| ∞–6.0 | 174.4 | 427 | 521.2 | 434 | 2.988 | 361.7 | 425 | 337.0 | 436 | 0.932 |
| 6.0–4.0 | 212.0 | 549 | 639.6 | 589 | 3.017 | 430.3 | 562 | 436.3 | 576 | 1.014 |
| 4.0–3.5 | 185.7 | 301 | 513.7 | 315 | 2.766 | 367.5 | 307 | 339.3 | 309 | 0.923 |
| 3.5–3.0 | 155.5 | 466 | 406.7 | 492 | 2.615 | 294.6 | 479 | 274.3 | 479 | 0.931 |
| 3.0–2.5 | 107.7 | 741 | 264.4 | 785 | 2.455 | 192.8 | 761 | 183.9 | 765 | 0.954 |
| 2.5–2.2 | 84.2 | 698 | 192.3 | 754 | 2.283 | 140.9 | 719 | 139.8 | 733 | 0.992 |
| 2.2–2.0 | 74.1 | 611 | 155.2 | 687 | 2.094 | 115.4 | 652 | 118.7 | 646 | 1.029 |
| 2.0–1.9 | 60.0 | 370 | 119.0 | 432 | 1.983 | 92.5 | 398 | 91.1 | 404 | 0.985 |
| 1.9–1.8 | 55.2 | 404 | 89.9 | 474 | 1.630 | 73.0 | 434 | 74.9 | 444 | 1.027 |
| 1.8–1.6 | 47.9 | 706 | 62.9 | 895 | 1.315 | 56.4 | 807 | 56.1 | 794 | 0.995 |
| | $H + L$ = odd | | $H + L$ = even | | | $K + L$ = odd | | $K + L$ = even | | |
| | $\langle F \rangle$ | # ref | $\langle F \rangle$ | # ref | $\langle F_e/F_o \rangle$ | $\langle F \rangle$ | # ref | $\langle F \rangle$ | # ref | $\langle F_e/F_o \rangle$ |
| ∞–6.0 | 340.4 | 428 | 357.9 | 433 | 1.051 | 349.5 | 425 | 348.9 | 436 | 0.998 |
| 6.0–4.0 | 413.8 | 554 | 451.8 | 584 | 1.092 | 419.3 | 572 | 447.5 | 566 | 1.067 |
| 4.0–3.5 | 354.1 | 302 | 352.8 | 314 | 0.996 | 362.5 | 305 | 344.5 | 311 | 0.950 |
| 3.5–3.0 | 272.1 | 475 | 296.7 | 483 | 1.090 | 285.4 | 480 | 283.5 | 478 | 0.993 |
| 3.0–2.5 | 181.6 | 761 | 195.0 | 765 | 1.074 | 188.2 | 766 | 188.5 | 760 | 1.002 |
| 2.5–2.2 | 134.4 | 713 | 146.0 | 739 | 1.087 | 140.6 | 724 | 140.1 | 728 | 0.997 |
| 2.2–2.0 | 111.9 | 641 | 122.1 | 657 | 1.092 | 115.9 | 647 | 118.2 | 651 | 1.019 |
| 2.0–1.9 | 88.8 | 401 | 94.8 | 401 | 1.067 | 90.8 | 403 | 92.8 | 399 | 1.022 |
| 1.9–1.8 | 70.5 | 438 | 77.4 | 440 | 1.097 | 73.2 | 440 | 74.7 | 438 | 1.021 |
| 1.8–1.6 | 55.2 | 792 | 57.3 | 809 | 1.039 | 56.0 | 797 | 56.6 | 804 | 1.011 |
| | $H$ = odd | | $H$ = even | | | $K$ = odd | | $K$ = even | | |
| | $\langle F \rangle$ | # ref | $\langle F \rangle$ | # ref | $\langle F_e/F_o \rangle$ | $\langle F \rangle$ | # ref | $\langle F \rangle$ | # ref | $\langle F_e/F_o \rangle$ |
| ∞–6.0 | 355.8 | 416 | 343.0 | 445 | 0.964 | 370.7 | 377 | 332.4 | 484 | 0.897 |
| 6.0–4.0 | 430.2 | 537 | 436.1 | 601 | 1.014 | 429.9 | 497 | 435.9 | 641 | 1.014 |
| 4.0–3.5 | 360.6 | 308 | 346.2 | 308 | 0.960 | 364.3 | 267 | 345.1 | 349 | 0.947 |
| 3.5–3.0 | 286.6 | 484 | 282.4 | 474 | 0.985 | 278.8 | 417 | 288.9 | 541 | 1.036 |
| 3.0–2.5 | 188.4 | 749 | 188.3 | 777 | 1.000 | 188.6 | 668 | 188.1 | 858 | 0.998 |
| 2.5–2.2 | 144.3 | 704 | 136.6 | 748 | 0.946 | 138.3 | 645 | 142.0 | 807 | 1.027 |
| 2.2–2.0 | 115.2 | 648 | 118.9 | 650 | 1.032 | 115.8 | 566 | 118.1 | 732 | 1.020 |
| 2.0–1.9 | 91.8 | 399 | 91.8 | 403 | 1.000 | 91.7 | 349 | 91.9 | 453 | 1.002 |
| 1.9–1.8 | 72.9 | 440 | 75.0 | 438 | 1.030 | 73.0 | 380 | 74.6 | 498 | 1.022 |
| 1.8–1.6 | 56.0 | 779 | 56.5 | 822 | 1.009 | 55.7 | 702 | 56.8 | 899 | 1.020 |
| | $L$ = odd | | $L$ = even | | | | | | | |
| | $\langle F \rangle$ | # ref | $\langle F \rangle$ | # ref | $\langle F_e/F_o \rangle$ | | | | | |
| ∞–6.0 | 355.4 | 422 | 343.3 | 439 | 0.966 | | | | | |
| 6.0–4.0 | 431.2 | 561 | 435.4 | 577 | 1.010 | | | | | |
| 4.0–3.5 | 362.5 | 310 | 344.2 | 306 | 0.949 | | | | | |
| 3.5–3.0 | 289.3 | 475 | 279.8 | 483 | 0.967 | | | | | |
| 3.0–2.5 | 192.3 | 758 | 184.4 | 768 | 0.959 | | | | | |
| 2.5–2.2 | 141.1 | 721 | 139.5 | 731 | 0.989 | | | | | |
| 2.2–2.0 | 118.2 | 639 | 116.0 | 659 | 0.982 | | | | | |
| 2.0–1.9 | 91.7 | 412 | 91.8 | 390 | 1.001 | | | | | |
| 1.9–1.8 | 72.8 | 444 | 75.1 | 434 | 1.032 | | | | | |
| 1.8–1.6 | 56.5 | 797 | 56.1 | 804 | 0.993 | | | | | |

† $\langle F \rangle$ is the averaged magnitude of observed structure factors. ‡ $F_o$ or $F_e$ is the averaged magnitude of reflections with odd or even indices. Only reflections with $K < 6$ were used in the statistics.

referred to twofold pseudo-translation. Similarly, $n$-fold pseudo-translation can be defined as $n$ copies of coordinates in the crystallographic asymmetric unit being related to each other by translation of one $n$th of the cell dimensions. Thus, if a crystal has $n$-fold pseudo-translation, its diffraction will show a pattern of one strong reflection at every $n$ spots. The above definition, which describes one-dimensional pseudo-translations, can be easily extended to three-dimensional cases such as pseudo-face-centered and pseudo-body-centered cases.

Pseudo-translation symmetry can be detected by statistics of diffraction intensity. In the absence of pseudo-translation, diffraction intensity in three-dimensional $hkl$ space will be distributed in a pattern in which the average magnitude for all parity groups of reflections will be comparable with one another. For example, the ratio of the average magnitudes of $(h + k + l)$ = even and $(h + k + l)$ = odd reflections will be close to 1 if there is no pseudo-body-centered symmetry. If the ratio deviates substantially from 1, pseudo body-centered symmetry most likely exists. Therefore, a comparison of average magnitudes for different parity groups of reflections can be used to detect pseudo-translation symmetry.

Table 2. *Superposition of molecules in the CyPC–CsA structure*

| Molecules† in comparison | Rotation matrix | | | Fractional translation | | | R.m.s. difference for Cα atoms (Å) |
|---|---|---|---|---|---|---|---|
| *A*1 over *B*1 | −1.0000 | 0.0044 | 0.0002 | 0.496 | 0.529 | 0.512 | 0.17 |
| | 0.0044 | 0.9999 | −0.0153 | | | | |
| | −0.0003 | −0.0153 | −0.9999 | | | | |
| *A*2 over *B*2 | −1.0000 | −0.0053 | 0.0006 | 0.508 | 0.481 | 0.512 | 0.21 |
| | −0.0053 | 0.9999 | −0.0149 | | | | |
| | −0.0005 | −0.0149 | −0.9999 | | | | |

† *A*1, *A*2, *B*1 and *B*2 are the four molecules in the crystallographic asymmetric unit.

Table 3. *Statistics on magnitudes of reflections in parity groups for CMase*

| | $H + K + L$ = odd | | $H + K + L$ = even | | | $H + K$ = odd | | $H + K$ = even | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Resolution (Å) | ⟨*F*⟩† | # ref | ⟨*F*⟩ | # ref | $\langle F_e/F_o \rangle$ | ⟨*F*⟩ | # ref | ⟨*F*⟩ | # ref | $\langle F_e/F_o \rangle$ |
| ∞–6.0 | 13.7 | 1184 | 13.7 | 1720 | 0.999 | 13.4 | 1462 | 14.0 | 1442 | 1.047 |
| 6.0–4.0 | 15.6 | 3050 | 15.8 | 4379 | 1.012 | 15.6 | 3711 | 15.8 | 3718 | 1.018 |
| 4.0–3.5 | 13.5 | 2034 | 14.0 | 2964 | 1.031 | 13.6 | 2500 | 13.9 | 2498 | 1.022 |
| 3.5–3.0 | 10.3 | 3449 | 10.3 | 5090 | 1.000 | 10.2 | 4267 | 10.3 | 4272 | 1.009 |
| 3.0–2.5 | 7.5 | 6086 | 7.5 | 9263 | 1.008 | 7.4 | 7669 | 7.6 | 7680 | 1.019 |
| 2.5–2.2 | 6.4 | 5497 | 6.5 | 8464 | 1.013 | 6.4 | 6981 | 6.5 | 6980 | 1.008 |
| 2.2–2.0 | 5.9 | 348 | 5.9 | 415 | 1.009 | 5.9 | 378 | 5.9 | 385 | 1.016 |
| | $H + L$ = odd | | $H + L$ = even | | | $K + L$ = odd | | $K + L$ = even | | |
| | ⟨*F*⟩ | # ref | ⟨*F*⟩ | # ref | $\langle F_e/F_o \rangle$ | ⟨*F*⟩ | # ref | ⟨*F*⟩ | # ref | $\langle F_e/F_o \rangle$ |
| ∞–6.0 | 13.9 | 1015 | 13.6 | 1889 | 0.979 | 13.7 | 929 | 13.7 | 1975 | 1.001 |
| 6.0–4.0 | 15.4 | 2556 | 15.9 | 4873 | 1.033 | 15.4 | 2338 | 15.8 | 5091 | 1.025 |
| 4.0–3.5 | 13.4 | 1683 | 14.0 | 3315 | 1.048 | 13.5 | 1658 | 13.9 | 3340 | 1.028 |
| 3.5–3.0 | 9.9 | 2819 | 10.5 | 5720 | 1.055 | 10.0 | 2877 | 10.4 | 5662 | 1.040 |
| 3.0–2.5 | 7.3 | 5164 | 7.6 | 10185 | 1.037 | 7.4 | 5257 | 7.6 | 10092 | 1.016 |
| 2.5–2.2 | 6.4 | 4866 | 6.5 | 9095 | 1.011 | 6.4 | 4878 | 6.5 | 9083 | 1.010 |
| 2.2–2.0 | 5.9 | 292 | 5.9 | 471 | 1.008 | 5.9 | 316 | 5.9 | 447 | 1.013 |
| | $H$ = odd | | $H$ = even | | | $K$ = odd | | $K$ = even | | |
| | ⟨*F*⟩ | # ref | ⟨*F*⟩ | # ref | $\langle F_e/F_o \rangle$ | ⟨*F*⟩ | # ref | ⟨*F*⟩ | # ref | $\langle F_e/F_o \rangle$ |
| ∞–6.0 | 6.7 | 1454 | 20.7 | 1450 | 3.105 | 14.1 | 1380 | 13.3 | 1524 | 0.944 |
| 6.0–4.0 | 11.9 | 3713 | 19.5 | 3716 | 1.648 | 15.9 | 3578 | 15.5 | 3851 | 0.969 |
| 4.0–3.5 | 12.5 | 2514 | 15.0 | 2484 | 1.200 | 14.0 | 2456 | 13.6 | 2542 | 0.974 |
| 3.5–3.0 | 10.3 | 4275 | 10.2 | 4264 | 0.992 | 10.3 | 4180 | 10.2 | 4359 | 0.987 |
| 3.0–2.5 | 7.9 | 7689 | 7.1 | 7660 | 0.906 | 7.5 | 7572 | 7.5 | 7777 | 0.994 |
| 2.5–2.2 | 6.7 | 6968 | 6.2 | 6993 | 0.918 | 6.5 | 6911 | 6.4 | 7050 | 0.999 |
| 2.2–2.0 | 6.0 | 389 | 5.8 | 374 | 0.953 | 5.9 | 391 | 5.9 | 372 | 1.011 |
| | $L$ = odd | | $L$ = even | | | | | | | |
| | ⟨*F*⟩ | # ref | ⟨*F*⟩ | # ref | $\langle F_e/F_o \rangle$ | | | | | |
| ∞–6.0 | 13.8 | 710 | 13.7 | 2194 | 0.987 | | | | | |
| 6.0–4.0 | 15.2 | 1723 | 15.8 | 5706 | 1.038 | | | | | |
| 4.0–3.5 | 13.1 | 1225 | 14.0 | 3773 | 1.070 | | | | | |
| 3.5–3.0 | 9.8 | 2112 | 10.4 | 6427 | 1.063 | | | | | |
| 3.0–2.5 | 7.2 | 4058 | 7.6 | 11291 | 1.061 | | | | | |
| 2.5–2.2 | 6.2 | 4017 | 6.5 | 9944 | 1.047 | | | | | |
| 2.2–2.0 | 5.7 | 276 | 6.0 | 487 | 1.052 | | | | | |

† ⟨*F*⟩ is the averaged magnitude of observed structure factors. $F_o$ or $F_e$ is the averaged magnitude of reflections with odd or even indices.

### 2.2. *Crystallization*

Crystals of murine cyclophilin C complexed with the immunosuppressive drug cyclosporin A (CyPC–CsA) were grown by microdialysis of 13 mg ml$^{-1}$ CyPC–CsA against a crystallization buffer of 20 m*M* Tris base, 2 m*M* dithiothreitol, 2 m*M* EDTA, 0.5 m*M* $NaN_3$, 14% polyethylene glycol 3350 (PEG 3350), 6% ethanol, 2% isopropanol, 0.2 *M* ammonium sulfate, pH 6.0 at 277 K (Ke *et al.*, 1993). It has the real-space group $P2_1$ with cell dimensions of $a$ = 60.3, $b$ = 74.0, $c$ = 97.9 Å, $\beta$ = 90°. However, the apparent symmetry of the CyPC–CsA crystal could be described in a pseudo-space group of $I2_12_12$. The asymmetric unit in the space group $P2_1$ contains four molecules of CyPC–CsA, which are related to one another by a pseudo-body-centered symmetry and a pseudo-rotational twofold axis close to the crystallographic $c$ axis as shown in the final structure (Ke *et al.*, 1993). Diffraction data of CyPC–CsA were collected on the Rigaku imaging-plate system to 1.64 Å resolution.

Chorismate mutase from *Bacillus subtilis* was crystallized by microdialysis of 12 mg ml$^{-1}$ protein against a buffer of 5 m*M* Tris–HCl, 1 m*M* 2-mercaptoethanol, 0.5 m*M* $NaN_3$, 0.1 m*M* EDTA, 12% PEG 3350 at pH 5.3 (Chook *et al.*, 1993). The crystal is in space group $P2_1$

Table 4. *Superposition of trimers in the chorismate mutase structure*

| Molecules† in comparison | Rotation matrix | | | Fractional translation | | | R.m.s. difference for Cα atoms (Å) |
|---|---|---|---|---|---|---|---|
| *T*1 over *T*2 | 0.9997 | −0.0189 | 0.0161 | 0.497 | −0.031 | 0.006 | 0.23 |
| | 0.0189 | 0.9998 | 0.0015 | | | | |
| | −0.0161 | −0.0012 | 0.9999 | | | | |
| *T*3 over *T*4 | 0.9998 | −0.0147 | −0.0168 | 0.522 | 0.009 | −0.006 | 0.24 |
| | 0.0143 | 0.9997 | −0.0220 | | | | |
| | 0.0171 | 0.0218 | 0.9996 | | | | |

† *T*1, *T*2, *T*3 and *T*4 are the four trimers in the crystallographic asymmetric unit.

with cell dimensions of $a$ = 102.4, $b$ = 68.3, $c$ = 102.8 Å, and $\beta$ = 105.6°. The crystallographic asymmetric unit contains 12 monomeric molecules or four trimers of chorismate mutase. Diffraction data were collected using a Hamlin multiwire detector system to 1.9 Å resolution.

## 3. Results and discussion

### 3.1. *Pseudo-translation symmetry of CyPC–CsA and chorismate mutase*

In the precession photos of the CyPC–CsA crystals, reflections with indices $(H + K)$ = odd in the $[HK0]$ zone and reflections with $(K + L)$ = odd in the $[0KL]$ zone appeared to be systematically absent in regions with $K < 6$, but are present when $K > 6$ (data not shown). Statistics using reflections with $K < 6$ in three-dimensional $hkl$ space revealed that the average magnitude of structure factors for the parity group of $(H + K + L)$ = even ($F_{hkl=e}$) was about 1.3 to 3.0 times the average magnitude for the parity group of $(H + K + L)$ = odd ($F_{hkl=o}$) (Table 1). In contrast, the average magnitudes for other parity groups were close to 1 with maximum deviation of 10% (Table 1). When all reflections to 1.64 Å resolution were used in the statistics, the ratios of $F_{hkl=e}/F_{hkl=o}$ are 2.1 for the >6 Å resolution shell, 1.2 for the 4–6 Å shell, and about 1 for remaining shells. Thus, the existence of systematically weak reflections of $F_{hkl=o}$ at 4 Å or lower resolution is evident, suggesting the presence of pseudo body-centered symmetry in the crystal of CyPC–CsA. Superposition of the four molecules in the asymmetric unit of the finally solved CyPC–CsA structure (Ke *et al.*, 1993) confirmed the pseudo-body-centered symmetry (Table 2). The largest deviation from the body-center of (0.5, 0.5, 0.5) is about 2 Å away from the translation of half a unit in the $Y$ dimension (Table 2).

Pseudo-translation in the chorismate mutase crystal appeared to be a twofold translation along the $a$ axis. The ratios of the average magnitudes between the parity groups of $h$ = even and $h$ = odd ($F_{h=e}/F_{h=o}$) are 3.1, 1.6, 1.2, and about 1, respectively for the resolution shells of 50–6, 6–4, 4–3.5 and <3.5 Å (Table 3). In contrast, the ratios for other parity groups of reflections were close to 1, with a maximum deviation of 10% (Table 3). The statistics of reflection magnitudes thus suggests that 12 molecules or four trimers of chorismate mutase in the crystallographic asymmetric unit are related to one another by translation of about half a unit cell along the $a$ axis. This prediction is consistent with the results of difference Patterson maps, in which pseudo-translation produced a large peak at (0.5, 0.0, 0.0). Also it is confirmed by the final chorismate mutase structure (Chook *et al.*, 1993) where the trimers of T1 and T3 are related to the trimers of T2 and T4 by translation of half a unit cell along the $a$ axis with deviations less than 3 Å (Table 4).

The above two examples showed that diffraction data at low resolution can be used to detect pseudo-translational symmetry in protein crystals. The statistics on intensity of reflections in three-dimensional $hkl$ space revealed that the ratios of the average magnitudes of structure factors were equal to 1 ± 0.1 for parity groups which are not related to the pseudo-translations. When pseudo-translation occurs, ratios of up to 3 were observed for the related parity groups (Tables 1 and 3). In short, the crystal structures of CyPC–CsA and chorismate mutase suggest that an average magnitude ratio of ≥1.2 for parity groups of structure factors may be sufficient to indicate the existence of pseudo-translation.

### 3.2. *Use of pseudo-translation in determination of the CyPC–CsA structure*

The preliminary characterization of the CyPC–CsA crystals revealed mirror symmetries in the zones of $[h0l]$, $[0kl]$, and $[hk0]$ of precession photos. In addition, the reduction of diffraction data according to the 222 symmetry yielded an $R_{\rm merge}$ of 0.047 for 47 292 independent reflections out of total 140 443 measurements. This excellent $R_{\rm merge}$ and the mirror symmetry in diffraction space led us to determine the CyPC–CsA structure in an orthorhombic space group, even though the true space group is $P2_1$ as observed in the final structure (Ke *et al.*, 1993).

The structure of CyPC–CsA was solved by molecular replacement, using the structure of cyclophilin A (Ke *et al.*, 1991) as the initial model. The *CCP*4 fast rotation function performed in the space group $P1$ revealed a unique 10$\sigma$ peak at ($\alpha$ = 110, $\beta$ = 75, $\gamma$ = −105°). This rotational orientation was used for the translation search by the program *X-PLOR* (Brünger *et al.*, 1987), respectively, in the four possible orthorhombic space

groups. After a series of trial, the space group $P2_12_12$ showed a superior result over the other orthorhombic space groups. For the translation search in the space group $P2_12_12$, the correlation coefficients were 0.342 for the strongest peak (0.10, 0.43, 0.13), 0.286 for the second peak (0.11, 0.44, 0.13 or 0.61, 0.94, 0.63), and 0.245 the third strongest peak (0.10, 0.45, 0.45). The strongest peak (0.10, 0.43, 0.13) had favorable intermolecular packing as displayed on a graphics system, and was, therefore, taken as one of the two molecules in the crystallographic asymmetric unit.

The statistics on intensity of reflections in three-dimensional *hkl* space revealed pseudo-body-center symmetry in the CyPC–CsA structure. This implies that the second molecule should have roughly the same rotational orientation as the first molecule, and a translational shift of about (0.5, 0.5, 0.5). Therefore, the knowledge of pseudo-body-center symmetry enabled us immediately to recognize the two strongest peaks from the translation search, (0.10, 0.43, 0.13) and (0.61, 0.94, 0.63), as the correct solutions. The model with these two molecules was refined first as rigid bodies to an $R$ factor of 0.488, and then refined using molecular dynamics protocols (Brünger *et al.*, 1987) to an $R$ factor of 0.383 against 13 951 reflections between 10 and 2.5 Å resolution.

Careful inspection of the electron density revealed that molecule $A$ in the asymmetric unit has a perfect symmetry of $P2_12_12$ [$(x, y, z)$, $(-x, -y, z)$, $(\frac{1}{2}+x, \frac{1}{2}-y, -z)$, and $(\frac{1}{2}-x, \frac{1}{2}+y, -z)$] while molecule $B$ only retains half the symmetries of $(x, y, z)$ and $(\frac{1}{2}-x, \frac{1}{2}+y, -z)$. Therefore, the structure was further refined in the space group $P2_1$ with the special origin at $(\frac{1}{4}, 0, 0)$ and the symmetries of $(x, y, z)$ and $(\frac{1}{2}-x, \frac{1}{2}+y, -z)$, to a final $R$ factor of 0.197 against 72 159 reflections between 6 and 1.64 Å resolution (Ke *et al.*, 1993).

### 3.3. *Application of pseudo-translation in the structure determination of chorismate mutase*

The crystallographic asymmetric unit of the $P2_1$ crystal of chorismate mutase contains 12 molecules of the enzyme, each of which has multiple heavy-atom sites in the five derivatives used in structure determination. A total of 24 heavy-atom sites in each of the four heavy-atom-soaked derivatives (ethyl mercury phosphate, iridium hexachloride, and two osmium hexachloride derivatives) and 72 selenium sites in the selenomethionine derivative were found in the crystallographic asymmetric unit.

The large number of heavy-atom sites coupled with low occupancy (low concentration of heavy-atom solutions were required to minimize non-isomorphism) resulted in noisy difference Patterson maps. Since many peaks were around $2\sigma$, all difference Patterson maps were plotted with initial contour of $1\sigma$ and increments of $0.5\sigma$ to ensure visualization of all heavy-atom–heavy-atom vector peaks. However, this low contour also brought up many noise peaks in the maps and thus the interpretation of the difference Patterson maps became extremely difficult. The difference Patterson map of the mercury derivative was manually interpreted after failure of automatic interpretation by direct methods (Fan *et al.*, 1990) or by *HASSP* (Terwilliger *et al.*, 1987). The prior identification of the twofold pseudo-translation along the $a$ axis played a key role in the identification of the true peaks and in the interpretation of the complicated Patterson map. First, only peaks that exhibited pseudo-translation were considered as candidates for Harker and cross peaks, and spurious noise peaks were therefore identified and disregarded. 12 mercury sites were located from the difference Patterson map and SIR phases from the refinement of these sites were then used in difference and cross-phase Fourier maps to locate more heavy-atom sites in the mercury derivative and in the other four derivatives. Second, only peaks that exhibited pseudo-translation along the $a$ axis were picked from the difference and cross-phase Fourier maps. These new sites were verified against the respective difference Patterson maps prior to heavy-atom parameter refinement. Therefore, application of pseudo translation in identification of heavy-atom sites in the chorismate mutase derivatives is twofold: (1) reduction of the number of 'unique' heavy-atom sites by half and (2) differentiation of real peaks from noise in the difference Patterson and difference Fourier maps.

## 4. Concluding remarks

(1) Pseudo-translation symmetry in the asymmetric unit of macromolecular crystals can be identified by statistics on the magnitudes of structure factors in three-dimensional *hkl* space. The protein structures of CyPC–CsA and chorismate mutase showed that pseudo-translation is most likely to exist when the ratios of the average magnitudes of certain parity groups are bigger than 1.2. Otherwise the ratio of average magnitude should be close to 1.0 for all parity groups of reflections.

(2) Pseudo-translation in atomic space yields a diffraction pattern with systematically weak reflections at 4 Å or lower resolution. This pattern is apparently similar to that of face-center or body-center, thus preventing the identification of the true crystal symmetry and space group. However, pseudo-translation can be differentiated from a true crystallographic symmetry by examination of diffraction at high resolution. Likewise, attention should be paid to possible loss of information when only strong reflections are used for structure determination.

(3) Pseudo-translation may hinder structure determination, as observed in many structures of small

molecules. On the other hand, knowledge of pseudo-translation may be useful in determination of protein structures. Crystal structures of CyPA–CsA and chorismate mutase are two examples of successful application of pseudo-translation in protein structure determination by molecular replacement and by multiple isomorphous replacement. Thus, examination of pseudo-translation is recommended in the early stages in determination of protein structures.

(4) Characterization of crystal system and space group by classic precession photos has been replaced at present by automatic data-collection software. The statistics presented in this paper could be an auxiliary procedure to data-collection software to detect pseudo-translation. A program *SUDOTR* which examines twofold pseudo-translation, pseudo-face-center, and pseudo-body-center is available upon request from hke@med.unc.edu.

We thank the NIH for support to WNL (GM06920) and to HK (AI33072), HHMI to YMC, and Life Sciences Research Foundation for a postdoctoral fellowship to YMC.

## References

Brünger, A. T., Kuriyan, J. & Karplus, M. (1987). *Science*, **235**, 458–460.

Chook, Y., Ke, H. & Lipscomb, W. N. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 8600–8603.

Fan, H. F., Qian, J. Z., Zheng, C. D., Gu, Y. X., Ke. H. & Huang, S. H. (1990). *Acta Cryst.* A**46**, 99–103.

Ke, H., Zhao, Y., Luo, F., Weissman, I. & Friedman, J. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 11850–11854.

Ke, H., Zydowsky, L. D., Liu, J. & Walsh, C. T. (1991). *Proc. Natl Acad. Sci. USA*, **88**, 9483–9487.

Terwilliger, T. C., Kim, S. H. & Eisenberg, D. (1987). *Acta Cryst.* A**43**, 1–5.