Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

X-ray diffraction of helices with arbitrary periodic ligand binding

Jin Gu and Leepo C. Yu*

National Institute of Arthritis, Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, MD 20892-2755, USA

Correspondence e-mail: lcyu@helix.nih.gov

Received 20 January 1999
Accepted 15 October 1999

The classical formalism for studying diffraction from helical structures extended to include ligand binding is presented. The diffraction from such a binding pattern is the convolution of the Fourier transforms of the helix and the one-dimensional binding distribution. It is shown in the present analysis that it is not necessary to assume that the binding distribution is strictly periodic, as long as its Fourier transform can be determined. Analysis of the convolution gives a general expression for the diffracted intensities and the selection rule for the layer-lines. It shows two groups of layer-lines: one group is the familiar layer-line set from the original helix, while the other group shows reciprocal spacings shifted by $1/a$ from the original helix layer-lines, where a is the average repeat of the binding distribution. This group of layer-lines is contributed by the ligand only. By way of examples, calculated diffraction patterns from muscle actin filaments with bound myosin heads in three different binding patterns are presented. This approach provides a method for determining the ligand-binding distribution along helices by an analysis of their X-ray diffraction patterns.

1. Introduction

The helix is one of the most common biological forms. The Fourier transform for a general helical structure was given by Cochran *et al.* (1952). The classical method of calculation of diffraction intensities from helices was provided by Franklin & Klug (1955). Worthington & Elliott (1989) reformulated it to allow the inclusion of paracrystalline disorder. Millane (1991) directly derived the helix-selection rule by generalizing the diffracted intensity of a non-crystalline fiber with atoms regularly distributed in one c repeat. In biological systems, helices with bound ligands occur frequently, *e.g.* in muscle fibers, the binding of myosin (cross-bridges) to the actin filament and the troponin molecules distributed along the actin filament. Holmes *et al.* (1980) modeled the X-ray diffraction patterns from insect flight muscle in rigor by introducing a periodic occupancy function for cross-bridges which modulates the diffraction intensity from the actin filament. Using a method similar to that of Holmes, layer-lines in the X-ray diffraction patterns from skeletal muscle in rigor, which has a symmetry different from that of the insect flight muscle, were explained by Yagi (1996). However, the distribution of ligand binding has been assumed to be periodic, although frequently the binding is marginally periodic, *i.e.* periodicity with perturbations ('perturbed periodicity'). In the present study, we introduce a general formalism for deriving selection rules and intensities originating from ligand binding distributed in a perturbed periodic function.

Table 1
Notation.

r, φ, z	Cylindrical coordinates in real space
R, ψ, Z	Cylindrical coordinates in reciprocal space
$\delta(z)$	Dirac delta function
σ	Electron density of the ligand partially decorated on the helix
σ_h	Electron density of the ligand fully decorated on the helix
$A(z)$	Distribution function of the ligand-binding sites in the helix
$F_A(Z)$	Fourier transform of $A(z)$
ξ_k	Ligand-binding site coordinate along the helix axis
$\Im$	A bound ligand containing a set of atoms
F	Structure factor of the ligand partially decorated helix
F_h	Structure factor of the ligand fully decorated helix
u_v	Helix with u repeat units in v turns
c	Helix repeat
c_b	Common repeat of the helix and its bound ligand
l	Index of the layer-line from the helix
l_b	Index of the layer-line from the bound ligand
$L(Z)$	One-dimensional distribution interference function
$J_n(R)$	The nth-order Bessel function of the first kind
$G_n(R, Z)$	Fourier–Bessel structure factor
$I(R, l)$	Cylindrical averaged intensity on layer-line l

2. Helical binding structure and its diffraction

The notation used is defined in Table 1.

Consider the case where ligands bind to the sites of a discrete u_v helix but do not fully decorate it. The ligand-binding sites in the z direction of the helical axis can be described by a distribution function $A(z)$ which is a set of planes at the binding sites, *i.e.* $A(z) = \sum\delta(z - \xi_k)$, where δ is the Dirac delta function and $k = 1, 2, 3, \ldots, N$. The 'ligand helix' (helix partially occupied by ligands) is equal to this set of planes multiplying the helix fully occupied by ligands (Fig. 1). If the subunit $\Im$ has M atoms, the fully decorated helix can be considered as the summation of helices decorated by those single atoms in $\Im$. The jth single atom in the helix has its origin at (r_j, φ_j, z_j). If the fully decorated helix is denoted by the electron-density function σ_h, the electron-density function σ of the 'ligand helix' can be expressed as $\sigma = \sum_j\sigma_j = \sum_j\sigma_h A_j(z)$, where $A_j(z) = \sum_{j,k}\delta(z - \xi_k - z_j)$ and $j \in \Im$. For simplicity, only the helix formed by the jth atom is considered. The diffraction pattern of such a structure is governed by its structure factor, which is the Fourier transform of σ_j, *i.e.*

$$F_j(R, \psi, Z) = \mathcal{F}(\sigma_j) = \mathcal{F}[\sigma_h A_j(z)] = F_h(R, \psi, Z)F_{A_j}(Z), \quad (1)$$

where $\mathcal{F}$ denotes the Fourier transform operation. In fact, $\mathcal{F}(\sigma_j)$ is the convolution of two structure factors, F_h and F_{A_j}. Biological systems are generally non-crystalline. The intensity diffracted by a non-crystalline fiber depends on the Fourier–Bessel structure factors (Franklin & Klug, 1955; Klug *et al.*, 1958). Following the derivation by Cochran *et al.* (1952) and by Millane (1991), the structure factor for the helix u_v is given by

$$F_h(R, \psi, l) = f_j \sum_{n=-\infty}^{\infty} \{J_n(2\pi Rr_j)\exp[in(\psi + \pi/2)] \times \exp[i(-n\varphi_j + 2\pi z_j l/c)]\} \times \sum_{p=0}^{u-1}\exp[i2\pi p(l - nv)/u], \quad (2)$$

which is non-zero only if $(l - nv)$ is a multiple of u, *i.e.* $l = nv + mu$. It is the layer-line selection rule of the helix u_v. By introducing the Dirac delta function δ to the layer-lines, (2) may be rewritten as

$$F_h(R, \psi, Z) = \sum_{n=-\infty}^{\infty} G'_n(R, \psi, l)\delta(Z - l/c), \quad (3)$$

where

$$G'_n(R, \psi, l) = f_j J_n(2\pi Rr_j)\exp[in(\psi + \pi/2)] \times \exp[i(-n\varphi_j + 2\pi z_j l/c)]. \quad (4)$$

Meanwhile, the structure factor of the binding distribution function A_j for the jth helix is given by

$$F_{A_j}(Z) = \mathcal{F}\left[\sum_{k=1}^{N}\delta(z - \xi_k - z_j)\right] = F_A(Z)\exp(2\pi i z_j Z). \quad (5)$$

Therefore, using the convolution theorem substituting (3), (4) and (5) into (1), the structure factor of the jth atom helix in the 'ligand helix' is given by

$$F_j(R, \psi, Z) = \int \sum_{n=-\infty}^{\infty} G'_n(R, \psi, l)\delta(t - l/c)F_A(Z - t) \times \exp[2\pi i z_j(Z - t)]\mathrm{d}t. \quad (6)$$

After simplifying, the structure factor becomes

$$F_j(R, \psi, Z) = f_j \sum_{n=-\infty}^{\infty} J_n(2\pi Rr_j)\exp[in(\psi + \pi/2)] \times \exp[i(-n\varphi_j + 2\pi z_j Z)]F_A(Z - l/c), \quad (7)$$

which is the summation of the product of layer-lines arising from the fully decorated helix structure factor and the binding-distribution structure factor F_A shifted by the corresponding layer-line. Furthermore, with summation of all the atoms over j in the ligand in (7), the entire structure factor of the 'ligand helix' is rewritten as

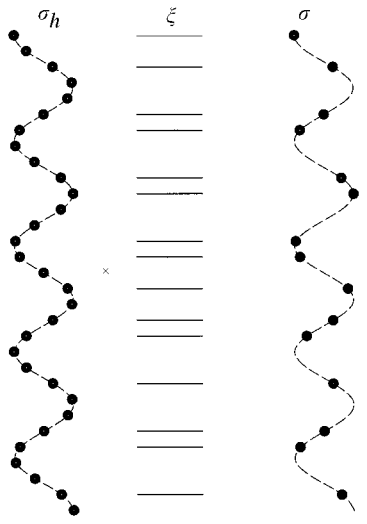

Figure 1
A 'ligand helix' σ is the product of a discontinuous helix σ_h and a set of planes ξ arrayed perpendicular to the helix axis.

$$F(R, \psi, Z) = \sum_{n\in\mathcal{H}} G_n(R, Z)\exp[in(\psi + \pi/2)]F_A(Z - l/c). \quad (8)$$

After cylindrical averaging and elimination of the cross terms of the different Bessel orders, the diffracted intensity of the 'ligand helix' is obtained by

$$I(R, Z) = \sum_{n\in\mathcal{H}} |G_n(R, Z)|^2 L(Z - l/c), \quad (9)$$

where

$$G_n(R, Z) - \sum_{j\in\Im} f_j J_n(2\pi R r_j)\exp[i(-n\varphi_j + 2\pi z_j Z)]. \quad (10)$$

$L(Z)$ [$= |F_A(Z)|^2$] is the one-dimensional interference function of the binding sites $A(z)$. The fiber-diffracted intensity formula in (9) is the extension from the case of the fully position-discrete helix in the classical method to the case of the partially positioned helix. $L(Z)$ can usually be obtained by the Fourier transform of the one-dimensional auto-correlation function (a.c.f.), *i.e.*

$$L(Z) = \mathcal{F}[A(z) * A(-z)]. \quad (11)$$

The a.c.f. of the binding distribution gives the binding-distribution characteristics in reciprocal space. Consider that the one-dimensional ligand distribution has a repeat of a that matches the repeat c of the original helix with a common repeat c_b, *i.e.* ($ka = qc = c_b$), where k and q are integers. The 'ligand helix' will exhibit layer-lines l_b ($= Zc_b$) only if $L(Z - l/c)$ has peaks at multiples of $1/a$ in reciprocal space, *i.e.* (9) has layer-lines following the selection rule

$$l_b = ql + ks = q(vn + um) + ks, \quad (12)$$

where l_b is the index of the layer-line from the 'ligand helix' and l is the index of the original helix layer-lines. All n, m and s are integers.

It is not necessary for the ligand to bind to the helix with an exact periodicity. In fact, (9) allows us to relax the strict periodic requirement and the ligand could bind to the helix in an arbitrary distribution. For example, if the separation of binding sites is in a Gaussian distribution with the mean distance a between neighbors, then

$$L(Z) = N + 2\sum_{m=1}^{N-1}(N - m)Q^m(Z)\cos(2\pi maZ), \quad (13)$$

where $Q(z) = \exp[-\pi(wZ)^2]$ (Vainshtein, 1966). Fig. 2 shows $L(Z)$ of the Gaussian distribution, which has periodic peaks with separation of $1/a$ in reciprocal space. Actually, in the special case when $Q^m(z) = 1$, the binding distribution is an equal separation distribution (Worthington & Elliott, 1989), *i.e.* the ligands bind to the helical sites with an exact separation 'a' along the helical axis, and its $L(Z)$ is equal to $\sin^2(\pi NaZ)/\sin^2(\pi aZ)$ (James, 1962).

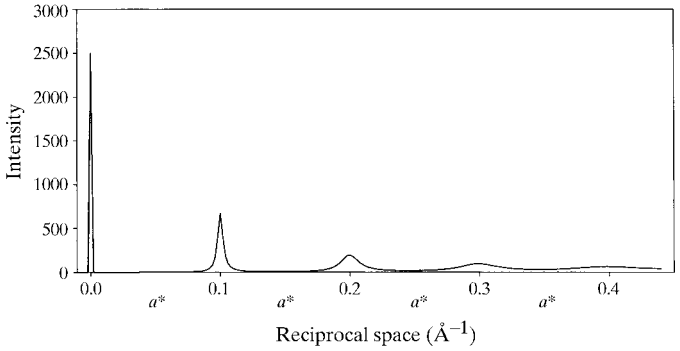

Figure 2
The interference function of binding sites with Gaussian distribution. The parameters of the distribution are the averaged period a = 10 Å ($aa^* = 1$), the standard deviation $\sigma = 0.08a$ [$w = (2\pi)^{1/2}\sigma$] and the lattice points number N = 50.

3. Examples of hypothetical binding of myosin heads to the actin filament

Here, we consider skeletal muscle as an example in order to demonstrate the relationship between the binding distribution and its X-ray diffraction pattern. In skeletal muscle, pairs of myosin heads have a helical symmetry distribution around the thick filament. The thick filament has three strands (Squire, 1972; Maw & Rowe, 1980) of 9_1 right-handed helices with a 14.3 nm interval (Huxley & Brown, 1967). The actin filament is a non-integer left-handed genetical helix close to 13_{-6} (Hanson & Lowy, 1963), with an interval of 2.75 nm between actin monomers. The true beat period between the myosin filament and the actin filament is about 216 nm (Squire & Harford, 1988). The ratio of myosin filaments to actin filaments is 1:2, so that on average there are three myosin heads from a single myosin filament in every 14.4 nm interval along actin (Huxley & Brown, 1967). Along the actin helix, there are 45 myosin heads from the myosin filament *versus* 78 actin monomers in 216 nm, if all myosin heads are bound to the actin filament. For the sake of simplicity, we assume that the relative orienta-

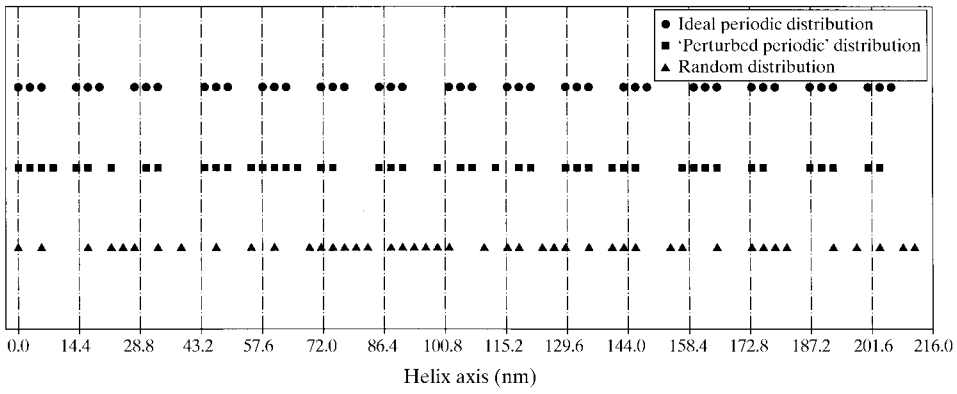

Figure 3
Distributions of the myosin heads bound to actin filament: (i) an ideal periodic distribution (filled circles), (ii) 'perturbed periodic' distribution (filled squares) and (iii) random distribution (filled triangles).

tions between the actin and the myosin filaments are all equivalent.

We assume that the 'decorated' actin filament has an ideal periodic distribution of the binding sites, *i.e.* there are 15 group levels of myosin heads bound to actin filaments with a 14.4 nm period in the range of 216 nm.

$$A(z) = \sum_{n=0}^{n=44} \delta(z - \xi_n),$$

where

$$\xi_{3j} = 2.75\text{rint}[216j/(15 \times 2.75)],$$
$$\xi_{3j+1} = \xi_{3j} + 2.75,$$
$$\xi_{3j+2} = \xi_{3j} + 5.5,$$

rint(x) is the integer nearest x and $j = 0, \ldots, 14$. The binding distribution and its interference function are shown in Figs. 3 and 4. The interference function shows a 1/14.4 nm^{-1} period in reciprocal space. The repeat c of the actin helix is about 36 nm, so that the common repeat between the actin helix and the binding distribution is 72 nm, *i.e.* $q = 2$ for actin period and $k = 5$ for the period of the binding distribution corresponding to the selection rule (12). The selection rule of layer-lines is then given by

$$l_b = 2(-6n + 13m) + 5s = -12n + 26m + 5s.$$

Fig. 4(a) also shows that $L(Z)$ has only one prominent periodic peak, *i.e.* $s = 0$ or ± 1 in the selection rule. The diffraction pattern of the ideal periodical binding is shown in Fig. 5(a). The layer-lines can be divided as two categories: l_{2j} where $s = 0$ and $l_{|2j\pm5|}$ where $s = \pm 1$. By indexing to common repeat 72 nm, layer-lines l_{2j} corresponding to reflections 36.0, 18.0, ..., 5.9 and 5.1 nm *etc*. are in the original actin layer-line positions. These layer-lines are not only enhanced in intensity, but the lateral position of the peaks on the layer-lines also moves closer to the inner radius because the mass center of the bound myosin has a larger radius than the actin monomer in the actin helix. In contrast, layer-lines $l_{|2j\pm5|}$ corresponding to reflections l_3 = 24.0 nm and l_7 = 10.4 nm *etc*. only come from bound myosin heads and the positions of these layer-lines are shifted from the actin layers to $l_{|2j\pm5|}$. These shifted layer-lines have the same order of the Bessel function as those original actin helix layer-lines. However, the intensities of these layer-lines are much lower than the original actin layer-lines, because the intensity is modulated by the $L(Z)$. The intensities of the original lines, *i.e.* modulated by $L(0)$, is proportional to the square of the bound ligand number N^2 and the intensities of the shifted lines is modulated by $L(Z \neq 0)$, which is much less than N^2.

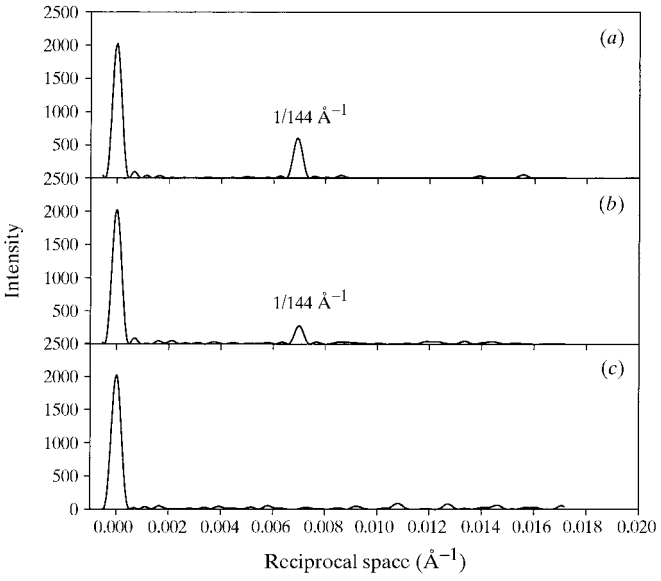

Figure 4
The binding interference functions corresponding to three difference distributions in Fig. 3. The binding-distribution interference function has only one prominent peak at 1/144 Å^{-1} in ideal periodic distribution (a), and in 'perturbed periodic' distribution (b). However, the intensities of peaks in (b) is lower than that in (a). There is no prominent peak when the binding distribution is random (c).

As a second example, we modify the ideal periodic distribution in the first example by randomly removing one of the three myosins in each periodic group and randomly redistributing them. This results in a 'perturbed periodic' distribution (Fig. 3). The diffraction pattern from this type of binding distribution is similar to that of the first example (Figs. 4a and 5b). The layer-lines l_{2j} are the same in both examples; the shifted layer-lines $l_{|2j\pm5|}$ are in the same position in both examples. However, the intensities of the layer-lines are weaker than those in the first example, since the amplitude of $L(1/14.4)$ is lower.

In contrast, if the myosin heads bind to the actin helix totally randomly (Fig. 3), the binding distribution interference function $L(Z)$ has no significant peak (Fig. 4c) besides $L(0) = N^2$. It is not surprising that the pattern in Fig. 5(c) is similar to that from the actin filament alone except that the lateral positions of peaks move inward in the layer-lines.

The three examples illustrate the feasibility of distinguishing binding distributions from one another by their diffraction patterns. By comparing Figs. 5(a) and 5(b) with the experimental data (Xu *et al.*, 1997), our preliminary analysis indicates that the myosin heads of skeletal muscle in the rigor state bind to actin in a periodic distribution rather than in a random distribution. Details of the analysis will appear elsewhere.

4. Discussion and concluding remarks

The advantage of the method presented here is its generality and simplicity. The distribution of ligands is not required to be strictly periodic as long as the interference function (Fourier transform) of the distribution can be determined. It follows that (9) can be used to interpret diffraction patterns. From the intensity distribution of the layer-lines indexed on the original helix, the mass distribution of the bound ligands can be determined using the methods normally applied to general helical structures. The second group of layer-lines are more informative about the distribution of the ligand binding. If ligands bind to the helix with some regularity, the interference

function $L(Z)$ should exhibit a peak in reciprocal space. This leads to at least one extra observable layer-line on the meridian. Conversely, the binding pattern on the helix can be derived by indexing the second group of layer-lines. By combining the information from the two groups of layer-lines, the distribution of the ligand binding on helices can be resolved.

The principle of this work also can be extended to helices with defects: *i.e.* helices with some subunits missing or with 'holes'. One may consider such a helix as an ideal one intersecting with planes which describe the remaining subunits. (9) can then be applied to analyze the diffraction pattern. Accordingly, the diffraction pattern will have two groups of layer-lines if the remaining subunits have some periodicity other than that of the original ideal helix. On the other hand, the diffraction pattern will show only one group of the familiar layer-lines of the unperturbed helix if the remaining subunits are randomly distributed. This approach was applied to the diffraction patterns from the relaxed rabbit skeletal muscle, where the myosin heads of the thick filament are distributed in two structural populations: one is ordered in a helical array and the other is disordered (Xu *et al.*, 1997, 1999; Malinchik *et al.*, 1997). When a myosin head becomes disordered, it leaves a 'hole' in the array. Using an approach similar to the present one, we concluded that the two populations are in dynamic equilibrium, since the diffraction patterns indicated a random distribution of 'holes'.

In conclusion, the formalism for the helical ligand binding described here provides a convenient way to interpret the X-ray diffraction patterns of helices with ligand binding and determine their structures.

We would like to thank Dr Sengen Xu of the Laboratory of Physical Biology, NIAMS, NIH for stimulating discussions.

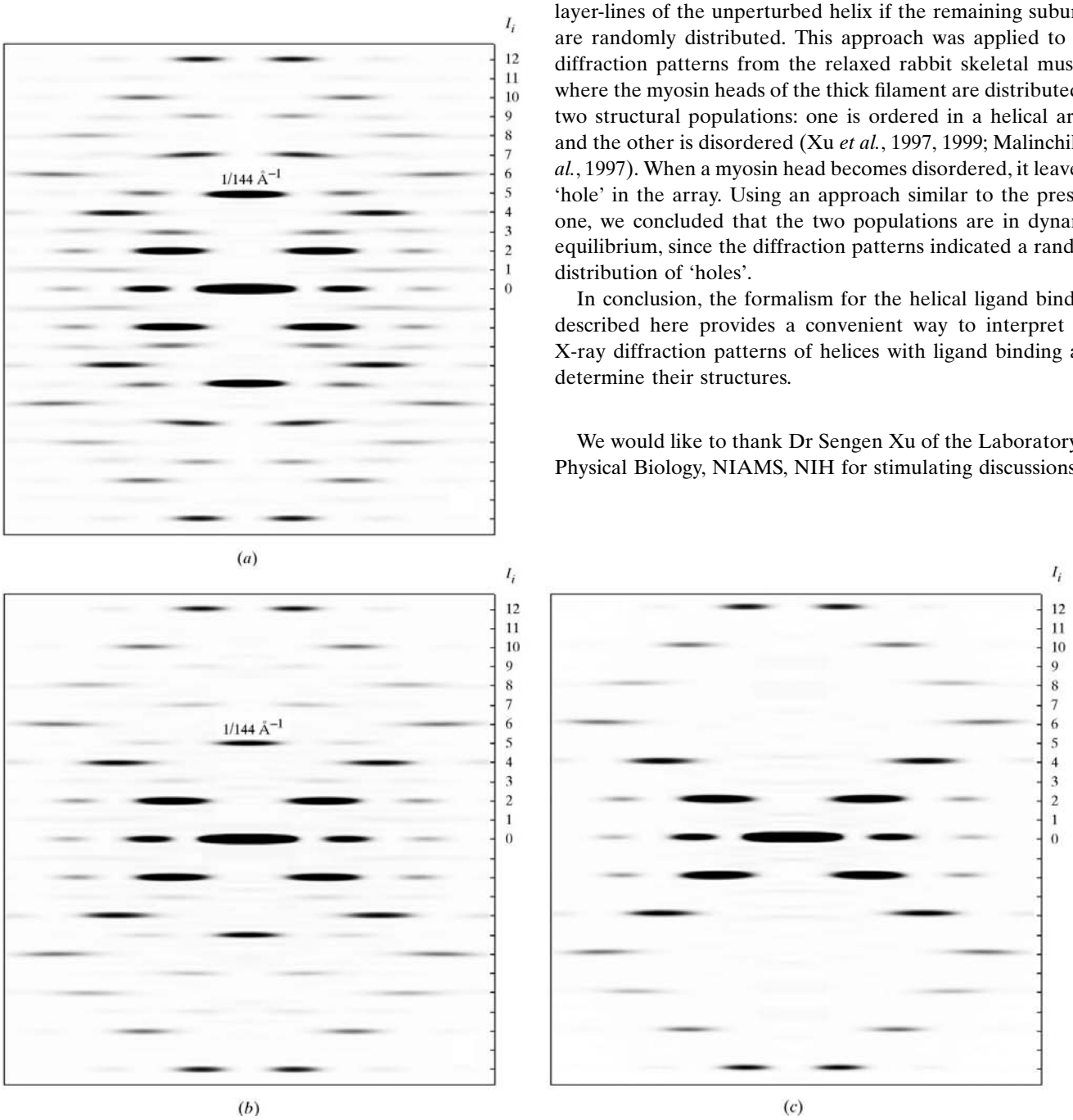

Figure 5
The diffraction patterns of three different binding distributions are simulated. The X-ray diffraction pattern (*a*) has two groups of layer-lines (see text) because the interference function (Fig. 4*a*) has only one prominent peak at 1/144 Å^{-1}. The diffraction pattern (*b*) is similar to the pattern (*a*) except that the intensities of the second group of layer-lines are much weaker. The diffraction pattern (*c*) looks like the diffraction pattern of actin helix alone if the myosin cross-bridges bind to actin helix randomly, since the interference function (Fig. 4*c*) does not have any significant peak. All the layer-lines l_i are indexed to the repeating length c_b = 720 Å.

References

Cochran, W., Crick, F. H. C. & Vand, V. (1952). *Acta Cryst.* **5**, 581–586.
Franklin, R. E. & Klug, A. (1955). *Acta Cryst.* **8**, 777–780.
Hanson, J. & Lowy, J. (1963). *J. Mol. Biol.* **6**, 46–60.
Holmes, K. C., Tregear, R. T. & Barrington Leigh, J. (1980). *Proc. R. Soc. London Ser. B*, **207**, 13–33.
Huxley, H. E. & Brown, W. (1967). *J. Mol. Biol.* **30**, 383–434.
James, R. W. (1962). *The Crystalline State*, Vol. II. London: Bell.
Klug, A., Crick, F. H. C. & Wyckoff, H. W. (1958). *Acta Cryst.* **11**, 199–213.
Malinchik, S., Xu, S. & Yu, L. C. (1997). *Biophys. J.* **73**, 2304–2312.
Maw, M. C. & Rowe, A. J. (1980). *Nature (London)*, **286**, 412–414.
Millane, R. P. (1991). *Acta Cryst.* A**47**, 449–451.
Squire, J. M. (1972). *J. Mol. Biol.* **72**, 125–135.
Squire, J. & Harford, J. (1988). *J. Muscle Res. Cell. Motil.* **9**, 831–851.
Vainshtein, B. K. (1966). *Diffraction of X-rays by Chain Molecules*. Amsterdam: Elsevier.
Worthington, C. R. & Elliott, G. F. (1989). *Acta Cryst.* A**45**, 645–654.
Xu, S., Gu, J., Rhodes, T., Belknap, B., Rosenbaum, G., Offer, G., White, H. & Yu, L. C. (1999). *Biophys. J.* **77**, 2665–2676.
Xu, S., Malinchik, S., Gilroy, D., Kraft, Th., Brenner, B. & Yu, L. C. (1997). *Biophys. J.* **73**, 2292–2303.
Yagi, N. (1996). *Acta Cryst.* D**52**, 1169–1173.