

Structural insights into the specific anti-HIV property of actinohivin: structure of its complex with the $\alpha(1-2)$ mannobiose moiety of gp120

M. Mominul Hoque,^{a,b} Kaoru Suzuki,^c Masaru Tsunoda,^a Jiandong Jiang,^d Fang Zhang,^d Atsushi Takahashi,^a Naomi Ohbayashi,^a Xiaoxue Zhang,^d Haruo Tanaka,^{a,d,e} Satoshi Ōmura^f and Akio Takénaka^{a,d,g,*}

^aFaculty of Pharmacy, Iwaki Meisei University, 5-5-1 Chuodai-lino, Iwaki, Fukushima 970-8551, Japan, ^bDepartment of Biochemistry and Molecular Biology, Rajshahi University, Rajshahi, Bangladesh, ^cCollege of Science and Engineering, Iwaki Meisei University, Iwaki, Fukushima 970-8551, Japan, ^dGraduate School of Science and Engineering, Iwaki Meisei University, 5-5-1 Chuodai-lino, Iwaki, Fukushima 970-8551, Japan, ^eKIIM Pharmaceutical Laboratories Inc., Fukushima 970-8551, Japan, ^fKitasato Institute for Life Sciences, Kitasato University, Tokyo 108-8641, Japan, and ^gGraduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, Yokohama 226-8501, Japan

Correspondence e-mail:
atakenak@sakura.email.ne.jp

Actinohivin (AH) is an actinomycete lectin with a potent specific anti-HIV activity. In order to clarify the structural evidence for its specific binding to the $\alpha(1-2)$ mannobiose (MB) moiety of the D1 chains of high-mannose-type glycans (HMTGs) attached to HIV-1 gp120, the crystal structure of AH in complex with MB has been determined. The AH molecule is composed of three identical structural modules, each of which has a pocket in which an MB molecule is bound adopting a bracket-shaped conformation. This conformation is stabilized through two weak C—H...O hydrogen bonds facilitated by the $\alpha(1-2)$ linkage. The binding features in the three pockets are quite similar to each other, in accordance with the molecular pseudo-threefold symmetry generated from the three tandem repeats in the amino-acid sequence. The shape of the pocket can accept two neighbouring hydroxyl groups of the O³ and O⁴ atoms of the equatorial configuration of the second mannose residue. To recognize these atoms through hydrogen bonds, an Asp residue is located at the bottom of each pocket. Tyr and Leu residues seem to block the movement of the MB molecules. Furthermore, the O¹ atom of the axial configuration of the second mannose residue protrudes from each pocket into an open space surrounded by the conserved hydrophobic residues, suggesting an additional binding site for the third mannose residue of the branched D1 chain of HMTGs. These structural features provide strong evidence indicating that AH is only highly specific for MB and would facilitate the highly specific affinity of AH for any glycoprotein carrying many HMTGs, such as HIV-1 gp120.

Received 17 July 2012
Accepted 25 September 2012

PDB Reference: actinohivin
in complex with MB, 4den

1. Introduction

HIV/AIDS is a major health concern, a global pandemic which remains a relatively uncontrolled infectious disease. Currently, over 20 kinds of inhibitors targeting HIV enzymes (*e.g.* reverse transcriptase, integrase and protease) are used as medicines to disturb the HIV life cycle after HIV entry into cells (Jegade *et al.*, 2008). These antiretroviral drugs have recently been evaluated further for their dual effects (Cohen, 2011) as treatments for and in the prevention of HIV infection (Sigal *et al.*, 2011; Cohen *et al.*, 2011). In addition, some proteins which are able to bind the surface glycoprotein of HIV are now expected to prevent HIV entry into cells (Balzarini, 2007), as shown in Fig. 1. This effect (entry inhibition) is also applicable to help suppress the spread of infection. Structurally, trimeric gp120 protruding from the HIV surface binds to human CD4⁺ to initiate entry and then to a chemokine receptor: CCR5 or CXCR4 (Berger *et al.*, 1999). Each gp120 is highly glycosylated

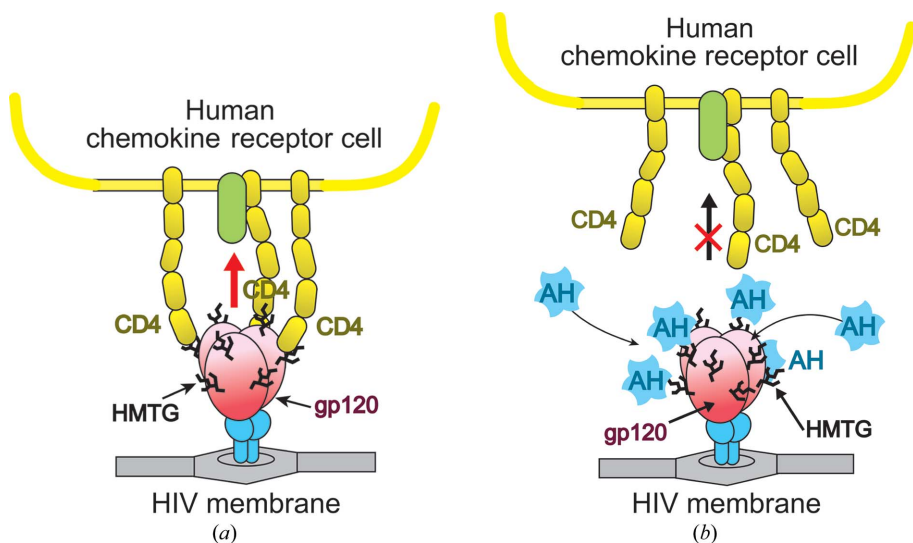


Figure 1
Schematic illustrations of HIV entry into the human cell (a) and its blocking by AH (b).

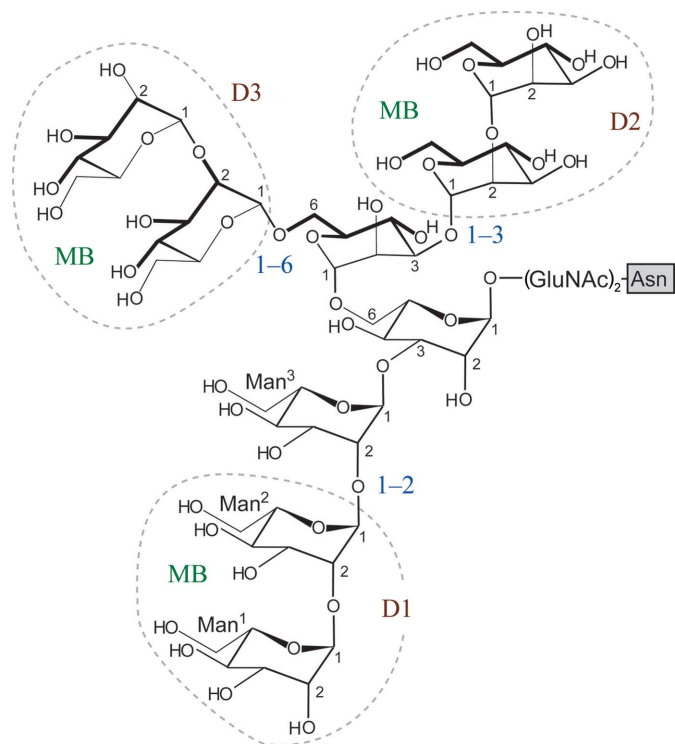


Figure 2
Chemical structure of HMTG. GlcNAc indicates *N*-acetylglucosamine, which is bound to an Asn residue. Here the numbering of mannose residues starts from the end of the branched chain.



Figure 3
The primary structure of AH. The conserved residues are shaded. The numbering system is shown on the right. The arrows indicate two residues which are crucially different between the three repeated sequences.

(Leonard *et al.*, 1990) to cover the surface with high-mannose-type glycans (HMTGs; see Fig. 2). Several carbohydrate-binding proteins (lectins) have been isolated and characterized as candidates for suppressing gp120 binding to susceptible cells. Among them, cyanovirin-N (CV-N) has already been intensively investigated to determine its structural properties, carbohydrate-binding potential and antiviral activity (Bewley & Otero-Quintero, 2001; Bewley *et al.*, 2002; Fromme *et al.*, 2008; Tsai *et al.*, 2004). Another lectin, griffithsin (GRFT), isolated from a red alga (*Griffithsia* sp.), also exhibits a high binding affinity for HMTG of gp120 (Ziółkowska *et al.*, 2007; Moulaei *et al.*, 2010). These are strong candidates for development as microbicides to prevent HIV transmission.

Similarly, we have independently discovered a new lectin, actinohivin (hereafter designated AH), from an actinomycete, *Longispora albida* K97-0003^T (Matsumoto *et al.*, 2003), which possesses a potent specific anti-HIV activity (Chiba *et al.*, 2001). This lectin inhibits the entry of various HIV-1 and HIV-2 strains into susceptible cells, as well as T-cell-tropic and macrophage-tropic syncytium formation, through binding to HMTGs of the HIV envelope glycoprotein gp120 (Chiba *et al.*, 2001, 2004). Previous research (Tanaka *et al.*, 2009; Takahashi *et al.*, 2010, 2011) suggested that the small lectin AH composed of 114 residues (Fig. 3) binds to three HMTGs at the D1 chain $\alpha(1-2)$ mannobiose (hereafter referred to as MB) moiety of the three branched chains of HMTG (Fig. 2). AH is thus expected to be developed as another candidate for a useful antiretroviral drug. X-ray analysis of the apo form of AH identified three binding pockets formed tandemly in the single peptide. This structural situation is quite different from those found in CV-N (Botos *et al.*, 2002) and GRFT (Ziółkowska *et al.*, 2007), which bind to one HMTG; in particular, GRFT has three binding sites to which the three chains are separately bound (Moulaei *et al.*, 2010). Therefore, detailed structural knowledge of the binding of AH to MB is essential to modify AH as a useful drug. In the present study, we have successfully determined the crystal structure of AH in complex with MB.

2. Materials and methods

2.1. Crystallization

AH was purified from a cultured broth of *L. albida* K97-0003^T as described previously (Tanaka *et al.*, 2009). MB was purchased from Sigma Chemical Co. (St Louis, Missouri, USA). Crystallization conditions of AH in complex with MB were surveyed by the hanging-drop vapour-diffusion method at 298 K using crystallization kits from Hampton Research (California, USA) and Emerald BioSystems (Washington,

Table 1

Data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

| | |
|--|------------------------------------|
| Data collection | |
| X-ray source | AR-NW12, PF |
| Wavelength (Å) | 1.00 |
| Resolution (Å) | 39.75–1.60 (1.63–1.60) |
| Observed reflections | 167074 |
| Unique reflections | 8104 (1160) |
| Completeness (%) | 99.9 (100.0) |
| $R_{\text{merge}}^{\dagger}$ | 0.089 (0.376) |
| $\langle I/\sigma(I) \rangle$ | 18.1 (6.7) |
| Crystal data | |
| Space group | $P2_13$ (disordered $P2_12_12_1$) |
| Unit-cell parameters (Å) | $a = b = c = 56.2$ |
| Z^{\ddagger} | 1/3 |
| V_M (Å ³ Da ⁻¹) | 3.54 |
| Structure refinement | |
| Resolution (Å) | 28.08–1.60 (1.64–1.60) |
| Reflections used | 7647 |
| R factor § | 0.145 (0.197) |
| $R_{\text{free}}^{\parallel}$ | 0.202 (0.290) |
| No. of protein atoms | 875 |
| No. of ions | 3 K ⁺ |
| No. of water molecules | 71 |
| R.m.s.d. | |
| Bond lengths (Å) | 0.021 |
| Bond angles (°) | 2.2 |
| Ramachandran plot (%) | |
| Most favoured regions | 97.3 |
| Additionally allowed regions | 2.7 |

$^{\dagger} R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ is the i th measurement of the intensity of reflection hkl and $\langle I(hkl) \rangle$ is its mean value. ‡ Number of proteins in the asymmetric unit. § R factor = $\sum_{hkl} ||F_{\text{obs}}| - |F_{\text{calc}}|| / \sum_{hkl} |F_{\text{obs}}|$, where $|F_{\text{obs}}|$ and $|F_{\text{calc}}|$ are the observed and calculated structure-factor amplitudes, respectively. $^{\parallel}$ R_{free} was calculated using a random data set containing 5% of the observations that were not included throughout refinement (Brünger, 1992).

USA). Protein droplets prepared by mixing 2 μl 10 mg ml⁻¹ AH solution containing 10 mg ml⁻¹ MB in pure water and 2 μl reservoir solution were equilibrated against 700 μl reservoir solution. Several conditions under which crystals appeared were further optimized by varying the concentrations of AH, precipitants and salts at different pH values. Crystals suitable for X-ray experiments grew during 20–30 d using a reservoir solution consisting of 20% (w/v) polyethylene glycol 1000, 0.2 M sodium chloride in 0.1 M sodium/potassium phosphate buffer pH 6.2.

2.2. X-ray data collection and processing

The crystals obtained were transferred into a cryoprotectant solution (a 1:1 mixture of the reservoir solution and 80% glycerol solution) for 30 s and mounted on a CryoLoop (Hampton Research) for flash-cooling. X-ray data were obtained at 100 K using synchrotron radiation of wavelength 1.00 Å on beamline AR-NW12 at the Photon Factory (PF), Tsukuba, Ibaraki, Japan. A crystal which showed higher order reflections with sharp spots was chosen. Diffraction patterns were taken at 1° oscillation steps with 20 s exposure per frame (a total of 180 frames were obtained) using an ADSC Quantum 210 CCD detector (Area Detector Systems Corp., California, USA). The Bragg spots were indexed and their integrated intensities were scaled between the frames and converted to amplitudes using *iMOSFLM* (Battye *et al.*, 2011)

with *SCALA* in *CCP4* (Winn *et al.*, 2011). Crystal data and diffraction processing statistics are summarized in Table 1. The unit-cell dimensions indicate that the most plausible Matthews coefficient (V_M) is 3.54 Å³ Da⁻¹ with the cell containing four AH molecules.

2.3. Structure determination and refinement

The approximate phase angles of the reflections were estimated by the molecular-replacement technique using *AMoRe* (Navaza, 1994) from *CCP4* with the apo form of AH (PDB entry 3a07; Tanaka *et al.*, 2009) as a phasing probe. The replaced atomic parameters of AH were refined by the restrained maximum-likelihood least-squares technique using *REFMAC5* (Murshudov *et al.*, 2011) from *CCP4*. The molecular occupancy was assumed to be one third in the cubic form. The $F_o - F_c$ maps showed electron densities assignable to bound MBs, potassium cations and water molecules. After several steps of *REFMAC* refinement with additional atoms, the molecular structures were revised by interpreting OMIT maps at every residue using *Coot* (Emsley & Cowtan, 2004). The stereochemistry of the protein structures was verified using *PROCHECK* (Laskowski *et al.*, 1993). Statistical data of the structure determination are summarized in Table 1. Fig. 4 shows electron-density maps visualized with *DINO* (Philippson, 2003) and Fig. 5 shows an overall view of the three disordered AH molecules depicted with *RasMol* (Sayle & Milner-White, 1995). Figs. 6(a), 6(b), 7, 8 and 9 were produced with *PyMOL* (DeLano, 2002). Models of AH with other isomers of mannobiose were constructed using *Coot* and that with HMTG was constructed using *QUANTA* (Accelrys, California, USA).

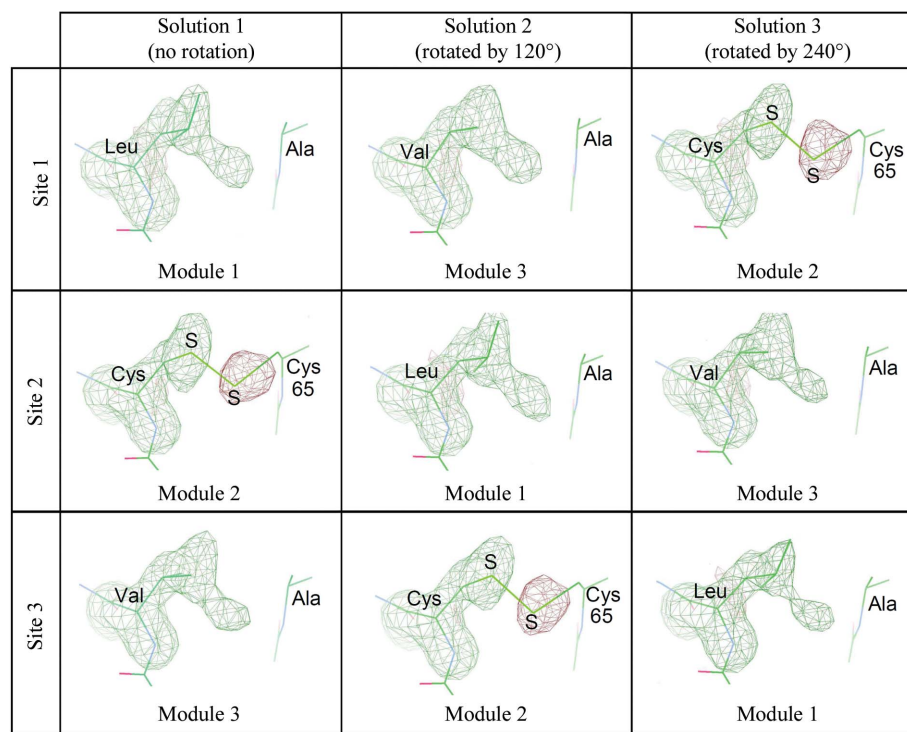
3. Results

3.1. Crystal structure

Data processing of the diffraction patterns indicated two possible space groups, $P2_12_12_1$ and $P2_13$, with the same unit-cell size. They differ in whether or not the unit cell contains threefold axes in the body-diagonal directions. In addition, the calculated V_M value predicted that the asymmetric unit in $P2_12_12_1$ contains a complete AH molecule. A similar example was reported for the crystal of the pseudosymmetric PSPC1–NONO heterodimer (Lee *et al.*, 2011). $R_{\text{p.i.m.}}$ (Evans, 2006) would be a reliable index to distinguish between these cases. The calculated values of R_{merge} and $R_{\text{p.i.m.}}$ are 0.087 and 0.034, respectively, for $P2_12_12_1$ data and 0.089 and 0.029, respectively, for $P2_13$ data. The lower $R_{\text{p.i.m.}}$ value for $P2_13$ suggests that the latter space group is more plausible. If the cubic space group is chosen, however, the asymmetric unit must contain one third of an AH molecule. This means that an AH molecule is disordered around the crystallographic threefold axis. For an initial examination, structure analyses were attempted using the $P2_12_12_1$ data. The application of molecular replacement gave three unique solutions that differed by a $\pm 120^\circ$ rotation around a body-diagonal axis of the $P2_12_12_1$ unit cell. Each solution structure was refined to an R factor of 0.213 and an

R_{free} of 0.235. Fig. 4(a) shows the resultant OMIT maps calculated without the Leu13, Cys51 and Val89 residues (see Fig. 3), which are related by the molecular pseudo-threefold symmetry within the molecule but differ in amino-acid species between the three modules (Tanaka *et al.*, 2009). In the second module Cys51 forms a disulfide bond to Cys65, while in the

first and the third modules the corresponding residues are Leu13 and Val89, which are exposed to Ala27 and Ala103, respectively. All of the shapes of the densities in Fig. 4(a) are similar to each other in the three sites and in the three solutions, suggesting a mixture of the three amino acids. The negative density coloured brown corresponds to the S atom of Cys65, which was not omitted, but corresponding densities appear in every map of the three solutions. This strongly suggests that Cys65 does not fully occupy one site and that it is distributed to the other sites rotated by $\pm 120^\circ$. The remaining densities appear as positive densities between Leu65 and Ala27 and between Val89 and Ala103 in the three solutions.



(a)

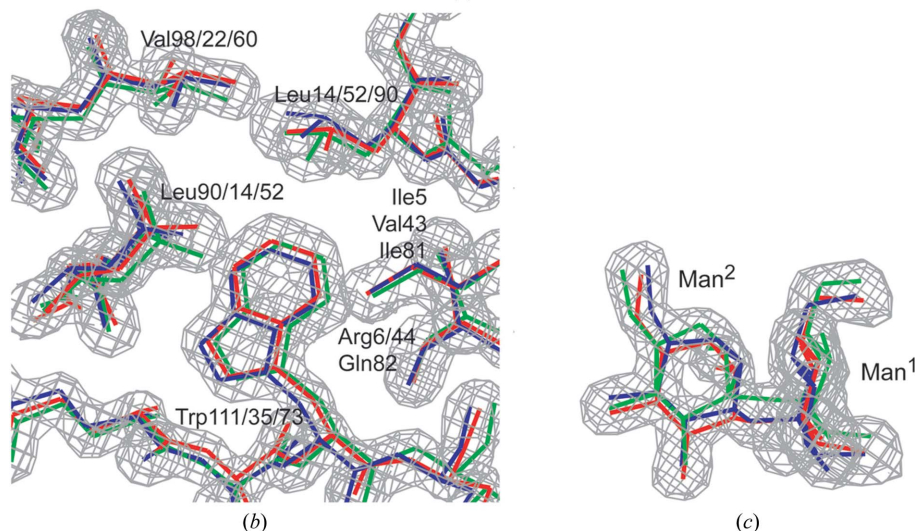


Figure 4

(a) $F_o - F_c$ OMIT maps at residues 13, 51 and 89 of AH when the three non-disordered crystal structures obtained by the molecular-replacement method were refined separately using the $P_{2_1}2_12_1$ processed data. From solution 1, the three structural modules of AH are found at three sites in the crystal. In the different solutions, the module rotates between the three sites. Therefore, the amino-acid residues in the same site vary between solutions, e.g. Leu13 \rightarrow Val89 \rightarrow Cys51 at site 1. The densities are contoured at the 3σ level. The negative density coloured brown indicates the S atom of Cys65, suggesting that it is not fully occupied. Its residual density appears between the two hydrophobic amino-acid residues in the other modules. Local $2F_o - F_c$ map (contoured at the 1.5σ level) of AH (b) and an OMIT map (contoured at the 3.0σ level) of MB (c). The three disordered molecules are coloured red, green and blue, respectively.

As described above, intensity distribution statistics and Fourier transformation of high-resolution $P_{2_1}2_12_1$ data consistently show that an AH molecule is disordered with the crystallographic threefold rotation symmetry in the crystal. Therefore, it is considered that the three AH rotamers generated by $\pm 120^\circ$ rotation around the molecular pseudo-rotation axis are packed randomly in the unit cell according to the $P_{2_1}2_12_1$ symmetry and that the whole crystal exhibits an apparent space group $P_{2_1}3$ as an averaged structure. This type of disorder would be a rare case in which an asymmetric macromolecule crystallizes according to the crystallographic symmetry based on its pseudo-molecular symmetry. For these reasons, the structure determination was performed under the constraints of space group $P_{2_1}3$ using the $P_{2_1}3$ data. The refined crystal structure gave a lower R factor of 0.145 ($R_{\text{free}} = 0.202$). Figs. 4(b) and 4(c) show the final $2F_o - F_c$ maps, in which the three disordered structures are well fitted, as well as those of the bound MB molecules. Whole views of the disordered three AH molecules are shown in Fig. 5. The atomic coordinates have been deposited in the PDB with accession code 4den. Although the crystal used contained disordered AH molecules, the X-ray diffraction at high resolution and the high symmetry of the AH molecule helped us to successfully reveal the detailed structure of the MB-bound state of AH. A Ramachandran plot of the polypeptide shows that all of the main-chain atoms fall within

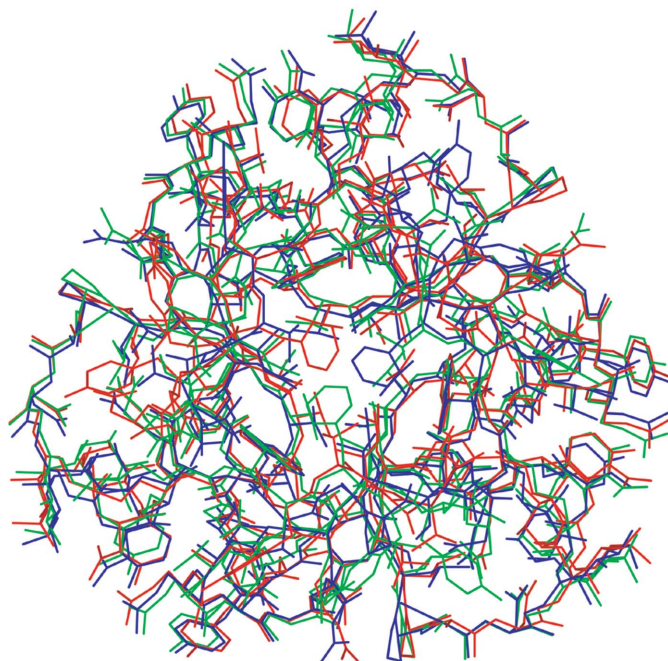


Figure 5
Three disordered AH molecules in the asymmetric unit viewed down the crystallographic threefold axis.

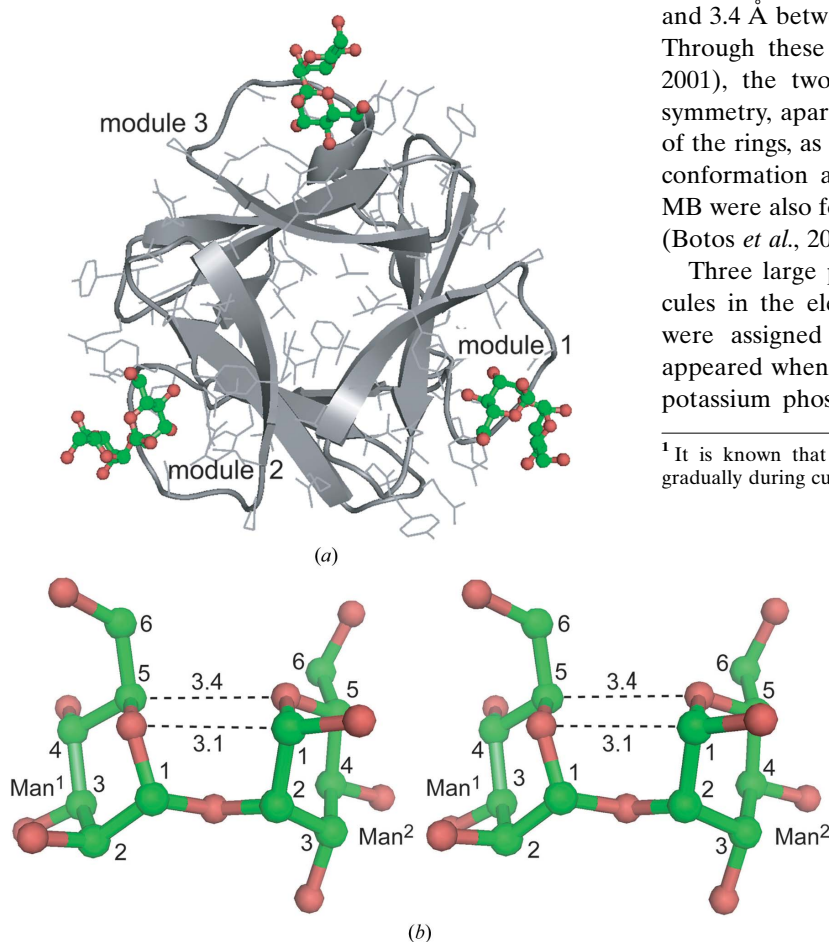


Figure 6
A overall view of AH and MBs bound in the three pockets of the modules (a), a stereo diagram (b) of MB with bracket-shaped conformation and its chemical structure (c). Dashed lines indicate possible C–H \cdots O hydrogen bonds with their average distances.

allowed regions, with 97.3% of the residues in favoured regions and 2.7% of the residues in allowed regions.

3.2. Overall structure of AH and MB

In the apo form (Tanaka *et al.*, 2009), the two AH structures found in the asymmetric unit of the crystal are essentially the same, with an r.m.s.d. of 0.14 Å between C α atoms when they are superimposed onto each other. While the r.m.s.d.s between the MB-bound state and the two apo forms are 0.27 and 0.31 Å, the values for all atoms including side chains are 0.75 and 0.84 Å. These data suggest that the overall conformation is rather rigid, but that the side chains are slightly responsive to target binding. In the apo forms, an additional extension¹ of two residues at the N-terminus is visible in the electron-density map. In the complex, however, it is difficult to assign these residues even in the final $F_o - F_c$ map, despite the protein sample being purified from the same batch. These two N-terminal residues are invisible, perhaps owing to the packing disorder of the AHs.

Three MB molecules are bound to an AH molecule, as shown in Fig. 6(a). The two mannose residues (Man¹ and Man²) of each MB are close together (see Fig. 6b), so that two C–H \cdots O interactions are made between them. The average C \cdots O distances are 3.1 Å between C¹(Man²) and O(Man¹) and 3.4 Å between C⁵(Man¹) and O(Man²) in the three MBs. Through these weak hydrogen bonds (Desiraju & Steiner, 2001), the two hexose rings are oriented to have mirror symmetry, apart from the atomic species and the side groups of the rings, as seen in Figs. 6(b) and 6(c), to form a compact conformation as a bracket shape. Similar conformations of MB were also found as a common feature in the lectins CV-N (Botos *et al.*, 2002) and GRFT (Moulaei *et al.*, 2010).

Three large peaks were found near the bound MB molecules in the electron-density map, to which potassium ions were assigned because the crystals that were used only appeared when the protein solution contained 0.1 M sodium/potassium phosphate buffer. Through the mediation of the

¹ It is known that AH is matured by reducing the N-terminal extension gradually during cultivation of the microorganism (Suzuki *et al.*, 2012).

two potassium ions, the three AH–MB complexes are associated with each other laterally to form a cluster, as seen in Fig. 7. The K^1 atom is bound to the three carbonyl O atoms of Asn28, Asn66 and Asn104 and to the three O^2 atoms of the Man^1 residues to form an octahedral coordination. Another K^2 atom is bound to the three O^4 atoms of the Man^1 residues. Therefore, each Man^1 is bridged between the two K atoms at the O^2 and O^4 atoms. These two potassium ions lie on another crystallographic threefold axis crossing the midpoint of the cubic edges, which differs from that passing the centre of every AH molecule. It appears that they facilitate crystallization of AH–MB in space group $P2_13$. Indeed, it was difficult to obtain crystals without potassium ions.

4. Discussion

The molecular structure of AH is composed of three structural modules associated with a pseudo-threefold symmetry based on tandem repeats in the sequence (Fig. 3). To detect the conformational rigidity of the three modules, their structures were compared by superimposition of the corresponding C^α atoms (Table 2). In the two apo-form AHs, the r.m.s.d.s between the three modules are as small as 0.25–0.34 Å, suggesting that the structural equivalency of the three modules is highly conserved. This trend is also observed in the complex form of AH, although the r.m.s. deviations are slightly higher. The structural equivalency between the

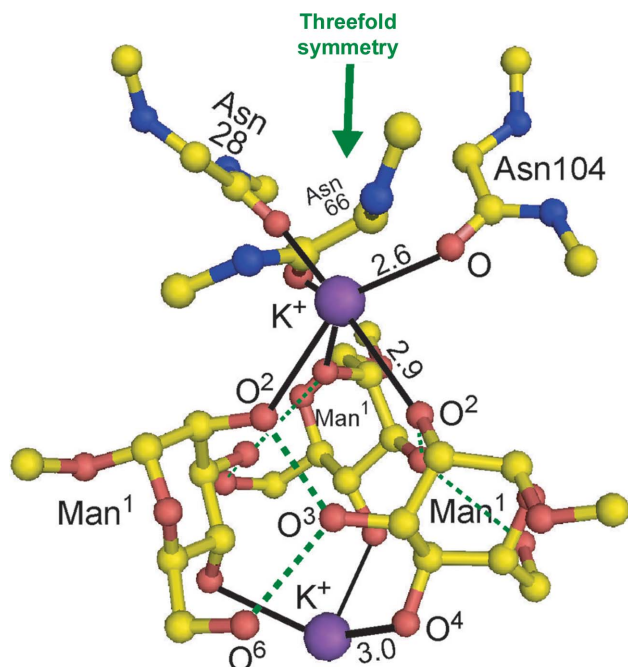


Figure 7
Two potassium ions which cluster the three AH molecules. One of the two potassium ions is surrounded by three Asn carbonyl O atoms and three O^2 atoms of Man^1 to form an octahedron. Another potassium ion is bound to three O^4 atoms of Man^1 , on the opposite side to which several water molecules are disordered. These ligands are related by the crystallographic threefold symmetry. Full black lines show coordination bonds around the two potassium cations (with their distances in Å) and dashed green lines indicate possible hydrogen bonds between the three Man^1 s: $O^3 \cdots HO^2$ and $O^3H \cdots O^6$ in each.

Table 2

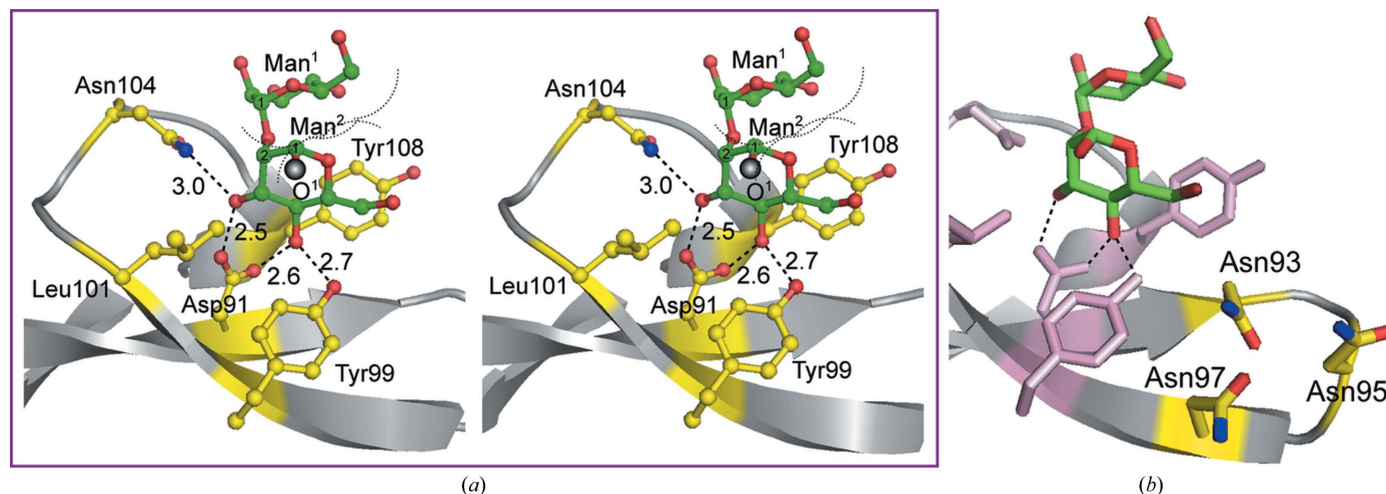
R.m.s.d.s (Å) between corresponding C^α atoms when they are superimposed on each other.

Module(*i,j*) indicates superimposition between modules *i* and *j*. ApoAH-1 and ApoAH-2 are the two independent AH molecules in the asymmetric unit of the apo-form crystal.

| | Module(1,2) | Module(1,3) | Module(2,3) |
|---------|-------------|-------------|-------------|
| ApoAH-1 | 0.30 | 0.29 | 0.29 |
| ApoAH-2 | 0.34 | 0.25 | 0.27 |
| AH–MB | 0.48 | 0.39 | 0.48 |

modules may support the high pseudo-symmetry of the molecule. In each module, the three β -strands form a convex slope on the lateral surface and a long loop forms the rim of a shallow pocket. Fig. 8 depicts an example of the geometry of the third pocket. In every pocket the first mannose residue (Man^1) of MB is located at the edge of the rim. The second mannose residue (Man^2) is accommodated by pushing one side (the C^3 , C^4 and C^5 atoms) of the hexose ring into the pocket and is trapped by four hydrogen bonds. Asp91 forms double hydrogen bonds to Man^2 , in which one of the two carboxyl O atoms accepts the hydroxyl group attached to the C^3 atom of Man^2 ($O \cdots O$ distance of 2.5 Å) and another O atom accepts the hydroxyl group at the C^4 atom of the same Man^2 ($O \cdots O$ distance of 2.7 Å). In addition, the O^3 atom of Man^2 accepts the carbonyl amide N atom of Asn104 ($N \cdots O$ distance of 2.9 Å) and the O^4 atom accepts the hydroxyl group of Tyr99 ($O \cdots O$ distance of 2.8 Å). It seems that the Tyr108 side chain wedges between the two Man residues to contact the C^5 and C^6 atoms through hydrophobic interaction. The methyl group of Leu101 faces the C^3 atom of Man^2 to block its movement, so that the hydroxyl group attached to the C^1 atom in the axial configuration protrudes into the outside in an upwards direction. These MB-binding features are the same as those in the other modules.

The three MB molecules, each of which adopts a bracket-shaped conformation stabilized through two weak $C-H \cdots O$ hydrogen bonds, are separately bound in the three pockets, which are similar to each other in accordance with the molecular pseudo-threefold symmetry generated from the three tandem repeats in the amino-acid sequence. The shape of the pocket for Man binding can accept the two neighbouring hydroxyl groups of the O^3 and O^4 atoms of the second Man residue, both of which are in an equatorial configuration. To recognize them, an Asp residue is located at the bottom of the pocket. The Tyr and Leu residues seem to block the movement of the Man^2 moiety. Furthermore, in each pocket, the O^1 atom of the Man^2 residue protrudes into an open space surrounded by conserved hydrophobic residues (Leu101 and Tyr99 in Fig. 8*a*). This is a suitable situation for interaction with the third Man residue of D1 (Fig. 2) when HMTGs are bound, because the hydrophobic regions of Man^3 could accommodate the binding through hydrophobic interactions. Practically, although AH itself exhibits low solubility in aqueous solution, when amphipathic molecules such as propanol, methylpentane-2,4-diol, MB *etc.* are added the solubility increases remarkably. This fact suggests that these exposed

**Figure 8**

A stereo diagram of MB bound in a pocket of the third module (a) and its right side (b), where three Asn residues are closely localized. MB adopts a bracket-shaped conformation stabilized by two C—H...O hydrogen bonds. The two hydroxyl groups in the equatorial configuration of the second mannose residue are recognized, each by two hydrogen bonds (the average distances are shown in Å). The O¹ atom is drawn as a grey sphere. Dashed and dotted lines indicate possible hydrogen bonds with their average distances and van der Waals spheres, respectively.

hydrophobic residues are involved in interactions with amphipathic MB.

AH exhibits high specificity for the $\alpha(1-2)$ -linked mannobiose moiety of HMTG. Here, it is interesting to examine how the AH pocket discriminates other isomers of mannobiose with different linkages using *in silico* structural modelling. In the case of 1-3 linkages, the two hexose rings cannot adopt a compact bracket-shape conformation: they are extended to have an L-shaped conformation so that it becomes difficult for the O⁴ (equatorial) and O⁵ atoms to enter the pocket. Similarly, other cases with 1-4, 1-6, 2-3, 2-4 and 2-6 linkages also extend the conformation. To maintain a bracket-shaped conformation, the two Man residues are required to be linked together between the hydroxyl groups in the axial configurations. Although such configurations can also possibly occur in 1-1 and 2-2 linkages, it is difficult to stabilize the compact conformation through the two C—H...O interactions.

Structural comparisons of AH with other microbicide lectins active towards gp120 provide an understanding of the high AH specificity. CV-N (Botos *et al.*, 2002) is a dimeric protein that is stabilized by domain swapping between the two subunits and possesses two separate pockets. Each pocket has a W-shaped convex in which the second and third Man residues, linked by an $\alpha(1-2)$ bond, are bound through several hydrogen bonds. Another lectin, GRFT (Moulaei *et al.*, 2010), that has a different tertiary structure is also a dimeric protein stabilized by N-terminal peptide swapping so that each domain has three binding pockets for one HMTG. These two lectins have neither three tandem repeats in their primary structure nor pseudo-threefold symmetry in their tertiary structure. In contrast, AH adopts a rather rigid structure constructed on a stable scaffold² folded by the molecular

pseudo-threefold symmetry, so that it does not change in overall conformation, and binds specifically to the $\alpha(1-2)$ -linked MB end of the D1 chain. In addition, it can bind three HMTGs of gp120 simultaneously as a trivalent microbicide, magnifying its affinity through a cluster effect in which AH binds three HMTGs (1:3 stoichiometry). This unique structure may bring about its highly specific HMTG binding. Furthermore, AH does not exhibit any mitogenic activity (Hoorelbeke *et al.*, 2010; Matoba *et al.*, 2010) in contrast to the other lectins, *e.g.* CN-V.

As shown in Fig. 2, HMTG protrudes from the surface of gp120 and branches into three chains: D1 (with 1-2 and 1-2 linkages), D2 (with 1-2 and 1-3 linkages) and D3 (with 1-2 and 1-6 linkages). Although their ends are commonly the same MB, D1 is preferentially bound to AH and D3 is weakly bound, as described previously (Tanaka *et al.*, 2009). In addition, AH binds three HMTGs (Tanaka *et al.*, 2009). The present X-ray structure of AH in complex with three MBs shows that the three pockets are equivalent major binding sites for the MB parts of three D1s. Between the three pockets of AH, it is noted that there are three open spaces (Fig. 8b) in which Asn/Asp residues are closely localized (Asn17, Asn19 and Asn21 in the first space, Asn55, Asp57 and Ala59 in the second space and Asn93, Asn95 and Asn97 in the third space). These hydrophilic residues could interact with other parts of the HMTGs through hydrogen bonding. By mutation experiments on AH (Takahashi *et al.*, 2010), it has been confirmed that these residues are responsible for its activity. Based on these structural features, an *in silico* structural model of three-HMTG-bound AH was constructed, as shown in Fig. 9, in which the ends of three D1 chains are accommodated in the major binding pockets and those of the D3 chains are in contact with the Asn/Asp residues in the open spaces. This binding feature is quite different from that of GRFT, in which the three branches D1, D2 and D3 of HMTG are separately bound in the three pockets. As HIVgp120 is covered by about

² The peptide folding is classified as CBM family 13 (Hazes, 1996; Gilkes *et al.*, 1991), which contains many proteins that differ in function but are similar to each other in tertiary structure with a stable scaffold.

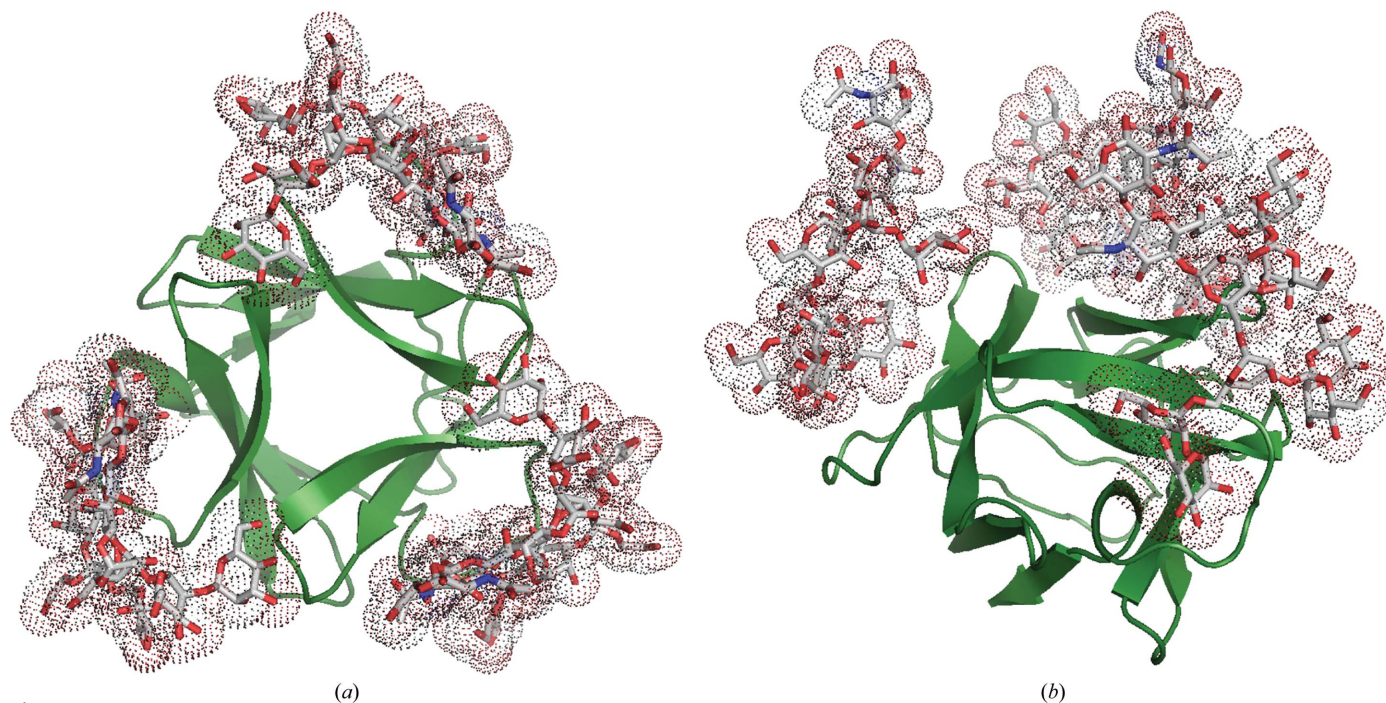


Figure 9
A model of AH bound to three HMTGs drawn in space-filling representation: (a) top view, (b) side view.

ten HMTGs, the three equivalent binding pockets of AH can bridge between HMTGs. This multivalent effect could magnify its potent specificity for gp120. Indeed, dimeric AH prepared by linking two AHs with a designed short peptide exhibits further magnification (Takahashi *et al.*, 2011). From the above-mentioned interpretations, it could be concluded that AH possesses extremely high specificity for MB binding.

Recently, it has been reported that infections involving cell-to-cell spread are markedly less sensitive to antiretroviral drugs and that cell-to-cell spread may provide a barrier to curing HIV infection (Sigal *et al.*, 2011). Because AH inhibits syncytium formation by Env-expressing HeLa cells and receptor (CD4 and CXCR4/CCR5) expressing cells (Matsumoto *et al.*, 2003; Chiba *et al.*, 2001), it is considered to be able to block the cell-to-cell spread of HIV. Therefore, AH would provide a candidate for investigation for the possibility of curing HIV infection. Together with the development of antiretroviral drugs, some types of microbicide lectins, including AH, should be developed as complementary drugs in order to compensate for disadvantages of the various types of drugs, which may induce multiple drug resistance and immunological problems. To overcome these difficulties, we need to continue intensive and extensive research to develop AH as a useful antiretroviral drug, as well as to examine drug combinations which could more effectively suppress the infectious expansion of HIV/AIDS and help to expedite an end to the HIV/AIDS pandemic in the near future.

We thank Y. Yamada, N. Matsugaki, N. Igarashi and S. Wakatsuki (Photon Factory, Tsukuba, Japan) for their assistance in data collection and with the synchrotron facility. This

work was supported in part by grants from the Tokyo Biochemical Research Foundation (A. Takénaka), the JAXA-GCF project (MT), the Science Research Promotion Fund, the Promotion and Mutual Aid Corporation for Private Schools of Japan (HT), the Japan Ministry of Education, Culture, Sports, Science and Technology (HT and SO for the 21st Century Program), the Japan Ministry of Health, Labour and Welfare Research: the Japan Health Science Foundation on Drug Innovation (HT) and the Iwaki City, Japan Strategic Industry Producing Project (HT).

References

- Balzarini, J. (2007). *Nature Rev. Microbiol.* **5**, 583–597.
 Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. W. (2011). *Acta Cryst.* **D67**, 271–281.
 Berger, E. A., Murphy, P. M. & Farber, J. M. (1999). *Annu. Rev. Immunol.* **17**, 657–700.
 Bewley, C. A., Kiyonaka, S. & Hamachi, I. (2002). *J. Mol. Biol.* **322**, 881–889.
 Bewley, C. A. & Otero-Quintero, S. (2001). *J. Am. Chem. Soc.* **123**, 3892–3902.
 Botos, I., O’Keefe, B. R., Shenoy, S. R., Cartner, L. K., Ratner, D. M., Seeberger, P. H., Boyd, M. R. & Wlodawer, A. (2002). *J. Biol. Chem.* **277**, 34336–34342.
 Brünger, A. T. (1992). *Nature (London)*, **355**, 472–475.
 Chiba, H., Asanuma, S., Okamoto, M., Inokoshi, J., Tanaka, H., Fujita, K. & Omura, S. (2001). *J. Antibiot.* **54**, 818–826.
 Chiba, H., Inokoshi, J., Nakashima, H., Omura, S. & Tanaka, H. (2004). *Biochem. Biophys. Res. Commun.* **316**, 203–210.
 Cohen, J. (2011). *Science*, **334**, 1628.
 Cohen, M. S. *et al.* (2011). *N. Engl. J. Med.* **365**, 493–505.
 DeLano, W. L. (2002). *PyMOL*. <http://www.pymol.org>.
 Desiraju, G. R. & Steiner, T. (2001). *The Weak Hydrogen Bond: In Structural Chemistry and Biology*. Oxford University Press.
 Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* **D60**, 2126–2132.

- Evans, P. (2006). *Acta Cryst.* **D62**, 72–82.
- Fromme, R., Katiliene, Z., Fromme, P. & Ghirlanda, G. (2008). *Protein Sci.* **17**, 939–944.
- Gilkes, N. R., Henrissat, B., Kilburn, D. G., Miller, R. C. & Warren, R. A. (1991). *Microbiol. Rev.* **55**, 303–315.
- Hazes, B. (1996). *Protein Sci.* **5**, 1490–1501.
- Hoorelbeke, B., Huskens, D., Férir, G., François, K. O., Takahashi, A., Van Laethem, K., Schols, D., Tanaka, H. & Balzarini, J. (2010). *Antimicrob. Agents Chemother.* **54**, 3287–3301.
- Jegede, O., Babu, J., Di Santo, R., McColl, D. J., Weber, J. & Quiñones-Mateu, M. (2008). *AIDS Rev.* **10**, 172–189.
- Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
- Lee, M., Passon, D. M., Hennig, S., Fox, A. H. & Bond, C. S. (2011). *Acta Cryst.* **D67**, 981–987.
- Leonard, C. K., Spellman, M. W., Riddle, L., Harris, R. J., Thomas, J. N. & Gregory, T. J. (1990). *J. Biol. Chem.* **265**, 10373–10382.
- Matoba, N., Husk, A. S., Barnett, B. W., Pickel, M. M., Arntzen, C. J., Montefiori, D. C., Takahashi, A., Tanno, K., Omura, S., Cao, H., Mooney, J. P., Hanson, C. V. & Tanaka, H. (2010). *PLoS One*, **5**, e11143.
- Matsumoto, A., Takahashi, Y., Shinose, M., Seino, A., Iwai, Y. & Omura, S. (2003). *Int. J. Syst. Evol. Microbiol.* **53**, 1553–1559.
- Moulaei, T., Shenoy, S. R., Giomarelli, B., Thomas, C., McMahon, J. B., Dauter, Z., O’Keefe, B. R. & Wlodawer, A. (2010). *Structure*, **18**, 1104–1115.
- Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* **D67**, 355–367.
- Navaza, J. (1994). *Acta Cryst.* **A50**, 157–163.
- Philippsen, A. (2003). *DINO*. <http://www.dino3d.org>.
- Sayle, R. A. & Milner-White, E. J. (1995). *Trends Biochem. Sci.* **20**, 374–376.
- Suzuki, K., Ohbayashi, N., Jiang, J., Zhang, X., Hoque, M. M., Tsunoda, M., Murayama, K., Tanaka, H. & Takenaka, A. (2012). *Acta Cryst.* **F68**, 1060–1063.
- Sigal, A., Kim, J. T., Balazs, A. B., Dekel, E., Mayo, A., Milo, R. & Baltimore, D. (2011). *Nature (London)*, **477**, 95–98.
- Takahashi, A., Inokoshi, J., Hachiya, A., Oka, S., Omura, S. & Tanaka, H. (2011). *J. Antibiot.* **64**, 551–557.
- Takahashi, A., Inokoshi, J., Tsunoda, M., Suzuki, K., Takenaka, A., Sekiguchi, T., Omura, S. & Tanaka, H. (2010). *J. Antibiot.* **63**, 661–665.
- Tanaka, H., Chiba, H., Inokoshi, J., Kuno, A., Sugai, T., Takahashi, A., Ito, Y., Tsunoda, M., Suzuki, K., Takénaka, A., Sekiguchi, T., Umeyama, H., Hirabayashi, J. & Omura, S. (2009). *Proc. Natl Acad. Sci. USA*, **106**, 15633–15638.
- Tsai, C. C., Emau, P., Jiang, Y., Agy, M. B., Shattock, R. J., Schmidt, A., Morton, W. R., Gustafson, K. R. & Boyd, M. R. (2004). *AIDS Res. Hum. Retroviruses*, **20**, 11–18.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.
- Ziółkowska, N. E., Shenoy, S. R., O’Keefe, B. R., McMahon, J. B., Palmer, K. E., Dwek, R. A., Wormald, M. R. & Wlodawer, A. (2007). *Proteins*, **67**, 661–670.