Identification, Optimization and Preliminary X-ray Diffraction of a New Crystal Form of the N-terminal Domain of the HIV-1 CA Protein

Ryan Van Woerkom[1], Andrew Dixon[1], Rob Oslund and Bruce R. Howard*
Department of Physical Science
Southern Utah University
Cedar City, Utah 84720 USA

Received: August 25, 2006 Accepted: January 18, 2007

ABSTRACT

In an effort to better understand the detailed intersubunit interactions of the N-terminal Domain of the CA (capsid) protein from HIV-1 within the conical core of the mature virus, we have identified a novel crystal form of this domain and have optimized conditions to grow single protein crystals suitable for x-ray analysis. These high quality crystals diffract to better than 1.8 Å resolution on a rotating anode generator.

I. INTRODUCTION

The CA protein of HIV-1 is initially produced as part of a larger GAG polyprotein before incorporation into newly-constructed virions. The GAG polyprotein is localized to the inner phospholipid surface of the budding virus particle by interactions of the MA (matrix) domain via a positively-charged surface as well as a lipid anchor. The packaged GAG molecules are arranged centrosymmetrically in the immature virus with the N-terminal matrix domains at the periphery and the C-termini toward the center. Upon budding, in a process termed maturation, the viral protease is activated, cutting the GAG polyprotein at specific cleavage sites to release the individual structural proteins including CA. The liberated CA proteins then rearrange, assembling together to form the capsid of the mature virus which encloses the nucleocapsid-bound viral RNA, and additional viral factors (for good reviews see [1] and [2]).

Of the approximately 5000 GAG molecules (and therefore processed CA proteins) packaged into the virus particle [3], between 1,000 and 1,500 of these assemble together to form the mature capsid which is based on a fullerene cone [4-6]. Cryoelectron microscopic reconstructions of helical assemblies of the CA protein show a hexagonal lattice [5]. The two domains of the CA protein [7,8] perform different roles in forming the capsid. The N-terminal domain is involved in hexameric interactions (and 12 pentameric defects which allow a closed surface) [4, 5], while the C-terminal domain forms dimers which hold the hexamers together in the mature capsid structure [5].

X-ray crystallography is a powerful technique used to solve the three-dimensional structures of molecules in a crystal. One requirement for a successful structure determination is a well-ordered, single crystal. Crystals that have a higher level of order within the crystal lattice diffract x-rays to higher scattering angles. The diffraction data at higher angles corresponds to more closely-spaced planes of electron density throughout the crystal and therefore represents higher resolution information. This higher resolution information provides a

* To whom correspondence should be addressed. E-mail howard@suu.edu
[1] These authors contributed equally to this work

Figure 1. Initial crystals of the HIV-1 CA A92E N-terminal domain in this new crystal form. This condition resulted in clusters of plates.

more precise location for each of the ordered atoms in the crystal lattice. Therefore the higher the angle of diffraction from a crystal, the higher the resolution of the collected data set, and the better-resolved the resulting atomic structure will be. For a more detailed description of x-ray crystallography and the techniques used to solve macromolecular crystal structures see references [9] and [10].

A high-resolution crystal structure is available showing the dimerization interface between the C-terminal domains of the HIV-1 CA protein [8], as well as a high-resolution structure of this domain in complex with an assembly inhibitor [11]. In contrast, the detailed interactions between the N-terminal domains that form the hexamers and pentamers remain elusive. A crystal structure of a hexameric assembly of the N-terminal domain of the capsid protein from Murine Leukemia Virus has been solved [12]. Although this provides a good overall model for a retroviral capsid, the detailed subunit interactions specific to the HIV-1 CA hexamer are still unclear.

It has been shown that compounds which bind this N-terminal domain and inhibit assembly of CA into the mature capsid also inhibit viral infectivity [13]. An intimate knowledge of the details of the subunit contacts will be useful in designing and optimizing assembly inhibitors that target this essential binding interaction. Such inhibitors may provide another avenue for treatment of HIV-infected individuals. While a crystal structure of the isolated N-terminal domain of CA from HIV-1 has recently been determined, neither a hexameric or pentameric assembly, or other relevant interfaces were observed [14]. We have therefore continued to search for a different crystal form which will provide a high-resolution view of these intersubunit interactions within the mature hexamer. As a result, we have identified a new crystal form of this domain which diffracts x-rays to high resolution, and is amenable to structural study.

II. MATERIALS, METHODS AND RESULTS

a. Protein Source

Purified HIV-1 CA N-terminal domain (1-146) A92E mutant protein was generously provided by Brian N. Kelly and Christopher P. Hill at the University of Utah School of Medicine, and was prepared as described [14]. We used this protein for our crystallization trials.

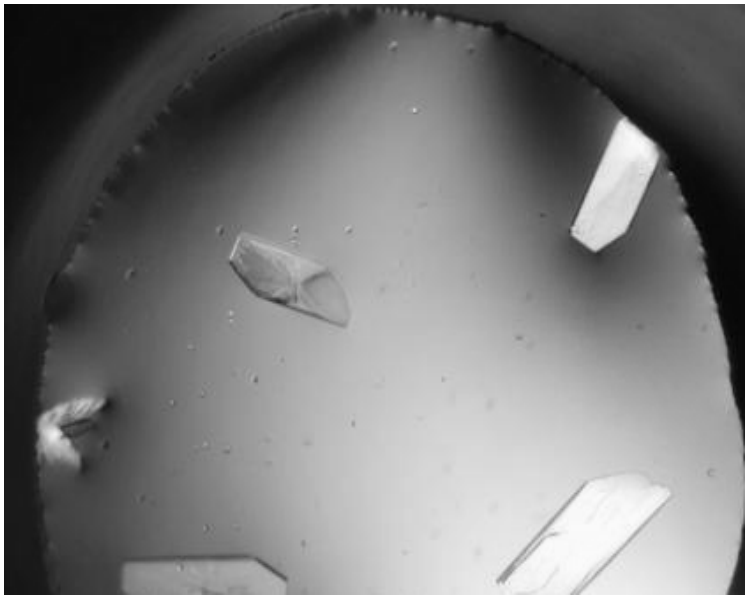

Figure 2. Single thin plates formed upon further optimization of the buffer pH, salt and precipitant concentrations.

b. Crystallization

Sparse-matrix crystal screens were purchased from Hampton Research for initial crystallization experiments [15]. We used the Crystal Screen, and the Crystal Screen Cryo kits. Each of these kits contains 10 ml samples of 50 unique solutions containing various salts, buffers and precipitants. The Crystal Screen Cryo kit includes in addition some sort of cryoprotectant in each of its 50 different conditions. These 100 buffer solutions were used as "well solutions" in our crystallization trays. The process of vapor diffusion was used to cause a controlled precipitation of the protein. The protein solution contained $CA_{1\text{-}146}$ (A92E) protein at 15 mg/ml, 10 mM Tris-HCl pH 8.0, and 2mM β-mercaptoethanol. 2 µl of this protein solution was mixed with an equal amount of each well solution and allowed to equilibrate through vapor diffusion with a 500 µl reservoir of that same well buffer. Because of the 1:1 mixing with well solution, the protein in the drop was initially 7.5 mg/ml – half the concentration of the stock solution. For the same reason, the precipitant from the well solution was also diluted to half the original concentration. Under these conditions (ideally), the protein would remain soluble. As the drop equilibrated with a much larger amount (~125 x) of well solution with a more concentrated precipitant, the drop would gradually shrink as the water diffused into the well. This causes all the solutes in the drop to gradually increase in concentration until precipitation of the protein occurs.

We screened these 100 conditions visually using a stereomicroscope after an incubation period of one week at room temperature. While many conditions resulted in precipitate, only a few had anything that looked remotely crystalline. After looking through the solution ingredients for these promising leads, we rejected anything containing citrate. Citrate was present in the previous crystallization conditions for this protein and was found to be involved in crystal contacts with well ordered citrate molecules holding the proteins together in the lattice [14]. Since we wanted a new crystal form, we avoided these conditions. Setting up trays around each of the remaining most promising conditions resulted in a few hopeful leads.

The condition that resulted in the growth of the current diffraction-quality crystals was not originally identified during our first inspection of the trays. After a second review (approximately one month

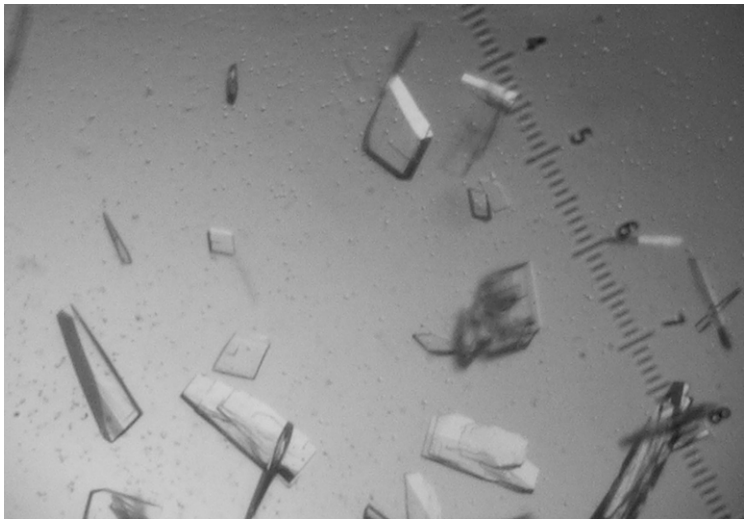

Figure 3. The final stage of optimization that resulted in high-quality well-ordered single crystals used for X-ray diffraction experiments. Varying the ratio of protein solution to well solution enabled the growth of single crystals with more thickness and volume (each of the smallest divisions on the scale corresponds to ~ 10 µm).

after the initial setup) with the stereomicroscope, this condition that was previously clear, without precipitation, contained clusters of roughly diamond-shaped crystals (Figure 1). The well solution for this condition contained 200 mM $MgCl_2$, 100 mM Tris-HCl pH 8.5, and 30% w/v polyethyleneglycol (PEG) 4,000. The protein solution contained $CA_{1\text{-}146}$ (A92E) protein at 15 mg/ml, 10 mM Tris-HCl pH 8.0, and 2mM β-mercaptoethanol. After optimization of the conditions using a slightly larger PEG (4.5 kDa), we were able to obtain larger single crystals (Figure 2). Nice crystals grew somewhat regularly in conditions that varied between 22 and 27% PEG 4,500 and a $MgCl_2$ concentration between 550 mM and 700 mM. A pH of 8.5 turned out to be optimal for the Tris-HCl buffer which was kept at 100 mM. The formation of thicker crystals was obtained by increasing the ratio of the volume of protein solution to that of the well solution in each drop. We found that a 2:1 ratio worked nicely, resulting in single crystal blocks in approximately 10% of the setups (Figure 3). In order to overcome this low frequency of drops containing nice single crystals, we set up a dozen identical, 24-well trays in anticipation of visiting an X-ray source to screen the crystals for diffraction quality.

c. Cryoprotection and X-ray Diffraction

Since the only X-ray diffraction equipment we currently have at SUU is designed for powdered mineral samples and not useful for single-crystal analysis, we traveled to the University of Utah in Salt Lake City to use their X-ray equipment to screen the crystals. In order to preserve the quality of the crystals from the damaging effects of the focused X-ray beam, we screened using cryoprotectants that would allow us to flash-cool the crystals in liquid nitrogen. After several different trials, we found that supplementation of the well solution with 10% glycerol allowed the crystals to remain intact, and also prevented the formation of ice crystals which could expand and destroy the fragile protein crystals upon freezing. Crystals were transferred briefly to this cryoprotectant solution, and after several seconds were suspended in a nylon loop and plunged into liquid nitrogen to cool them rapidly. Crystals were then mounted on a Rigaku rotating anode X-ray generator equipped with mirror boxes and an Oxford cryostream to keep the crystals at ~100 K during diffraction experiments. The diffraction data were collected on an Raxis4 image plate detector.

The crystals diffracted beautifully, and an example of an oscillation image is shown

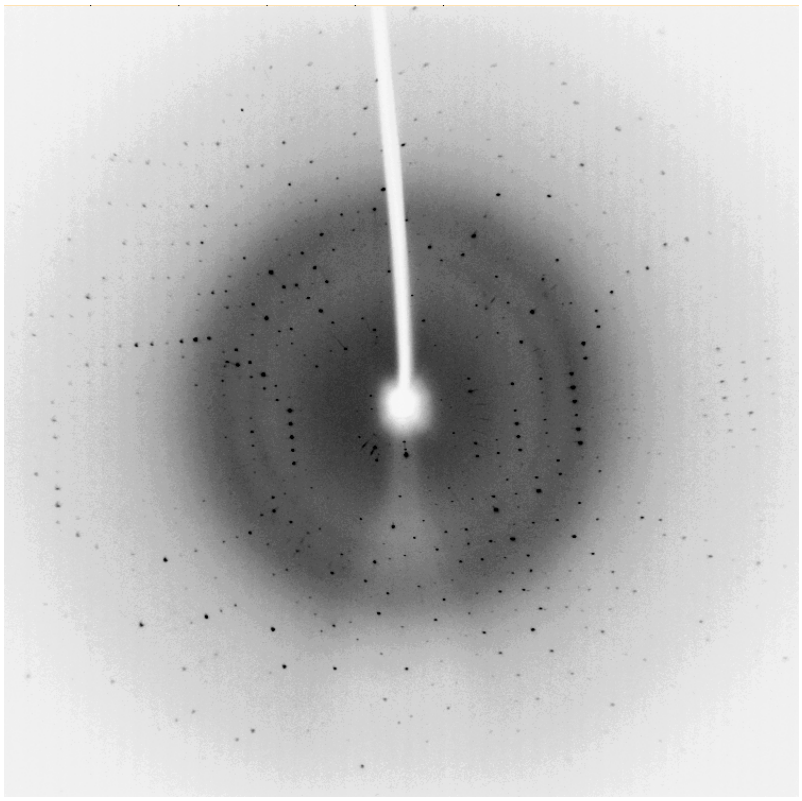

Figure 4. X-ray diffraction from one of the thicker plates measuring approximately 100 x 50 x 15 μm^3. The oscillation range is one degree, with an exposure time of one hour. The edge of the image plate represents a diffraction angle corresponding to ~ 2.0 Å resolution. Spots extend into the corners of the image, with several spots at ~1.8 Å.

in figure 4. The protein crystallizes in the space group P1 (triclinic) with cell dimensions of a = 48.42 Å, b = 59.11 Å, c = 92.50 Å; α = 71.42°, β = 87.40°, γ = 83.09°. There is a pseudosymmetry apparent along one axis (the 92.50 Å cell edge) of the crystal, causing every other layer of spots in the data to be much weaker than the intervening layers. Based on the volume of the unit cell (2.491 x 10^5 $Å^3$) and the size of the N-terminal domain (16,165.5 Da) we can calculate the probable number of molecules in the unit cell using the Matthews coefficient (V_M) [16]. Based on the 15,641 protein crystal structures deposited in the Protein Data Bank [17] by 2002, the most likely number of molecules in the unit cell is six which would correspond to a hexamer and give a V_M of 2.57 $Å^3$/Da [18, 19]. But, based on the 1.8 Å or better diffraction resolution of our crystals, the most likely number of molecules in the unit cell is seven which would give a V_M of 2.20, although the actual number could with a reasonable probability be five, six, seven or eight [18, 19].

III. CONCLUSION

In conclusion, we have identified a new crystal form of the N-terminal domain of HIV-1 CA protein which diffracts to high resolution and is amenable to a detailed x-ray crystallographic analysis. This should result in a high quality structure showing the arrangement of this domain in relation to that of its neighbors in a new environment. Ideally, this will show us the intersubunit contacts within the hexamers or pentamers which make up the mature capsid structure of infectious virus particles. In any case, it will likely yield a high-resolution structure which can be analyzed to learn more about the interactions between these domains, and will hopefully yield additional insight into the assembly process of the HIV-1 capsid.

ACKNOWLEDGEMENTS

We thank the Provost Faculty Development Grant program at Southern Utah University for generous support of this project. We also thank Dr. Christopher P. Hill, Dr. Wesley I. Sundquist, and Brian N. Kelly for supplying the protein for these experiments, and for the use of the x-ray generator and other equipment in the biochemistry department at the University of Utah School of Medicine in Salt Lake City.

REFERENCES

1. Frankel, A. D., and Young, J. A. (1998) HIV-1: Fifteen proteins and an RNA, *Annu Rev Biochem* 67, 1-25.
2. Göttlinger HG. (2001) The HIV-1 assembly machine. *AIDS* 15:S13-S20.
3. Briggs, J. A., Simon, M. N., Gross, I., Krausslich, H. G., Fuller, S. D., Vogt, V. M., Johnson, M. C. (2004) The stoichiometry of Gag protein in HIV-1 *Nat Struct Mol Biol.* 11(7):672-5.
4. Ganser, B. K., Li, S., Klishko, V. Y., Finch, J. T., and Sundquist, W. I. (1999) Assembly and analysis of conical models for the HIV-1 core, *Science* 283, 80-3.
5. Li, S., Hill, C. P., Sundquist, W. I., and Finch, J. T. (2000) Image reconstructions of helical assemblies of the HIV-1 CA protein, *Nature* 407, 409-13.
6. Briggs, J. A., Wilk, T., Welker, R., Krausslich, H. G., and Fuller, S. D. (2003) Structural organization of authentic, mature HIV-1 virions and cores, *Embo J* 22, 1707-15.
7. Gitti, R. K., Lee, B. M., Walker, J., Summers, M. F., Yoo, S., and Sundquist, W. I. (1996) Structure of the amino-terminal core domain of the HIV-1 capsid protein, *Science* 273, 231-5.
8. Gamble, T. R., Yoo, S., Vajdos, F. F., von Schwedler, U. K., Worthylake, D. K., Wang, H., McCutcheon, J. P., Sundquist, W. I., and Hill, C. P. (1997) Structure of the carboxyl-terminal dimerization domain of the HIV-1 capsid protein, *Science* 278, 849-53.
9. Rhodes, G. (2000) Crystallography Made Crystal Clear, 2nd Ed., Academic Press 24-28 Oval Road, London NW17X, UK.
10. Stout, G. H. and Jensen, L. H. (1989) X-ray Structure Determination, A Practical Guide, John Wiley & Sons, Inc.
11. Ternois F, Sticht J, Duquerroy S, Krausslich HG, Rey FA. (2005) The HIV-1 capsid protein C-terminal domain in complex with a virus assembly inhibitor, *Nat Struct Mol Biol.* 12(8):678-82.
12. Mortuza, G. B., Haire, L. F., Stevens, A., Smerdon, S. J., Stoye, J. P., and Taylor, I. A. (2004) High-resolution structure of a retroviral capsid hexameric amino-terminal domain, *Nature* 431, 481-5.
13. Tang C, Loeliger E, Kinde I, Kyere S, Mayo K, Barklis E, Sun Y, Huang M, Summers M.F. (2003) Antiviral inhibition of the HIV-1 capsid protein, *J Mol Biol*. 327(5):1013-20.
14. Kelly, B. N., Howard, B. R., Wang, H. E., Robinson, H., Sundquist, W. I., and Hill, C. P. (2006) Implications for viral capsid assembly from crystal structures of HIV-1 Gag_{1-278} and $CA^{N}_{133-278}$, *Biochemistry* 45(38), 11257-11266.
15. Hampton Research, 34 Journey, Aliso Viejo, CA 92656-3317. http://www.hamptonresearch.com/
16. Matthews, B.W. (1968) Solvent content of protein crystals, *J. Mol. Biol.* 33: 491–497
17. The Protein Data Bank, http://www.rcsb.org/pdb
18. Kantardjieff, K. A. and Rupp, B. (2003) Matthews coefficient probabilities: Improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals, *Protein Sci.* 12(9):1865-71.
19. Matthews Probability Calculator at http://www.ruppweb.org/Mattprob/