Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Crystallization and preliminary X-ray diffraction analysis of vaccinia virus H1L phosphatase

Laura Roces,[a,b] Phillip P. Knowles,[b] Gavin Fox,[c] Jordi Juanhuix,[c]‡ Nicki Scaplehorn,[d] Michael Way[d] and Neil Q. McDonald[b,e]*

[a]Departamento de Química Física y Analítica, Universidad de Oviedo, Julián Clavería 8, 33006 Oviedo, Spain, [b]Structural Biology Laboratory, London Research Institute, Cancer Research UK, 44 Lincoln's Inn Fields, London WC2A 3PX, England, [c]ESRF, 6 Rue Jules Horowitz, 38043 Grenoble CEDEX 9, France, [d]Cell Motility Laboratory, London Research Institute, Cancer Research UK, 44 Lincoln's Inn Fields, London WC2A 3PX, England, and [e]School of Crystallography, Birkbeck College, Malet Street, London WC1E 7HX, England

‡ Current address: ALBA Synchrotron, Experiments Division Ed. Cn, Campus UAB, 08193 Barcelona, Spain.

Correspondence e-mail: neil.mcdonald@cancer.org.uk

Received 12 December 2007
Accepted 2 February 2008

The cysteine-based protein phosphatase H1L was the first reported dual-specificity protein phosphatase. H1L is encapsidated within the vaccinia virus and is required for successful host infection and for the production of viable vaccinia progeny. H1L has therefore been proposed as a target candidate for antiviral compounds. Recombinant H1L has been expressed in a catalytically inactive form using an *Escherichia coli* host, leading to purification and crystallization by the microbatch method. The crystals diffract to 2.1 Å resolution using synchrotron radiation. These crystals belong to space group *P*422, with unit-cell parameters $a = b = 98.31$, $c = 169.15$ Å, and are likely to contain four molecules in the asymmetric unit. A sulfur SAD data set was collected to 2.8 Å resolution on beamline BM14 at the ESRF to facilitate structure determination. Attempts to derivatize these crystals with xenon gas changed the space group to *I*422, with unit-cell parameters $a = b = 63.28$, $c = 169.68$ Å and a single molecule in the asymmetric unit. The relationship between these two crystal forms is discussed.

1. Introduction

Dual-specificity phosphatases (DUSPs) are a subset of protein tyrosine phosphatases (PTPs) that are capable of dephosphorylating threonine and tyrosine residues either in different substrates or within the same substrate (Fauman & Saper, 1996; Zhang, 1998). Like the PTP superfamily to which they belong, they have a canonical $C(X)_5R$ motif in their active sites bearing the nucleophilic cysteine (Zhou *et al.*, 1994; Kolmodin & Åqvist, 2001). The first identified DUSP was the vaccinia virus protein H1L, which was identified more than 15 years ago; however, its substrates remain poorly characterized (Guan *et al.*, 1991). Furthermore, the evidence that it is a DUSP is based on *in vitro* data (Guan *et al.*, 1991), which may not reflect its physiologically relevant substrates. H1L is encapsidated in the vaccinia virus particle and is required for the viral genome to be transcribed by the viral transcriptase (Liu *et al.*, 1995; Condit *et al.*, 2006). In fact, H1L is not required for transcription per se but is necessary for the assembly of transcriptionally competent particles and is thus required for virion infectivity. It is therefore well validated as a poxvirus drug target (Harrison *et al.*, 2004). The presence of H1L within the cytoplasm of infected cells has two additional known consequences. Firstly, H1L blocks the IFN-γ signalling pathway by inhibiting the phosphorylation and nuclear translocation of Stat1 (Najarro *et al.*, 2001). Secondly, H1L appears to mediate the destruction of the centrosome through an unknown mechanism (Scaplehorn, 2004). As a first step towards understanding the substrate specificity of the H1L DUSP, we have undertaken a structural and functional analysis. Here, we report the production of two crystal forms of recombinant H1L-(C110S) mutant, one of which will lead to a high-resolution structure determination and facilitate structures of complexes with H1L substrates.

2. Materials and methods

2.1. Cloning and expression

The expression vector pMW172 containing the full-length cDNA sequence of vaccinia virus H1L harbouring a cysteine 110 to serine (C110S) mutant was generated by subcloning the mutant H1L gene at the unique *Eco*RI and *Nde*I restriction sites (Way *et al.*, 1990). H1L-(C110S) protein was expressed in the host strain *Escherichia coli* BL21(DE3)RecA$^-$ pLysS using LB media. Cells were grown at 310 K until a cell density of OD_{600} of 0.6 was reached, at which point protein expression was induced by the addition of 50 µ*M* isopropyl *β*-D-1-thiogalactopyranoside (ITPG) for an additional 3 h at 310 K. For the preparation of soluble protein fractions, the cells from a 1 l culture were first pelleted and then resuspended in 20 ml cold lysis buffer containing 50 m*M* Tris pH 8.0, 500 m*M* NaCl, 10 m*M* benzamidine, 5 m*M* 2-mercaptoethanol, 1 m*M* phenylmethylsulfonyl fluoride (PMSF). Cell pellets were then lysed by sonication on ice. The clear supernatant containing soluble protein was collected by centrifugation. All the following purification steps were performed at 277 K.

2.2. Purification

The vector-derived N-terminal hexahistidine tag (His_6) was used to purify recombinant H1L-(C110S) using Ni–NTA IMAC resin (Novagen). The Ni–NTA affinity matrix loaded with H1L-containing cell lysate was washed extensively with 20 m*M* Tris pH 8.0, 500 m*M* NaCl, 0.5 m*M* 2-mercaptoethanol and 20 m*M* imidazole. Recombinant His_6-H1L-(C110S) was eluted from the affinity column using 100 m*M* imidazole along with 20 m*M* Tris pH 8.0, 500 m*M* NaCl, 0.5 m*M* 2-mercaptoethanol. In order to remove the imidazole, the protein was extensively dialysed against 20 m*M* Tris pH 8.0, 250 m*M* NaCl, 1 m*M* DTT. Dialysed protein was concentrated to 9–10 mg ml^{-1} using a 30 kDa molecular-weight cutoff Centricon (Vivaspin). The yield of recombinant protein was approximately 25 mg per litre of cell culture. The purity of H1L-(C110S) was judged by SDS–PAGE to be close to 95% and its identify was confirmed by mass spectrometry.

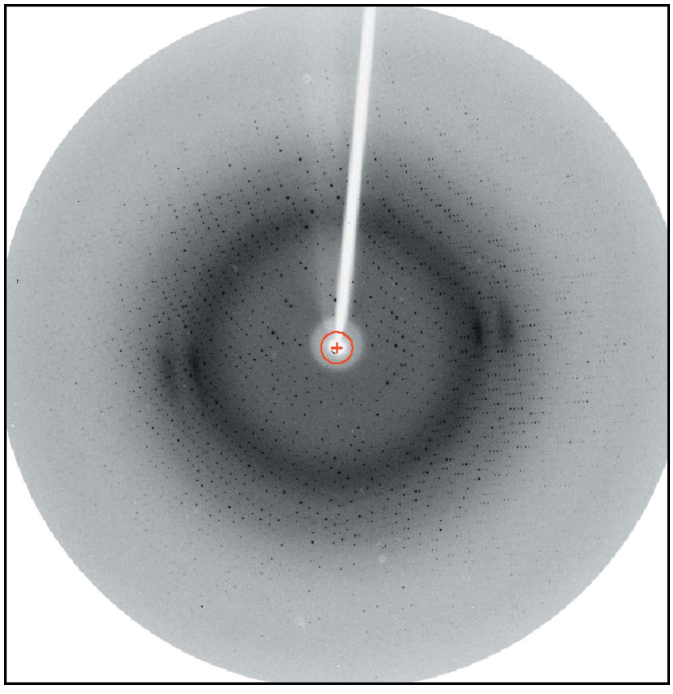

Figure 1
Oscillation X-ray image taken from a crystal of H1L-(C110S) showing diffraction to a Bragg spacing of 2.1 Å.

Table 1
Data-collection statistics.

Values in parentheses are for the highest resolution shell.

Data set	I. Native	II. Sulfur	III. Xenon derivative
Wavelength (Å)	0.976	1.851	1.851
Space group	*P*422	*P*422	*I*422
Unit-cell parameters (Å)	$a = b = 98.31$, $c = 169.15$	$a = b = 98.35$, $c = 169.07$	$a = b = 63.28$, $c = 169.68$
Resolution (Å)	22.96–2.10 (2.21–2.10)	30.00–2.80 (2.95–2.80)	30.00–3.00 (3.16–3.00)
No. of measured reflections	458423	516889	20201
No. of unique reflections	48131	20605	3636
Completeness (%)	98.2 (96.7)	97.5 (86.2)	98.3 (96.7)
Multiplicity	9.5 (9.3)	25.1 (3.6)	5.6 (3.4)
Mean $I/\sigma(I)$	19.2 (5.5)	40.0 (5.8)	18.0 (2.6)
R_{meas} (%)	7.9 (36.6)	7.0 (25.4)	7.1 (41.3)
$R_{p.i.m.}$ (%)	2.4 (10.1)	1.7 (15.3)	3.7 (26.6)
R_{anom} (%)		2.2 (14.2)	4.3 (25.1)

2.3. Crystallization

Initial crystallization experiments for H1L-(C110S) screened a variety of standard commercially available screens using the microbatch method. Briefly, 1 µl protein solution and 1 µl crystallization reagent were pipetted under a layer of paraffin oil (Hampton Research) into a hydrophilic Microbatch Plate (Hampton Research). Diffraction-quality square plate-shaped crystals were observed at 291 K in condition No. 41 (0.1 *M* Na HEPES pH 7.5, 10% 2-propanol, 20% polyethylene glycol 4000) of Crystal Screen I (Hampton Research). Diffraction-quality crystals were finally obtained by mixing 2 µl protein solution at 10 mg ml^{-1} and 2 µl crystallization reagent (Crystal Screen I condition No. 41) under a layer of Al's oil (Hampton Research) using a hydrophobic Vapour-Batch Plate (Hampton Research). After 5–7 days, when the first crystals were observed, a layer of paraffin oil (Hampton Research) was dispensed on the top of the plate to avoid further evaporation of the solvent and to allow full-size crystals to grow.

2.4. Data collection

A 1:1(*v*:*v*) mixture of Paratone-N (Hampton Research) and mineral oil (Hampton Research) was used as cryoprotectant. The crystals were transferred in a small amount of mother liquor (∼0.5 µl) into a drop of oil (∼100 µl) with a pipette. The crystal in the hydrocarbon oil was loop-mounted and transferred as quickly as possible into a stream of nitrogen gas maintained at 100 K. Data were collected from three different crystals at beamline BM14, European Synchrotron Radiation Facility (ESRF), Grenoble (Table 1). The diffraction data were recorded on a MAR CCD detector with a diameter of 135 mm. Data were collected according to the *STRATEGY* option within the *MOSFLM* program (Leslie, 1992). Data were processed and scaled using *MOSFLM* (Leslie, 1992), *TRUNCATE* and *SCALA* (Evans, 2006) from the *CCP*4 program suite (Collaborative Computational Project, Number 4, 1994). Data set II was collected to optimize the sulfur anomalous signal by tuning the wavelength to 1.851 Å and collecting a highly redundant data set from 360° of data close to the sulfur *K* edge. Data set III was prepared as a xenon derivative. Briefly, a single native H1L-(C110S) crystal mounted on a cryo-loop was transferred to a Xcell xenon pressure chamber (Oxford Cryosystems) and pressurized under 4 MPa xenon atmosphere for 1 min. The pressure was released slowly over 30 s and

the crystal was immediately flash-cooled in a stream of nitrogen gas maintained at 100 K.

3. Results and discussion

Our initial expression trials for recombinant wild-type H1L protein frequently resulted in low yields of purified protein (data not shown). To circumvent this problem, we assumed that the catalytic activity of H1L was interfering with protein expression and thus we produced a catalytically inactive form of H1L in which the nucleophilic cysteine (Cys110) was changed to serine. This gave greater than tenfold higher yields of purified recombinant protein and offered the potential to generate crystals of enzyme–substrate complexes. Crystals of recombinant H1L-(C110S) mutant with good diffraction properties were obtained after about 5–7 days and grew to approximate dimensions of 0.3 × 0.3 × 0.1 mm after 10 days. A complete data set was collected to 2.1 Å resolution at ESRF (Grenoble, France) on beamline BM14 (Fig. 1). Data-processing statistics are reported in Table 1. Careful inspection of the data indicated no systematic absences along 00*l* and established these crystals as belonging to the tetragonal space group *P*422. From solvent considerations, the asymmetric unit is most likely to contain four molecules; the corresponding crystal volume per unit molecular weight, $V_{\rm M}$, was calculated to be 2.48 Å^3 Da^{-1}, with a solvent content of 50.4% (Matthews, 1968). The self-rotation function showed no evidence of noncrystallographic symmetry (NCS). However, inspection of a native Patterson function showed a 23σ non-origin peak at (0.50, 0.026, 0.50), indicating a translational vector of 0.5 along the *c* axis.

To obtain experimental phases, we considered using the S atoms present in the native crystals to conduct a sulfur SAD experiment. The theoretical anomalous signal ($\langle\Delta F\rangle/\langle F\rangle$) for native H1L is estimated to be 1.4% for 40 S atoms based on the assumption of four molecules in the asymmetric unit (Hendrickson & Ogata, 1997). We collected a highly redundant data set by tuning close to the sulfur *K* edge at 1.851 Å on beamline BM14 at the ESRF (data set II, Table 1). With the high redundancy of the data collected, the error or noise in the observed intensities can be assessed by the precision-indication merging *R* factor ($R_{\rm p.i.m.}$). $R_{\rm p.i.m}$ is smaller than the observed anomalous signal ($R_{\rm anom}$) over the low-resolution shells, suggesting that a possible anomalous signal is present.

We also prepared a xenon derivative as a backup strategy for phasing. The usage of xenon has advantages over conventional heavy-atom derivative soaks as Xe-derivatized crystals are usually isomorphous with the native owing to its ability to bind by dispersion forces to hydrophobic pockets within the protein. Furthermore, the xenon anomalous signal is appreciable even at energies remote from its *K* (0.36 Å) or *L* (2.27–2.59 Å) absorption edges. We thus collected a data set from the H1L xenon derivative by tuning close to its *L* edge at 1.85 Å (data set III, Table 1). Surprisingly, we found that exposure to xenon gas caused the crystals to change to a tetragonal body-centred Bravais lattice, with space group *I*422 and unit-cell parameters $a = b = 63.28$, $c = 169.68$ Å. Analysis of the Matthews coefficient revealed that the smaller unit cell is likely to contain one molecule per asymmetric unit. The crystal volume per unit molecular weight, $V_{\rm M}$, was calculated to be 2.07 Å^3 Da^{-1}, corresponding to a solvent content of 40.66%, which is markedly lower than the native crystal. As we had not previously identified crystals grown from native protein in space group *I*422, we presumed that either the pressurization within the chamber and/or the inclusion of xenon gas in the native crystals modified the crystal packing, leading to the space-group change.

We have used the X-ray data described here to solve the structure of the *P*422 crystal form of H1L, thus confirming the space-group assignment and the existence of four copies of H1L within the asymmetric unit (Roces *et al.*, manuscript in preparation). The H1L molecules pair as two dimers that associate as a tetramer with approximate 222 symmetry. Each of the tetramer dyad axes lie approximately parallel to crystallographic axes. The unit cells of the different crystal forms are related by placing the *ab* diagonal of the *I*422 cell along the *a* or *b* unit-cell axis of the larger *P*422 cell. When combined with crystallographic symmetry, the NCS twofold axes approximately perpendicular and parallel to *c* generate the translational vector seen in the native Patterson function. This also creates an NCS 2_1 screw axis almost parallel to *c*. When this 2_1 screw axis becomes exactly parallel, this converts the space group from a *P*422 lattice to *I*422 and reduces the size of the asymmetric unit to a single H1L protomer.

Whilst preparing this crystallization report, the structure of the related *Variola major* H1 phosphatase was published (Phan *et al.*, 2007; PDB code 2p4d). At the same time, a structural genomics consortium deposited the coordinates for a wild-type vaccinia virus H1L protein (PDB code 2q05), albeit in a different crystal form to those described here for the C110S H1L mutant. The crystals of the wild-type vaccinia virus H1L diffracted to a lower resolution (2.6 Å) than the *P*422 crystal form. Each crystal form has four copies of H1L that pair as equivalent dimers which then associate into quite different tetrameric arrangements. A similar dimer is observed within the variola virus H1L structure, implicating the dimer to be the biologically relevant species (Phan *et al.*, 2007). To date, no information is available as to how host-cell or viral phosphoprotein substrates for H1L phosphatases are recognized. The crystals described here will allow co-complexes with phosphopeptides to be prepared to reveal the basis for H1L substrate discrimination.

We would like to thank the Spanish Ministerio de Educación y Ciencia (MAT2006-01997 and Factoría de Cristalización – Consolider Ingenio 2010) and Gobierno del Principado de Asturias (PCTI 2006-2009) and Cancer Research UK for financial support.

References

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Condit, R. C., Moussatche, N. & Traktman, P. (2006). *Adv. Vir. Res.* **66**, 31–124.
Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.
Fauman, E. B. & Saper, M. A. (1996). *Trends Biochem. Sci.* **21**, 413–417.
Guan, K. L., Broyles, S. S. & Dixon, J. E. (1991). *Nature (London)*, **350**, 359–362.
Harrison, S. C., Alberts, B., Ehrenfeld, E., Enquist, L., Fineberg, H., McKnight, S. L., Moss, B., O'Donnell, M., Ploegh, H., Schmid, S. L., Walter, K. P. & Theriot, J. (2004). *Proc. Natl Acad. Sci. USA*, **101**, 11178–11192.
Hendrickson, W. A. & Ogata, C. M. (1997). *Methods Enzymol.* **276**, 494–523.
Kolmodin, K. & Åqvist, J. (2001). *FEBS Lett.* **498**, 208–213.
Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **26**.
Liu, K., Lemon, B. & Traktman, P. (1995). *J. Virol.* **69**, 7823–7834.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Najarro, P., Traktman, P. & Lewis, J. A. (2001). *J. Virol.* **75**, 3185–3196.
Phan, J., Tropea, J. E. & Waugh, D. S. (2007). *Acta Cryst.* D**63**, 698–704.
Scaplehorn, N. J. (2004). PhD thesis. University College London, England.
Way, M., Pope, B., Gooch, J., Hawkins, M. & Weeds, A. G. (1990). *EMBO J.* **9**, 4103–4109.
Zhang, Z. Y. (1998). *Crit. Rev. Biochem. Mol. Biol.* **33**, 1–52.
Zhou, G., Denu, J. M., Wu, L. & Dixon, J. E. (1994). *J. Biol. Chem.* **269**, 28084–28090.