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

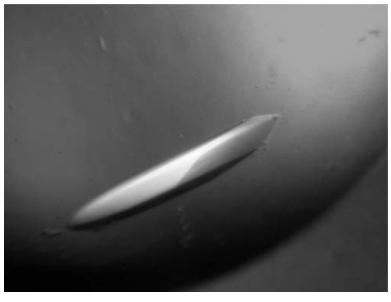
Crystallization and X-ray crystallographic analysis of human STAT1

Xiang Mao and Xiaomin Chen*

Department of Biochemistry and Molecular Biology, The University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Boulevard, Box 117, Houston, TX 77030, USA

Correspondence e-mail: xiaomin_chen@odin.mdacc.tmc.edu

Received 3 March 2005
Accepted 2 June 2005
Online 15 June 2005



Unphosphorylated human STAT1 (1–683) has been crystallized in the presence of a phosphopeptide derived from the α-chain of human interferon γ (IFNγ) receptor. A complete data set has been collected from a $KAu(CN)_2$-derivatized and dehydrated crystal. The crystal belonged to space group $P6_122$, with unit-cell parameters $a = b = 102.6$, $c = 646.5$ Å, $\alpha = \beta = 90$, $\gamma = 120^\circ$.

1. Introduction

STAT (signal transducer and activator of transcription) proteins are a family of SH2-containing transcription factors that play an important role in cytokine and growth-factor signaling. Latent STATs normally reside in the cytoplasm as unphosphorylated molecules. Upon stimulation, they are recruited to the receptor *via* SH2–pTyr interactions and become tyrosine-phosphorylated by Jak kinases, receptors with intrinsic kinase activities or non-receptor tyrosine kinases. After phosphorylation, activated STATs form homodimers or heterodimers through their reciprocal SH2–pTyr interactions, translocate into the nucleus, bind to specific DNA elements and activate gene transcription (Schindler & Darnell, 1995; Levy & Darnell, 2002).

There are seven STAT members in mammals: STAT1, STAT2, STAT3, STAT4, STAT5A, STAT5B and STAT6. They all contain between 750 and 850 amino-acid residues with similar domain structures. Naturally occurring human STAT1β protein has 712 residues and can be divided into three structural regions. From the N- to C-terminus, these are called the N-domain (residues 1–123), the core fragment (136–683; coiled-coil domain, DNA-binding domain, linker domain and SH2 domain) and the flexible tail segment (684–712) containing the tyrosine phosphorylation site (Tyr701). Crystal structures of core fragments of tyrosine phosphorylated human STAT1 (Chen *et al.*, 1998) and mouse STAT3 (Becker *et al.*, 1998) bound to DNA have been solved. Crystal structures of the dimeric N-terminal domain of STAT4 (Vinkemeier *et al.*, 1998; Chen *et al.*, 2003) and a phosphorylated *Dictyostelium* STATa core fragment without DNA (Soler-Lopez *et al.*, 2004) have also been determined. However, no crystal structure of an unphosphorylated STAT has been reported.

Unphosphorylated STAT1 was initially reported to be monomeric, whereas phosphorylated STAT1 was dimeric (Shuai *et al.*, 1994). More recent studies on the STAT proteins prior to tyrosine phosphorylation suggest dimeric or higher order oligomeric structures (Novak *et al.*, 1998; Lackmann *et al.*, 1998; Ndubuisi *et al.*, 1999; Braunstein *et al.*, 2003). Furthermore, isolated STAT N-terminal domains were shown to exist as homodimers in solution (Baden *et al.*, 1998; Byrd *et al.*, 2002; Chen *et al.*, 2003; Ota *et al.*, 2004). To explore the nature of STAT1 prior to activation, we crystallized unphosphorylated human STAT1 (1–683).

2. Experimental

2.1. Protein expression, purification and crystallization

Human STAT1 (1–683) was overexpressed in *Escherichia coli* strain BL21-CodonPlus (DE3)-RP (Stratagene) and was purified to homogeneity as described in Bromberg & Chen (2001) with minor modifications. Briefly, cells were grown at 310 K in LB and induced with 0.1 m*M* IPTG when OD_{600} was between 0.8 and 1.0. The

Table 1
Data-processing statistics from *SCALEPACK* for the $KAu(CN)_2$-derivative data set measured at ALS 822 for STAT1.

Resolution limits (Å)	$\langle I \rangle$	$\sigma(I)$	R_{sym}†	Completeness (%)
30.00–6.44	1823.7	39.8	0.040	97.0
6.44–5.12	585.8	13.6	0.062	99.4
5.12–4.48	798.3	16.2	0.058	99.2
4.48–4.07	557.5	12.7	0.073	99.2
4.07–3.78	372.9	12.0	0.101	99.3
3.78–3.56	257.9	12.3	0.142	99.0
3.56–3.38	168.1	11.6	0.197	99.1
3.38–3.23	110.7	11.1	0.274	99.1
3.23–3.11	74.4	10.8	0.383	98.9
3.11–3.00	52.8	10.5	0.508	98.9
Overall	497.3	15.3	0.083	98.9

† $R_{sym} = \sum_{hkl} |I(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} I(hkl)$, where $\langle I(hkl) \rangle$ is the mean of the symmetry-equivalent reflections of $I(hkl)$.

induction was allowed to continue overnight at 291 K. Cells were harvested and lysed with an EmulsiFlex-C5 high-pressure homogenizer (Avestin). Nucleic acid contamination was removed by adding 0.1%(*v*/*v*) (final) polyethyleneimine to the supernatant of the lysate. STAT1 was separated from other proteins by 55% ammonium sulfate precipitation. An alkylation reaction was carried out with 20 m*M* *N*-ethylmaleimide (NEM) to prevent aggregation caused by cysteine cross-linking. The STAT protein was then purified on an ÄKTA Purifier system (Amersham Pharmacia) using high-performance phenyl-Sepharose, SP Sepharose and Superdex 200 columns. Purified STAT1 was concentrated and aliquots in Eppendorf tubes were flash-frozen in liquid nitrogen and stored at 193 K.

For crystallization, 40 mg ml^{-1} protein (in 10 m*M* HEPES pH 7.2, 100 m*M* KCl, 2 m*M* DTT and 0.5 m*M* EDTA) was mixed with the phosphopeptide pYDKPH (in 40 m*M* HEPES pH 7.4) derived from the STAT1-docking site of interferon γ receptor α chain in a 1:1.5 molar ratio. Crystallization trials were carried out at 277 K by the hanging-drop vapor-diffusion method using 24-well VDXm plates (Hampton Research). Each drop was prepared by mixing 1 µl protein solution with the same volume of reservoir solution. The best crystals (~500 × 50 × 50 µm) were obtained in 1–2 weeks in 100 m*M* HEPES pH 7.0–7.1, 10–12% PEG 400 (Fig. 1). Heavy-atom derivatives were obtained by soaking crystals in stabilization solution (100 m*M* HEPES pH 7.1, 15% PEG 400) with 10 m*M* $KAu(CN)_2$ for 27 h or 1 m*M* Na_2OsCl_6 for 15 h.

2.2. Data collection and processing

Crystals were transferred at 277 K through cryoprotection/dehydration solutions with increasing concentrations of PEG 4000 (100 m*M* HEPES pH 7.1, 10.5% PEG 400 and 10–30% PEG 4000) using 5% PEG 4000 increments and 15–30 min per step. The crystals then were frozen at 100 K in a stream of liquid nitrogen and recovered into liquid nitrogen. X-ray diffraction data were measured at Advanced Light Source (ALS) beamline 8.2.2 using an ADSC Quantum-315 CCD detector. Because of the long unit-cell edge (c = 646.5 Å; Fig. 2), diffraction data were collected using a horizontal beam divergence of 0.31 and a crystal-to-detector distance of 425 mm. 96 images were collected with 1° oscillation and 90 s exposure per image. Fortuitously, the long axis of the crystals coincides with the c axis of the lattice and the crystals mounted in the loops usually had the long axis between 5 and 20° from the rotation axis. This enabled us to resolve the reflections sufficiently well for indexing even with the seemingly large oscillation range. Data processing and reduction were carried out using the *HKL*2000 and *SCALEPACK* programs (Otwinowski & Minor, 1997). Table 1 summarizes the data-processing statistics.

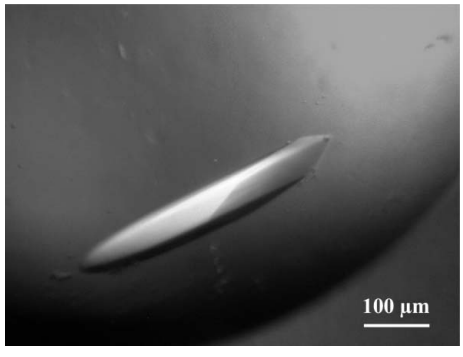


Figure 1
A crystal of STAT1 (1–683) in complex with a receptor-derived phosphopeptide.

3. Results and discussion

The best native crystals diffracted only to ~3.7 Å on a home source (Rigaku RU-H3R with an R-AXIS IV^{++} detector). The dehydration

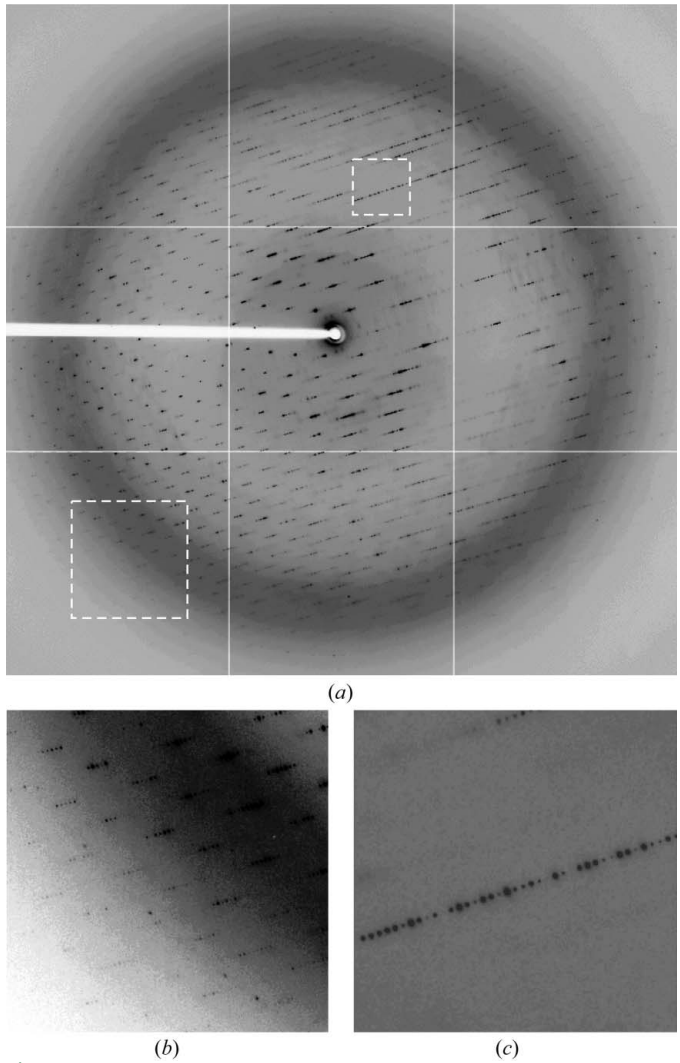


Figure 2
(*a*) X-ray diffraction image of a Na_2OsCl_6-derivatized crystal of unphosphorylated human STAT1 (1–683) in complex with the phosphopeptide; (*b*) and (*c*) magnified views of the boxed regions in (*a*).

procedure using increasing concentrations of PEG 4000 increased the diffraction limit to ∼3 Å at home. This was inspired by several recent reports on using dehydration to improve crystal diffraction (Heras *et al.*, 2003 and references therein). Most heavy-atom soaking of crystals led to physical damage to the crystals or diminished diffraction, whereas soaking with 10 m*M* $KAu(CN)_2$ or 1 m*M* Na_2OsCl_6 gave rise to similar or improved diffraction limits. The best gold-derivatized crystal diffracted to ∼2.7 Å at home. The combination of crystal dehydration and heavy-atom soaking increased the diffraction limit by 1 Å. Freezing of the crystals following the reported protocol gave rise to very low mosaicity (∼0.24) which, in combination with the use of the large detector at a synchrotron source, enabled us to resolve the closely spaced reflections (Fig. 2). By examining the *SCALEPACK* output according to the software manual, the crystals were found to belong to space group $P6_122$ or $P6_522$, with unit-cell parameters $a = 102.6$, $b = 102.6$, $c = 646.5$ Å. The structure was determined by the single anomalous dispersion method (Mao *et al.*, 2005) and the space group was determined to be $P6_122$. Two molecules of STAT1 protein and two peptide molecules are present in each asymmetric unit. The Matthews coefficient (V_M) is 3.1 $Å^3$ Da^{-1} and the solvent content is 59.6%.

We wish to thank Dr Gregory N. Parker for obtaining the initial crystals under a different condition for this project, Michael A. Pastorello and Dr Zhiyong Ren for critical reading of the manuscript, and staff at beamlines 8.2.2 and 8.3.1 of Advanced Light Source (Berkeley, CA, USA) and beamlines F1 and F2 of Cornell High Energy Synchrotron Source (Ithaca, NY, USA) for expert support. XC thanks Dr John Kuriyan for starting him on the crystallographic studies of STAT proteins and Dr David Jeruzalmi for helpful discussions. Research is supported by grants from the NIH and the Welch Foundation (XC).

References

Baden, H. A., Sarma, S. P., Kapust, R. B., Byrd, R. A. & Waugh, D. S. (1998). *J. Biol. Chem.* **273**, 17109–17114.
Becker, S., Groner, B. & Muller, C. W. (1998). *Nature (London)*, **394**, 145–151.
Braunstein, J., Brutsaert, S., Olson, R. & Schindler, C. (2003). *J. Biol. Chem.* **278**, 34133–34140.
Bromberg, J. & Chen, X. (2001). *Methods Enzymol.* **333**, 138–151.
Byrd, R. A., Gaponenko, V., Sarma, S. P., Tarasova, N., Li, J. J. & Altieri, A. S. (2002). *Proc. Am. Assoc. Cancer Res.* **43**, 139.
Chen, X., Bhandari, R., Vinkemeier, U., Van Den Akker, F., Darnell, J. E. Jr & Kuriyan, J. (2003). *Protein Sci.* **12**, 361–365.
Chen, X., Vinkemeier, U., Zhao, Y., Jeruzalmi, D., Darnell, J. E. Jr & Kuriyan, J. (1998). *Cell*, **93**, 827–839.
Heras, B., Edeling, M. A., Byriel, K. A., Jones, A., Raina, S. & Martin, J. L. (2003). *Structure*, **11**, 139–145.
Lackmann, M., Harpur, A. G., Oates, A. C., Mann, R. J., Gabriel, A., Meutermans, W., Alewood, P. F., Kerr, I. M., Stark, G. R. & Wilks, A. F. (1998). *Growth Factors*, **16**, 39–51.
Levy, D. E. & Darnell, J. E. (2002). *Nature Rev. Mol. Cell Biol.* **3**, 651–662.
Mao, X., Ren, Z., Parker, G. N., Sondermann, H., Pastorello, M. A., Wang, W., McMurray, J. S., Demeler, B., Darnell, J. E. Jr & Chen, X. (2005). *Mol. Cell*, **17**, 761–771.
Ndubuisi, M. I., Guo, G. G., Fried, V. A., Etlinger, J. D. & Sehgal, P. B. (1999). *J. Biol. Chem.* **274**, 25499–25509.
Novak, U., Ji, H., Kanagasundaram, V., Simpson, R. & Paradiso, L. (1998). *Biochem. Biophys. Res. Commun.* **247**, 558–563.
Ota, N., Brett, T. J., Murphy, T. L., Fremont, D. H. & Murphy, K. M. (2004). *Nature Immunol.* **5**, 208–215.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Schindler, C. & Darnell, J. E. Jr (1995). *Annu. Rev. Biochem.* **64**, 621–651.
Shuai, K., Horvath, C. M., Huang, L. H., Qureshi, S. A., Cowburn, D. & Darnell, J. E. Jr (1994). *Cell*, **76**, 821–828.
Soler-Lopez, M., Petosa, C., Fukuzawa, M., Ravelli, R., Williams, J. G. & Muller, C. W. (2004). *Mol. Cell*, **13**, 791–804.
Vinkemeier, U., Moarefi, I., Darnell, J. E. Jr & Kuriyan, J. (1998). *Science*, **279**, 1048–1052.