FOR THE RECORD

A reinterpretation of the dimerization interface of the N-terminal Domains of STATs

XIAOMIN CHEN,^{1,7} RASHNA BHANDARI,^{2,7} UWE VINKEMEIER,³
FOCCO VAN DEN AKKER,⁴ JAMES E. DARNELL, JR.,⁵ AND JOHN KURIYAN^{2,6}

¹Department of Biochemistry and Molecular Biology, The University of Texas, M.D. Anderson Cancer Center, Houston, Texas 77030, USA

²Howard Hughes Medical Institute, Departments of Molecular and Cell Biology and of Chemistry, University of California, Berkeley, Berkeley, California 94720, USA

³Forschungsinstitut für Molekulare Pharmacologie, 13125 Berlin, Germany

⁴Department of Molecular Biology, Lerner Research Institute, Cleveland Clinic Foundation, Cleveland, Ohio 44195, USA

⁵The Rockefeller University, New York, New York 10021, USA

⁶Physical Bioscience Division, Lawrence Berkeley National Lab, Berkeley, California 94720, USA

(RECEIVED June 7, 2002; ACCEPTED June 13, 2002)

Abstract

The crystal structures of the N-terminal domain (N-domain) and the core region of the STAT family of transcription factors have been determined previously. STATs can form cooperative higher order structures (tetramers or higher oligomers) while bound to DNA. The crystal packing in the STAT4 N-domain crystal structure, determined at 1.5 Å resolution, suggests two alternate organizations of the N-domain dimer. We now present the results of site directed mutagenesis of residues predicted to be involved at each dimer interface. Our results indicate that the dimer interface suggested earlier as being physiologically relevant is, in fact, unlikely to be so. Given the alternative model for the N-domain dimer, the ability of the N-domain to mediate interactions of two STAT dimers on DNA remains unchanged.

Keywords: STAT; dimerization; cooperative DNA binding

The STAT (signal transducers and activators of transcription) proteins are a family of transcription factors involved in the activation of target genes in response to cytokines and growth factors (Ihle and Kerr 1995; Darnell 1997). The binding of these ligands to their cognate receptors leads to tyrosine kinase activation and phosphorylation of latent STAT monomers in the cytoplasm (Darnell et al. 1994). Tyrosine phosphorylated STATs undergo homo- or heterodimerization via reciprocal SH2-phosphotyrosine interactions, followed by translocation to the nucleus and activa-

tion of gene expression (Darnell 1997). The canonical STAT recognition site on DNA is the palindromic sequence TTCN₃₋₄GAA (Horvath et al. 1996; Ihle 1996). It has been shown that STAT1, STAT4, and STAT5 are able to form higher order complexes (dimer:dimer or higher) on promoters containing two or more neighboring STAT binding sites (Vinkemeier et al. 1996; Xu et al. 1996; John et al. 1999). This interaction between STAT dimers is cooperative, and is lost upon deletion of the N-domain of the STATs (Vinkemeier et al. 1996; Xu et al. 1996; Zhang and Darnell 2001).

Earlier work on the crystal structure of the N-domain of STAT4 (residues 1–124) (Vinkemeier et al. 1998) and of the core (residues ~130 to ~715; lacking the N-domain) STAT1 and STAT3β dimers bound to DNA (Becker et al. 1998; Chen et al. 1998), has led to our current understanding of the molecular architecture of STAT proteins. The N-domain

Reprint requests to: John Kuriyan, 401 Barker Hall MC 3202, Molecular and Cell Biology, University of California, Berkeley, CA 94720-3202, USA; e-mail: kuriyan@uclink.berkeley.edu; fax: (510) 643-2352.

⁷These authors contributed equally to this work.

Article and publication are at <http://www.proteinscience.org/cgi/doi/10.1110/ps.0218903>.

of STAT is linked to the core via a flexible linker of ~24 residues, and it was suggested that dimerization of the N-domains of adjacent STAT dimers on DNA leads to the formation of higher order STAT complexes on DNA (Chen et al. 1998). The N-domain of STAT4, which is highly similar to STAT1 (51% sequence identity) was crystallized with one molecule in the asymmetric unit. Mutation of Trp 37, a residue located between two molecules at a crystal packing interface, led to the loss of cooperative STAT binding to tandem sites on DNA (Vinkemeier et al. 1998; John et al. 1999). Consequently, we interpreted our structure in terms of this putative dimer interface seen in the crystal (Vinkemeier et al. 1998).

In this communication we demonstrate that our earlier interpretation of the dimer interface in the STAT4 N-domains may not be valid. We present an alternate interpretation of the crystal packing in the same crystal form, indicating that another dimer interface suggested by the crystal structure may be relevant in solution.

Results and Discussion

Experimental studies in our laboratories have focussed mainly on STAT1, but repeated attempts to obtain high quality crystals of the N-domain of STAT1 have failed. We have also failed to obtain crystals of the STAT4 N-domain in a different crystal form than that originally reported (Vinkemeier et al. 1998). We, therefore, continue to interpret experiments on STAT1 in terms of the original crystal structure of the N-domain of STAT4. Crystal packing in the STAT4 crystal suggests two interfaces that are potentially relevant for dimer formation. Interface I (Fig. 1A), originally analyzed by Vinkemeier et al. (1998), is essentially polar, with 1458 Å² of total surface area buried (calculated using a 1.4-Å probe radius). Interface II is more extensive (2030 Å² total surface area buried), and contains hydrophobic residues (Fig. 1B).

We introduced point mutations in STAT1 at several sites at each interface. The dimerization properties of these mu-

Figure 1. Two alternate organizations of the STAT4 N-domain dimer. (A) The STAT4 N-domain dimer suggested earlier by Vinkemeier et al. (1998). (B) The alternate STAT4 N-domain dimer suggested by crystal packing. Close-up views of dimer interface I (A) and interface II (B) are presented, and the residues involved in dimer formation are indicated. The structures were drawn using Ribbons (Carson 1991), and the PDB coordinates 1BGF for the STAT4 N-domain (Vinkemeier et al. 1998). Starting from the deposited coordinates of the protomer (x,y,z ; PDB entry 1BGF), the coordinates for the second molecule in the new dimer (x',y',z') can be generated by $x' = 0.5000*x + 0.8660*y + -0.0000*z + -39.7485$; $y' = 0.8660*x + -0.5000*y + 0.0000*z + 68.8534$; $z' = 0.0000*x + 0.0000*y + -1.0000*z + 70.5855$.

tant proteins are shown in Table 1. At interface I (Vinkemeier et al. 1998) residues Trp37, Gln41, Gln36, and Arg70 were mutated to Ala. STAT1 (W37A) was expressed very poorly and we were unable to study the properties of this protein. A low level of expression of this mutant STAT protein has also been reported in another study (Murphy et al. 2000). The production of full-length STAT1 (W37A) frequently leads to proteolytic degradation of the protein (U. Vinkemeier unpubl.). We were, however, able to obtain and purify sufficient amounts of the N-domain of STAT1 (W37F), and this protein is a dimer, as shown by analytical ultracentrifugation (Fig. 2) and gel filtration analysis (Table 1). Trp 37 was thought to mediate dimer formation by participating in direct and water-mediated hydrogen bonds (Vinkemeier et al. 1998), interactions that would be disrupted in the W37F mutant. The N-domain of STAT1 (W37F) is stable and is still a dimer, suggesting that W37 is perhaps not a part of the dimer interface. The fact that dimer formation is unimpeded in the W37F mutant suggests that the loss of tetramer formation on tandem sites on DNA seen for the full-length STAT1 (W37A) mutant (Vinkemeier et al. 1988) is probably not due to a specific disruption of the N-domain dimer interface. Three other residues implicated in dimer formation at interface I were mutated individually to Ala in STAT1, and the mutants are all dimeric (Table 1).

An alternate dimer interface suggested by crystal packing (interface II) is shown in Figure 1B. Not only is interface II more extensive than interface I, it also involves interactions between hydrophobic residues (unlike the essentially polar nature of interface I). Certain residues at interface II were individually replaced by Ala (Table 1), and the mutant STAT N-domains were examined for dimerization. Proteins containing mutations at one side of the interface, F77A and L78A, were monomers as seen by analytical ultracentrifugation (Fig. 2). To ensure that the mutant proteins (F77A and L78A) are folded properly, CD scans of these proteins were carried out and were found to be identical to wild-type STAT1 N-domain (Fig. 3). Results for mutations at the other side of the interface were not as clear, but nevertheless provide evidence for interference with dimer formation. M28A migrated as an intermediate between dimer and monomer on gel filtration analysis, but appeared as a dimer by analytical ultracentrifugation analysis. S12A showed mainly aggregates and a small monomer population. The hydrophobic nature of the residues at positions 77 and 78 is conserved between STAT1 and STAT4, which has leucine residues at both positions. Likewise, Met 28 is conserved in STAT4.

These results indicate that interface II may indeed be relevant to dimer formation in solution. In contrast to interface I, for which none of the mutations we introduced had a significant effect on dimer formation, several mutations at interface II clearly interfered with the stability of the dimer. It should be stressed, however, that a definitive conclusion

Table 1. Properties of the wild-type and mutant STAT1 N-domain proteins

STAT1	Sedimentation equilibrium	Gel filtration
Wild type	28 kD	dimer
Interface I		
W37A	—	—
W37F	28 kD	dimer
Q41A	27 kD	dimer
Q36A	29 kD	dimer
R70A	27 kD	dimer
Interface II		
Q8A	27 kD	dimer
S12A	17 kD + aggregates	not examined
L15A	28 kD	dimer
M28A	26 kD	monomer
E29A	27 kD	dimer
F77A	15 kD	monomer
L78A	15 kD	not examined

The point mutation introduced in each of the STAT1 N-domain mutant proteins is indicated in the first column. The approximate molecular weight estimated by sedimentation equilibrium experiments, and migration as a monomer or dimer species on gel filtration analysis, is shown for each mutant protein.

regarding the mode of dimer formation by the N-domain awaits further structural analysis.

A key conclusion that emerged from the previous analysis of the N-domain dimer was that the distance between the C-terminal residues in the dimer was consistent with the placement of the N-domain dimer between two adjacent STAT core dimers on tandem DNA sites (Chen et al. 1998). Our reinterpretation of the N-domain dimer interface does not alter this conclusion. The original N-domain dimer had its C-termini located 30 Å apart (Vinkemeier et al. 1998). The original N-domain dimer could be positioned between two STAT core dimers modeled on adjacently located sites on DNA so that the C-terminal region of each N-domain monomer was located about 27 Å away from an N-terminal region of the adjacent STAT core dimer, to which it would be connected by a flexible 24 residue tether (Chen et al. 1998). The C-terminal residues of the newly proposed dimer are located ~64 Å apart. The increased span between the C-termini means that this dimer can be positioned between two adjacent STAT core dimers modeled on DNA with essentially no gap at the junction points.

In conclusion, our results demonstrate that the earlier interpretation of the dimer interface of the N-domains of STATs needs to be reconsidered. There have been several studies wherein mutation of Trp 37 to Ala led to a decrease in transcriptional activation from tandem STAT binding sites (Vinkemeier et al. 1998; John et al. 1999; Zhang and Darnell 2001). It has also been reported that the W37A STAT4 mutant failed to be tyrosine phosphorylated upon interferon- α stimulation (Murphy et al. 2000). These results may reflect an overall destabilization of the N-domain rather than a specific defect in dimerization. We also sug-

Figure 2. Analytical ultracentrifugation sedimentation equilibrium data. Each STAT N-domain protein was analyzed at three different concentrations at 25,000 rev/min (see Materials and Methods). Representative results for the wild-type protein and some of the STAT1 N-domain mutant proteins are shown. In each case, the upper panel shows the residual difference between experimental and fitted values by its standard deviation, and the lower panel shows the equilibrium profile. The variances (V) between the fitted and experimental values, and calculated molecular mass (M) in daltons are indicated. The theoretical molecular weight of the STAT1 N-domain monomer is 15,223 Daltons.

gest that future studies investigating the role of STAT N-domains in dimerization keep in mind the alternate STAT dimer proposed in this article, and consider mutating residues corresponding to Phe 77 and Leu 78 in STAT1.

Materials and methods

The N-domain of human STAT1 (amino acid residues 1–124) was cloned as a C-terminal fusion to glutathione S-transferase (GST),

in a pGEX2T vector (Amersham Biosciences) that had been modified to replace the thrombin protease cleavage site with a cleavage site for tobacco etch virus (TEV) protease. Site-directed mutagenesis was carried out using the Quikchange method (Stratagene). The construct and mutations were confirmed by sequencing.

The constructs were expressed in the *Escherichia coli* strain BL21(ΔDE3). Cells were resuspended in buffer A (50 mM Tris pH 8.0, 150 mM NaCl, and 1 mM DTT) and lysed in a French press. The lysate was clarified by high-speed centrifugation and the supernatant fraction was purified on a glutathione sepharose column on the Amersham Biosciences AKTA FPLC system. After wash-

Figure 3. Circular dichroism spectra of STAT1 N-domain proteins. The spectra for wild-type STAT1 N-domain (blue), STAT1 F77A (green), and STAT1 L78A (red) were determined as described in Materials and Methods.

ing the column with five column volumes of buffer A, the fusion protein was eluted using 20 mM reduced glutathione in buffer A. TEV protease was added to the pooled fractions and the digestion was carried out at 15°C overnight. The N-domain and GST were separated on a HiTrap Q column (Amersham Biosciences), in buffer A using a 0%–70% gradient of buffer B (50 mM Tris pH 8.0, 800 mM NaCl, and 1 mM DTT) over 30 column volumes. The pooled fractions of the peak containing the STAT1 N-domain were concentrated and passed over a Superdex 75 column to separate any remaining GST, which migrates as a dimer of about 52 kDa. In the case of mutant proteins F77A and L78A, there was very poor separation between GST and STAT1 N-domain on a Q column. These proteins were well separated from GST on a Superdex 75 column.

For gel filtration analysis, 1.5 mg of purified STAT N domain protein in a volume of 500 μ L was run on a 120-mL Superdex 75 column at a flow rate of 0.5 mL/min, in 50 mM Tris, pH 8.0, 100 mM NaCl, and 1 mM DTT. Equilibrium sedimentation experiments were performed using a Beckman Optima XL-A analytical ultracentrifuge with an An-60 Ti rotor and six-sector cells. STAT N-domain proteins at concentrations of 0.65, 0.32, and 0.16 mg/mL were centrifuged in the gel filtration buffer, at 25,000 rev/min at 4°C for 20 h. Subsequently, absorbance measurements at 280 nm were taken in 0.001 cm radial steps and equilibrium was ascertained by comparing scans taken at 1-h intervals. The Optima XL-A/XL-I data analysis software from Beckman Coulter was used for data processing and curve fitting. A partial specific vol-

ume of 0.73 cm³/g was used and background absorbance was corrected empirically by allowing the baseline to float during the fitting calculations.

CD measurements were performed on an Aviv Model 215 Circular Dichroism Spectrometer at 25°C using a 0.02-cm pathlength cuvette. The purified proteins were dialysed against PBS (10 mM sodium phosphate buffer, pH 7.4, 140 mM NaCl, 10 mM KCl) and diluted to a concentration of 40 μ M. Spectra were recorded from 250 to 190 nm using a step of 0.5 nm and an averaging time of 4 sec.

Acknowledgments

We thank Xiaokui Zhang, Melissa Henrikson, Mike Goger, and David Cowburn for helpful discussions, Natalia Rodionova for ultracentrifugation analysis, and Lore Leighton and Nahed Shahabi for preparation of the manuscript and figures.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

- Becker, S., Groner, B., and Muller, C.W. 1998. Three-dimensional structure of the Stat3 β homodimer bound to DNA. *Nature* **394**:145–151.
- Carson, M. 1991. RIBBONS 2.0. *J. Appl. Crystallogr.* **24**: 958–961.
- Chen, X., Vinkemeier, U., Zhao, Y., Jeruzalmi, D., Darnell, Jr., J.E., and Kuriyan, J. 1998. Crystal structure of a tyrosine phosphorylated STAT-1 dimer bound to DNA. *Cell* **93**: 827–839.
- Darnell, Jr., J.E. 1997. STATs and gene regulation. *Science* **277**: 1630–1635.
- Darnell, Jr., J.E., Kerr, I.M., and Stark, G.R. 1994. Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. *Science* **264**: 1415–1421.
- Horvath, C.M., Stark, G.R., Kerr, I.M., and Darnell, Jr., J.E. 1996. Interactions between STAT and non-STAT proteins in the interferon-stimulated gene factor 3 transcription complex. *Mol. Cell. Biol.* **16**: 6957–6964.
- Ihle, J.N. 1996. STATs: Signal transducers and activators of transcription. *Cell* **84**: 331–334.
- Ihle, J.N. and Kerr, I.M. 1995. Jaks and stats in signaling by the cytokine receptor superfamily. *Trends Genet.* **11**: 69–74.
- John, S., Vinkemeier, U., Soldaini, E., Darnell, Jr., J.E., and Leonard, W.J. 1999. The significance of tetramerization in promoter recruitment by Stat5. *Mol. Cell. Biol.* **19**: 1910–1918.
- Murphy, T.L., Geissal, E.D., Farrar, J.D., and Murphy, K.M. 2000. Role of the Stat4 N domain in receptor proximal tyrosine phosphorylation. *Mol. Cell. Biol.* **20**: 7121–7131.
- Vinkemeier, U., Cohen, S.L., Moarefi, I., Chait, B.T., Kuriyan, J., and Darnell, Jr., J.E. 1996. DNA binding of in vitro activated Stat1 α , Stat1 β and truncated Stat1: Interaction between NH2-terminal domains stabilizes binding of two dimers to tandem DNA sites. *EMBO J.* **15**: 5616–5626.
- Vinkemeier, U., Moarefi, I., Darnell, Jr., J.E., and Kuriyan, J. 1998. Structure of the amino-terminal protein interaction domain of STAT-4. *Science* **279**: 1048–1052.
- Xu, X., Sun, Y.L., and Hoey, T. 1996. Cooperative DNA binding and sequence-selective recognition conferred by the STAT amino-terminal domain. *Science* **273**: 794–797.
- Zhang, X. and Darnell, Jr., J.E. 2001. Functional importance of Stat3 tetramerization in activation of the α 2-macroglobulin gene. *J. Biol. Chem.* **276**: 33576–33581.