FEBS Letters 585 (2011) 281–285

FEBS Letters

journal homepage: www.FEBSLetters.org

An adjacent arginine, and the phosphorylated tyrosine in the c-Met receptor target sequence, dictates the orientation of c-Cbl binding

Qingxiang Sun [a], Cherlyn Ng [b], Graeme R. Guy [b,*], J. Sivaraman [a,*]

[a] Department of Biological Sciences, National University of Singapore, 14 Science Drive 4, Singapore 117543, Singapore
[b] Institute of Molecular and Cell Biology, 61 Biopolis Drive, 03-01 Proteos, Singapore 138673, Singapore

ARTICLE INFO

Article history:
Received 21 October 2010
Revised 19 November 2010
Accepted 27 November 2010
Available online 15 December 2010

Edited by Christian Griesinger

Keywords:
c-Casitas B-lineage lymphoma tyrosine kinase binding domain
Hepatocyte growth factor receptor
Reverse binding
Surface plasmon resonance
Crystal structure

ABSTRACT

Previously, we have demonstrated that the tyrosine phosphorylated hepatocyte growth factor receptor (Met) binds to the c-Cbl phosphotyrosine-recognition, tyrosine kinase binding (TKB) domain in a reverse orientation compared to other c-Cbl binding partners. A Met peptide with the DpYR motif changed to RpYD (MetRD) retains a similar TKB binding affinity as the native Met peptide. However, the TKB: MetRD complex crystal structure reveals a complete reversal of the binding orientation. Collated data indicates that both binding and orientation is dictated by the phosphorylated tyrosine and an adjacent arginine forming intra-peptide hydrogen bonds and aligning unidirectionally with complementary charges in the phosphotyrosine binding pocket of c-Cbl.

Structured summary:
c-Cbl and **MetRD** *bind*: shown by *x-ray crystallography* (view interaction)
c-Cbl and **MetRD** *bind*: shown by *mass spectrometry studies of complexes* (view interaction)
c-Cbl *bind* to **Met**: shown by *surface plasmon resonance* (view interactions 1,2)

1. Introduction

The casitas B-lineage lymphoma (Cbl) family of proteins, epitomized by c-Cbl, has a well-documented role in a number of aspects of cell signaling. Cbl was first known as a scaffold protein and thereafter characterized as a Ubiquitin E3 ligase [1,2]. It plays a major role in the downregulation of a number of receptor and non-receptor tyrosine kinases. Cbl proteins have multiple domains and have been demonstrated to interact with a number of other proteins via proline-rich sequences and phosphotyrosine residues [3]. To enable its role in the ubiquitin-targeting system, Cbl contains a novel embedded Src homology 2 (SH2) domain that recognizes particular phoshotyrosine containing sequences on target proteins [4].

SH2 domains are well-characterized phosphotyrosine recognizing domains that seemed to co-evolve with tyrosine phosphorylation in metazoans, with relatively strict recognition sequences specific to each SH2 domain contained within different proteins [5]. The association of 4H and EF domains with the SH2 domain on Cbl, however, appears to be unique. It was somewhat surprising that the embedded SH2 domain of Cbl appeared to have a rather wide specificity in sequence recognition. We recently addressed the question as to how the Cbl tyrosine kinase binding (TKB) domain could recognize a 'consensus' sequence (N/D)XpY(S/T)XXP as well as an apparently unrelated DpYR sequence [6]. We demonstrated that in recognizing the DpYR sequence on the hepatocyte growth factor receptor (Met), Cbl flips over to bind in a reverse orientation relative to targets with the canonical sequence. Binding affinity and perhaps orientation were mediated by an intrapeptidyl bond between the phosphorylated tyrosine residue and an adjacent Arg or a nearby Asp residue [7]. It is noteworthy that the epidermal growth factor receptor (EGFR), a well characterized Cbl target, had an Arg in the pY−1 position that would resemble the DpYR motif found in the Met family of receptors, but in a reversed direction. The EGFR sequence also contained a pY+1 Ser and pY+4 Pro unlike Met's sequence.

We asked the question in this study: what determines the orientation of target peptides binding to the Cbl TKB domain? To accomplish this, we took the most minimal DpYR Met binding sequence and tested the binding of derivative peptides and crystallized candidates that bound, in order to understand factors governing binding motif orientation of the Cbl-TKB.

Abbreviations: TKB, tyrosine kinase binding; SH2, Src homology 2; SPR, surface plasmon resonance; EGFR, epidermal growth factor receptor; Met, hepatocyte growth factor receptor; Cbl, casitas B-lineage lymphoma

* Corresponding authors. Fax: +65 6779 1117 (G.R. Guy), +65 6779 5671 (J. Sivaraman).

E-mail addresses: mcbgg@imcb.a-star.edu.sg (G.R. Guy), dbsjayar@nus.edu.sg (J. Sivaraman).

0014-5793/$36.00
doi:10.1016/j.febslet.2010.11.060

2. Materials and methods

We have designed two peptides based on the sequence of wild type Met (MetWT), (1) by reversing core residues of the peptide (hereafter referred as MetRD) and (2) by changing pYRA residues of MetWT peptide to pYAN (hereafter referred as MetN) (Table 1). These peptides were purchased from GL Biochem (Shanghai, China).

2.1. Biophysical interaction studies

Surface plasmon resonance (SPR) experiments were performed with Biacore 3000 (Biacore AB, Uppsala, Sweden). c-Cbl protein (50 ng/μL in 5 mM sodium citrate pH 6.5) was immobilized onto a CM 5 Chip as per the manufacturer's recommendations. The running buffer consisted of 20 mM Na Hepes pH 7.0 and 200 mM NaCl. Different concentrations of peptides (300 nM to 27 μM in running buffer) were applied to the chip surface at a flow rate of 20 μl/min at 25 °C. Regeneration using running buffer wash for 2 min resulted in a stable baseline corresponding to the starting baseline level. Reference cells are immobilized with inactive c-Cbl TKB domain resulting from 1 min flow of 10 mM H_2SO_4 through the cell. The equilibrium constant (Kd) was determined by the 1:1 Langmuir binding fitting model provided by the Biacore 3000 instrument software.

2.2. Complex formation, crystallization and structure determination

Cloning, expression and purification of c-Cbl was performed as described previously [6]. Phosphopeptides were reconstituted in 20 mM Hepes, 200 mM NaCl and pH 7.0. MetRD was incubated with purified c-Cbl TKB in fivefold molar excess, and concentrated to 5 mg/ml using Amicon ultrafiltration devices (Millipore, Billerica, MA). TKB: MetRD complex crystallized in 100 mM Bis–tris propane pH 6.5, 50 mM ammonium sulfate by mixing 1 μl of reservoir solution with 2 μl of protein using hanging drop vapor diffusion method at room temperature. The mother liquor supplemented with 10% glycerol as cryo-protectant. The X-ray diffraction data was collected using in-house Bruker X-8 PROTEUM system and processed with HKL2000 [7]. The structure was solved by using molecular replacement method with the program MolRep [8] using c-Cbl–TKB domain as a search model (pdb code: 3BUX). The resulting model with the electron density map was examined in the program COOT [9], and necessary manual model building was performed. Several cycles of map fitting and alternated with refinement using the program Refmac5 [10] led to the convergence of R-values (Table 2).

2.3. Protein databank accession code

Coordinates and structure factors of the MetRD: Cbl-TKB complex was deposited at RCSB Protein Data Bank with the code 3PLF.

3. Results and discussion

A peptide from the Cbl–TKB atypical binding motif (DpYR) in the Met receptor (MetWT) chosen in our previous study was crystallised in complex with the Cbl–TKB domain [6]. This peptide had the sequence SNESVDpYRATFPE, and bound in a reverse direction on the TKB domain relative to other peptides of the typical binding motif [NXpY(S/T)XXP]. In an initial experiment to determine factors responsible for the orientation reversal of the Met atypical sequence, two more peptides: (1) SNESVRpYDATFPE (MetRD), where conserved residues flanking the phosphorylated tyrosine were reversed, (2) SNESVDpYANTFPE (MetN) where the conserved arginine in the wildtype sequence was replaced by alanine, were constructed.

The affinities of these peptides binding to the TKB domain were assessed by SPR and the data obtained is shown in Table 1 and Fig. 1. Two distinct observations were apparent: (1) The MetRD binding affinity was similar to the wildtype peptide affinity, and (2) no binding was observed when the conserved arginine in the wild type Met sequence was substituted by alanine. It was previously determined that the minimum TKB binding sequence for MetWT is DpYR, suggesting little contribution by any other residues flanking the DpYR motif. The lack of binding of the MetN peptide to the TKB domain indicates that the conserved arginine is essential for binding and possibly orientation of the peptide. While the MetRD and MetWT peptides bind with similar affinity, we did not know the structural changes implicated by reversal of the DpYR motif in the MetRD peptide.

To ascertain the binding orientation of MetRD and to investigate its mode of binding to the Cbl-TKB, we proceeded to crystallize the MetRD: Cbl-TKB complex and compared it with the previously determined MetWT: Cbl-TKB structure [6]. The c-Cbl-TKB: MetRD complex was crystallized in $P2_1$ spacegroup with two complex molecules in an asymmetric unit. The structure was refined to a

Table 1
Comparison of binding parameters for the interactions between c-Cbl and the phosphorylated peptides.

Peptide	Sequence	k_{on} (M^{-1} s^{-1})	k_{off} (s^{-1})	*Kd* (μM)	χ^2
MetWT	SNESVDpYRATFPE	2.7×10^5	1.17	4.4	1.55
MetRD	SNESVRpYDATFPE	2.0×10^5	1.38	6.7	0.42
MetN	SNESVDpYANTFPE			No binding	

χ^2 is the statistical error between the experimental and theoretical models. χ^2 is defined as $\Sigma(R_f - R_x)^2/(n-p)$, where R_f is the fitted value at a given point, R_x is the experimental value at the same point, n is the number of data points, and p is the number of fitted parameters.

Table 2
Crystallographic refinement statistics of c-Cbl:MetRD complex.

Data set Data collection*	c-Cbl:MetRD
Cell parameters (Å, °)	a = 52.7, b = 104.5, c = 60.6, β = 90.2
Space group	$P2_1$
Resolution range (Å)	50–1.92 (1.99–1.92)
Wavelength (Å)	1.5418
Observed reflections	187 613
Unique reflections	48 545
Completeness (%)	97.6 (92.5)
Overall *(I/σI)*	13.3 (3.1)
R_{sym}[a] (%)	0.065 (0.203)
Solvent content (%)	61
Refinement statistics	
Resolution range (Å)	20–1.92
R_{work}[b] (no. of reflections)	0.159 (43 601)
R_{free}[c] (no. of reflections)	0.180 (2462)
RMSD bond lengths (Å)	0.022
RMSD bond angles (°)	1.627
No. of protein atoms/ligand atoms/water molecules	4942/80/575
B-factors (Å²)	
Average B-factors of protein atoms	29.2
Rms B-factor of protein atoms	1.8
Average B-factors of ligand atoms	34.1
Rms B-factor of ligand atoms	3.3
Average B-factors of water atoms	33.6
Ramachandran plot	
Most favored regions (%)	99.4
Generously allowed regions (%)	0.6
Disallowed regions (%)	0.0

[a] $R_{sym} = \sum |I_i - \langle I \rangle| / \sum |I_i|$, where I_i is the intensity of the ith measurement, and $\langle I \rangle$ is the mean intensity for that reflection.
[b] $R_{work} = \sum |F_{obs} - F_{calc}| / \sum |F_{obs}|$, where F_{calc} and F_{obs} are the calculated and observed structure factor amplitudes, respectively.
[c] R_{free} = as for R_{work}, but for 5.0% of the total reflections chosen at random and omitted from refinement for all datasets.
* Statistics for the high resolution bin.