

## Research

## Open Access

# Mapping of immunogenic and protein-interacting regions at the surface of the seven-bladed $\beta$ -propeller domain of the HIV-1 cellular interactor EED

Dina Rakotobe<sup>1</sup>, Sébastien Violot<sup>1,2</sup>, Saw See Hong<sup>1</sup>, Patrice Gouet<sup>2</sup> and Pierre Boulanger\*<sup>1,3</sup>

Address: <sup>1</sup>Laboratoire de Virologie & Pathologie Humaine, Université Lyon I & CNRS FRE-3011, Faculté de Médecine Laennec, 7 rue Guillaume Paradin, 69372 Lyon Cedex 08, France, <sup>2</sup>Laboratoire de BioCristallographie, IBCP, Institut Fédératif de Recherche IFR128 BioSciences Lyon-Gerland, 7 passage du Vercors, 69367 Lyon Cedex 07, France and <sup>3</sup>Laboratoire de Virologie Médicale, Centre de Biologie & Pathologie du Pôle Est, Hospices Civils de Lyon, 59 Boulevard Pinel, 69677 Bron Cedex, France

Email: Dina Rakotobe - dinaraktb@yahoo.fr; Sébastien Violot - sebastien.violot@free.fr; Saw See Hong - sawsee.hong@sante.univ-lyon1.fr; Patrice Gouet - p.gouet@ibcp.fr; Pierre Boulanger\* - Pierre.Boulanger@sante.univ-lyon1.fr

\* Corresponding author

Published: 27 February 2008

Received: 22 January 2008

*Virology Journal* 2008, **5**:32 doi:10.1186/1743-422X-5-32

Accepted: 27 February 2008

This article is available from: <http://www.virologyj.com/content/5/1/32>

© 2008 Rakotobe et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## Abstract

**Background:** The human EED protein, a member of the superfamily of *Polycomb* group proteins, is involved in multiple cellular protein complexes. Its C-terminal domain, which is common to the four EED isoforms, contains seven repeats of a canonical WD-40 motif. EED is an interactor of three HIV-1 proteins, matrix (MA), integrase (IN) and Nef. An antiviral activity has been found to be associated with isoforms EED3 and EED4 at the late stage of HIV-1 replication, due to a negative effect on virus assembly and genomic RNA packaging. The aim of the present study was to determine the regions of the EED C-terminal core domain which were accessible and available to protein interactions, using three-dimensional (3D) protein homology modelling with a WD-40 protein of known structure, and epitope mapping of anti-EED antibodies.

**Results:** Our data suggested that the C-terminal domain of EED was folded as a seven-bladed  $\beta$ -propeller protein. During the completion of our work, crystallographic data of EED became available from co-crystals of the EED C-terminal core with the N-terminal domain of its cellular partner EZH2. Our 3D-model was in good congruence with the refined structural model determined from crystallographic data, except for a unique  $\alpha$ -helix in the fourth  $\beta$ -blade. More importantly, the position of flexible loops and accessible  $\beta$ -strands on the  $\beta$ -propeller was consistent with our mapping of immunogenic epitopes and sites of interaction with HIV-1 MA and IN. Certain immunoreactive regions were found to overlap with the EZH2, MA and IN binding sites, confirming their accessibility and reactivity at the surface of EED. Crystal structure of EED showed that the two discrete regions of interaction with MA and IN did not overlap with each other, nor with the EZH2 binding pocket, but were contiguous, and formed a continuous binding groove running along the lateral face of the  $\beta$ -propeller.

**Conclusion:** Identification of antibody-, MA-, IN- and EZH2-binding sites at the surface of the EED isoform 3 provided a global picture of the immunogenic and protein-protein interacting regions in the EED C-terminal domain, organized as a seven-bladed  $\beta$ -propeller protein. Mapping of the HIV-1 MA and IN binding sites on the 3D-model of EED core predicted that EED-bound MA and IN ligands would be in close vicinity at the surface of the  $\beta$ -propeller, and that the occurrence of a ternary complex MA-EED-IN would be possible.

## Background

Human EED protein, the human ortholog of the mouse embryonic ectoderm development (*eed*) gene product, is a member of the superfamily of WD-40 repeat proteins which belongs to the highly conserved *Polycomb* group (*PcG*) family of proteins [1-7]. The human EED protein has been found to interact with several cellular proteins in both cytoplasmic and nuclear compartments. At the inner side of the plasma membrane, EED interacts with the cytoplasmic tail of integrin  $\beta 7$  subunit [8], a domain involved in major integrin functions [9,10]. Within the nucleus, EED participates in Polycomb Repressive Complexes (PRCs), multiprotein edifices which have been identified in *Drosophila* and in mammals (reviewed in [11]). Several types of PRCs have been described and referred to as PRC1, PRC2 and PRC3 [12]. PRC2/3 content includes, among other components, EED, EZH2, SUZ12 and RbAp46/48 [12-14].

In the context of HIV-1-infected cells, EED has been found to interact with three viral proteins, the structural protein matrix (MA) [15], the enzyme integrase (IN) [16] and the regulatory protein Nef [17]. These interactions involved the C-terminal domain of EED, or EED core, common to the four isoforms. It has been suggested that the nuclear depletion of EED which resulted from the EED-Nef interaction occurring at the plasma membrane of HIV-1-infected cells would be responsible for the release of an EED-mediated transcriptional block and for an indirect transcriptional activation of the virus [17]. This hypothesis was consistent with the reported functions of *PcG* proteins, which act as transcriptional repressors of homeotic genes (reviewed in [11,18-20]), and contribute to the maintenance of the silent state of chromatin in upper eukaryotes [21]. It was also consistent with the finding that HIV-1 preferentially integrates into transcriptionally active regions of the host genome [22-25]. Thus, at the early phase of the HIV-1 life cycle, EED might play a role in targeting the regions of proviral DNA integration into the host chromatin. At the late steps of the virus replication cycle, we found that overexpression of isoforms EED3 and EED4 had a significant negative effect on virus production, and that virus assembly and genome packaging were the major targets of this EED inhibitory activity [26].

The finding that EED was an interactor of three HIV-1 components and an intracellular factor possibly involved in antiviral innate immunity prompted us to analyse the three-dimensional (3D) structure of EED. Crystallogene- sis of EED was therefore undertaken to better understand the nature of the multiple interactions and functions of EED in the HIV-1 life cycle. Unfortunately, none of our attempts to obtain diffracting crystals of EED alone, or in complex with its viral partners MA, IN or Nef was success-

ful, and we therefore analyzed the 3D structure of EED using indirect approaches. They consisted of (i) three-dimensional modelling based on computer-assisted methods of sequence alignment and determination of homology with a prototype of seven-bladed  $\beta$ -propeller protein previously crystallized [27,28]; (ii) mapping of accessible regions of the EED protein, using anti-EED antibodies and a phage display technique.

During the completion of this work, crystallographic data of the EED protein core, co-crystallised with a peptide from the N-terminal domain of EZH2, was deposited in the protein data bank (PDB code #2QXV) [29] and later published [30]. Our predictive model determined by indirect methods was in good consistency with the crystal structure of EED, except for the region 267-295 which comprises a unique  $\alpha$ -helix facing a short  $\beta$ -strand in the crystallographic structure. Major immunogenic regions in EED were found to correspond to flexible loops and  $\beta$ -strands which were accessible at the surface of the  $\beta$ -propeller. In addition, EED modelling suggested that HIV-1 MA and IN bound to two contiguous sites forming a continuous protein-interacting domain localized in a groove running along the lateral face of the EED  $\beta$ -propeller.

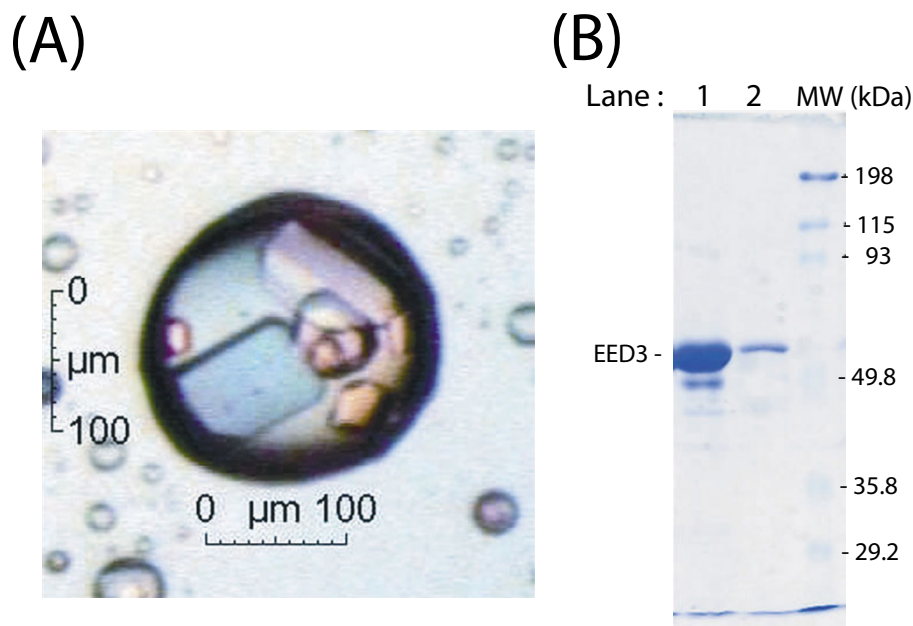
## Results and Discussion

### EED Crystallogenesis

The coding sequence for the His-tagged EED protein of 441 residues representing isoform 3 (EED3-H<sub>6</sub>) was expressed in *E. coli* [16]. EED3 corresponded to the sequence spanning residues Met95-Arg535 in the EED1 isoform [12]. EED3-H<sub>6</sub> protein was found to be highly soluble and was purified to homogeneity, using affinity chromatography followed by a gel filtration step. Solutions of EED3-H<sub>6</sub> titrating 5 to 10 mg/ml were subjected to more than a thousand of different conditions for crystallization. EED3-H<sub>6</sub> protein crystals, appearing as thin platelets of  $0.1 \times 0.07 \times 0.01$  mm<sup>3</sup>, were observed after 30 days at 19°C under certain buffer conditions (0.1 M MES buffer, pH 6.0, 40 % MPD ; Fig. 1A). One single crystal was removed from the well-buffer, washed and dissolved in SDS-sample buffer. SDS-PAGE analysis showed that this crystal was really constituted of EED3-H<sub>6</sub> protein (Fig. 1B; lane 2). However, the crystals obtained under these conditions failed to generate X-ray diffraction patterns. We then tried to co-crystallise EED3-H<sub>6</sub> with its viral protein partners MA, IN or Nef, respectively, but all these attempts were unsuccessful. We then used alternative methods for structure determination of EED, as described below.

### 3D-modelling of the EED core domain

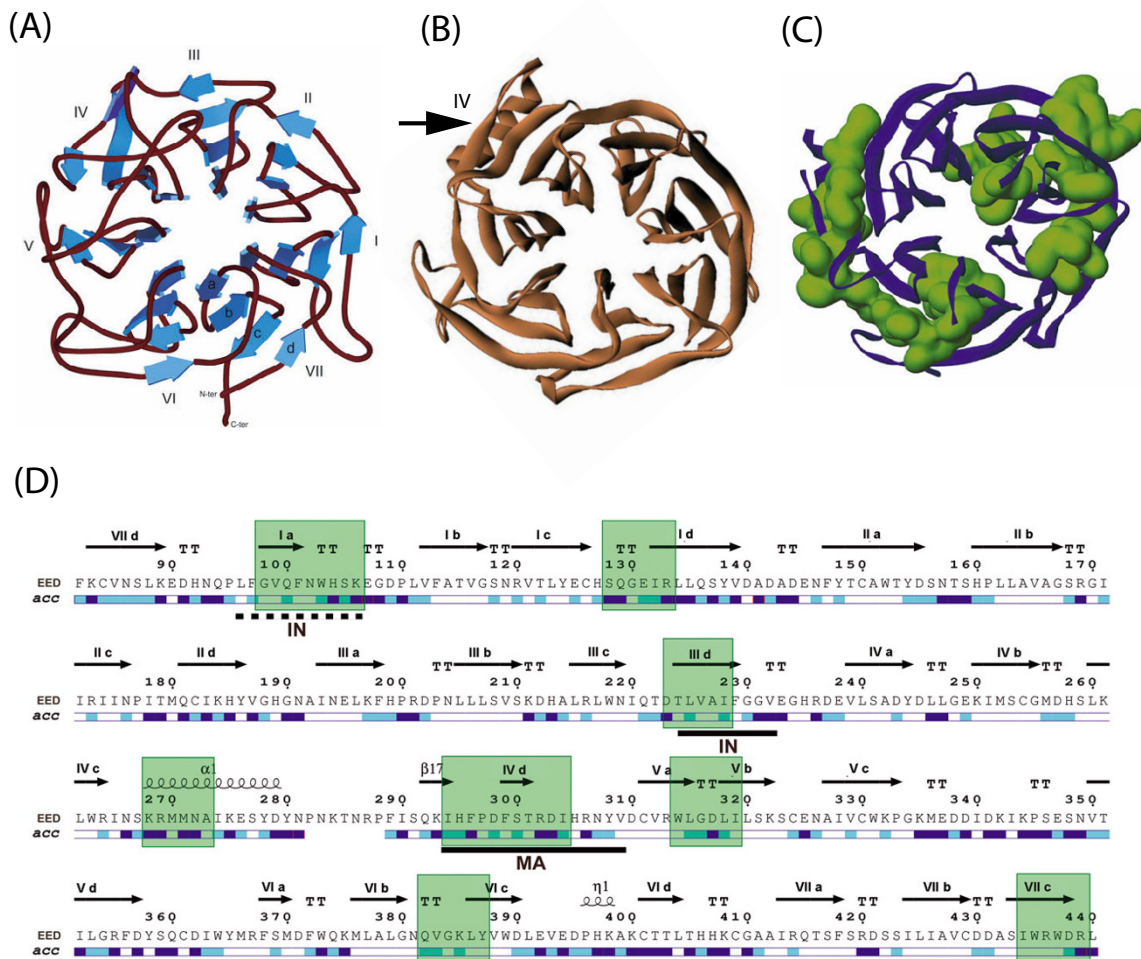
Seven repeats of a canonical WD-motif have been identified in the C-terminal core of the EED protein, shared by the four isoforms EED1, EED2, EED3 and EED4 [12,15]. It was therefore possible to build a three-dimensional

**Figure 1**

**Crystallogenes of histidine-tagged isoform 3 of EED.** (A), Platelet-like crystals of EED3-H<sub>6</sub> protein of 441 residues, obtained in suspended drop in 0.1 M MES buffer pH 6.0, 40 % MPD. (B), Solubilization of the crystals and analysis by SDS-PAGE and Coomassie blue staining. Lane 1 : solution of purified EED3-H<sub>6</sub> (10 mg/mL) used for crystallogenesis (protein load : 50 μg). Lane 2 : protein content of solubilized single crystal. MW : markers of protein molecular mass, indicated in kilodaltons (kDa) on the right side of the panel.

model for the C-terminal core domain of EED spanning residues 84–441 (roughly corresponding to the EED3 isoform), using homology modelling by sequence alignment and homology with protein(s) of known structures, and assessment of accessible motifs and epitopes at the surface of the EED protein. The template used was the  $\beta$  subunit of the bovine signal-transducing G protein ( $G\beta$ ), of which crystal structure has been determined [27,28]. However, due to the limited degree of identity between their primary structures (only 20 % amino acid residues identical between EED3 and  $G\beta$ ), the sequence alignment of both proteins was manually optimized to improve the correspondence between consensus residues in the WD repeats. The model obtained for EED3 corresponded to a typical seven-bladed  $\beta$ -propeller structure (Fig. 2A). Each WD-40 repeat was folded as 3  $\beta$ -strands referred to as a, b and c, respectively. The sequence connecting every WD-40 repeat also folded as an additional  $\beta$ -strand, called d. Thus, a WD-40 repeat formed a structural unit made of 4 antiparallel  $\beta$ -strands referred to as  $\beta$ -blade, and the seven  $\beta$ -blades defined in EED were folded as a  $\beta$ -propeller structure (Fig. 2A).

Our  $\beta$ -propeller model was confirmed by the refined crystal model of the EED core domain recently published [30], and depicted in Fig. 2 (panels B and D). EED and EZH2 proteins are partners involved in PRC2/3 complexes, along with SUZ12 and RbAp46/RbAp48 [13]. The proposed structure represented the co-crystallized complex of a fragment of the N-terminal domain of EZH2 (residues 39 – 68) with the C-terminal domain of EED (residues 82 – 440). EED-EZH2 interaction took place via the insertion of both ends of the EZH2  $\alpha$ -helical peptide into two peptide-binding hydrophobic pockets in EED formed by the side chains of V112, L123, W152 and P161, and by residues L318, L353, L391 and P396, respectively [30]. The 3D structure reconstructed from crystallographic data was globally similar to our 3D-model of seven  $\beta$ -bladed propeller, with three exceptions. In  $\beta$ -blade IV, there were two structures at the junction of  $\beta$ -strands IVc and IVd that were unique among representatives of WD-40 proteins, (i) an  $\alpha$ -helix encompassing region 267–280 ( $\alpha$ 1) and (ii) an outer  $\beta$ -strand referred to as  $\beta$ 17. (iii) In  $\beta$ -blade VI, a short  $3_{10}$ -helix (termed  $\eta$ 1) was found on the N-terminal side of  $\beta$ -strand VIc (Fig. 2D).

**Figure 2**

**Structural models and immunogenic regions of EED isoform 3.** (A), Seven-bladed  $\beta$ -propeller model of the EED core domain, based on sequence homology with the beta subunit of the bovine G protein ( $G\beta$ ; [27, 28]). Shown is a ribbon representation of the polypeptide backbone atoms of EED3 isoform (amino acid residues 84–441), with secondary and tertiary structures of the different  $\beta$ -blades. (B), 3D-model of the EED3 seven-bladed  $\beta$ -propeller, deduced from crystallographic data (modified, from [30]). The black arrow indicates the major difference between our putative model (A) and the crystal model (B), consisting of the  $\alpha$ I helical region facing the  $\beta$ -strand  $\beta$ 17 in  $\beta$ -blade IV. (C), Position of immunogenic epitopes (depicted in green) on the 3D-model of EED polypeptide backbone (represented in blue). (D), Primary and secondary structures of EED3, deduced from crystallographic data [30]. The amino acid sequence was numbered according to the accepted nomenclature [12]: Met95 in EED1 isoform represented MetI in EED3; thus, the C-terminal residue L440 in EED3 corresponded to L535 in EED1. Regions in  $\beta$ -strand structure are represented by horizontal arrows, with reference to the blade number and  $\beta$ -strand letter a, b, c or d;  $\alpha$ -helices are represented by spirals, and turns by TT. Helical regions marked  $\alpha$ I and  $\eta$ I, and the  $\beta$ -strand region marked  $\beta$ 17, were structured domains of EED which were unique among representatives of WD-40 proteins. The relative accessibility of each residue (acc) in the 3D structure was extracted from the dictionary of protein structure [45], and indicated as coloured bars under the sequence with the following colour code: dark blue, highly accessible; light blue, accessible; white, buried. Discrete regions recognized by anti-EED IgG are indicated by green boxes. The binding sites of HIV-1 matrix protein (MA) and integrase (IN) are underlined by solid black lines.

### Surface-exposed regions in the seven-bladed $\beta$ -propeller domain of EED

#### Theoretical considerations

An important feature of the  $\beta$ -propeller structure of the EED core was that most of the accessible surfaces should

be confined to the outer  $\beta$ -strands d, and to flexible loops connecting  $\beta$ -strands of the same blade (Fig. 2D). These accessible regions would be potential sites of protein-protein interaction, as shown by X-ray diffraction analysis of protein complexes involving other  $\beta$ -propeller proteins

[31,32]. E.g. in the case of bovine G $\beta$  protein, several residues belonging to loops d-a and b-c were found to be involved in hydrophobic contacts with the  $\alpha$  subunit [32]. These accessible regions would also contain putative immunogenic epitopes, responsible for the induction of EED antibodies in animals in response to administration of human EED. The next experiments were designed to test this hypothesis.

#### *Mapping of accessible regions in EED using anti-EED antibodies and phage biopanning*

We raised in rabbit a highly reactive and specific antiserum against affinity chromatography-purified, His-tagged EED3 isoform also used for crystallization trials. IgG were isolated from this antiserum and used in a screen with recombinant phages [15,16]. We reasoned that anti-EED IgG immobilized on a solid support would preferentially bind to phages which are mimotopes of accessible and antigenic regions of EED [33,34]. The phages selected on these anti-EED antibodies mapped to eight discrete regions on EED, numbered 1 to 8 (Fig. 2C, D). These regions spanned residues 98–106 (1), 128–133 (2), 223–228 (3), 268–273 (4), 294–305 (5), 314–319 (6), 382–387 (7) and 434–439 (8), respectively. Five of these regions, n° 2, 4, 5, 6 and 7, corresponded to predicted accessible loops separating  $\beta$ -strands in the EED 3D-structure (Fig. 2C, D). Interestingly, the N-terminal portion of the  $\alpha$ -helix K268-D279 coincided with the immunogenic region 4 that we identified (Fig. 2D).

The question raised however, for the accessibility of three regions, numbered 1, 3 and 8, which were partially or totally folded as  $\beta$ -sheet (Fig. 2B–D). Region 1 overlapped with  $\beta$ -strand Ia and the adjacent loop a-b forming the junction with  $\beta$ -strand Ib (Fig. 2B–D). Its motif 103-WHS-105 was included in the EZH2 binding site, and was accessible in a groove oriented towards the lower face of the propeller [30]. Likewise, the reactivity of region 3, which coincided with  $\beta$ -strand IIId, was in good consistency with the 3D-model, as it was oriented outwards and accessible at the surface of the  $\beta$ -propeller. However, our data concerning region 8 were more intriguing : this region corresponded to the  $\beta$ -strand VIIc which was close to the C-terminus and was accessible to antibodies in our experimental screening. This suggested that EED in solution adopted a 3D structure which was less tightly closed than shown in the 3D model. Thus, our mapping of major immunogenic regions of EED was in good consistency with the position of accessible loops and surface exposed portions of  $\beta$ -stands predicted by the EED 3D-model.

#### *3D structure and protein interacting regions in the EED core domain*

The binding site of the HIV-1 MA protein has been mapped to position 294–309 on the linear sequence of EED [15]. The newly established conformation of this

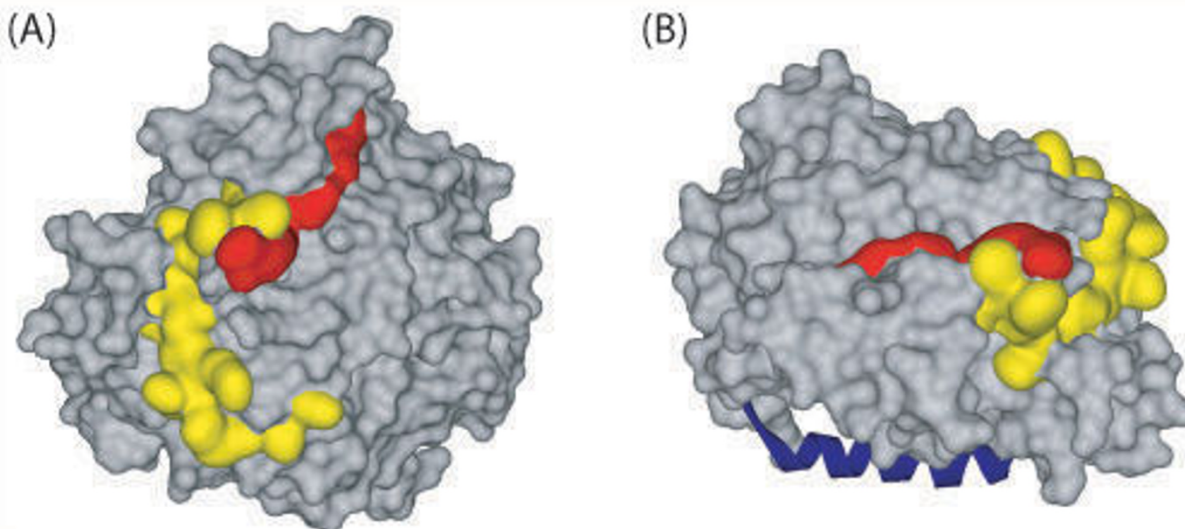
region implied that the region of interaction with the MA protein was not only confined to the flexible loop IVd-Va on the upper face of the  $\beta$ -propeller, but also included the short, rigid  $\beta$ -strand IVd and the neighboring loop IVd-Va, located on the lateral face of the  $\beta$ -propeller. This was not contradictory to our mapping of the MA binding site on the EED linear sequence, since  $\beta$ -strands d were the most exposed  $\beta$ -strands at the periphery of the  $\beta$ -propeller (Fig. 2D). Of note, the upper face of the  $\beta$ -propeller was narrower in surface, compared to its lower face.

However, there was some ambiguity in the determination of the IN binding domain in EED, as two potential binding sites (bs) were identified by phage display, one at position 96–105 (bs1), the other one at position 224–232 (bs2) [16]. In the light of the EED crystal structure, it appeared that in bs1, residues 97–102 were buried in the  $\beta$ -propeller central tunnel, and amino acids 103–105 were part of the groove on the lower face of the EED  $\beta$ -propeller which homed the N-terminal fragment of EZH2 in co-crystals [30]. F96 was the only residue of bs1 which was oriented upwards and accessible on the top of EED. By contrast with bs1, bs2 mapped to the  $\beta$ -strand IIId and the neighbouring turn included in loop IIId-IVa (Fig. 2D). This region lied at the periphery of the  $\beta$ -propeller and was therefore highly accessible, as determined from crystallographic data (Fig. 2B, D).

It was therefore difficult to conceive how one single IN molecule could bind simultaneously to bs1 and bs2, as these sites were far from each other and in different orientation with respect to the  $\beta$ -propeller plane. Although the possibility existed that one molecule of EED would bind to two IN molecules (e.g. dimeric or tetrameric forms of IN), this was unlikely for the following reasons : (i) only one single EED-binding site has been identified in the HIV-1 IN sequence [16], and it is unlikely that the same IN motif would bind to two different sequences in EED ; (ii) mutant EED-103A3, in which the tripeptide motif 103-WHS-105 was replaced by the tripeptide AAA, was still binding to IN with significant efficiency [16]. Taken together, these results suggested that region 224–232 (bs2) was the most probable and unique binding site for IN on EED.

Although the IN and MA binding sites were found to be located at significant distance from each other on the EED linear sequence (224–232 and 294–309, respectively; Fig. 2D), they appeared to be in close vicinity in the 3D structure : both were located on the lateral face of the EED  $\beta$ -propeller, as shown by surface representation, but they did not overlap (Fig. 3A). This was corroborated by the absence of competition between MA and IN for binding to EED3-H $_6$  protein in vitro in histidine pull-down assays (data not shown). In addition, the possibility of occur-





**Figure 3**

**Surface representation of the  $\beta$ -propeller domain of EED and protein-interacting regions.** The binding residues of HIV-1 proteins are represented with the following colour code : yellow for the matrix protein (MA), red for integrase (IN).

**(A)**, Top view of the  $\beta$ -propeller showing the MA and IN binding sites laterally oriented. Note the absence of overlapping between the MA and IN binding sites, which form a continuous binding groove. **(B)**, Side view of the  $\beta$ -propeller showing the MA+IN-binding groove on the lateral face, and the position of the EZH2  $\alpha$ -helical peptide 39–68 (represented in blue), bound to the EZH2-binding pocket facing downwards.

rence of ternary complex involving EED, MA and IN has previously been suggested by their colocalization observed by immuno-electron microscopy of HIV-1-infected cells at early steps of the virus life cycle [16].

The EZH2 binding groove, which was oriented downwards with respect to the EED  $\beta$ -propeller plane, was totally independent of the continuous MA-IN binding groove (Fig. 3B). Interestingly, although  $\alpha$ -helices represent privileged domains of protein-protein interaction, none of the newly identified helices in the EED core,  $\alpha 1$  or  $\eta 1$ , represented binding domains of known cellular or viral partners of EED, e.g. EZH2 [30], MA [15], or IN [16].

## Conclusion

The refined structural model of the EED C-terminal core as a seven-bladed  $\beta$ -propeller determined from crystallographic data provided structural support to our mapping of immunogenic epitopes recognized by our anti-EED polyclonal antibodies, and of the binding sites of HIV-1 MA and IN [15,16]. Several immunoreactive regions coincided with the MA, IN and EZH2 binding sites, confirming the accessibility of these regions at the surface of EED. According to the EED 3D-model, the domain of interaction with the HIV-1 MA protein would be localised on the lateral face of the  $\beta$ -propeller, and be comprised of two

loops separated by the short  $\beta$ -strand IVd (Fig. 2 and Fig. 3). The region of interaction with IN would be assigned to  $\beta$ -strand IIId and its neighboring turn, also located on the peripheral area of the  $\beta$ -propeller. When represented on the surface of the EED molecule, the two discrete prints of MA and IN interaction were contiguous but did not overlap, and formed a continuous protein-interacting groove running along the lateral face of the EED  $\beta$ -propeller. This groove slightly opened towards the lower face of the  $\beta$ -propeller (Fig. 3). The absence of overlapping of the MA and IN binding sites and the possible occurrence of ternary complex involving EED, MA and IN raised the issue of the biological parameters of a simultaneous binding of EED to MA and IN, and of the role that such a ternary complex might play in the HIV-1 life cycle.

## Methods

### Plasmids, proteins and cells

Plasmids coding for GST-fused or His-tagged proteins EED, MA, IN and Nef and protein expression in bacterial cells have been described in previous studies [15,16,26,35].

### EED crystallogenesis

The commercial kits used (Crystal screen 1 and 2 ; Grid Screen Ammonium sulfate ; Grid Screen Sodium Chloride

; Grid Screen MPD ; Grid Screen PEG 6000 ; Grid Screen PEG/LiCl ; Natrix ; SaltRX ; Index & PEG ; Ion Screen) were purchased from Hampton Research (Aliso Viejo, CA, USA) and NeXtal (PEG Suite ; Anions & Cations ; Qiagen SA). The screen was carried out in 96-well plates (Greiner Bio-One GmbH, Divison BioScience, Les Ulis, Courtaboeuf, France), using the hanging-drop vapor-diffusion method. The drops were generated using the Mosquito® Crystal technology (TTP LabTech Ltd, Melbourne, Hertfordshire, UK). Crystals were obtained by the vapor diffusion method from a solution containing 0.1 M 2-(N-morpholino)ethanesulfonic acid (MES) buffer, pH 6.0, 40 % 2-Methyl-2,4-pentanediol (MPD) at a temperature of 292 K. A 1:1 ratio of protein to reservoir solution was used.

### Protein purification

Isolation of His-tagged proteins from bacterial cell lysates was carried out as follows. *E. coli* BL21 DE3 transformed with pT7.7 plasmid [16] were lysed by resuspension in TBS containing lysozyme (1 mg/mL) and a cocktail of protease inhibitors (Roche Diagnostics Corp., Meylan, France), with five cycles of ultrasonication (20 sec each). Cell debris were removed by centrifugation at 10 krpm for 30 min at 4°C in the Biofuge centrifuge (Heraeus, Kendro Laboratory Products, IMLAB Sarl, Lille, France). Affinity purification of His-tagged protein was performed on HiTrap column (1 mL total volume ; GE Healthcare BioSciences, Saclay, France). The column was first loaded with  $\text{Ni}_2^+$  (0.1 M  $\text{Ni}_2\text{SO}_4$ ) prior to affinity chromatography, using a high-performance liquid chromatography (HPLC) system (BioLogic DuoFlow ; Bio-Rad France, Marnes-la-Coquette, France). Proteins were adsorbed in 50 mM Tris-HCl buffer, pH 7.5, 150 mM NaCl (TBS) containing 20 mM imidazole, and eluted with TBS containing 1 M imidazole. Further purification was achieved by gel filtration chromatography, using a Superdex-200 column (GE Healthcare Bio-Sciences, Saclay, France) equilibrated in TBS buffer.

### Antibodies and immunological analysis

Anti-EED rabbit antiserum was laboratory-made. Affinity chromatography-purified, His-tagged EED3 isoform was used as the immunogen. Anti-oligohistidine tag polyclonal antibody was purchased from Qiagen SA (Courtaboeuf, France). For isolation of anti-EED IgG, rabbit antiserum against EED (1 mL) was precipitated by ammonium sulfate at 33 % saturation, and pH 6.5. The IgG precipitate (12–15 mg) was resuspended in TBS (1 mL) and adsorbed on protein G-Sepharose gel. IgG elution was carried out with two gel volumes of 0.1 M Tris-glycine pH 2.2, and the eluate dialyzed against TBS. Proteins were analyzed by electrophoresis in SDS-containing 12 % polyacrylamide gels in the discontinuous Laemmli's buffer system (SDS-PAGE) and Coomassie blue staining, or

Western blotting using the above-mentioned antibodies, as previously described [16,26].

### Phage biopanning

Biopanning of the 6-mer phage library and the ligand elution technique have been described in detail in previous studies [15,16,33,34,36]. In brief, for identification of antigenic regions on the EED protein, recombinant bacteriophages were adsorbed onto anti-EED IgG coated on plates. After extensive rinsing, phages were recovered by three successive cycles of acid buffer elution, followed by final elution by affinity chromatography-purified EED-His<sub>6</sub> protein used as competing ligand [33]. Phagotopes were determined by DNA sequencing.

### Protein homology modelling

The choice of the protein print for EED was determined by sequence comparison using the CLUSTALW program [37] and the PDB [29]). The beta chain of the bovine G protein ( $\text{G}\beta$ ), a WD motif-containing protein, was then obtained (PDB code #1TBG). After preliminary sequence alignment of EED with  $\text{G}\beta$ , alignment was optimized using the following programs : MLRC [38], DSC [39] and PHD [40], all of them available on the NPS@ server [41]. The construction of the 3D-model of EED from the  $\text{G}\beta$  structure was carried out by substitution of the amino acid side-chains using the CALPHA program [42]. Reorientation of the side-chains as well as construction of reinserted polypeptide chain fragments were both performed using the TURBO-FRODO program [43]. Final optimization of the EED 3D-model was achieved using the conjugated gradient method and the CNS program [44].

### Competing interests

The author(s) declare that they have no competing interests.

### Authors' contributions

SV and DR performed the laboratory work and contributed equally to this study. PB and SSH conceived the strategies and designed the experiments. PG contributed to EED homology modelling and data analysis. PB wrote the manuscript. All authors read and approved the final manuscript.

### Acknowledgements

This work has been supported by the Agence Nationale de Recherche sur le SIDA ANRS, AC14-2 and Grant AO2007-DendrAde). SV was the recipient of a fellowship from ANRS (2003–2005). DR was financially supported by ANRS (2004–2006) and by SIDACTION (2007), successively. We are grateful to Cathy Berthet for her efficient secretarial aid.

### References

1. Ng J, Li R, Morgan K, Simon J: **Evolutionary conservation and predicted structure of the *Drosophila* extra sex combs repressor protein.** *Mol Cell Biol* 1997, **17**:6663-6672.

2. Schumacher A, Faust C, Magnuson T: **Positional cloning of a global regulator of anterior-posterior patterning in mice.** *Nature* 1996, **383**:250-253.
3. Schumacher A, Lichtarge O, Schwartz S, Magnusson T: **The murine polycombgroup *eed* and its human orthologue: functional implications of evolutionary conservation.** *Genomics* 1998, **54**:79-88.
4. Simon J: **Locking in stable states of gene expression: transcriptional control during *Drosophila* development.** *Curr Opin Cell Biol* 1995, **7**:376-385.
5. Sewalt RGAB, van der Vlag J, Gunster MJ, Hamer KM, den Blaauwen JL, Satijn DPE, Hendrix T, van Driel R, Otte AP: **Characterization of interactions between the mammalian Polycomb-group proteins *Enx1/EZH2* and *EED* suggests the existence of different mammalian Polycomb-group protein complexes.** *Mol Cell Biol* 1998, **18**:3586-3595.
6. Neer EJ, Schmidt CJ, Nambudripad R, Smith TF: **The ancient regulatory-protein family of WD-repeat proteins.** *Nature* 1994, **371**:297-300.
7. Denisenko O, Bomsztyk C: **The product of the murine homolog of the *Drosophila* extra sex combs gene displays transcriptional repressor activity.** *Mol Cell Biol* 1997, **17**:4707-4717.
8. Rietzler M, Bittner M, Kolanus W, Schuster A, Holzmann B: **The human WD repeat protein WAIT-1 specifically interacts with the cytoplasmic tails of b7-integrins.** *J Biol Chem* 1998, **273**:27459-27466.
9. Crowe DT, Chiu H, Fong S, Weissmann IL: **Regulation of the avidity of integrin alpha4 beta7 by the beta7 cytoplasmic domain.** *J Biol Chem* 1994, **269**:14411-14418.
10. Manié SN, Astier A, Wang D, Phifer JS, Chen J, Lazarovits AI, Morimoto C, Freedman AS: **Stimulation of tyrosine phosphorylation after ligation of beta7 and beta1 integrins on human B cells.** *Blood* 1996, **87**:1855-1861.
11. Otte AP, Kwaks THJ: **Gene repression by Polycomb group protein complexes: a distinct complex for every occasion?** *Curr Opin Genet Dev* 2003, **13**:448-454.
12. Kirmizis A, Bartley SM, Kuzmichev A, Margueron R, Reinberg D, Green R, Farnham PJ: **Silencing of human polycomb target genes is associated with methylation of histone H3 Lys27.** *Genes & Dev* 2004, **18**:1592-1605.
13. Kuzmichev A, Jenuwein T, Tempst P, Reinberg D: **Different Ezh2-containing complexes target methylation of histone H1 or nucleosomal histone H3.** *Mol Cell* 2004, **14**:183-193.
14. Cao R, Zhang Y: **SUZ12 is required for both the histone methyltransferase activity and the silencing function of the EED-EZH2 complex.** *Molec Cell* 2004, **15**:57-67.
15. Peytavi R, Hong SS, Gay B, Dupuy d'Angéac A, Selig L, Bénichou S, Benarous R, Boulanger P: **H-EED, the product of the human homolog of the murine *eed* gene, binds to the matrix protein of HIV-1.** *J Biol Chem* 1999, **274**:1635-1645.
16. Violot S, Hong SS, Rakotobe D, Petit C, Gay B, Moreau K, Billaud G, Priet S, Sire J, Schwartz O, et al.: **The human Polycomb group EED protein interacts with the integrase of human immunodeficiency virus type 1.** *J Virol* 2003, **77**:12507-12522.
17. Witte V, Laffert B, Rosorius O, Lischka P, Blume K, Galler G, Stilper A, Willbold D, D'Aloja P, Sixt M, et al.: **HIV-1 Nef mimics an integrin receptor signal that recruits the polycomb group protein Eed to the plasma membrane.** *Mol Cell* 2004, **13**:179-190.
18. Pirrotta V: **Polycomb the genome: Pcg, trxG, and chromatin silencing.** *Cell* 1998, **93**:333-336.
19. Brock HW, van Lohuizen M: **The Polycomb group – no longer an exclusive club?** *Curr Opin Genet Dev* 2001, **11**:175-181.
20. Orlando V: **Polycomb, epigenomes, and control of cell identity.** *Cell* 2003, **112**(5):599-606.
21. Mager J, Montgomery ND, de Villena FP, Magnusson T: **Genome imprinting regulated by the mouse Polycomb group protein Eed.** *Nat Genet* 2003, **33**:433-434.
22. Schroder AR, Shinn P, Chen H, Berry C, Ecker JR, Bushman F: **HIV-1 integration in the human genome favors active genes and local hotspots.** *Cell* 2002, **110**:521-529.
23. Wu X, Li Y, Crise B, Burgess SM: **Transcription start regions in the human genome are favored targets for MLV integration.** *Science* 2003, **300**:1749-1751.
24. Engelman A: **The ups and downs of gene expression and retroviral DNA integration.** *Proc Natl Acad Sci USA* 2005, **102**:1275-1276.
25. Van Maele B, Debyser Z: **HIV-1 integration: an interplay between HIV-1 integrase, cellular and viral proteins.** *AIDS Reviews* 2005, **7**:26-43.
26. Rakotobe D, Tardy J-C, Andre P, Hong SS, Darlix J-L, Boulanger P: **Human Polycomb group EED protein negatively affects HIV-1 assembly and release.** *Retrovirology* 2007, **4**:37.
27. Lambright DG, Sondek J, Bohm A, Skiba NP, Hamm HE, Sigler PB: **The 2.0 Å crystal structure of a heterotrimeric G protein.** *Nature* 1996, **379**:311-319.
28. Sondek J, Bohm A, Lambright DG, Hamm HE, Sigler PB: **Crystal structure of a Gprotein beta gamma dimer at 2.1 Å resolution.** *Nature* 1996, **379**:369-374.
29. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE: **The Protein Data Bank.** *Nucleic Acids Res* 2000, **28**:235-242.
30. Han Z, Xing X, Hu M, Zhang Y, Liu P, Chai J: **Structural basis of EZH2 recognition by EED.** *Structure* 2007, **15**:1306-1315.
31. Chen L, Mathews FS, Davidson VL, Huizinga EG, Vellieux FM, Hol WG: **Three-dimensional structure of the quinoprotein methylamine dehydrogenase from *Paracoccus denitrificans* determined by molecular replacement at 2.8 Å resolution.** *Proteins* 1992, **14**:288-299.
32. Wall MA, Coleman DE, Lee E, Iniguez-Lluhi JA, Posner BA, Gilman AG, Sprang SR: **The structure of the G protein heterotrimer Gi alpha 1 beta 1 gamma 2.** *Cell* 1995, **83**:1047-1058.
33. Hong SS, Boulanger P: **Protein ligands of human Adenovirus type 2 outer capsid identified by biopanning of a phage-displayed peptide library on separate domains of WT and mutant penton capsomers.** *EMBO J* 1995, **14**:4714-4727.
34. Hong SS, Karayan L, Tournier J, Curiel DT, Boulanger PA: **Adenovirus type 5 fiber knob binds to MHC class I alpha2 domain at the surface of human epithelial and B lymphoblastoid cells.** *EMBO J* 1997, **16**:2294-2306.
35. Fournier C, Cortay J-C, Carbonnelle C, Ehresmann C, Marquet R, Boulanger P: **The HIV-1 Nef protein enhances the affinity of reverse transcriptase for RNA in vitro.** *Virus Genes* 2002, **25**:255-269.
36. Huvent I, Hong SS, Fournier C, Gay B, Tournier J, Carriere C, Courcoul M, Vigne R, Spire B, Boulanger P: **Interaction and co-encapsulation of HIV-1 Vif and Gag recombinant proteins.** *J Gen Virol* 1998, **79**:1069-1081.
37. Higgins DG, Thompson JD, Gibson TJ: **Using CLUSTAL for multiple sequence alignments.** *Methods Enzymol* 1996, **266**:383-402.
38. Guermeur Y, Geourjon C, Gallinari P, Deléage G: **Improved performance in protein secondary structure prediction by inhomogeneous score combination.** *Bioinformatics* 1999, **15**:413-421.
39. King RD, Sternberg MJ: **Identification and application of the concepts important for accurate and reliable protein secondary structure prediction.** *Protein Sci* 1996, **5**:2298-2310.
40. Rost B, Sander C: **Prediction of protein secondary structure at better than 70% accuracy.** *J Mol Biol* 1993, **232**:584-599.
41. Combet C, Blanchet C, Geourjon C, Deléage G: **NPS@: network protein sequence analysis.** *Trends Biochem Sci* 2000, **25**:147-150.
42. Esnouf RM: **Polyalanine reconstruction from C $\alpha$  positions using the program CALPHA can aid initial phasing of data by molecular replacement procedures.** *Acta Cryst D* 1997, **53**(6):665-672.
43. Roussel A, Fontecilla-Camps JC, Cambillau C: **CRYSTALLIZE: a crystallographic symmetry display and handling subpackage in TOM/FRODO.** *J Mol Graph* 1990, **8**:86-88.
44. Brünger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL: **Crystallography & NMR system: A new software suite for macromolecular structure determination.** *Acta Crystallogr D* 1998, **54**:905-921.
45. Kabsch W, Sander C: **Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features.** *Biopolymers* 1983, **22**:2577-2637.