

# 1 Characterisation and structure determination of a llama-derived nanobody 2 targeting the J-base binding protein 1

3 **Bart Van Beusekom<sup>a</sup>, Tatjana Heidebrecht<sup>a</sup>, Athanassios Adamopoulos<sup>a</sup>, Alexander Fish<sup>a</sup>, Els  
4 Pardon<sup>bc</sup>, Jan Steyaert<sup>bc</sup>, Robbie P. Joosten<sup>a</sup> and Anastassis Perrakis<sup>a\*</sup>**

5 <sup>a</sup>Biochemistry, Netherlands Cancer Institute, Plesmanlaan 121, Amsterdam, 1066 CX, The Netherlands

6 <sup>b</sup>VIB Center for Structural Biology, Vrije Universiteit Brussel, Pleinlaan 2, Brussels, 1050, Belgium

7 <sup>c</sup>Structural Biology Brussels, Vrije Universiteit Brussel, Pleinlaan 2, Brussel, 1050, Belgium

8 Correspondence email: [a.perrakis@nki.nl](mailto:a.perrakis@nki.nl)

9 **Keywords: nanobody, JBP1, llama, immune system**

10 **PDB reference: 6FPV**

## 11 Synopsis

12 A camelid single domain antibody fragment (nanobody) is shown to have high affinity towards its recognition  
13 target, the J-base binding protein 1 (JBP1). The serendipitous crystallisation of this nanobody alone, and its crystal  
14 structure solution and refinement to 1.64Å resolution are described. Ensemble refinement suggests that on the ligand-free  
15 state one of the complementarity determining regions (CDRs) is flexible while the other two adopt well-defined  
16 conformations.

## 17 Abstract

18 The J-base Binding Protein 1 (JBP1) contributes to biosynthesis and maintenance of base J ( $\beta$ -D-glucosyl-  
19 hydroxymethyluracil), a modification of thymidine confined to some protozoa. Camelid (llama) single domain antibody  
20 fragments (nanobodies) targeting JBP1 were produced for use as crystallization chaperones. Surface plasmon resonance  
21 (SPR) screening identified Nb6 as a strong binder, recognising JBP1 with a 1:1 stoichiometry and high affinity  
22 ( $k_D=30$ nM). Crystallisation trials of JBP1 in complex with Nb6, yielded crystals diffracting to 1.47Å resolution.  
23 However, the asymmetric unit dimensions and molecular replacement with a nanobody structure, clearly showed that the  
24 crystals of the expected complex with JBP1 were of the nanobody alone. Nb6 crystallizes in spacegroup P3<sub>1</sub> with two  
25 molecules in the asymmetric unit; its crystal structure was refined to a final resolution of 1.64Å. Ensemble refinement  
26 suggests that on the ligand-free state one of the complementarity determining regions (CDRs) is flexible while the other  
27 two adopt well-defined conformations.

## 28 1. Introduction

29 In kinetoplastid flagellates such as Trypanosoma, Leishmania and Crithidia, one percent of thymidine is replaced by a  
30 modified base called base-J ( $\beta$ -D-glucosyl-hydroxymethyluracil) (Gommers-Ampt *et al.*, 1993). In Leishmania, 99  
31 percent of J is located in telomeric repeats (van Leeuwen *et al.*, 1996; van Leeuwen *et al.*, 1997; Genest *et al.*, 2007).  
32 The remaining one percent resides in internal chromosomal positions (iJ). In Trypanosomes this iJ is associated with  
33 transcription initiation sites (Cliffe *et al.*, 2010), while in Leishmania iJ is associated with transcription termination (van  
34 Luenen *et al.*, 2012).

35 A two-step mechanism is responsible for the biosynthesis of base-J. In the first step, hydroxymethyluracil (hmU) is  
36 formed by hydroxylation of the 5-methyl group of specific thymines in the genome of kinetoplastida. Second, a glucose

37 molecule is transferred to hmU, creating base-J (Bullard *et al.*, 2014; Iyer *et al.*, 2013; Sekar *et al.*, 2014). Two proteins  
38 are responsible for the first step, J-DNA binding protein 1 and 2 (JBP1, JBP2) that both contain an N-terminal  
39 hydroxylase domain. This hydroxylation reaction, which is dependent on the presence of Fe(II) and 2-oxoglutarate  
40 (2OG) (Cliffe *et al.*, 2012), converting T to hMU, provided the inspiration for suggesting and proving the mechanism for  
41 the conversion of methylcytosin (mC) to hydroxymethylcytosine (hmC) in humans by the TET proteins (Tahiliani *et al.*,  
42 2009; Ito *et al.*, 2010). While the structure of the hydroxylase domain of TET proteins has been resolved (Hu *et*  
43 *al.*, 2013) the structure of the JBPs hydroxylase domain remains elusive.

44 JBP1 specifically recognises J-base containing DNA (J-DNA), through a short ~150 residue autonomous folding unit in  
45 the middle of JBP1: the J-DNA binding domain (JDBD). Equilibrium kinetics experiments revealed that JDBD has a high  
46 affinity for base-J (10nM) and remarkable specificity for J-DNA over normal DNA (~10,000 fold). Pre-steady state  
47 kinetics experiments together with small angle X-ray and neutron scattering (SANS and SAXS) experiments, suggest that  
48 JBP1 has increased flexibility and undergoes a conformational change upon recognition and binding of J-DNA  
49 (Heidebrecht *et al.*, 2011; Heidebrecht *et al.*, 2012). This flexible nature of JBP1 has proven to be the most likely  
50 bottleneck that has prevented us to crystallise JBP1 over countless efforts.

51 To decrease JBP1 flexibility, we decided to use nanobodies as crystallization chaperones. In the early 1990s it was  
52 discovered (Hamers-Casterman *et al.*, 1993) that camelids possess a type of immunoglobulin (Ig) with only one heavy  
53 chain and and three Ig domains, compared to the classical Ig composed of a light chain with two Ig domains and a heavy  
54 chain with four Ig domains. Nanobodies are the recombinant variable domains of camelid heavy-chain antibodies,  
55 produced by immunization of llamas, and have been a valuable tool in structural biology. They are often used as  
56 crystallization chaperones, which will bind with high affinity to target molecules decreasing their flexibility and  
57 providing the necessary crystal contacts (Korotkov *et al.*, 2009; Rasmussen *et al.*, 2007).

58 We have raised nanobodies that bind specifically to JBP1. Here, we present the characterisation of nanobody-6 (Nb6)  
59 which binds JBP1 with high affinity, and its serendipitous and somewhat unfortunate crystallisation in the absence of  
60 JBP1. This is therefore also a cautionary tale on nanobody-aided crystallisation: although it rarely happens, excellent  
61 nanobodies against their targets can crystallise alone. The structure of Nb6 has been refined, and we present the ensemble  
62 refinement which showed that only one of the complementarity determining regions (CDRs) is flexible while the other  
63 two adopt well-defined conformations.

## 64 2. Materials and methods

### 65 2.1. Macromolecule production

66 Constructs encoding for nanobodies were transformed in *Escherichia coli* (WK6 strain) (Table 1) and were grown and  
67 purified as described (Pardon *et al.*, 2014). In short: cells were grown in terrific broth (TB) media supplemented with  
68 0.1% glucose, 2 mM MgCl<sub>2</sub> and 0.1 mg/mL ampicillin at 37°C, until they reach an optical density of ~1.2. Addition of  
69 1mM IPTG induced nanobody overexpression over a period of 18 hours at 25 °C. Cells were collected by centrifugation  
70 at 4000 g for 15 min. Osmotic shock was performed by adding 15 mL of TES buffer (200 mM Tris/HCl pH 8.0, 500 mM  
71 sucrose, 0.5 mM EDTA) to the pellet and incubating at 4 °C with agitation overnight. Next, 30 mL of a 4x diluted TES  
72 buffer was added to the culture followed by incubation on ice for 60 min. The culture was centrifuged at 9000 g for 40  
73 min and the supernatant was subjected to immobilized metal affinity chromatography using Ni-chelating sepharose beads.  
74 Binding to the beads was performed for 60 min at 4°C followed by two washing steps, using 10 mL Wash Buffer 1 (50  
75 mM Na phosphate pH 7, 1M NaCl) and 30 mL Wash Buffer 2 (50 mM Na phosphate pH 6, 1M NaCl). Elution was  
76 performed in 5 mL of Elution-Buffer (50 mM Na phosphate pH 7, 150 mM NaCl, 300 mM Imidazole). Imidazole was

77 removed by dialysis and fractions containing nanobodies, were analyzed using SDS-PAGE.

78 JBP1 with a strep-tag was purified as previously described (van Luenen *et al.*, 2012).

## 79 2.2. Surface plasmon resonance

80 The binding affinity towards JBP1 was determined by surface plasmon resonance using a Biacore T200 (Life Sciences  
81 General Electric). JBP1 containing a strep tag was immobilized on a Series S Sensor Chip SA (Life Sciences General  
82 Electric). The experiment was performed in a buffer consisting of 20 mM HEPES/HCl (pH 7.5), 140 mM NaCl, 0.05%  
83 Tween-20, 1mM TCEP, and 1mg/mL BSA. Increasing concentration of Nb6 was injected consecutively over the chip.  
84 The data was analyzed using the Biacore T200 Evaluation software and GraphPad Prism 7 (GraphPad Software Inc.).  
85 Equilibrium dissociation constant ( $K_D$ ) was calculated using equation  $RU=B_{max} * Conc / (Conc + K_D)$ .

## 86 2.3. Crystallization, data collection and processing

87 The protein-nanobody complex was set up in a 96-well 3-drop Swissci plates sitting-drop format (Table 2). The drops  
88 consisted of 100 nL protein/nanobody complex with 100 nL of reservoir. After two days crystals appeared at 277 K in  
89 25% w/v PEG3350 and 0.3 M citric acid (pH 3.5). For vitrifying the crystals for data collection 30% v/v glycerol was  
90 added as cryoprotectant. Diffraction data were collected at 100K at the European Synchrotron Radiation Facility on the  
91 MASSIF beamline (Bowler *et al.*, 2015; Svensson *et al.*, 2015). Details of data collection and processing are given in  
92 Table 3.

## 93 2.4. Structure solution and refinement

94 An initial single chain model was obtained by molecular replacement with *PHASER 2.7* (McCoy *et al.*, 2007) using PDB  
95 entry 3ezj (Korotkov *et al.*, 2009) as template. Subsequently, automated building and refinement was performed by  
96 ARP/wARP (Langer *et al.*, 2008), which built 116 residues in chain A and 103 residues in chain B; and 6 additional  
97 residues that were not assigned to a chain, which were manually attached to the N-terminus of chain B. The model was  
98 further optimized using the PDB-REDO webserver (Joosten *et al.*, 2014), which also completed the missing loops (van  
99 Beusekom *et al.*, 2018). The optimal resolution cut-off was determined at 1.64 Å by PDB-REDO, which does a paired  
100 refinement as described by Karplus and Diederichs (2012).

101 Thereafter, the model was alternately refined with REFMAC5 (Murshudov *et al.*, 2011), manually rebuilt using COOT  
102 (Emsley *et al.*, 2010) and validated with *MolProbity* (Chen *et al.*, 2010). The N-terminus of chain B and the C-termini of  
103 both chains were extended by one residue, yielding a final total of 120 residues in chain B and 119 residues in chain A.  
104 The TLS groups were optimized to two groups per chain: one from the N-terminus up to residue number 90, and the  
105 second from 91 to the C-terminus. Side-chain alternates were modelled for 17 amino acids in total. Additionally, two  
106 main-chain alternates were built for the Gly29 and Gly30 of the first chain. Notably, the conserved cysteine bridge shows  
107 two clearly distinct conformations in chain B.

108 To analyse the conformation flexibility of the CDR loops, the structure was also refined with  
109 *phenix.emsemble\_refinement* (Burnley *et al.*, 2012).

110 Molecular graphics figures were prepared with *CCP4mg* (McNicholas *et al.*, 2011).

## 111 3. Results and discussion

112 Nb6 binds very tightly to JBP1 (Fig. 1): the  $K_D$  was determined to be 30 nM. However, despite the high affinity for its  
113 target Nb6 unexpectedly crystallised without its intended binding partner JBP1, when set-up for crystallisation trials. It is  
114 therefore important to realize that despite very high binding affinity, separate crystallization of the nanobody is perfectly  
115 possible. The overall structure of Nb6 forms the expected immunoglobulin fold of a nanobody (Fig. 2A) (Duhoo *et al.*,

116 2017). The three CDR domains are located at residues 48-56 (CDR1), 71-77 (CDR2), and 119-129 (CDR3). After  
117 superposition, the  $C\alpha$  RMSD between the two chains is 0.56 Å. The RMSD between the CDR1, CDR2, and CDR3 loops  
118 themselves in the different chains is only 0.41 Å, 0.27 Å, and 0.51 Å, respectively. Instead, a  $\beta$ -hairpin (residue 61-64) and  
119 some residues near the termini adopt distinct conformations.

120 The average protein B factor was 30 Å<sup>2</sup>. Surprisingly, the average B factors for the CDR1 and CDR2 are mostly below  
121 average: in chain A and B, CDR1 has average B factors of 31 and 25 Å<sup>2</sup>, respectively. For CDR2, the average B-factor  
122 values were still lower at 21 and 19 Å<sup>2</sup>. CDR3 is much more flexible with average B factors of 47 and 58 Å<sup>2</sup>. The middle  
123 of CDR3 in chain B has the highest B factors in the entire structure model: average B factors of 96 and 110 Å<sup>2</sup> were  
124 found for Tyr125 and Arg126, respectively. The flexibility of CDR3 was further confirmed by ensemble refinement (Fig.  
125 2B), in which this loop adopts a wide range of different conformation. In contrast, CDR1 and CDR2 maintain largely the  
126 same conformation as in the single structure model. Furthermore, the electron density is very well-defined for CDR1 and  
127 CDR2, while density for CDR3 is only clear for the main chain at lower contour levels (Fig. 3).

128 It should be noted that crystal contacts or non-crystallographic contacts may cause loops to appear more rigid than they  
129 are in solution, while intramolecular contacts such as interactions within the  $\beta$ -sheet are also present *in vivo*. In fact, apart  
130 from intramolecular contacts, all three CDR domains are also forming (non-)crystallographic contacts to some extent.

131 CDR1 forms a tight  $\beta$ -hairpin with the outer edges anchored in the  $\beta$ -sheet, which explains why it is ordered in general.  
132 In chain A, the tip is further stabilized by a hydrogen bond between the side-chain of Asn53 and a crystallographically  
133 related Gly63 backbone. This is not the case for chain B; however, there the first residues are stabilized further by strong  
134  $\beta$ -sheet-like hydrogen bonding backbone of Phe49 with Ser38 from chain A.

135 The stabilization of the first half of CDR2 is due to intramolecular  $\beta$ -sheet hydrogen bonding: the largest part of this  
136 variable domain is an extension of the conserved  $\beta$ -strand. In contrast, the last few residues form a  $\beta$ -hairpin and these  
137 residues form intermolecular contacts in the protein crystal. In both chains, extensive crystal contacts are formed for the  
138 last three residues in the CDR (residues 75-77), and the succeeding residues (78-80) even form a  $\beta$ -sheet with a crystal  
139 copy of Nb6.

140 The most flexible loop, CDR3, is far less influenced by crystal contacts. The tip of CDR3 forms some crystal contacts  
141 with its side-chains in chain A; however, these side-chains are not well-ordered so these contacts may not be so strong.

142 The analysis of RMSD, B factors, ensemble refinement and electron density consistently show that CDR1 and CDR2  
143 have a very rigid conformation in Nb6. In contrast, CDR3 is much more flexible and there is only clear density for this  
144 loop at lower contour levels, implying that CDR3 only gets ordered upon binding to its target, JBP1. However, CDR1 and  
145 CDR2 are also involved in (non-)crystallographic contacts, which could cause them to appear more ordered than they are  
146 in solution. The extensive  $\beta$ -sheet interactions between different Nb6 monomers in the protein crystal may explain why it  
147 has crystallized in the absence of JBP1.

#### 148 Table 1

---

149 Source organism	E. coli
150 DNA source	E. coli
151 Forward primer	TTATGCTTCCGGCTCGTATG
152 Reverse primer	CCACAGACAGCCCTCATAG
153 Cloning vector	pMESy4 (Genbank KF415192)
154 Expression vector	pMESy4
155 Expression host	E. coli WK6

---

156 Complete amino acid sequence of the construct produced†  
 MKYLLPTAAAGLLLLAAQPAMAQVQLVESGGGLVQAGGSLRLSCAAS  
 GSFFSINDMGWYRQAPGKQRELVAVISSGGSTN YADSV EGRSTISSD  
 NAKNTVYLQLSSLKPEDTAVYYCNANVRLREYRTTSYHYWGQGTQ  
 VTVSSH HHHHHEPEA

157 †presignal in italics

158 **Table 2**

159 Crystallization

160 Method	Sitting drop crystallisation
161 Plate type	96-well 3-drop Swissci plates from Molecular Dimensions
162 Temperature (K)	277
163 Protein concentration	12 mg/mL complex of JBP1:Nb6
164 Buffer composition of protein solution	140 mM NaCl; 20mM HEPES/NaOH pH 7.5; 1mM TCEP
165 Composition of reservoir solution	0.3 M citric acid pH 3.5; polyethylene glycol 335
166 Volume and ratio of drop	200 nL 1:1
167 Volume of reservoir	50 $\mu$ L

168 **Table 3**

169 Data collection and processing

170 Diffraction source	ESRF beamline massif-1
171 Wavelength ( $\text{\AA}$ )	0.96600
172 Temperature (K)	100
173 Detector	Dectris PILATUS3 2M pixel
174 Crystal-detector distance (mm)	171.65
175 Rotation range per image ( $^{\circ}$ )	0.20
176 Total rotation range ( $^{\circ}$ )	180
177 Exposure time per image (s)	0.1
178 Space group	$P3_1$
179 $a, b, c$ ( $\text{\AA}$ )	57.68, 57.68, 64.75
180 $\alpha, \beta, \gamma$ ( $^{\circ}$ )	90, 90, 120
181 Mosaicity ( $^{\circ}$ )	0.07
182 Resolution range ( $\text{\AA}$ )	49.95–1.47 (1.52–1.47)
183 Total No. of reflections	210220
184 No. of unique reflections	40857
185 Completeness (%)	99.7 (98.4)
186 Redundancy	5.100 (5.10)
187 $\langle I/\sigma(I) \rangle$	14.3 (1.1)
188 $R_{i.m.}$	0.057 (1.541)
189 Overall $B$ factor from Wilson plot ( $\text{\AA}^2$ )	23.2

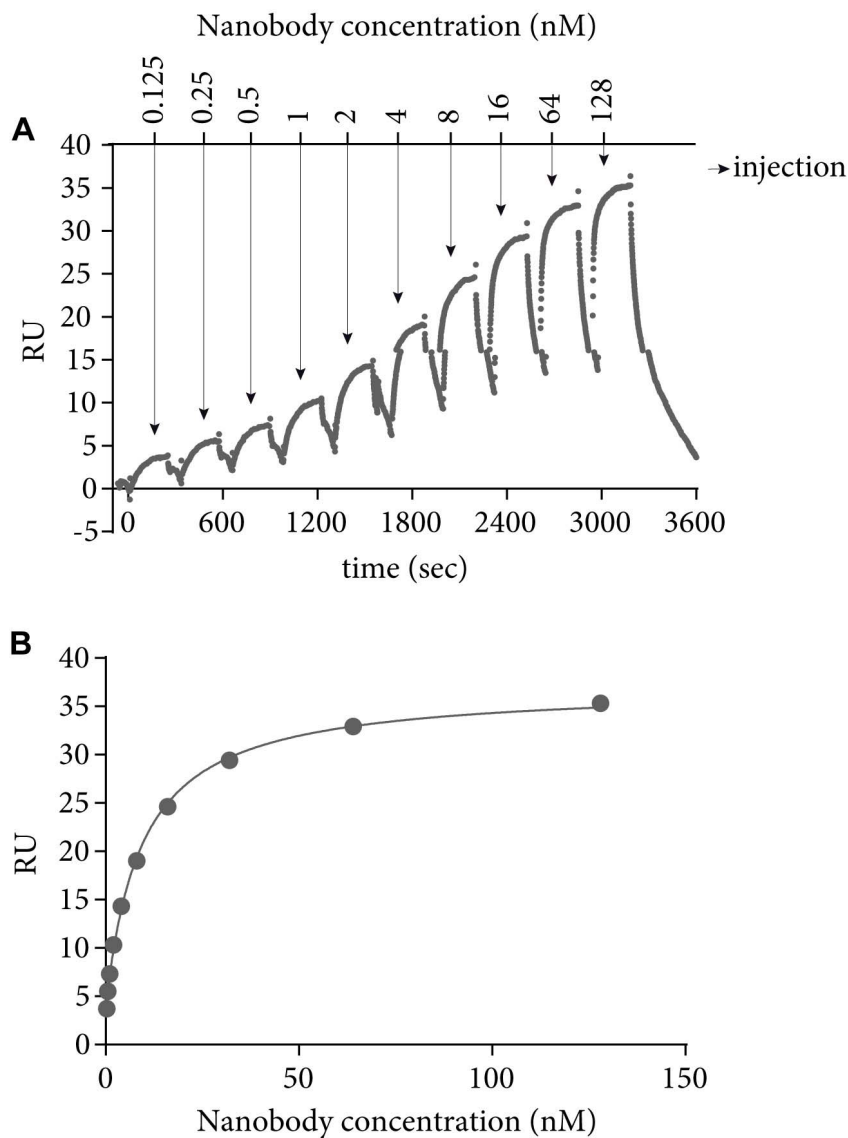
190 **Table 4**

191 Structure refinement

192 Resolution range ( $\text{\AA}$ )	49.95–1.64 (–)
193 Completeness (%)	99.96
194 $\sigma$ cutoff	$F > 0.0\sigma(F)$
195 No. of reflections, working set	28065
196 No. of reflections, test set	1463
197 Final $R_{\text{crist}}$	0.163

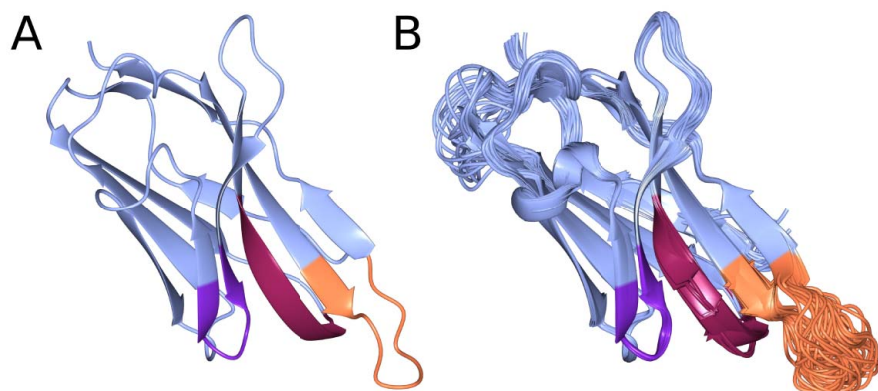
198	Final $R_{\text{free}}$	0.195
199	Cruickshank DPI	0.08
200	No. of non-H atoms: protein / water and solvent / total	1844 / 168 / 2012
201	R.m.s. Z deviations: bonds / angles	0.85 / 0.88
202	Average $B$ factors ( $\text{\AA}^2$ ): protein / water and solvent / total	30.0 / 36.5 / 30.5
203	Ramachandran plot: favoured / allowed / outlier (%)	98.8 / 1.2 / 0.0
204	Molprobit score	1.07 (99 <sup>th</sup> percentile)
205	Molprobit clash score	2.83 (99 <sup>th</sup> percentile)
206	Rotamers: favoured / outlier (%)	99.1 / 0.9

74\_original.tif



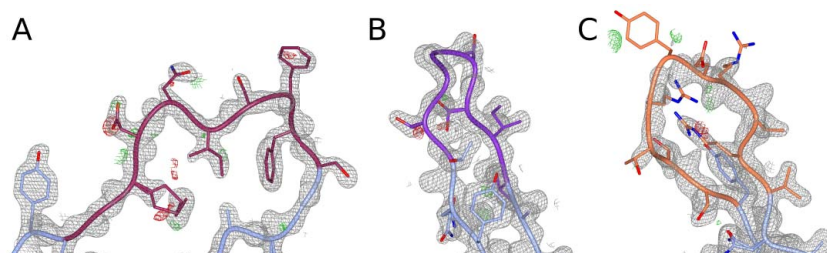
207 **Figure 1**

208 (A) A sensogram showing the subtraction of the empty flow cell from the binding profile of the flow cell with  
209 immobilized JBP1. (B) The equilibrium response versus concentration. The equilibrium binding fit is shown as a black  
210 line.



**Figure 2**

An overview of the nanobody structure model. (A) The nanobody structure model and (B) the same model after ensemble refinement. Chain A is shown in ice blue; chain B is shown in white. The CDR regions are shown in purple-red (CDR1), violet (CDR2), and orange (CDR3). From the ensemble refinement model, it is obvious that CDR1 and CDR2 are in a very fixed conformation, while CDR3 is much more flexible.



**Figure 3**

Details from the CDR regions of chain B: (A) CDR1, (B) CDR2 and (C) CDR3. The  $2mF_o - DF_c$  map is shown at  $1.5 \sigma$  for CDR1 and CDR2, and at  $1.0 \sigma$  for CDR3. The  $mF_o - DF_c$  map is shown at  $3.0 \sigma$  in all cases. The backbone is shown as a tube for clarity; for side-chains, all atoms are shown. The density for CDR1 and CDR2 is very well-defined; for CDR3, however, the contour level had to be decreased to show clear density. Still, there is hardly any density observed for the side-chains in the tip of CDR3. Partly, this can be explained because there are two arginines and a glutamic acid in the tip, which are often flexible side-chains in general.

### Acknowledgements

The authors would like to thank Instruct-ERIC part of the European Strategy Forum on Research Infrastructures (ESFRI), and the Research Foundation Flanders (FWO) for their support to the Nanobody discovery. We thank Katleen Willibal for technical assistance during Nanobody discovery; we would also like to express our deep gratitude to the involved llamas.

### Funding information:

Nederlandse Organisatie voor Wetenschappelijk Onderzoek (grant No. VIDI 723.013.003 to R.P. Joosten; grant No. TOP.13.B2.009 to A. Perrakis).

230 **References**

- 231 van Beusekom, B., Joosten, K., Hekkelman, M. L., Joosten, R. P. & Perrakis, A. (2018). *bioRxiv*,  
232 <https://doi.org/10.1101/329219>.
- 233 Bowler, M.W., Nurizzo, D., Barrett, R., Beteva, A., Bodin, M., Caserotto, H., Delagenière, S., Dobias, F., Flot, D.,  
234 Giraud, T., Guichard, N., Guijarro, M., Lentini, M., Leonard, G. A., McSweeney, S., Oskarsson, M., Schmidt, W.,  
235 Snigirev, A., von Stetten, D., Surr, J., Svensson, O., Theveneau, P. & Mueller-Dieckmann, C. (2015). *J. Synchrotron*  
236 *Rad.* **22**, 1540-1547
- 237 Bullard, W., Lopes da Rosa-Spiegler, J., Liu, S., Wang, Y. & Sabatini, R. (2014). *Journal of Biological Chemistry*, **289**,  
238 20273–20282.
- 239 Burnley, B. T., Afonine, P. V., Adams, P. D. & Gros, P. (2012). *eLife* 1:e00311
- 240 Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J.  
241 S. & Richardson, D. C. (2010). *Acta Cryst.* **D66**, 12–21.
- 242 Cliffe, L.J., Hirsch, G., Wang, J., Ekanayake, D., Bullard, W., Hu, M., Wang, Y. & Sabatini, R. (2012). *Journal of*  
243 *Biological Chemistry*, **287**, 19886–19895.
- 244 Cliffe, L.J., Siegel, T.N., Marshall, M., Cross, G.A.M. & Sabatini, R. (2010). *Nucleic Acids Res.* **38**, 3923–3935.
- 245 Duhoo, Y., Roche, J., Trinh, T. T. N., Desmyter, A., Gaubert, A., Kellenberger, C., Cambillau, C., Roussel, A., & Leone, P.  
246 (2017). *Acta Cryst.* **F73**, 286-293.
- 247 Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* **D66**, 486-501.
- 248 Genest, P.-A., ter Riet, B., Cijssouw, T., van Luenen, H.G.A.M. & Borst, P. (2007). *Nucleic Acids Res.* **35**, 2116–2124.
- 249 Gommers-Ampt, J.H., van Leeuwen, F., De Beer, A.L.J., Vliegthart, J.F.G., Dizdaroglu, M., Kowalak, J.A., Crain, P.F.  
250 & Borst, P. (1993). *Cell*, **75**, 1129–1136.
- 251 Hamers-Casterman, C., Atarhouch, T., Muyldermans, S., Robinson, G., Hammers, C., Songa, E. B., Bendahman, N. &  
252 Hammers, R. (1993). *Nature*, **363**, 446–448.
- 253 Heidebrecht, T., Christodoulou, E., Chalmers, M. J., Jan, S., Riet, ter, B., Grover, R. K., Joosten, R. P., Littler, D., van  
254 Luenen, H., Griffin, P. R., Wentworth, P., Borst, P. & Perrakis, A. (2011). *Nucleic Acids Research*, **39**, 5715–5728.
- 255 Heidebrecht, T., Fish, A., Castelmur, von, E., Johnson, K. A., Zaccari, G., Borst, P. & Perrakis, A. (2012). *J. Am. Chem.*  
256 *Soc.*, **134**, 13357–13365.
- 257 Hu, L., Li, Z., Cheng, J., Rao, Q., Gong, W., Liu, M., Shi, Y. G., Zhu, J., Wang, P. & Xu, Y. (2013). *Cell*. **155**, 1545–1555.
- 258 Ito, S., D'Alessio, A. C., Taranova, O. V., Hong, K., Sowers, L. C. & Zhang, Y. (2010). *Nature*. **466**, 1129–1133.
- 259 Iyer, L.M., Zhang, D., Maxwell Burroughs, A. & Aravind, L. (2013). *Nucleic Acids Res.* **41**, 7635–7655.
- 260 Joosten, R. P., Long, F., Murshudov, G. N., & Perrakis, A. (2014). *IUCrJ*, **1**, 213-220.
- 261 Karplus, P. A. & Diederichs, K. (2012). *Science*, **336**, 1030-1033.
- 262 Korotkov, K. V., Pardon, E., Steyaert, J. & Hol, W. G. (2009). *Structure*, **17**, 255-265.
- 263 Langer, G., Cohen, S. X., Lamzin, V. S. & Perrakis, A. (2008). *Nature protocols*, **3**, 1171.
- 264 van Leeuwen, F., Wijsman, E.R., Kieft, R., van der Marel, G.A., van Boom, J.H. & Borst, P. (1997). *Genes Dev.* **11**, 3232–  
265 3241.
- 266 van Leeuwen, F., Wijsman, E.R., Kuyl-Yeheskiely, E., van der Marel, G.A., van Boom, J.H. & Borst, P. (1996). *Nucleic*  
267 *Acids Res.* **24**, 2476–2482.



- 268 van Luenen, H. G. A. M., Farris, C., Jan, S., Genest, P.-A., Tripathi, P., Velds, A., Kerkhoven, R.M., Nieuwland, M.,  
269 Haydock, A., Ramasamy, G., Vainio, S., Heidebrecht, T., Perrakis, A., Pagie, L., van Steensel, B., Myler, P.J. & Borst,  
270 P. (2012). *Cell*, **150**, 909–921.
- 271 McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.*,  
272 **40**, 658–674.
- 273 McNicholas, S., Potterton, E., Wilson, K. S. & Noble, M. E. M. (2011). *Acta Cryst. D***67**, 386-394
- 274 Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. &  
275 Vagin, A. A. (2011). *Acta Cryst. D***67**, 355–367.
- 276 Pardon, E., Laeremans, T., Triest, S., Rasmussen, S. G. F., Wohlkönig, A., Ruf, A., Muyldermans, S., Hol, W. G. J.,  
277 Kobilka, B. K. & Steyaert, J (2014). *Nature Protocols*, **9**, 674-693.
- 278 Rasmussen, S. G. F., Choi, H.-J., Rosenbaum, D. M., Kobilka, T. S., Thian, F. S., Edwards, P. C., Burghammer, M.,  
279 Ratnala, V. R. P., Sanishvili, R., Fischetti, R. F., Schertler, G. F. X., Weis, W. I. & Kobilka & B. K. (2007). *Nature*, **450**,  
280 383–387.
- 281 Sekar, A., Merritt, C., Baugh, L., Stuart, K. & Myler, P.J. (2014). *Mol Biochem Parasitol*, **196**, 9–11.
- 282 Svensson, O., Monaco, S., Popov, A. N., Nurizzo, D. & Bowler, M. W. (2015). *Acta Cryst. D***71**, 1757-1767
- 283 Tahiliani, M., Koh, K. P., Shen, Y., Pastor, W. A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L. M., Liu, D. R.,  
284 Aravind, L. & Rao, A. (2009). *Science*. **324**, 930–935.