Notum deacylates Wnts to suppress signalling activity

Satoshi Kakugawa[1]*, Paul F. Langton[1]*, Matthias Zebisch[2, 6]*@, Steve Howell[1], Tao-Hsin Chang[2], Yan Liu[5], Ten Feizi[5], Ganka Bineva[4], Nicola O'Reilly[4], Ambrosius P. Snijders[3], E. Yvonne Jones[2@], Jean-Paul Vincent[1@]

SUPPLEMENTARY INFORMATION

Detailed description of crystal growth conditions and cryoprotection

Crystal form I of hNotum appeared over a wide pH range with 10-30% (w/v) PEG2000-6000 as precipitant, but grew best at 100 mM BisTris/HCl pH 6, 20% PEG3350, 200 mM ammonium sulfate (AS) with a protein concentration of 6-12mg/ml. The Heparin oligomer DP06 (iduron) was used in co-crystallization at a concentration of 1mM. Crystals of form II were obtained using full length protein under similar conditions as form I but were not reproducible. Crystal form III appeared in 1.5 M AS, 0.1 M NaAcetate pH 4.6 and sometimes under form I conditions if AS was present. The crystals of form IV grew in 10% (w/v PEG4000), 10 mM $CaCl_2$, 50 mM Sodium Cacodylate pH 6.0 when 1mM of the DP06 Heparin oligo (Iduron) was present. Yet, the structure revealed no electron density for Heparin. Crystal forms V-VII depended on the addition of 20 mM sucrose octasulfate (SOS, Toronto Research Chemicals Inc.) and grew at a protein concentration of 7 mg/ml. Form V appeared in 20% (w/v) PEG3350, 0.200 M AS; form VI in 20% (v/v) glycerol, 24% (w/v) PEG4000, 160 mM $MgCl_2$, 80 mM Tris-HCl pH 8.5 and form VII in 200 mM AS, 100 mM MES 6.5, 35% (w/v) pentaerythritol type PEP 629. The non-deglycosylated surface variant R144S, R145S gave rise to crystal form VIII at protein concentrations of 12.5 g/l and 1.8 M Na/K phosphate pH 6.3. Co-crystallization of hNotum, with a Wnt7-derived palmitoylated peptide (commercially synthesized by Eurogentec) that was used at approximately 100 μM concentration, gave rise to form IX in 3.0 NaCl, 100mM citrate pH 5.0 at a protein concentration of

5.5 mg/mL. Due to limited solubility of the peptide the complex had to be prepared salt free in 10mM Hepes/NaOH pH 7.5. dNotum crystallized in condition D10 of the Morpheus Screen (crystal form I; [63]) and in 100 mM NaCl, 100 mM Li_2SO_4, 100 mM CAPSO pH 9.5, 12.0 % (w/v) PEG 4000 (form II). Cryoprotection was achieved using standard methods and was generally unproblematic. Crystals were transferred to mother liquor supplemented with PEG200 (PEG conditions), glycerol (AS conditions) or sodium malonate (NaCl condition) so that the final water content was lower than 65%. Crystals were then flash-cooled by dipping into liquid nitrogen. Diffraction data were collected at DIAMOND synchrotron light source at the beamlines indicated in Table S1 at the end of this document.

Structure determination

The structure of hNotum was solved with AUTOSOL in PHENIX [64] using SAD data from a crystal of form I that was soaked for 1 day with Mersalyl acid and diffracted to 3.1Å. Four strong sites could be identified that were later found to be binding sites to the free cysteine C330 and H92. Phases were extended to an initial resolution of 2.5Å and a first model was built. However, refinement stalled at approximately 50% model completeness. This problem was overcome by using the initial model to phase a lower resolution structure in crystal form II, which contained four instead of two copies in the asymmetric unit. Solvent modification and NCS averaging then boosted the building and refinement process.

All other structures were solved by molecular replacement with PHASER [65] and completed by manual rebuilding in COOT [66] and refinement with REFMAC5. Models were validated with MOLPROBITY [67]. Residues G127, W128, E390, I391 of hNotum and the corresponding residues of dNotum, which all directly contribute to

the active site, were repeatedly found to lie outside of allowed regions of the Ramachandran plot. Superpositions were performed within CCP4 or COOT using the SSM algorithm. Electrostatics potentials were generated using APBS [68]. The volume of the active site pocket was assessed using CASTp [69]. Figures were produced in PYMOL and assembled in PHOTOLINE32. Diffraction data collection and refinement statistics are provided a the end of this document

Molecular docking

Carboxylic acids were docked into the active site of the high resolution structure of $hNotum_{core}$ in crystal form III using AUTODOCK VINA [70] and LeadIT (BioSolveIT, Germany) with standard settings.

Detailed description of structural and biophysical GAG binding data

From a multitude of crystals grown in presence of 1mM heparin oligo-saccharides we obtained only a single crystal structure which showed clear density for a single disaccharide heparin fragment. The binding site for the heparin fragment is very small including the single residue R115 (Fig. 3c). The mutation R115S did not change the affinity for Heparin (Fig. 3d) suggesting that the observed binding mode is not relevant or that R115 represents only a minor fraction of a larger epitope.

In the crystal form III structure of $hNotum_{core}$ we identified several sulfate ions bound to surface residues. We mutated these binding sites and measured the impact on Heparin binding in SPR and affinity chromatography (Extended Data Fig. 6). The three consecutive basic residues R409, H412 and R416 forming sulfate site 1 (Fig. 3c) lie on the C-terminal end of helix αK. This sulfate binding site is preserved as a phosphate binding site in crystal form VIII. For the knockout variant R409Q H412N

R416Q, the dissociation constant was substantially increased (40-fold) (Fig. 3c,d and Extended Data Fig. 6a). Knockouts of additional sulfate and chloride sites did not show major changes in affinity for Heparin (Extended Data Fig. 6). Two variants (Y297F+R133S and R361S) appeared to have a slightly reduced affinity in SPR (Extended Data Fig. 6b and e) but failed to show a reduced binding in heparin affinity chromatography (Extended Data Fig. 6a).

The use of sucrose octasulfate (SOS) as a chemically defined heparin mimic in crystallization led to the formation of three additional crystal forms. In all three crystal structures SOS binds to the same groove on the Notum surface. This groove is characterized by high electrostatic surface potential, and is located approximately midway between the heparin fragment binding site and sulfate site 1. The density corresponding to SOS could be modelled in two of the three crystal forms. A multitude of hydrogen bonds are formed between five of the eight sulfate groups of SOS and Arg, Lys, His, Asn and Ser side chains of Notum (Figure 3c). Compared to the triple knockout sulfate site 1 variant, the knockout variant including all nine residues involved in binding of the Heparin disaccharide, SOS and sulfate 1 reveals an even more reduced affinity to heparin (Extended Data Fig. 6a). For the crystal grown in the presence of 200 mM KI we noted strong binding of two iodide ions to the SOS binding groove, further highlighting the ability of this groove to attract negative charges (data not shown). Altogether the complex structures with heparin disaccharide, SOS and sulfate 1 define a large patch on the enzyme surface that is centred around a basic groove between the top of the β-sheet and helix αK.

Synthesis of oxidised palmitoleoylated Wnt3A peptide

The following peptide was synthesized: NH2- Cys(**S**-)-His-Gly-Leu-Ser(-C=0-(CH2)7CH=CH(CH2)5CH3)-Gly-Ser-Cys(**S**-)-Glu-Val-Lys-COOH. As a general strategy, palmitoleoylation was carried out with the peptide still attached to the synthesis resin and with N terminal and side chain protection groups intact to ensure palmitoleoylation of the natively modified serine without additional unwanted modification of other hydroxyls, thiols or amines. After cleavage and deprotection the peptide was oxidised with clear-ox^{TM} resin (Peptides International Inc) before purification. The peptide was synthesised on an Intavis Multipep CF peptide synthesiser using continuous flow and uv monitoring of Fmoc deprotection with concomitant extension of coupling where required. HCTU (O-(1H-6-Chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate; Merck) was used as a coupling reagent to couple amino acids on to Fmoc-Lys(Boc)-Wang resin LL (100-200 mesh) 0.31 mmol/g. The N terminal amino acid was incorporated as Boc-Cys(Trt)-OH. Fmoc-Ser-OH was incorporated at the site to be palmitoleoylated . Other amino acids, incorporated as required, were Fmoc-His(Trt)-OH, Fmoc-Gly-OH, Fmoc-Leu-OH, Fmoc-Val-OH, Fmoc-Ser(tBu)-OH, Fmoc-Met-OH, Fmoc-Thr(tBu)-OH, Fmoc-Cys(Trt)-OH, Fmoc-Glu(OtBu)-OH. Following assembly of the peptide chain, palmitoleoylation was achieved by treating 50 umol of peptidyl resin with palmitoleoyl chloride (33 eqs; Santa Cruz Biotechnology) dissolved in 4 mls of DMF (dimethyl formamide). DMAP (dimethylaminopyridine) (0.1eqs) in 1.5 mls pyridine was added to the reaction mixture and the reaction left on a roller overnight. Next the peptidyl resin was washed with 3 x 10 mL DMF, then 2 x 10 mL DCM (dichloromethane) followed by cleavage and deprotection with 10 mL of 92.5% trifluoroacetic acid (TFA), 2.5% ethanedithiol, 2.5% triisopropyl silane, and

2.5% H_2O. After 1 h, the resin was removed by filtration and peptides were precipitated with diethyl ether on ice. Peptides were isolated by centrifugation, then dissolved in H_2O and freeze-dried overnight. Oxidation of the peptides was achieved as follows. 18 umol of each peptide was dissolved in 5.5 ml H20:MeOH (10:1) and added to clear-ox^{TM} resin (3eqs) which had been pretreated by being swollen in DCM for 30 mins, then washed with 10 mL of DCM, MeOH and finally H20. After two days on a roller at RT the peptide solution was removed from the resin and the peptide purified on a C8 reverse phase HPLC column (Agilent PrepHT Zorbax 300SB-C8, 21.2x250 mm). Buffer A was 1% ACN (acetonitrile), 0.08% TFA in H_2O, buffer B was 90% ACN, 0.08% TFA in H_2O. The elution gradient was from 20% to 90% over 30 min at a flow rate of 8 ml/min. The peak fractions were analyzed by LC–MS on an Agilent 1100 LC-MSD.

Synthesis of palmitoylated Hedgehog peptide

The following peptide, Palm-CGPGRGFGKRR-COOH, was synthesised on an Activotec P11 peptide synthesiser. HCTU was used as a coupling reagent to couple amino acids on to Fmoc (Fluorenylmethyloxycarbonyl)-Arg(Pbf)-Wang resin LL (100-200 mesh) 0.37 mmol/g. Amino acids, incorporated where required, were Fmoc-Cys(Trt)-OH, Fmoc-Gly-OH, Fmoc-Pro-OH, Fmoc-Arg(Pbf)-OH, Fmoc-Phe-OH, Fmoc-Lys(Boc)-OH. Following chain assembly and removal of Fmoc from the N-terminal cysteine, 50 umole of peptidyl resin was palmitoylated as follows: palmitic acid (4eqs) was dissolved in 500 uL of DMSO:NMP (4:1) 40 ^{o}C. HOBt (Hydroxybezatriazole) in NMP (n-methyl pyrrolidone, 200 uL, 1M) was added to the solution then after 3 min DIC (diisopropylcarbodiimide, 50 uL) was added. After 30 min the activated solution of palmitic acid was added to the peptidyl resin (pre-

swollen with NMP (1.5 mL)) and reacted overnight at 40 ^{o}C. Next the peptidyl resin was washed with 10 mL each of DMSO:NMP (4:1) , NMP x 3, DCM x 2. Cleavage and deprotection and purification of the peptide was as previously described for the palmitoleoylated peptides. MeOH was used as a solvent prior to purification.

Supplementary Information references

63 Gorrec, F. The MORPHEUS protein crystallization screen. *Journal of Applied Crystallography* **42**, 1035-1042 (2009).

64 Terwilliger, T. C. *et al.* Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. *Acta crystallographica. Section D, Biological crystallography* **65**, 582-601 (2009).

65 McCoy, A. J. *et al.* Phaser crystallographic software. *Journal of Applied Crystallography* **40**, 658-674 (2007).

66 Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta crystallographica. Section D, Biological crystallography* **66**, 486-501 (2010).

67 Chen, V. B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta crystallographica. Section D, Biological crystallography* **66**, 12-21 (2010).

68 Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. Electrostatics of nanosystems: application to microtubules and the ribosome. *Proc Natl Acad Sci USA* **98**, 10037-10041 (2001).

69 Liang, J., Edelsbrunner, H. & Woodward, C. Anatomy of protein pockets and cavities: measurement of binding site geometry and implications for ligand design. *Protein science* **7**, 1884-1897 (1998).

70 Trott, O. & Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry* **31**, 455-461 (2010).

Table S1: Data collection and refinement statistics for hNotum and *d*Notum crystal structures. Values in brackets refer to the highest resolution shell.

PDB code	**4wbh**	**- phasing -**	**4uyu**	**4uyw**	**4uzl**	**4uyz**	**4uz1**	**4uz5**
Species	human	human	human	human	human	human	human	human
Crystal form	I	I	I	I	I	II	III	IV
identifier	apo	mersalylic acid	iodide	heparin disaccharide	myristoleic acid	apo	apo	apo
Data collection								
X-ray source	i04	i03	i24	i04-1	i04-1	i24	i03	i04
wavelength (Å)	0.98	1.90	0.97	0.92	0.92	1.02	0.98	0.98
space group	$P2_12_12$	$P2_12_12$	$P2_12_12$	$P2_12_12$	$P2_12_12$	$P2_1$	$P2_12_12_1$	$P4_132$
unit cell dimensions								
a, b, c (Å)	72.7, 175.4, 60.8	74.5, 181.9, 60.3	74.2, 177.6, 61.0	74.1, 178.6, 61.2	74.1, 170.7, 61.6	60.8, 193.9, 75.7	60.0, 69.8, 77.5	138.4, 138.4, 138.4
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 91.9, 90	90, 90, 90	90, 90, 90
Wilson B-factor (Å^2)	35.7	75.7	35.9	19.8	44.6	60.8	12.6	39.9
Resolution range (Å)	67.16-2.20 (2.32-2.20)	74.5-3.1 (3.31-3.10)	47.12-2.30 (2.38-2.30)	47.17-1.70 (1.73-1.70)	74.12-2.10 (2.16-2.10)	58.01-2.80 (2.92-2.80)	77.54-1.40 (1.42-1.40)	69.21-2.10 (2.16-2.10)
Unique reflections	40 214 (5 790)	15 562 (2 733)	36 714 (3 537)	89 279 (4 537)	46 049 (3 928)	38 300 (4 496)	63 440 (3 194)	27 077 (2 165)
Average multiplicity	4.0 (4.1)	12.8 (12.7)	7.0 (7.2)	3.7 (3.8)	7.2 (7.2)	2.1 (2.0)	3.5 (3.6)	12.8 (13.4)
Completeness (%)	99.7 (100.0)	99.8 (99.2)	100 (100)	99.2 (99.9)	99.2 (99.3)	89.1 (85.6)	98.0 (98.6)	100 (100)
$\langle I/\sigma I \rangle$	10.1 (1.9)	13.9 (2.6)	10.2 (2.2)	12.9 (1.8)	16.7 (1.5)	6.6 (1.3)	14.1 (1.8)	18.3 (2.1)
R_{merge} (all) (%)	8.0 (60.8)	15.1 (112.8)	13.6 (86.4)	6.0 (76.9)	7.1 (155.6)	15.9 (117.7)	5.2 (67.6)	10.9 (139.1)
R_{pim} (all) (%)	5.0 (38.9)	4.4 (32.6)	5.6 (34.8)	3.6 (45.6)	2.9 (61.6)	12.9 (96.8)	3.2 (41.0)	3.2 (39.2)
Refinement								
Num. of monomers	2	2	2	2	2	4	1	1
R_{work} / R_{free} (%)	21.5 / 26.4	-	20.8 / 24.4	18.2 / 22.3	21.4 / 25.6	24.9 / 29.3	14.4 / 18.1	18.8 / 23.7
No. of non H atoms								
Protein/N-glycan	5605		5569	5642	5549	10990	2905	2816
Water	59		62	493	45	0	287	84
Ligands	9		13	35	32	1	20	1
Average B-factor (Å^2)								
Protein	44		47	28	64	48	18	44
Water	33		33	36	46	-	30	43
Ligands	82		73	43	59	72	30	35
r.m.s.d. from ideality								
bond lengths (Å)	0.007		0.010	0.010	0.011	0.008	0.013	0.013
bond angles (°)	1.172		1.329	1.382	1.364	1.301	1.591	1.504

Table S1 - continued

PDB code	**4uz6**	**4uz7**	**4uz9**	**4uza**	**4uzq**	**4uzj**	**4uzk**
Species	human	human	human	human	human	Drosophila	Drosophila
Crystal form	V	VI	VII	VIII	IX	I	II
identifier	SOS	SOS (unmodelled)	SOS	phosphate	Ser-palmitoleate	apo	apo
Data collection							
X-ray source	i04	i03	i03	i04	i04/i02	i02	i03
wavelength (Å)	0.98	0.97	0.97	0.98	1.00 / 1.07	0.98	0.97
space group	$P2_12_12_1$	I2	$P4_12_12$	F432	I2	$P2_1$	$P2_12_12_1$
unit cell dimensions							
a, b, c (Å)	66.5, 70.0, 195.8	62.9, 79.9, 160.0	69.6, 69.9, 195.1	232.3, 232.3, 232.3	50.5, 69.4, 121.0	59.4, 81.4, 86.9	59.3, 142.1, 31.7
α, β, γ (°)	90, 90, 90	90, 94.0, 90	90, 90, 90	90, 90, 90	90, 92.3, 90	90, 105.5, 90	90, 90, 90
Wilson B-factor (Å^2)	27.6	21.9	26.2	56.1	11.1	53.0	31.7
Resolution range (Å)	34.43-1.90 (1.94-1.90)	79.80-2.20 (2.27-2.20)	97.57-2.20 (2.27-2.20)	70.06-2.40 (2.50-2.40)	60.15-1.50 (1.52-1.50)	58.38-2.40 (2.51-2.40)	142.12-1.90 (1.94-1.90)
Unique reflections	72 075 (4 340)	40 126 (3 475)	25 372 (2 273)	21 598 (2 412)	64 352 (2 805)	30 929 (3 742)	62 554 (4 126)
Average multiplicity	3.9 (3.1)	3.7 (3.7)	7.1 (7.3)	21.2 (22.2)	3.4 (2.2)	3.0 (3.1)	6.5 (6.6)
Completeness (%)	99.2 (98.0)	99.7 (99.9)	100 (100)	100 (100)	96.4 (85.4)	98.9 (99.3)	99.7 (99.5)
$\langle I/\sigma I \rangle$	7.7 (1.9)	8.5 (2.2)	13.4 (2.4)	25.0 (1.7)	11.6 (2.2)	13.8 (1.9)	13.9 (2.3)
R_{merge} (all) (%)	8.6 (75.9)	11.9 (68.8)	15.8 (116.8)	11.6 (223.2)	10.1 (92.0)	5.0 (67.6)	8.1 (95.6)
R_{pim} (all) (%)	4.5 (49.0)	7.2 (41.3)	6.4 (46.3)	2.6 (48.2)	5.7 (67.4)	3.5 (45.3)	3.4 (39.9)
Refinement							
No. of monomers	2	2	1	1	1	2	2
R_{work} / R_{free} (%)	19.5 / 23.5	18.3 / 22.8	19.0 / 23.7	20.7 / 27.1	12.8 / 17.7	19.4 / 24.5	17.6 / 20.4
No. of non H atoms							
Protein/N-glycan	5725	5712	2889	2859	2938	5483	5668
Water	206	208	115	23	308	14	219
Ligands	112	1	56	5	28	0	0
Average B-factor (Å^2)							
Protein	38	34	39	64	17	68	36
Water	37	32	38	55	26	59	39
Ligands	57	60	82	74	40	-	-
r.m.s.d. from ideality							
bond lengths (Å)	0.015	0.011	0.014	0.011	0.018	0.013	0.015
bond angles (°)	1.744	1.455	1.674	1.537	1.877	1.601	1.594