Electronic Supplementary Material (ESI) for Chemical Communications

Electronic Supplementary Information

Clicked Europium Dipicolinates Complexes for Proteins X–Ray Structural Determination.

Romain Talon,[a,b,c] Lionel Nauton,[d] Jean–Louis Canet,[d] Richard Kahn,[a,b,c] Eric Girard*[a,b,c] and Arnaud Gautier,*[d]

[a] *CEA, DSV, Institut de Biologie Structurale (IBS), 41 rue Jules Horowitz, Grenoble F–38027, France.*
[b] *CNRS, UMR 5075, Grenoble, France.*
[c] *Université Joseph Fourier, Grenoble 1, F–38000, France.*
[d] *Institut de chimie de Clermont–Ferrand, CNRS UMR 6296, Clermont–Université, 24 Ave des Landais, F–63177 Aubière, France.*

SI 1- Materials and chemicals.

SI 2- Experimental procedures.

SI 1- Materials and chemicals.

NMR spectra were recorded in Fourier Transform mode with a Bruker AVANCE 400 (^{1}H at 400 MHz, ^{13}C at 100 MHz), at 298K. Data are reported as chemical shifts (δ) in ppm. Residual solvent signals were used as internal references (^{1}H, ^{13}C).

Elemental analyses were performed at the Service de Microanalyse, Université Henry Poincaré,Vandoeuvre–les–Nancy, France.

Hen egg white lysozyme (HEWL) was purchased from Roche® (Ref. 10 837 059 001). *Thaumatococcus danielli* thaumatine (TdTHAU) was purchased from Sigma-Aldrich® (Ref. T7638). Both proteins were directly used for crystallization without any further purification

SI 2- Experimental procedures

General procedure for the synthesis of the ligands:

I– CuAAC reaction: to a suspension of dimethyl 4–azidopyridine–2,6–dicarboxylate (1.0 mmol) in MeOH (10 mL) was added 1.5 equivalents of alkyne under stirring. Then, 2 mol–% of [(SIMes)(4.7dichloro 1,10–phenanthroline)CuCl][1] was added and the reaction was stirred for 3 hours during which the precipitation of the product occurred. A filtration afforded the products in 85% to 95% yield.

II– Saponification: to a suspension of the diester (1.0 mmol) in water (5 mL) was added aqueous NaOH (1M, 6 mL). The reaction mixture was heated at 80°C until complete dissolution and still continued until 2 hours before being cooled to room temperature. A solution of HCl (3.5M) was then added until pH = 2.0 that lead to a precipitation of the diacid. A filtration affords the expected compounds in quantitative yields.

Ligand 4, HCl, 1/2H_2O: synthesized according to literature.[2]

Ligand 5, HCl, 1/2H_2O. ^{1}H RMN (400 MHz, D_2O+NaOD) δ 9.18 (s, 1H), 8.75 (s, 2H), 3.59 (t, J=7Hz, 2H), 2.73 (t, J=7Hz, 2H). ^{13}C RMN (100 MHz, D_2O+NaOD) δ 164.8, 150.6, 146.6, 144.8, 121.3, 116.4, 59.8, 29.1. E. A. Calcd for $C_{11}H_{10}N_4O_5$, HCl, ½ H_2O: C 40.82, H: 3.74, N: 17.31; found C: 40.60, H: 3.52, N: 17.27.

Ligand 6, ½ HCl. ^{1}H RMN (400 MHz, D_2O) δ 9.12 (s, 1H), 8.55 (s, 2H), 4.90 (s, 2H), 4.18 (brs, 2H), 3.61 (brs, 2H), 3.25 (s, 6H). ^{13}C RMN (100 MHz, DMSO–d_6) δ. 164.7, 150.6, 144.6, 137.0, 127.6, 117.1, 64.5, 58.2, 54.9, 50.4. E. A. Calcd for $C_{14}H_{18}ClN_5O_5$, ½ HCl: C 43.11, H: 4.78, N: 17.96; found C: 43.37, H: 3.90, N: 17.85.

General procedure for the preparation of the solution of the lanthanide complexes.

To a suspension of the diacid (0.2 mmol, 3.0 eq.) in 10.0 mL of water was added 0.27 mmol (4.5 equivalents) of Na_2CO_3. Then, 1.0 equivalent of $LnCl_3$, $6H_2O$ was added and the solution was stirred at least one hour to afford a 10^{-2} M stock solution.

Alternatively, the corresponding complexes can be obtain by the concentration of the reaction solution to its 1/3 under reduced pressure (temperature below 40°C) followed by the addition of 15–20 mL of ethanol to precipitates the expected complexes (yield 75–85%).

Heavy atom derivative crystals preparation:

HEWL:

Derivative crystals of HEWL were grown by vapour diffusion at 293 K in a crystallization solution of 0.9 to 1.5 M NaCl with 100 mM sodium acetate buffer pH 4.6 and by mixing 1.5 µL of protein solution at 20 mg.mL^{-1} (1.4 mM) or 30 mg.mL^{-1} (2.1 mM), 1.5 µL of crystallization solution and 1.5 µL of lanthanide complex solution at 75 mM. Useful crystals appeared in a few days (5 days for 1.5 mM NaCl and 15 days for 0.9 M NaCl).

Prior to data collection, derivative crystals were cryocooled using 25 % glycerol and parafine oil as cryoprotectants.

Thaumatine de *Thaumatococcus danielli* :

Derivative crystals of TdTHAU were grown by vapour diffusion at 293 K in a crystallization solution of 0.3 to 0.9 M sodium potassium tartrate with 100 mM bis-tris propane buffer pH 6.5 and by mixing 1.5 µL of protein solution at 40 mg.mL^{-1} (1.8 mM), 1.5 µL of

crystallization solution and 1.5 µL of complex solution at 30 mM. Useful crystals appeared in a few days (5 days for 0.9 mM Na/K tartrate and 10 days for 0.3 M Na/K tartrate). Derivative crystal were cryocooled using 25 % ethylene glycol or paratone as cryoprotectants.

Data collection and data statistics:
Diffraction data were collected on beamlines FIP–BM30A and PROXIMA1, at ESRF (Grenoble) and SOLEIL (Saint-Aubin) synchrotrons respectively and were integrated with XDS[3]. All HEWL crystals belong to the $P4_32_12$ spacegroup with one molecule per asymmetric unit leading to a solvent content of 27 %. TdTHAU crystals belong to the $P4_12_12$ spacegroup with one molecule per asymmetric unit leading to a solvent content of 47 %. Integrated intensities were scaled and merged using SCALA and TRUNCATE from the CCP4 programs suite.[4] Data collection conditions as well as data statistics are summarized in Table 1.

Table 1. Data collection and processing statistics

Complex number	**7**		**8**		**9**
Protein name	**HEWL**	**TdTHAU**	**HEWL**	**TdTHAU**	**HEWL**
Synchrotron Source	SOLEIL	ESRF			
Beamline	PROXIMA 1	BM30A			
[Complex] in drop (mM)	2.7	10	1.4	1.4	12
[Protein] in drop (mM)	0.25	0.6	0.46	0.6	0.67
[Complex]/[Protein] ratio	10.8	16.67	3.0	2.3	17.91
λ (Å)	0.886	0.979	0.979	0.979	0.979
Space group	$P4_32_12$	$P4_12_12$	$P4_32_12$	$P4_12_12$	$P4_32_12$
Cell parameter (Å)	a= 77.40, c= 38.47	a= 57.99, c=150.18	a= 77.76, c= 37.90	a= 57.85, c= 150.21	a= 77.24, c= 38.36
Resolution (Å) **Highest resolution shell**	19.24 - 1.35 (1.42 - 1.35)	45.90 - 1.30 (1.37 - 1.30)	38.88 - 1.51 (1.59 - 1.51)	45.83 - 1.20 (1.26 - 1.20)	38.75 - 1.21 (1.27 - 1.21)
Unique reflexions	26192 (3738)	64239 (9169)	18899 (2644)	78373 (11008)	36469 (5190)
R_{merge} (%) [(a)]	3.6 (19.3)	4.9 (45.4)	5.0 (32.3)	5.1 (45.7)	5.1 (59.8)
R_{pim} (%) [(b)]	2.1 (11.9)	3.1 (29.2)	2.0 (13.3)	1.9 (15.7)	6.5 (34.4)
R_{ano} (%) [(c)]	2.7 (9.4)	2.6 (21.2)	2.0 (11.0)	2.5 (15.6)	3.5 (27.3)
I/σ(I) [(d)]	12.8 (3.8)	12.5 (1.7)	10.3 (2.4)	10.4 (1.7)	10.8 (1.3)
Completness (%)	99.7 (99.2)	99.9 (99.4)	99.7 (98.0)	98.0 (95.8)	99.7 (98.8)
Multiplicity	7.5 (6.8)	6.6 (6.5)	13.6 (12.7)	9.9 (10.0)	7.6 (7.2)

[a] $R_{merge} = \sum_h \sum_i \left| \bar{I}(h) - I_i(h) \right| / \sum_h \sum_i \left| I_i(h) \right|$ where $I_i(h)$ is the i[th] measurement of reflection *h* and $\bar{I}(h)$ is the mean measurement of reflection *h*.

[b] $R_{p.i.m.} = \sum_h \left(\frac{1}{(N-1)} \right)^{1/2} \sum_i \left| I_i(h) - \bar{I}(h) \right| / \sum_h \sum_i I_i(h)$. This indicator, which describes the precision of the averaged measurement, is most relevant.[5]

[c] $R_{ano} = \sum_h \left| \bar{I}^+(h) - \bar{I}^-(h)) \right| / \sum_h \left| \bar{I}^+(h) + \bar{I}^-(h)) \right|$ where $\bar{I}^+(h)$ and $\bar{I}^-(h)$ are the mean intensities of a Friedel mate.

[d] I/σ(I) is the signal-to-noise ratio for merged intensities.

Structure refinement:
Both HEWL and TdTHAU model were manually improved in COOT[6] prior to refinement with Phenix.refine.[7] Models were then optimized through iterative rounds of refinement and model building. At the end stages of the refinement, TLS was used with TLS-groups determined with the TLSMD server.[8]
As shown in Table ESI 2, the verification (Phenix.validate) of the 1.35 Å resolution complete HEWL-complex7 final model showed no residues in disallowed regions of the Ramachandran

plot (98.6 % in preferred regions, 1.4 % in allowed regions). Similar refinement statistics are observed for the 1.51 Å resolution complete HEWL-complex8 final model (no residues in disallowed regions of the Ramachandran plot, 98.5 % in preferred regions, 1.5 % in allowed regions) as well as for the 1.21 Å resolution complete HEWL-complex9 final model (no residues in disallowed regions of the Ramachandran plot, 98.6 % in preferred regions, 1.4 % in allowed regions)
The analysis of the complete 1.30 Å resolution TdTHAU-complex7 final model showed no residues in disallowed regions of the Ramachandran plot (99.0 % in preferred regions, 1.0 % in allowed regions). Similar refinement statistics are observed for the 1.20 Å resolution TdTHAU final model with complex 8: no residues were in disallowed regions of the Ramachandran plot (99.0 % in preferred regions, 1.0 % in allowed regions).

Table 2: Refinement statistics.

Complex number	**7**		**8**		**9**
Protein name	**HEWL**	**TdTHAU**	**HEWL**	**TdTHAU**	**HEWL**
PDB Code					
Resolution (Å)	19.74 - 1.35	28.99 - 1.30	34.78 - 1.51	31.50 - 1.30	38.75 - 1.21
R_{work} (%) [(a)]	12.34	13.56	15.86	13.70	14.85
R_{free} (%) [(a)]	15.11	15.78	18.19	15.55	17.45
Number of reflexion used	48869	64138	18847	78332	68533
Atomic composition					
Protein	1091	1592	1048	1592	1100
Water	159	376	148	411	146
Ions	11	1	12	1	11
Ligands	80	57	62	60	4
Global standard deviation					
Bond length (Å)	0.017	0.009	0.011	0.013	0.018
Bond angle (°)	1.741	1.269	1.347	1.613	1.687
Bfactor values					
Mean protein Bfactor (Å^2)	13.46	11.90	16.49	11.77	12.86
Min protein Bfactor (Å^2)	6.08	5.27	6.43	5.42	7.02
Max protein Bfactor (Å^2)	37.62	38.27	51.47	32.02	34.40
Mean water Bfactor (Å^2)	27.70	28.40	29.80	28.29	27.20
Mean Ln ions Bfactor (Å^2)	18.70	15.77	11.22	17.77	11.12
Mean ligand Bfactor (Å^2)	19.70	21.17	14.55	20.51	15.56

[a] $R = \sum_h |F_o - F_c| / \sum_h |F_o|$ where F_o and F_c are the observed and calculated structure factor amplitudes of reflection *h* respectively. R_{free} [9] is the *R* for the test reflection data set for cross validation (5% of excluded reflections). R_{work} is the *R* for the working reflection data set.

Experimental phasing:
HEWL-Complex7 structure was determined with single-wavelength diffraction (SAD) phasing method. SAD data were collected at high energy in order to get high resolution. Lanthanide atoms positions within the asymmetric unit were determined using the program SHELXD.[10] Heavy atom refinement, phasing and density modification calculations were performed using SHARP[11] program. Regarding HEWL-complex 7 structure, this process yielded to a 1.35 Å resolution interpretable electron density map (Figure of merit of 0.334 and 0.770 after SHARP and density modification respectively). Automatic model building, using Arp/warp[12], led to a model consisting in 109 (over 129) residues (Table ESI 3).
The same procedure was applied for all complexes and proteins described herein and led to similar phasing statistics (Talon et al. manuscript in preparation).

Table 3: Phasing statistics for HEWL-complex7 structure determination.

Complexe number	**7**
Protein name	**HEWL**
Anomalous phasing power	1.41
Arp/wARP (built/total)	109/129

References:

[1] (a) M.–L. Teyssot, L. Nauton, J.–L. Canet, F. Cisnetti, A. Chevry, A. Gautier *Eur. JOC*, 2010, 3507–3515.(b) M.–L. Teyssot, A. Chevry, M. Traïkia, M. El–Ghozzi, D. Avignant, A. Gautier *Chemistry, European J*, 2009, *15*, 6322–6326.
[2] Z. E. Chamas, X. Guo, J.–L. Canet, A. Gautier, D. Boyer, R. Mahiou *Dalton Trans.*, 2010, **39**, 7091–7097.
[3] W. Kabsch. XDS. Acta Cryst. (2010). D66, 125-132.
[4] M. D. Winn et al.. *Acta. Cryst.* 2011, **D67** , 235-242.
[5] M. S. Weiss.. *J. Appl. Cryst*. 2001. **34**, 130-135.
[6] P. Emsley, B. Lohkamp, W. G. Scott and K. Cowtan. *Acta Cryst.* 2010, **D66**, 486-501.
[7] Afonine PV, Grosse-Kunstleve RW, Echols N, Headd JJ, Moriarty NW, Mustyakimov M, Terwilliger TC, Urzhumtsev A, Zwart PH, Adams PD. *Acta Crystallogr D Biol Crystallogr*. 2012, **68**:352-367.
[8] (a) J Painter & E A Merritt Acta Cryst.2006, **D62**, 439-450. (b) J Painter & E A Merritt J. Appl. Cryst. 2006, **39**, 109-111.
[9] Brünger AT. *Nature*. 1992, 355(6359), 472-5.
[10] Sheldrick, G.M. *Acta Cryst*.2010, **D66**, 479-485.
[11] Bricogne, G., Vonrhein, C., Flensburg, C., Schiltz, M. & Paciorek, W. *Acta Cryst*. 2003, **D59**, 2023-2030.
[12] Langer, G., Cohen, S.X., Lamzin, V.S. & Perrakis, A. *Nature Protocols.* 2008, **3**, 1171-1179