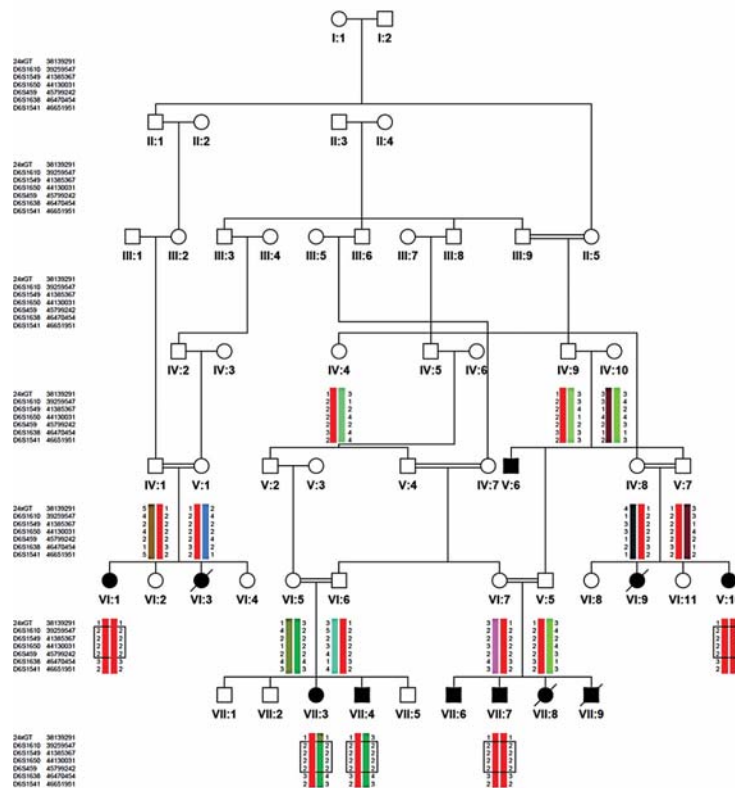


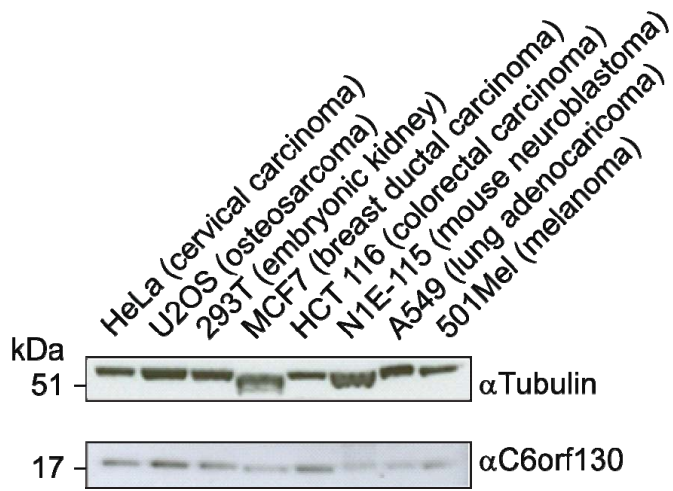
Deficiency of Terminal ADP-Ribose protein Glycohydrolase TARG1/C6orf130 in neurodegenerative disease

Reza Sharifi, Rosa Morra, C. Denise Appel, Michael Tallis, Barry Chioza, Gytis Jankevicius, Michael A. Simpson, Ivan Matic, Ege Ozkan, Barbara Golia, Matthew J. Schellenberg, Ria Weston, Jason G. Williams, Marianna N. Rossi, Hamid Galehdari, Juno Krahn, Alexander Wan, Richard C. Trembath, Andrew H. Crosby, Dragana Ahel, Ron Hay, Andreas G. Ladurner, Gyula Timinszky, R. Scott Williams, and Ivan Ahel

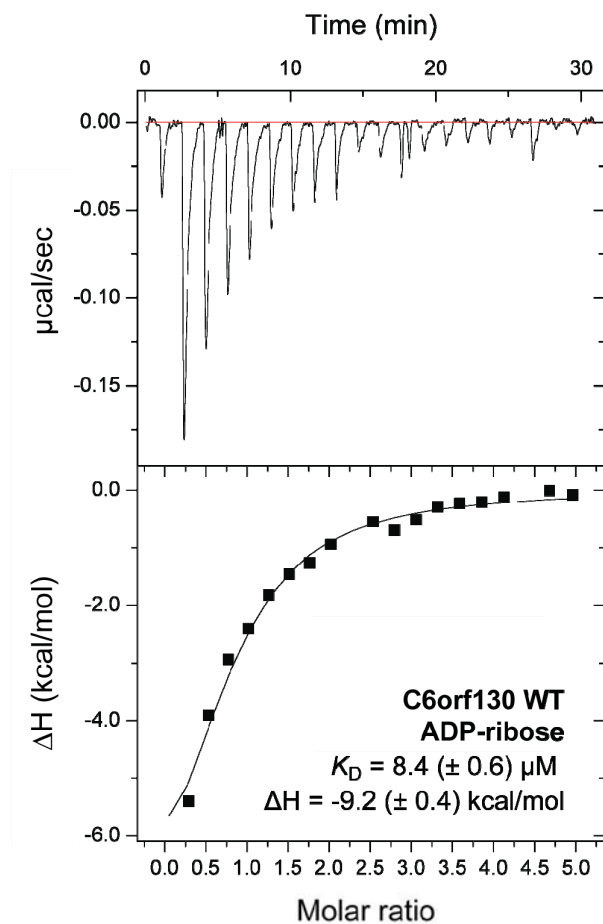
SUPPLEMENTARY FIGURES AND TABLES



Supplementary Figure 1. Refinement mapping of the region on chromosome 6 using microsatellite markers.

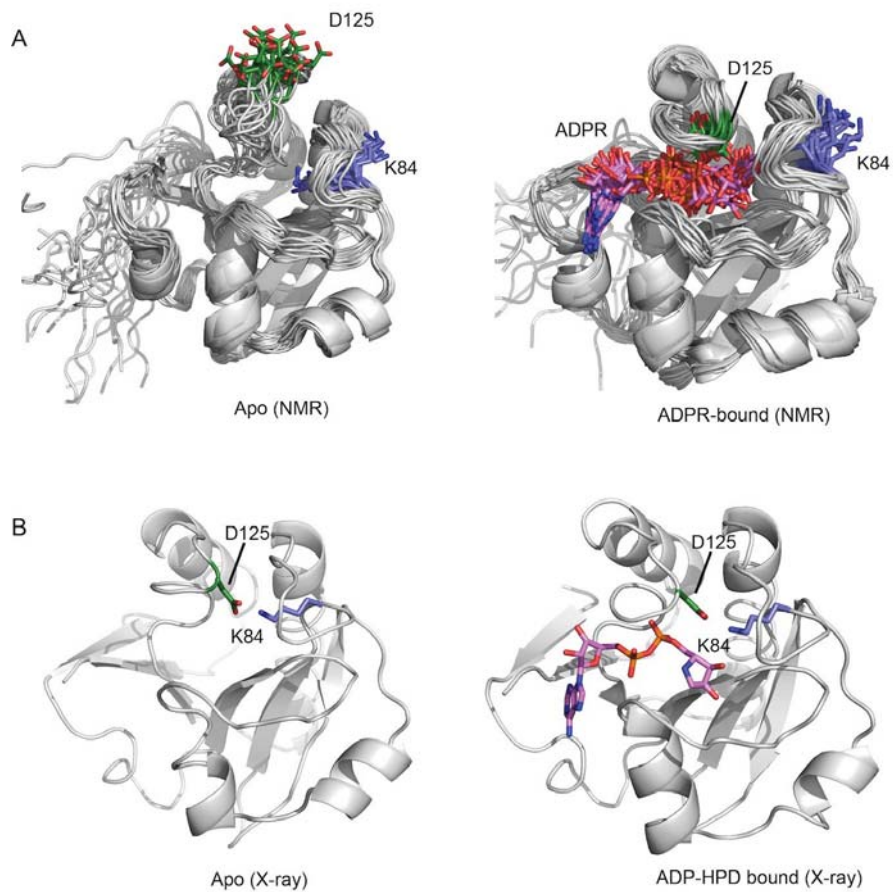


Supplementary Figure 2. Expression profile showing ubiquitous expression of C6orf130 protein across a range of cell lines derived from different human and mouse tissues, as tested by Western blot.

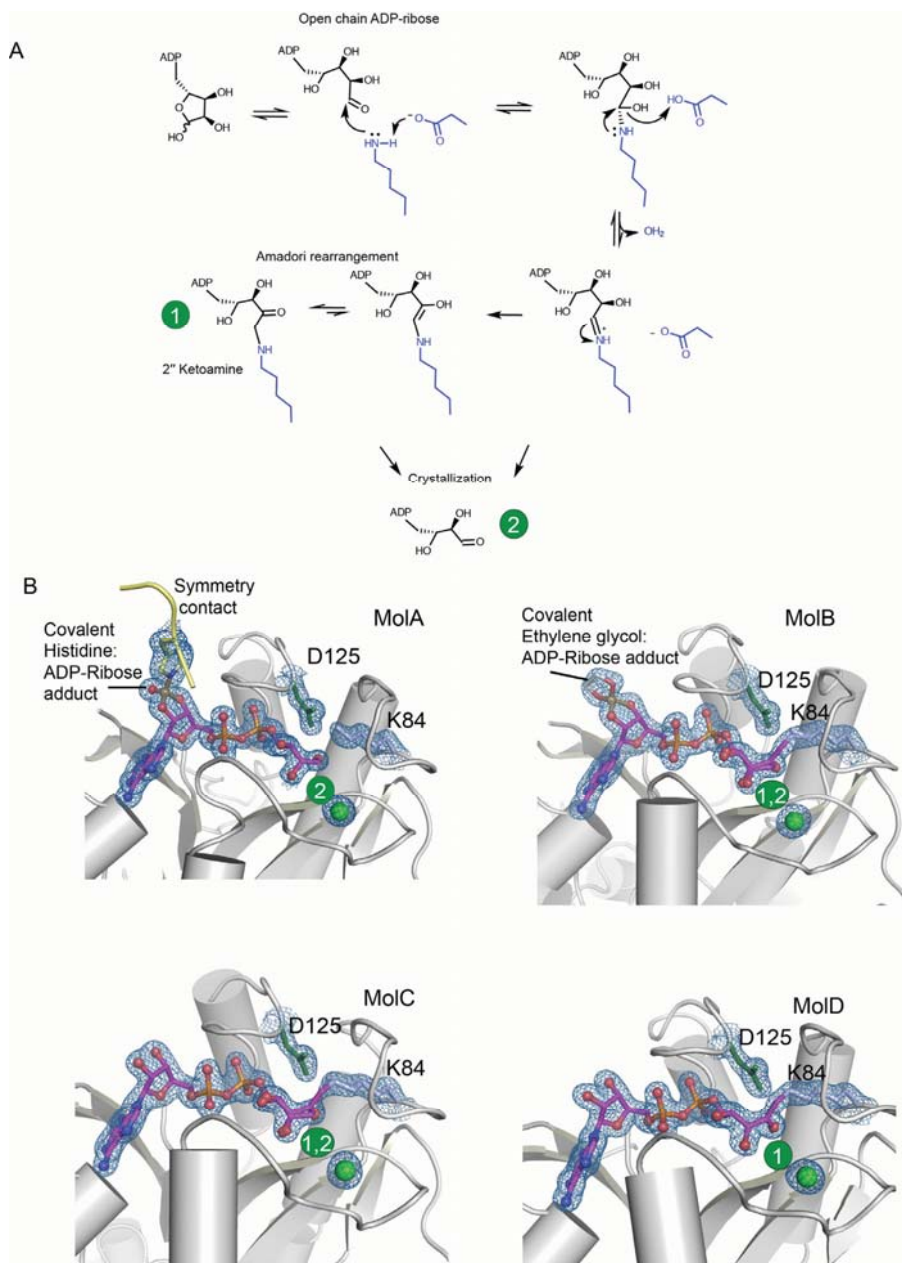


| C6orf130 | K_d (μM) | N |
|----------|-------------------------|------|
| WT | 8.4 | 0.76 |
| K84A | no binding | |
| G123E | no binding | |
| D125A | 2.6 | 0.77 |

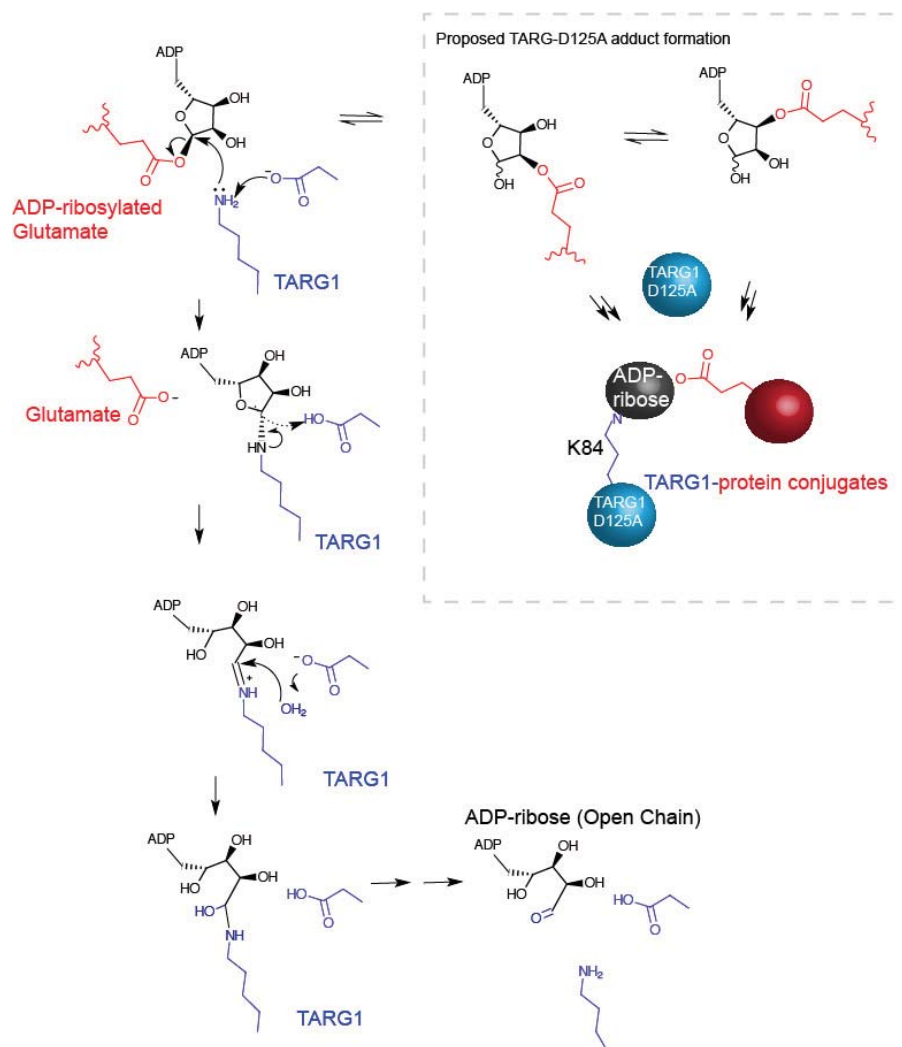
Supplementary Figure 3. Summary of ADP-ribose dissociation constants for C6orf130 protein and its mutant forms as determined by Isothermal Titration Calorimetry. Fitted dissociation constant and number of binding sites are indicated in the table.



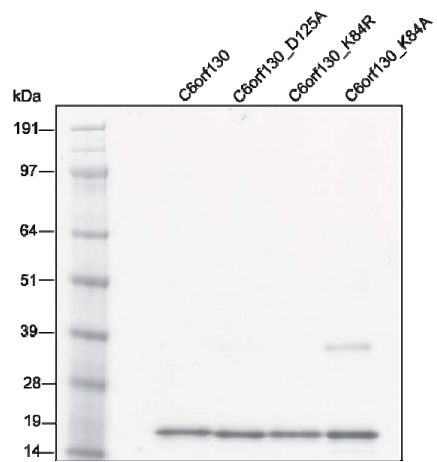
Supplementary Figure 4. Structural comparisons of C6orf130 NMR (Petersen *et al*, 2011) and X-ray structures (This study). **(A)** Apo (RCSB:2LGR) C6orf130 and ADP-ribose bound (RCSB 2L8R) NMR ensembles are displayed, showing an active site excluded conformation of K84. **(B)** X-ray structures of C6orf130/TARG1 show K84 in close proximity to D125 and bound ligand.



Supplementary Figure 5. TARG1/C6orf130 reaction with ADP-ribose. **(A)** Lys 84 reacts with the C1 position of ADP-ribose. Following dehydration, an imine intermediate converts to form a 2''- ketoamine (species "1", green circle) via an Amadori rearrangement (Cervantes-Laurean et al, 1993), and similar to that observed in the X-ray structure of DraG-ADP-ribose complex (Berthold et al, 2009). **(B)** Sigma-A weighted 2Fo-Fc electron density maps (contoured at 1.0 σ for the four C6orf130-ADP-ribose complexes in the crystallographic asymmetric unit. For Molecule D, geometry of the K84 covalent adduct is most consistent for the 2''-ketoamine (species "1"). A second co-crystallized ADP-ribose derived species arising during the three-month crystal growth (MW=528.12 Da, species "2") is apparent for MolA. Molecules B and C, have been modeled with mixed occupancy of species 1 and 2. Molecules A and B additionally are modified by covalent adducts to symmetry related histidine or ethylene glycol on the distal ribose sugars. The nature of these later linkages is uncertain, but the electron density is consistent with a boron bridge, suggesting they arose from contaminating borate in ADP-ribose, crystallization reagents or protein samples.



Supplementary Figure 6. Proposed reaction mechanism for the hydrolysis of a terminal ADP-ribose from PARP. A single ADP-ribose is attached to a glutamate side chain of PARP (red). When the PARP side chain is linked to the 1''-hydroxyl, the Asp-Lys side chains of C6orf130 displace the PARP glutamate, forming a Schiff-base intermediate, which is resolved by water-mediated hydrolysis. Acyl migration (Kasamatsu et al, 2011, Sauve et al, 2001) of the PARP side chain may facilitate 2'' or 3''-ADP-ribose linkages. In this case, C6orf130/TARG1 D125A forms adduct without releasing the PARP side chain.



Supplementary Figure 7. Purification of human C6orf130 proteins. Wild-type and mutant proteins were purified as His⁶-tagged proteins by affinity chromatography.

| Clinical presentation: | |
|---------------------------------------|---|
| Onset and course of the disease | 4-6 months after the birth, progressive, expiring by age 10-11 years mainly because of aspiration pneumonia |
| Neurological findings | |
| Neurodevelopment | Severe delay |
| Seizure | Generalised tonic-clonic (grand mal) |
| Motor neuron | Quadriplegia, absence of tendon reflexes |
| Visual, Hearing | Not assessable |
| Other neurological findings | Aphasia, partial absence of swallowing reflex |
| Other systems | No evidence for cardiovascular, renal, gastrointestinal or endocrine abnormalities |
| Laboratory tests | |
| Blood, urine, and CSF tests | Normal results for: blood chemistry, blood sugar, complete blood count (CBC), kidney function tests, liver function tests, cerebro-spinal fluid, cellular and biochemistry tests, urine biochemistry, blood gas and amino acids profile |
| Skin fibroblast cultured analysis | Negative for ketolysis defect |
| Brain and spinal cord CT and MRI scan | Normal |
| EEG | Abnormal indicating general seizure without brain cortex atrophy |
| EMG & NCV | Abnormal indicating motor-sensorial neuropathy |

Supplementary Table 1. Clinical and paraclinical assesment of the 6 affected cases age from 2 to 9 years.

| | Heterozygous | Homozygous |
|---|---------------------|-------------------|
| Single nucleotide variants (SNVs): | | |
| Coding | 10869 | 7547 |
| Synonymous | 5586 | 3925 |
| Non-synonymous missense | 5058 | 3458 |
| Stopgain | 61 | 20 |
| Stoploss | 17 | 16 |
| Splice site (+/-10bp) | 1332 | 945 |
| Insertion and deletions (indels): | | |
| Coding | | |
| Frameshift deletion | 42 | 18 |
| Frameshift insertion | 20 | 38 |
| In-frame deletion | 44 | 30 |
| In-frame insertion | 28 | 35 |

Supplementary Table 2. The number of variants for SNVs and indels for the affected case number VII7 identified by new generation sequencing (NGS).

| Gene | Description | High expression |
|------------------------|---|------------------------|
| <i>DAAM2</i> | dishevelled associated activator of morphogenesis 2 | Whole Brain |
| <i>MOCS1</i> | molybdenum cofactor synthesis 1 | Hypothalamous |
| <i>UBR2</i> | ubiquitin protein ligase E3 component n-recognin 2 | Cortex |
| <i>TBCC</i> | tubulin folding cofactor C | Whole brain |
| <i>KIAA0240</i> | uncharacterized protein KIAA0240 | Cerebellum |
| <i>KLHDC3</i> | kelch domain containing 3 | Whole brain |
| <i>C6orf108</i> | chromosome 6 open reading frame 108 | Whole brain |
| <i>TTBK1</i> | tau tubulin kinase 1 | Cortex and cerebellum |
| <i>C6orf154</i> | chromosome 6 open reading frame 154 | Occipital lobe |
| <i>AARS2</i> | alanyl-tRNA synthetase 2, mitochondrial (putative) | Whole brain |

Supplementary Table 3. The list of genes located on the putative region (Chr6p21) which are highly expressed in the central nervous system and were screened for two affected cases by Sanger sequencing. They did not harbour any pathological variant that could cause the disease phenotype.

| | C6orf130 Apo | C6orf130 ADP-HPD | C6orf130 ADP-ribose |
|---|---------------------|---|---------------------|
| Data collection | | | |
| Space group | C2 | P2 ₁ 2 ₁ 2 ₁ | P2 ₁ |
| Cell dimensions | | | |
| <i>a</i> , <i>b</i> , <i>c</i> (Å) | 63.32, 75.37, 35.00 | 50.80, 72.00, 81.92 | 73.38, 56.22, 73.79 |
| α , β , γ (°) | 90, 100.63, 90 | 90, 90, 90 | 90, 94.47, 90 |
| Resolution (Å) | 50-1.35 (1.40-1.35) | 50-1.25 (1.27-1.25) | 50-1.55 (1.58-1.55) |
| <i>R</i> _{sym} (%) | 9.9 (44.5) | 7.6 (40.2) | 4.3 (53.4) |
| <i>I</i> / σ <i>I</i> | 11.7 (2.2) | 19.3 (2.6) | 24.3 (2.5) |
| Completeness (%) | 98.1 (96.2) | 94.3 (64.6) | 97.9 (96.2) |
| Redundancy | 3.5 (2.9) | 5.0 (3.8) | 3.8 (3.8) |
| Refinement | | | |
| Resolution (Å) | 50-1.35 | 50-1.25 | 50-1.55 |
| No. reflections | 32,791 | 75,062 | 85,113 |
| <i>R</i> _{work} / <i>R</i> _{free} | 12.6 / 16.5 | 12.0 / 15.4 | 13.9 / 17.8 |
| No. atoms | | | |
| Protein | 1236 | 2570 | 5824 |
| Ligand/ion | n/a | 70 | 272 |
| Water | 239 | 628 | 658 |
| <i>B</i> -factors | | | |
| Protein | 19.8 | 10.2 | 24.9 |
| Ligand/ion | n/a | 19.8 | 27.7 |
| Water | 34.7 | 26.1 | 35.2 |
| R.m.s. deviations | | | |
| Bond lengths (Å) | 0.015 | 0.015 | 0.008 |
| Bond angles (°) | 1.72 | 1.91 | 1.19 |

Each dataset was collected from a single crystal.

Values in parentheses are for highest-resolution shell.

Supplementary Table 4. Data collection and refinement statistics

Supplementary Table 5. Data collection and refinement statistics

| | TARG1/C6orf130 Apo | TARG1/C6orf130 ADP-HPD | TARG1/C6orf130 ADP-ribose |
|---|---------------------|---|---------------------------|
| Data collection | | | |
| Space group | C2 | P2 ₁ 2 ₁ 2 ₁ | P2 ₁ |
| Cell dimensions | | | |
| <i>a</i> , <i>b</i> , <i>c</i> (Å) | 63.32, 75.37, 35.00 | 50.80, 72.00, 81.92 | 73.38, 56.22, 73.79 |
| α , β , γ (°) | 90, 100.63, 90 | 90, 90, 90 | 90, 94.47, 90 |
| Resolution (Å) | 50-1.35 (1.40-1.35) | 50-1.25 (1.27-1.25) | 50-1.55 (1.58-1.55) |
| <i>R</i> _{sym} (%) | 9.9 (44.5) | 7.6 (40.2) | 4.3 (53.4) |
| <i>I</i> / σ <i>I</i> | 11.7 (2.2) | 19.3 (2.6) | 24.3 (2.5) |
| Completeness (%) | 98.1 (96.2) | 94.3 (64.6) | 97.9 (96.2) |
| Redundancy | 3.5 (2.9) | 5.0 (3.8) | 3.8 (3.8) |
| Refinement | | | |
| Resolution (Å) | 50-1.35 | 50-1.25 | 50-1.55 |
| No. reflections | 32,791 | 75,062 | 85,113 |
| <i>R</i> _{work} / <i>R</i> _{free} | 12.6 / 16.5 | 12.0 / 15.4 | 13.7 / 17.5 |
| No. atoms | | | |
| Protein | 1236 | 2527 | 4867 |
| Ligand/ion | n/a | 137 | 248 |
| Water | 239 | 628 | 668 |
| <i>B</i> -factors | | | |
| Protein | 19.1 | 9.8 | 27.3 |
| Ligand/ion | n/a | 10.7 | 27.7 |
| Water | 34.0 | 26.1 | 35.0 |
| R.m.s. deviations | | | |
| Bond lengths (Å) | 0.015 | 0.015 | 0.008 |
| Bond angles (°) | 1.72 | 1.91 | 1.19 |

Each dataset was collected from a single crystal.

Values in parentheses are for highest-resolution shell.