The cryo-EM structure of the neurofibromin dimer reveals the molecular basis for von Recklinghausen disease.

Christopher J. Lupton1,8, Charles Bayly-Jones1,2, Laura D’Andrea1,8, Cheng Huang3, Ralf B. Schittenhelm3, Hari Venugopal4, James C. Whisstock1,2,5,6, Michelle L. Halls7, Andrew M. Ellisdon1

1Biomedicine Discovery Institute, Monash University, Clayton 3800, Victoria, Australia
2ARC Centre of Excellence in Advanced Molecular Imaging, Monash University, Clayton 3800, Victoria, Australia
3Monash Proteomics & Metabolomics Facility, Monash University, Clayton 3800, Victoria, Australia
4Ramaciotti Centre for Cryo-Electron Microscopy, Monash University, Clayton 3800, Victoria, Australia
5EMBL Australia, Monash University, Melbourne, 3800, Victoria, Australia. 6ACRF Department of Cancer Biology and Therapeutics, John Curtin School of Medical Research, Australian National University, Canberra 2601, ACT, Australia. 7Drug Discovery Biology Theme, Monash Institute of Pharmaceutical Sciences, Monash University, Parkville 3052, Victoria, Australia
8These authors contributed equally.

Correspondence: andrew.ellisdon@monash.edu

Neurofibromin (NF1) is a tumour suppressor mutated in neurofibromatosis type 1 (von Recklinghausen disease), one of the most common human genetic diseases(1). NF1 regulates cellular growth through suppressing the Rat Sarcoma (RAS) pathway and, accordingly, mutations in this protein drive numerous cancers, including melanoma, ovarian, breast and brain cancer(2, 3). Currently, however, the molecular basis for NF1 function remains to be understood. Here we address this problem and use cryogenic Electron Microscopy (cryo-EM) to determine the structure of full-length NF1. The 640 kDa NF1 homodimer forms an extraordinary lemniscate (∞) shaped molecule that is ~30 nm in length and ~10 nm wide. Each NF1 monomer comprises an N-terminal HEAT-repeat domain (N-HEAT), a guanosine triphosphatase activating protein (GAP)-related domain (GRD), a Sec14 homologous and pleckstrin homologous module (SEC-PH), and a C-terminal HEAT domain (C-HEAT). The core NF1 scaffold is formed via a head-to-tail dimer of the N- and C-HEAT domains. This platform, which is responsible for interacting with more than 10 regulatory binding partners, comprises an extraordinary array of over 150 α-helices. Analysis of these EM data revealed that the GRD and SEC-PH domain are highly mobile with respect to the core scaffold and could not initially be accurately placed in electron density. Strikingly, however, using 3D variability analysis we were able to identify a significant subpopulation of NF1 particles and determine the complete NF1 structure to 5.6 Å resolution. These data revealed that the catalytic GRD and lipid binding SEC-PH domain are positioned against the core scaffold in a closed, autoinhibited conformation. We postulate that interaction with the plasma membrane may release the closed conformation in order to promote RAS inactivation. Our structural data further allow us to map the location of disease-associated NF1 variants and provide a long sought-after structural explanation for the extreme susceptibility of the molecule to loss-of-function mutations. Finally, it is suggested that approaches to combat NF1-linked diseases may include release of the autoinhibited state in order to improve NF1 catalytic efficiency.

Neurofibromin | von Recklinghausen disease | cancer | neurofibromatosis | cryoEM | autoinhibition | RAS | SPRED1 | tumor suppressor

The oncogenic RAS family of small GTP-binding proteins (Neuroblastoma RAS, Harvey RAS, and Kirsten RAS) have crucial roles in coordinating eukaryotic cell growth(4). Growth factor signalling promotes the cycling of RAS proteins from their inactive GDP-bound state to their active GTP-bound state. GTP-bound RAS proteins activate fundamental signalling pathways such as the RAF-MEK-ERK cascade and the PI3K-Akt-mTOR pathway(5). NF1 is a highly conserved RAS GAP that catalyses the hydrolysis of RAS-bound GTP(6). As such, NF1 acts as a potent off-switch for RAS-family signalling(2) (Extended Data Fig. 1).

Germline and sporadic mutations of NF1 cause the common genetic condition neurofibromatosis type 1 (von Recklinghausen disease) with variable clinical presentations, including benign cutaneous and plexiform neurofibromas1. NF1 is also a common driver gene in numerous cancers(7) with frequent mutations, deletions, or structural rearrangements in melanoma(8), breast cancer(9), lung cancer(10), and glioblastoma(11). Mutations that cause neurofibromatosis type 1 and cancer-associated somatic mutations are distributed throughout the NF1 gene(12, 13) without widespread clustering to the catalytic GAP domain, suggesting an acute susceptibility of the entire protein to dysregulation. NF1 inactivation increases GTP-bound RAS levels driving downstream activation of MAPK and PI3K pathways and aberrant cell growth(14–16). NF1 loss further dysregulates many additional processes that lead to cognitive defects(17), conferral of drug resistance in cancer therapy(18), and learning defects(19).

NF1 interacts with a very large number of cellular binding proteins and effectors (Extended Data Fig. 1). Notably, NF1 binds Sprouty Related EVH1 Domain Containing 1 (SPRED1), an event that enables its membrane recruitment(20). The SEC-PH module is also postulated to play an additional role in membrane recruitment through lipid and phosphoinositide (PI) binding(21). Once at the membrane the catalytic NF1 GRD stabilises GTP-bound...
RAS to promote GTP-hydrolysis(22, 23). SEC-PH binding to G-protein βγ can also inhibit NF1, modulating RAS activation and neural pathways central to opioid addiction(24).

Multiple further binding sites are distributed throughout NF1 including sites for tubulin binding in mitosis(25), 14-3-3 binding(26), as well as protein kinase A and C phospho-regulation(27).

To better understand NF1 function we expressed recombinant protein in insect cells and determined the cryo-EM structure of the full-length 640 kDa homodimer. The structure presented significant challenges with respect to particle heterogeneity, however, after symmetry expansion one wing of the NF1 homodimer could be reconstructed to a nominal resolution of 4.6 Å (Fig 1a, Extended Movie 1, Extended Data Fig. 2, 3, 4 and Extended Data Table 1). The maps were of sufficient quality to completely build and sequence de novo the N- and C-HEAT domains (Fig. 1b and Extended Data Fig. 5). In contrast, however, the central GRD and SEC-PH domains were poorly resolved and showed a high degree of mobility within the majority of particles. To address this, we used 3D variability analysis to identify a subpopulation of NF1 particles and determined the 5.6 Å structure of the complete NF1 molecule (Extended Data Fig. 4). In this latter reconstruction, clear secondary structure density was observed for the entire molecule (Fig. 1c-e) and crystal structures of both the GRD(20) and SEC-PH modules(21) could be fitted unequivocally into the density with each domain packed against the N- and C-HEAT domains. Extensive crosslinking mass spectrometry with BS^3 and BS^2G lysine crosslinkers confirm the placement of the central GRD and SEC-PH domains into the cryo-EM maps (Fig. 1f, g) as well as the sequence assignment of the dimeric N- and C-HEAT repeat core (Extended Data Fig. 5 and Extended Data Table 2). The GRD also crosslinks to known phospho-regulatory regions in NF1 (Extended Data Fig. 6).

Analysis of these structural data reveal that the NF1 scaffold comprises a large lemniscate-shaped assembly that forms as a consequence of head-to-tail dimerization of the N- and C-HEAT domains (Fig. 1a and Extended Movie 1). Each wing of the structure curves away from a central interface with an outer dimension spanning a remarkable ∼30 nm between each tip (Fig. 1b). Despite being well resolved in the cryo-EM maps, the central scaffold is remarkably flexible with a large degree of continuous conformational heterogeneity detected across the dataset (Extended Movie 2). Docking sites of more than 10 known NF1 binding partners map across the entire NF1 structure (Extended Data Fig. 7). These data suggest, that like the karyopherin HEAT-repeat family(28), the core scaffold acts as a flexible interaction platform coordinating the many functions of NF1 that span from cell growth(2) to cell division(25).
NF1 dimer assembly is mediated by five interfaces. The central interface formed by the symmetrical and anti-parallel alignment of the C-HEAT domains, two symmetrically equivalent N- and C-terminal interfaces that close each end of the lemniscate HEAT-repeat fold, and two entry and exit interfaces that enable the flexible GRD and SEC-PH domains to align on the edge of the core scaffold (Fig. 2a, b). The central interface is extensive (∼1670 Å² buried surface area) and formed by multiple C-HEAT domain helices as well as the highly conserved αC2-3 loop that is sandwiched at each end of the symmetrical interface (Fig. 2a, c). The central interface is particularly well conserved and likely the main energetic driver of NF1 dimer assembly (Extended Data Fig. 8 and Extended Data Fig. 9).

The two smaller (∼905 Å² buried surface area) equivalent N- and C-HEAT interfaces are formed by the packing of helices αN1 and αN2 of the N-HEAT domain with helices αC32 and αC33 of the C-HEAT domain to form a continuous helicoid topology (Fig. 2d). The centre of the interface is predominantly hydrophobic, with the conserved N-HEAT tryptophan (W9) deeply buried into a hydrophobic pocket formed by helices αC32 and αC33 (Fig. 2d). The two entry and exit interfaces are formed by the N-HEAT helical pair αN44 and αN45 that form a continuous HEAT-repeat like array with the C-HEAT helical pair αC1 and αC2. A total of ∼850 Å² is buried in each interface with predominantly hydrophobic packing between helices (Fig. 2e). This interface is central to supporting the integrity of the central scaffold as well as the flexible positioning of the GRD and SEC-PH domains on the scaffold edge.

In the 5.7 Å structure of the complete NF1 molecule, the GRD and SEC-PH modules pack tightly against the N- and C-HEAT domains. Whilst resolution limited detailed analysis, the GRD sits buried in the cleft formed by the N- and C-HEAT domains making contacts with both conserved C-HEAT domain α-helices as well as surface loops of the N-HEAT repeats (Fig. 3a-c). Analysis of the closed NF1 structure reveals that the RAS-binding interface is oriented towards the central scaffold in a conformation that is sterically incompatible with direct RAS binding (Fig. 3a, b). Similarly, the lipid binding sites of both the SEC and PH domains(21) are also occluded and would need to undertake significant rotation to enable membrane binding (Fig. 3d, e). In contrast, the modelled SPRED1 binds intimately to NF1 surface potentially interfacing with a previously unrecognized site in the C-HEAT domain (Fig. 3b, c) providing a clear route for SPRED1-driven membrane recruitment(29).

Collectively these data suggest that we have identified a closed conformation of NF1. Our data further suggest that significant rearrangement of both the GRD and SEC-PH domains must take place upon SPRED1 membrane recruitment in order to enable RAS turnover and lipid binding, respectively (Fig. 3e). Consistent with this idea, our structural data reveal that GRD and SEC-PH domains are highly mobile and can thus readily transition between the closed
Fig. 3. The GRD and SEC-PH domains are in a closed conformation. a. RAS binding to the GRD is sterically occluded by the N-HEAT domain. b. Boxed region from a rotated to highlight steric clashes of RAS with the N-HEAT domain and the SPRED1 binding site exposed on the GRD domain. c. NF1 oriented to highlight the close association of the modelled SPRED1 domain to the central scaffold. d. The NF1 structure oriented to highlight the horizontal alignment of the SEC-PH lipid binding pockets as well as the SPRED1 binding region on each GRD. The predicted plasma membrane position is indicated to highlight the symmetrical alignment of known NF1 membrane binding regions. e. Boxed region from d highlighting the buried PI binding site in the PH domain. f. Proposed regulatory model of NF1. A subpopulation of NF1 is in a stable closed conformation marked by the steric occlusion of the GRD interface and lipid-binding sites. The SPRED1 binding site is exposed on the surface of the closed GRD conformation and is known to recruit NF1 to the membrane.

Lastly, in order to gain understanding of the molecular basis of neurofibromatosis type 1 we analysed the frequency of disease-associated mutations across the NF1 gene (13, 30) (Fig. 4a.). Neurofibromatosis type 1 mutations are evenly spread throughout the gene with only sporadic clustering evident along its length. Analysis of the COSMIC (12) database also highlights that somatic cancer-associated mutations are evenly distributed throughout the NF1 gene (Fig. 4b). We mapped these disease-associated mutations onto the NF1 structure to gain greater insight into the mechanistic basis of loss-of-function of this key tumour suppressor protein (Fig. 4c). For example, the clinically significant neurofibromatosis type 1 mutants (ΔM991(31)) and (ΔL844-G848(32)) map to the N-HEAT domain and would likely interfere with helical packing directly beneath the SEC-PH interface (Fig. 4d). In addition, the common cancer-associated mutation R1849Q (33) packs intricately against the central interface (αC2-3) loop, a crucial region for NF1 dimer assembly (Fig. 4e). However, globally, NF1 mutations do not cluster to any one interface or region of the protein but are uniformly distributed throughout the structure (Fig. 4c). It is clear from the intricate NF1 fold that missense, deletion, or nonsense mutations are likely to have an acute effect on the stability of the entire NF1 protein fold. For example, any mutation that disrupts the intricate HEAT-repeat packing of the core scaffold may result in an inability of the lemniscate-fold to close and downstream loss of function.
This sensitivity of the NF1 fold to mutations also rationalises the lack of mutational hot spots within the tumour suppressor.

In conclusion, we have shown that NF1 assembles into a large homodimer characterised by a central lemniscate-shaped core formed by the assembly of the N- and C-HEAT domains. NF1 dimer assembly requires the coordinated packing of eight domains, which contain over 150 α-helices, to form the final topology. This highly mobile molecular scaffold provides an elongated binding platform for both the GRD and SEC-PH domains as well as the many NF1 cellular effector molecules that have been identified to date. We have further identified a sub-population of NF1 with the NF1 (isoform 1) was subcloned from R777-E139 Hs.NF1 (a gift from Dominic Esposito, Addgene plasmid #70423) as an N-terminal His₆-tagged protein in pFastBac1 (Invitrogen). NF1 was expressed in SF9 cells for 2 days following baculovirus infection (Bac-to-Bac, Invitrogen). Cells were harvested by centrifugation and lysed by sonication in 20 mM Tris-HCl pH 8.0, 500 mM NaCl, 10 % (v/v) glycerol, and 5 mM β-mercaptoethanol. Lysates were cleared by centrifugation, filtered at 0.45 µm, and incubated with Ni-NTA resin (Qiagen) and 20 mM imidazole at 4 °C with agitation. The resin was washed with lysis buffer, and NF1 eluted with 500 mM imidazole in the same buffer. NF1-containing fractions were further purified on a Superose 6 10/300 GL size-exclusion column (Cytiva) equilibrated in 20 mM HEPES pH 8.0, 150 mM NaCl, and 2 mM DTT.

Site-directed mutagenesis was performed on KRAS 1-169 (Q61H, isoform b, a gift from Cheryl Arrowsmith, Addgene plasmid #25153) to obtain the wild-type sequence. N-terminally His₆-tagged KRAS was expressed in BL21-CodonPlus(DE3)-RIL cells (Stratagene) by isopropyl-β-D-1-thiogalactopyranoside induction at 18 °C. Cells were lysed by sonication in 20 mM Tris-HCl pH 8.0, 500 mM NaCl, 10 % (v/v) glycerol, and 3 mM β-mercaptoethanol, with two complete EDTA-free protease inhibitor tablets (Roche). Cells were clarified by centrifugation, filtered through a 0.45 µm membrane, and incubated with Ni-NTA resin (Qiagen) and 20 mM imidazole at 4 °C with agitation. The resin was washed with lysis buffer and KRAS was eluted in lysis buffer containing 300 mM imidazole. KRAS-containing fractions were further purified on a HiLoad Superdex 75 16/60 (Cytiva) pre-equilibrated with 20 mM HEPES pH 8.0, 150 mM NaCl, 10 % (v/v) glycerol, and 2 mM DTT.

Cryo-EM sample preparation and data collection. Quantifoil R1.2/1.3 200 mesh copper grids were glow-discharged for 90 s at 15 mA using a Pelco easiGlow instrument. Freshly purified NF1 (3 µL at 0.5 mg ml⁻¹) was applied to the freshly discharged grid and vitrified in liquid ethane using a Vitrobot Mk IV (Thermo Fisher Scientific) with a blot of force of –2 for 2.5 s. Temperature and relative humidity were maintained at 4 °C and 100 %, respectively. Data were collected on a Talos Arctica (Thermo Scientific) operating at 200 kV with a 50 µm C2 aperture. Micrographs were acquired using a bottom mounted Falcon 3 direct electron detector in counting mode at a nominal magnification of 120,000×, corresponding to a calibrated physical pixel size of 1.194 Å. The electron dose rate was set to 0.8 electrons
pixel$^{-1}$ s$^{-1}$ with a total exposure time of 71.68 s, yielding a total dose of 40 electrons Å$^{-2}$. Automated collection was carried out using EPU with beam-shift to collect 9 images per stage movement. Defocus range was set between −0.5 to −2.2 μm.

Cryo-EM data processing.
All initial data monitoring and micrograph denoising were performed in Warp(34) (v1.0.9) to assess grid quality and select appropriate regions for data collection. An initial pass of the data was performed in cryoSPARC(35) (v2.14.2). Here, particle picking was performed by the elliptical blob method, these 23,480 particles were extracted in a box of 432 pixels, down sampled to 2.388 Å pixel$^{-1}$ and subjected to 2D classification. An ab initio model was generated from 1,818 particles without imposing symmetry. 

Upon completion of data collection, 2,675 dose-fractionated movies were corrected for beam induced motion and compensated for radiation damage within MotionCor2(36) (v1.1.0). All aligned movie frames were subsequently summed into dose-weighted averages for further processing. The contrast transfer function parameters were re-estimated with CTFFIND(37) (v4.1.8). Curated particles from the initial pass were used to train a cryOLO(38) (v1.6.1) model which was used to pick 192,225 particles from the full dataset. These were subjected to multiple rounds of 2D classification in both RELION(39) (v3.1) and cryoSPARC, yielding 95,564 particles of sufficient quality and homogeneity. A consensus reconstruction was generated by refining this subset in RELION with C2 symmetry yielding an 8.2 Å map after polishing and re-sampling to 1.194 Å pixel$^{-1}$. These particles were 3D classified in RELION without further alignments. A single subset was refined using RELION and SIDESPLITTER(40) yielding a map of 7.5 Å resolution. Further local refinement in cryoSPARC lead to a final dimer reconstruction at 6.2 Å resolution.

CryoSPARC 3D variational analysis and RELION multi-body analysis was performed on the consensus reconstruction to assess the degree of flexibility and guide signal subtraction. These analyses revealed significant movement of the two NF1 wings, as well as flexibility within individual lobes. Therefore, symmetry expansion of the consensus reconstruction was performed, and each wing was re-extracted after signal subtraction of the other wing. This gave rise to 191,128 sub-particles in a 256-pixel box. An additional round of 2D classification was performed to remove empty and junk particles. Refinement in RELION yielded a 6.4 Å map after symmetry expansion and signal subtraction, indicating flexibility of the NF1 homodimer had significantly limited the resolution of the consensus reconstruction. Further 3D classification of these particles yielded two homogenous subsets. Refinement of the CTF parameters was performed in cisTEM(41) with a high-resolution cut-off maintained above the fall-off of the FSC. Lastly, a final homogenous refinement was performed with RELION and SIDESPLITTER yielding a 4.56 Å resolution map. Conversions between software were performed with EMAN(42) (v2.2), with code written in-house or by pynm. The FSC was used to estimate resolution at the 0.143 threshold. Local resolution was estimated by the windowed blocres FSC method as implemented in cryoSPARC. Map sharpening was performed in RELION, an ad hoc B-factor of −80 Å$^{-2}$ was selected by inspection of the map.

A second round of 3D variation analysis was performed on the symmetry expanded particles, which after 3D principal component analysis, revealed a small discrete cluster of particles corresponding to an autoinhibited state of NF1. This subpopulation of 9,283 particles was re-extracted at the original pixel size and locally refined with gaussian priors to prevent diverging from previous alignment parameters. This analysis resulted in a 5.6 Å resolution map of the autoinhibited NF1 with GRD and SEC-PH domains with clearly resolved secondary structure.

The model of the N-HEAT and C-HEAT domains was built manually into the cryo-EM map using Coot(43), with iterative cycles of refinement carried out using PHENIX(44) and Isoloide(45). Secondary structure and adaptive network constraints were employed until the final round of refinement, where only secondary structure constraints were imposed. The structure had excellent final geometry (Extended Data Table 1). Structures of both the GRD(20) (PDB: 6V65) and SEC-PH(21) (PDB: 3PG7) domains were rigid body fit into the 5.7 Å cryo-EM reconstruction with excellent agreement between the EM and crystal structures to produce an autoinhibited NF1 model. A final, full NF1 dimer model was generated by rigid body fitting the autoinhibited NF1 model into the 6.2 Å dimer reconstruction.

CLMS. Crosslinking mass spectrometry was performed by adding either BS$^3$ (Thermo Fisher) or BS$^2$G (Thermo Fisher) crosslinkers at a 1:100 molar ratio to 4 μM NF1 in 20 mM HEPES pH 8.5, 150 mM NaCl, and 2 mM DTT. For BS$^3$, samples were incubated at room temperature for 20 min prior to the addition of 50 mM Tris-HCl pH 8.0 to quench the reaction. For BS$^2$G, samples were incubated at room temperature for 30 min prior to the addition of 20 mM NH$_4$HCO$_3$ to quench the reaction. For both crosslinkers, samples were snap-frozen in liquid nitrogen prior to further processing. Samples were subsequently denatured for 30 min at 65 °C in the presence of 10 mM DTT. Subsequently, 40 mM chloroacetamide was then added to the samples prior to incubation for 20 min at room temperature in the dark. A 1:100 (w/v) ratio of trypsin was added to the samples and further incubated at 37 °C overnight. Digestion was stopped with the addition of 1% (v/v) formic acid. Samples were subsequently cleaned using OMIX C18 pipette tips (Agilent Technologies) and stored in 0.1% (v/v) formic acid prior to mass spectrometry.

Samples were analysed by LC-MS/MS using a Dionex Ultimate 3000 RSLCnano system coupled onto an Orbi-
trap Fusion Tribrid instrument (Thermo Fisher). An Acclaim PepMap RSLC analytical column (75 μm × 50 cm, nanoViper, C18, 2 μm, 100 Å; Thermo Scientific) and an Acclaim PepMap 100 trap column (100 μm × 2 cm, nanoViper, C18, 5 μm, 100 Å; Thermo Scientific) were used to separate tryptic peptides by increasing concentrations of 80 % (v/v) acetonitrile (can) / 0.1 % (v/v) formic acid at a flow of 250 nl min⁻¹ for 90 min. The mass spectrometer was operated in data-dependent mode with the following parameters. The three-dimensional cryo-EM density maps are deposited into the Electron Microscopy, a Node of Microscopy Australia. We also acknowledge the office of the Vice-Provost for Research and Research Infrastructure (VPRIR) at Monash University and of Bioplatforms Australia (BPA) as part of the National Collaborative Research Infrastructure Strategy (NCRIS). This preprint was formatted in LATEX using an adaptation of Ricardo Henríquez’s template.

**AUTHOR CONTRIBUTIONS**
A.M.E. and M.L.H. conceived the study. C.J.L. and L.D. performed cloning, protein expression, and purification. C.J.L. prepared cryo-EM grids. C.J.L. and H.V. collected and processed cryo-EM data. C.B.L., C.J.L., and A.M.E. prepared cryo-EM maps. BRM and D.M. provided refined models. C.J.L. and L.D. performed GAP activity assays. L.D., C.H. and R.B.S. performed crossing experiments. All authors contributed to analyses of data and writing of the manuscript.

**Bibliography**

1. D. H. Gutmann, R. E. Ferber, R. H. Listernick, B. R. Korf, P. L. Wolters, and K. J. Johnston. Neurofibromatosis type 1. Nat Rev Dis Primers, 3(1):7004, 2017. ISSN 2056-676X (Electronic) 2056-6767 (Linking). doi:10.1038/nrdp.2017.4.
2. 7. Icgc Tcga Pan-Cancer Analysis of Whole Genomes Consortium. Pan-cancer analysis of whole genomes. Nature, 578(7793):82–93, 2020. ISSN 1476-4867 (Electronic) 0028-0863 (Linking). doi:10.1038/s41586-019-0991-8.
3. 8. K. Cichowski and T. Jacks. Nf1 tumor suppressor gene function: narrowing the gap. Cell, 104(4):593–604, 2001. ISSN 0092-8674 (Print) 0092-8674 (Linking). doi:10.1016/S0092-8674(01)00245-8.
4. 1. D. H. Gutmann, R. E. Ferner, R. H. Listernick, B. R. Korf, P. L. Wolters, and K. J. Johnston. Neurofibromatosis type 1. Nat Rev Dis Primers, 3(1):7004, 2017. ISSN 2056-676X (Electronic) 2056-6767 (Linking). doi:10.1038/nrdp.2017.4.
5. 14. T. N. Basu, D. H. Gutmann, J. A. Fletcher, T. W. Glover, F. S. Collins, and J. Downward. Nf1 tumor suppressor critically regulates tsc2 and mtor. Nature, 490(7417):519–25, 2012. ISSN 1476-4687 (Electronic) 0042-8691 (Print). doi:10.1038/nature11412.
6. 13. E. Kang, Y. M. Kim, G. H. Seo, A. Oh, H. M. Yoon, Y. S. Ra, E. K. Kim, H. Kim, S. H. Heo, G. H. Kim, J. M. Osborn, J. Tolar, H. W. Yoo, and B. H.Lee. Phenotype categorization of neurofibromatosis type i and correlation to nf1 mutation types. Proc Natl Acad Sci U S A, 109(18):7173–8, 2012. ISSN 0027-084X (Print) 0027-0844 (Linking). doi:10.1073/pnas.1106564109.
7. 12. J. G. Tate, S. Bamford, H. C. Gedze, M. J. Sondka, D. M. Nair, M. B. U. R. V. C. Monnas, and D. R. E. M. H. P. S. S. J. C. L. and H. V. performed crosslinking and mass spectrometry. All authors contributed to analyses of data and writing of the manuscript.
and learning. Cell 135(3):549–60, 2008. ISSN 1097-4172 (Electronic) 0092-8674 (Linking). doi: 10.1016/j.cell.2008.09.060.

18. M. Holzel, S. Huang, J. Koster, I. Ora, A. Lakanen, H. Caron, W. Nikjam, J. Xie, T. Callens, S. Asgharzadeh, R. C. Seeger, L. Messiaen, R. Versteeg, and R. Bernards. NF1 is a tumor suppressor in neuroblastoma that determines retinoid acid response and disease outcome. Cell 142(2):289–210, 2010. ISSN 1097-4172 (Electronic) 0092-8674 (Linking). doi: 10.1016/j.cell.2010.04.004.

19. M. A. Wolman, E. D. de Grish, S. M. McBride, T. A. Jongens, M. Granato, and J. A. Epstein. Mutation of a ras pathway signaling pathway causes distinct behavioral deficits in a zebrafish model of neurofibromatosis type 1. Cell 85(5):1265–70, 2014. ISSN 2212-1247 (Electronic). doi: 10.1016/j.cleareye.2014.07.054.

20. W. Yan, E. Markegard, S. Dharmiah, A. Urisman, M. Drew, D. Esposito, K. Schef- feld, and M. Walz. NF1 encodes a novel Ras GTPase-activating protein that modulates growth factor signaling. J Cell Sci 119(20):4373–84, 2006. ISSN 1477-9137. doi: 10.1242/jcs.20060945.

21. X. Liu, L. Calgaro, M. T. Dao, B. S. Muntean, L. P. Sutton, C. Orlandi, S. L. Boye, S. E. Boye, C. C. Shi, Y. Li, B. Xu, R. G. Smith, Y. Yasuda, and K. A. Martemyanov. NF1 is a direct G protein effector for opioid signaling to ras in the striatum. Curr Biol 26 (22):2982–3003, 2016. ISSN 0960-9822 (Linking). doi:10.1016/j.cub.2016.08.029.

22. X. Koliou, C. Fedorinis, T. Kalpachidis, and D. Mangoura. Nuclear import mechanism of neurofibromin for localization on the spindle and function in chromosome congression. J Neurochem 136(1):78–91, 2016. ISSN 1471-4159 (Electronic) 0022-3042 (Linking). doi:10.1111/jn.13441.

23. Lping Feng, Shunjuns Unosite, Hiroshi Tokou, Tatsuya Owase, Dongwei Zhang, Siporn Patratikitkomjorn, Toru Ichimura, Hideyuki Saya, and Norie Araki. PKA phosphorylation and 14-3-3 interaction regulate the function of neurofibromin type 1 suppressor of neurofibromatosis. FEBS Lett 557(1-3):275–282, 2004. ISSN 0046-2672 (Linking). doi:10.1016/S0014-5793(04)00310-2.

24. D. Mangoura, Y. Sun, C. Li, D. Singh, D. H. Gutmann, A. Flores, M. Ahmed, and G. Val- larodriguez. Phosphorylation of neurofibromin by p21 is a possible molecular switch in the receptor signaling to ras in neurons. Oncogene. 25(5):735–46, 2006. ISSN 0950-9232 (Print) 1476-5958 (Electronic). doi:10.1038/sj.onc.1209372.

25. S. H. Yoshimura and T. Hirano. Heat repeats - versatile arrays of amphiphilic helices working in crowded environments? J Cell Sci 129(21):3893–3970, 2016. ISSN 1477-9137. doi:10.1242/jcs.160533.

26. T. Dunzendorfer-Matt, E. L. Mercado, K. Maly, F. McCormick, and K. Scheffzek. The neurofibromin-2/SDH complex and the balance of Krebs cycle and Ras signaling in neural cells. J Cell Sci. 130(22):4313–4326, 2017. ISSN 1471-4159 (Electronic) 0022-3042 (Linking). doi:10.1242/jcs.189991.

27. D. Mangoura, Y. Sun, C. Li, D. Singh, D. H. Gutmann, A. Flores, M. Ahmed, and G. Val- larodriguez. Phosphorylation of neurofibromin by p21 is a possible molecular switch in the receptor signaling to ras in neurons. Oncogene. 25(5):735–46, 2006. ISSN 0950-9232 (Print) 1476-5958 (Electronic). doi:10.1038/sj.onc.1209372.
Fig. S1. Purification and biochemical characterisation of the NF1 dimer. 

a. Schematic of the inhibitory role of NF1 in RAS signalling. 

b. Select NF1 binding partners and cellular effectors are listed.

- RAS
- SPRED1
- FAK
- PI
- ETEA
- LRRP/CUL1/FBXW7
- Kinesin-1
- Cullin 3
- APP
- Caveolin-1
- LIMK2
- Syndecan
- Tubulin
- DDHAH
- VCP

2. ** p<0.01 versus Ras-GTP alone, paired two-tailed t-test.

3. Sedimentation velocity analytical ultracentrifugation analysis indicates that NF1 likely forms a homodimer in solution (calculated molecular weight of 640 kDa, r.m.s.d. 0.004).

NF1 accelerates the hydrolysis of GTP by KRAS. Hydrolysis was measured using a GTPase-Glo kit (Promega) to detect the amount of GTP remaining in the reaction after incubation. ** p<0.01 versus Ras-GTP alone, paired two-tailed t-test.

NF1 purifies as a single peak by size-exclusion chromatography.

Purified recombinant NF1 is highly pure by SDS-PAGE.

Sedimentation velocity analytical ultracentrifugation analysis indicates that NF1 likely forms a homodimer in solution (calculated molecular weight of 640 kDa, r.m.s.d. 0.004).
Fig. S2. Cryo-EM summary and validation for the NF1 wing localised reconstruction. a. A representative denoised micrograph of NF1 collected on a Talos Arctica 200 kV. b. Selected 2D class averages of the NF1 dimer. Diffuse signal can be observed at the central interface corresponding to flexible GRD/SEC/PH domains. c. Select 2D class averages of symmetry expanded, localised reconstructions of a single NF1 dimer lobe. d. Angular distribution of the final localised reconstruction, shown beside is the corresponding orientation of the reconstruction for clarity. e. The NF1 dimer lobe coloured by local resolution. f. Distribution of voxels as a function of local resolution. The majority of voxels are resolved to a mode value of ~4.5 Å. g. The Fourier shell correlation plot estimates a global resolution of 4.5 Å.
Fig. S3. Cryo-EM summary and validation for the reconstructions of NF1 autoinhibited and dimeric states. 

a. Angular distribution of the autoinhibited NF1 localised reconstruction, b., and local resolution analysis. c. Frequency distribution of voxels as a function of local resolution for the autoinhibited NF1 localised reconstruction. d. The Fourier shell correlation plot estimates a global resolution of 5.6 Å. e, f, g. As in a, b, c. for the full NF1 homodimer. h. The Fourier shell correlation plot estimates a global resolution of 6.3 Å.
Initial cryo-EM analysis is indicated in purple. Subsequent analysis of the whole dataset gave rise to a 7.5 Å reconstruction of the NF1 dimer (green). Symmetry expansion and signal subtraction were utilized to generate a focused refinement of a single NF1 lobe (light blue) to overcome significant continuous conformational heterogeneity. The final 4.6 Å reconstruction was later fitted into the NF1 dimer reconstruction to yield a composite map. Variational analysis revealed a small discrete population of NF1 particles with resolved GRD and SEC-PH domains (red) in an autoinhibited conformation.
Fig. S5. Map to model fit.  

a. Map-to-model fit of the SEC-PH domain (blue). 

b. Map-to-model fit of the GRD domain, clearly resolved secondary structure is evident. 

c. The autoinhibited conformation of NF1, GRD (green) and SEC-PH (blue) domains are well resolved against the scaffold. 

d. The sharpened 4.6 Å reconstruction of the core NF1 scaffold and select regions of density illustrating clearly resolved side chains.
Fig. S6. Crosslinking mass spectrometry of NF1.  

a. Crosslinking mass spectrometry identified a total of 174 BS^2G crosslinks and 72 BS^3 crosslinks within the NF1 dimer.

b. Mapping of the BS^2G crosslinks onto the NF1 structure revealed that the majority of all crosslinks mapped to allowable regions of the NF1 model. Cα carbons of crosslinked residues are shown as spheres and lines indicate crosslinked residues.

c. Circos plot highlighting all BS^3 and BS^2G crosslinks observed between the GRD domain and the rest of the NF1 molecule. GRD crosslinked regions map remarkably well to the two known NF1 phosphorylation clusters(51). Known NF1 phosphorylation sites(52) are indicated as dashed red lines on the inside of the circos plot.

d. Although these regions are too flexible to appear in the cryo-EM maps, they are each situated adjacent to the GRDs and appear poised to further regulate either the conformation of the GRD or membrane binding capacity. Structural representation of the NF1 surface with the GRD and SEC-PH domains in cartoon format. The Cα carbon of the exit (C2432) and entry (P2596) residue of the main phospho-regulatory loop are indicated as yellow spheres. The Cα carbon of the final resolved residue of the C-terminus (L2726) is also indicated as a yellow sphere. Cα carbons on the GRD that crosslink to the phosphoregulatory loops are indicated as pink spheres.
Fig. S7. Mapping of known binding sites to the NF1 structure. The position of select binding partners of NF1 (grey) are mapped onto the structure (blue)(2).

Fig. S8. Surface conservation of NF1. NF1 surface conservation analysed using ConSurf(53). a. The NF1 dimer with surface conservation mapped to a single chain. b. NF1 rotated 180° to show surface conservation of the second chain. c. A single NF1 chain displayed to show the conserved central interface.
Fig. S9. NF1 multiple sequence alignment and secondary structure assignment. Multiple sequence alignment of NF1. Conserved regions are highlighted in grey with the secondary assignment from the NF1 structure indicated above the alignments. Key sites of phosphorylation(51) and dimer interfaces are indicated.
Table S1. Cryo-electron microscopy data collection, processing and analysis statistics.

| Data collection and processing | NF1 |  
|-------------------------------|-----|
| **Microscope**                | Talos Arctica |  
| **Voltage (keV)**             | 200 |  
| **Camera**                    | Falcon III |  
| **Mode**                      | Counting |  
| **Exposure navigation**       | Beam till |  
| **Set defocus range (μm)**    | -0.5 to -2.0 |  
| **Measured defocus (μ ± σ) (μm)** | -1.2 ± 0.7 |  
| **Exposure rate (e⁻ / pixel / s)** | 0.8 |  
| **Frames**                    | 50 |  
| **Exposure duration (s)**     | 71.68 |  
| **Total exposure (e⁻ / Å²)**  | 40.22 |  
| **Magnification**             | 120k |  
| **Starting pixel size (Å / pixel)** | 1.194 |  
| **Micrographs**               | 2,675 |  
| **Particles (Total extracted)** | 193,225 | 191,128<sup>+</sup> |  
| **Particles (Final reconstruction)** | 95,584 | 34,993<sup>+</sup> | 9,238<sup>+</sup> |  
| **Reconstruction Strategy**   | **Standard** | **Symmetry expand / Localised reconstruction** | **3DPCA clustering / Local refinement** |  
| **Binning factor**            | 2 | 1 | 1 |  
| **Imposed symmetry**          | C2 | C1 | C1 |  
| **Global Resolution (Å)**     | 0.143 (Unmasked / Masked) | 7.6/5.6 | 5.18/4.56 | 8.0/6.3 |  
| **Local resolution variation** | 5.3-19.2 | 7.3-4.2 | 4.58-15.8 |  
| **Global B-factor (Å<sup>2</sup>) (manual)** | -200 | -80 | -100 |  
| **EMDB identifier**<sup>*</sup> | EMBD XXXX | EMBD XXXX | EMBD XXXX |  

<sup>*</sup>As determined by windowed FSC at 0.5 criterion in Relion 3.1 or cryoSPARC 3.0.  
<sup>+</sup>Mask, half-maps, unsharpened map, final map where applicable.  
<sup>+</sup>Sub/symmetry expanded particles

| Model statistics | NF1 homodimer core | NF1 wing | NF1 autoinhibited state |  
|------------------|---------------------|----------|-------------------------|
| **Model-to-map** | **Correlation Coefficient (volume)** | 0.71 | 0.70 | 0.62 |  
|                  | **Resolution Estimates (Å, FSC 0.143 / 0.5)** | 5.6/7.6 | 3.7/5.2 | 4.6/7.4 |  
|                  | **B factors (Å<sup>2</sup>)** |  
|                  | **Protein** | 125.47 | 127.23 | 115.23 |  
|                  | **Ligand** | N/A | N/A | N/A |  
| **R.M.S. deviation** | **Bonds (Å)** | 0.005 | 0 | 0.005 |  
|                  | **Angles (°)** | 1.067 | 0.01 | 1.08 |  
| **All-atom clash score** | **Ramachandran plot (%)** | 33.9 | 31.6 | 16.1 |  
|                  | **Favoured** | 94.74 | 94.76 | 95.27 |  
|                  | **Allowed** | 5.26 | 5.24 | 4.68 |  
|                  | **Outliers** | 0 | 0 | 0.05 |  
|                  | **CaBLAM outlier (%)** | 3.97 | 2.48 | 3.10 |  
|                  | **Rotamer outliers (%)** | 0 | 0.14 | 0.96 |  
|                  | **C-beta deviations (%)** | 0 | 0 | 0 |  
|                  | **Molprobity Score** | 2.21 | 2.34 | 2.04 |  
| **PDB identifier** | **PDB XXXX** | **PDB XXXX** | **PDB XXXX** |  

<sup>+</sup>Mask, half-maps, unsharpened map, final map where applicable.  
<sup>+</sup>Sub/symmetry expanded particles
Table S2. NF1 crosslinking mass spectrometry sites.

See Extended_data_table_2.xlsx
| Residue 1 | Protein 1 | Residue 2 | Protein 2 | # XLs |
|-----------|-----------|-----------|-----------|-------|
| 33 N      | 77 N      |           |           | 12    |
| 48 N      | 99 N      |           |           | 6     |
| 48 N      | 2468 C    |           |           | 3     |
| 48 N      | 2515 C    |           |           | 6     |
| 63 N      | 111 N     |           |           | 9     |
| 63 N      | 2499 C    |           |           | 2     |
| 77 N      | 469 N     |           |           | 2     |
| 77 N      | 471 N     |           |           | 2     |
| 77 N      | 478 N     |           |           | 3     |
| 77 N      | 2499 C    |           |           | 2     |
| 99 N      | 111 N     |           |           | 11    |
| 99 N      | 189 N     |           |           | 21    |
| 99 N      | 2499 C    |           |           | 3     |
| 99 N      | 2515 C    |           |           | 3     |
| 99 N      | 2521 C    |           |           | 2     |
| 111 N     | 2499 C    |           |           | 6     |
| 111 N     | 2515 C    |           |           | 6     |
| 111 N     | 2521 C    |           |           | 3     |
| 195 N     | 248 N     |           |           | 2     |
| 206 N     | 261 N     |           |           | 366   |
| 261 N     | 305 N     |           |           | 9     |
| 263 N     | 305 N     |           |           | 19    |
| 449 N     | 469 N     |           |           | 4     |
| 449 N     | 471 N     |           |           | 6     |
| 449 N     | 478 N     |           |           | 4     |
| 449 N     | 480 N     |           |           | 3     |
| 449 N     | 490 N     |           |           | 2     |
| 469 N     | 480 N     |           |           | 9     |
| 469 N     | 490 N     |           |           | 11    |
| 469 N     | 584 N     |           |           | 2     |
| 469 N     | 662 N     |           |           | 2     |
| 469 N     | 789 N     |           |           | 2     |
| 471 N     | 490 N     |           |           | 13    |
| 471 N     | 584 N     |           |           | 3     |
| 476 N     | 490 N     |           |           | 4     |
| 478 N     | 490 N     |           |           | 20    |
| 490 N     | 810 N     |           |           | 2     |
| 505 N     | 2515 C    |           |           | 5     |
| 609 N     | 2468 C    |           |           | 7     |
| 609 N     | 2499 C    |           |           | 4     |
| 609 N     | 2514 C    |           |           | 2     |
609 N 2515 C 4
609 N 2531 C 3
609 N 2547 C 2
613 N 2468 C 3
613 N 2515 C 2
613 N 2521 C 2
615 N 662 N 3
615 N 1436 GRD 5
615 N 2468 C 4
615 N 2499 C 4
615 N 2514 C 3
615 N 2515 C 2
662 N 874 N 4
662 N 1419 GRD 4
757 N 2468 C 2
757 N 2514 C 2
757 N 2521 C 2
789 N 810 N 19
789 N 874 N 7
789 N 893 N 43
796 N 810 N 14
796 N 874 N 3
796 N 908 N 9
798 N 908 N 12
810 N 908 N 2
874 N 938 N 32
874 N 943 N 10
874 N 1036 N 6
874 N 1408 GRD 12
874 N 1419 GRD 10
874 N 1423 GRD 6
874 N 1560 SECPH 9
874 N 1567 SECPH 5
918 N 938 N 2
918 N 1014 N 7
938 N 1036 N 2
1036 N 1062 N 2
1036 N 1560 SECPH 39
1036 N 1704 SECPH 7
1040 N 1062 N 4
1345 GRD 2521 C 3
1419 GRD 2468 C 3
1448 GRD 2515 C 3
1496 GRD 1508 GRD 6
1496 GRD 1567 SECPH 5
1496 GRD 1704 SECPH 2
1496 GRD 2468 C 3
1496 GRD 2499 C 6
1496 GRD 2515 C 5
1496 GRD 2802 C 3
1508 GRD 2139 C 13
1508 GRD 2286 C 3
1508 GRD 2401 C 18
1508 GRD 2407 C 14
1508 GRD 2468 C 5
1508 GRD 2499 C 6
1508 GRD 2514 C 3
1508 GRD 2515 C 5
1508 GRD 2521 C 7
1508 GRD 2531 C 3
1531 GRD 1634 SECPH 10
1567 SECPH 1704 SECPH 33
1634 SECPH 1823 C 30
1634 SECPH 1827 C 11
1634 SECPH 1946 C 21
1634 SECPH 2515 C 2
1634 SECPH 2802 C 5
1634 SECPH 2810 C 2
1670 SECPH 1728 SECPH 26
1670 SECPH 2802 C 3
1670 SECPH 2810 C 2
1680 SECPH 2802 C 3
1693 SECPH 1735 SECPH 10
1728 SECPH 2810 C 2
1731 SECPH 2802 C 5
1731 SECPH 2810 C 10
1823 C 1946 C 22
1827 C 1946 C 43
1915 C 1956 C 4
1940 C 2028 C 10
1940 C 2810 C 2
1946 C 2802 C 5
1946 C 2810 C 5
1956 C 2006 C 7
1956 C 2802 C 3
1956 C 2810 C 4
| 2515 C | 2810 C | 1  |
| 2521 C | 2547 C | 12 |
| 2531 C | 2547 C | 9  |
| 2699 C | 2802 C | 5  |