



HAL
open science

An Atypical Mechanism of SUMOylation of Neurofibromin SecPH Domain Provides New Insights into SUMOylation Site Selection

Mohammed Bergoug, Christine Mosrin-Huaman, Amandine Serrano, Fabienne Godin, Michel Doudeau, Iva Susic, Stéphane Goffinont, Thierry Normand, Marcin J. Suskiewicz, Béatrice Vallée, et al.

► To cite this version:

Mohammed Bergoug, Christine Mosrin-Huaman, Amandine Serrano, Fabienne Godin, Michel Doudeau, et al.. An Atypical Mechanism of SUMOylation of Neurofibromin SecPH Domain Provides New Insights into SUMOylation Site Selection. *Journal of Molecular Biology*, 2024, 436 (22), pp.168768. 10.1016/j.jmb.2024.168768 . hal-04771131

HAL Id: hal-04771131

<https://hal.science/hal-04771131v1>

Submitted on 7 Nov 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

An Atypical Mechanism of SUMOylation of Neurofibromin SecPH Domain Provides New Insights into SUMOylation Site Selection

Mohammed Bergoug[#], Christine Mosrin[#], Amandine Serrano, Fabienne Godin, Michel Doudeau, Iva Susic, Stephane Goffinont, Thierry Normand, Marcin J. Suskiewicz, Béatrice Vallée, and Hélène Bénédetti*

Centre de Biophysique Moléculaire, CNRS, UPR 4301, University of Orléans and INSERM, 45071 Orléans Cedex 2, France

[#] Contributed equally to the work and are co-first authors

* Corresponding author

Hélène Bénédetti

Centre de Biophysique Moléculaire, CNRS, UPR 4301, University of Orléans and INSERM, 45071 Orléans Cedex 2, France

Tel : +33 2 38 25 55 84

E-mail: helene.benedetti@cns-orleans.fr

Highlights

Partial co-localisation of Nf1 with PML NB suggested Nf1 SUMOylation

Nf1 isoform 2 SUMOylation by SUMO2/3 was established

Two lysines of Nf1 SecPH domain, K1634 and K1731, were shown to be SUMOylated

K1731 SUMOylation did not depend on the SCM, nor on SecPH SIM domain or Ubc9 auto-SUMOylation

Instead, K1731 SUMOylation required two SecPH structural elements

A new example of SUMO site selection depending on the tertiary target protein structure was revealed

ABSTRACT limit of 250 words

Neurofibromin (Nf1) is a giant multidomain protein encoded by the tumour-suppressor gene *NF1*. *NF1* is mutated in a common genetic disease, neurofibromatosis type I (NF1), and in various cancers. The protein has a Ras-GAP (GTPase activating protein) activity but is also connected to diverse signalling pathways through its SecPH domain, which interacts with lipids and different protein partners. We previously showed that Nf1 partially colocalized with the ProMyelocytic Leukemia (PML) protein in PML nuclear bodies, hotspots of SUMOylation, thereby suggesting the potential SUMOylation of Nf1. Here, we demonstrate that the full-length isoform 2 and a SecPH fragment of Nf1 are substrates of the SUMO pathway and identify a well-defined SUMOylation profile of SecPH with two main modified lysines. One of these sites, K1731, is highly conserved and surface-exposed. Despite the presence of an inverted SUMO consensus motif surrounding K1731, and a potential SUMO-interacting motif (SIM) within SecPH, we show that neither of these elements is necessary for K1731 SUMOylation, which is also independent of Ubc9 SUMOylation on K14. A 3D model of an interaction between SecPH and Ubc9 centred on K1731, combined with site-directed mutagenesis, identifies specific structural elements of SecPH required for K1731 SUMOylation, some of which are affected in reported *NF1* pathogenic variants. This work provides a new example of SUMOylation dependent on the tertiary rather than primary protein structure surrounding the modified site, expanding our knowledge of mechanisms governing SUMOylation site selection.

Keywords (5)

Neurofibromin (Nf1) and Nf1-SecPH domain SUMOylation; Model of Nf1-SecPH/Ubc9 interaction; Ubc9 fusion-directed SUMOylation, Mechanism of SUMOylation site selection, SCM- and SIM-independent SUMOylation; SUMOylation site selection dependent of 3D structure.

Footnote Abbreviations

CDM, Central Dimerisation Module

E-64, Proteinase Inhibitor E 64

FBS, Fetal Bovine Serum

GAP, GTPase-activating protein

GRD, GAP-related domain

LIM, domain name is an acronym of the three genes in which it was first identified (LIN-11, Isl-1 and MEC-3)

LIMK2, LIM kinase-2

NEM, N-Ethylmaleimide

NTD, N terminal Domain

PDSM, Phospho-dependent SUMOylation Motif

PH, Pleckstrin Homology

PtdEtn, Phosphatidylethanolamine

PEI, polyethylenimine

PTM, Post Translational Modification

SCM, SUMO Consensus Motif

Sec, yeast Sec-14 phospholipid interaction module

SENP, Sentrin-specific protease

SIM, SUMO-Interacting Motif

Ubc9, Sumo E2 Enzyme

Ubl, Ubiquitin like

E2-25K, ubiquitin-conjugating enzyme E2-25K

UFDS, Ubc9 fusion-directed SUMOylation

VCP, Valosin Containing Protein

5-HT₆, serotonin 5 hydroxytryptamine 6

INTRODUCTION

SUMOylation is a protein post-translational modification (PTM) that involves the covalent enzymatic conjugation of the SUMO (Small Ubiquitin-like Modifier) protein to specific lysine residues of substrate proteins [1, 2]. Several SUMO paralogs can be found in human cells, the main ones being SUMO2 and SUMO3, which are almost identical at the amino-acid sequence level, and SUMO1, which is less than 50% identical to SUMO2/3. SUMOylation involves a dedicated enzymatic pathway similar to that for ubiquitylation, comprising the heterodimeric

E1-activating enzyme SAE1-SAE2, the sole E2 conjugating enzyme Ubc9, and, optionally, various E3 ligases that can enhance the yield and specificity of the reaction [3]. SUMOylation is a highly dynamic modification due to the opposing activities of its ligation and removal, the latter being performed by Sentrin-specific proteases or SENPs (where Sentrin is an alternative SUMO name) [4]. This dynamics results in a low apparent level of SUMO modification on most substrates, which nevertheless can induce major alterations in protein functions. Different hypothesis have been put forward to explain this paradox, one of them being that SUMO attachment promotes a change in the properties of the protein that persists upon deSUMOylation [5, 6]. SUMOylation effectors are involved in all fundamental cellular processes and are also over-represented in proteins implicated in neurodegenerative and cancer-related processes [7, 8]. At the molecular level, SUMOylation can modulate the activity, stability, solubility, structure, subcellular localization, and interaction network of target proteins [7, 9]. Furthermore, the SUMO system has been shown to be a key organizer of assembly and dynamics of PML (ProMyelocytic Leukemia) bodies, multiproteic nuclear structures nucleated by the PML protein, which are implicated in a wide range of biological processes [10].

This work is focused on SUMOylation of Neurofibromin (Nf1), which our group has previously shown to partially colocalize with the PML protein in PML nuclear bodies, thereby suggesting its SUMOylation [11]. Nf1 is encoded by the *NF1* gene, the mutations in which are responsible for neurofibromatosis type I (NF1), a genetic disease that affects approximately one individual in 3,500. The disease is characterized by a wide spectrum of manifestations such as development of peripheral and central nervous system tumours, pigmentary lesions, skeletal abnormalities and learning disabilities (for reviews see [12-15]). *NF1* is also mutated in various types of cancer [16-20]. More than 3,000 different germline *NF1* mutations have been reported in the Human Gene Mutation Database [21], revealing a substantial diversity of pathogenetic mutations with very few hotspots [13, 22, 23].

Nf1 is a giant dimeric [24] multidomain protein (for a review of Nf1 structure, function, and regulation, see [25]). It consists of an N-terminal domain, a central Ras-GAP (GTPase-activating protein)-related domain (GRD), followed by a SecPH domain (a bipartite module composed of a segment homologous to the yeast Sec-14 phospholipid interaction module and a pleckstrin homology (PH)-like domain), a Central Dimerisation Module (CDM) and a C-terminal domain (CTD). Recently, the three-dimensional (3D) structure of the entire Nf1 has been elucidated by cryo-EM, revealing the formation of a homodimer with two distinct conformations: a closed auto-inhibited conformation with occluded Ras-binding site and an asymmetric open and active conformation with an exposed Ras-binding site [26-28].

Prior to the recent cryo-EM work, GRD and SecPH domains had long been the only Nf1 domains with resolved 3D structures [29, 30], and they remain the most investigated regions of Nf1 in terms of activity and function. Nf1 GRD domain is responsible for the best studied function of Nf1, its Ras-GTPase activating protein (Ras-GAP) activity, which consists in stimulating Ras-GTPase activity, thereby converting active Ras-GTP into inactive Ras-GDP. However, Nf1 carries out a number of other functions, the main ones being the regulation of cAMP levels in both a Ras-dependent and Ras-independent manner [31-33] and the modulation of actin cytoskeletal dynamics through the negative regulation of the Rho/ROCK/LIMK2/cofilin [34] and Rac1/Pak1/LIMK1/cofilin pathways [35]. Nf1 SecPH domain, on the other hand, is able to bind phospholipids [30, 36] and actors of different signalling pathways (LIMK2, 5-HT₆ Receptor, VCP) [34, 37, 38]. From a functional point of view, it has been hypothesized that SecPH allosterically regulates the Ras-GAP activity of the GRD and mediates new functions of Nf1 [34, 37, 38].

Different transcript variants resulting from the alternative splicing of *NF1* pre-mRNA have been identified. The most abundant form of *NF1* mRNA is the one containing 57 exons and encoding a 2818 amino-acids Nf1 protein called isoform 2 [39]. Isoform 1 contains 21 additional amino-acids in the GRD due to the unspliced 30alt31 exon (formerly exon 23a) corresponding to 63 nucleotides. The resulting protein comprises 2839 amino-acids and has a reduced Ras-GAP activity [39, 40].

Mass spectrometry-based global studies of SUMOylation in cultured human cells have identified Nf1 isoform 1 as a potential SUMO target protein [5], and two SUMOylated lysines belonging to the isoform 1-specific 30alt31 exon were identified [41, 42]. However, the SUMOylation of Nf1 isoform 2, the most studied Nf1 isoform which harbours the higher Ras-GAP activity, has never been demonstrated.

Here, we show that the full-length isoform 2 and an isolated SecPH domain of Nf1 are SUMOylated with the SUMO2 paralog. SecPH exhibits a specific SUMOylation profile in immunoblots, with two main bands corresponding to SUMOylation on K1731 and K1634. K1731, which appears to be the major of the two sites, is highly conserved and solvent-exposed in both known Nf1 conformations. Our analysis reveals the dependence of K1731 SUMOylation on the tertiary rather than primary structure surrounding the site, providing new insights into mechanisms of SUMOylation site selection.

RESULTS

1. The SecPH domain of Nf1 is SUMOylated

Nf1 isoform 2 is a large multidomain protein of 2818 residues. Two SUMOylation prediction software programs, SUMOPlot (<https://www.abgent.com/sumoplot>) and JASSA [43], were used to analyse its sequence, resulting in the prediction of 13 instances of a SUMOylation consensus motif (SCM), (Ψ KxE/D) [44] or inverted (E/DxK Ψ) consensus motif [45] (where Ψ stands for a hydrophobic residue with high preference for I or V) with a high score for at least one software program (Figure 1a).

We decided to focus on the SecPH domain of Nf1 to further characterize its SUMOylation for several reasons : (i) the SecPH domain contains the highest number of the predicted SCMs and inverted consensus motifs, (ii) its 3D structure has been studied in detail [36, 46] (iii) the recently resolved Cryo-EM structures of Nf1 dimer show that it is exposed at the surface of the dimer and, together with its adjacent GRD domain, is subject to a structural rearrangement [26-28], and (iv) pleckstrin homology (PH) domains have been reported to be enriched among SUMOylation substrates [47].

We first studied SecPH SUMOylation by co-transfecting HEK293 cells with three plasmids to overproduce Flag-tagged SecPH, Ubc9, and 6His-tagged SUMO2. After lysis in denaturing conditions, SUMOylated proteins were enriched by performing a pull down on cobalt beads. Western blot analysis of eluates using anti SUMO2/3 antibodies showed the 6His pull down to be strongly enriched in SUMOylated proteins (Figure 1b).

Eluates were also analysed by western blotting using anti Flag antibodies. SUMOylated Flag SecPH showed a very reproducible and discrete profile of four bands: three (bands 1, 2, and 3) migrated at approximately 50 kDa, whereas band 4 migrated with a slightly higher molecular weight (Figure 1c). The lower band around 37 kDa (marked by * in Figure 1c) represents the nonSUMOylated Flag SecPH non specifically bound to the cobalt beads. Of note, the detection of nonSUMOylated overproduced proteins in SUMO-enriched fractions is a frequently observed issue with such approaches [48].

2. K1634 and K1731 contribute to the SecPH SUMOylation profile

We next wanted to identify the SUMO conjugation sites within SecPH. Lysine residues belonging to bioinformatically-predicted putative SUMOylation sites were independently mutated into arginine, creating five SecPH mutants: SecPH K1593R, SecPH K1634R, SecPH K1717R, SecPH K1731R, and SecPH K1735R (Figure 2a). The strongest SUMOylation band, band 2, disappeared in SecPH K1731R (Figure 2b). For unknown reasons, other mutants exhibited a stronger signal than WT SecPH, especially SecPH K1634R. In this latter mutant, the resulting broader band 2 partially occluded band 3, hindering the observation of any potential changes in band 3 intensity. Thus, we constructed a SecPH K1634R K1731R double

mutant and compared its SUMOylation profile to that of each corresponding SecPH single mutant. Western blot in Figure 2c shows that band 2 completely disappeared and band 3 was much fainter in the double SecPH K1634R K1731R mutant, strongly suggesting that K1731 and K1634 were required for the presence of bands 2 and 3, respectively.

To confirm that bands 1, 2, 3, and 4 corresponded to SUMOylated SecPH, we additionally performed an alternative enrichment approach, which relied on the immunoprecipitation of Flag-tagged SecPH using anti Flag antibody rather than cobalt pull down of 6His-tagged SUMO2-conjugated proteins. Thus, Flag-tagged SecPH and SecPH K1731R were immunoprecipitated from lysates of HEK293 cells co-transfected with the corresponding p3X Flag SecPH vectors, pcDNA3 Ubc9, and pcDNA3 6His-SUMO2. The eluted proteins were revealed by western blotting with anti Flag and anti SUMO2/3 antibodies. First, the anti Flag antibodies revealed the same pattern of bands for immunoprecipitated WT SecPH and pulled down WT SecPH (Figure 3a, upper panel). Second, the four-band profile revealed for WT SecPH with the anti Flag and anti SUMO2/3 antibodies around 50 kDa was very similar (Figure 3a, lower panel), strongly suggesting that these bands correspond to SUMO-modified Flag SecPH. Third, blotting the SecPH K1731R immunoprecipitates with the anti Flag and anti SUMO2/3 antibodies showed that the SUMOylated band migrating as band 2 disappeared for this mutant with both antibodies (Figure 3a, upper and lower panel). Band 2 was the strongest of the four bands and western blot quantification obtained from three independent immunoprecipitation experiments analysed with anti Flag antibodies indicated that it corresponded to 16% of the total Flag SecPH (bands 1,2,3,4 and unSUMOylated band were summed up and then the fraction of the total that corresponded to the SUMOylated band 2 was calculated).

Immunoprecipitation of WT SecPH from cells with endogenous levels of SUMO2 and Ubc9 and long exposure of the anti Flag immunoblot showed a band migrating at the same molecular weight as band 2, indicating that this major SUMOylation, dependent on K1731, may take place at endogenous SUMO levels (Figure 3b). In these conditions, western blot quantification obtained from three independent immunoprecipitation experiments analysed with anti Flag antibodies indicated that it corresponded to a fraction of 8% of the total Flag SecPH (band 2 and unmodified band were summed up and then the fraction of the total that corresponded to the SUMOylated band 2 was calculated). This is consistent with the level of steady-state SUMOylation described for several proteins with endogenous expression level of the SUMO system [49].

Although our analysis provided strong candidates for bands 2 and 3, bands 1 and 4 remained to be identified. We hypothesised that they might come from combinations of SUMOylation

sites, thus precluding their identification using single K-to-R mutants. Therefore, we created a mutant of SecPH in which all lysines in the predicted SUMOylation consensus motifs were simultaneously mutated to arginine (SecPH K1593R, K1634R, K1717R, K1731R, K1735R, called SecPH 5K-R) and tested it for SUMOylation. The SUMOylation profile of this mutant was very similar to that of the Flag SecPH K1634R K1731R double mutant, showing stronger bands 1 and 4 (supplementary Figure **S1**). This result suggests that none of the five lysines predicted as potential SUMO sites plays a role in the SUMOylation events that gave rise to bands 1 and 4.

3. The Flag tag in Flag SecPH is SUMOylated

We delved further into the identification of the 4 bands depicted in Figure 1c. For this purpose, we decided to construct a SecPH mutant with all its lysines (a total of 21) mutated into arginines (SecPH K0). When analysed with our pull down approach, SecPH K0 did not show band 4 anymore. However, surprisingly, a broad SUMOylation band migrating at the same position as band 1 could still be observed (Figure 4a). This result suggests that band 1 may be due to the SUMOylation of a lysine present in the 3X Flag tag. Indeed, Fuseya et al. [50] previously showed that a lysine of the second Flag tag could be ubiquitylated. JASSA software predicted the two lysines of the last Flag tag to be putative SUMOylation sites (Figure 4b). Therefore, we replaced these two lysines with arginines in the constructs encoding Flag SecPH WT (2KR Flag SecPH WT) and Flag SecPH K0 (2KR Flag SecPH K0). Band 1 disappeared from 2KR Flag SecPH WT, whereas 2KR Flag SecPH K0 showed no SUMOylation band (Figure 4c, right panel). Altogether, these additional experiments show that one or both of the lysines of the third Flag tag are SUMOylated and contribute to band 1, whereas band 4 is likely due to the SUMOylation of one or more of the non-consensus lysines present within SecPH. We did not further investigate the SUMO site(s) corresponding to band 4, rather focusing on sites responsible for bands 2 and 3.

4. K1634 is a minor and K1731 a major site of SUMOylation in SecPH

To reinforce our results and confirm that K1731 and K1634 are SUMOylated, giving rise to bands 2 and 3, we reintroduced these sites into the 2KR Flag SecPH K0 mutant and studied the SUMOylation profile of these SecPH derivatives using our pull down approach.

In line with our expectations, the reintroduction of K1731 into 2KR Flag SecPH K0 mutant resulted in the recovery of the SUMOylation band migrating like band 2 (Figure 5a), while further reintroduction of K1634 into this mutant resulted in the recovery of the SUMOylation band migrating like band 3 (Figure 5a). Thus, the two lysines are not only necessary, but also sufficient for the production of the two bands, which is best explained by them being direct

targets of SUMOylation. As a negative control, another lysine, K1735, was reintroduced into 2KR Flag SecPH K0, and, as expected, we did not observe any SUMOylation bands (Figure 5b), which validated the specificity of the effects observed with K1731 and K1634. Concerning the faint band migrating as band 3 in SecPH K1634R K1731R and SecPH 5K-R, we think it might result from a compensatory phenomenon, where another SecPH lysine, not among the five with a consensus SUMOylation sequence, undergoes faint SUMOylation

We then examined the 3D structure of SecPH [36] and Nf1 isoform 2 dimer in its two conformations to visualize the location of K1634 and K1731 [26]. Both K1634 and K1731 are located on the surface of the SecPH 3D structure but on opposite sides (Figure 6a). In the context of the Nf1 dimer, K1731 is located on the surface and accessible whether the molecule is in the open or closed conformation. In contrast, K1634 is only accessible when the molecule is in the open conformation (Figure 6b).

5. Endogenous Nf1 and exogenously produced Nf1 isoform 2 are SUMOylated by SUMO2

Global studies of SUMOylation based on bottom-up mass spectrometry have identified Nf1 isoform 1 as a potential SUMO target protein [5] and two SUMOylated lysines belonging to the isoform 1-specific exon 30alt31 (formerly 23a) were shown to be SUMO modified [5]. However, no sites have been identified to date within the main Nf1 variant, isoform 2. While bottom-up mass spectrometry is a powerful technique for SUMO site identification, it is known to provide a limited coverage of the proteome, necessitating the use of other techniques in some cases. Since our data obtained for the isolated SecPH domain strongly suggests its SUMOylation on multiple sites, we wondered if the full-length isoforms 1 and 2 harbouring this domain are SUMOylated in cells.

In a first approach, we tested if we could biochemically detect SUMOylation of endogenous Nf1 naturally produced by HEK cells in conditions of Ubc9 and SUMO2 overproduction. For this purpose, HEK cells were co-transfected with two plasmids allowing the overproduction of the SUMO-conjugating enzyme, Ubc9, and Flag-tagged SUMO2. Mock-transfected cells were used as a negative control. In both types of cells, equal levels of a band migrating at a molecular mass of approximately 300 kDa, the size corresponding to Nf1 (isoforms 1 and 2), was detected by western blotting (Figure 7a). Flag-SUMO2 (18 kDa) was well expressed and integrated into several different SUMOylated proteins (Figure 7b).

Subsequently, we first enriched SUMOylated or SUMO-interacting proteins from cell lysates by performing immunoprecipitation with anti Flag antibodies. Unfortunately, we could not reveal any Nf1 in the eluates, probably due to it being below the detection threshold

(unpublished results). We then decided to enrich these eluates in Nf1-derived proteins by performing a second round of immunoprecipitation, this time using anti Nf1 antibodies. The second immunoprecipitation eluate was analysed by western blotting with anti Flag and anti Nf1 antibodies. A band corresponding to the molecular mass of Nf1 was revealed by both antibodies only in the presence of Flag-SUMO2 and Ubc9 (Figure 7c,d). The detection of a band with the same apparent molecular mass as Nf1 using the anti SUMO2,3 antibodies demonstrated that Nf1 was conjugated to Flag-SUMO2 and not simply interacted with SUMO2-conjugated proteins. However, since isoforms 1 and 2 of Nf1 have similar sizes, we were not able to conclude if one or both of these isoforms appear to be SUMOylated.

Isoform 2 was of a particular interest to us because it is the major endogenous Nf1 isoform with a higher Ras-GAP activity and lacks the exon 30alt31 where SUMO sites have been previously detected. Therefore, isoform 2 SUMOylation would indicate the existence of previously undetected SUMOylation sites within Nf1, possibly including those we identified above using the isolated SecPH domain. Given that Nf1 isoform 2 contains 169 lysines, including 13 within SUMO consensus sequences, we tested its SUMOylation and compared it to that of Nf1 K1731R and K1634R K1731R mutants to investigate the role of these specific lysines in Nf1 isoform 2 SUMOylation. For these experiments, we transiently expressed tagged forms of Nf1 isoform 2 derivatives from plasmids. Plasmids were constructed based, in part, on a plasmid containing the codon optimized *NF1* cDNA developed by Sherekar and colleagues [24]. As we had shown above that the Flag tag can be SUMOylated (see § 3), we decided to work with 2KR Flag Nf1 constructs in which the lysines of the Flag tag that are prone to being SUMOylated were mutated to arginines. These plasmids were co-transfected into HEK293 cells along with plasmids allowing the overproduction of Ubc9 and 6His-tagged-SUMO2. After anti Flag immunoprecipitation, western blot analysis of eluates was performed using anti Flag and anti SUMO2/3 antibodies (Figure 7e,f). Anti SUMO2/3 antibodies revealed bands only with SUMO2 and Ubc9 overproduction, one corresponding to the molecular mass of Nf1 for wild-type Nf1 isoform 2 and the two mutants (black arrow). Lower molecular mass bands were also detected by these antibodies but not by anti Flag antibodies suggesting they might be SUMO-modified proteins co-immunoprecipitated with Nf1. These results demonstrate that exogenously produced 2KR Flag wild-type and mutated Nf1 isoform 2 are modified by SUMO2. Although this does not confirm that SecPH is SUMOylated within the full Nf1 isoform 2 protein, given its many additional putative SUMO acceptor lysines, the observation of isoform 2 SUMOylation suggests that Nf1 has a previously unreported potential to be SUMOylated outside exon 30alt31, consistent with SUMOylation at SecPH sites K1731 and K1634.

6. K1731 is highly conserved across evolution and prevention of its SUMOylation is associated with NF1 pathogenicity

To shed more light on whether the proposed sites are biologically relevant, the sequence conservation across the Nf1 SecPH domain was analysed and mapped onto the structure using the ConSurf 2016 web server [6]. ConSurf first collects homologous sequences from the UniProt database using the hidden Markov model (HMM)-based HMMER method with default settings. A random set of 200 homologues with at least 65% identity, which corresponds to a selection of Nf1s from various metazoan species, were selected and conservation scores were calculated with the Bayesian method. The projection of these scores onto SecPH crystal structure (PDB ID: 2E2X) [36] via the ConSurf web server shows that K1731 is highly conserved in most metazoan Nf1 orthologues, thereby highlighting its importance (Figure 8).

Moreover, the physiological importance of SecPH SUMOylation at K1731 is hinted at by the existence of a c.5192A>G (p.Lys1731Arg) missense variant of *NF1* found in the ClinVar and gnomAD database associated with two NF1-patients for which the *in silico* predictions of deleterious effects by SIFT [51] and PolyPhen-2 [52] software were inconclusive. Indeed, our results showing K1731 SUMOylation could provide an unexpected molecular explanation for the physiopathology of this mutant. However, further studies are needed to elucidate if and, if yes, how SUMOylation on K1731 might contribute to the normal function of Nf1.

7. The apparent SUMO consensus motif is dispensable for the SUMOylation of SecPH on K1731

K1731 belongs to a predicted consensus motif, DTKV, evaluated by JASSA [43] to be a strong inverted consensus motif fitting the pattern E/DxKΨ [45]. Moreover, we observed that there is a serine residue in position 1733, which could potentially enhance the SUMOylation of the nearby lysine in a phosphorylation-dependent manner, similarly to what has been described for the phosphorylation-dependent SUMOylation motifs (PDSMs) of MEF2 and HSF1 [53]. In order to investigate if the surrounding sequence might be important for directing SUMOylation to K1731, we constructed three different SecPH mutants that target the mentioned elements and investigated their SUMOylation using our pull down and western blotting approach (Figure 9a). The first mutant, D1729N, which is expected to result in the loss of the negative charge of this acidic residue, did not prevent K1731 SUMOylation, as judged by the presence of the strong band 2 (Figure 9a,b). We then constructed D1729N S1733A, thus combining the first mutant with the ablation of the potential phosphorylation site. Again, K1731 was still SUMOylated (Figure 9a,b). Finally, we made a triple mutant, D1729N V1732G S1733A, with a completely abolished consensus site, but K1731 was still robustly SUMOylated (Figure 9a,b).

The apparent inverted SUMOylation consensus motif around K1731 was responsible for including this residue in the initial pool of potential SUMOylation sites and identifying it as the major site within SecPH. However, our biochemical data presented in this paragraph suggest that, paradoxically, the sequence around K1731 has only accidental resemblance to a consensus motif and does not direct K1731 SUMOylation. This observation is consistent with K1731 being located on the surface of a folded domain, while a part of the apparent consensus sequence (V1732) is buried (supplementary Figure **S2**).

8. K1731 SUMOylation does not involve SecPH SIM motif nor Ubc9 autoSUMOylation on K14

JASSA software predicts, with a high score, a SUMO-interaction motif (SIM) in SecPH between residues 1621 and 1625 (VVDLT), which is consistent with the type b consensus sequence ([P/I/L/V/M/F/Y]-[I/L/V/M]-[D]-[L]-[T]) [54, 55]. Therefore, an alternative explanation for preferential K1731 SUMOylation could involve a SIM-dependent mechanism [56]. We tested this possibility by modelling the interaction of the predicted SIM with SUMO using the published crystal structure of Nf1 SecPH (PDB ID: 2E2X) [36] and, as a comparison, the structure of the complex between SUMO3 and the SIM of CAF1 (PDB ID: 2RPQ) [57] (Figure 10). While a SIM always interacts with SUMO as a beta-strand that complements the central beta-sheet of SUMO in either a parallel (Figure 10b) or an anti-parallel manner, the predicted SIM of SecPH is located in a middle beta-strand of a beta-sheet (Figure 10c), and thus neighbouring beta-strands on either side would clash with SUMO (Figure 10d). This modelling therefore strongly suggests that the predicted SIM appears to be inaccessible and therefore non-functional.

Since Knipscheer et al. [58] have demonstrated that SUMO-modified Ubc9 on K14 has enhanced conjugation activity for certain target proteins containing a SIM motif, we further explored the role of the predicted SIM in K1731 SUMOylation by testing a mutant of Ubc9 that could not be SUMOylated on K14, Ubc9 K14R. For this purpose, we used the UFDS (Ubc9 fusion-directed SUMOylation) technique which, by directly fusing Ubc9 to the target protein facilitates the analysis of protein SUMOylation even in the absence of an E3 ligase [59]. This approach furthermore ensured that our substrate was targeted for SUMOylation by Ubc9 K14R rather than the endogenous wild-type Ubc9 [59].

In the first step, we tested if K1731 was SUMOylated when using the UFDS system and identified the corresponding band on western-blot. Two plasmids encoding Ubc9-SecPH-3xHA and Ubc9-SecPH K1731R-3xHA proteins were constructed and co-transfected with pcDNA3-6His-SUMO2. Pull down was performed as previously described and analyzed along lysates using western blotting with anti HA antibodies (Figure 11a left and right panel

respectively). One of the bands labelled with a red arrow was specifically observed with the Ubc9-SecPH-3xHA protein but not its K1731R mutant, clearly demonstrating the SUMOylation of K1731 in the fusion protein.

In the second step, we introduced the K14R mutation into Ubc9-SecPH-3xHA and conducted a pull down analysis (Figure 11b). This mutation did not affect K1731 SUMOylation, as indicated by the persistence of the previously identified band. These experiments strongly suggest that the predicted SIM motif within SecPH, along with Ubc9 K14 SUMOylation, is dispensable for K1731 SUMOylation.

9. Loop 1723-1729 on SecPH contributes to the recognition of K1731 for SUMOylation

We sought to obtain further insights into the structural elements required for K1731 SUMOylation. Indeed, considering that K1731 is located on the surface of a folded domain, specific interaction interfaces might be created between SecPH and Ubc9 to position K1731 for SUMOylation. Previously, some protein phosphorylation sites on surfaces of folded domains have been shown to be surrounded in space by non-contiguous structural motifs that mimicked linear motifs surrounding canonical phosphorylation sites [60]. We hypothesised that analogous mimicry of a linear SCM by a tertiary structural epitope might explain K1731 SUMOylation.

The best studied example of SUMOylation site recognition by Ubc9 is the one in RanGAP1. In this case, the modified lysine, K524, belongs to a linear SCM that is complementary, in structural terms, to the active site of Ubc9, with L523 and E526 on RanGAP1 docking into specific pockets on Ubc9 [61]. Therefore, we wondered whether K1731 might be surrounded in space by residues that mimic the leucine and glutamate residues in the RanGAP1 SCM. Analysing the spatial environment around K1731 in the crystal structure of SecPH (PDB ID: 2E2X) [36], we identified two surface residues, L1723 and E1747, which, although not adjacent in sequence, are located near K1731 in space, aligned with it in an apparent straight line. A structural model of a possible SecPH/Ubc9 interaction created by superposing L1723, K1731, and E1747 of SecPH onto L523, K524, and E526 of RanGAP1 (from the crystal structure of RanGAP1 in complex with SUMO, Ubc9, and an E3 ligase; PDB ID: 5D2M, [62]), revealed that the three residues are indeed superposing well between the two SUMO substrates (Figure 12b). To test whether the identified residues within SecPH might constitute a functional “3D SCM”, we constructed single and double L1723A E1747A mutants that would abolish this configuration. In all cases, K1731 was still SUMOylated (Figure 12c). This attractive assumption was thus refuted.

We then hypothesized that K1731 SUMOylation may depend on a different interplay of neighbouring hydrophobic and acidic residues than those already investigated by us when testing the potential inverted linear sequence motif or the potential “3D SCM”. Thus, we combined mutations of the acidic and hydrophobic residues of the two sets. When the E1747A mutation was combined with V1732G, K1731 was still SUMO-modified (Figure 12c). By contrast, combination of the L1723A mutation with D1729N resulted in the complete disappearance of the K1731 SUMOylation band (Figure 12c). If the two apparent consensus motifs were individually functioning and compensating each other, both sets of mixed mutants should have had the same detrimental effect on K1731 SUMOylation. Therefore, these results speak against the compensation hypothesis. However, the effect of the combined mutations in both L1723 and D1729, which are structurally but not sequentially adjacent to each other, underlines the importance of the structural surroundings of K1731 for its recognition by Ubc9. Of note, the relative positions of L1723, K1731, and D1729 in space (Figure 12) are such that they do not mimic in a simple manner the positions of key residues in the classical SCM of RanGAP1, suggesting a mode of engaging Ubc9 that goes beyond SCM mimicry.

Since L1723 and D1729 belong to the same short loop (Figure 12d), we thus extended our structural investigations to this loop. It is possible that its conformation is stabilised by a hydrogen bond involving the peptide amide between L1725 and A1726 and the carbonyl of D1729, which would likely be preserved in the D1729N mutant tested by us previously (Figure 9b). We constructed mutants D1729A and A1726P to abolish this hydrogen bond. The introduction of a proline in a geometrically constrained loop environment in mutant A1726P likely had an additional disruptive effect. Indeed, K1731 SUMOylation appeared to be affected in the A1726P but not D1729A mutant (Figure 12e), demonstrating the dispensability of the hydrogen bond between A1726 and D1729 but, nonetheless, confirming the importance of the loop between residues 1723-1729. We went on to test the importance of residues located at the tip of the loop, H1727 and K1728, by mutating them to A. These substitutions had no effect on K1731 SUMOylation, suggesting that it may be the conformation of the loop rather than the specific side chains – particularly not those at the tip of the loop – that contribute to recognition by Ubc9 (Figure 12e).

Overall, our analysis, starting from a presumed structural mimic of an SCM surrounding K1731, finally identified the loop 1723-1729 as an important element for efficient SUMOylation of this particular lysine residue.

10. Positively-charged residues on SecPH β protrusion are essential for SecPH SUMOylation on K1731

In order to find further features that could contribute to K1731 SUMOylation, we first attempted to use AlphaFold2-Multimer to model an interaction between Ubc9 and SecPH [63]. However, the obtained models had low interface predicted template modelling (iPTM) scores (around or below 0.2), suggesting unreliable prediction, and none of them positioned K1731 near Ubc9 active site (unpublished data). Poor predictability with AlphaFold2 is expected for an interaction that is likely very transient and involves two partners that interact with many different proteins/sites.

Therefore, we manually created a new structural model of the SecPH/Ubc9 interaction, by positioning SecPH (from PDB ID: 2E2X, [36]) and Ubc9 (from PDB ID: 5D2M, [62]) in such a way that i) K1731 is placed in the active site of Ubc9, ii) clashes are minimised, iii) the interaction surface is large and juxtaposes opposite electrostatic charges and potential hydrogen bond donors and acceptors on the two partners. The purpose of this manual modelling was not to predict the precise positioning of Ubc9 and SecPH – which would be difficult with such an approximate procedure – but rather to propose new potential determinants of K1731 SUMOylation that could then be tested experimentally. Indeed, based on a model obtained in this way, we predicted two more distant structural elements at the SecPH surface that could contribute to Ubc9 interaction (Figure 13a). One of them was constituted by two short surface loops (composed of residues E1775 to Q1777 and M1792 to E1795) that contain polar residues (E1775, N1776, and Q1794), which could interact with polar residues of an Ubc9 α helix (Y134, Q139) (Figure 13b). However, the triple mutant SecPH E1775A N1776A Q1794A had a WT SUMOylation pattern with a clear K1731 SUMOylation band (Figure 13c), speaking against the importance of these two loops for SUMOylation on this site. The second distal structural element that we identified was the β protrusion, with its two positively charged residues R1748 and K1750, which could interact with a cluster of negatively charged residues on Ubc9 (E98, E99, D100) (Figure 13d). When each of these two residues was individually replaced by alanine producing two single mutants, R1748A and K1750A, the SUMOylation pattern matched the one of the K1731R mutant, with the band corresponding to K1731 SUMOylation disappearing (Figure 13c). This result strongly implies that both residues of the β protrusion are crucial for K1731 SUMOylation, possibly through transiently engaging the mentioned acidic patch on Ubc9. Interestingly, both residues have been previously proposed to be suitably positioned for binding to an unknown ligand [46]. Furthermore, three pathogenic *NF1* missense variants are reported, with missense mutations replacing these residues by amino-acids with very different physicochemical properties (c.5306G>C, p.Arg1748Pro, c.5306G>A, p.Arg1748Gln and c.5311A>G, p.Lys1750Glu). Indeed, K1731 SUMOylation was affected in these three mutants (unpublished results), again hinting at the importance of K1731 SUMOylation for proper Nf1 function.

11. K1731 SUMOylation relies on SecPH loop 1723-1729 and β -protrusion residues when SecPH is fused to Ubc9

The UFDS technique, which involves fusing Ubc9 to the target protein for SUMOylation, enhances the interaction between the two proteins and reduces the dependency on E3 ligases [59]. E3 ligases facilitate the SUMOylation of target lysines by Ubc9 in two main ways: promoting the proximity and favorable orientation of the Ubc9–SUMO complex and the protein substrate [64–66], and stabilizing an active conformation of Ubc9-SUMO, thereby accelerating SUMO transfer [62, 64, 65, 67, 68].

Our previous experiments demonstrated that K1731 is SUMOylated when Ubc9 is fused to SecPH. To further explore the role of SecPH structural elements in K1731 SUMOylation by the UFDS technique, we introduced mutations affecting the 1723-1729 loop (A1726P), positively charged residues of the β -protrusion (R1748A and K1750A), and a control mutation (K1634R) into plasmids encoding Ubc9-SecPH-3xHA. These plasmids were co-transfected with a pcDNA3-6His-SUMO2 plasmid into HEK293 cells. Cell lysates and pull-down assays were then analyzed using anti-HA antibodies.

As shown in Figure 14, the band corresponding to K1731 SUMOylation, present in Ubc9-SecPH-3xHA and Ubc9-SecPH K1634R-3xHA, disappears in the Ubc9-SecPH-3xHA K1731R, Ubc9-SecPH-3xHA A1726P, Ubc9-SecPH-3xHA R1748A, and Ubc9-SecPH-3xHA K1750A mutants (in pull-downs and lysates). These results indicate that, under conditions favouring SUMOylation of K1731 independently of SUMO E3 ligases, the 1723-1729 loop and the positively charged residues of the β -protrusion are crucial for this process.

DISCUSSION

In this work, we showed that the exogenously expressed SecPH domain of Nf1 is SUMOylated with the SUMO2 paralogue in human cells and identified two specific lysines, K1634 and K1731, as major SUMOylation sites. Based on the SUMOylation pattern obtained by western blotting, at least one more site within SecPH is significantly SUMOylated but has not been mapped. Furthermore, in our context of N-terminal tagging of the SecPH domain, we could unexpectedly demonstrate that two lysines of the 3X Flag tag were also SUMOylated to a significant extent. The last observation might by itself be interesting to the field, perhaps warning users of the properties of this tag in future SUMOylation studies.

Our study is also the first one which demonstrates that Nf1 isoform 2, the most studied Nf1 isoform, is SUMOylated with SUMO2. Of note, mass spectrometry-based global studies of SUMOylation had previously identified Nf1 isoform 1 as a potential SUMO target protein, but

the two SUMOylated lysines identified were localized in the 30alt31exon, which is spliced out from isoform 2 [42]. Thus, our demonstration of isoform 2 SUMOylation suggests the existence of Nf1 SUMOylation sites outside 30alt31exon, which have remained undetected with mass spectrometry approaches. Our results suggest that K1634 and K1731 of SecPH are good candidates for these extra sites of SUMOylation. Although we did not show their SUMOylation on the full length Nf1 isoform 2, we believe there is no reason why they should be SUMOylated only in the isolated SecPH domain, considering that both of these sites are at least temporarily surface-exposed in the full-length Nf1 based on cryo-EM structures.

Interestingly, K1731 remains surface-exposed in the Nf1 dimer structure regardless of its conformation (closed or open). When Ubc9 and SUMO2 are overexpressed, K1731 SUMOylation generates a very intense SUMO signal corresponding to 16% of total SecPH. Under endogenous expression levels of these proteins this signal drops to 8%, which still corresponds to the standard protein SUMOylation level under native conditions. An important role of K1731 and its SUMOylation in Nf1 function is further suggested by the high conservation of this residue in most metazoans and the reported pathogenicity of different *NF1* missense variants: c.5192A>G (p.Lys1731Arg), c.5306G>C (p.Arg1748Pro), c.5306G>A (p.Arg1748Gln) and c.5311A>G (p.Lys1750Glu), all of which are devoid of K1731 SUMOylation. SUMOylation of the SecPH domain on K1731 might exert a biological function by regulating the Ras-GAP activity of the adjacent GRD domain. We assessed this hypothesis, but we failed to confirm it in a pilot experiment (supplementary Figure S3). More precisely, we transfected HeLa cells knocked out for *NF1* [69] with plasmids encoding WT, K1731R, or R1276Q Nf1 isoform 2 (where R1276Q is a control pathogenic variant of *NF1* unrelated to K1731 SUMOylation [70]), and evaluated Ras-GAP activity by measuring the pERK/ERK ratio and Ras-GTP level. We could not detect any differences upon the K1731 mutation. Considering this result, we hypothesise that K1731 SUMOylation might instead affect Ras-independent functions of Nf1 by regulating the interactions of SecPH with actors of different signalling pathways (5-HT₆ receptor, LIMK2, VCP) or phospholipids. In this regard, it is worth pointing out that K1731 is located close to the base of the β protrusion of the PH domain, a crucial element of the PH subdomain, which plays a role in the interconnexion between the Sec and PH subdomains and SecPH structural integrity [71] and has been proposed to regulate the gating of a lipid-binding pocket present within Sec [30].

Beyond its biological importance, SecPH SUMOylation on K1731 can also be seen as a model example of an atypical SUMOylation process that does not conform to the best described mechanisms governing SUMOylation site selection. Firstly, whereas SUMOylation sites, like most other PTMs, are enriched within intrinsically disordered and other flexible regions [72],

K1731 of Nf1 is located at an end of a beta strand within a tightly folded SecPH domain. Secondly, while around 50% of all known SUMOylation sites are surrounded by one of the specific linear consensus motifs [42], the similarity of the sequence around K1731 to the so-called inverted SUMOylation consensus motif turns out to be merely accidental and functionally irrelevant. Thirdly, whereas the SUMOylation of some non-consensus SUMOylation sites has been shown to rely on a SIM motif present in the substrate – which could help to recruit SUMOylated or SUMO-loaded Ubc9 – the predicted SIM motif present in Nf1 SecPH does not seem to be available for SUMO binding. Fourthly, instead of relying on a linear consensus motif, a SIM, or Ubc9 auto-SUMOylation, K1731 SUMOylation appears to depend on the nearby tertiary structural elements within the substrate: the conformation of the loop in the vicinity of the modified lysine and two positively charged residues of SecPH β protrusion. In this regard, SecPH SUMOylation on K1731 is similar to E2-25K and Ubc9 SUMOylation on K14, which relies in both cases on a secondary structural element (an α helix) rather than on the flanking primary structure [73], [58]. In the case of Nf1, the structural elements required for specific SUMOylation are located on different secondary structural elements that together form a tertiary arrangement. It cannot be excluded that other yet unidentified SecPH structural elements, beyond the two revealed in our work, may also contribute to this process. If, as some of our observations suggest, K1731 SUMOylation is involved in Nf1 function, identification of these potential further structural elements might open the door to the categorization of various *NF1* pathogenic variants.

In canonical SUMO substrates such as RanGAP1, SUMOylation site selection depends on a direct physical interaction between the substrate SCM and Ubc9-SUMO, allowing the proper positioning of the catalytic site of Ubc9 (involving C93 and D127) close to the target lysine for the duration of the SUMOylation reaction. In the case of the SecPH domain of Nf1, the structural determinants of K1731 SUMOylation might also function by directly recruiting Ubc9. However, unlike in some cases known from the protein phosphorylation field where 3D motifs structurally emulate known linear consensus motifs [60], the 3D motif required for K1731 SUMOylation does not appear to be a straightforward structural mimic of an SCM. Instead, the identified elements around K1731 seem to form a more complex composite “structural motif” for recruiting Ubc9 and positioning it on K1731.

We believe that E3 ligases might play a secondary role in the dependence of K1731 SUMOylation on the SecPH structural motif for two reasons: First, using the UFDS system, which enables target SUMOylation even in the absence of SUMO ligases, we demonstrated that K1731 SUMOylation also relies on the SecPH structural motif (comprising loop 1723-1729 and β -protrusion residues). Second, SUMO E3 ligases are generally known to accelerate

SUMOylation at sites recognized by Ubc9 alone. **Indeed**, as far as we know, there are no known examples of E3 ligases imposing a specificity different from that imposed by Ubc9 except for yeast PCNA SUMOylation at K164, which requires the SUMO E3 ligase Siz1 **[65, 74]**.

However, since our experiments were conducted in cells containing numerous proteins and components, we cannot exclude the possibility that additional factors might influence K1731 SUMOylation.

In summary, although structural details of K1731 recognition by the SUMOylation machinery remain to be fully determined, it is clear that SecPH K1731 SUMOylation is an SCM- and SIM-independent process reliant on a tertiary structural environment around the modified site. This new noncanonical example expands our knowledge of the ways by which SUMOylation sites are selected (Figure 15).

MATERIALS AND METHODS

Recombinant DNA and site-directed mutagenesis

pcDNA3 Ubch9/SV5 tagged (thereafter named pcDNA3 Ubc9) and pcDNA3-His6-SUMO-2 (thereafter named pcDNA3 6His-SUMO2) were kindly provided by R. Hay and have been previously described **[75, 76]**. pcDNA3 6His-SUMO2 amplification with Flag SUMO2 primers (all primers are listed in Table S1) generated an amplicon which was eventually cleaved with HindIII and BamHI and inserted into the same sites of p3X Flag-Myc-CMV™-24 expression Vector (E9283, Sigma-Aldrich) to create p3X Flag SUMO2 encoding a Flag-tagged SUMO2 fusion protein. p3X Flag SecPH was created by amplification of the *NF1* DNA sequence encoding SecPH from a Marathon-Ready™ bank of human lung cDNA with appropriate oligonucleotides containing restriction sites for cloning into the p3X Flag-Myc-CMV™-24 Expression Vector (E9283, Sigma-Aldrich) **[34]**. This plasmid was then used to create most of p3X Flag SecPH plasmids derivatives with one, two, or three amino acid substitutions in SecPH with Q5 PCR methodology (Q5@Site-Directed Mutagenesis Kit, New England Biolabs, E0554) or PCR site-directed mutagenesis strategy (List of oligonucleotides in Table S1 and list of p3X Flag SecPH derivatives in Table S2). To construct p3X Flag SecPH K1634R K1731R, K1634R substitution was introduced by PCR into p3X Flag SecPH K1731R with oligonucleotides K1634R For and Rev. To construct p3X Flag SecPH L1723A D1729N, L1723A was introduced into p3X Flag SecPH D1729N with oligonucleotides L1723A. To construct p3X Flag SecPH L1723A E1747A, E1747A was introduced into p3X Flag SecPH

L1723A with oligonucleotides E1747A. To construct p3X Flag SecPH V1732G E1747A, p3X Flag SecPH E1747A was used as template to introduce the amino acid substitution with oligonucleotides V1732G. To construct p3X Flag SecPH 5K-R, K1731R was first introduced into p3X Flag SecPH K1735R with appropriate oligonucleotides; then the three other KR substitutions (K1593R, K1634R, K1717R) were introduced one after the other by successive rounds of mutagenesis. Plasmids p3X Flag SecPH A1726P, p3X Flag SecPH R1748P, p3X Flag SecPH R1748Q, p3X Flag SecPH K1750E, p3X Flag SecPH H1727A K1728A, p3X Flag SecPH E1775A N1776A Q1794A, p3X Flag SecPH K0 and p3X 2KR Flag SecPH derivatives (p3X 2KR Flag SecPH, p3X 2KR Flag SecPH K0, p3X 2KR Flag SecPH K0 K1731, p3X 2KR Flag SecPH K0 K1634 K1731 and p3X 2KR Flag SecPH K0 K1735) were synthesised at GenScript (Table S2). Plasmids pUbc9-SecPH-3xHA, pUbc9-SecPH-K1731R-3xHA, pUbc9-K14R-SecPH-3xHA, pUbc9-K14R-SecPH-K1731R-3xHA were synthesised at GeneCust [plasmid backbone is pcDNA3.1(+)]. Plasmids pUbc9-SecPH K1634R-3xHA, pUbc9-SecPH A1726P-3xHA, pUbc9-SecPH R1748A-3xHA and pUbc9-SecPH K1750A-3xHA were constructed by transferring the EcoRI/BspEI fragment of the p3X Flag SecPH K1634R, p3X Flag SecPH A1726P, p3X Flag SecPH R1748A and p3X Flag SecPH K1750A plasmids into the pUbc9-SecPH-3xHA plasmid hydrolysed with the same enzymes.

p3X Flag GRD was constructed by amplification of the *NF1* DNA sequence (without 30alt31 exon) encoding GRD from the Marathon-Ready™ bank described above with the Taq Advantage polymerase (BD Biosciences) using the oligonucleotides listed in Table S1. The amplicon generated was inserted into the NotI and Sall sites of p3X Flag-Myc-CMV™-24 Expression Vector (E9283, Sigma-Aldrich).

To create p3X Flag Nf1 constructs, we decided to use a codon-optimized version (opt) of *NF1* cDNA developed by Sherekar et al. [24] (Addgene plasmid # 70423 encoding human Nf1 isoform 2) but also to take advantage of the p3X Flag SecPH derivative plasmids already constructed which possess DNA sequence encoding SecPH of native *NF1* (Nat). In a first approach, four DNA amplicons corresponding to the nucleotidic sequences of Nter opt - GRD opt, SecPH Nat, Cter opt and the entire nucleotidic sequence of the p3X Flag plasmid were generated with adequate overlapping oligonucleotides and were assembled with the NEBuilder® HiFi DNA Assembly kit to create a p3X Flag Nf1 SecPH Nat construct (*alias* p3X Flag Nf1). The same strategy was used to construct a p3X Flag Nf1 GRD Nat plasmid using amplicons corresponding to the nucleotidic sequences of Nter opt - GRD Nat - SecPH opt - Cter opt and the p3X Flag GRD plasmid (created with DNA sequence encoding GRD of native *NF1*).

In a second approach, p3X Flag Nf1 and p3X Flag SecPH K1731R were used to generate, with adequate overlapping nucleotides, two overlapping amplicons that were assembled with the NEBuilder® HiFi DNA Assembly kit to give p3X Flag Nf1 K1731R. p3X Flag Nf1 R1276Q was generated from p3X Flag Nf1 GRD Nat construct using Q5 PCR methodology and two oligonucleotides listed in Table S1.

p3X 2KR Flag Nf1 isoform 2 was constructed by the assembly of two amplicons using the NEBuilder® HiFi DNA Assembly kit. One amplicon corresponded to the *NF1* cDNA nucleotide sequence of Nter - GRD opt - SecPH Nat - Cter opt. The other corresponded to the entire nucleotide sequence of the p3X Flag plasmid in which the codon sequence of the two SUMOylated lysines of the last Flag tag were mutated into arginine using oligonucleotide p3X 2KR Flag-rev (Table S1).

At all steps of their construction, the full sequences of our constructs were verified.

Cell growth and Transfection

HEK293T cells from the American Type Culture Collection (ATCC ref: CRL1573, thereafter named HEK293) were cultured under 5% CO₂ at 37°C in low-glucose Dulbecco's modified Eagle's medium (DMEM) (Sigma-Aldrich, #D6046) supplemented with 10% foetal bovine serum (FBS-Gibco). Transient transfection using the calcium-phosphate method was carried out with 15 or 20 µg of plasmid in 100-mm dish or 4 µg in 20-mm 6 well plates. Further experiments were conducted 48 h after transfection.

HeLa KO (NF1^{-/-}) cells (BioCat, knockout was achieved by using CRISPR/Cas9, Homozygous: 1 bp deletion in exon 1) and HeLa NF1 WT were cultured under 5% CO₂ at 37°C in high-glucose DMEM (Sigma-Aldrich, #D5796) supplemented with 10% FBS. Transient transfection was performed using PEI (0.35 µM/µg DNA) (Invitrogen™ PEI-Transferrinfection Kit # BMS1003) and 5 µg of plasmid DNA in 6 well plates. After 24 h of growth, cells were serum-starved (0.1% FBS) in high-glucose DMEM for 18 h to decrease to a basal level the activation of the Ras pathway. After 18 h of growth-factor deprivation, the cells were restimulated for 15 min with fresh medium supplemented with 10% FBS immediately before the preparation of cell lysates (adapted from Cichowski [77, 78]).

Cells were regularly checked for mycoplasma contamination using a luminescence-based kit (Lonza, MycoAlert™ Mycoplasma Detection Kit).

Immunoprecipitation for the detection of endogenous Nf1 SUMOylation

HEK293 cells were transiently transfected with p3X Flag SUMO2 and pcDNA3 Ubc9 (or the corresponding empty vectors) using the calcium phosphate method. After 48 h of growth cells were lysed for 10 min on ice in 500 μ l lysis buffer containing 50 mM Tris-HCl (pH 7.5), 100 mM NaCl, 5 mM EDTA, 0.1% Triton-X100, 1 mM PMSF, 1 μ M E-64, 20 mM NEM, and EDTA-free protease inhibitor cocktail (Sigma-Aldrich, #S8830). After centrifugation at 10,000 x g for 10 min at 4 °C, the clarified lysates were incubated with equilibrated anti Flag M2 affinity gel (ANTI-FLAG M2 Affinity Gel Sigma-Aldrich, #A2220) on a rotating wheel for 3 h at 4°C. The beads were subsequently washed five times with ice-cold lysis buffer freshly supplemented with 20 mM NEM and subsequently incubated with Flag peptide (2 mg/mL, Sigma-Aldrich, #F3290) for 30 min at 4°C. The eluted SUMOylated or SUMO-interacting proteins were then mixed with 10 μ l undiluted anti Nf1 antibody (Novus Biologicals, #NB100-418) and incubated for 3 h at 4°C on a rotating wheel. Then, the Nf1-bound proteins were captured by an additional incubation for 1 h at 4°C with Protein G beads (Pierce, #2851), followed by extensive washing in ice-cold lysis buffer freshly supplemented with 20 mM NEM. The captured proteins were eluted in Laemmli buffer by incubation at 95°C for 5 min, followed by a centrifugation step at 10,000 x g for 5 min.

Immunoprecipitation for the detection of exogenously produced Nf1 isoform 2 SUMOylation

HEK293 cells were co-transfected with p3X 2KR Flag Nf1 (unSUMOylable Flag tag in p3X 2KR Flag SecPH), pcDNA3 6His-SUMO2 and pcDNA3 Ubc9 plasmids (these two later plasmids were replaced by pcDNA3 empty plasmid for the negative control) and total lysates were prepared as described in the previous section. The clarified lysates were incubated with equilibrated anti Flag M2 affinity gel (ANTI-FLAG M2 Affinity Gel Sigma-Aldrich, #A2220) on a rotating wheel for 3 h at 4°C. The beads were subsequently washed five times with ice-cold lysis buffer freshly supplemented with 20 mM NEM and subsequently incubated with Flag peptide (2 mg/mL, Sigma-Aldrich, #F3290) for 30 min at 4°C. The eluted proteins were recovered after a centrifugation step at 10,000 x g for 2 min, and stored at -80°C in 1X Laemmli buffer.

Immunoprecipitation and affinity pull down of Flag SecPH derivatives

HEK293 cells were co-transfected with p3X Flag SecPH derivatives, along with pcDNA3 6His-SUMO2 and pcDNA3 Ubc9 or the corresponding empty vectors. As a negative control, HEK293 cells were co-transfected with the MDH1-Flag plasmid, along with pcDNA3 6His-SUMO2 and pcDNA3 Ubc9.

For the Immunoprecipitation analyses, the Flag tagged SecPH and MDH1 proteins were captured with anti Flag M2 affinity gel and further eluted with the Flag peptide as described above.

For the enrichment of SUMOylated proteins by affinity pull down, HEK293 cells co-transfected for 48 h with p3X Flag SecPH derivatives along, or not, with pcDNA3 6His-SUMO2 and pcDNA3 Ubc9. Cells were recovered by low-speed centrifugation at 4°C, washed two times in 3 ml ice-cold PBS buffer, and then resuspended in 1 ml PBS. One tenth of the cell resuspension was used to prepare whole-cell extracts with 100 µL ice-cold lysis buffer described above. After 10 min incubation on ice, lysates were cleared by centrifugation at 10,000 x g for 10 min at 4°C. Then, 60 µL of the lysate was mixed with 20 µL of 4X Laemmli buffer and samples stored at -80°C. The second part of the cell resuspension was centrifuged at low speed and cell extracts prepared under denaturing conditions by incubation in 500 µL 6 M guanidium hydrochloride buffer pH 7.6 (6 M guanidium hydrochloride prepared in 100 mM phosphate buffer) at 95°C for 5 min followed by extensive vortexing to remove chromatin-bound proteins. Insoluble material was removed by centrifugation at 16,000 x g for 30 min at room temperature (RT). The supernatant was then mixed with 50 µL cobalt beads (TALON® Metal Affinity beads, Clontech, #635502) pre-equilibrated and the samples incubated at RT for 3 h with rotation. Pulled down His-tagged SUMO-conjugated proteins were washed twice in guanidium hydrochloride buffer, twice in buffer containing guanidium hydrochloride buffer and imidazole buffer (20 mM Imidazole, 25 mM Tris-HCl, pH 6.8) at a 1:3 ratio, and, finally, three times in imidazole buffer. The SUMO-conjugated proteins were eluted by incubation at 95°C for 5 min in Laemmli buffer, followed by centrifugation for 10 min at 16,000 x g. Samples were subsequently analysed on 8% SDS-PAGE alongside whole-cell extracts.

Immunoblotting

Proteins were transferred onto Immobilon-P PVDF membranes (Merck, #IPVH00010) using a wet-transfer system (Bio-Rad) at 100 V and 4°C for 90 min (Flag SecPH and Ubc9-SecPH-3xHA protein derivatives, and 6His-SUMO-conjugated proteins) or 180 min (Nf1, 2KR Flag Nf1 protein, and Flag-SUMO2 conjugated proteins). Then, membranes were blocked in 5% non-fat milk or 5% BSA (according to the primary antibody) prepared in Tris-buffered saline with 0.1% Tween-20 (TBS-T). After 1h at RT, the buffer was discarded and the primary antibody diluted into fresh TBS-T + 5% non-fat milk or 5% BSA was added (listed in Table S3). In the case of GAPDH visualisation, anti GAPDH primary antibody coupled to horseradish peroxidase was used. After incubation overnight at 4°C, the membrane was washed three times with TBS-T and GAPDH was visualized by chemiluminescence (SuperSignal™ West Dura Extended Duration Substrate, Pierce™) using the PXi imaging system (Syngene). In the

other cases, the primary antibody was discarded after the overnight incubation at 4°C, the membranes were washed three times in TBS-T for 20 min; then, horseradish peroxidase-conjugated secondary antibody were added to the membrane (1:50 000 in TBS-T+ 5% non-fat milk) which were further incubated at RT for 2 h (Invitrogen, rabbit #65-6120, mouse #61-6520, rat #62-9520). The membranes were then washed as above and the proteins were visualized using the PXi imaging system (Syngene) or the Invitrogen™ iBright™ FL1500 Imaging system.

Quantification of western blot bands was performed by densitometry using GeneTools from Syngene.

Ras-GAP activity measurement

pERK and ERK levels were measured in HeLa cells transfected and grown as described above. Cells were lysed in 100 µL ice-cold buffer for 5 min in lysis buffer (from RAS-G-LYSA kit, purchased at Cytoskeleton, Inc.) containing protease inhibitors (50 mM NaF and 1 mM PMSF). After centrifugation at 10,000 x g for 10 min at 4°C, the total protein concentration was determined by colorimetry using the Uptima BC Assay Protein Quantitation Kit (# UP40840B) and serial dilutions of bovine serum albumin as a calibration standard. After determination of protein concentration, the lysates were separated in two batches and stored at -80°C. An equal amount of protein was separated on a SDS-Polyacrylamide Gel, immunoblotted, and analysed using anti ERK, anti pERK, anti Nf1, and anti GAPDH antibodies. The pERK and ERK ratios were calculated after quantification of western blots as above.

On the second batch of cells lysates, the RAS-G-LYSA kit (Cytoskeleton, Inc.) was used according to the manufacturer's instructions to determine Ras-GTP level. Of note, in this experiment Ubc9 and SUMO2 were not overexpressed in order to prevent the SUMO-dependant Ras proteins activation [79].

Statistical analyses

Statistical analyses were performed using StatEL software (www.adscience.eu). The non-parametric Mann-Whitney U test was used (criterion of significance: 5%) and the results were considered statistically significant at P-value <0.05.

Structural models and figures

Structural models and figures were produced in PyMOL Molecular Graphics System, Version 2.0 (Schrödinger, LLC). For all analyses, we used the crystal structure of Nf1 SecPH bound to phosphatidylethanolamine (PDB ID: 2E2X) [36]. AlphaFold2-Multimer modelling was performed using the on-line ColabFold platform (v1.5.5: AlphaFold2 using MMseqs2) with default settings by inputting the sequence of Nf1 SecPH (residues 1566-1837 according to isoform 1 numbering) and the full-length sequence of human Ubc9 separated by a colon.[63, 80, 81]

Sequence conservation analysis

We performed sequence conservation analysis with the HMMER method using the ConSurf 2016 web server [6] with default settings except for focusing only on homologues with at least 65% identity in the SecPH domain, which corresponds to Nf1 from most metazoan species except for distant outliers. For the homology search with the HMMER algorithm, 200 different sequences with a minimal sequence identity of 65% were automatically selected in the UNIREF-90 protein database by the server in order to provide a ConSurf conservation score for each residue. In this group of 200 homologous sequences, K1731 appears strictly conserved. For the representation of sequence conservation, ConSurf conservation scores were projected onto the SecPH crystal structure (PDB ID: 2E2X), via the ConSurf web server.

Figure preparation

Graphs were prepared using Microsoft Excel 2016 and the statistical significance evaluated with the non-parametric Mann-Whitney U test (two-tailed) with StatEL, a software operating on the spreadsheet of Microsoft Excel (www.adscience.eu). Editing and assembly of Figures was performed with Power Point.

Data availability

All data generated or analysed during this study are included in this published article. The materials used in this study are available from the corresponding author, HB, upon reasonable request.

Acknowledgements

We thank Dr. Ronald Thomas Hay (Centre for Gene Regulation and Expression, College of Life Sciences, University of Dundee, Dundee, Scotland DD1 5EH, UK) for providing the plasmids. We are grateful to laboratory alumni for their technical support (Elodie Villalonga, Bojan Žunar, Veronika Buršić, Manon Julien, Navid Barakzoy, Aurélie Gombault, Marie-

Ludivine de Tauzia, and Khadija Aziz). We thank Alessia Zamborlini, Stéphane Martin, Eric Pasmant for stimulating discussions.

Author Contributions

M.B, C.M, A.S, F.G, M.D, I.S, S.G. conceived and performed most of the experiments. M.J.S prepared structural models and figures and proposed structure-guided mechanistic hypotheses. A.S, F.G, M.B, S.G, T.N and M.D contributed to the final figures and provided supporting data. B.V, C.M, and M.J.S reviewed and edited the manuscript. H.B. conceived the project and H.B and B.V supervised the study. H.B wrote the original draft with the assistance of B.V, M.J.S, and C.M. All authors contributed to revising the manuscript.

Funding and additional information

The CNRS, the French “Association Neurofibromatose et Recklinghausen” (research grant to H.B.), the Ligue Contre le Cancer (research grant to H.B. and B.V.), and the French Agence Nationale de la Recherche (CliNeNF1, grant number: ANR-19-CE18-0016-02) provided financial support. M.B. was supported by a fellowship from the French Ministère de l’Enseignement Supérieur, de la Recherche et de l’Innovation.

Conflict of interest

The authors declare that they have no conflicts of interest with the contents of this article.

REFERENCES

- [1] Mahajan R, Delphin C, Guan T, Gerace L, Melchior F. A Small Ubiquitin-Related Polypeptide Involved in Targeting RanGAP1 to Nuclear Pore Complex Protein RanBP2. *Cell*. 1997;88:97-107.
- [2] Matunis MJ, Coutavas E, Blobel G. A novel ubiquitin-like modification modulates the partitioning of the Ran-GTPase-activating protein RanGAP1 between the cytosol and the nuclear pore complex. *The Journal of cell biology*. 1996;135:1457-70.
- [3] Johnson ES. Protein modification by SUMO. *Annu Rev Biochem*. 2004;73:355-82.
- [4] Hickey CM, Wilson NR, Hochstrasser M. Function and regulation of SUMO proteases. *Nature reviews Molecular cell biology*. 2012;13:755-66.
- [5] Schimmel J, Eifler K, Sigurðsson JO, Cuijpers SA, Hendriks IA, Verlaan-de Vries M, et al. Uncovering SUMOylation dynamics during cell-cycle progression reveals FoxM1 as a key mitotic SUMO target protein. *Molecular cell*. 2014;53:1053-66.

- [6] Ashkenazy H, Abadi S, Martz E, Chay O, Mayrose I, Pupko T, et al. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic acids research*. 2016;44:W344-50.
- [7] Celen AB, Sahin U. Sumoylation on its 25th anniversary: mechanisms, pathology, and emerging concepts. *The FEBS journal*. 2020;287:3110-40.
- [8] Mandel N, Agarwal N. Role of SUMOylation in Neurodegenerative Diseases. *Cells*. 2022;11.
- [9] Gareau JR, Lima CD. The SUMO pathway: emerging mechanisms that shape specificity, conjugation and recognition. *Nature reviews Molecular cell biology*. 2010;11:861-71.
- [10] Patra U, Müller S. A Tale of Usurpation and Subversion: SUMO-Dependent Integrity of Promyelocytic Leukemia Nuclear Bodies at the Crossroad of Infection and Immunity. *Frontiers in cell and developmental biology*. 2021;9:696234.
- [11] Godin F, Villette S, Vallée B, Doudeau M, Morisset-Lopez S, Ardourel M, et al. A fraction of neurofibromin interacts with PML bodies in the nucleus of the CCF astrocytoma cell line. *Biochemical and biophysical research communications*. 2012;418:689-94.
- [12] Koczkowska M, Chen Y, Callens T, Gomes A, Sharp A, Johnson S, et al. Genotype-Phenotype Correlation in NF1: Evidence for a More Severe Phenotype Associated with Missense Mutations Affecting NF1 Codons 844-848. *American journal of human genetics*. 2018;102:69-87.
- [13] Koczkowska M, Callens T, Gomes A, Sharp A, Chen Y, Hicks AD, et al. Expanding the clinical phenotype of individuals with a 3-bp in-frame deletion of the NF1 gene (c.2970_2972del): an update of genotype-phenotype correlation. *Genet Med*. 2019;21:867-76.
- [14] Karaconji T, Whist E, Jamieson RV, Flaherty MP, Grigg JRB. Neurofibromatosis Type 1: Review and Update on Emerging Therapies. *Asia-Pacific journal of ophthalmology (Philadelphia, Pa)*. 2019;8:62-72.
- [15] Wilson BN, John AM, Handler MZ, Schwartz RA. Neurofibromatosis type 1: New developments in genetics and treatment. *J Am Acad Dermatol*. 2021;84:1667-76.
- [16] Eisfeld AK, Kohlschmidt J, Mrózek K, Mims A, Walker CJ, Blachly JS, et al. NF1 mutations are recurrent in adult acute myeloid leukemia and confer poor outcome. *Leukemia*. 2018;32:2536-45.
- [17] Krauthammer M, Kong Y, Bacchiocchi A, Evans P, Pornputtapong N, Wu C, et al. Exome sequencing identifies recurrent mutations in NF1 and RASopathy genes in sun-exposed melanomas. *Nature genetics*. 2015;47:996-1002.
- [18] Pearson A, Proszek P, Pascual J, Fribbens C, Shamsheer MK, Kingston B, et al. Inactivating NF1 Mutations Are Enriched in Advanced Breast Cancer and Contribute to Endocrine Therapy Resistance. *Clin Cancer Res*. 2020;26:608-22.
- [19] Redig AJ, Capelletti M, Dahlberg SE, Sholl LM, Mach S, Fontes C, et al. Clinical and Molecular Characteristics of NF1-Mutant Lung Cancer. *Clin Cancer Res*. 2016;22:3148-56.
- [20] Verhaak RG, Hoadley KA, Purdom E, Wang V, Qi Y, Wilkerson MD, et al. Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer Cell*. 2010;17:98-110.

- [21] Stenson PD, Mort M, Ball EV, Shaw K, Phillips A, Cooper DN. The Human Gene Mutation Database: building a comprehensive mutation repository for clinical and molecular genetics, diagnostic testing and personalized genomic medicine. *Human genetics*. 2014;133:1-9.
- [22] Young LC, Goldstein de Salazar R, Han SW, Huang ZYS, Merk A, Drew M, et al. Destabilizing NF1 variants act in a dominant negative manner through neurofibromin dimerization. *Proceedings of the National Academy of Sciences of the United States of America*. 2023;120:e2208960120.
- [23] Lois S, Báez-Flores J, Isidoro-García M, Lacal J, Triviño JC. Identification of Germinal Neurofibromin Hotspots. *Biomedicines*. 2022;10.
- [24] Sherekar M, Han S-W, Ghirlando R, Messing S, Drew M, Rabara D, et al. Biochemical and structural analyses reveal that the tumor suppressor neurofibromin (NF1) forms a high-affinity dimer. *Journal of Biological Chemistry*. 2020;295:1105-19.
- [25] Bergoug M, Doudeau M, Godin F, Mosrin C, Vallée B, Bénédicti H. Neurofibromin Structure, Functions and Regulation. *Cells*. 2020;9.
- [26] Chaker-Margot M, Werten S, Dunzendorfer-Matt T, Lechner S, Ruepp A, Scheffzek K, et al. Structural basis of activation of the tumor suppressor protein neurofibromin. *Molecular cell*. 2022;82:1288-96.e5.
- [27] Naschberger A, Baradaran R, Rupp B, Carroni M. The structure of neurofibromin isoform 2 reveals different functional states isoform 1 nouvelle nomenclature 2838 aa UniProt: P21359-2 et Genbank: NM_000267.3. *Nature*. 2021;599:315-9.
- [28] Lupton CJ, Bayly-Jones C, D'Andrea L, Huang C, Schittenhelm RB, Venugopal H, et al. The cryo-EM structure of the human neurofibromin dimer reveals the molecular basis for neurofibromatosis type 1 2818 AA forme 2. *Nature structural & molecular biology*. 2021;28:982-8.
- [29] Scheffzek K, Ahmadian MR, Wiesmüller L, Kabsch W, Stege P, Schmitz F, et al. Structural analysis of the GAP-related domain from neurofibromin and its implications. *The EMBO journal*. 1998;17:4313-27.
- [30] D'angelo I, Welti S, Bonneau F, Scheffzek K. A novel bipartite phospholipid-binding module in the neurofibromatosis type 1 protein. *EMBO reports*. 2006;7:174-9.
- [31] Hannan F, Ho I, Tong JJ, Zhu Y, Nurnberg P, Zhong Y. Effect of neurofibromatosis type I mutations on a novel pathway for adenylyl cyclase activation requiring neurofibromin and Ras. *Human molecular genetics*. 2006;15:1087-98.
- [32] Anastasaki C, Orozco, Gutmann DH. RAS and beyond: the many faces of the neurofibromatosis. *Human molecular genetics*. 2022;23:6712-21.
- [33] Anastasaki C, Gutmann DH. Neuronal NF1/RAS regulation of cyclic AMP requires atypical PKC activation. *Human molecular genetics*. 2014;23:6712-21.
- [34] Vallée B, Doudeau M, Godin F, Gombault A, Tchalikian A, de Tauzia ML, et al. Nf1 RasGAP inhibition of LIMK2 mediates a new cross-talk between Ras and Rho pathways. *PloS one*. 2012;7:e47283.
- [35] Starinsky-Elbaz S, Faigenbloom L, Friedman E, Stein R, Kloog Y. The pre-GAP-related domain of neurofibromin regulates cell migration through the LIM kinase/cofilin pathway. *Mol Cell Neurosci*. 2009;42:278-87.

- [36] Welti S, Fraterman S, D'Angelo I, Wilm M, Scheffzek K. The sec14 homology module of neurofibromin binds cellular glycerophospholipids: mass spectrometry and structure of a lipid complex. *J Mol Biol.* 2007;366:551-62.
- [37] Deraredj Nadim W, Chaumont-Dubel S, Madouri F, Cobret L, De Tautzia ML, Zajdel P, et al. Physical interaction between neurofibromin and serotonin 5-HT₆ receptor promotes receptor constitutive activity. *Proceedings of the National Academy of Sciences of the United States of America.* 2016;113:12310-5.
- [38] Wang H-F, Shih Y-T, Chen C-Y, Chao H-W, Lee M-J, Hsueh Y-P. Valosin-containing protein and neurofibromin interact to regulate dendritic spine density. *The Journal of clinical investigation.* 2011;121:4820-37.
- [39] Andersen LB, Ballester R, Marchuk DA, Chang E, Gutmann DH, Saulino AM, et al. A conserved alternative splice in the von Recklinghausen neurofibromatosis (NF1) gene produces two neurofibromin isoforms, both of which have GTPase-activating protein activity. *Molecular and cellular biology.* 1993;13:487-95.
- [40] Hinman MN, Sharma A, Luo G, Lou H. Neurofibromatosis type 1 alternative splicing is a key regulator of Ras signaling in neurons. *Molecular and cellular biology.* 2014;34:2188-97.
- [41] Hendriks Ivo A, Treffers Louise W, Verlaan-de Vries M, Olsen Jesper V, Vertegaal Alfred CO. SUMO-2 Orchestrates Chromatin Modifiers in Response to DNA Damage. *Cell Reports.* 2015;10:1778-91.
- [42] Hendriks IA, Lyon D, Young C, Jensen LJ, Vertegaal AC, Nielsen ML. Site-specific mapping of the human SUMO proteome reveals co-modification with phosphorylation. *Nature structural & molecular biology.* 2017;24:325-36.
- [43] Beauclair G, Bridier-Nahmias A, Zagury J-F, Saïb A, Zamborlini A. JASSA: a comprehensive tool for prediction of SUMOylation sites and SIMs. *Bioinformatics.* 2015;31:3483-91.
- [44] Sampson DA, Wang M, Matunis MJ. The Small Ubiquitin-like Modifier-1 (SUMO-1) Consensus Sequence Mediates Ubc9 Binding and Is Essential for SUMO-1 Modification *. *Journal of Biological Chemistry.* 2001;276:21664-9.
- [45] Ivanov AV, Peng H, Yurchenko V, Yap KL, Negorev DG, Schultz DC, et al. PHD Domain-Mediated E3 Ligase Activity Directs Intramolecular Sumoylation of an Adjacent Bromodomain Required for Gene Silencing. *Molecular cell.* 2007;28:823-37.
- [46] Welti S, Kühn S, D'Angelo I, Brügger B, Kaufmann D, Scheffzek K. Structural and biochemical consequences of NF1 associated nontruncating mutations in the Sec14-PH module of neurofibromin. *Human mutation.* 2011;32:191-7.
- [47] Zhou F, Xue Y, Lu H, Chen G, Yao X. A genome-wide analysis of sumoylation-related biological processes and functions in human nucleus. *FEBS letters.* 2005;579:3369-75.
- [48] Panse VG, Kressler D, Pauli A, Petfalski E, Gnädig M, Tollervey D, et al. Formation and Nuclear Export of Preribosomes Are Functionally Linked to the Small-Ubiquitin-Related Modifier Pathway. *Traffic.* 2006;7:1311-21.
- [49] Geiss-Friedlander R, Melchior F. Concepts in sumoylation: a decade on. *Nature reviews Molecular cell biology.* 2007;8:947-56.

- [50] Fuseya Y, Fujita H, Kim M, Ohtake F, Nishide A, Sasaki K, et al. The HOIL-1L ligase modulates immune signalling and cell death via monoubiquitination of LUBAC. *Nature cell biology*. 2020;22:663-73.
- [51] Vaser R, Adusumalli S, Leng SN, Sikic M, Ng PC. SIFT missense predictions for genomes. *Nature Protocols*. 2016;11:1-9.
- [52] Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. *Nature methods*. 2010;7:248-9.
- [53] Mohideen F, Capili AD, Bilimoria PM, Yamada T, Bonni A, Lima CD. A molecular basis for phosphorylation-dependent SUMO conjugation by the E2 UBC9. *Nature structural & molecular biology*. 2009;16:945-52.
- [54] Sun H, Hunter T. Poly-small ubiquitin-like modifier (PolySUMO)-binding proteins identified through a string search. *The Journal of biological chemistry*. 2012;287:42071-83.
- [55] Miteva M, Keusekotten K, Hofmann K, Praefcke GJ, Dohmen RJ. Sumoylation as a signal for polyubiquitylation and proteasomal degradation. *Sub-cellular biochemistry*. 2010;54:195-214.
- [56] Chang PC, Izumiya Y, Wu CY, Fitzgerald LD, Campbell M, Ellison TJ, et al. Kaposi's sarcoma-associated herpesvirus (KSHV) encodes a SUMO E3 ligase that is SIM-dependent and SUMO-2/3-specific. *The Journal of biological chemistry*. 2010;285:5266-73.
- [57] Sekiyama N, Ikegami T, Yamane T, Ikeguchi M, Uchimura Y, Baba D, et al. Structure of the small ubiquitin-like modifier (SUMO)-interacting motif of MBD1-containing chromatin-associated factor 1 bound to SUMO-3. *The Journal of biological chemistry*. 2008;283:35966-75.
- [58] Knipscheer P, Flotho A, Klug H, Olsen JV, van Dijk WJ, Fish A, et al. Ubc9 Sumoylation Regulates SUMO Target Discrimination. *Molecular cell*. 2008;31:371-82.
- [59] Jakobs A, Koehnke J, Himstedt F, Funk M, Korn B, Gaestel M, et al. Ubc9 fusion-directed SUMOylation (UFDS): a method to analyze function of protein SUMOylation. *Nature methods*. 2007;4:245-50.
- [60] Duarte ML, Pena DA, Nunes Ferraz FA, Berti DA, Paschoal Sobreira TJ, Costa-Junior HM, et al. Protein folding creates structure-based, noncontiguous consensus phosphorylation motifs recognized by kinases. *Science signaling*. 2014;7:ra105.
- [61] Bernier-Villamor V, Sampson DA, Matunis MJ, Lima CD. Structural Basis for E2-Mediated SUMO Conjugation Revealed by a Complex between Ubiquitin-Conjugating Enzyme Ubc9 and RanGAP1. *Cell*. 2002;108:345-56.
- [62] Cappadocia L, Pichler A, Lima CD. Structural basis for catalytic activation by the human ZNF451 SUMO E3 ligase. *Nature structural & molecular biology*. 2015;22:968-75.
- [63] Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;596:583-9.
- [64] Eisenhardt N, Chaugule VK, Koidl S, Droescher M, Dogan E, Rettich J, et al. A new vertebrate SUMO enzyme family reveals insights into SUMO-chain assembly. *Nature structural & molecular biology*. 2015;22:959-67.

- [65] Streich Jr FC, Lima CD. Capturing a substrate in an activated RING E3/E2-SUMO complex. *Nature*. 2016;536:304-8.
- [66] Pichler A, Gast A, Seeler JS, Dejean A, Melchior F. The Nucleoporin RanBP2 Has SUMO1 E3 Ligase Activity. *Cell*. 2002;108:109-20.
- [67] Reverter D, Lima CD. Insights into E3 ligase activity revealed by a SUMO-RanGAP1-Ubc9-Nup358 complex. *Nature*. 2005;435:687-92.
- [68] Varejão N, Lascorz J, Codina-Fabra J, Bellí G, Borràs-Gas H, Torres-Rosell J, et al. Structural basis for the E3 ligase activity enhancement of yeast Nse2 by SUMO-interacting motifs. *Nat Commun*. 2021;12:7013.
- [69] Long A, Liu H, Liu J, Daniel M, Bedwell DM, Korf B, et al. Analysis of patient-specific NF1 variants leads to functional insights for Ras signaling that can impact personalized medicine. *Human mutation*. 2021 43:30-41.
- [70] Ahmadian MR, Stege P, Scheffzek K, Wittinghofer A. Confirmation of the arginine-finger hypothesis for the GAP-stimulated GTP-hydrolysis reaction of Ras. *Nat Struct Biol*. 1997;4:686-9.
- [71] Rizza F, Vertemara J, Tisi R. Molecular Dynamics Simulations Reveal Structural Interconnections within Sec14-PH Bipartite Domain from Human Neurofibromin. *International journal of molecular sciences*. 2022;23:5707.
- [72] Bludau I, Willems S, Zeng WF, Strauss MT, Hansen FM, Tanzer MC, et al. The structural context of posttranslational modifications at a proteome-wide scale. *PLoS Biol*. 2022;20:e3001636.
- [73] Pichler A, Knipscheer P, Oberhofer E, van Dijk WJ, Körner R, Olsen JV, et al. SUMO modification of the ubiquitin-conjugating enzyme E2-25K. *Nature structural & molecular biology*. 2005;12:264-9.
- [74] Parker JL, Bucceri A, Davies AA, Heidrich K, Windecker H, Ulrich HD. SUMO modification of PCNA is controlled by DNA. *The EMBO journal*. 2008;27:2422-31.
- [75] Desterro JM, Rodriguez MS, Hay RT. SUMO-1 modification of I κ B α inhibits NF- κ B activation. *Molecular cell*. 1998;2:233-9.
- [76] Vertegaal AC, Ogg SC, Jaffray E, Rodriguez MS, Hay RT, Andersen JS, et al. A proteomic study of SUMO-2 target proteins. *The Journal of biological chemistry*. 2004;279:33791-8.
- [77] Cichowski K, Jacks T. NF1 tumor suppressor gene function: narrowing the GAP. *Cell*. 2001;104:593-604.
- [78] Cichowski K, Santiago S, Jardim M, Johnson BW, Jacks T. Dynamic regulation of the Ras pathway via proteolysis of the NF1 tumor suppressor. *Genes & development*. 2003;17:449-54.
- [79] Choi BH, Chen C, Philips M, Dai W. RAS GTPases are modified by SUMOylation. *Oncotarget*. 2018;9:4440-50.
- [80] Evans R, O'Neill M, Pritzel A, Antropova N, Senior A, Green T, et al. Protein complex prediction with AlphaFold-Multimer. *bioRxiv*. 2022:2021.10.04.463034.
- [81] Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: making protein folding accessible to all. *Nature methods*. 2022;19:679-82.

SUPPORTING INFORMATION: SUPPLEMENTAL MATERIAL

Table S1: List of oligonucleotides

Table S2: List of plasmids

Table S3: List of primary antibodies

Table S4: Software and Databases

Table S1: List of oligonucleotides

All oligonucleotides sequences are from 5' to 3' end

Oligonucleotides for SUMO2 cloning into p3X Flag-Myc-CMV™-24 Expression Vector (plasmid p3X Flag SUMO2)

Flag SUMO2-for CGCGCGAAGCTTATGTCCGAGGAGAAGCCCAAG HindIII

Flag SUMO2-rev CGCGCGGGATCCCTAACCTCCCGTCTGCTGCT BamHI

Italics in the sequence denote the indicated restriction sites

Oligonucleotides for GRD and SecPH mutagenesis

R1276Q For GACCCTGTTCCaAGGCAACTCCCTGG

Rev TGCATGGAGTCGGCCAGC

K1593R For GTTGACCGGAGGTTCCgAACTGGTCAAATCAATGGT

Rev ATTGATTTGACCAGTTcgGAACCTCCGTGCAACATA

K1634R, For CAATCGCTTTAgAACAGACTTTCTCTCTAAG

Q5

Rev CTAGGCCCGGTATGGGTA

K1634R For CCTAGCAATCGCTTTcgAACAGACTTTCTCTCTAAGTG

Rev GAGAGAAAGTCTGTTcgAAAGCGATTGCTAGGCC

K1717R For TTAGAAGAGGACCTGAgGGTATTCCACAATGCTCTC

Rev GAGCATTGTGGAATACCcTCAGGTCCTCTTCTAAAGC

L1723A For GTATTCCACAATGCTgcCAAGCTAGCTC

Rev CTCAGGTCCTCTTCTAAAGCCAAGGT

D1729N	For	AGCTCACAAAaACACCAAAGTTC
	Rev	AGCTTGAGAGCATTGTGG
D1729A	For	GCTCACAAAGcCACCAAAGTTTC
	Rev	TAGCTTGAGAGCATTGTG
K1731R	For	AAAGACACCAgAGTTTCTATTAAGTTG
	Rev	GTGAGCTAGCTTGAGAGC
K1731R	For	GCTCACAAAGACACCAgAGTTTCTATTAAGTTGGT
	Rev	AACTTTAATAGAAACTcTGGTGTCTTTGTGAGCTAG
V1732G	For	GACACCAAAGgTTCTATTAAG
	Rev	TTTGTGAGCTAGCTTGAG
K1735R	For	GTTTCTATTAgAGTTGGTTCTACTG
	Rev	TTTGGTGTCTTTGTGAGC
E1747A	For	ACTTCAGCAGcGCGAACAAAAG
	Rev	TACTTGGACAGCAGTAGAAC
R1748A	For	GCTGTCCAAGTAACTTCAGCAGAGgcAACAAAAGTCCTAGGGCAATCA GTC
	Rev	GACTGATTGCCCTAGGACTTTTGTGgcCTCTGCTGAAGTTACTTGGACA GC
K1750A	For	AGAGCGAACAgctGTCCTAGGGC
	Rev	GCTGAAGTTACTTGGACAG
D1729N S1733A	For	AGTTgCTATTAAGTTGGTTCTACTGC
	Rev	TTGGTGTtTTTGTGAGCTAGCTTGAG
D1729N V1732G S1733A	For	aACACCAAAGgTgCTATTAAG
	Rev	TTTGTGAGCTAGCTTGAG
K1731, 1735R	For	GCTCACAAAGACACCAgAGTTTCTATTcgAGTTGGT
	Rev	AACTcgAATAGAAACTcTGGTGTCTTTGTGAGCTAG

Nucleotides substitutions are in red

Oligonucleotides for GRD cloning into p3X Flag-Myc-CMV™-24 Expression Vector (p3X Flag GRD)

GRD-
SecPH-
EX-for TACTGCGGCCGCGGGTTACCACAAGGATCTCCAGACAAGAGCTACATTTATG
NotI

Nf1 Flag GCTGACGTCGACTCATGTATCTGCCACAGGTTTGTGCTCT
GRD2-rev Sall

Italics in the sequence denote the indicated restriction sites

Oligonucleotides for *NF1* cloning into p3X Flag-Myc-CMV™-24 Expression Vector

- p3X Flag Nf1 SecPH Nat *alias* p3X Flag Nf1 [p3X Flag (Nter opt* - GRD opt - SecPH Nat** - Cter opt)
- p3X Flag Nf1 GRD Nat [p3X Flag (Nter opt - GRD Nat - SecPH opt - Cter opt)]

Amplicon	Name	Sequence
p3X Flag to assemble with <u>Nter opt</u> and <u>Cter opt</u>	p3X Flag-for	GTCGACTCTAGAGGATCCG
	p3X Flag-rev	ctgcaccattccacaggctctgtgggctg cGGCCGCAAGCTTGTC
<u>Nter opt</u> to assemble with <u>GRD Nat</u> and p3xFlag	Nter opt-for	GCCGCCACAGACCTGTG
	Nter opt-rev GRD Nat	agctctgtctggagatcctgtgtaacc CAGGCCGATGCTGTGCAT
<u>GRD Nat</u> to assemble with <u>SecPH opt</u> and <u>Nter opt</u>	GRD-for	GGTTACCACAAGGATCTCCA
	GRD Nat-SecPH opt-rev	gctgctgtaaggtaaggctggaccagtg TGTATCTGCCACAGGTTTGT
<u>SecPH Nat</u> to assemble with <u>Cter opt</u> and <u>GRD opt</u>	◇ SecPH-for	CACTGGTCCAGCCTTAACC TTACCAGTTCAAAGTTTGAG GAATTTATGACTAGG
	SecPH-rev	gggccgatcttggtgtgctgggggatgc tGTCGGGCTGTGACAGTTC

	Cter opt-for	AGCATCCCCCAGCACACC
<u>Cter opt</u> to assemble with p3X Flag and <u>SecPH Nat</u>	Cter opt-rev	gagttttgttcggatcctctagagtcgac ATTACACGATTTTCTTGATG CTGTTC
<u>Nter-GRD opt</u> to assemble with <u>SecPH Nat</u>	Nter opt-for	See above
	GRD opt-SecPH Nat-rev	tgaactggtaaggtaaggctggaccag tgGGTGTCCGCCACAGGCTT
<u>SecPH-Cter opt</u> to assemble with p3X Flag and <u>GRD Nat</u>	◇ SecPH opt-for	CACTGGTCCAGCCTTAACC TTACCAGCAGCAAGTTCGA GGAATTCATG
	Cter opt-rev	See above
	Nter opt-for	See above
	Cter opt-rev	See above
<u>Nter-GRD opt -SecPH Nat -Cter opt</u> to assemble with p3X 2KR Flag	p3X Flag-for	See above
	p3X 2KR Flag -rev	ctgcaccattccacaggtctgtggcgg cGGCCGCAAGC <u>g</u> TGTCATCG TCATCC <u>g</u> TGTAATCGATGTC ATG

* Nf1 domains with optimised sequences and amplicons to create them are highlighted with wavy lines and the name of the corresponding oligonucleotides and the domain created harbours the suffix "opt". **Nf1 domains with native sequences and the corresponding oligonucleotides to create them are denoted "Nat" and underlined. The oligonucleotides sequences in small letters emphasize the homology between overlapping primers, capital letters show the sequence that hybridise with the template and oligonucleotides labelled ◇ contain both native and optimised sequences encoding residues between GRD and SecPH domains (bold letters). The nucleotide substitutions to create K to R mutation of the last two lysines of the Flag tag in p3X Flag are in small letters, bold and red.

Table S2: List of plasmids

p3X Flag SecPH derivatives	Features	Source
p3X Flag SecPH	WT SecPH	Ref 32
p3X Flag SecPH K1593R	K1593R	This work
p3X Flag SecPH K1717R	K1717R	This work
p3X Flag SecPH L1723A	L1723A	This work
p3X Flag SecPH A1726P	A1726P	GenScript
p3X Flag SecPH D1729A	D1729A	This work
p3X Flag SecPH D1729N	D1729N	This work
p3X Flag SecPH K1731R	K1731R	This work
p3X Flag SecPH K1735R	K1735R	This work
p3X Flag SecPH E1747A	E1747A	This work
p3X Flag SecPH R1748A	R1748A	This work
p3X Flag SecPH K1750A	K1750A	This work
p3X Flag SecPH L1723A E1747A	Oligonucleotide E1747A into p3X Flag SecPH L1723A	This work
p3X Flag SecPH V1732G E1747A	Oligonucleotide V1732G into p3X Flag SecPH E1747A	This work
p3X Flag SecPH H1727A K1728A	H1727A K1728A	GenScript
p3X Flag SecPH K1634R K1731R	Oligonucleotide K1634R was introduced into p3X Flag SecPH K1731R	This work
p3X Flag SecPH L1723A D1729N	L1723A was introduced into p3X Flag SecPH D1729N with primer L1723A	This work
p3X Flag SecPH D1729N S1733A	D1729N S1733A	This work
p3X Flag SecPH D1729N V1732G S1733A	D1729N V1732G S1733A	This work

p3X Flag SecPH 5K-R	K1731R was introduced into p3X Flag SecPH K1735R with primers K1731,1735R; then K1593R, K1634R, K1717R were introduced one after the other by successive rounds of mutagenesis	This work
p3X Flag SecPH E1775A N1776A Q1794A	E1775A N1776A Q1794A	GenScript
p3X Flag SecPH K0	21 lysines in SecPH mutated into arginines	GenScript
p3X 2KR Flag SecPH	The two lysines of the last Flag tag replaced by arginine	GenScript
p3X 2KR Flag SecPH K0	21 lysines in SecPH mutated into arginines	GenScript
p3X 2KR Flag SecPH K0 K1731	K0 K1731	GenScript
p3X 2KR Flag SecPH K0 K1731 K1634	K0 K1731 K1634	GenScript
p3X 2KR Flag SecPH K0 K1735	K0 K1735	GenScript
pUbc9-SecPH-3xHA	Ubc9 and SecPH 3xHA fused	GeneCust
pUbc9-SecPH-K1731R-3xHA	Ubc9 and SecPH K1731R 3xHA fused	GeneCust
pUbc9-SecPH-A1726P-3xHA	Ubc9 and SecPH A1726P 3xHA fused	This work
pUbc9-SecPH-K1634R-3xHA	Ubc9 and SecPH K1634R 3xHA fused	This work
pUbc9-SecPH-R1748A-3xHA	Ubc9 and SecPH R1748A 3xHA fused	This work
pUbc9-SecPH-K1750A-3xHA	Ubc9 and SecPH K1750A 3xHA fused	This work
pUbc9-K14R-SecPH-3xHA	Ubc9-K14R and SecPH-3xHA fused	GeneCust
pUbc9-K14R-SecPH-K1731R-3xHA	Ubc9-K14R and SecPH K1731R 3xHA fused	GeneCust

p3X Flag Nf1 derivatives*		
p3X Flag Nf1 SecPH Nat <i>alias</i> p3X Flag Nf1		This work
p3X Flag Nf1 GRD Nat		This work
p3X Flag Nf1 R1276Q	R1276Q	This work
p3X Flag Nf1 K1731R	K1731R	This work
p3X 2KR Flag Nf1	The two lysines of the last Flag tag replaced by arginine	This work
p3X 2KR Flag Nf1 K1731R	K1731R	This work
p3X 2KR Flag Nf1 K1634R K1731R	K1634R K1731R	This work

*Domains which were created from native sequences harbour the Nat suffix; by default, the other domains were generated with optimised codons

Miscellaneous plasmids		
Addgene plasmid # 70423		R777-E139 Hs.NF1
pcDNA3		Invitrogen
pcDNA3 Ubc9	pcDNA3 Ubch9/SV5	kindly provided by R. Hay
pcDNA3 6His-SUMO2	pcDNA3-His6-SUMO-2	kindly provided by R. Hay
MDH1-Flag		GeneCopoeia (EX-C0441-M13-10)

Table S3: List of primary antibodies

Target	Dilution	Biological source	Company	Reference
ERK	1:1000	Rabbit	Cell Signaling Technology	4695

pERK	1:2000	Rabbit	Cell Signaling Technology	4370
Flag	1:5000	Mouse	Sigma-Aldrich	F3165
GAPDH horseradish peroxidase conjugate	1:100 000	Mouse	Sigma-Aldrich	G9295
HA	1:5000	Rat	Sigma-Aldrich	3F10
NF1	Undiluted or diluted	Rabbit	Novus Biologicals	NB100-418
SUMO2/3	1:2000	Mouse	Abcam	ab81371

Table S4: Software and Databases

PyMOL	http://www.pymol.org/pymol
SUMOPlot™ Analysis Program	https://www.abcepta.com/sumoplot
PolyPhen-2	http://genetics.bwh.harvard.edu/pph2/
SIFT	https://sift.bii.a-star.edu.sg
StatEL	https://www.adscience.fr/logiciels/statel
ClinVar	https://www.ncbi.nlm.nih.gov/clinvar/
gnomAD	https://gnomad.broadinstitute.org/gene/ENSG00000196712?dataset=gnomad_r2_1

