Computational and Structural Biotechnology Journal 19 (2021) 5622–5636

COMPUTATIONAL AND STRUCTURAL BIOTECHNOLOGY JOURNAL

journal homepage: www.elsevier.com/locate/csbj

A conserved Neurite Outgrowth and Guidance motif with biomimetic potential in neuronal Cell Adhesion Molecules

Giorgia Scapin [a,1], Matteo Gasparotto [a,1], Daniele Peterle [b,1], Simone Tescari [b], Elena Porcellato [a,b], Alberto Piovesan [a,b], Irene Righetto [a], Laura Acquasaliente [b,*], Vincenzo De Filippis [b,*], Francesco Filippini [a,*]

[a] Synthetic Biology and Biotechnology Unit, Department of Biology, University of Padua, 35131, Italy
[b] Department of Pharmaceutical and Pharmacological Sciences, University of Padua, 35131, Italy

ARTICLE INFO

Article history:
Received 7 July 2021
Received in revised form 27 September 2021
Accepted 3 October 2021
Available online 12 October 2021

Keywords:
Neurite outgrowth
Axon guidance
Protein-protein interaction
NOG motif
Homophilic binding
Ig-like fold
Biomimetic peptide
Regenerative medicine
Neuronal differentiation

ABSTRACT

The discovery of conserved protein motifs can, in turn, unveil important regulatory signals, and when properly designed, synthetic peptides derived from such motifs can be used as biomimetics for biotechnological and therapeutic purposes. We report here that specific Ig-like repeats from the extracellular domains of neuronal Cell Adhesion Molecules share a highly conserved Neurite Outgrowth and Guidance (NOG) motif, which mediates homo- and heterophilic interactions crucial in neural development and repair. Synthetic peptides derived from the NOG motif of such proteins can boost neuritogenesis, and this potential is also retained by peptides with recombinant sequences, when fitting the NOG sequence pattern. The NOG motif discovery not only provides one more tile to the complex puzzle of neuritogenesis, but also opens the route to new neural regeneration strategies via a tunable biomimetic toolbox.

1. Introduction

In the last decades, computational biology has strongly improved discoveries in life sciences and biotechnological design via the identification of functional motifs. Thanks to an ever-growing number of sequenced genomes and hence of inferred protein sequences, proteome-wide comparative analyses allowed to discover and define many conserved domains, functional and signalling motifs and protein family signatures. Once such functional information is progressively deposited, as either matrix-based profiles or regular expression patterns, in knowledgebases and databases such as e.g. PROSITE, CDD, Pfam, and InterPro [1–4], it can improve the functional annotation of genomes, metagenomes and proteomes, as well as provide new insights on a high number of individual records in bio databases. This is particularly important with newly sequenced and still uncharacterized genes and inferred protein products. Soon after gene prediction and ORF inference, domain profiles and motif patterns allow for *in silico* functional annotation, for instance, by defining domain architecture, classifying families, or predicting interactions. Motifs can also improve ORF prediction by suggesting the proper translation frame [5].

When proteins endowed with a newly identified interaction motif are found to share a specific function, new regulatory mechanisms can be unveiled, opening the route to further computational prediction of interactors and interactome nodes. Interaction maps can help the design of experiments to better characterize protein complexes and their roles in biological pathways. For newly discovered motifs, it is crucial to consider their 3D position: when located at the protein surface, they are likely to mediate/regulate interactions, whereas when positioned in the protein core, they are likely structural determinants. Advancements in structural biology and bioinformatics greatly simplified the retrieval of such data. Large scale structural biology projects have provided the scientific community with a sufficient number of reference structural templates [6], and even when the original resolution is low, it can be reconstructed and improved by computational methods [7]. Moreover, improved structural modelling and refinement methods allow for obtaining reliable models of

* Corresponding authors.
E-mail addresses: laura.acquasaliente@unipd.it (L. Acquasaliente), vincenzo.defilippis@unipd.it (V. De Filippis), francesco.filippini@unipd.it (F. Filippini).
[1] Equal contributions.

https://doi.org/10.1016/j.csbj.2021.10.005

most target proteins/domains [8], especially when concerning well-characterized model organisms and, of course, human proteins.

Relevance of the identification of interaction motifs goes beyond basic research and genome/proteome annotation, as the development of synthetic peptides derived from such motifs can be a cheap and easy-to-use alternative to producing and using the whole recombinant parent proteins. Chemical synthesis of peptides proved to have several advantages compared to using entire recombinant proteins, such as: (i) low immunogenic activity, (ii) increased stability, (iii) low production costs and (iv) simplified preparation and immobilization onto substrates. Furthermore, peptides can be: (v) presented to cells at surface densities significantly higher than those possibly achieved with entire proteins or domains and (vi) tailored in composition for each tissue-specific application [9].

Recently, we have shown that synthetic peptide L1-A can boost neuritogenesis, with potential application in regenerative medicine [10]. L1-A is derived from the Ig2 repeat in the extracellular domain (ED) of human Cell Adhesion Molecule (CAM) L1, a membrane protein acting as a positive regulator in neuronal differentiation, neuritogenesis and axon regeneration [11,12]. L1CAM is a prototypical member of the L1 family and other neuronal CAMs sharing large EDs with four to six Ig-like repeats, followed by a variable number of fibronectin type III regions. The L1CAM Ig2 repeat belongs to the immunoglobulin-like C2-type (Ig-like C2) subclass and, not surprisingly, CAMs with Ig-like C2 repeats play a fundamental role in cell-cell and cell-extracellular matrix (ECM) interactions in both mature and developing nervous system, as well as in axonal regeneration and neural repair [13]. The homo/heterophilic ED-ED binding is prototyped for L1 family proteins by the homodimer structure of Neurofascin, in which the Ig1 to Ig4 domains form a horseshoe structure within each monomer, whereas the two Ig2 domains bind each other [14]. Notably, L1-A sequence partially corresponds to the homophilic binding region of Ig2.

Evidence that the L1-A peptide is part of the conserved Ig-like C2 fold, shared among several neuronal CAMs, prompted us to investigate the existence of a conserved Neurite Outgrowth and Guidance (NOG) motif and to develop several natural and chimeric NOG-derived peptides with neuritogenic potential. By means of computational, molecular, cellular, and biophysical analyses, we opened the route for the development of a NOG-based biomimetic toolbox both for regenerative medicine and developmental biology studies.

2. Materials and methods

2.1. Sequence-based analyses

Homology searches were carried out using BLASTp [15] with default settings and L1CAM Ig2 or L1CAM family proteins Ig sequences as queries for defining an initial dataset for multiple alignment. Then, regular expressions for positions centred around the peptide regions were inferred and merged to derive a single pattern, which was written in PROSITE syntax [16]. Pattern refinement was performed by iterative scanning with ScanProsite [17] of Vertebrata and Mammalia proteome sections of the UniProtKB database [18], excluding splice isoforms and protein fragments. The precision index was calculated following PROSITE rules, i.e., as true positive hits / (true positive + false positive hits) ratio. Pairwise sequence comparison by global William Pearson's alignment for sequences corresponding to the different Ig regions that were superimposed (see section below) was performed using LALIGN (https://www.ebi.ac.uk/Tools/psa/lalign/), which implements the algorithm of Huang and Miller [19]. LALIGN was run with default settings, except for matrix set to BLOSUM62.

2.2. Structural comparison, modeling, electrostatics

The following structures from the Protein Data Bank (PDB) were used for comparison and/or as templates for homology modeling: 3P3Y (Neurofascin, UniProt AC O94856) [14], 3S97 (Contactin 1, CNTN1, AC Q12860) [20], 2OM5 (Contactin 2, CNTN2, AC Q02246) [21], 2V9R (Roundabout receptor 1, ROBO1, AC Q9Y6N7) [22], 3LAF (Deleted in Colon Cancer/Netrin receptor, DCC/NetrinR, AC P43146) [23]. Structural superpositions were performed and Root Mean Square Deviation (RMSD) was calculated using UCSF Chimera [24,25] v. 1.15. Structural models were obtained via SWISSMODEL [26] for target human proteins: L1CAM (UniProt AC P32004), Close Homolog of L1 (CHL1, AC O00533), NgCAM-related (NrCAM, AC Q92823), contactins CNTN3 (AC Q9P232), CNTN4 (AC Q8IWV2), CNTN5 (AC O94779), CNTN6 (AC Q9UQ52), Roundabout receptors ROBO2 (AC Q9HCK4) and ROBO3 (AC Q96MS0). Next, modeled structures were refined using SCWRL [27,28]. Model quality was checked via QMEAN server [29].

Electrostatic maps were obtained via DelPhi webserver [30,31]. Isopotential contours were calculated using UCSF Chimera1.15 and plotted at ± 2kBT/e. PDB2PQR was used to assign partial charges and van der Waals radii according to the PARSE force field [32]. Dielectric constant values for protein interior (ε_p) and solvent (ε_s) were set as $\varepsilon_p = 2$ and $\varepsilon_s = 78.54$ [33–35], T = 298.15 K. Probe radius for dielectric surface and ion accessibility surface were set as r = 1.4 Å and r = 2.0 Å, respectively.

2.3. Docking simulations and molecular dynamics

Docking of synthetic peptides to their targets was simulated using PepDock at GalaxyWeb [36] and CABSdock standalone [37]. Docking results were then refined using GalaxyRefineComplex [38] at GalaxyWeb server. Prior to performing docking runs, targets were submitted to UCSF Chimera Dockprep routine. Interactions between peptides and target proteins were evaluated using UCSF Chimera. Molecular Dynamics (MD) was performed with Gromacs 2020.3 [39] using the Amber03 force field [40], in order to investigate the role of mutations on the structural organization of the prototypical Neurofascin Ig2 structure. The models were solvated with the TIP3P water model in a rectangular box with a minimum distance of 1 nm between the protein and the border. 0.15 M of NaCl was added to simulate a realistic ionic strength. System energy was minimized by 5000 steps of energy minimization, with a tolerance of 1000 kJ $mol^{-1}nm^{-1}$. Subsequently, a 200 ps NVT MD simulation was used to heat the system from 0 to 100 K with restraints lowered to 400 kJ mol^{-1} nM^{-2} and then the system was heated up to 310 K in 400 ps during a NPT simulation with further lowered restraint (200 kJ mol^{-1} nM^{-2}). Finally, the system was equilibrated during a NPT simulation for 1 ns with backbone restraints lowered to 50 kJ mol^{-1} nM^{-2}. All restraints were removed for the production run at 310 K. The Berendsen thermostat [41], was used to equilibrate the temperature. The average pressure was kept at 0.01 atm [42] using the Parrinello-Rahman barostat [43]. Newton's equation of motion was integrated using a leapfrog algorithm with a 2 fs time step. The particle mesh Ewald (PME) method [44] was used to compute the long-range electrostatic forces. Rotational and translational motions of the system were removed, and H-bonds were constrained with the LINCS algorithm. The simulation was run for 300 ns, until the model reached a stable conformation, as confirmed by RMSD.

2.4. Peptide synthesis, purification and characterization

All peptides were synthesized in our laboratory by the solid-phase method using the flurenylmethyloxycarbonyl(Fmoc)-chemis try [45,46] on a model PS3 automated synthesizer from Protein Technologies International (Tucson). The peptides were assembled stepwise on a Wang Resin (Novabiochem) derivatized with the desired corresponding C-terminal amino acid. Removal of N^{a}-Fmoc-protecting groups was achieved by treatment for 20 min at room temperature with a deprotection solution (20% piperidine in N-methylpirrolidone - NMP). Standard coupling reactions were performed with an equal molar ratio of 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate and 1H-hydroxy-benzotriazole as activating agents, with a fourfold molar excess of N^{α}-Fmoc-protected amino acids in activation solution. For double coupling at peptide bonds involving Val, Ile, Leu and Phe, the stronger activator 2-(7-aza-1H-benzotriazol-1-yl)-1,1,3,3 -tetramethyluronium hexafluorophosphate was used. Once the peptide assembly was completed, the side chain-protected peptidyl resin was treated for 90 min at room temperature with the following mixture: 92.5% TFA, 2.5% H_2O, 2.5% ethandithiol, 2.5% triisopropylsilane. The resin was removed by filtration, and the acidic solution, containing the unprotected peptide, was precipitated with ice-cold tertbutyl-methylether and then lyophilized.

Peptides were purified to homogeneity (>98%) by semi-preparative RP-HPLC (Jasco HPLC Pu-1575 equipped with 1575 UV–Vis detector) on a Grace-Vydac (Hesperia) C18 column (4.6 × 150 mm, 5 μm particle size, 300 Å porosity) equilibrated with 0.1% (v/v) aqueous TFA and eluted with a linear 0.078% (w/w) TFA-acetonitrile gradient at a flow rate of 0.8 ml/min. Absorbance of the effluent was monitored at 226 nm.

Purified peptides were analyzed using a Waters (Milford) Xevo-G2S Q-TOF instrument, which yielded mass values in agreement with the theoretical mass within 2 ppm accuracy. Concentration of peptides with aromatic amino acids was determined on a V-630 spectrophotometer (Jasco, Tokyo, Japan) by measuring the absorbance at 257 nm, for Phe-containing peptides, or at 280 nm for Tyr-containing peptides, using a molar absorptivity of 200 M^{-1}·cm^{-1} or 1280 M^{-1}·cm^{-1} for Phe or Tyr, respectively. For peptides lacking any suitable chromophore, the concentration was determined by analytic scaling (E/50 Gibertini) [47].

Peptide secondary structures were analyzed by circular dichroism (CD) in the far-UV region. CD spectra were recorded on a J-810 spectropolarimeter (Jasco), equipped with a thermostated cell holder and a Peltier PTC-423S temperature control system. The spectra were recorded at 25 °C at a peptide concentration of 0.1 mg/ml in 10 mM phosphate buffer, pH 7.4, 0.15 M NaCl (PBS) using a 1-mm 0.1-cm pathlength cuvette. Each spectrum resulted from the average of four accumulations after baseline subtraction. CD signal was expressed as the mean residue ellipticity: $[\theta] = \theta_{obs}$·MRW/10·l·c (deg·cm^2·$dmol^{-1}$) where θ_{obs} is the observed ellipticity in deg, MRW is the mean residue weight, l is the path length in cm and c is the peptide concentration in g/ml [48,49].

2.5. Peptide stability assays

Each peptide was added at 20 μM final concentration in 600 μl of differentiation medium, whose composition is reported in the Molecular and Cellular biology section. At each time point, 500 μl of medium was withdrawn and high MW molecules from fetal bovine serum (FBS) or secreted by cells were eliminated by ultrafiltration at 10,000×*g* on a Vivaspin 10 kDa cutoff filter (Sartorius). 350 μl of the ultrafiltrated medium was immediately frozen at −20 °C for the RP-HPLC analysis. RP-HPLC runs were performed by loading 300 μl of the sample acidified to 0.1% H_2O/TFA onto a Zorbax 300SB-C18 column (4.6 × 300 mm, 5 μm particle size, 300 Å porosity) (Agilent) using a linear acetonitrile/0.078% TFA gradient from 5 to 34% in 30 min at 0.8 ml/min flow rate, monitoring the absorbance of the effluent at 226 nm. Peptide concentration at each time point was inferred from its peak area, normalised vs a reference peak showing a constant area during the whole experiment.

2.6. Surface plasmon resonance (SPR) experiments

Binding affinity was carried out using a multi-cycle injection strategy on a dual flow cell Biacore-X100 instrument (GE-Healthcare). Homophilic binding was validated as follows: 5 μg/ml purified L1CAM ED (Thermofisher Scientific) (i.e. the ligand) was immobilized (3300 RU) at pH 4.0 on a carboxymethylated-dextran chip (CM5) using the amine coupling chemistry; then, increasing concentrations of L1CAM ED solutions (i.e. the analyte) were injected in the mobile phase [50]. Titration was performed at 25 °C in 10 mM Hepes, pH 7.5, 150 mM NaCl, 3 mM EDTA, 0.05% polyoxyethylene sorbitan (HBS-EP$^+$), at a flow rate of 10 μl/min. The regeneration step was performed using 0.05% (w/w) SDS solution for 20 sec. Each binding curve was subtracted for the corresponding baseline, accounting for non-specific binding (<2% of RU_{max}). The equilibrium dissociation constant (K_d) of the homophilic interaction was obtained as a fitting parameter by plotting the value of the response units at the steady state (RU_{max}) for each concentration of the analyte (i.e., L1CAM ED). Data analysis was performed using the BIAevalution software and OriginPro 2018b, using the following equation, describing the one-site binding model:

$$RU_{eq} = RU_{max}\frac{[A]_F}{[A]_F + K_d}$$

where $[A]_F$ is the concentration of the free analyte in equilibrium with the analyte-ligand complex present on the sensor chip surface, while RU_{eq} and RU_{max} are the RU values measured at the steady state with intermediate or saturating A_F concentrations. Competition assays with peptides in solution were carried out by incubating at constant concentration of L1CAM ED (50 nM) with increasing concentrations of peptides up to 100 μM for 30 min, followed by injection over the sensor chip. All measurements were carried out at a flow rate of 10 μl/min, at 25 °C using HBS-EP$^+$ as running buffer. The SPR signal was taken at the end of the association phase and plotted against each peptide concentration.

2.7. Cell culture and differentiation

Exponential growing human neuroblastoma cell line SH-SY5Y [51], was cultured with Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12 with GlutaMAX™ supplement (DMEM/F-12, Invitrogen Life Technologies) supplemented with 10% heat-inactivated foetal bovine serum (FBS, Euroclone) and 25 μg/ml of gentamicin (Sigma) (growth medium), in a humidified atmosphere of 5% of CO_2 in air at 37 °C. Cultures were maintained by sub-culturing cells into 25 cm^2 flasks (Sarstedt) once they reached roughly 80% confluence. Cell differentiation was induced by treating cells with all-*trans*-retinoic acid (RA, Sigma) at 10 μM concentration and lowering the FBS in the culture medium to 2% (differentiation medium) 24 h after seeding. In undifferentiated control samples, Dimethyl sulfoxide (DMSO) was added as the equivalent amount (in which RA is dissolved). In experiments with peptides added to the culture medium, cells were seeded in a 24-well plate (15000 cells/well) coated with a gelatine (Sigma)/poly-L-lysine (Invitrogen) solution. Poly-L-lysine is widely used as a good substrate for neural cell adhesion and growth. 24 h after cell seeding (day 0), the growth medium was replaced by the differentiation medium; then, 24 h

after RA induction (day 1) peptides were added to the culture medium, except for control samples. Cell viability and proliferation were assessed at reported time points, while neurite numbers and lengths were measured 24 h after peptide addition.

2.8. Neuritogenesis assay

Neurite outgrowth was measured after staining cells with Calcein-AM (Biotium, 2 μM in HBSS, Hank's Balanced Salt Solution, Invitrogen) and Hoechst 33258 (Invitrogen Life Technologies, 10 μg/ml) for 30 min in the dark at 37 °C and 5% CO_2. Medium was then replaced with fresh HBSS and cells were observed with a Leica DMI4000 microscope at 10X magnification with GFP and DAPI filter. Ten images per well were recorded; the first two fields were set to correspond to the centre of the well. Next fields were then selected in the periphery of the well (N, NE, E, SE, S, SW, W, NW, respect the centre), so that images could be representative of the whole well. Images were evaluated with Fiji [52]. Cells were counted by manually counting nuclei. Overall, between 1000 and 2000 cells were recorded per well. Neurite length was measured by tracing the trajectory of the neurite from the tip to the junction between the neurite and cell body. If a neurite exhibited branching, the measure from the end of the longest branch to the soma was recorded, then each branch was measured from the tip of the neurite to the neurite branch point. Only neurites longer than 50 μm were considered [53,54]. The neuritogenic properties were analysed in terms of total neurite length/no. of cells (aggregate length of all cellular processes divided by cell number) and no. of neurites/no. of cells. Values were then normalized to the untreated proliferative control and reported as percentage. Each experiment was performed in at least three independent replicates. As not all peptides were tested at the same moment, negative (proliferative and differentiative untreated) and positive (proliferative and differentiative L1-A treated) controls were always included in the assay to ensure reproducibility.

2.9. Immunofluorescence

Cells were fixed in 4% paraformaldehyde (PFA) for 15 min and permeabilized by treatment with ice cold 100% methanol for 10 s. Samples were then blocked in 0.5% BSA in PBS for 45 min at room temperature (RT). Staining was performed for 90 min at RT using MAP2 primary antibody (Synaptic Systems) diluted in 3% BSA in PBS. Secondary antibody (Alexa Fluor 488, Molecular Probes-ThermoFisher) was diluted in 0.5% BSA and incubated for 45 min at RT. Finally, nuclei were stained with 1 μg/ml HOECHST 33258 (Invitrogen) for 20 min and coverslip were mounted with mowiol mounting medium. After 24 h of polymerization samples were observed using a Leica SP5 confocal microscope.

2.10. Cell proliferation assays

Resazurin reduction assay was performed to quantify metabolically active living cells and thus to monitor the effects of peptides on cell proliferation. Briefly, the culture medium was replaced by 500 μl of resazurin solution (Resazurin Sigma 15 μg/mL in growth medium without phenol red) and cells were incubated for 4 h in the dark at 37 °C, 5% CO_2. Then, 200 μl of resazurin solution was removed twice from each well and transferred to a 96 well plate (technical duplicates). Fluorescence, directly correlated with cell numbers, was detected using a plate reader (Ascent Fluoroscan, excitation 540 nm, emission 590 nm). Background values from blank samples were subtracted and average values for the duplicates calculated. Cell proliferation was calculated from a calibration curve by linear regression using Microsoft Excel.

2.11. Statistical analysis

Statistical analysis was performed using GraphPad Prism v8. Unpaired, two-tailed Student's *t*-test and results were considered significant when $p < 0.05$.

3. Results

3.1. Identification of a neurite outgrowth and guidance (NOG) motif

The amino acid sequences of Ig2 repeats from human L1CAM and the other three proteins of the L1 family (i.e., CHL1, NrCAM and Neurofascin) were used as BLASTp queries to retrieve a dataset of animal orthologs. A regular expression (pattern) was derived from a multiple alignment including these sequences, according to the PROSITE syntax [55]. In particular, the seed region of the pattern was derived by a 14 residues block of the multiple alignment including the sequence of the biomimetic L1-A peptide [10]. However, since the shorter is a regular expression, the higher is the number of retrieved false positive hits, to improve scanning specificity the pattern was extended C-terminally, including the complete β-hairpin region of the Ig2 domain and a conserved Asp residue. Early ScanProsite [17] scannings retrieved further neuronal CAMs, other than L1 family; refinement by iteration resulted in the following pattern (standard PROSITE synthax): [FHIKLPRST]-{CHMNWY}(2)-{CFHVWY}-[DEGINRSTV]-[DEGAKLS]-R-{CDENPQW}-[EFINSTVY]-[AIKLMQTV]-[ADFGKLMRS]-{CEFIMWY}-[DEGNKST]-[GKNST]-[DFGILNTYA]-[LIMN]-[YEFMTILQ]-[FVIGSY]-[HAINST]-x(4,6)-D. Scanning of UniProtKB release 2021_03 (SwissProt + TrEMBL, sequence fragments and splice variants excluded) confirmed the taxonomic restriction of this motif to Metazoa (Taxon ID: 33208). Almost ten thousand positive hits were retrieved, mainly from Vertebrata (Taxon ID: 7742), in agreement with the specificity for neuronal proteins involved in Neurite Outgrowth and Guidance, prompting us to name such newly identified conserved pattern as 'NOG motif'. The prediction power of a PROSITE pattern is named "Precision Index" and indicates specificity of pattern presence in true positive hits (i.e., those proteins having both the pattern and the associated motif function). Precision Index is calculated as true positive/true + false positive ratio, and thus it is 100% when only true positives are retrieved, while for some patterns, this index shows acceptable, high values lower than 100% because of recognition of some false positive hits (i.e., proteins in which a sequence fragment, non-associated to the motif function, is compatible with the pattern). The Precision Index is calculated using only entries from the SwissProt section of UniProtKB, because protein sequences in the TrEMBL subsection are often computationally inferred and/or uncharacterized. When calculated this standard way, on a few tens positive SwissProt entries, NOG motif precision was found to be 100%; however, we decided to check whether such an absolute specificity could be confirmed also in the much broader dataset of >3600 positive hits in TrEMBL (Mammalia section, Taxon ID 40674), as small numbers may hide false positive hits, which are eventually found in extended datasets. In TrEMBL section, roughly two thirds of positive hits were identified soon as orthologues of the aforementioned CAMs, whereas further sequence analysis was needed for more than one thousand positive hits named "Uncharacterized protein". Analysis by domain architecture and BLASTp screening confirmed that they are neuronal CAM as well, thus confirming 100% precision on the wide dataset and suggesting the NOG motif can be used with high confidence for *in silico* annotation of deposited proteomes and protein sequences.

Intriguingly, the NOG motif is specific to the Ig2 repeat in all four L1-family CAMs, all six Contactins and in DCC/NetrinR,

whereas in Roundabout receptors ROBO1, ROBO2 and ROBO3 it can be found both in Ig2 and/or Ig4 (see Fig. 1, showing the ED architecture of human NOG-positive proteins). More specifically, some ROBO sequences are NOG positive in both Ig domains, while most of them only have either Ig2 or Ig4 positive for NOG. However, further investigation on the seemingly negative domains showed that in ROBO sequences, when a perfectly matching NOG motif is detected in Ig2, a slightly degenerated version of it is present in Ig4, and *vice versa*. Notably, Ig2 mediates attractive homo- and heterophilic interactions in the horseshoe conformation of ectodomains from the first three CAM families. ROBO receptors mediate repulsive interactions via the Ig1 domain, while [56] Ig2 and Ig4 are known to mediate homophilic interactions involved in ROBO1 dimerization [57]. Strikingly, NOG containing regions in ROBO1 crystal structure overlap with already characterized interface area [56].

The only NOG motif residues showing 100% conservation are the arginine at position 7, and the aspartic acid at the extreme C-terminus. It must be mentioned that conserved tryptophan residues are located upstream to the NOG motif. Even though they were initially considered for extending N-terminally the seed region, they were not included in the final NOG regular expression. Like the conserved cystein pair of the disulphide bond stabilizing the Ig-like fold, such tryptophans are conserved in all Ig repeats of neuronal CAMs, hence representing structural determinants of the Ig fold rather than an interaction motif. Structure inspection and comparison of multiple Ig domains from neuronal CAMs confirms the involvement of tryptophans in conserved aromatic stacks (not shown).

Starting from these findings, we decided to investigate on individual Ig sequence, their structural conservation, and a possible role for the conserved R-D pair.

3.2. Structural and sequence comparison among Ig domains from NOG positive CAM proteins

The crystal structure of a representative member of each family of NOG proteins was used for structural comparison. All but the DCC/NetrinR are human proteins: for this latter, the rat protein structure was included as the closest available ortholog. The presence of four to six Ig repeats in each CAM ED allows to compare intra-molecular conservation (i.e., among Ig repeats from the same ED) and inter-molecular conservation (Ig domains from different proteins), which were evaluated by superposition and Cα-atoms RMSD comparison. Matching sequences were also compared in terms of percentage of identity by global alignment using the

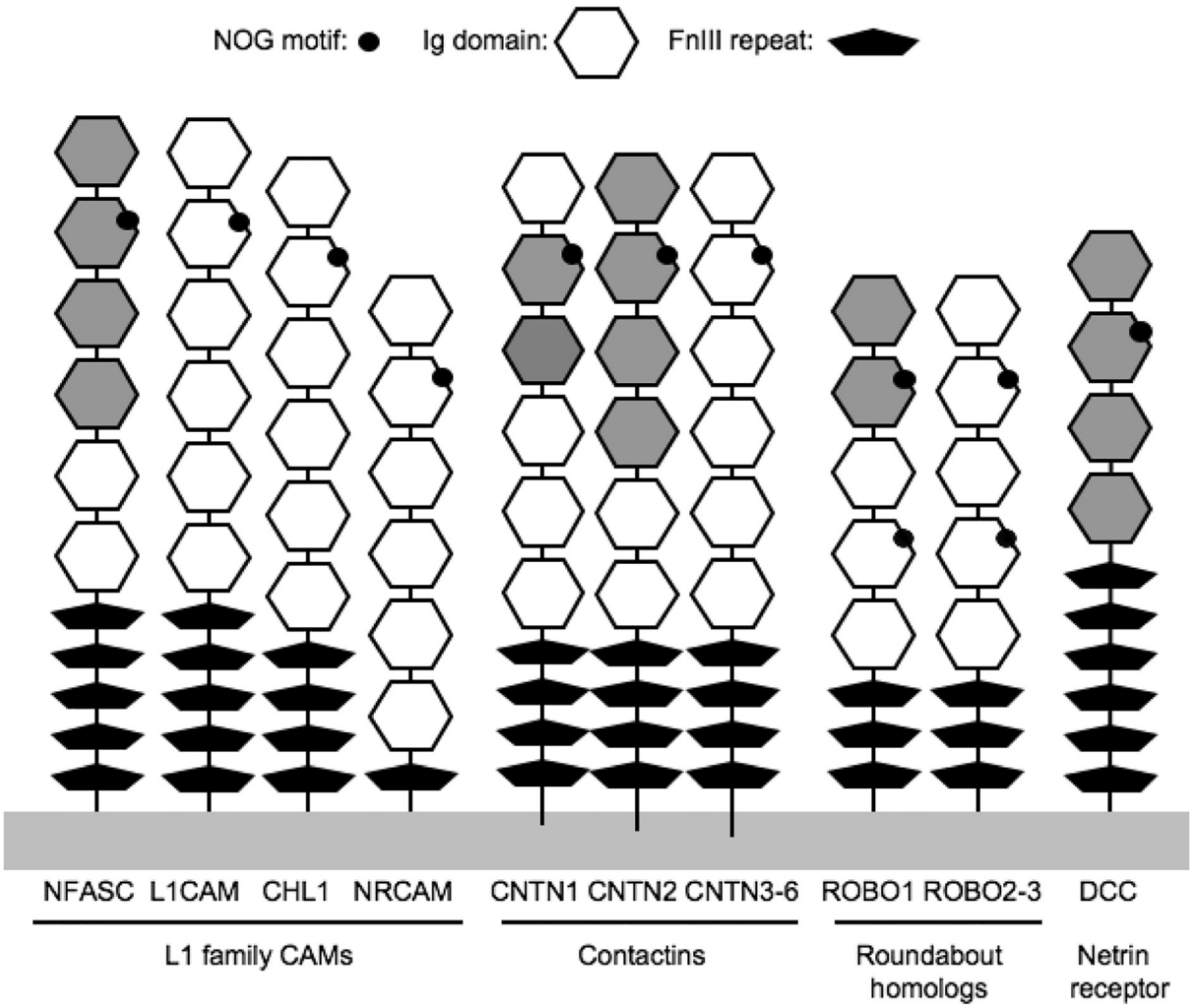

Fig. 1. Extracellular domain (ED) architecture of neuronal CAMs positive for the NOG motif. Each ED of a NOG protein protruding from the plasma membrane (grey horizontal bar) shows a variable number of Immunoglobulin-like (Ig, hexagon) and Fibronectin III-like (FnIII, black pentagon) domains. Ig domain hexagons have grey background when a corresponding X-ray structures are available, and NOG motifs are depicted by full black circles. See methods for accession numbers for structures in Protein Data Bank (PDB) and sequences in UniprotKB.

LALIGN software. Interestingly, each Ig from any CAM is closer to the corresponding one in other proteins, than to other Ig repeats from the same protein (Table 1). Ig structures are more conserved than the corresponding sequences; for instance, when Neurofascin Ig2 is superposed to any Ig2 domain from another CAM, RMSD values fall in the 1.5 Å to 3.9 Å range (average: 2.93 Å). Instead, when it is superposed to Neurofascin Ig1, Ig3 or Ig4, higher values (even > 8 Å) are found (Fig. 2). Such structural conservation is observed for all four Ig repeats, as shown by the average values in Table 1. Evidence that each Ig (1 to 4) from NOG proteins shows a peculiar structural conservation is in agreement with their role in specific intra- and inter-molecular interactions. As an example, intramolecular Ig1:Ig4 and Ig2:Ig3 interactions are needed for horseshoe formation in L1CAM-family, whereas intermolecular Ig2:Ig2 and Ig4:Ig4 interactions are crucial to ROBO dimerization [56,57].

Strikingly, when NOG regions are structurally compared, RMSD values are roughly two units lower than those of corresponding Ig domains (Table 2).

Starting from the observation that arginine (R) and aspartic acid (D) are the only two NOG-residues showing absolute conservation, and thus likely to play a pivotal role, we wondered if they were also conserved in NOG negative Ig domains of NOG positive CAMs (like the aforementioned W and C residues). However, multiple sequence alignment and structural superposition of such Ig repeats of a number of representative CAMs indicates that this is not the case (data not shown), suggesting the R-D pair is relevant (and specific) to Ig2 function rather than being a structural determinant of the Ig-like fold. Fig. 3 and Supplementary Fig. 3 compare five NOG-positive Ig from the four aforementioned crystal structures. Notably, side chains in the R-D pair are juxtaposed and closer than 4 Å (i.e., a distance compatible with the formation of a salt bridge [58]).

The biomimetic L1-A peptide was originally compared with a mutant version (R184A) reported by Zhao and co-workers [59].

Table 2
Cα-RMSD (Å) between the NOG regions of indicated CAM proteins.

	Neurofascin	Contactin1	Robo3	DCC
Neurofascin	–	0,72	0,89	0,70
Contactin1	0,72	–	0,87	0,78
Robo3	0,89	0,87	–	0,87
DCC	0,70	0,78	0,87	–

Table 1
Structural and sequence comparison among the Ig-like domains of NOG positive CAM and ECM proteins. In the upper right part of the larger table, Cα-RMSD values are shown in Å, while % identity values of corresponding sequences are reported in the lower left part. The smaller table below resumes the average RMSD and identity values for homogeneous Ig type comparison or for comparison with the other Igs from the same protein.

	Structure comparison		L1 family				Contactin family						Roundabout homologs		Netrin receptor			
Sequence comparison		(rmsd)	Neurofascin				CNTN1		CNTN2				ROBO1		DCC			
	(identity)		Ig1	Ig2	Ig3	Ig4	Ig2	Ig3	Ig1	Ig2	Ig3	Ig4	Ig1	Ig2	Ig1	Ig2	Ig3	Ig4
L1 family	Neurofascin	Ig1		8.19	6.69	4.81	10.13	4.15	7.04	8.51	3.35	8.17	4.64	6.98	3.76	4.02	2.69	4.35
		Ig2	25.9		5.33	4.03	2.17	4.67	6.32	2.65	5.04	3.38	4.43	3.15	7.92	3.92	4.65	5.14
		Ig3	29.0	23.7		4.33	5.25	3.64	7.73	5.12	2.05	3.09	3.92	4.36	5.00	4.17	3.45	4.43
		Ig4	30.5	19.8	41.4		4.76	2.74	4.38	6.23	1.73	1.97	4.46	3.28	3.24	5.84	3.57	2.31
Contactin family	CNTN1	Ig2	19.4	33.0	26.6	16.3		4.51	7.52	1.54	5.33	3.55	4.65	3.08	4.83	3.77	7.18	4.25
		Ig3	25.8	24.2	33.7	23.6	18.7		3.79	4.70	1.77	2.91	2.84	4.07	2.92	4.06	2.52	2.13
	CNTN2	Ig1	23.7	19.3	25.8	28.1	28.1	21.1		4.82	5.83	4.67	3.02	2.32	3.98	6.83	2.30	2.59
		Ig2	17.3	28.0	21.7	18.3	60.0	19.8	25.8		5.44	5.30	4.92	3.39	4.66	3.76	5.52	4.15
		Ig3	21.6	20.9	31.5	21.3	17.2	48.8	24.2	20.4		3.17	4.99	2.07	5.78	6.44	2.03	1.62
		Ig4	27.8	26.7	30.3	36.4	17.6	32.6	29.2	23.1	26.7		4.53	4.21	3.93	4.80	4.92	1.42
Roudabout homologs	ROBO1	Ig1	30.0	24.2	29.0	26.8	14.4	29.6	30.9	19.2	26.5	28.9		3.02	2.47	2.81	2.41	1.98
		Ig2	23.5	23.3	26.4	23.3	14.4	29.6	28.9	22.8	24.7	28.4	39.8		4.15	1.87	3.88	2.17
Netrin receptor	DCC	Ig1	19.4	21.1	30.6	17.9	22.7	26.8	29.2	21.9	24.0	23.2	30.6	24.7		3.52	2.36	2.19
		Ig2	28.6	25.8	31.2	23.1	20.4	31.9	29.7	26.9	26.4	30.8	29.6	26.1	34.6		2.00	2.40
		Ig3	19.2	20.4	26.7	20.9	25.5	22.2	16.3	25.0	21.1	33.3	29.0	31.5	29.5	28.8		1.22
		Ig4	24.7	20.0	31.1	30.7	31.1	29.9	30.0	23.1	29.9	34.9	30.9	30.7	28.2	32.9	40.0	

	Ig1	Ig2	Ig3	Ig4
Ig1	4.15 (27.3%)			
Ig2		2.93 (28.1%)		
Ig3			2.58 (30.7%)	
Ig4				1.90 (34.0%)
other Igs from the same protein	4.61 (29.7%)	4.41 (26.7%)	4.09 (28.2%)	3.57 (30.2%)

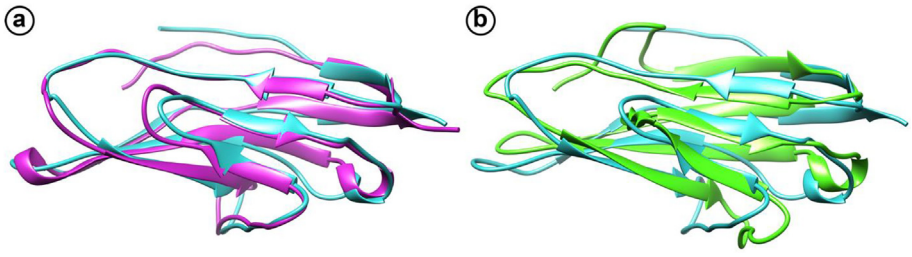

Fig. 2. Ig2 domain superposition. Neurofascin Ig2 (cyan) superimposed to (a) Contactin 1 Ig2 (magenta) or (b) Neurofascin Ig3 (green). Structure similarity is closer between Ig2 from different molecules than between different Ig domains from the same molecule. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

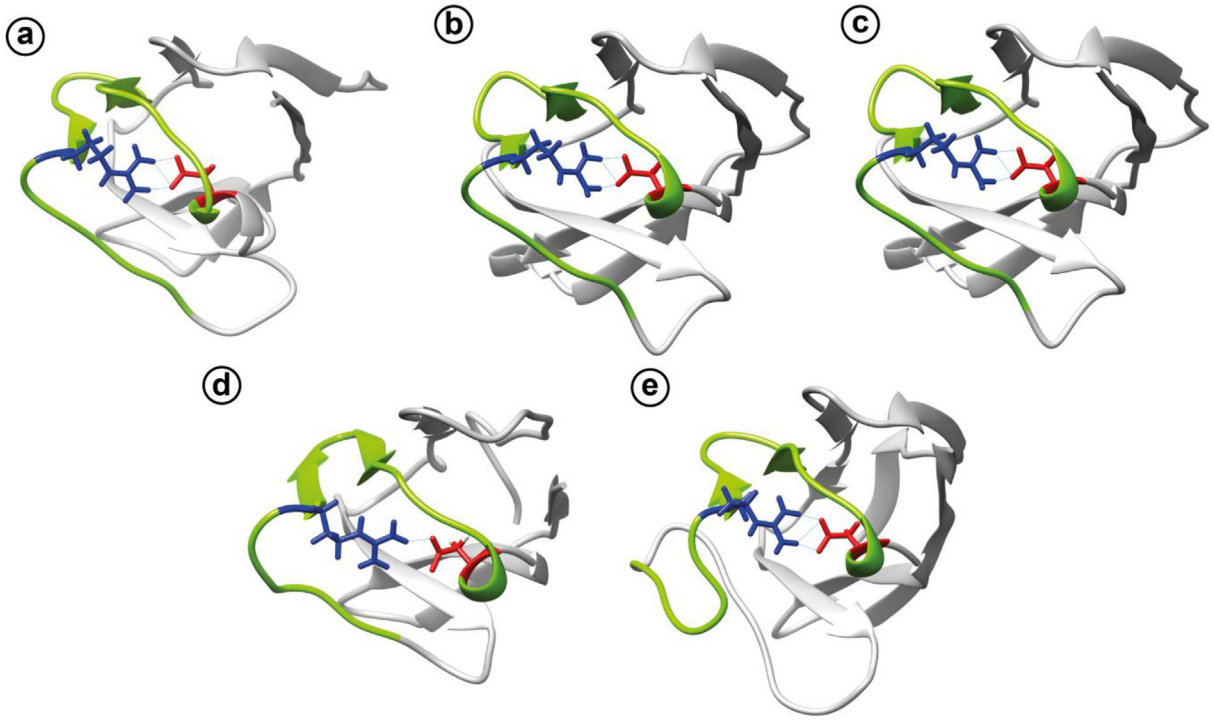

Fig. 3. Comparison of NOG positive domains. (a) Neurofascin Ig2, (b) Contactin2 Ig2, (c) DCC-Netrin receptor Ig2, (d) ROBO2 Ig2, (e) ROBO1 Ig4. The NOG region (green) is highlighted, as well as the side chains of the two residues forming the salt bridge: Arg (blue) and Asp (red). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Replacement of central arginine by alanine resulted in complete loss of neuritogenic ability. Since this amino acid corresponds to the absolutely conserved arginine at position 7 of the NOG motif, we speculated that disruption of the R-D salt bridge might lead to Ig2 misfolding, impairing its binding capacity. Hence, using the Neurofascin Ig2 crystal structure, we built *in silico* the R188A (as R188 in Neurofascin corresponds to R184 in L1CAM) and D206A mutants, which were compared to the wild type in MD simulations. Data in Fig. 4 suggest that the R-D salt bridge helps to stabilize proper folding of the Ig2 domain, as confirmed by Cα-RMSD variation in the Arg → Ala or Asp → Ala mutants. Focusing on the β-hairpin region which includes the NOG motif, both mutations impair the spatial distribution of the β-strands within the Ig2 domain; however, the Arg → Ala mutation has a stronger impact because one of the β-strands is no longer predicted to retain its secondary structure. Indeed, the overall Ig2 fold is kept, according to evidence the Ig-like fold depends on a conserved disulfide bridge, hydrophobic bonds and aromatic stacks [60]. Instead, local conformation of the binding site is impaired, in agreement with the Ig2-specific conservation of the R-D pair. Relevance of such salt bridge to proper binding features of the domain is further highlighted by changes in the surface electrostatic potential of Ig2. Isocontours in Fig. 5 and Supplementary Fig. 2 show that, at the end of the MD simulation, the Arg → Ala mutation causes a relevant charge redistribution, with local concentration of positive (blue) charges on one side (bottom in figure) of the domain, whereas negative (red) charge "spotting" occurs at another side (top in figure), also resulting in local displacement of a positive area that was present in the wild type isocontour. Noteworthy, the Asp → Ala mutation involves milder shifts in charge distribution as isocontours are mostly comparable to the wild type. Indeed, mutations not resulting in complete unfolding or severe misfolding of a domain, can anyway strongly impair or even abolish its binding capacity, when altering the proper orientation of the binding region.

3.3. Design of NOG-based peptides

As explained above, the seed region of the NOG motif was elongated to improve the pattern precision, and to this aim, Ig2-specific C-terminal conservation was preferred over less specific N-terminal residues. Even though the aforementioned R-D salt bridge is crucial for proper orientation and surface charge of the homo/heterophilic binding region in the context of a whole Ig2 domain, this is neither necessarily true, nor can be excluded *a priori*, for Ig-2 derived peptides, as their activity could mimic either a "conformational" or "linear" signaling motif. Indeed, since the L1-A peptide is disordered in solution and biomimetic *per se* in boosting neuritogenesis [10], we could argue that, at least in the instance of L1CAM, the seed region of the NOG motif is functional as a linear motif. Therefore, we wondered if the R-D bridge could be eventually relevant to the function of "conformational", NOG-derived peptides. To this aim, we performed a series of molecular dynamics unfolding simulations with the following fragments of Neurofascin Ig2: (i) peptide NFASC-B (aa 182–208), i.e. the complete NOG motif plus two additional residues at the C-terminus and (ii) NFASC-C (aa 161–216), i.e. Ig2 subregion in between the two Cys residues involved in the disulfide bridge.

Data in Supplementary Fig. 3 suggest NFASC-B is unstable in solution, as simulation RMSD values swiftly shift around 1 Å (within 100 ns) and then stabilize around 0.75 Å. Most importantly, the R-D salt bridge is predicted to be rapidly lost, probably due to the much higher dielectric constant of water. However, even if NFASC-B undergoes extensive conformational variation, it is not predicted to unfold completely, as the β-hairpin, albeit distorted, can partially stabilize the main chain fold. Conversely, NFASC-C, which includes both the salt and the disulfide bridges, appears to be much more stable, as RMSD values rapidly stabilize around 0.5 Å. Notably, the salt bridge is lost even in this simulation, as it is not buried anymore inside the Ig2. Peptide relative stability

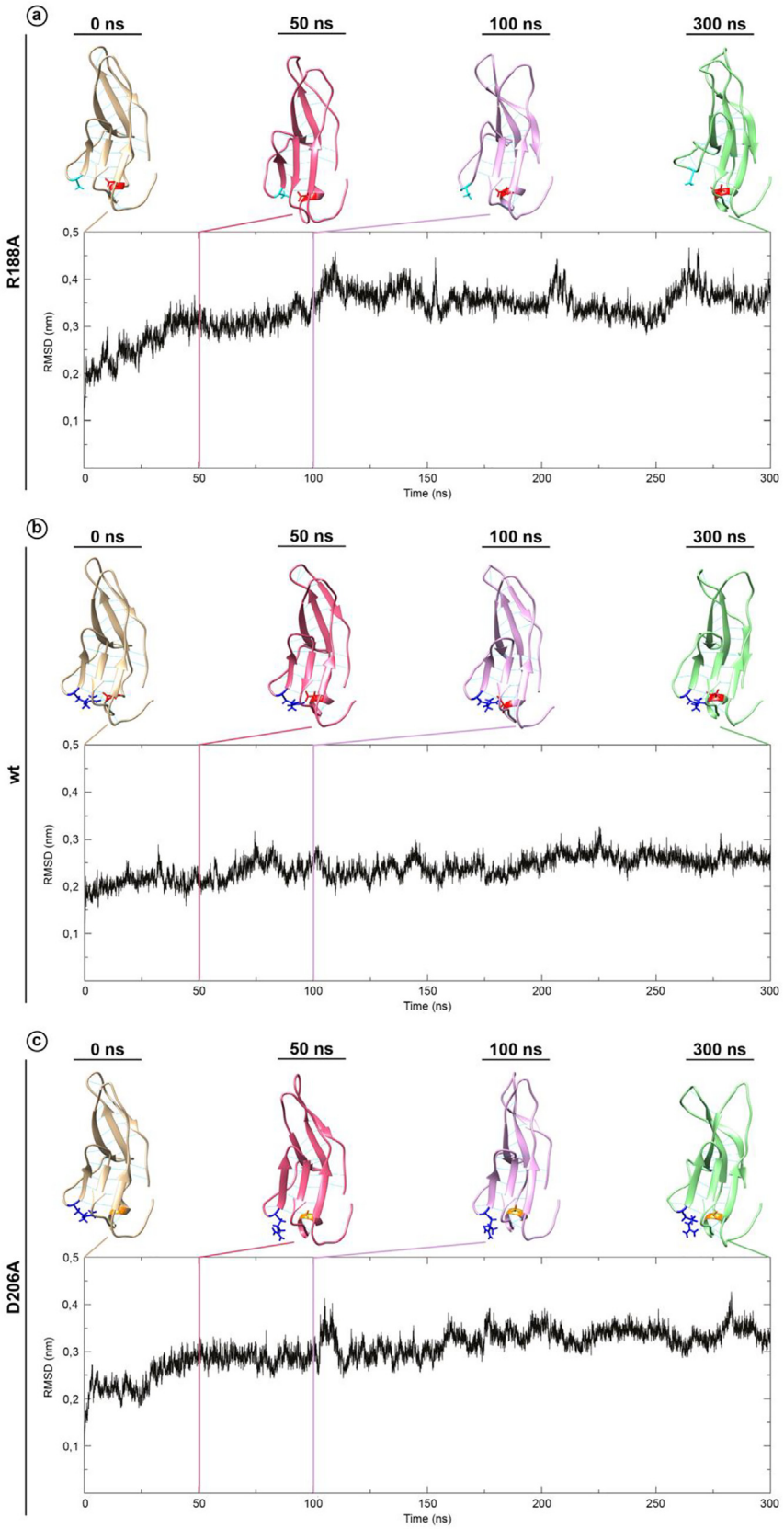

Fig. 4. Molecular dynamics simulation of wild type Neurofascin and mutants R188A and D206A. Conformation of (a) R188A mutant, (b) wild type and (c) D206A mutant of the Ig2 domain in human Neurofascin at different time points during the simulation. Plots of the Cα-RMSD calculated during the simulation are highlighted.

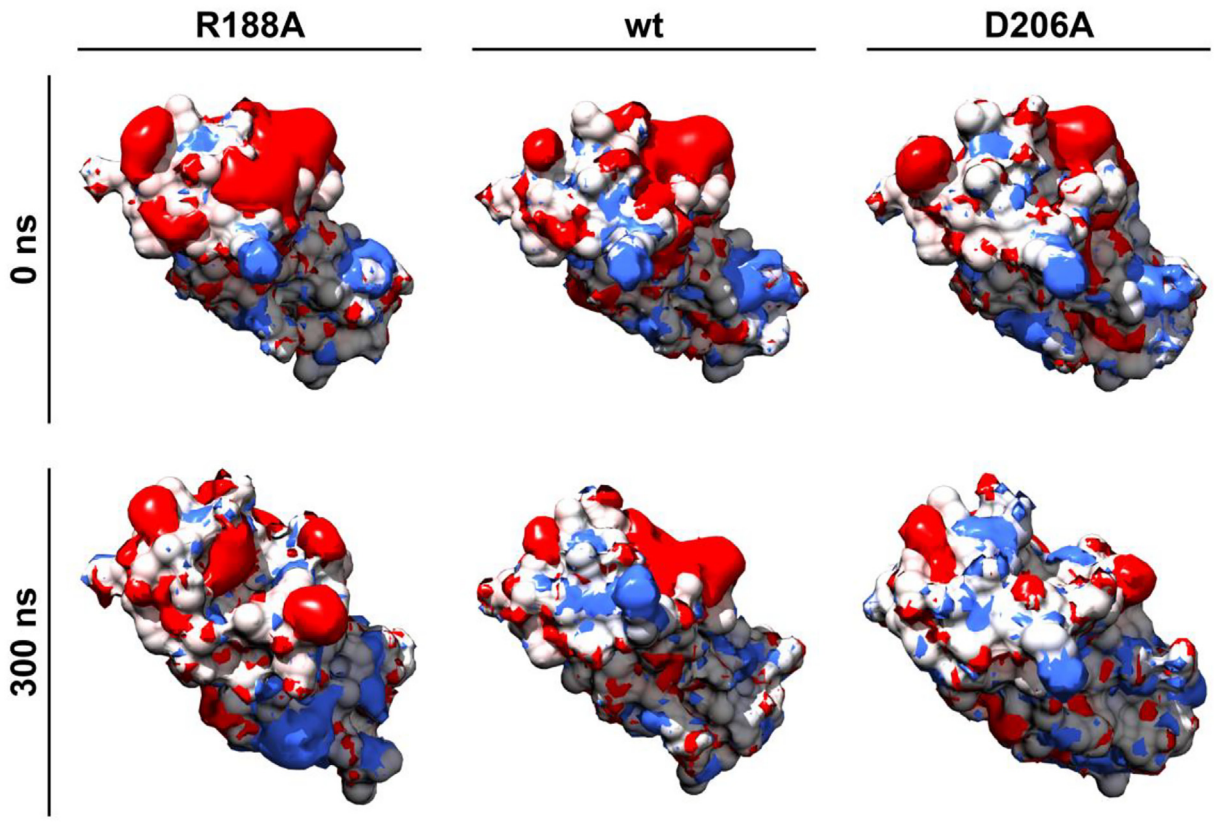

Fig. 5. Electrostatic isocontours of Neurofascin Ig2 at the beginning and the end of the MD simulation. Surface potential of the wt Ig2 remains roughly constant during the simulation, whereas R188A mutant displays an increase in positive charge (blue, bottom half of the domain) and a separation of the negative charge (red) occurs at the top half. At the same time, dispersion of the positive charge results in the disappearance of a negative area visible on the top half of the wt. Variation in surface charge, even if at a lesser extent, also concern the D206A mutant. For space constraints, only isocontours obtained at 150 mM concentration with 90° orientation are presented here, while complete information (isocontours at both 0 mM and 150 mM concentration, with 0°, 90°, 180° and 270° orientations) is reported in Supplementary Fig. 2. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3
Synthetic peptides used in this work. Sequence type: wt, wild-type, scr, scrambled. Notes: (*) neuronal isoform; (**) amino acids 1–6 from L1-A, conserved R, amino acids 8–14 CT from CNTN2-A; (***) amino acids 1–6 from CNTN2-A, conserved R, amino acids 8–14 CT from L1-A.

Peptide name	Human protein	UniProt AC	Protein fragment	Sequence type	Sequence
L1-A (*)	L1CAM	P32004	178–191	wt	HIKQDERVTMGQNG
L1-A_scr	L1CAM	P32004	178–191	scr	IVDQGNREMGTKHQ
$L1_{NT}$_$CNTN2_{CT}$ (**)	L1CAM + Contactin 2	P32004 + Q02246	178–184 + 183–190	chimeric	HIKQDERHFVSQTTG
$CNTN2_{NT}$_$L1_{CT}$ (***)	Contactin 2 + L1CAM	Q02246 + P32004	176–182 + 185–191	chimeric	FIPTDGRVTMGQNG
NFASC-A	Neurofascin	O94856	182–195	wt	PITQDKRVSQGHNG
CHL1-A	CHL1	O00533	173–186	wt	HIEQDERVYMSQKG
NrCAM-A	NrCAM/Bravo	Q92823	187–200	wt	RLPQSERVSQGLNG
DCC-A	DCC/Netrin receptor	P43146	182–195	wt	PIPGDSRVVVLPSG
CNTN1-A	Contactin 1	Q12860	180–194	wt	FITMDKRRFVSQTNG
CNTN2-A	Contactin 2	Q02246	176–190	wt	FIPTDGRHFVSQTTG
CNTN5-A	Contactin 5	Q94779	238–252	wt	FVAEDSRRFISQETG
ROBO2-A	Roundabout receptor 2	Q9HCK4	174–187	wt	IDDKEERISIRGGK
ROBO3-A	Roundabout receptor 3	Q96MS0	207–220	wt	LKEEEGRITIRGGK

can also be inferred from Root Mean Square Fluctuation values in Supplementary Fig. 3**c**: NFASC-B residues display values much higher (up to two-fold) than the same region of NFASC-C or the whole Neurofascin Ig2, which was used as reference. Taken together, data indicate if NOG-derived peptides acted in a "conformation-based" fashion, >50 residues of Ig2 would be required to stabilize their structure. Increasing the number of residues would be inconvenient, as it would impair benefits of using small peptides as biotechnological tools [9]. Since the already proven pro-neuritogenic activity of L1-A might be limited to this L1CAM derived peptide, we decided to check whether other peptides derived from the seed region of NOG-containing proteins could influence neuritogenesis as well.

3.4. Structure-activity relationships (SAR) studies of NOG-based peptides

Nine peptides were thus derived from representative members of the L1, contactin, ROBO and DCC families (Table 3). These peptides were synthesized by the solid-phase Fmoc-chemistry, puri-

fied to homogeneity (>98%) by semi-preparative RP-HPLC, and chemically characterized by high-resolution mass spectrometry (MS) as reported in the methods section (see also Supplementary Fig. 4). The conformational properties of the synthetic peptides were investigated by circular dichroism (CD) in the far-UV region, showing a minimum at about 200 nm and another shallow negative band at 220 nm, indicative of a mainly disordered conformation (Supplementary Fig. 5).

Peptide resistance to proteolytic degradation, possibly occurring during biological assays, was assessed by running control experiments under experimental conditions identical to those used for neuritogenesis experiments [10]. In order to have a better signal to noise ratio, 20 μM of each synthetic peptide was first added in separate experiments to the differentiation medium and then high-MW molecules were eliminated by ultrafiltration, while the peptides could be reliably identified and quantified by RP-HPLC and MS (not shown). A single peptide, representative of each family (L1-A, CNTN5-A, ROBO2-A and DCC-A), was used in degradation cellular tests, together with L1-A_scr as NOG-negative control. Medium was refreshed, and peptides were added at 1 μM concentration every two days. Ultrafiltrated media were analyzed after 0, 8, 24, and 48 h, with a final recovery of residual intact peptides from 60 to 85% (average yield: 74%), after 48-h incubation (Fig. 6**a**). Since peptides are added to cells at the optimal concentration of 1 μM, and the final concentration around 0.74 μM is known to be almost as neuritogenic as the optimal one [10], peptide-mediated stimulation is expected to be roughly constant during treatment.

The neuritogenic potential of the synthetic peptides was thus compared to that of the already characterized L1-A and its scrambled version L1-A_scr. Experiments were performed on human neuroblastoma derived SH-SY5Y cells (an established model for dopaminergic neurons [61]), evaluating the total neurite length per number of cells in the presence or absence of all-*trans*-retinoic acid (RA). Noteworthy, all nine new NOG-derived peptides proved to boost neuritogenesis (Fig. 6**bc** and Supplementary Fig. 6), with a potency comparable or even higher than that determined for L1-A, and, noteworthy, without significantly affecting cell proliferation (Fig. 6**d**). Capacity of all such peptides to boost outgrowth and elongation of neurite-like cellular processes proved to be reliable overtime, in several biological replicates, however, we aimed to go deeper into the nature of such cellular processes. SH-SY5Y cells are known to express high levels of MAP2, a marker for nascent neurites [62]. To further confirm the pro-neuritogenic activity of the peptides immunofluorescence against MAP2 was performed on cells treated with L1-A as reference peptide. Data presented in Supplementary Fig. 7 indicate the developed cell processes are in fact neurites.

Since the NOG motif belongs to the neuronal CAMs homo- and heterophilic binding region, we investigated on the effects of NOG-derived peptides on homophilic binding using a commercially available recombinant L1CAM ectodomain (L1CAM-ED) and surface plasmon resonance (SPR) technology. Notably, this L1CAM-ED is expressed in mammalian cells, and corresponds to amino acids 1–1120 of the human protein (~125 kDa), with a 6xHis tag at the C-terminus. We determined the equilibrium dissociation constant (K_d) of the L1CAM-ED dimer by first immobilizing L1CAM-ED on a CM5 sensor chip and then injecting incremental concentrations of the same receptor solutions in the mobile phase, to yield a K_d of 17.0 ± 1.3 nM (Fig. 7**ab**), which is six-fold lower (i.e. higher affinity) than that estimated by Gouveia and colleagues [63] using a recombinant L1CAM-ED expressed in insect cells, with a difference in the free energy change of binding ($\Delta\Delta G_b$ = 1.1 kcal/mol) only two-fold higher than the internal energy of a solution at 25 °C, i.e. ~ 0.6 kcal/mol. This moderate difference in binding strength likely reflects the difference in (i) glycosylation processing that recombinant L1CAM-ED undergoes in the two expression systems, where mammalian and insect cells yield recombinant proteins with high and low mannose content, respectively, and in (ii) the experimental settings and instrumentation used for the two sets of analyses. Thereafter, we attempted to estimate the affinity of L1-A and other synthetic peptides in Table 3 to immobilized L1CAM-ED by SPR. However, the % mass increment expected for a 1:1 peptide-L1CAM-ED interaction (1%, approximately) is too low to yield a meaningful signal and, indeed, we could not observe any significant signal change (for more details, see Biacore X-100 user's manual). Therefore, we performed competition experiments by incubating soluble L1CAM-ED (50 nM) with increasing concentrations (1–100 μM) of each synthetic peptide, and then injected this solution to immobilized L1CAM-ED. Even though the affinity of L1-A and its synthetic analogues for L1CAM-ED is not known, we assume that at the highest competitor peptide concentrations explored (i.e., 100 μM), the homophilic binding site on the Ig2 domain of L1CAM-ED is fully saturated. A NOG-derived peptide representative for each neuronal CAM family was analysed, and the decrease of SPR signal, compared to L1CAM-ED alone, was taken as an indication that the peptide could productively interact with L1CAM-ED and impair homophilic Ig2-Ig2 interaction. Data in Fig. 7**c** show that L1-A and other NOG-motif based peptides (i.e. L1-A, CNTN2-A, DCC-A, e ROBO3-A) can reduce SPR signal (expressed as maximal response units, RU) of L1CAM-ED binding only at high concentrations (50–100 μM). Among tested peptides, L1-A showed the strongest blocking effect, as it reduces up to 40% the RU signal, in a dose dependent manner. This effect was taken as an indication that the peptide in solution can efficiently compete with the Ig2-Ig2 homophilic interaction. From the neuritogenic function of L1-A peptides, shown in Fig. 6, we would have expected a much higher displacing effect in SPR competition experiments. However, such a hypothesis is valid only when assuming ED-ED interaction to be mediated by Ig2 alone, and this is not the case. Conversely, the high baseline of residual ED-ED binding hence the low-moderate displacing effect exerted by L1-A peptide is in agreement with published evidence. Indeed, homophilic L1CAM-ED binding is mediated not only by Ig2, as other domains are involved [63,64] which can even assume a variable conformation within the L1CAM molecule [65]. Therefore, a peptide competitor such as L1-A might remarkably affect function without dramatically impairing ED-ED interaction. Intriguingly, the addition of incremental concentrations of CNTN2, DCC and ROBO3, only partially (i.e., 20%) reduced the RU_{max} of L1CAM-L1CAM interaction. Then, we compared the displacing capacity of the wild type L1-A peptide with its R184A and scrambled mutants, which lack neuritogenic capacity. It seems that the ability of the mutant peptides to interfere with the homophilic interaction is comparable with that of the other peptides tested (Fig. 7**d**). This is a rather interesting result that can be explained considering that ED-ED interactions are mediated by further L1CAM domains [63,64], i.e. not only by the Ig2 domain for which L1-A peptide and analogues have affinity, according to evidence that in some instances, homo- and/or heterophilic binding among NOG-containing proteins is mediated by horseshoe conformation and, in others, by an open one [65]. Notably, the same trend was observed performing the competition assays using a lower concentration of L1CAM-ED (i.e., 15 nM) in the presence of the highest competitor peptides concentration (Supplementary Fig. 8).

3.5. Non-natural sequence peptides fitting the NOG motif retain the biomimetic potential

Evidence gathered so far demonstrates that peptides from NOG positive CAMs display neuritogenic activity independently from the CAM family they are derived from. However, all such peptides,

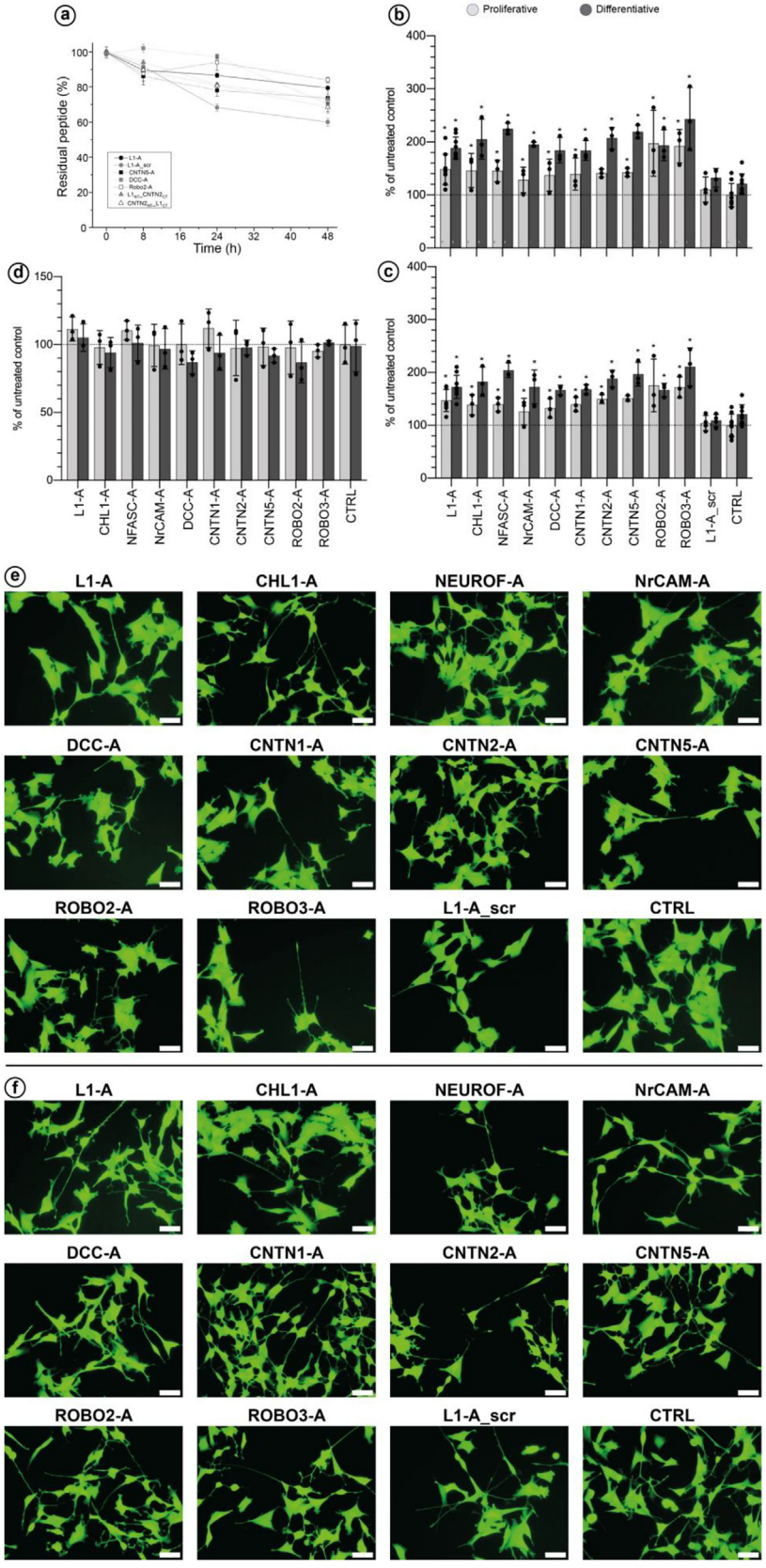

Fig. 6. Chemical stability and neuritogenic effect of synthetic peptides derived from the NOG motif of indicated neuronal CAMs. (a) Chemical stability of NOG peptides in SH-SY5Y cells differentiating culture medium. (b) Total neurite length and (c) neurites / number of cells of SH-SY5Y cells treated with indicated NOG peptides, * indicates $p < 0.05$ respect to sample without peptide addition (CTRL). (d) Cell proliferation 2 days after treatment with NOG peptides. All data represent the mean ± SD of at least three independent experiments; sample size is indicated at the bottom of each column, whereas the orizontal line represents the untreated CTRL reference value. (e) Micrograph of proliferating SH-SY5Y treated cells. (f) Micrograph of RA-differentiated SH-SY5Y treated cells. Images are taken with a Leica DMI4000 at 20X magnification, scalebar is 50 μM.

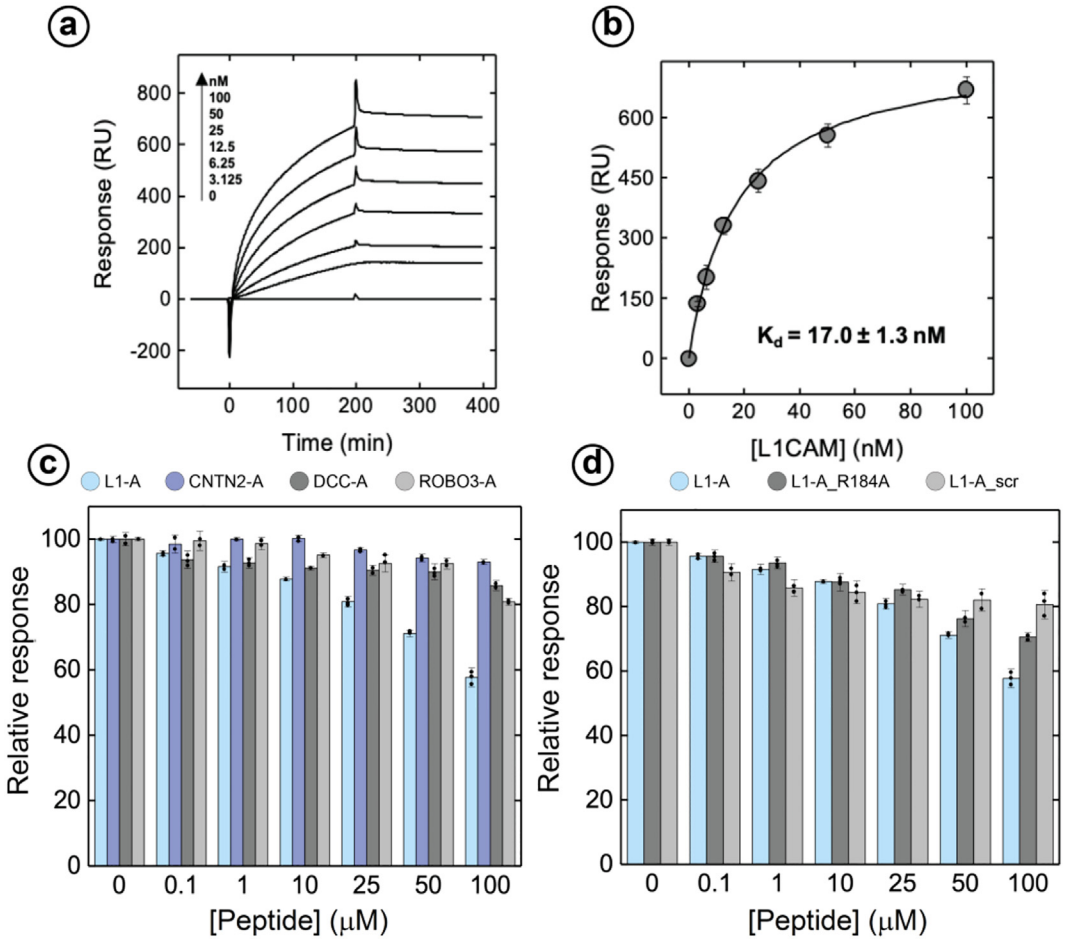

Fig. 7. Surface plasmon resonance analysis of L1CAM-ED - L1CAM-ED interaction. (a) Increasing concentration of L1CAM-ED were sequentially injected over the sensor chip, at a flow rate of 10 μl/min at 25 °C, using HBSEP + as running buffer. Each SPR trace was subtracted for unspecific binding (i.e.,2% of RUmax). (b) The response units (RU) at the steady state were plotted as a function of [L1CAM] and fitted to the Langmuir equation (equation 1) to yield the dissociation constant K_d. (c-d) Competition experiment in the presence of L1A peptide and analogues. A 50 nM solution of L1CAM-ED was incubated with increasing concentrations of peptides (0 to 100 μM) for at least 15 min before injection over the same L1CAM-ED-coated sensor chip. Results are shown as the maximal association RU (expressed as the percentage relative to the response measured without peptide) achieved at increasing concentrations of competitor, as indicated. All data represent the mean ± SD of three independent experiments.

in addition to fit the N-terminal, seed sequence of the NOG motif, are "sequence based", in that perfectly reproducing the sequences of the corresponding proteins. Therefore, we wondered if "motif based" peptides, i.e., still fitting the NOG motif (in terms of being compatible with the PROSITE-like regular expression) while not corresponding to the sequence of any specific CAM, could be effective as well. To this aim, two chimeric peptides between L1-A and CNTN2-A were designed around the 100% conserved Arg at position 7 of the NOG motif. In chimeric peptide $L1_{NT}$_$CNTN2_{CT}$, the six N-terminal residues are derived from L1-A and the seven C-terminal ones from CNTN2-A, while in chimeric peptide $CNTN2_{NT}$_$L1_{CT}$, just the reverse occurs. This way, both peptides retain full sequence compatibility with the NOG regular expression (as all amino acids of each peptide are still listed among those accepted at each specific position), while showing instead a recombinant sequence, which no longer corresponds to a real NOG sequence from any specific CAM. The resistance of both peptides to proteolysis was found not meaningfully different from that of other peptides (not shown). In neuritogenesis assays, the two chimeras proved to have a pro-neuritogenic effect comparable to that of parent peptide L1-A, both in proliferative and differentiative conditions (Fig. 8**abc** and Supplementary Fig. 9), whereas SPR data indicate that their competitive effects are closer to parent peptide CNTN2-A than to L1-A (Fig. 8**d**).

3.6. Docking simulations

To gain insights on the binding of NOG-derived peptides to NOG-containing proteins, reducing possible biases of algorithms sampling engines, we performed flexible docking simulations with two methods: template independent CABSDock [37] and template based Galaxy Pep Dock [36]. Even though docking simulations are often error prone and hence are not presented in detail, results we obtained here are in general agreement with a model in which the overall NOG motif is crucial for binding, but the absence of the absolutely conserved arginine makes this binding unable to trigger the neuritogenic signal. Indeed, during simulations with CABSdock (Supplementary Figs. 10 and 11), a similar poses enrichment, spatially close to Ig2, is shared by neuritogenic peptides and non-neuritogenic peptide L1-A_R184A, which also shares with neuritogenic ones the binding capacity in SPR experiments. Instead, L1-A_scr, a peptide missing both binding and neuritogenic capacities, shows more dispersed poses, all around the horseshoe. Simulation runs with GalaxyPepDock (not shown) did not contradict this gen-

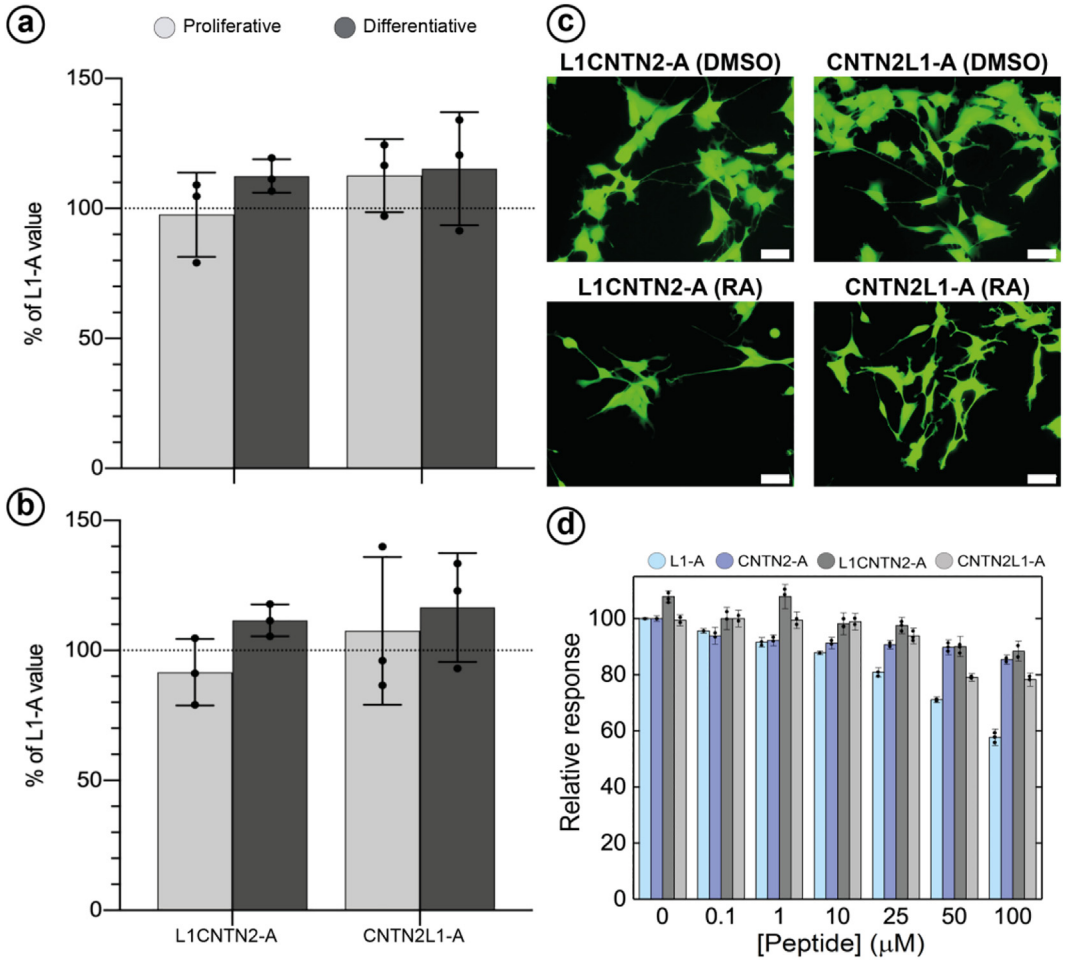

Fig. 8. Neuritogenic potential of chimeric peptides. (a)Total neurite length and (b) neurites / number of cells of SH-SY5Y cells treated with chimeric peptides. Data are normalized for the corresponding effect of the L1-A peptide. All data represent the mean ± SD of at least three independent experiments; sample size is indicated at the bottom of each column, the orizontal line represents the L1-A reference value. (c) Micrograph of SH-SY5Y treated cells. Images are taken with a Leica DMI4000 at 20X magnification, scalebar is 50 μM. (d) Competition experiment carried out in the presence of L1-A peptide and its analogues. A 50 nM solution of L1CAM-ED was incubated with different concentrations of peptides (0 to 100 μM) for at least 15 min before injection over the same L1CAM-ED-coated sensor chip. Results are shown as the maximal association RU (expressed as the percentage relative to the response measured without peptide) achieved at increasing concentrations of competitor.

eral model, which mandatorily will need further experimental work to be exhaustively validated.

3.7. Concluding remarks

The newly discovered NOG motif is specific to the neuronal CAMs and it is widely conserved in Metazoa and especially in Vertebrata, with most positive hits from Mammals.

It is already written in standard PROSITE syntax and shows 100% Precision Index, making it suitable for deposit into the PROSITE database. Since its 100% precision is confirmed even in the much larger dataset of the whole UniProtKB database, the NOG pattern is suitable for use with high confidence for improving UniProtKB and other databases in genomes/proteomes annotation programmes. The NOG motif is part of an interaction region by which the Ig2 repeats from neuronal CAM horseshoes mediate homo- and heterophilic interactions, which in turn are crucial to proper neural development.

Indeed, the formation of neural circuits in the nervous system requires that a correct connectivity is established between neurons during development. To this aim, neural precursors need to migrate and properly locate their cell bodies for projecting their axons to synaptic targets. Failure to achieve correct connectivity results in a dysfunctional nervous system. Every neural circuit is made of axons and dendrites, with individual axons stimulating multiple targets, and single dendrites assimilating input signals. During development, neuronal CAMs with immunoglobulin (Ig) repeats provide defined 'roadway' surfaces to which receptors of the growth cone can adhere, hence playing a fundamental role in cell–cell and cell-ECM interactions in both mature and developing nervous system. The NOG motif, put forward in this study, is a key element in the homo/heterophilic interactions among neuronal CAMs and its identification opens the route to further investigations and will help shedding more light on the complex signalling and regulatory system underlying neuritogenesis and axon guidance.

From a biotechnological point of view, NOG motif derived synthetic peptides lay the basis for the creation of a biomimetics toolbox for regenerative medicine and neurodevelopment experiments, and evidence that non-natural sequences fitting the NOG motif can retain the neuritogenic potential opens the route to design and tuning of peptides with improved or more specific activity. Moreover, evidence that some mutant peptides (such as e.g. L1-A_R184A) retain their binding activity, while missing the neuritogenic one, suggest for design and use of NOG inhibitors. Given the relatively low molecular weight of NOG-based peptides

and the feasibility of their chemical synthesis in high amounts and elevated purity, in the next future extensive and systematic structure–activity relationship studies will be undertaken by exploiting the possibility of varying peptide length and introducing unnatural and noncoded amino acids with tailored physicochemical properties [66–69].

Author's contributions

GS, MG, DP performed most biochemical, molecular, and cellular biology experiments. ST, EP, DP, and AP performed peptide synthesis and biochemical characterization. LA performed SPR analysis. MG and IR performed bioinformatic analysis. VDF coordinated the biochemistry tasks and co-supervised the project. FF conceived and coordinated the project. GS, MG, VDF and FF jointly wrote the manuscript.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by funding from the Padua University (Project PRAT2015 CPDA151948 to FF, PRID Junior 2019 Project to LA and PhD fellowship to GS) and the CaRiPaRo foundation (Excellence Research Project-2018 "BiPTA" to VDF and PhD fellowship to MG). We thank Tommaso De Re for help with ScanProsite scannings, Sara Schiavon for technical help with microscopy, Marika Tonellato for help with immunofluorescence, Davide Righetti for help with data analysis, Elisabetta Bergantino, Roberto Battistutta, Alessandro Grinzato, and Yuriko Suemi Hernandez Gomez for helpful discussions and Elisa Greggio for help with cell cultures.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.csbj.2021.10.005.

References

[1] SIB Swiss Institute of Bioinformatics Members. The SIB Swiss Institute of Bioinformatics' resources: focus on curated databases. Nucleic Acids Res 2016;44:D27–37. https://doi.org/10.1093/nar/gkv1310.
[2] Yang M, Derbyshire MK, Yamashita RA, Marchler-Bauer A. NCBI's conserved domain database and tools for protein domain analysis. Curr Protoc Bioinform 2020;69:. https://doi.org/10.1002/cpbi.90e90.
[3] Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, Sonnhammer ELL, et al. Pfam: The protein families database in 2021. Nucleic Acids Res 2021;49: D412–9. https://doi.org/10.1093/nar/gkaa913.
[4] McDowall J, Hunter S. InterPro protein classification. Methods Mol Biol 2011;694:37–47. https://doi.org/10.1007/978-1-60761-977-2_3.
[5] Ribrioux S, Brüngger A, Baumgarten B, Seuwen K, John MR. Bioinformatics prediction of overlapping frameshifted translation products in mammalian transcripts. BMC Genomics 2008;9(1):122. https://doi.org/10.1186/1471-2164-9-122.
[6] Tai C-H, Sam V, Gibrat J-F, Garnier J, Munson PJ, Lee B. Protein domain assignment from the recurrence of locally similar structures. Proteins Struct Funct Bioinform 2011;79:853–66. https://doi.org/10.1002/prot.22923.
[7] Badaczewska-Dawid AE, Kolinski A, Kmiecik S. Computational reconstruction of atomistic protein structures from coarse-grained models. Comput Struct Biotechnol J 2020;18:162–76. https://doi.org/10.1016/j.csbj.2019.12.007.
[8] Adiyaman R, McGuffin LJ. Methods for the Refinement of Protein Structure 3D Models. Int J Mol Sci 2019;20. https://doi.org/10.3390/ijms20092301.
[9] Chen H, Yuan L, Song W, Wu Z, Li D. Biocompatible polymer materials: Role of protein–surface interactions. Prog Polym Sci 2008;33:1059–87. https://doi.org/10.1016/j.progpolymsci.2008.07.006.
[10] Scapin G, Salice P, Tescari S, Menna E, De Filippis V, Filippini F. Enhanced neuronal cell differentiation combining biomimetic peptides and a carbon nanotube-polymer scaffold. Nanomed Nanotechnol, Biol Med 2015;11 (3):621–32. https://doi.org/10.1016/j.nano.2014.11.001.
[11] Maness PF, Schachner M. Neural recognition molecules of the immunoglobulin superfamily: signaling transducers of axon guidance and neuronal migration. Nat Neurosci 2007;10(1):19–26. https://doi.org/10.1038/nn1827.
[12] Colombo F, Meldolesi J. L1-CAM and N-CAM: from adhesion proteins to pharmacological targets. Trends Pharmacol Sci 2015;36(11):769–81. https://doi.org/10.1016/j.tips.2015.08.004.
[13] Zhang Y, Yeh J, Richardson PM, Bo X. Cell adhesion molecules of the immunoglobulin superfamily in axonal regeneration and neural repair. Restor Neurol Neurosci 2008;26:81–96.
[14] Liu H, Focia PJ, He X. Homophilic adhesion mechanism of neurofascin, a member of the L1 family of neural cell adhesion molecules. J Biol Chem 2011;286(1):797–805. https://doi.org/10.1074/jbc.M110.180281.
[15] Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 1997;25:3389–402. https://doi.org/10.1093/nar/25.17.3389.
[16] Sigrist CJA, Cerutti L, de Castro E, Langendijk-Genevaux PS, Bulliard V, Bairoch A, et al. PROSITE, a protein domain database for functional characterization and annotation. Nucleic Acids Res 2009;38:D161–6. https://doi.org/10.1093/nar/gkp885.
[17] de Castro E, Sigrist CJA, Gattiker A, Bulliard V, Langendijk-Genevaux PS, Gasteiger E, et al. ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. Nucleic Acids Res 2006;34(Web Server):W362–5. https://doi.org/10.1093/nar/gkl124.
[18] Boutet E, Lieberherr D, Tognolli M, Schneider M, Bansal P, Bridge AJ, et al. Uniprotkb/swiss-prot, the manually annotated section of the uniprot knowledgebase: How to use the entry view. Methods Mol Biol., vol. 1374, Humana Press Inc.; 2016, p. 23–54. https://doi.org/10.1007/978-1-4939-3167-5_2.
[19] Huang X, Miller W. A time-efficient, linear-space local similarity algorithm. Adv Appl Math 1991;12(3):337–57. https://doi.org/10.1016/0196-8858(91)90017-D.
[20] Lamprianou S, Chatzopoulou E, Thomas J-L, Bouyain S, Harroch S. A complex between contactin-1 and the protein tyrosine phosphatase PTPRZ controls the development of oligodendrocyte precursor cells. Proc Natl Acad Sci U S A 2011;108(42):17498–503. https://doi.org/10.1073/pnas.1108774108.
[21] Mörtl M, Sonderegger P, Diederichs K, Welte W. The crystal structure of the ligand-binding module of human TAG-1 suggests a new mode of homophilic interaction. Protein Sci 2007;16(10):2174–83. https://doi.org/10.1110/(ISSN)1469-896X10.1110/ps.072802707.
[22] Morlot C, Hemrika W, Romijn RA, Gros P, Cusack S, McCarthy AA. Cloning, expression, crystallization and preliminary X-ray analysis of the first two Ig domains from human roundabout 1 (Robo1). Acta Crystallogr Sect F, Struct Biol Cryst Commun 2007;63:689–91. https://doi.org/10.1107/S1744309107033027.
[23] Chen Q, Sun X, Zhou X, Liu J, Wu J, Zhang Y, et al. N-terminal horseshoe conformation of DCC is functionally required for axon guidance and might be shared by other neural receptors. J Cell Sci 2013;126:186–95. https://doi.org/10.1242/jcs.111278.
[24] Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF chimera – A visualization system for exploratory research and analysis. J Comput Chem 2004;25(13):1605–12. https://doi.org/10.1002/jcc.v25:1310.1002/jcc.20084.
[25] Huang CC, Meng EC, Morris JH, Pettersen EF, Ferrin TE. Enhancing UCSF Chimera through web services. Nucleic Acids Res 2014;42:W478–84. https://doi.org/10.1093/nar/gku377.
[26] Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, et al. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Res 2014;42:W252–8. https://doi.org/10.1093/nar/gku340.
[27] Canutescu AA, Shelenkov AA, Dunbrack RL. A graph-theory algorithm for rapid protein side-chain prediction. Protein Sci 2003;12(9):2001–14. https://doi.org/10.1110/(ISSN)1469-896X10.1110/ps.03154503.
[28] Wang Q, Canutescu AA, Dunbrack RL. SCWRL and MolIDE: Computer programs for side-chain conformation prediction and homology modeling. Nat Protoc 2008;3(12):1832–47. https://doi.org/10.1038/nprot.2008.184.
[29] Benkert P, Biasini M, Schwede T. Toward the estimation of the absolute quality of individual protein structure models. Bioinformatics 2010;27:343–50. https://doi.org/10.1093/bioinformatics/btq662.
[30] Smith N, Witham S, Sarkar S, Zhang J, Li L, Li C, et al. DelPhi web server v2: incorporating atomic-style geometrical figures into the computational protocol. Bioinformatics 2012;28:1655–7. https://doi.org/10.1093/bioinformatics/bts200.
[31] Sarkar S, Witham S, Zhang J, Zhenirovskyy M, Rocchia W, Alexov E. DelPhi Web Server: A comprehensive online suite for electrostatic calculations of biological macromolecules and their complexes. Commun Comput Phys 2013;13 (1):269–84. https://doi.org/10.4208/cicp.300611.201011s.
[32] Sitkoff D, Sharp KA, Honig B. Accurate calculation of hydration free energies using macroscopic solvent models. J Phys Chem 1994;98(7):1978–88. https://doi.org/10.1021/j100058a043.
[33] Guo T, Gong L-C, Sui S-F, Laggner P. An electrostatically preferred lateral orientation of SNARE complex suggests novel mechanisms for driving membrane fusion. PLoS ONE 2010;5(1):e8900.
[34] Schutz CN, Warshel A. What are the dielectric "constants" of proteins and how to validate electrostatic models? Proteins 2001;44(4):400–17. https://doi.org/10.1002/prot.1106.

[35] Gorham RD, Kieslich CA, Morikis D. Electrostatic clustering and free energy calculations provide a foundation for protein design and optimization. Ann Biomed Eng 2011;39(4):1252–63. https://doi.org/10.1007/s10439-010-0226-9.
[36] Lee H, Heo L, Lee MS, Seok C. GalaxyPepDock: A protein–peptide docking tool based on interaction similarity and energy optimization. Nucleic Acids Res 2015;43(W1):W431–5. https://doi.org/10.1093/nar/gkv495.
[37] Kurcinski M, Pawel Ciemny M, Oleniecki T, Kuriata A, Badaczewska-Dawid AE, Kolinski A, et al. CABS-dock standalone: a toolbox for flexible protein–peptide docking. Bioinformatics 2019;35:4170–2. https://doi.org/10.1093/bioinformatics/btz185.
[38] Heo L, Lee H, Seok C. GalaxyRefineComplex: Refinement of protein-protein complex model structures driven by interface repacking. Sci Rep 2016;6:1–10. https://doi.org/10.1038/srep32153.
[39] Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, et al. Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX 2015;1-2:19–25. https://doi.org/10.1016/j.softx.2015.06.001.
[40] Duan Y, Wu C, Chowdhury S, Lee MC, Xiong G, Zhang W, et al. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. J Comput Chem 2003;24(16):1999–2012. https://doi.org/10.1002/jcc.10349.
[41] Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR. Molecular-dynamics with coupling to an external bath. J Chem Phys 1984;81(8):3684–90. https://doi.org/10.1063/1.448118.
[42] Steiner LA, Andrews PJD. Monitoring the injured brain: ICP and CBF. Br J Anaesth 2006;97(1):26–38. https://doi.org/10.1093/bja/ael110.
[43] Parrinello M, Rahman A. Polymorphic transitions in single crystals: A new molecular dynamics method. J Appl Phys 1981;52(12):7182–90. https://doi.org/10.1063/1.328693.
[44] Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A smooth particle mesh Ewald method. J Chem Phys 1995;103(19):8577–93. https://doi.org/10.1063/1.470117.
[45] Atherton E. Solid phase peptide synthesis. A Pract Approach; 1989.
[46] De Filippis V, Quarzago D, Vindigni A, Di Cera E, Fontana A. Synthesis and characterization of more potent analogues of hirudin fragment 1–47 containing non-natural amino acids. Biochemistry 1998;37(39):13507–15. https://doi.org/10.1021/bi980717n.
[47] Peterle D, Pontarollo G, Spada S, Brun P, Palazzi L, Sokolov AV, et al. A serine protease secreted from Bacillus subtilis cleaves human plasma transthyretin to generate an amyloidogenic fragment. Commun Biol 2020;3(1). https://doi.org/10.1038/s42003-020-01493-0.
[48] De Filippis V, Polverino de Laureto P, Toniutti N, Fontana A. Acid-induced molten globule state of a fully active mutant of human interleukin-6. Biochemistry 1996;35(35):11503–11. https://doi.org/10.1021/bi9604587.
[49] Pontarollo G, Acquasaliente L, Peterle D, Frasson R, Artusi I, De Filippis V. Non-canonical proteolytic activation of human prothrombin by subtilisin from Bacillus subtilis may shift the procoagulant-anticoagulant equilibrium toward thrombosis. J Biol Chem 2017;292(37):15161–79. https://doi.org/10.1074/jbc.M117.795245.
[50] Acquasaliente L, Peterle D, Tescari S, Pozzi N, Pengo V, De Filippis V. Molecular mapping of α-thrombin (αT)/β2-glycoprotein I (β2GpI) interaction reveals how β2GpI affects αT functions. Biochem J 2016;473:4629–50. https://doi.org/10.1042/BCJ20160603.
[51] Ross RA, Spengler BA, Biedler JL. Coordinate morphological and biochemical interconversion of human neuroblastoma Cells2. JNCI J Natl Cancer Inst 1983;71:741–7. https://doi.org/10.1093/jnci/71.4.741.
[52] Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. Nat Methods 2012;9(7):676–82. https://doi.org/10.1038/nmeth.2019.
[53] Hu H, Ni Y, Montana V, Haddon RC, Parpura V. Chemically functionalized carbon nanotubes as substrates for neuronal growth. Nano Lett 2004;4(3):507–11. https://doi.org/10.1021/nl035193d.
[54] Munnamalai V, Weaver CJ, Weisheit CE, Venkatraman P, Agim ZS, Quinn MT, et al. Bidirectional interactions between NOX2-type NADPH oxidase and the F-actin cytoskeleton in neuronal growth cones. J Neurochem 2014;130(4):526–40. https://doi.org/10.1111/jnc.12734.
[55] Sigrist CJA, Cerutti L, Hulo N, Gattiker A, Falquet L, Pagni M, et al. PROSITE: A documented database using patterns and profiles as motif descriptors. Brief Bioinform 2002;3:265–74. https://doi.org/10.1093/bib/3.3.265.
[56] Reichert MC, Brown HE, Evans TA. In vivo functional analysis of Drosophila Robo1 immunoglobulin-like domains. Neural Dev 2016;11:15. https://doi.org/10.1186/s13064-016-0071-0.
[57] Aleksandrova N, Gutsche I, Kandiah E, Avilov SV, Petoukhov MV, Seiradake E, et al. Robo1 forms a compact dimer-of-dimers assembly. Structure 2018;26(2):320–328.e4. https://doi.org/10.1016/j.str.2017.12.003.
[58] Donald JE, Kulp DW, DeGrado WF. Salt bridges: Geometrically specific, designable interactions. Proteins Struct Funct Bioinforma 2011;79:898–915. https://doi.org/10.1002/prot.22927.
[59] Zhao X, Yip PM, Siu C-H. Identification of a homophilic binding site in immunoglobulin-like domain 2 of the cell adhesion molecule L1. J Neurochem 1998;71(3):960–71. https://doi.org/10.1046/j.1471-4159.1998.71030960.x.
[60] Wang J-H. The sequence signature of an Ig-fold. Protein Cell 2013;4(8):569–72. https://doi.org/10.1007/s13238-013-3903-2.
[61] Xie H, Hu L, Li G. SH-SY5Y human neuroblastoma cell line: in vitro cell model of dopaminergic neurons in Parkinson's disease. Chin Med J (Engl) 2010;123:1086–92.
[62] Paik S, Somvanshi RK, Kumar U. Somatostatin-mediated changes in microtubule-associated proteins and retinoic acid-induced neurite outgrowth in SH-SY5Y cells. J Mol Neurosci 2019;68(1):120–34. https://doi.org/10.1007/s12031-019-01291-2.
[63] Gouveia RM, Gomes CM, Sousa M, Alves PM, Costa J. Kinetic analysis of L1 homophilic interaction: Role of the first four immunoglobulin domains and implications on binding mechanism. J Biol Chem 2008;283(42):28038–47. https://doi.org/10.1074/jbc.M804991200.
[64] Samatov TR, Wicklein D, Tonevitsky AG. L1CAM: Cell adhesion and more. Prog Histochem Cytochem 2016;51:25–32. https://doi.org/10.1016/j.proghi.2016.05.001.
[65] Wei CH, Ryu SE. Homophilic interaction of the L1 family of cell adhesion molecules. Exp Mol Med 2012;44:413–23. https://doi.org/10.3858/emm.2012.44.7.050.
[66] De Filippis V, Acquasaliente L, Pontarollo G, Peterle D. Noncoded amino acids in protein engineering: Structure-activity relationship studies of hirudin-thrombin interaction. Biotechnol Appl Biochem 2018;65:69–80. https://doi.org/10.1002/bab.1632.
[67] De Filippis V, Pozzi N, Acquasaliente L, Artusi I, Pontarollo G, Peterle D. Protein engineering by chemical methods: Incorporation of nonnatural amino acids as a tool for studying protein folding, stability, and function. Pept Sci 2018;110(5):e24090. https://doi.org/10.1002/pep2.24090.
[68] Henninot A, Collins JC, Nuss JM. The current state of peptide drug discovery: Back to the future? J Med Chem 2018;61(4):1382–414. https://doi.org/10.1021/acs.jmedchem.7b00318.
[69] George KL, Horne WS. Foldamer tertiary structure through sequence-guided protein backbone alteration. Acc Chem Res 2018;51(5):1220–8. https://doi.org/10.1021/acs.accounts.8b00048.