# $^{1}H$, $^{13}C$ and $^{15}N$ assignments of EGF domains 8 to 11 of human Notch-1

**Author list**

Philip C. Weisshuhn, Penny A. Handford* and Christina Redfield*

*Corresponding authors:

PAH – Tel: +44 1865 613260; FAX: +44 1865 613201;

email: penny.handford@bioch.ox.ac.uk;

CR - Tel: +44 1865 275330; FAX: +44 1865 613201;

email: christina.redfield@bioch.ox.ac.uk

Department of Biochemistry, University of Oxford,

South Parks Road, Oxford OX1 3QU, UK

**Abstract**

The Notch receptor is part of a core cell-cell signaling system crucial for development and tissue homeostasis in Metazoa. Structural information is available for the negative regulatory region (NRR), the ligand-binding region and the intracellular domain (NICD) of Notch, but data for the remaining portions of the extracellular region which determine its overall shape at the cell surface are still lacking. This region consists of 36 EGF-like domains arranged as multiple tandem repeats. Most EGF-like domains near the ligand binding domains EGF 11 and 12 are of the calcium-binding type, with well-described, rigid and near-linear interdomain interfaces. However, EGF10 is a conserved, non-calcium-binding domain which may confer flexibility or a non-linear organization to the receptor. To probe this, we have expressed and purified a four-domain construct, EGF8-11, from human Notch-1, and report here the $^{1}H$, $^{13}C$ and $^{15}N$ resonance assignments. Differences in EGF11 chemical shifts between this construct and a previously assigned construct, EGF11-13, confirm the presence of hydrophobic interdomain contacts between the hairpin turn of the major β-sheet in EGF11 and the conserved aromatic residue within the C-terminal region of EGF10. This suggests that the EGF10-11 interface is rigid.

**Keywords**

Human Notch-1, NMR resonance assignments, EGF domain

**Biological context**

The Notch pathway is a core signal transduction pathway in Metazoa. Both the Notch receptor and its ligands are cell-surface transmembrane proteins. The receptor becomes activated when ligands from the Delta and Serrate/Jagged families bind to the extracellular EGF11-12 region. This results in two regulated proteolysis steps and release of the intracellular domain of Notch (NICD). The NICD fragment subsequently translocates to the nucleus and, following interaction with the transcription factor CSL and a co-activator Mastermind, promotes transcription of a specific sub-set of genes (for review see Bray 2006).

Canonical signaling via ligand-mediated activation of the Notch receptor is involved in most cell fate decisions which occur both during embryonic development and in the maintenance of the adult organism. Dysregulation of the pathway arises as a consequence of heritable genetic disorders such as Alagille-syndrome and CADASIL, but also in acquired disease including various cancers such as T-ALL. The direction and intensity of the Notch signal is often context specific (Bray 2006, Guruharsha, Kankel, and Artavanis-Tsakonas 2012, Kopan and Ilagan 2009). In disease, Notch signaling can be oncogenic, as in hematopoietic malignancies, but can also have tumour suppressor functions, such as in solid tumours (Koch and Radtke 2007).

Whereas high resolution structures of the negative regulatory region (NRR), NICD and EGF11-12 ligand-binding region (Figure 1) have facilitated our understanding of receptor activation and signaling (for review see Chillakuri et al. 2012), little structural information is available for the majority of the extracellular region. This information is necessary to understand how the receptor is activated and regulated by O-glycosylation at the cell surface. The extracellular region comprises 36 EGF-like domains organized as multiple tandem repeats (Figure 1). Early deletion studies in cell aggregation assays revealed a critical role for EGF11 and EGF12 in ligand binding (Rebay et al. 1991). Further mapping indicated a key role for the central β-hairpin of human Notch-1 (hN-1) EGF12 (Whiteman et al. 2013), and suggested that O-fucosylation and Fringe extension of a threonine residue located on the hairpin extended the binding site (Taylor et al. 2014). Very recently the structure of a complex of Notch ligand Delta-like 4 and Notch-1 has confirmed and extended these observations, indicating an extended antiparallel binding surface (Luca et al. 2015). High resolution X-ray and NMR structures for the hN-1 EGF11-13 construct (Cordle et al. 2008a,

Hambleton et al. 2004, Luca et al. 2015) all indicate this region is elongated with well-defined interdomain interfaces stabilized by hydrophobic packing and $Ca^{2+}$ binding. These structural features, of an elongated and rigid structure, may be present in other regions of Notch, since most of the EGF-like domains in Notch are $Ca^{2+}$-binding (red in Figure 1) and also have a conserved aromatic residue likely to be involved in interdomain packing. However, three EGF domains in the extracellular region, EGF6, EGF10 and EGF22, do not bind calcium and could form flexible, rigid, or atypical interfaces as found for other EGF-like domain containing proteins (Brandstetter et al. 1995, Kettle et al. 1999, Lee et al. 2004). This information is critical to our understanding of the conformation of the extracellular region, which could project out from the cell surface in a near linear organization, or adopt a more compact jack-knife structure depending upon the properties of these interfaces.

Here, we focus on EGF10 (Figure 1) which is of particular importance because of its proximity to the ligand-binding domains EGF11 and EGF12. The functional role of the EGF domains N-terminal to EGF11 is not known, but the addition of EGF10 to the construct hN-1 EGF11-13 has been shown to cause a reduction of binding to human Delta-like-1 (Cordle et al. 2008b). Furthermore, an additional binding site in *Drosophila* Notch for the ligand Serrate was recently identified in EGF8 (Yamamoto et al. 2012). EGF10, as a non-calcium binding EGF domain, is conserved in *Drosophila* and in both species there are only 5 residues in between the last cysteine of EGF9 and the first cysteine in EGF10. All other EGF domains in Notch contain 6 residues in this interdomain linker. The presence of an arginine (in *Drosophila* it is a lysine) at the position of the normally conserved aromatic packing residue in EGF9 further suggests an atypical interaction between EGF9 and EGF10.

In order to investigate the role of EGF10 in the Notch receptor, we have cloned a construct in which EGF10 is placed in a native-like context with flanking calcium-binding EGF domains, EGF11 at its C-terminus and EGF8 and EGF9 at its N-terminus. In this report we present $^{1}H$, $^{13}C$ and $^{15}N$ assignments for the four domain EGF-like domain construct EGF8-11 from human Notch-1 expressed in *E. coli*.

**Methods and experiments**

Protein expression, purification and refolding

Protein expression and purification was carried out as described previously for similar constructs (Weisshuhn et al., 2014, Knott et al. 1996, Whiteman, Redfield, and Handford 2014). Constructs were expressed in *E. coli* BL21 cells transformed with a pQE30 (Qiagen) based protein expression construct and a pREP4 plasmid for control of expression via the Lac Repressor. The expression vectors contained an N-terminal His6 tag for purification. The four-domain construct, EGF8 to EGF11 (EGF8-11), contained an enterokinase protease recognition site (DDDL), for removal of the His6 tag, and a SA spacer followed by D295 of EGF8. The last residue in the EGF8-11 construct was I451. A second construct, containing EGF9 to EGF11, was also used to help with resonance assignment. This contained a factor Xa protease recognition site (IEGR) and a SA spacer followed by N335 of EGF9. The His6 tag was not cleaved in this construct due to the presence of a cryptic factor Xa cleavage site in EGF9.

$^{15}N$ and $^{15}N/^{13}C$-double-labelled protein was produced by growing cells in M9 medium containing 0.1% (w/v) $^{15}NH_4Cl$ and 0.5% (w/v) $^{13}C$-glucose (Cambridge Isotope Laboratories), 100 μg/ml ampicillin and 25 μg/ml kanamycin as described previously (Muranyi et al. 2004). 100 ml of starter culture, grown in labeled medium at 37°C for ~18 hours, was added to 0.5 l of medium. Cells were grown to $OD_{600}$ ~0.8 and expression was induced with isopropyl-β-D-thiogalactopyranoside (IPTG) at a final concentration of 100 μM. Cells were incubated at 18°C for at least 6 hours. Cells were spun down at 5000 rpm for 10 minutes and lysed with 6 M guanidine-HCl, 50 mM sodium phosphate, 5 mM β-mercaptoethanol for 1 h at room temperature (RT). DNA was sheared by sonication for 1 min at 20 Watt (Jencons Ultrasonic Processor) and the cell lysate was spun down at 40,000 rpm for 45 minutes (Beckman L7-55). The supernatant was loaded onto a $Ni^{2+}$ chelating Sepharose column (GE Healthcare) and eluted with buffer containing 50 mM EDTA and 100 mM Tris pH 8.3. The protein was reduced for 1h at RT by addition of dithiothreitol (DTT) to a final concentration of 0.1 M. The pH was then lowered to pH ~2, by addition of concentrated HCl, and the solution was dialysed against 0.1% (v/v) trifluoroacetic acid (TFA) overnight. The solution was filtered through 0.2 μM filters (Millex-GP) and the soluble

fraction purified by reverse-phase HPLC using a Beckman Gold system with reverse phase C8 column. Purified, reduced protein was refolded under the following conditions: 100 mM Tris-HCl pH 8.3, 3 mM L-cysteine, 0.3 mM L-cystine, 0.2 mg/ml protein at 37°C for ~48 hours. The refolding mixture was acidified to pH ~2 and dialysed against 0.1% (v/v) TFA for at least 5 hours. Protein was then concentrated by ultrafiltration, purified by HPLC and then lyophilized. For the EGF 8 to EGF11 construct, the His6 tag was removed by incubation overnight with enterokinase at an enzyme:protein ratio by weight of 1:1,000,000 at RT in 50 mM Tris-HCl pH 7.5, 0.1 M NaCl, 1 mM $CaCl_2$. Proteins were further purified by cation-exchange fast liquid protein chromatography and HPLC. The final products were analysed by SDS-PAGE in the presence and absence of 5% (v/v) 2-mercaptoethanol. Electrospray ionisation mass spectrometry was performed to check the correct mass of the constructs (14223.4 Dalton for the unlabeled EGF9-11 construct, 16608 Dalton for the unlabeled EGF8-11 construct).

NMR spectroscopy

Protein concentrations for resonance assignment experiments were 1 mM for unlabeled and $^{15}N$-single-labelled, and 0.5 mM for $^{15}N/^{13}C$-double-labelled material, in 95% $H_2O$/5% $D_2O$ (v/v) at pH 6.1. A $CaCl_2$ concentration of 30 mM and 40 mM were used for EGF8-11 and EGF9-11, respectively. All NMR experiments were carried out at 298 K on home-built spectrometers with triple-resonance probes and GE/Omega data acquisition systems operating at $^{1}H$-operating frequencies of 600 and 950 MHz. In addition, a Bruker Avance 500 MHz spectrometer with a Cryoplatform, equipped with a TCI CryoProbe, was used for most triple-resonance experiments. Initially, sequential backbone assignment was attempted using 3D $^{15}N$-edited TOCSY-HSQC and NOESY-HSQC experiments for EGF9-11 (at 600 MHz) and EGF8-11 (at 950 MHz). More extensive backbone assignments were obtained using 3D HNCA, HNCO, HN(CA)CO, CBCA(CO)NH and HBHA(CBCACO)NH experiments collected at 500 MHz for $^{15}N/^{13}C$-double labeled EGF8-11. Side-chain assignments were obtained using these experiments, supplemented with 3D HCCH-TOCSY, 3D $^{13}C$-edited NOESY-HSQC and 2D $^{1}H$-$^{1}H$ NOESY experiments. NMRPipe was used for processing NMR data (Delaglio et al. 1995) and CcpNmr Analysis for subsequent analysis (Vranken et al. 2005).

**Assignments and data deposition**

$^{1}H^{N}$ and $^{15}N$ resonance assignments for residues from EGF11 were easily obtained from 3D $^{15}N$-edited TOCSY-HSQC and NOESY-HSQC spectra of EGF9-11 since most of the chemical shifts were similar to those for EGF11 in the previously assigned EGF11-13 construct (Muranyi et al. 2004). However, the majority of peaks belonging to EGF9 and EGF10 were weak or missing from the HSQC spectrum of EGF9-11 due to severe line broadening. Much better quality HSQC spectra were obtained from the larger EGF8-11 construct in which EGF9 is stabilized by the additional EGF domain at its N-terminus. More complete resonance assignment was carried out using the EGF8-11 construct and triple resonance NMR experiments. Resonances from residues D295 to I451 were then assigned for EGF8-11; Figure 2 shows a $^{1}H$-$^{15}N$ HSQC spectrum of EGF8-11. All non-proline backbone $^{1}H^{N}$ and $^{15}N$ resonances were assigned except H316 and L369. Further assignments were 92.7% of Hα, 92.0% of Hβ, 69.4% of Hγ,60.2% of Hδ and 50.0% of Hε, 95.0% of $^{13}C'$, 99.4% of $^{13}C\alpha$, 95.7% of $^{13}C\beta$, 50.0 % of $^{13}C\gamma$, 40.3% of $^{13}C\delta$ and 15.8% of $^{13}C\varepsilon$. The chemical shift assignments for EGF8-11 have been deposited in the BioMagResBank (http://www.bmrb.wisc.edu) under the accession number 25533.

Assignments have been obtained for EGF11 in both the EGF8-11 and EGF11-13 constructs (Muranyi et al., 2004). The chemical shifts for EGF11 in these two contexts are compared in Figure 3. Most of the measured backbone chemical shifts for EGF11 are similar in the two constructs, suggesting that the overall structure of EGF11 is independent of interdomain contacts with either EGF10 or EGF12. Chemical shift differences are observed for D412, V413 and E450; these residues, located at the N- and C-terminus of EGF11 reflect their different environment in EGF8-11 and EGF11-13. In the former, D412 is preceded by Q411 and E450 is adjacent to the C-terminal residue, I451. In the latter, D412 is preceded by the SA linker and I451 if followed by D452. Interestingly, large changes in chemical shifts are observed for residues that are directly involved in interdomain contacts. Large chemical shift differences are observed for Y444, and residues adjacent to it in the sequence, which is the conserved aromatic residue that mediates packing with EGF12 in the EGF11-13 construct, but not in the EGF8-11 construct where EGF11 is the C-terminal domain. Large chemical shift differences are also observed for L433, G434 and S435; these residues are located in the turn of the major β-sheet in EGF11 that is involved in packing with EGF10 in the EGF8-11

construct but not in the EGF11-13 construct where EGF11 is the N-terminal domain. These data suggest that calcium-binding EGF11 forms an interdomain packing interaction with non-calcium-binding EGF10 that involves residues from the β–hairpin of EGF11 and the conserved aromatic consensus residue between the 5$^{th}$ and 6$^{th}$ cysteine of EGF10.

**Acknowledgements**

PAH and CR acknowledges a Cancer Research UK-funded PhD studentship for support of PCW.

**Figure captions**

Fig. 1 Modular organization of the extracellular domain of the human Notch-1 receptor. The negative regulatory region (NRR) and the transmembrane domain (TM) are indicated. Individual domains belonging to the Notch intracellular domain (NICD) are not indicated separately. Calcium-binding EGF domains are indicated in red. Schematic pictures of the EGF8-11, EGF9-11 and EGF11-13 constructs, discussed in this study, are shown. Each EGF domain contains six highly-conserved cysteine residues paired in a 1-3, 2-4, 5-6 arrangement to stabilize domain structure. There are typically 6 residues between the 6$^{th}$ cysteine of an EGF domain and the 1$^{st}$ cysteine of the following domain. EGF9 and EGF 10 have only 5 residues in this linker. Since EGF9 and EGF11 have a calcium-binding consensus sequence, the linker length for EGF8/EGF9 and EGF10/EGF11 comprises 2 residues; the other 4 residues are constrained by calcium binding. In the case of EGF9/EGF10, the conformation of the five amino acids is unknown and will be the subject of a future structural study.

Fig. 2 500 MHz $^{1}$H-$^{15}$N HSQC spectrum of hN-1 construct EGF8-11 at pH 6.1, 30 mM $Ca^{2+}$, 298 K, in 95% $H_2O$/5% $D_2O$. Some weak peaks not visible at the contour level chosen are indicated with circles.

Fig. 3 Chemical shift differences for residues from EGF11 between the EGF8-11 and EGF11-13 constructs. EGF11 is the C-terminal domain in EGF8-11, whereas it is the N-terminal domain in EGF11-13. Chemical shift differences centered around L433/G434 reflect the involvement of these residues in an interdomain packing interaction with EGF10 in the EGF8-11 construct but not in EGF11-13. Chemical shift differences centered around Y444 reflect the involvement of this conserved aromatic residue in an interdomain packing

interaction with EGF12 in EGF11-13 but not in EGF8-11. The 6 cysteine residues are shown in yellow. Residues involved in calcium binding (D412/V413/E415/N431/T432/S435) are shown in red. The consensus aromatic, F436, and the packing aromatic, Y444, are shown in green. The black arrows below the sequence indicate the location of the β-hairpin in EGF11.

**References**


Brandstetter H, Bauer M, Huber R, Lollar P, Bode W (1995) X-ray structure of clotting factor IXa: Active site and module structure related to xase activity and hemophilia b. Proc Natl Acad Sci U S A 92(21): 9796-800

Bray SJ (2006) Notch signalling: A simple pathway becomes complex. Nat Rev Mol Cell Biol 7(9): 678-89

Chillakuri CR, Sheppard D, Lea SM, Handford PA (2012) Notch receptor-ligand binding and activation: Insights from molecular studies. Semin Cell Dev Biol 23(4): 421-8

Cordle J, Johnson S, Tay JZ, Roversi P, Wilkin MB, de Madrid BH, Shimizu H, Jensen S, Whiteman P, Jin B, Redfield C, Baron M, Lea SM, Handford PA (2008a) A conserved face of the jagged/serrate dsl domain is involved in notch trans-activation and cis-inhibition. Nat Struct Mol Biol 15(8): 849-57

Cordle J, Redfield C, Stacey M, van der Merwe P, Willis AC, Champion BR, Hambleton S, Handford PA (2008b) Localization of the Delta-like-1-binding Site in Human Notch-1 and Its Modulation by Calcium Affinity. J Biol Chem 283(17):11785-93.

Delaglio F, Grzesiek S, Vuister G, Zhu G, Pfeifer J, Bax, A (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes, J Biomol NMR 6(3): 277-93

Guruharsha KG, Kankel MW, Artavanis-Tsakonas S (2012) The notch signalling system: Recent insights into the complexity of a conserved pathway. Nat Rev Genet 13(9): 654-66

Hambleton S, Valeyev NV, Muranyi A, Knott AV, Werner JM, McMichael AJ, Handford PA, Downing AK (2004) Structural and functional properties of the human notch-1 ligand binding region. Structure 12(12): 2173-83

Kettle S, Yuan X, Grundy G, Knott V, Downing AK, Handford PA (1999) Defective calcium binding to fibrillin-1: Consequence of an N2144S change for fibrillin-1 structure and function. J Mol Biol 285(3): 1277-87

Knott V, Downing AK, Cardy CM, Handford PA (1996) Calcium binding properties of an epidermal growth factor-like domain pair from human fibrillin-1. J Mol Biol 255(1): 22-7

Koch U, Radtke F (2007) Notch and cancer: A double-edged sword. Cell Mol Life Sci 64(21): 2746-62

Kopan R, Ilagan MX (2009) The canonical notch signaling pathway: Unfolding the activation mechanism. Cell 137(2): 216-33

Lee SSJ, Knott V, Jovanovic J, Harlos K, Grimes JM, Choulier L, Mardon HJ, Stuart DI, Handford PA (2004) Structure of the integrin binding fragment from fibrillin-1 gives new insights into microfibril organization. Structure 12(4): 717-29

Luca VC, Jude KM, Pierce NW, Nachury MV, Fischer S, Garcia, KC (2015) Structural basis for Notch1 engagement of Delta-like 4. Science 347:847-853

Muranyi A, Hambleton S, Knott V, McMichael A, Handford PA, Downing AK (2004) $^{1}$H, $^{13}$C, and $^{15}$N resonance assignments of human Notch-1 calcium binding EGF domains 11-13. J Biomol NMR 29(3): 443-4

Rebay I, Fleming RJ, Fehon RG, Cherbas L, Cherbas P, Artavanis-Tsakonas S (1991) Specific EGF repeats of Notch mediate interactions with Delta and Serrate: Implications for Notch as a multifunctional receptor. Cell 67(4): 687-99

Taylor P, Takeuchi H, Sheppard D, Chillakuri C, Lea SM, Haltiwanger RS, Handford PA (2014) Fringe-mediated extension of O-linked fucose in the ligand-binding region of Notch1 increases binding to mammalian Notch ligands. Proc Natl Acad Sci U S A 111(20): 7290-5

Vranken W, Boucher W, Stevens T, Fogh R, Pajon A, Llinas M, Ulrich E. Markley J, Ionides J, Laue, E (2005) The CCPN data model for NMR spectroscopy: development of a software pipeline. Proteins, 59(4), 687-96

Weisshuhn PC, Handford PA, Redfield C (2014) $^{1}$H, $^{13}$C and $^{15}$N assignments of EGF domains 4 to 7 of human Notch-1, Biomol NMR Assign, DOI 10.1007/s12104-014-9591-x

Whiteman P, de Madrid BH, Taylor P, Li D, Heslop R, Viticheep N, Tan JZ, Shimizu H, Callaghan J, Masiero M, Li JL, Banham AH, Harris AL, Lea SM, Redfield C, Baron M, Handford PA. (2013). Molecular basis for Jagged-1/Serrate ligand recognition by the Notch receptor. J Biol Chem 288(10): 7305-12

Whiteman P, Redfield C, Handford PA (2014) Bacterial expression and in vitro refolding of limited fragments of the Notch receptor and its ligands. Methods Mol Biol 1187: 193-208

Yamamoto S, Charng WL, Rana NA, Kakuda S, Jaiswal M, Bayat V, Xiong B, Zhang K, Sandoval H, David G, Wang H, Haltiwanger RS, Bellen HJ (2012) A mutation in EGF repeat-8 of Notch discriminates between Serrate/Jagged and Delta family ligands. Science 338(6111): 1229-32

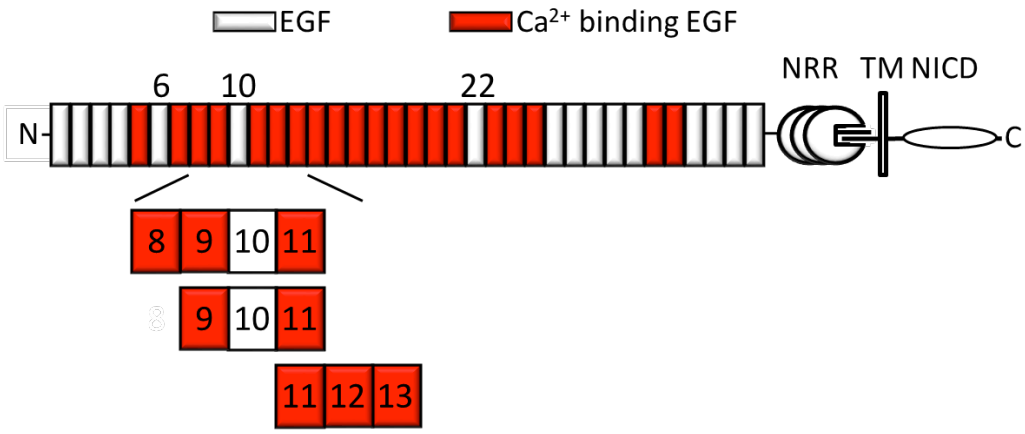


Figure 1

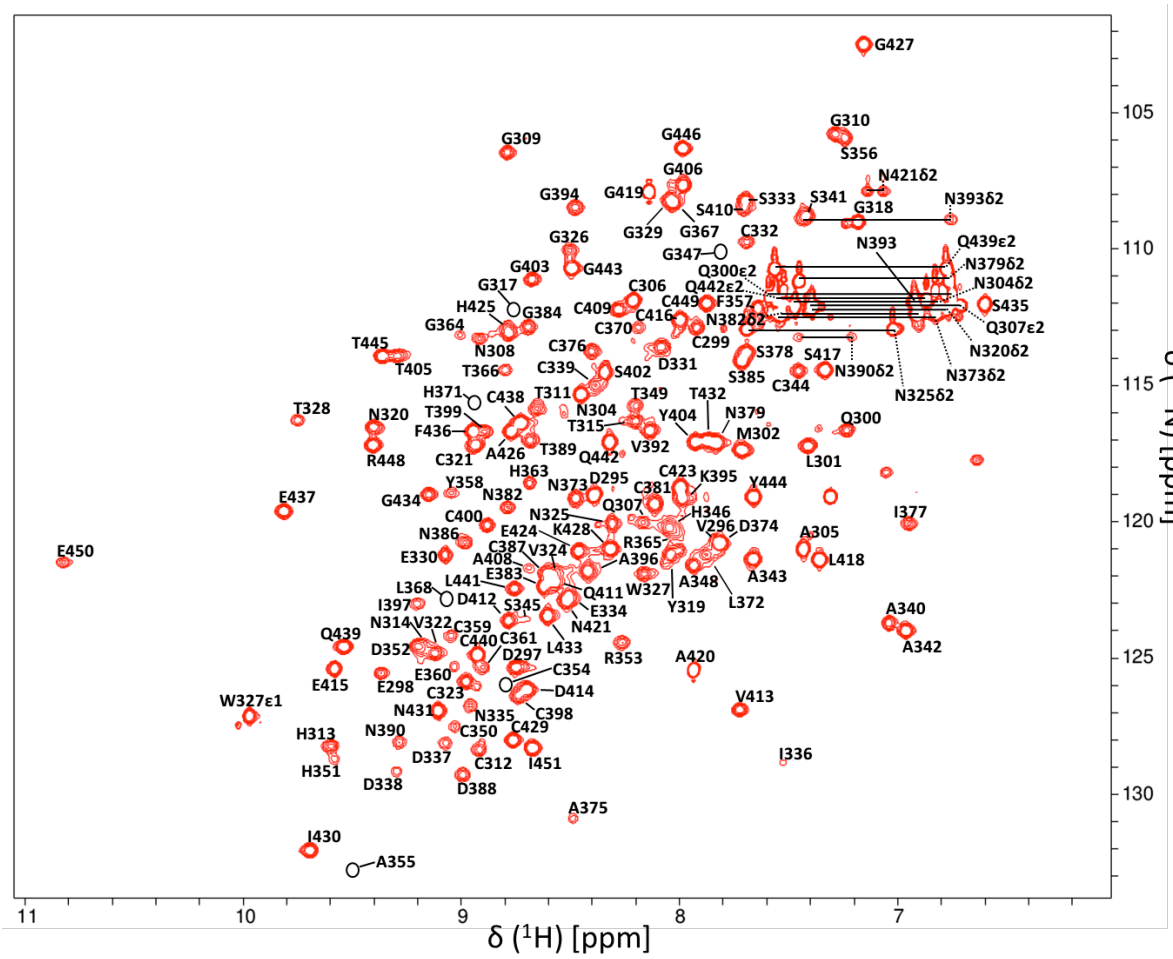
δ (15N) [ppm]
δ (1H) [ppm]
105
110
115
120
125
130
11
10
9
8
7
G427
G310
S356
G309
G446
G406
N421δ2
G394
G419
S333
S341
G318
N393δ2
S410
G329
G367
C332
G326
N393
Q439ε2
G347
N379δ2
G403
G443
Q300ε2
N304δ2
G317
C306
Q442ε2
S435
H425
G384
C409
C449
F357
C416
N382δ2
Q307ε2
G364
C370
N320δ2
T445
N308
C376
C299
S378
S417
T405
T366
C339
S402
D331
S385
C344
N390δ2
N373δ2
N325δ2
T328
H371
C438
T311
T349
T432
N320
T399
N304
Y404
N379
Q300
F436
T315
M302
R448
C321
A426
T389
Q442
V392
L301
H363
C423
K395
Y444
E437
Y358
D295
C381
G434
N382
Q307
I377
C400
N325
H346
E450
N386
E424
K428
V296
D374
A305
C387
V324
R365
E330
A408
A396
L418
E383
A343
L441
A348
L368
Q411
W327
D412
S345
E334
Y319
L372
I397
A340
N314
V322
C359
N421
Q439
C440
C361
A342
D352
D297
L433
R353
A420
E360
C354
E415
E298
C323
D414
V413
W327ε1
N431
N335
C398
H313
N390
C350
C429
I336
D337
C312
I451
H351
D338
D388
A375
I430
A355

Figure 2

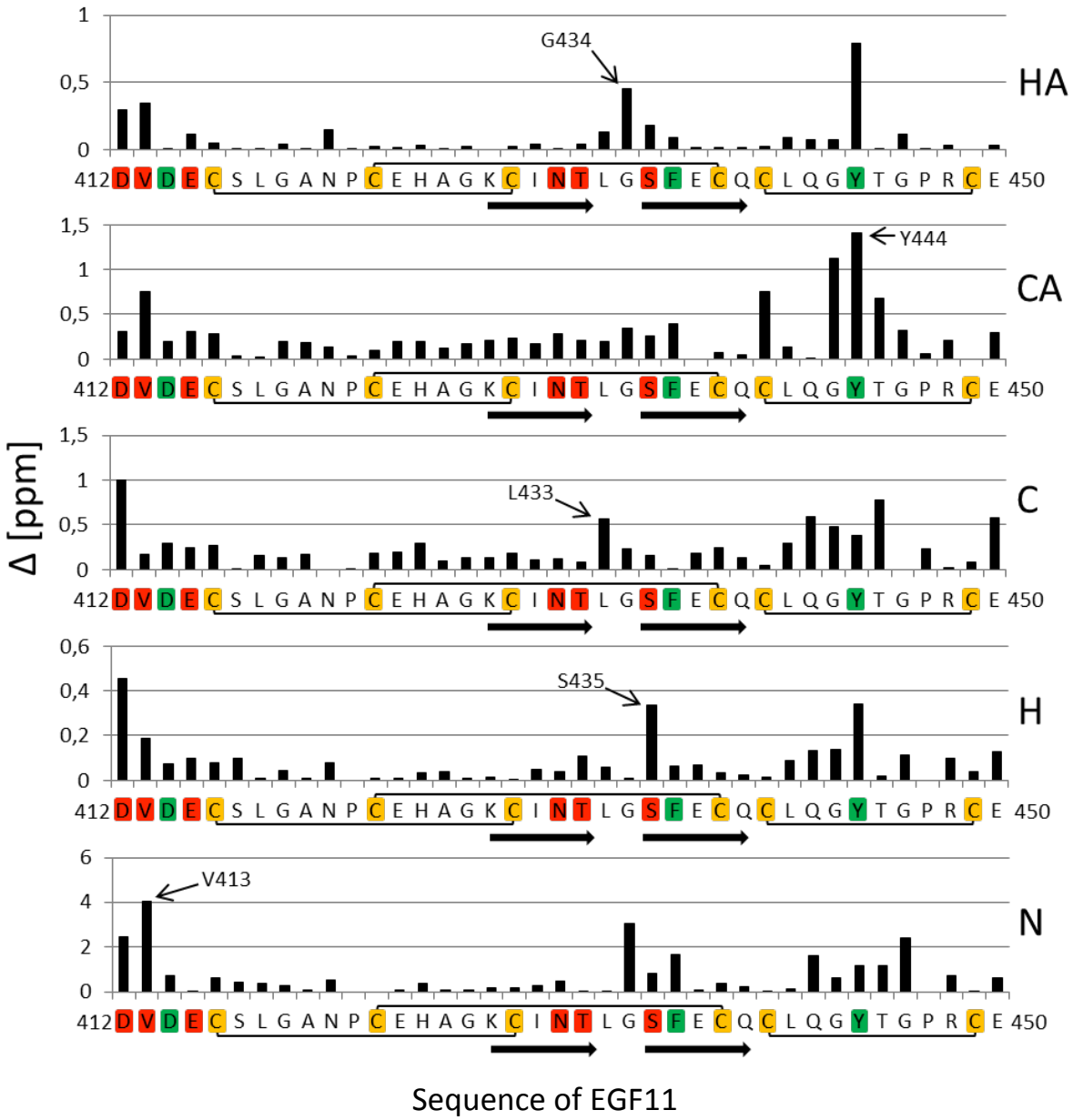
Δ [ppm]
HA
G434
CA
Y444
C
L433
H
S435
N
V413
412 D V D E C S L G A N P C E H A G K C I N T L G S F E C Q C L Q G Y T G P R C E 450
Sequence of EGF11

Figure 3