**MASL1: a neglected ROCO protein**

Sybille Dihanich*
Department of Molecular Neuroscience, UCL Institute of Neurology, Queen Square, London WC1N 3BG, U.K.

**Abstract**

The human ROCO proteins are a family of four proteins characterized by a conserved supradomain: a Ras-like GTPase domain. This domain consists of ROC (Ras of complex proteins) occurring in tandem with a COR (C-terminal of ROC) domain. Together, these proteins are linked to various pathologies including cancer and PD (Parkinson’s disease). Despite an increasing research focus on these proteins, their functions in general, and their specific roles in disease, are still unknown. In the case of MASL1 (malignant fibrous histiocytoma amplified sequences with leucine-rich tandem repeats 1), a predicted oncoprotein in MFHs (malignant fibrous histiocytomas), there is a particular lack of information available in the literature. The aim of the present review is therefore to summarize the existing information on MASL1 and also to compile data that could be linked to MASL1 and thus help our understanding of this neglected ROCO protein.

**Introduction**

The term MFH (malignant fibrous histiocytoma) has been used to classify soft tissue sarcomas since the 1960s. However, changes in diagnosis and classifications over the last 50 years have resulted in a re-evaluation of MFH, which have led to the revised terminology of ‘undifferentiated pleomorphic sarcomas’ [1–5]. Studies on MFH have brought forward a candidate oncogene termed MASL1 (MFH amplified sequences with leucine-rich tandem repeats 1) [also known as MFHAS1 (MFH amplified sequence 1)] [6], which belongs to a group of multidomain proteins termed ROCO proteins [7,8]. Numerous studies have linked this protein family to human pathologies, but it is the link to PD (Parkinson’s disease) through one of its family members [LRRK (leucine-rich repeat kinase) 2] that has put these proteins in the spotlight [6,9–11]. Despite an increasing body of literature, the cellular functions of the ROCO protein family remain unknown. In particular, MASL1, the least understood member of the ROCO protein family, needs to be studied and characterized in order to provide information on a biochemical and protein level, which is currently non-existent in the literature.

**MASL1 and MFHs**

MFHs were first described in the 1960s and characterize soft tissue sarcomas, assumed to derive from histiocytes capable of transforming into fibroblasts [1–3]. Over several decades, this terminology has been applied to classify these tumours, regarded as the most common soft tissue sarcomas in adults, leading to improved diagnosis over the years [3]. Not only did this increased diagnosis result in subdivisions of MFHs, but it also raised concerns about misdiagnoses, an overview of the term MFHs and that it has become a diagnosis of exclusion [2,3]. Together, these developments in MFH classification and diagnosis led to a re-evaluation of MFHs, which subsequently resulted in the new terminology of ‘undifferentiated pleomorphic sarcoma’ from 2002 [4,5]. These changes were driven by an improved collection of diagnostic tools such as immunohistochemistry or electron microscopy and as such the diagnosis of MFHs reduced significantly [12].

Despite discrepancies in terminology, sarcomas (including MFHs) are known to encompass alterations in karyotypes (chromosomal losses/gains), translocations and fusion genes and are generally divided into sarcomas with specific or unspecific genetic alterations [13]. Although to date no specific genetic alterations have been attributed to MFHs, a few studies have focused on investigating such changes, resulting in the identification of *MASL1* as a potential oncogene [6,14]. In 1999, Sakabe et al. [6] analysed primary MFH tumours and found an amplification of DNA at six loci, one of which is found on chromosome 8 (8p23) and contains MASL1. In this study, comparative genomic hybridization revealed the loss and gain of numerous regions on different chromosomes in 16 out of 19 tumours studied. The subsequent focus on a single locus on chromosome 8 was based on the fact that this location (8p23) was identified as a novel amplified region in gastric cancers as well as oesophageal/gastro-oesophageal cancers before the study by the Sakabe group [15,16]. A subsequent refinement of this amplicon led to the finding of *MASL1*, a candidate (proto-) oncogene [6]. In addition to solid tumours, it has also been suggested that *MASL1* could play a role in haematological malignancies and that it might be a possible target gene for translocation as investigated in a B-cell lymphoma cell line [14].

---

**Key words:** leucine-rich repeat kinase 2 (LRRK2), malignant fibrous histiocytoma amplified sequences with leucine-rich tandem repeats 1 (MASL1), Parkinson’s disease, ROCO protein.

**Abbreviations used:** CDK, C-terminal of Ras of complex proteins; DAPK1, death-associated protein kinase 1; LRR, leucine-rich repeat; LRRK, leucine-rich repeat kinase; MASL1, malignant fibrous histiocytoma amplified sequences with leucine-rich tandem repeats 1; MFH, malignant fibrous histiocytoma; PD, Parkinson’s disease; ROC, Ras of complex proteins.

1 email s.dihanich@ucl.ac.uk
MASL1: beyond a candidate oncogene

Considering the existing literature on MASL1, it is unsurprising that our understanding of this gene/protein is very limited. However, by looking at MASL1 from different angles, we might be able to learn more about processes or pathologies in which it could be involved. For example, even though MASL1 is not linked to other pathologies directly, we know that the chromosomal location where MASL1 can be found, the short arm of chromosome 8 (8p23), is linked to congenital heart defects, microcephaly and behavioural problems [17–19], and has also been mentioned in connection with other disorders such as CdLS (Cornelia de Lange syndrome), NCL (neuronal ceroid lipofuscinosis), asthma, diabetes or ASDs (autism spectrum disorders) [20–24]. Considering this wide spectrum of diseases, it is unlikely that they are all linked to or caused by MASL1 solely, as several other genes can be found in this location of the chromosome (NCBI table view of genes on chromosome 8: http://www.ncbi.nlm.nih.gov/mapview/maps.cgi?TAXID=9606&CHR=8&MAPS=ideogr,ugHs,genes[1.00%3A146364022.00]&CMD=TXT#2). Nevertheless, MASL1 might play a role for some of these processes/diseases, and investigating whether it does could give us a better understanding of its functions. Furthermore, it may establish a role for MASL1 beyond the one of an oncogene. As the above-mentioned pathologies are predominantly attributed to 8p23 deletions, both a lack of MASL1 as well as an increase of MASL1 could be linked to disease. This would suggest that maintaining the correct levels of MASL1 is important for its normal function and alterations either way, increases as well as decreases, could have negative effects.

In the case of two immune-related disorders, psoriasis and Crohn’s disease, structural alterations at 8p23 (changes in copy numbers) have been linked to a protection against Crohn’s disease or susceptibility towards psoriasis [26,27]. These changes were accompanied by altered copy numbers of DEFB4 (β-defensin 4), a peptide important for an innate immune response. Although these studies did not investigate changes in MASL1, they could imply 8p23 as a candidate locus for immune-related disorders. A recent study investigating the gene expression profile of macrophages upon bacterial challenges revealed a strong induction of MASL1 by four out of five different pathogens [28]. This study places MASL1 downstream of TLR (Toll-like-receptor) signalling, implying a potential role in inflammatory responses of innate immunity [28]. Furthermore, this bioinformatics study showed that classification of LRR (leucine-rich repeat) proteins, which include MASL1, could lead to functional insights and help to elucidate molecular pathways of these proteins [28].

MASL1 and RCO proteins

With the bioinformatics study by Ng et al. [28] in mind, classifications of LRR proteins and the literature available on these related proteins might provide further insights into the role of MASL1. Interestingly, MASL1 belongs to a distinct group of four human proteins termed RCO proteins [7,29] (Figure 1A), characterized by a GTPase supradomain [8]. In contrast with MASL1, its family members LRRK1 and LRRK2 and DAPK1 (death-associated protein kinase 1) have been studied considerably more. We therefore know that RCO proteins are linked to disorders such as Crohn’s disease, cancer or PD [6,9–11,30]. In particular, the link of LRRK2 to PD, known since 2004 [9,10], has put these proteins in the spotlight which resulted in an increasing pool of information about their function, structure and the role they play in disease [7,29].

The RCO protein domain organization itself, in particular their Ras-like GTPase domain, implies a role in intracellular signalling. This GTPase/ROC (Ras of complex proteins)–COR (C-terminal of ROC) domain is the unifying factor within this protein family and as such a detailed understanding of its role becomes of great importance and interest. Although the GTPase domains of LRRK2 and DAPK1 are known for a reciprocal interaction with effector domains of these proteins, a kinase and death domain respectively [31–33], its role in MASL1, which does not contain an effector domain, remains unclear (Figure 1A). Even though studies on DAPK1 and LRRK2
have linked these proteins to numerous functions (e.g. vesicle formation, autophagy or immune-related functions), it is unclear whether MASL1 may be involved in similar processes [28,34–36]. Therefore it will be important to determine the role of MASL1 to see whether it functions in a similar fashion. However, the lack of an effector domain would imply that MASL1 interacts with another protein to control its effector domain (Figure 1B). It is therefore crucial to find out whether MASL1 has such a binding partner, what that binding partner is and in what functions/pathways it might be involved.

Before determining binding partners, functions or signalling pathways, very basic questions about GTP binding, GTPase activity or structure, as carried out for other ROCO proteins [31,32,37], need to be addressed in the case of MASL1. Besides filling this void within the MASL1 literature, such studies would show how MASL1 compares with its family members and determine whether it is indeed a GTP-binding protein with GTPase activity. Additionally, in an indirect way, understanding MASL1 would aid our understanding of ROCO proteins in general. Relatively high homology within this protein family, in particular within the ROC–COR domain (Figure 1A), could also result in functional convergence. Thus getting a better understanding of the entire protein family could shed more light upon the function of individual proteins. Some residues that are linked to PD when mutated in LRRK2 (Ile1371/Arg1441) are, for example, homologous between MASL1 and LRRK2, but not between LRRK2 and LRRK1 (Figures 2 and 3). Whether this means that these residues have the same function in both MASL1 and LRRK2 is currently unclear, but could be investigated in MASL1 and subsequently compared with the literature available on these residues in LRRK2 [32,38]. As such, MASL1 might even be used as a tool to investigate PD-causing mutations in a much simpler system, considering its size and domain organization (Figure 1A). Furthermore, clues about MASL1 complex formation might derive from turning our attention to other ROCO proteins such as the prokaryotic ROCO protein in *Chlorobium tepidum* [37]. This protein is the prokaryotic homologue of LRRK2, but also resembles MASL1 in size, domain organization.

**Figure 2** | Sequence alignment of ROCO proteins

Residues/mutations linked to LRRK2-PD (I1371V and R1441C) can also be found in MASL1, whereas no homologues are found in LRRK1. Whether these residues play convergent roles in LRRK2 and MASL1 is currently unclear, but this might provide a platform for investigating disease-causing mutations in a simpler system such as MASL1.

**Sequence alignment: LRRK2/MASL1 vs. LRRK2/LRRK1**

![Sequence alignment of ROCO proteins](image1)

- **A** Chlorobium tepidum ROCO and MASL1
  - 27% sequence identity & E-value of 1E-29

- **B** Chlorobium tepidum ROCO and LRRK2
  - 25% sequence identity & E-value of 4E-37

- **C** MASL1 and LRRK2
  - 25% sequence identity & E-value of 5E-31

**Figure 3** | Sequence alignment of the prokaryotic ROCO protein (*C. tepidum*) and human ROCO proteins

A protein-protein BLAST search comparing the prokaryotic ROCO protein with MASL1 (A) and LRRK2 (B) revealed that both human ROCO proteins have a similar degree of sequence alignments with their prokaryotic counterpart. The dot matrix shows similarities between subject (*C. tepidum*) presented on the x-axis and the query (MASL1 or LRRK2) presented on the y-axis. Alignments between the two are shown as lines slanted from the bottom to the upper right corner. (C) Similarly, MASL1 and LRRK2 also have a high degree of sequence alignment.
and sequence homology (Figure 3). Additionally, structural information obtained from this protein shows that the COR domain plays an important role in dimerization which, in turn, is important for GTPase activity and its regulation [37]. Even though it is currently unclear whether the same will hold true for MASL1, comparing it with closely related ROCO proteins could help to characterize this protein.

The biology of MASL1 is currently a blank canvas: almost nothing is known about the cellular function of this protein, with what we do know derived by analogy or through bioinformatics. Learning more about the function of MASL1 is likely to be beneficial both in terms of understanding how MASL1 is linked to disease and in the context of a generic role for the ROCO proteins, providing us with a simple model for the ROC-COR supradomain in human cells. With studies suggesting an involvement of these domains in important basic cellular functions such as vesicle formation, autophagy or immune responses, it will be important to investigate the very same functions with regard to MASL1 in the future. Combining cellular studies with structural approaches will give us a powerful insight into the correlation between structure and function for MASL1, information that may well have implications for our understanding of the role of LRRK2 in disease. This neglected ROCO protein has the potential to offer us a valuable window on the role of this protein family.

Acknowledgements

I thank P. Lewis, C. Manzioni and M. Soutar for critical discussion.

Funding

This work was supported in part by the Wellcome Trust/Medical Research Council Joint Call in Neurodegeneration award [grant number WT089698] to the UK Parkinson’s Disease Consortium (UKPDC) whose members are from the UCL Institute of Neurology, the University of Sheffield and the MRC Protein Phosphorylation Unit at the University of Dundee.

References

25. Reference deleted


Received 1 May 2012
doi:10.1042/BST20120127