Biochimica et Biophysica Acta 1783 (2008) 1467–1479

Contents lists available at ScienceDirect

Biochimica et Biophysica Acta

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

Review

The RCC1 superfamily: From genes, to function, to disease

Ouadah Hadjebi, Eduard Casas-Terradellas, Francesc R. Garcia-Gonzalo [1], Jose Luis Rosa *

Departament de Ciències Fisiològiques II, IDIBELL, Campus de Bellvitge, Universitat de Barcelona, E-08907, L'Hospitalet de Llobregat, Barcelona, Spain

ARTICLE INFO

Article history:
Received 3 December 2007
Received in revised form 19 March 2008
Accepted 20 March 2008
Available online 10 April 2008

Keywords:
RCC1
RLD
GRF
GEF
HERC
RCBTB

ABSTRACT

The Regulator of Chromosome Condensation 1 (RCC1) was identified over 20 years ago as a critical cell cycle regulator. By analyzing its amino acid sequence, RCC1 was found to consist of seven homologous repeats of 51–68 amino acid residues, which were later shown to adopt a seven-bladed β-propeller fold. Since the initial identification of RCC1, a number of proteins have been discovered that contain one or more RCC1-like domains (RLDs). As we show here, these RCC1 superfamily proteins can be subdivided in five subgroups based on structural criteria. In recent years, a number of studies have been published regarding the functions of RCC1 superfamily proteins. From these studies, the emerging picture is that the RLD is a versatile domain which may perform many different functions, including guanine nucleotide exchange on small GTP-binding proteins, enzyme inhibition or interaction with proteins and lipids. Here, we review the available structural and functional data on RCC1 superfamily members, paying special attention to the human proteins and their involvement in disease.

1. Introduction

The RCC1 superfamily of proteins is characterized by a 350–500 residue domain, known as the RCC1-like domain or RLD, that was first reported in 1987 in the Regulator of Chromosome Condensation 1, RCC1 [1,2]. A mutated RCC1 allele was found to be responsible for the temperature-sensitive phenotypes seen in the ts-BN2 hamster cell lines, including premature chromosome condensation and arrest in the G1 phase of the cell cycle. These phenotypes could be rescued by the introduction of wild-type RCC1 [1–3]. It was not until the mid-nineties that another human protein, HERC1, was found that contained not one, but two RLDs [4]. Since then, sixteen additional human proteins have been described that have at least one RLD in their amino acid sequences.

In this review, we show that these proteins can be divided into five subgroups based on their structural characteristics (Fig. 1): (1) the RCC1 subgroup, (2) the HERC subgroup, (3) the RCBTB subgroup, (4) the kinase subgroup and (5) the miscellaneous subgroup, encompassing the proteins that cannot be classified into any of the other subgroups.

In addition, the following sections will present the available knowledge regarding the structure and function of RCC1 superfamily genes and proteins, with an emphasis on the human members. Finally, the involvement of RCC1 superfamily genes and proteins in pathological conditions will also be discussed.

2. RCC1 superfamily genes and proteins

The genes encoding human RLD-containing proteins (see Table 1) are located in different chromosomes, though closer homologs tend to be located in the vicinity of each other. Thus, the genes for RCC1 and TD-60 are close to each other in chromosome 1. Similarly, HERC1 and HERC2 are both on the long arm of chromosome 15, whereas HERC3, HERC5 and HERC6 are in the same region in chromosome 4 [5]. Likewise, both RCBTB1 and RCBTB2 are located on the long arm of chromosome 13. The other genes are dispersed throughout the genome, either on autosomes or, in the case of RPGR, in the X chromosome.

All these genes have different exon numbers and the proteins encoded by them range from little more than 400 amino acid residues to close to 5000 (see Table 1). Splicing isoforms have been studied in some cases, such as those of RCC1, DelGEF, RPGR, PAM and some HERC proteins [6–10]. While alternative splicing of RCC1, HERC6 and PAM does not affect the sequence of their RLDs [6,7,10], the same cannot be said of DelGEF, RPGR and HERC4, where some of the isoforms have truncated RLDs [7,8,11,12]. Apart from that, the Ensembl database also predicts alternatively spliced transcripts for some of the other proteins described here, but whether or not these isoforms actually exist has not yet been examined [13].

Regarding the architectures of RCC1 superfamily proteins (Fig. 1), there are instances where the RLD domain makes up almost the whole protein, as is the case for RCC1 subgroup members (RCC1, TD-60, DelGEF and WBSCR16), or it can be the only highly conserved region of a protein, as in RPGR. In other cases, the RLD is part of more complex, multidomain proteins. For example, RLDs can be found in proteins involved in ubiquitination, where they are associated with HECT or

* Corresponding author. Fax: +34 934024268.
E-mail address: joseluisrosa@ub.edu (J.L. Rosa).
[1] Present address: Gene Expression Laboratory, The Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, CA 92037, United States.

0167-4889/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbamcr.2008.03.015

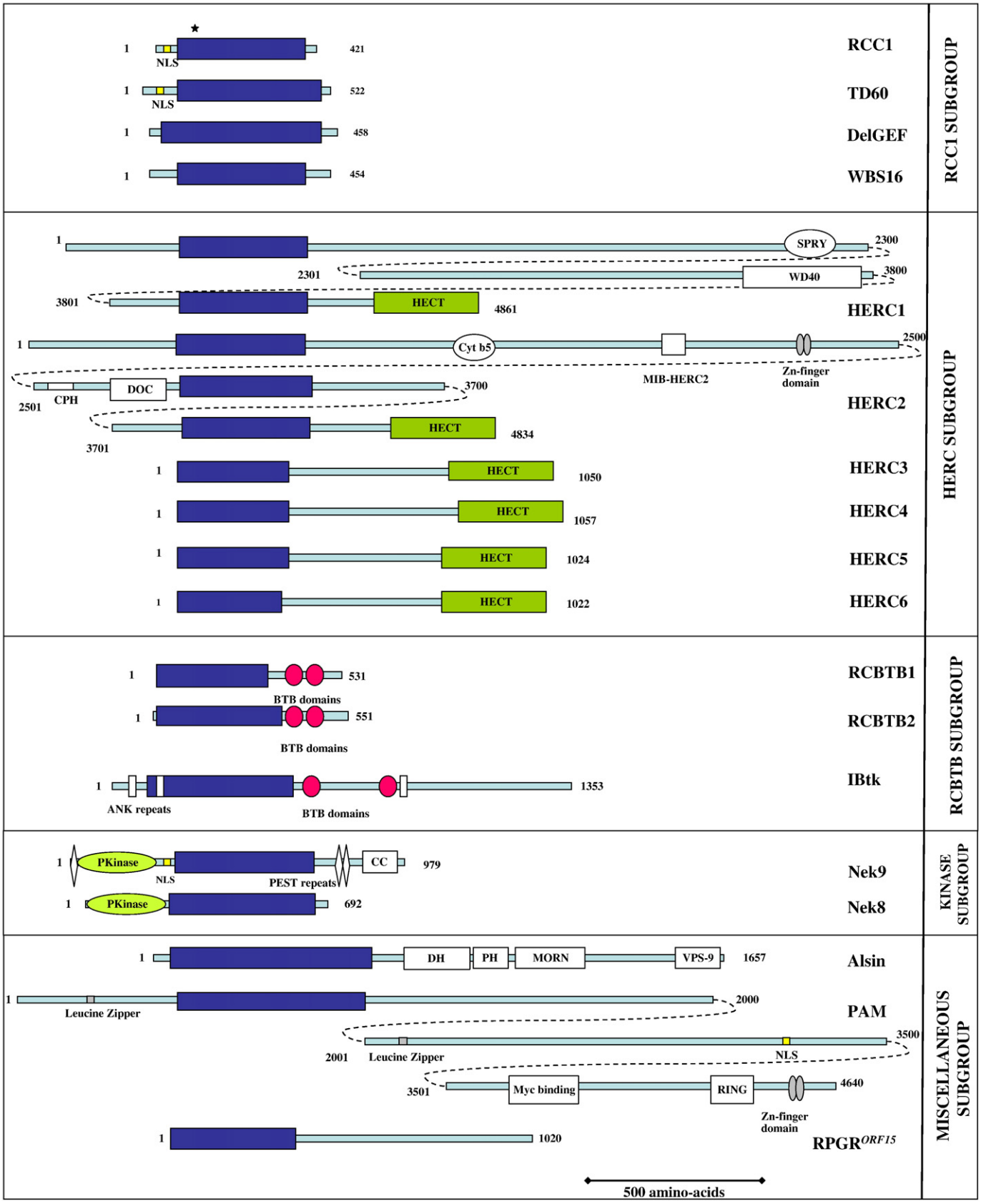

Fig. 1. Schematic representation of human RCC1 superfamily proteins. The eighteen proteins have been classified into five subgroups: (1) the RCC1 subgroup (including those proteins whose RLD spans almost the whole length of the protein), (2) the HERC subgroup (proteins containing RLD and HECT domains), (3) the RCBTB subgroup (proteins containing RLD and BTB domains), (4) the kinase subgroup (proteins containing RLD and kinase domains) and (5) the miscellaneous subgroup, encompassing those proteins that do not fit into any of the previous categories. Dark blue rectangles represent the RLDs. All proteins and domains are shown to scale (scale shown at bottom right side). Big proteins such as PAM, HERC1 and HERC2 are drawn in more than one line, with amino acid numbers shown at each end. A star on top of RCC1's RLD shows the location of the β-wedge, essential for RCC1's activity as a Ran-GEF.