Structure-based mechanistic principles for the paradoxical effects of pathological RET mutations

Anna Fassler Bakhman[1], Michal Cohen[2], Rachel Kolodny[3], Mickey Kosloff[1]

[1]Department of Human Biology, Faculty of Natural Science, University of Haifa, Haifa, 3103301, Israel.

[2]Department of Pediatrics, Ruth Rappaport Children's Hospital, Rambam Health Care Campus, Haifa, Israel

[3] Department of Computer Science, University of Haifa, Haifa, 3103301, Israel

* Correspondence to:

Mickey Kosloff, Department of Human Biology, Faculty of Natural Sciences, University of Haifa, 199 Aba Khoushy Ave., POB 3338, Haifa, 3103301, Israel. Telephone: +972-4-8288996.

E-mail: kosloff@sci.haifa.ac.il. Orcid ID 0000-0003-1807-4000

ABSTRACT

The RET receptor is a transmembrane protein that belongs to the receptor tyrosine kinase family. RET activates signaling pathways regulating cell growth, differentiation and survival in diverse tissues that include the thyroid and enteric nervous system. Mutations that reduce RET levels or function can cause Hirschsprung's (HSCR) disease, characterized by abnormal distal colon innervation in the developing embryo. In contrast, mutations that constitutively activate RET can lead to tumorigenesis, most notably in multiple endocrine neoplasia type 2 (MEN2) syndromes usually identified in the thyroid. Paradoxically, some RET mutations can both reduce RET activity in enteric ganglia, and increase signaling in other tissues, resulting in co-occurrence of both HSCR and MEN2A. Although extensive research has been conducted on RET mutations, the structural and mechanistic bases underlying these paradoxical effects remain unclear. Here, we assimilated data on 70 positions in RET extracellular domains where point-mutations were associated with HSCR, MEN2A, or both. Using 3D structure-based approaches, we predict the potential structural effects of mutations in these positions. Our analysis suggests that approximately 90% of positions associated with HSCR disease can, upon mutation, disrupt intramolecular interactions that stabilize RET tertiary structure: residues buried in the protein core, calcium-binding sites, or residues participating in stabilizing intramolecular electrostatic/covalent bonds. A smaller subset of mutations involves substitutions to/from glycines or prolines in key positions. Only a small minority of HSCR-associated positions affect protein-protein interactions needed for signal activation. On the other hand, our analysis showed that ~75% of mutations in positions that cause MEN2A lead to an unpaired cysteine that can form an intermolecular disulfide bond between two RET monomers. Other mutations that cause MEN2A are also predicted to enhance RET homodimerization via extracellular domains that are proximal to the membrane. Importantly, a substitution leading to an unpaired cysteine that concurrently destabilizes RET tertiary structure is predicted to cause the paradoxical co-occurrence of MEN2A and HSCR. Our findings suggest a mechanistic basis for almost all identified pathological mutations in RET and that therapeutic strategies for targeting RET activity in HSCR and MEN2A may need to be orthogonal.

INTRODUCTION

The receptor tyrosine kinase RET gene encodes for a transmembrane protein that belongs to the receptor tyrosine kinase family. RET initiates signaling pathways that regulate cell growth, differentiation, and survival in diverse tissues that include the embryonic kidney, thyroid, and enteric nervous system (1-3). Abnormal RET signaling can lead to disease, including Hirschsprung's disease (HSCR) – a congenital malformation characterized by agangliosis, the absence of enteric ganglia in segments of the intestine, which is typically diagnosed in newborns and infants (4-6). Conversely, sustained activation of RET signaling later in life contributes to tumorigenesis and in particular to multiple endocrine neoplasia type 2 (MEN2) syndromes (7-10). The RET receptor activates downstream signaling cascades via PI3K/AKT, RAS/RAF/MEK/ERK, JAK2/STAT3, and PLCγ (10). RET activation is typically initiated by binding of any of the following four ligands: Glia cell line-derived growth factor (GDNF), Neurturin (NRTN), Artemin (ARTN), or Persephin – usually combined with binding to one of four GFRa co-receptors (GFRa1-4) (11-13). X-ray structures of RET, which were solved in complex with these four ligands and co-receptors, demonstrated that the extracellular part of the RET receptor is composed of four (D1-D4) cadherin-like domains (CLDs), followed by one cysteine rich domain (CRD) designated D5 (9, 14, 15). In the tetrameric complex of RET with GFRα2 and NTRN, all five extracellular RET domains interact with either ligands or co-receptors, while in the dimeric complexes of RET with the different ligands and co-receptors, only domains D1-D3, and D5 contribute to these interactions (9, 14, 15). The structure of the tetrameric complex further showed that two hexameric 2:2:2 NRTN/GFRα2/RET complexes can dimerize, forming a 4:4:4 complex that suppresses RET endocytosis. This suggested that the RET receptor has two distinct interfaces with ligands and two distinct interfaces with the co-receptors.

The clinical relevance of RET was demonstrated across studies that mapped loss of function (LOF) mutations in the RET extracellular region that can lead to Hirschsprung's disease, presumably by affecting the stability of the RET protein, reducing functional enteric ganglia in the gut tissue (6, 9, 16-18). In contrast, gain of function (GOF) mutations in cysteine residues that are located in the RET CRD were suggested to trigger RET dimerization without requiring ligand and co-receptor binding and thereby lead to MEN2 syndromes (19, 20, 21. Such cysteine mutations presumably result in an unpaired cysteine that induces ligand-independent dimerization and constitutive activation of RET by forming an inter-chain disulfide bond {Li, 2019 #1). Paradoxically, four RET "Janus mutations" in residues C609, C611, C618, and C620 were shown to lead to both LOF and GOF effects and the co-occurrence of both HSCR and MEN2A in the same individual or family (6, 22-29).

In this study, we analyzed computationally structures of the extracellular regions of RET, seeking mechanistic explanations for missense mutation that lead to HSCR, MEN2A, or to the paradoxical occurrence of both diseases. To this end, we identify which positions may perturb RET stability or protein-

protein interactions upon mutation or might form intermolecular bonds that can lead to constitutive activation. Our findings map 70 RET positions reported to lead to clinical phenotypes of HSCR or MEN2A and predict the potential structural effects of mutations in all 594 positions across RET extracellular regions. These insights suggest potential therapeutic avenues for particular mutations, such as stabilizing RET structure for some HSCR-associated positions or inhibiting aberrant dimerization with covalent inhibitors to specific cysteine residues for some MEN2A-associated positions.

RESULTS

Dataset of RET mutations that lead to diseases

We collected a dataset of 70 positions in the extracellular region of RET that were linked to diseases; particularly, documented as leading to HSRC and/or MEN2A upon missense mutations (Table 1). To this end, we first searched the Clinvar database (30) for "pathogenic" and "likely pathogenic" RET variants associated with HSRC disease or MEN2A and found 19 HSRC-associated RET positions and seven MEN2A-associated positions. One of these positions, F555, is associated with both diseases. A search in the COSMIC and cBioPortal databases did not identify additional positions of pathological relevance. We then searched the literature for additional RET positions reported to lead to HSRC or MEN2A (5, 24, 31-35). When the pathological classification of specific positions conflicted between Clinvar and the literature, we prioritized the latter. Specifically, positions C609, C611, C618, and C620 were classified in Clinvar only as GOF mutations that lead to MEN2A. However, several papers showed that mutations in these four positions lead to both HSRC and MEN2A (26, 28, 35, 36). Overall, we collected 70 RET extracellular positions with pathological outcomes upon mutation. Of these, 54 positions are linked exclusively to HSCR disease, 11 are associated only with MEN2A, and five positions are linked to both HSCR and MEN2A.

RET position	RET domain	Reported mutation	Reported effect	Source
S32	D1	S32L	HSCR	Clinvar VCV000013923.5 and (9)
L40	D1	L40P	HSCR	(9, 37)
L56	D1	L56M	HSCR	Clinvar VCV000036723.76 and (31)
R57	D1	R57W	HSCR	(32)
P64	D1	P64L	HSCR	Clinvar VCV001782480.2 and (9)
R77	D1	R77C	HSCR	Clinvar VCV001069381.8 and (9)
G93	D1	G93S	HSCR	Clinvar VCV003065760.1 and (9)
Y96	D1	Y96C	HSCR	(32)
H114	D1	H114C	HSCR	(9, 38)
C142	D1	Unk	HSCR	(9)
V145	D1	V14G	HSCR	(9, 38)
F147	D1	F147S	HSCR	Clinvar VCV000599631.3
P155	D2	P155L	HSCR	(9, 38)
C157	D2	C157Y	HSCR	(9, 31)
F174	D2	Unk	HSCR	(9)
R175	D2	R175P	HSCR	(9, 38)
R180	D2	Unk	HSCR	(9)
C197	D2	C197Y	HSCR	(9, 38)
V202L	D2	V202L	HSCR	Clinvar VCV000599632.3
R231	D2	R231H	HSCR	(9, 24)
E235	D2	E235K	HSCR	(31)
E251	D2	E251K	HSCR	(9, 32)
V262	D2	V262A	HSCR	(32)
D264	D3	D264K	HSCR	(24)
D267	D3	D267N	HSCR	(32)
T278	D3	T278N	HSCR	Clinvar VCV000041846.35 and (9)
R287	D3	Unk	HSCR	Clinvar VCV000560625.3 and (9)
E289	D3	E289Q	HSCR	(32)
D300	D3	D300K	HSCR	(9, 24)
R313	D3	R313Q	HSCR	(9, 32, 38)
S316	D3	S316I	HSCR	(9, 38)
R330	D3	R330W	HSCR	(9, 32)
W324	D3	W324C	HSCR	(32)
S339	D3	S339L	HSCR	(9, 38)
D353	D3	D353Y	HSCR	(9, 38)
N359	D3	Unk	HSCR	(9)
C570	D3	C570W	HSCR	(32)
R360	D3	R360Q	HSCR	(9, 38)
N361	D3	N359K	HSCR	(31)
F393	D3	F393L	HSCR	(9, 32)
N394	D3	Unk	HSCR	(9)
V397	D3	Unk	HSCR	(9)
P399	D3	P399L	HSCR	(9, 32)
V412	D4	V412M	HSCR	(9, 38)

G423	D4	G423R	HSCR	(9, 38)
D469	D4	D469N	HSCR	(5)
R475	D4	R475Q/W	HSCR	(9, 32)
E480	D4	E480K	HSCR	(9, 38)
C558	D5	C558Y	HSCR	(29)
Q576	D5	Q576P	HSCR	(31)
D584	D5	D584G	HSCR	(32)
C585	D5	C585R	HSCR	Clinvar VCV000695031.3
G588	D5	G588D	HSCR	(32)
E595	D5	E595G	HSCR	(9)
F555	D5	F555C	HSCR+MEN2A	Clinvar VCV001777553.4 and (9)
C609	D5	C609R/S/G/F/Y/W	HSCR+MEN2A	**ME2A:** , C609R Clinvar VCV000013944.13, C609S Clinvar VCV001372611.7, C609G Clinvar VCV003254419.1, C609F Clinvar VCV000038284.17, C609Y Clinvar VCV000013933.45, Clinvar VCV002497944.1 and (36) **HSCR+MEN2A:** C609G (24), C609W (9) , , ,
C611	D5	C611R/S/G/F/Y/W	HSCR+MEN2A	**MEN2A:** C611R Clinvar VCV000024896.9, C611S Clinvar VCV001780971.4, C611G Clinvar VCV000024897.12, C611F Clinvar VCV000024899.13, C611Y Clinvar VCV000024898.31, C611W ClinvarVCV000013913.6 and (9, 36) **HSCR+MEN2A:** C611S (22, 27, 36)
C618	D5	C618R/S/G/F/Y	HSCR+MEN2A	**MEN2A:** C618R Clinvar VCV000013929.33, C618S Clinvar VCV000038601.25 and VCV000013914.31, C618G Clinvar VCV000013905.89, C618F Clinvar VCV000024902.18, C618Y Clinvar VCV000024901.21 and (9, 36) **MEN2A+HSCR:** C618R (22), C618S (23)
C620	D5	C620R/S/G/F/Y/W	HSCR+MEN2A	**MEN2A:** C620R Clinvar VCV000013915.37, C620R Clinvar VCV000013943.20, C620S Clinvar VCV000038602.12, C620G, Clinvar VCV000024905.20, C620F Clinvar VCV000013928.32, C620Y Clinvar VCV000013916.24, C620W Clinvar VCV000013934.14 and (9, 36) **HSCR+MEN2A:** C620R (22, 28), C620S (25), C620W (23)
V292	D3	V292M	MEN2A	Clinvar VCV000024880.54 and (33, 38)
G321	D3	G321R	MEN2A	(35)
C515	D5	C515S	MEN2A	(35)
C531	D5	Unk	MEN2A	(9)
G533	D5	G533C	MEN2A	Clinvar VCV000013950.30 and (3, 35)
C541	D5	C541Y	MEN2A	Clinvar VCV002133665.2
R600	D5	R600Q	MEN2A	(35)
K603	D5	K603E	MEN2A	(35)
Y606	D5	Y606C	MEN2A	(35)
C630	D5	C630/F/R/S/Y	MEN2A	(9, 36)
C634	D5	C634F/R/W/S/Y/G	MEN2A	(3, 9, 36)

Table 1. RET missense mutations that lead to HSCR and/or MEN2A. 70 RET positions linked to HSCR, MEN2A, or both upon mutation. Mutations to uncharacterized residues are marked as "Unk".

We mapped all 70 positions onto the 3D structure of the extracellular region of RET, classified as LOF, GOF, or both (Fig. 1A). The 59 positions associated with LOF mutations are distributed across all five domains, but with fewer positions in D4 and in the CRD D5 (Fig. 1B). Notably, the five RET mutations

that can lead to both diseases are only in D5 (Fig. 1C) while GOF-only mutations are predominantly in the CRD D5 – 14 out of 16 (Fig. 1D). We note that almost all of the 16 GOF mutations (Fig. 1E) seem to be clustered in positions located in a specific region of D5 (Fig. 1C,D).

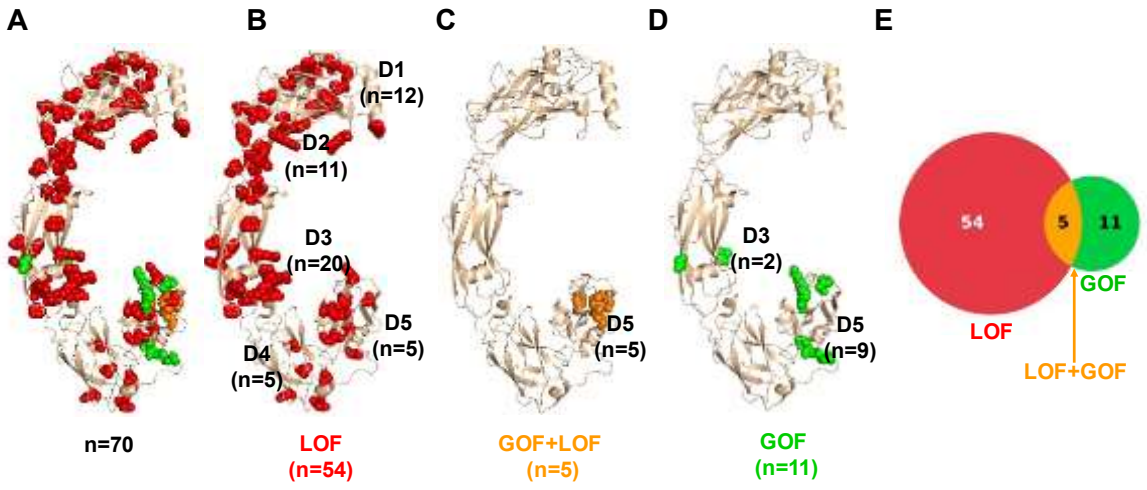

Figure 1. Positions in the extracellular side of RET that are associated with HSRC and/or MEN2A. A. GOF and LOF mutations across 70 RET positions that are associated with HSRC and MEN2A, respectively. Positions that can lead only to HSRC (LOF) are shown as red spheres, positions that lead only to MEN2A (GOF) are shown as green spheres, and positions associated with both diseases are shown as orange spheres. **B.** LOF-positions that are associated with HSRC only, colored as in A**. C.** "Janus" positions that are associated with both HSRC (LOF) and MEN2A (GOF), colored as in A. **D.** GOF positions associated only with MEN2A, colored as in A. **E.** A Venn diagram showing the number of the LOF- and GOF-associated positions shown in panels A-D.

Predicted mechanistic principles for LOF effects in RET positions associated with HSCR

We hypothesized that mutations that reduce downstream signaling by RET or that destabilize the protein and thereby reduce its amount in cells can lead to RET LOF and thereby to HSCR disease. Mutations in RET positions involved in protein-protein interactions with co-receptors or ligands may impair RET activation and reduce downstream signaling. On the other hand, mutations that, by diverse mechanisms, destabilize RET, can lead to LOF. One such mechanism may involve calcium-binding, as calcium ions were shown to stabilize RET tertiary structure. Therefore, mutations in positions crucial for calcium-binding may destabilize the receptor and disrupt its folding. Alternatively, substitutions of RET positions that are part of its protein core and stabilize its tertiary structure could also disrupt folding and structural integrity. Lastly, mutations in RET that disrupt strong intramolecular interactions may also affect its tertiary structure. Therefore, we mapped which residues mediate interactions between RET and its co-receptors and ligands, participate in the calcium-binding sites, are buried in the protein core, introduce a specific disruption of RET

structures via amino acids with unique physico-chemical properties such as glycines or prolines, or might affect RET stability by disrupting strong intramolecular interactions (Supplementary Fig. 1, Fig. 2).

To map positions that can affect protein-protein interactions, we analyzed the experimental structures of RET in complex with each of the following four co-receptors and four ligands: RET with GFRα1 and GDNF (PDB ID 6Q2J), RET with GFRα3 and Artemin (PDB ID 6Q2S), RET with GFRAL and GDF15 (PDB ID 6Q2N), RET with GFRα2 and Neurturin (PDB ID 6Q2O), and the 4:4:4 structure of RET with GFRα2 and Neurturin (PDB ID 6Q2R) (9). RET residues located within 5 Å of the interface with co-receptors (Fig. 2A) or ligands (Fig. 2B) were classified as "interacting residues". We identified 43 RET residues that are within 5Å of a co-receptor and 24 RET residues that are within 5Å of a ligand (Supplementary Fig. 2, Fig. 2A,B). RET residues that interact with co-receptors are mostly in domains D1 and D2, with only a few residues in D5. In contrast, RET residues that interact with ligands are only in domains D4 and D5. However, among the 59 HSCR-associated RET positions that presumably reduce activity upon mutation, only four are in proximity to RET co-receptors, R77, H114, R175, and D353 (Fig. 2A) and only one position associated with HSCR, E595, is within 5Å of a ligand (Fig. 2B).

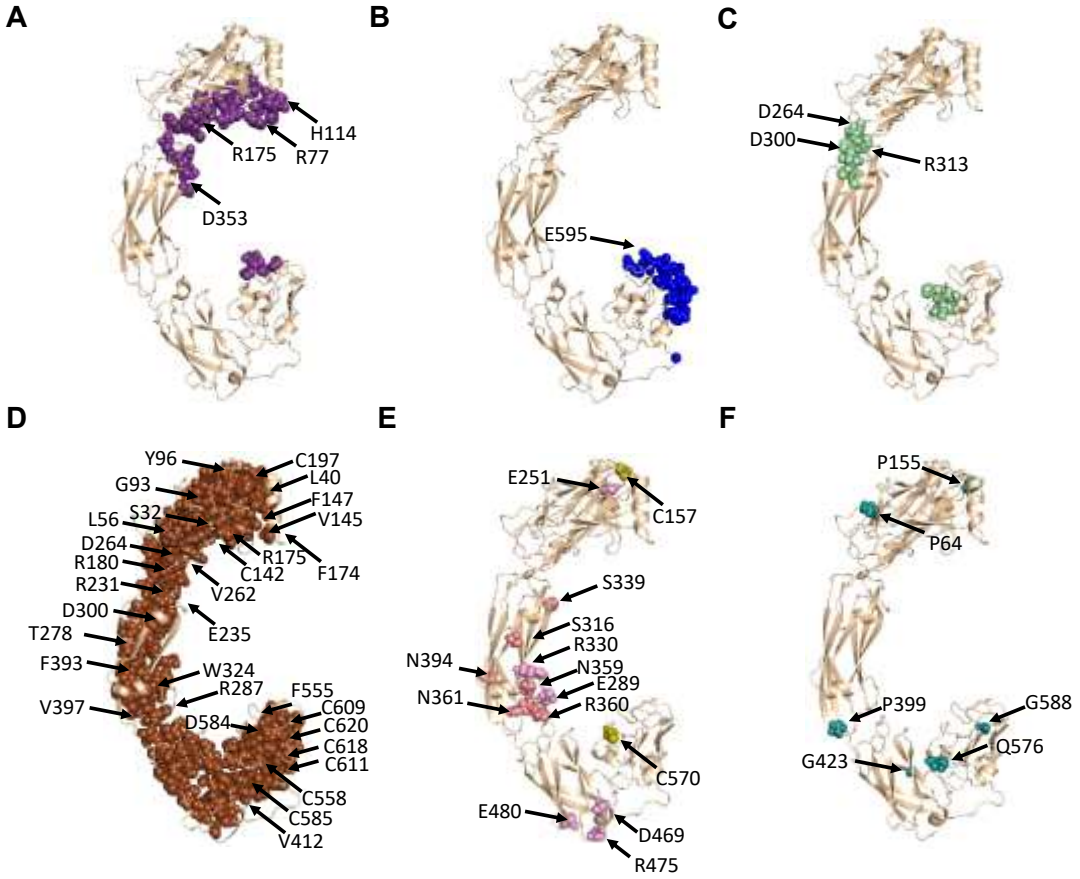

Figure 2. Predicted structural basis for RET LOF mutations leading to HSCR disease. A. 43 RET residues predicted to interact with RET co-receptors, shown as purple spheres. The four mutated positions among these residues that lead to HSCR are labeled with black arrows. **B.** 24 RET residues predicted to interact with RET ligands, shown as blue spheres. The mutated position that leads to HSCR is labeled with a black arrow. **C.** 19 RET residues that bind calcium ions, shown as light green spheres. The three mutated positions among these residues that lead to HSCR are labeled with black arrows. **D.** 222 RET residues that are classified as "buried" in the protein core (rASA ≤ 5% or BSA ≥ 200 Å^2)**,** shown as brown spheres. The 31 mutated positions among these residues that lead to HSCR are labeled with black arrows. **E.** 14 RET residues that lead to HSCR and stabilize RET tertiary structure via strong intramolecular interactions, labeled with black arrows and shown as spheres colored light pink (participating in hydrogen bonds), pink (participating in salt-bridges), and olive (forming intermolecular disulfide bonds). **F.** Six RET positions that involve mutations to/from prolines/glycines, thereby destabilizing RET tertiary structure. labeled with black arrows and shown as teal spheres.

We next mapped positions involved in calcium-binding (Fig. 2C), which were shown to be important for RET stability (9). Out of the 22 RET positions involved in stabilizing the tertiary structure of RET through calcium-binding, only two positions, D264 and D300, were associated with HSCR. A third HSCR-associated residue, R313, is immediately adjacent to the calcium-binding Y314, and was therefore also classified as part of the calcium binding site. Therefore, only three of the 59 LOF RET positions participate in calcium binding and likely lead to HSCR through destabilization of RET tertiary structure.

To map RET residues that are part of the core of the protein, we identified positions that are sufficiently buried in the protein hydrophobic core, where a mutation is more likely to perturb RET folding and destabilize it. To quantify the burial of each RET residue, we used an approach developed in our previous study (39). We measured the accessible surface area (ASA) of each RET residue using surfv (40). Then, we calculated the relative accessible surface area (rASA) for each residue by dividing the ASA value by the maximal empirical ASA for each residue, taken from (41). We also calculated the buried surface area (BSA) for each residue – as the difference between the maximal empirical ASA value and the calculated ASA for each residue. We classified RET residues as "buried" when either their rASA values are ≤15% or their BSA values are > 200Å^2, with all other residues considered "exposed". This burial-based analysis showed that out of 594 residues in RET extracellular region, 222 residues were classified as "buried" and 372 were classified as "exposed" (Fig. 2D). We assume that these 222 buried residues are much more likely to destabilize RET tertiary structure upon mutations that change the physico-chemical properties or size of the residue and thereby decrease RET levels in the relevant tissues. Indeed, 31 out of 59 RET LOF positions involve buried residues. We also note that two calcium-binding residues and one co-receptor binding residue are also classified as buried, meaning that a mutation of these residues can lead to LOF due to multiple causes simultaneously. Taken together, 32 of the 59 LOF mutations in RET destabilize its tertiary structure by affecting the protein core and/or calcium-binding, and only four residues reduce its activity solely via protein-protein interactions with ligands or co-receptors. This leaves 23 RET positions that were not classified under the categories above.

14 out of these 23 RET positions associated with HSCR participate in intramolecular interactions that can affect RET stability and folding (Fig. 2E). Visual inspection identified that two of these 14 HSCR-associated RET positions (C157 and C570) are cysteine residues that participate in intramolecular disulfide bonds with residues C197 and C585, respectively. The other 12 RET residues participate in intramolecular hydrogen bonds or salt-bridges. E251, S316, and E480 form intramolecular hydrogen bonds with R205, H333, and T503, respectively. R330 forms an intramolecular salt-bridge network with three negatively charged residues – E289, D290, and E332. Another salt-bridge is between D469 and R475, and two hydrogen bond networks are formed between S339, N359, R360, and N361, and between N394, Q371, and H392.

Six of the remaining nine HSCR-associated RET positions involve substitutions either to or from proline or glycine (Fig. 2F). Such mutations are likely to affect the RET stability because of the extreme backbone flexibility or rigidity associated with glycines and prolines, respectively (42, 43). P64 is in the middle of a loop next to a buried residue, F66. Consistent with previous findings (5), we predict that the HSCR-associated P64L mutation may lead to steric interference with F66, thereby affecting RET tertiary structure and stability. Similar to P64, P155 is in a loop and is also located between two buried residues, L160 and Y41. Likewise, P339 is adjacent to the buried residue V400.

Mutations in either of these prolines will likely perturb these regions. Reciprocally, Q576, which is also positioned in the middle of a loop in the D4-D5 interface, is substituted by proline in an HSCR-associated mutation, thereby increasing conformational rigidity and likely affecting RET tertiary structure negatively. In contrast, the G423R mutation inserts both a charge and a significant steric bulk next to two large residues, K424 and Q421, thereby likely perturbing RET tertiary structure. We note that five additional residues, that were classified above as buried or involved in protein-protein interactions, also involve mutations to or from prolines/glycines. L40P, G93S, and D584P are mutations in buried residues, R175P is a mutation in a residue that is both buried and predicted to bind co-receptors, while E595P is a mutation in a residue predicted to bind ligands. Therefore, these positions can also contribute to HSCR through two or three distinct mechanisms.

Overall, our analysis reveals that perturbing RET residues that stabilize its folding and tertiary structure seem to be the main structural basis for HSCR-associated mutations – providing mechanistic explanations for 52 out of the 59 HSCR-associated LOF mutations. Only four mutations appear to contribute to HSCR by impacting solely protein-protein interactions. Only three HSCR-associated positions out of the 59 were not predicted by our analysis to affect protein-protein interactions nor protein stability – R57, V202, and D267. These are "exposed" residues located in D2, D3, or D4, not predicted to interact with ligands or co-receptors. R57 is positioned in the middle of a loop, directly adjacent to proline and arginine residues. It is possible that a mutation of this charged arginine to a bulky tryptophan will alter local intra-molecular interactions and thereby global stability. In contrast, V202 and D267 are immediately adjacent to residues that are buried within the protein core. Specifically, V202 is immediately adjacent to the buried residue I200, while D267 is immediately adjacent to D266 and E178, and is also close to E265, all of which are buried yet charged residues. It is therefore plausible that mutations in these two residues will indirectly affect these buried residues and thereby RET tertiary structure and stability.

Predicted mechanistic principles for GOF effect in in RET positions associated with MEN2A

Out of the 16 gain-of-function (GOF) positions identified in RET (Table 1), 12 mutations involve substitutions to or from cysteine residues that participate in disulfide bonds within the RET CRD (D5). We hypothesize that these substitutions therefore lead to an unpaired cysteine in D5 that can form an intermolecular disulfide bond with its unpaired counterpart across the RET dimer interface. Such an intermolecular disulfide bond will lead to aberrant RET homodimerization and thus to increased signaling that presumably underlies oncogenesis. Nine of these 12 mutations change a cysteine residue to a non-cysteine residue (Fig. 3A). Interestingly, the remaining three out of 12 GOF mutations, F555C, G533C, and Y606C, also lead to an unpaired cysteine but via a reciprocal route; instead of mutating half of an existing intramolecular disulfide bridge, they introduce a novel cysteine, which will therefore be unpaired (Fig. 3B).

In most GOF cases in patients, we find substitutions in either cysteine of the same disulfide bond. However, in the disulfide bond between C541 and C519 only mutations in C541 have been observed, resulting in C519 becoming unpaired (Fig. 3C). All of these unpaired cysteines can potentially form an intermolecular disulfide bond with its counterpart across the dimer interface, as detailed above – similarly leading to RET homodimerization that can be independent of ligands and co-receptors. Indeed, we observed that all of the 12 unpaired cysteines that result from a GOF mutation are clustered around the same outer-facing side of D5 (Fig, 3A,B), supporting the intermolecular dimerization hypothesis. To test this hypothesis, we modeled the D5-D5 dimer of RET using AlphaFold3 (44). This model shows that all 12 unpaired cysteine positions are in the region of D5 that faces the predicted homo-dimerization interface – 11 of the 12 unpaired GOF cysteine residues mentioned above are ≤ 9Å of the dimer interface and F555C is ~12Å from the dimer interface (Fig. 3C). In contrast, almost all cysteine residues in D5 that have not been associated with MEN2A are farther away from the predicted D5-D5 dimer interface, at distances of 13-18Å (Fig. 3D). The only exception is C528 that is 7Å from the predicted D5-D5 dimer interface (Fig. 3E); this residue forms an intramolecular disulfide bond with C558 but has not been reported to be associated with MEN2A. Therefore, this location suggests that a mutation of C558 may lead to a GOF and to MEN2A.

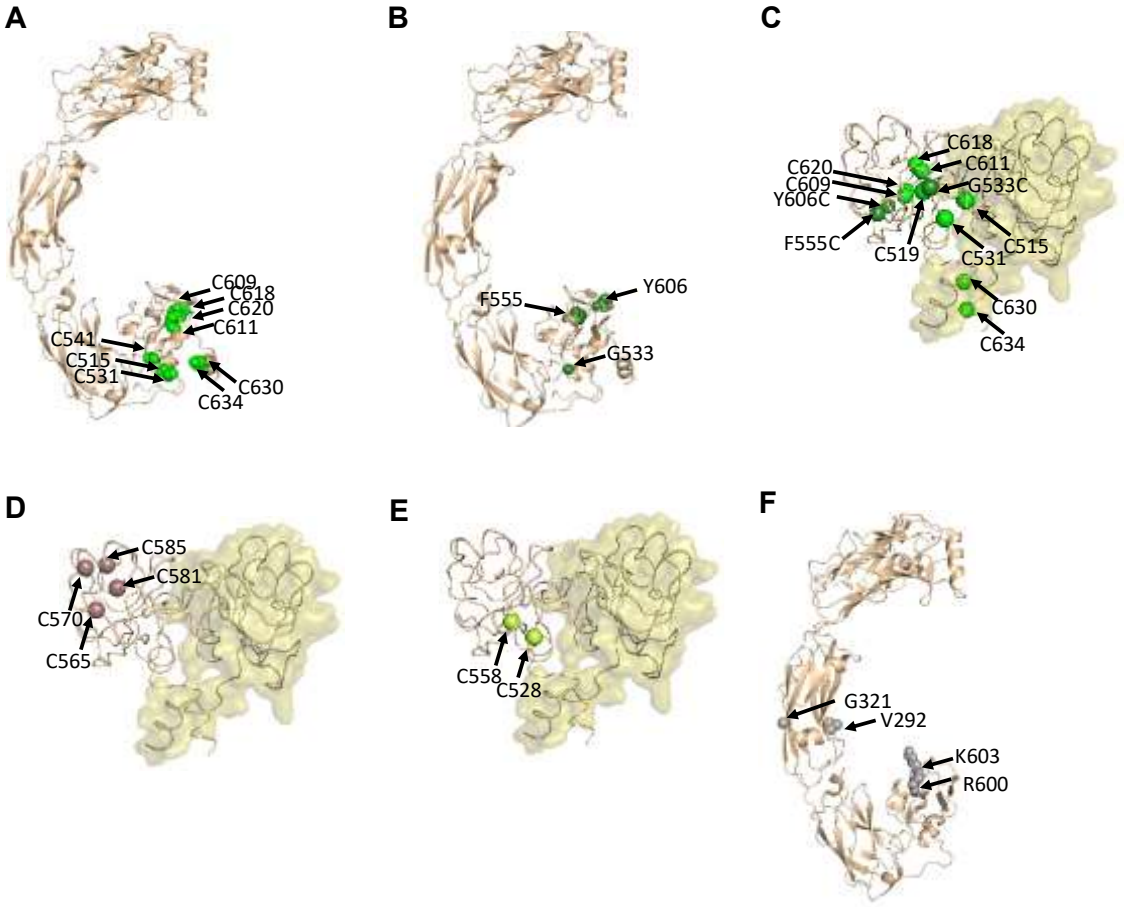

Figure 3. Predicted structural basis for RET GOF mutations that lead MEN2A. A. Nine cysteine residues which, upon mutation to a non-cysteine residue, result in an unpaired cysteine – shown as green spheres. **B.** Three non-cysteine residues which, upon mutation to a cysteine, also result in an unpaired cysteine – shown as dark green spheres. **C.** Positions of the 12 unpaired cysteines resulting from the mutations described in A and B, mapped onto one of the monomers in an Alphafold3 model of the D5-D5 dimer. Cα atoms are shown as spheres and colored as in A and B. Note that eight of these 12 mutations involve both cysteines across the same intramolecular disulfide bond. **D.** Four cysteine residues in D5 that have not been associated with MEN2A, mapped onto one of the monomers in the model (from C) of the D5-D5 dimer, with the Cα atoms shown as dark pink spheres. **E.** Positions of the two cysteine residues in D5 that are predicted to be close to the predicted D5-D5 interface but have not been associated with MEN2A, with their Cα atoms shown as light green spheres. **F.** Four RET positions that lead to GOF upon mutation but do not involve substitutions to or from cysteines, shown as grey spheres.

The remaining four out of the reported 16 GOF mutations do not lead to unpaired cysteines and therefore contribute to MEN2A through a different mechanism (Fig. 3F). These mutations are in four positions located in domains D3 (V292M, G321R) and D5 (R600W, K603E). V292 and G321 are exposed residues that are far from both the ligand-binding and co-receptor binding regions. The V292M mutation does not change the physico-chemical properties of this residue sufficiently to explain the GOF phenotype. However, the G321R mutation, as it does involve a dramatic change in the physico-chemical properties of the residue, might induce an intramolecular conformational change that can increase RET activation. As for the two residues in D5, both R600 and K603 are close to the predicted D5-D5 dimer interface (Fig. 3C), and mutations of these positions might therefore also promote D5-D5 dimerization, albeit through a non-

covalent mechanism that probably requires a conformational change. We note that K603 is within 5Å of the co-receptor, but visual inspection did not identify RET/co-receptor interactions that might be enhanced as a result of this mutation and can explain the GOF phenotype. We also note that the R600 residue, which is also positioned in close proximity to co-receptor binding residues, has been reported to be substituted in extremely rare cases (45), so perhaps the GOF phenotype is less pronounced. Overall, our analysis suggests that 14 out of the 16 GOF mutations associated with MEN2A may enhance D5-D5 dimerization and thereby enhance downstream signaling, thereby promoting tumorigenesis.

Predicted mechanistic explanation for the paradoxical LOF and GOF effects in RET Janus positions associated with both MEN2A and HSCR

Five disease-associated RET positions – F555, C609, C611, C618, and C620 – are linked to both HSCR and MEN2A (Table 1, Fig. 1). Our analysis revealed that all of these positions contribute to the stability of the tertiary structure of RET, as they are part of the protein core (Fig. 1D). Interestingly, it has been documented in patients that F555 is substituted exclusively to cysteine, while C609, C611, C618, and C620 are consistently substituted to the same specific amino acids – arginine, serine, glycine, phenylalanine, tyrosine, and tryptophan (Table 1). Therefore, all five Janus mutations involve very substantial changes in the physico-chemical properties following the mutation that are quite likely to affect protein folding. Overall, these substantial substitutions, particularly in core residues, may destabilize the tertiary structure of RET, impairing the embryonic development of the enteric nervous system and contributing to HSCR. On the other hand, mutations in all five residues, which are all located at the periphery of D5, result in an unpaired cysteine (Fig. 3). F555C introduces a new and unpaired cysteine to the D5-D5 interface, whereas the C609, C611, C618, and C620 positions are all part of intramolecular disulfide bonds, so their mutation leads to an unpaired cysteine. These unpaired cysteines can lead to aberrant RET dimerization as detailed above, contribute to MEN2A, which unlike HSCR manifests later in life.

DISCUSSION

We curated a comprehensive dataset of 70 positions in the extracellular domains of RET that have been associated with either HSCR disease, MEN2A, or both – integrating entries from both the Clinvar database and curated data from the literature (Table 2). This dataset provides a reconciled and clarified view of RET mutations that is more comprehensive than previously published and, to our knowledge, is the largest dataset of disease-associated RET mutations. We also provide the predicted structural and mechanistic based for the LOF and GOF phenotypes for almost all of these disease-associated mutations across these 70 positions.

Hirschsprung	MEN2A		
LOF	LOF+GOF	GOF	
S32^{b}, L40^{b}, L56^{b}, R57, P64$^{G/P}$, R77^{CR}, G93^{b}, Y96^{b}, H114^{cr}, C142^{b}, V145^{b}, F147^{b}, P155$^{G/P}$, C157^{s-s}, F174^{b}, R175$^{cr/b}$, R180^{b}, C197^{b}, V202, R231^{b}, E235^{b}, E251^{SB},V262^{b}, D264$^{Cal/b}$, D267, T278^{b},R287^{b}, E289^{SB}, D300$^{Cal/b}$, R313^{Cal}, S316^{HB}, W324^{b}, R330^{SB}, S339^{HB},D353^{CR}, N359^{HB}, R360^{HB}, N361^{HB}, F393^{b}, N394^{HB}, V397^{b}, P399$^{G/P}$, V412^{b}, G423$^{G/P}$, D469^{SB}, R475^{SB}, E480^{HB}, C558^{b}, C570^{s-s}, Q576$^{G/P}$, D584^{b},C585^{b}, G588$^{G/P}$, ,E595^{L}	F555$^{b/cys}$, C609$^{b/cys}$, C611$^{b/cys}$, C618$^{b/cys}$, C620$^{b/cys}$		Enteric ganglia
		V292, G321, C515^{cys}, C531^{b}, G533^{b}, C541^{b}, R600, K603, Y606, C630^{cys}, C634^{cys}	Thyroid

b – buried residue
G/P – glycine/proline in key positions
CR – interactions with co-receptor
S-S – intramolecular disulfide bond
Cal – calcium binding
HB – intramolecular hydrogen bond
SB – intramolecular salt-bridge
L – interactions with ligand
Cys – mutation of/to cysteine

Table 2. RET missense mutations that lead to HSCR, MEN2A, and both diseases, categorized by their predicted effect. Mutations are classified based on intra- and intermolecular interactions, burial status, calcium-binding participation, and type of residue substitution.

Our structure-based analysis of this dataset suggests that ~90% of RET extracellular positions that are associated with HSCR may destabilize the tertiary structure of the RET receptor. Only a small minority of mutations apparently lead to LOF by attenuating protein-protein interactions with the co-receptors or ligands that activate RET. A previous analysis of disease-associated missense mutations across different proteins reached similar rations, estimating that ~80% of such mutations impair protein stability, with the majority of these involving residues buried in the protein core or disruption of strong intramolecular interactions (46). Indeed, our predictions suggest that the most common reason for RET destabilization and therefore LOF is perturbation of the protein core. However, disruption of strong intramolecular interactions that include salt-bridges, hydrogen bonds, and disulfide bonds, or the calcium- binding sites that were shown

to stabilize RET tertiary structure, can also lead to LOF and HSCR disease. An additional mechanism for RET destabilization can occur when residues in the protein core or in intramolecular interfacial positions are substituted to or from glycines or prolines, due to their unique effects on protein 3D structure. Mutations involving glycine, with its extreme flexibility, or proline, which rigidifies the protein backbone and can adopt unique conformations, can therefore disrupt local folding and global stability. Notably, some HSCR-associated positions can lead to LOF via two or even three different mechanisms. Specifically, mutations in D264 and D300 can destabilize RET not only by disrupting calcium binding, but also by compromising the protein core. Similarly, mutations in the buried residues L40, G93, and E584 can thus affect both the protein core directly, or through the effects of a substitution to or from glycine. The buried residue R175 is predicted to affect interactions with co-receptors but also involves a substitution to proline. Likewise, the E595P mutation is also predicted to affect ligand binding. Overall, we predict that mutations destabilizing the protein tertiary structure are most likely to lead to HSCR.

In contrast to LOF mutations observed in HSCR, GOF mutations associated with MEN2A predominantly involve mutations to or from cysteine residues in the CRD D5 domain of RET. Such mutations lead to an unpaired cysteine, close enough to the predicted D5-D5 dimer interface so to presumably form an intermolecular disulfide bond across the homodimer. This, in turn, will lead to aberrant RET dimerization and activation that is less dependent on co-receptors and ligands. Additional GOF mutations that do not involve unpaired cysteines are nevertheless also located in membrane-proximal domains, suggesting a common GOF mechanism of enhancing RET homodimerization via these membrane-proximal domains. Indeed, these membrane-proximal interactions resemble other homotypic protein-protein interactions, and in particular the juxta-membrane homodimerization shown to occur between other RTKs – PDGFR, KIT, and VEGFR receptors (47-49).

We also identified four cysteine residues in D5 that are not known to lead to GOF upon mutation. Our analysis suggested this is due to their more distant location from the predicted dimer interface, presumably not allowing the formation of an intermolecular disulfide bond across the dimer interface. On the other hand, we predict that a mutation in C558 can lead to the resulting unpaired C528 contributing to MEN2A, which has not been identified in patients to date. Notably, four GOF mutations do not involve unpaired cysteines, yet still lead to MEN2A. However, two of these are also located in D5 where they can enhance D5-D5 dimerization and a third might involve activity-enhancing conformational changes, also supporting a general mechanism of GOF mutations leading to constitutive activation by enhancing RET dimerization proximal to the membrane.

Our structural analysis suggests how the five RET “Janus positions” lead to both HSCR and MEN2A. These positions are all buried within D5 and the substitutions found in HSCR involve dramatic changes to the physico-chemical properties of these residues, explaining the LOF phenotype by the

substantial destabilization of RET. On the other hand, as all five mutations result in an unpaired cysteine that can form a D5-D5 intermolecular disulfide bond, the GOF phenotype is explained as detailed above. The concurrent prediction though is that the LOF phenotype is partial, allowing sufficient levels of RET to facilitate abnormal dimer formation later in life. More broadly, the dual pathogenic phenotype of mutations in RET highlights the need for precisely tailored therapeutic strategies to address RET-associate diseases in the thyroid and enteric ganglia. Stabilizing the RET protein might help mitigate the LOF HSCR phenotype in the intestine, while preventing aberrant receptor activation is crucial for treating MEN2A. Furthermore, the pivotal role of cysteine residues within D5 in RET regulation underscores the potential of developing covalent inhibitors targeting intermolecular disulfide bonds as a therapeutic strategy for MEN2A.

MATERIALS AND METHODS

Protein structures

The following 3D structures were used in our analysis and visualization of RET with different co-receptors and ligands (with PDB IDs): RET with GFRα1 and GDNF (PDB ID 6Q2J), RET with GFRα3 and Artemin (PDB ID 6Q2S), RET with GFRAL and GDF15 (PDB ID 6Q2N), RET with GFRα2 and Neurturin (PDB ID 6Q2O), and 4:4:4 structure of RET with GFRα2 and Neurturin (PDB ID 6Q2R). Short segments of RET (residues 129-137, 207-211, 246-251, and 379-387) that are missing in PDB entries 6Q2R, 6Q2O, 6Q2S, and 6Q2S were modeled using Loopy (50) and partial or missing side chains were modeled using Scap (50). We used the Alpahfold3 server (https://alphafoldserver.com/) (44) to predict the 3D structure of the RET D5-D5 dimer. We used PyMol (https://www.pymol.org/) for 3D structural visualization.

Structure-based prediction of mutation effect

RET residues ≤ 5 Å of the interface with co-receptors and ligands were classified as "interacting residues". Calcium-binding residues were classified according to (9). To map residues that can affect the tertiary structure of a protein due to being buried in the protein core, we followed the methodology described previously (39). We measured the accessible surface area (ASA) of each residue using surfv (40). We calculated the relative accessible surface area (rASA) for each residue by dividing its ASA by the maximal empirical ASA, taken from the values calculated from a large dataset of structures culled from the PDB by (41), who followed the approaches laid out by Rose et al., (51) and Miller et al. (52). Buried surface area (BSA) was calculated by subtracting the ASA of each residue from the maximal empirical ASA value for that residue. A residue was classified as "buried" when its rASA ≤ 15% or when its BSA ≥ 200 $Å^2$. Residues involved in intramolecular salt bridges or hydrogen bonds were identified by visual inspection.

REFERENCES

1. Schuchardt A, D'Agati V, Larsson-Blomberg L, Costantini F, Pachnis V. Defects in the kidney and enteric nervous system of mice lacking the tyrosine kinase receptor Ret. Nature. 1994;367(6461):380-3.
2. Durbec PL, Larsson-Blomberg LB, Schuchardt A, Costantini F, Pachnis V. Common origin and developmental dependence on c-ret of subsets of enteric and sympathetic neuroblasts. Development. 1996;122(1):349-58.
3. Takahashi M. RET receptor signaling: Function in development, metabolic disease, and cancer. Proc Jpn Acad Ser B Phys Biol Sci. 2022;98(3):112-25.
4. Swenson O. Hirschsprung's disease: a review. Pediatrics. 2002;109(5):914-8.
5. Anders J, Kjar S, Ibanez CF. Molecular modeling of the extracellular domain of the RET receptor tyrosine kinase reveals multiple cadherin-like domains and a calcium-binding site. J Biol Chem. 2001;276(38):35808-17.
6. Takahashi M. The GDNF/RET signaling pathway and human diseases. Cytokine Growth Factor Rev. 2001;12(4):361-73.
7. Amiel J, Sproat-Emison E, Garcia-Barcelo M, Lantieri F, Burzynski G, Borrego S, et al. Hirschsprung disease, associated syndromes and genetics: a review. J Med Genet. 2008;45(1):1-14.
8. Santoro M, Carlomagno F. Central role of RET in thyroid cancer. Cold Spring Harb Perspect Biol. 2013;5(12):a009233.
9. Li J, Shang G, Chen YJ, Brautigam CA, Liou J, Zhang X, et al. Cryo-EM analyses reveal the common mechanism and diversification in the activation of RET by different ligands. Elife. 2019;8.
10. Regua AT, Najjar M, Lo HW. RET signaling pathway and RET inhibitors in human cancer. Front Oncol. 2022;12:932353.
11. Durbec P, Marcos-Gutierrez CV, Kilkenny C, Grigoriou M, Wartiowaara K, Suvanto P, et al. GDNF signalling through the Ret receptor tyrosine kinase. Nature. 1996;381(6585):789-93.
12. Kotzbauer PT, Lampe PA, Heuckeroth RO, Golden JP, Creedon DJ, Johnson EM, Jr., et al. Neurturin, a relative of glial-cell-line-derived neurotrophic factor. Nature. 1996;384(6608):467-70.
13. Baloh RH, Tansey MG, Lampe PA, Fahrner TJ, Enomoto H, Simburger KS, et al. Artemin, a novel member of the GDNF ligand family, supports peripheral and central neurons and signals through the GFRalpha3-RET receptor complex. Neuron. 1998;21(6):1291-302.
14. Goodman KM, Kjaer S, Beuron F, Knowles PP, Nawrotek A, Burns EM, et al. RET recognition of GDNF-GFRalpha1 ligand by a composite binding site promotes membrane-proximal self-association. Cell Rep. 2014;8(6):1894-904.
15. Bigalke JM, Aibara S, Roth R, Dahl G, Gordon E, Dorbeus S, et al. Cryo-EM structure of the activated RET signaling complex reveals the importance of its cysteine-rich domain. Sci Adv. 2019;5(7):eaau4202.
16. Manie S, Santoro M, Fusco A, Billaud M. The RET receptor: function in development and dysfunction in congenital malformation. Trends Genet. 2001;17(10):580-9.
17. Iwashita T, Murakami H, Asai N, Takahashi M. Mechanism of ret dysfunction by Hirschsprung mutations affecting its extracellular domain. Hum Mol Genet. 1996;5(10):1577-80.

18. Carlomagno F, De Vita G, Berlingieri MT, de Franciscis V, Melillo RM, Colantuoni V, et al. Molecular heterogeneity of RET loss of function in Hirschsprung's disease. EMBO J. 1996;15(11):2717-25.

19. Chappuis-Flament S, Pasini A, De Vita G, Segouffin-Cariou C, Fusco A, Attie T, et al. Dual effect on the RET receptor of MEN 2 mutations affecting specific extracytoplasmic cysteines. Oncogene. 1998;17(22):2851-61.

20. Kjaer S, Kurokawa K, Perrinjaquet M, Abrescia C, Ibanez CF. Self-association of the transmembrane domain of RET underlies oncogenic activation by MEN2A mutations. Oncogene. 2006;25(53):7086-95.

21. Asai N, Iwashita T, Matsuyama M, Takahashi M. Mechanism of activation of the ret proto-oncogene by multiple endocrine neoplasia 2A mutations. Mol Cell Biol. 1995;15(3):1613-9.

22. Mulligan LM, Eng C, Attie T, Lyonnet S, Marsh DJ, Hyland VJ, et al. Diverse phenotypes associated with exon 10 mutations of the RET proto-oncogene. Hum Mol Genet. 1994;3(12):2163-7.

23. Borst MJ, VanCamp JM, Peacock ML, Decker RA. Mutational analysis of multiple endocrine neoplasia type 2A associated with Hirschsprung's disease. Surgery. 1995;117(4):386-91.

24. Pelet A, Geneste O, Edery P, Pasini A, Chappuis S, Atti T, et al. Various mechanisms cause RET-mediated signaling defects in Hirschsprung's disease. J Clin Invest. 1998;101(6):1415-23.

25. Decker RA, Peacock ML. Occurrence of MEN 2a in familial Hirschsprung's disease: a new indication for genetic testing of the RET proto-oncogene. J Pediatr Surg. 1998;33(2):207-14.

26. Takahashi M, Iwashita T, Santoro M, Lyonnet S, Lenoir GM, Billaud M. Co-segregation of MEN2 and Hirschsprung's disease: the same mutation of RET with both gain and loss-of-function? Hum Mutat. 1999;13(4):331-6.

27. Nishikawa M, Murakumo Y, Imai T, Kawai K, Nagaya M, Funahashi H, et al. Cys611Ser mutation in RET proto-oncogene in a kindred with medullary thyroid carcinoma and Hirschsprung's disease. Eur J Hum Genet. 2003;11(5):364-8.

28. Arighi E, Popsueva A, Degl'Innocenti D, Borrello MG, Carniti C, Perala NM, et al. Biological effects of the dual phenotypic Janus mutation of ret cosegregating with both multiple endocrine neoplasia type 2 and Hirschsprung's disease. Mol Endocrinol. 2004;18(4):1004-17.

29. Kim JH, Yoon KO, Kim JK, Kim JW, Lee SK, Kong SY, et al. Novel mutations of RET gene in Korean patients with sporadic Hirschsprung's disease. J Pediatr Surg. 2006;41(7):1250-4.

30. Landrum MJ, Lee JM, Riley GR, Jang W, Rubinstein WS, Church DM, et al. ClinVar: public archive of relationships among sequence variation and human phenotype. Nucleic Acids Res. 2014;42(Database issue):D980-5.

31. Hofstra RM, Wu Y, Stulp RP, Elfferich P, Osinga J, Maas SM, et al. RET and GDNF gene scanning in Hirschsprung patients using two dual denaturing gel systems. Hum Mutat. 2000;15(5):418-29.

32. Kashuk CS, Stone EA, Grice EA, Portnoy ME, Green ED, Sidow A, et al. Phenotype-genotype correlation in Hirschsprung disease is illuminated by comparative analysis of the RET protein sequence. Proc Natl Acad Sci U S A. 2005;102(25):8949-54.

33. Castellone MD, Verrienti A, Magendra Rao D, Sponziello M, Fabbro D, Muthu M, et al. A novel de novo germ-line V292M mutation in the extracellular region of RET in a patient with phaeochromocytoma and medullary thyroid carcinoma: functional characterization. Clin Endocrinol (Oxf). 2010;73(4):529-34.

34. Castellone MD, Melillo RM. RET-mediated modulation of tumor microenvironment and immune response in multiple endocrine neoplasia type 2 (MEN2). Endocr Relat Cancer. 2018;25(2):T105-T19.

35. Krampitz GW, Norton JA. RET gene mutations (genotype and phenotype) of multiple endocrine neoplasia type 2 and familial medullary thyroid carcinoma. Cancer. 2014;120(13):1920-31.

36. Wells SA, Jr., Asa SL, Dralle H, Elisei R, Evans DB, Gagel RF, et al. Revised American Thyroid Association guidelines for the management of medullary thyroid carcinoma. Thyroid. 2015;25(6):567-610.

37. Yin L, Barone V, Seri M, Bolino A, Bocciardi R, Ceccherini I, et al. Heterogeneity and low detection rate of RET mutations in Hirschsprung disease. Eur J Hum Genet. 1994;2(4):272-80.

38. So MT, Leon TY, Cheng G, Tang CS, Miao XP, Cornes BK, et al. RET mutational spectrum in Hirschsprung disease: evaluation of 601 Chinese patients. PLoS One. 2011;6(12):e28986.

39. Bakhman A, Rabinovich E, Shlamkovich T, Papo N, Kosloff M. Residue-level determinants of angiopoietin-2 interactions with its receptor Tie2. Proteins. 2019;87(3):185-97.

40. Nicholls A, Sharp KA, Honig B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins. 1991;11(4):281-96.

41. Tien MZ, Meyer AG, Sydykova DK, Spielman SJ, Wilke CO. Maximum allowed solvent accessibilites of residues in proteins. PLoS One. 2013;8(11):e80635.

42. Levitt M. Conformational preferences of amino acids in globular proteins. Biochemistry. 1978;17(20):4277-85.

43. Bajaj K, Madhusudhan MS, Adkar BV, Chakrabarti P, Ramakrishnan C, Sali A, et al. Stereochemical criteria for prediction of the effects of proline mutations on protein stability. PLoS Comput Biol. 2007;3(12):e241.

44. Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. Nature. 2024;630(8016):493-500.

45. Saez ME, Ruiz A, Cebrian A, Morales F, Robledo M, Antinolo G, et al. A new germline mutation, R600Q, within the coding region of RET proto-oncogene: a rare polymorphism or a MEN 2 causing mutation? Hum Mutat. 2000;15(1):122.

46. Wang Z, Moult J. SNPs, protein structure, and disease. Hum Mutat. 2001;17(4):263-70.

47. Yang Y, Yuzawa S, Schlessinger J. Contacts between membrane proximal regions of the PDGF receptor ectodomain are required for receptor activation but not for receptor dimerization. Proc Natl Acad Sci U S A. 2008;105(22):7681-6.

48. Yang Y, Xie P, Opatowsky Y, Schlessinger J. Direct contacts between extracellular membrane-proximal domains are required for VEGF receptor activation and cell signaling. Proc Natl Acad Sci U S A. 2010;107(5):1906-11.

49. Yuzawa S, Opatowsky Y, Zhang Z, Mandiyan V, Lax I, Schlessinger J. Structural basis for activation of the receptor tyrosine kinase KIT by stem cell factor. Cell. 2007;130(2):323-34.

50. Xiang Z, Soto CS, Honig B. Evaluating conformational free energies: the colony energy and its application to the problem of loop prediction. Proc Natl Acad Sci U S A. 2002;99(11):7432-7.

51. Rose GD, Geselowitz AR, Lesser GJ, Lee RH, Zehfus MH. Hydrophobicity of amino acid residues in globular proteins. Science. 1985;229(4716):834-8.

52. Miller S, Janin J, Lesk AM, Chothia C. Interior and surface of monomeric proteins. J Mol Biol. 1987;196(3):641-56.