COMPUTATIONAL TARGET PROFILE

SAMD9L

A 1,584-Residue Disease Effector, Characterised Without A Released Structure

Three inherited diseases, one structureless giant. A huge multidomain effector behind ataxia-pancytopenia, monosomy-7 myelodysplasia and SCA49 — Astra maps its architecture, places its **disease variants** in context, and predicts an effector-core pocket.

UniProt Q8IVG5 · 1,584 aa · Soluble · IDG tier: T_{bio} · PDB: none released · July 2026

Have your own under-characterised target? Send a UniProt ID to contact@orbion.life for a preview like this one.

At A Glance

Class	Soluble (p=0.92), **non-enzyme** (p=1.00); a large **immune / signalling effector** — an antiviral, growth-restricting multidomain protein.
Disease Map	**Three Mendelian diseases, mapped.** Gain-of-function variants at 880 (ataxia-pancytopenia / monosomy-7 MDS) and 626 (SCA49) fall within the predicted effector core.
Effector Pocket	**High-confidence pocket (0.89).** A residue-level site in the central effector core — a candidate ligand / interaction site on a 1,584-residue protein.
Architecture	N-terminal SAM domain (14–79); a large, predominantly ordered (∼6% disorder) multidomain effector core (STAND / NTPase-like → TPR → OB).
Clean Signal	No membrane or amyloid signal.

PREDICTION CONFIDENCE

Figure 1. Model-reported confidence for the headline calls. The effector-core pocket (amber) is the load-bearing prediction — a targetable site on an unwieldy 1,584-residue disease protein.

1 The Gap

SAMD9L carries an outsized clinical burden for a protein no one has drugged. Germline **gain-of-function** variants cause a spectrum of severe inherited disease: **ataxia-pancytopenia syndrome** [3], **myelodysplasia with monosomy 7 / bone-marrow-failure predisposition** [4], and **spinocerebellar ataxia 49 (SCA49)** [5]. It is a large, antiviral, growth-restricting multidomain effector whose molecular mechanism is still being worked out [6]. And it is enormous — **1,584 residues** — which is exactly why it is hard to reason about.

Crucially, **no experimental structure is available in the public PDB** as of mid-2026 (the first cryo-EM structures of human SAMD9L were reported in a 2026 preprint, but the coordinates are not yet released). It remains understudied (IDG T_{bio}) [2]. So the practical problem is the same as for any dark giant: *where are the domains, where do the disease variants act, and is there a targetable site?* Everything below is computed from the canonical 1,584-residue sequence and derived structural predictions, with no experimental SAMD9L structure used as input.

2 Architecture And The Disease-Variant Map

Astra reads SAMD9L as a soluble, predominantly ordered multidomain effector: an N-terminal SAM domain, then a large central effector core. Placed on that map, the gain-of-function disease variants fall *within the effector core* — consistent with mutations that dial up the protein's growth-restricting activity — and the predicted pocket sits in the same region.

DOMAIN ARCHITECTURE, DISEASE VARIANTS AND PREDICTED POCKET

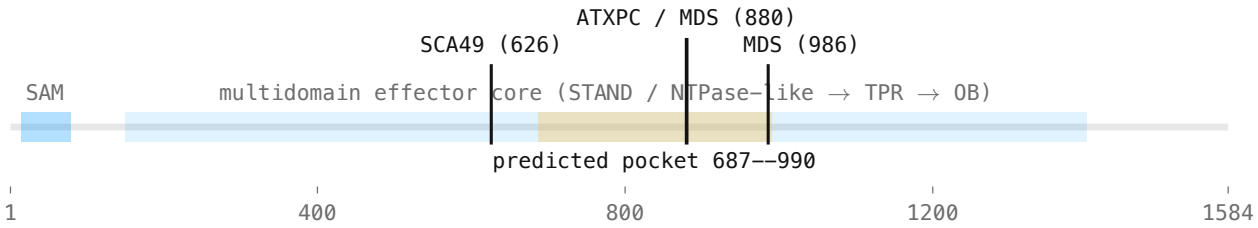

Figure 2. Architecture and disease map. An N-terminal SAM domain (blue) precedes a large multidomain effector core. The gain-of-function variants for ataxia-pancytopenia / monosomy-7 MDS (880) and SCA49 (626) fall within that core, and the predicted pocket (amber, 687–990) sits in the same region — a single, testable structural context for the disease biology.

3 Reading The Giant

For a 1,584-residue protein with no released structure, the most useful outputs are a domain map, a place to put the disease variants, and a targetable site. Astra delivers all three. It confirms a soluble, non-enzyme **immune/signalling effector**; it resolves the SAM domain and a large ordered effector core; and it places a **high-confidence pocket (0.89) in that core** (residues 687–990) — the region that also contains the disease variants. That gives a first, concrete hypothesis for *where* to intervene on a protein that is otherwise too large to reason about by eye.

PREDICTED EFFECTOR–CORE POCKET

Location	Central effector core, residues 687–990
Pocket residues	687–696, 743–749, 989–990
Confidence	HIGH (0.89); no known binder

A fold is coming; Astra adds the annotation. Even as the first SAMD9L structures arrive, a set of coordinates does not by itself tell you which domain carries a targetable pocket or how each clinical variant maps onto function. This profile is that annotation layer, available now from sequence. As a control, the same predictors recover the domain organisation and ligand sites on multidomain proteins whose structures *are* solved, so the SAMD9L calls are like-for-like predictions. The net effect: an unwieldy disease giant becomes a **ranked, testable map**.

4 Post-Translational And Structural Features

Astra's supporting calls are consistent with a large, ordered cytosolic effector:

- **Predominantly ordered (~6% disorder).** Despite its size, most of SAMD9L is foldable — so domain-by-domain expression is realistic.
- **N-terminal SAM domain (14–79).** A protein-interaction module; a clean, small unit to express and assay first.
- **No membrane or amyloid signal.** A soluble effector; the biology is intracellular.

5 Recommended Experimental Follow-Up

This turns an unwieldy giant into a ranked, executable plan: each prediction is paired with the experiment that would test it and the readout to watch for.

Prediction	Experiment	Readout
Effector-core pocket (687–990)	Fragment / ligand screen at the pocket	A chemical starting point on the core
Variants act in the effector core	Growth-restriction assay: 626 / 880 vs WT	Structural mechanism of gain-of-function
Domain map (SAM + core)	Domain-by-domain expression	Which domains fold independently
Immune / growth-restricting role	Antiviral / proliferation assays ± variants	Function of the core and its pocket

Its inherited bone-marrow-failure and ataxia phenotypes [3, 4, 5] indicate where a validated modulator would have the most clinical value.

6 Where This Fits In A Discovery Workflow

Team	What they get from this profile
Target biology / genetics	Structural context for each clinical variant
Structural biology	A domain map, construct boundaries, and a site to focus on
Medicinal chemistry	An effector-core pocket for docking and fragment work
BD / platform evaluation	Evidence Orbion can prioritise dark targets quickly

7 Methods

All predictions were generated with Orbion's Astra suite from the canonical SAMD9L sequence (UniProt Q8IVG5 [1]). Modules used, with method publications cited where available: **AstraSUIT2** and **AstraROLE2** [7] (membrane and functional class, oligomeric state), **AstraUNFOLD** (topology, disorder, aggregation), **AstraPTM2** [8] (modification sites) and **AstraBIND** [9] (binding pocket). Reported values are model outputs; model internals are out of scope. Confidence scores are model-estimated probabilities (0–1) that rank and gate each call — not calibrated rates of experimental success.

PREDICTION PROVENANCE

Target & input	SAMD9L · UniProt Q8IVG5, canonical isoform (1,584 aa)
Structure source	AlphaFold2 predicted model (no released experimental PDB entry)
Feature inputs	ESM-2 embeddings + AlphaFold2-derived structural features [7–9]
Models	AstraSUIT2 · AstraROLE2 · AstraUNFOLD · AstraPTM2 · AstraBIND
Binding-site model	GATv2 detector trained on >250k protein–ligand complexes [9]
Confidence & calibration	Per-call model probability; per-model calibration, thresholds and benchmarks reported in [7]–[9]

Scope And Limitations

- **Prediction, not experiment.** These are computational hypotheses to prioritise experiments — not a substitute for a structure or an assay. No result here has been validated in the wet lab.
- **Structures are imminent.** No SAMD9L coordinates are in the public PDB today, but the first cryo-EM structures have been reported in a 2026 preprint (release pending). This profile is the *functional / variant*

annotation, which a set of coordinates does not by itself provide.

- **Paralog caution.** SAMD9L's paralog SAMD9 already has experimental structures (e.g. a DNA-binding-domain entry) — these are a different gene and must not be read as SAMD9L structures.

8 References

[1] UniProt Consortium. UniProtKB entry Q8IVG5 (SAMD9L, human). uniprot.org.

[2] Pharos / Illuminating the Druggable Genome. SAMD9L target record — T_{bio}. pharos.nih.gov/targets/Q8IVG5.

[3] Chen D.-H., Below J.E., Shimamura A., et al. Ataxia-pancytopenia syndrome is caused by missense mutations in SAMD9L. *Am. J. Hum. Genet.* **98**(6), 1146–1158 (2016). doi:10.1016/j.ajhg.2016.04.009.

[4] Tesi B., Davidsson J., Voss M., et al. Gain-of-function SAMD9L mutations cause a syndrome of cytopenia, immunodeficiency, MDS and neurological symptoms. *Blood* **129**(16), 2266–2279 (2017). doi:10.1182/blood-2016-10-743302.

[5] Corral-Juan M., Serrano-Munuera C., Rábano A., et al. New spinocerebellar ataxia subtype caused by SAMD9L mutation triggering mitochondrial dysregulation (SCA49). *Brain Commun.* **4**(2), fcac030 (2022). doi:10.1093/braincomms/fcac030.

[6] Peng S., Meng X., Zhang F., et al. Structure and function of an effector domain in antiviral factors and tumor suppressors SAMD9 and SAMD9L. *Proc. Natl. Acad. Sci. USA* **119**(4), e2116550119 (2022). doi:10.1073/pnas.2116550119.

[7] Bozkurt Ç., Vasilyeva A., Goteti A. AstraROLE2 & AstraSUIT2: Multi-Task Annotation Models for Functional Profiling of Proteins. *bioRxiv* 2025.06.21.660734 (2025). doi:10.1101/2025.06.21.660734.

[8] Bozkurt Ç., Vasilyeva A., Goteti A. AstraPTM2: A Context-Aware Transformer for Broad-Spectrum PTM Prediction. *bioRxiv* 2025.10.03.680341 (2025). doi:10.1101/2025.10.03.680341.

[9] Goteti A., Vasilyeva A., Bozkurt Ç. AstraBIND: Graph Attention Network for Predicting Ligand Binding Sites. *bioRxiv* 2025.11.10.687555 (2025). doi:10.1101/2025.11.10.687555.

Turn a dark target into an experiment-ready plan.

Have a dark or under-characterised target — membrane or soluble? Send us a UniProt ID and Orbion returns a characterisation preview like this one: topology or fold, disorder, PTMs, pocket or active-site hypotheses, and recommended validation experiments.

Request a target-profile preview · **Book a 20-minute target review**

contact@orbion.life · orbion.life

Characterise. Engineer. Execute.