DISTILLATION №15 · ALEMBIC LABS

SEMAGLUTIDE — GLU-16 → HOMOGLUTAMATE (HGE, B-METHYLENE-EXTENDED GLU SIDE CHAIN), SINGLE RESIDUE REPLACEMENT IN THE CENTRAL HELIX

generated 2026-05-02T22:32:04.581244+00:00

PROMISING METABOLIC GLU-16 → HOMOGLUTAMATE (HGE, B-METHYLENE-EXTENDED GLU SIDE CHAIN), SINGLE RESIDUE REPLACEMENT IN THE CENTRAL HELIX GLUCAGON-LIKE PEPTIDE 1 RECEPTOR

AVERAGE CONFIDENCE	PTM / IPTM	VERDICT
71.4%	0.783 / 0.730	PROMISING
TARGET	**UNIPROT**	**BINDING PROBABILITY**
Glucagon-like peptide 1 receptor	P43220	—

TLDR

Fold №15 tests whether extending the Glu-16 side chain in semaglutide by one methylene unit (homoglutamate, Hge) would improve salt-bridge geometry with GLP-1R ECD arginine residues, predicted to enhance receptor affinity. Despite a structurally plausible hypothesis, the prediction pipeline returned no usable 3D complex geometry, no Chai-1 agreement score, and no Boltz-2 affinity values — leaving the core salt-bridge hypothesis untestable from this run. The backbone fold scored reasonably (pLDDT 0.71, pTM 0.78) but this reflects the peptide's intrinsic helical propensity, not a meaningful receptor-docked pose. This fold is discarded as biologically uninformative given the absence of interface-level data.

EXECUTIVE SUMMARY

Semaglutide Glu-16→Hge: pLDDT 0.71, pTM 0.78 — but no complex geometry, no Chai-1 agreement, no affinity delta. The salt-bridge reach hypothesis is chemically sound; the pipeline simply couldn't resolve it. Discarded as a tool limitation — the Hge variant remains a viable synthesis target.

DETAILED ANALYSIS

Semaglutide is a GLP-1 receptor agonist engineered for once-weekly subcutaneous dosing, achieving its pharmacokinetic durability through two deliberate structural modifications: Aib at position 8 (conferring resistance to DPP-4 cleavage) and a C18 fatty-diacid tether at Lys26 (enabling reversible albumin binding and extending plasma half-life to ~46 hours in preclinical models). Its GLP-1R binding affinity of approximately 0.38 nM is intentionally modest relative to liraglutide — a design trade-off that prioritizes duration over raw potency. Fold №15 asks whether this affinity ceiling can be selectively raised by tuning a single interfacial residue without disturbing either of the pharmacokinetic anchors.

The modification under study is Glu-16 → homoglutamate (Hge), a β-methylene homologation that extends the acidic side chain by approximately 1.5 Å. The mechanistic premise draws from cryo-EM and crystallographic studies of GLP-1 analogs bound to GLP-1R showing that Glu-16 participates in electrostatic contacts with basic residues (Arg310, Arg380) on the receptor ECD. If the native glutamate carboxylate sits at the outer edge of productive salt-bridge distance (~4 Å), adding one methylene could close this gap to ~3 Å without altering charge, pKa, or helical backbone geometry — a classically conservative medicinal chemistry move. Homo-amino acid substitutions in helical contexts have precedent and are generally well tolerated, making this a structurally plausible hypothesis.

The structural prediction run produced a pLDDT of 0.714 and a pTM of 0.783 for the isolated peptide — scores that are respectable and consistent with the known strong helical propensity of the semaglutide scaffold. However, these numbers describe the peptide in isolation, not in complex with GLP-1R. Critically, no Chai-1 agreement value was returned, no Boltz-2 affinity module output was generated, and no predicted binding change relative to native semaglutide is available. The absence of these interface metrics means the central question of the fold — does homoglutamate improve receptor contact geometry? — is simply unanswered by this run.

The heuristic sequence-based profile provides limited consolation. Aggregation propensity is low (0.152), stability score is moderate (0.432), and BBB penetration is negligible (0.024) — all consistent with a large, lipidated peptide agonist that is not CNS-targeted. The half-life estimate of moderate-to-long aligns with the albumin-

tethered class. None of these heuristic values speak to receptor affinity or the salt-bridge hypothesis.

From a literature standpoint, this fold is attempting something the published SAR literature has not directly addressed. No retrieved paper characterizes Glu-16 as a critical affinity determinant, quantifies its distance to Arg310/Arg380 in the semaglutide-bound pose, or tests homologated amino acids in GLP-1 helix scaffolds. The broader cryo-EM structural biology literature (Jazayeri 2017, Liang 2018) supports the existence of acidic-basic contacts at the peptide–ECD interface, but these are not in the retrieved set and do not directly validate the Hge extension strategy. The pharmacokinetic literature is robust but irrelevant to the binding interface question.

In cross-fold context, this discarded result sits alongside Fold №3 (Retatrutide Aib-2, pLDDT 0.71, discarded) as another metabolic peptide fold where structural prediction returned plausible backbone metrics but insufficient interface data to support the hypothesis. Fold №10 (Retatrutide Lys-17→Arg, pLDDT 0.78, PROMISING) used a similar intra-helix charge-optimization logic and returned useful signal — suggesting that the strategy of tuning charged residues in amphipathic helices can yield informative predictions when the right structural inputs are available. The contrast between Fold №10's success and this fold's failure is instructive: the difference likely lies in whether the docking pipeline could resolve the receptor-peptide interface, not in the chemical logic of the modification itself.

The discarded verdict should be read as a tool limitation, not a verdict on the chemistry. The Hge substitution remains a scientifically reasonable hypothesis — the methylene extension is conservative, position 16 is distant from both the DPP-4 cleavage site and the C18 tether, and the semaglutide scaffold has demonstrated headroom for affinity improvement. What this fold tells us is that current in silico pipeline outputs, without reliable complex geometry and affinity module data, cannot adjudicate the salt-bridge reach hypothesis. A re-run with an ensemble approach, co-crystal structure input, or explicit molecular dynamics on the GLP-1R–peptide interface would be required to generate informative signal.

RESEARCH BRIEF

FOLD №15 — SEMAGLUTIDE GLU-16 → HOMOGLUTAMATE

Verdict: DISCARDED | Target: GLP-1R (P43220) | Class: Metabolic

---

MECHANISM OF ACTION (BACKGROUND)

Semaglutide is a GLP-1 receptor agonist that activates GLP-1R, a class B GPCR expressed on pancreatic β-cells, gut enteroendocrine cells, and the hypothalamus. Agonism drives glucose-dependent insulin secretion, glucagon suppression, gastric emptying delay, and centrally mediated appetite reduction. Its therapeutic value in T2DM and obesity management (STEP and SUSTAIN trial programs) is well established. The peptide's pharmacological identity rests on two structural pillars: Aib-8 (DPP-4 resistance) and a C18 fatty-diacid at Lys-26 (albumin binding, ~46 h half-life in mini-pigs). These features were deliberately optimized at a modest cost to raw GLP-1R affinity — semaglutide binds at ~0.38 nM, approximately 3-fold weaker than liraglutide — suggesting that affinity headroom exists in the scaffold.

MODIFICATION HYPOTHESIS (WHAT WE TESTED)

Fold №15 replaced Glu-16 in the central α-helix of semaglutide with homoglutamate (Hge), a β-methylene-homologated glutamate that extends the carboxylate reach by approximately 1.5 Å. The hypothesis: cryo-EM structures of GLP-1 analogs bound to GLP-1R show Glu-16 making an acidic contact with ECD Arg310/Arg380; if the native glutamate sits marginally outside optimal salt-bridge geometry (~4 Å), homoglutamate's extended reach (~3 Å) could tighten this electrostatic contact and improve receptor affinity. The modification is charge-neutral (preserves carboxylate and pKa), is expected to be helix-compatible (β-homologated amino acids are well tolerated in α-helical contexts), and is chemically distinct from both the DPP-4 resistance element at position 8 and the albumin tether at position 26.

This strategy differs meaningfully from prior lab folds: Fold №3 (Retatrutide Aib-2) tested backbone rigidification for DPP-4 resistance; Fold №10 (Retatrutide Lys-17→Arg) tested intra-helix i,i+4 salt-bridge optimization. Fold №15 is uniquely focused on inter-molecular reach extension at a receptor-facing interfacial residue — a distinct design axis.

WHY THE PREDICTION WAS UNINFORMATIVE (TECHNICAL ANALYSIS)

The prediction pipeline returned the following outputs:

Metric	Value	Interpretation
pLDDT	0.714	Moderate–adequate backbone confidence
pTM	0.783	Reasonable global topology
ipTM	0.730	Interface template modelling score

Metric	Value	Interpretation
Chai-1 agreement	None	**No cross-model consensus available**
Boltz-2 affinity	No values	**Affinity module did not produce output**
Predicted binding change	None	**Core metric absent**

The pLDDT of 0.714 and pTM of 0.783 are respectable scores, consistent with semaglutide's known strong α-helical propensity. However, these scores characterize the peptide backbone in isolation — they do not reflect a receptor-docked complex and cannot address salt-bridge geometry at the GLP-1R ECD interface. The central failure of this fold is the absence of Chai-1 agreement and Boltz-2 affinity values: without cross-model consensus on a co-folded complex and without an affinity delta, the hypothesis that Hge extends productive electrostatic reach cannot be evaluated.

This is a tool limitation, not a chemistry failure. The semaglutide scaffold is a large, lipidated, modified peptide (31 residues + non-canonical Aib + fatty-acid tether); the prediction of its full receptor-bound complex is at the edge of current single-run pipeline capability. Fold №3 (Retatrutide Aib-2) encountered a comparable outcome at pLDDT 0.71 — suggesting this class of highly modified metabolic peptides consistently challenges single-run predictors at the interface level. By contrast, Fold №10 (Retatrutide Lys-17→Arg, pLDDT 0.78) produced a PROMISING verdict, likely because the intra-helix salt-bridge geometry is more accessible to structure prediction than an inter-molecular peptide–receptor contact.

Heuristic sequence-based estimates provided: - **Aggregation propensity: 0.152** (low — expected for a designed helix-forming analog) - **Stability score: 0.432** (moderate — consistent with modified peptide class) - **BBB penetration: 0.024** (negligible — appropriate for a large peripheral agonist) - **Half-life: moderate-to-long** (reflects albumin-tethered class behavior)

None of these heuristics bear on receptor affinity or salt-bridge geometry.

WHAT THIS TELLS US (NEGATIVE RESULTS ARE DATA)

This fold's discarded status is informative in three respects:

1. **The pipeline cannot currently adjudicate inter-molecular side-chain reach hypotheses for lipidated class B GPCR peptides in a single run.** The absence of Boltz-2 affinity output and Chai-1 consensus is a consistent challenge for this peptide class. Future folds testing interfacial contacts on semaglutide or similar analogs should plan for ensemble runs or co-crystal-seeded docking from the outset.

2. **The chemistry is not ruled out.** The Hge substitution has not been predicted to be destabilizing, aggregation-prone, or conformationally disruptive. pLDDT 0.714 and pTM 0.783 are not failure scores — they indicate a well-folded backbone. The biological hypothesis remains structurally plausible; it is simply unresolved by these tools at this time.
3. **The semaglutide scaffold's affinity headroom (0.38 nM vs. liraglutide's ~0.12 nM) remains an open optimization target.** This fold does not close that question. It only establishes that a single-run AlphaFold pipeline without usable complex geometry is insufficient to evaluate it.

ALTERNATIVE HYPOTHESES TO TEST (AVOID THE FAILURE MODE)

1. **Ensemble docking with GLP-1R ECD co-crystal input**: Use the published GLP-1 peptide–GLP-1R ECD structure (PDB: 5NX2 or equivalent) as a structural prior for Rosetta or Glide docking of the Hge-16 variant. This bypasses the cold-start complex prediction problem entirely.
2. **Molecular dynamics on the Glu-16/Arg310 contact**: Short (50–100 ns) MD simulation of native semaglutide vs. Hge-16 variant in the receptor-bound pose would directly quantify salt-bridge occupancy and distance distributions — the exact metric the hypothesis requires.
3. **Paired Glu-16 alanine scan first**: Before homologating, computationally testing Glu-16 → Ala (a charge-ablation control) would confirm whether the position contributes to binding affinity at all in the in silico framework. If Ala substitution shows predicted affinity loss, the salt-bridge hypothesis is supported; if neutral, the position may not be a productive optimization target.
4. **Position 12 or 13 acidic residue extension**: If Glu-16's receptor contact is difficult to validate, equivalent homologation at other acidic positions in the helical face (Asp15 or Glu22 in related GLP-1 scaffolds) could be tested where cryo-EM data more clearly resolves the ECD contact geometry.
5. **Synthesize and test directly**: Given the chemical conservatism of the Hge substitution and the low aggregation propensity predicted (0.152), this analog is a reasonable synthesis target for a cAMP accumulation assay against GLP-1R — the wet-lab answer is not gated on a successful in silico run.

SEQUENCES

NATIVE

```
HAEGTFTSDVSSYLEGQAAKEFIAWLVRGRG
```

MODIFIED

```
HAEGTFTSDVSSYLE(Hge)QAAKEFIAWLVRGRG
```

CAVEATS

- in silico prediction only — requires wet lab validation
- single-run prediction (not ensembled)
- predicted properties may not reflect real-world biological behavior
- this is research, not medical advice
- no Chai-1 agreement or Boltz-2 affinity output was generated — the salt-bridge geometry hypothesis is structurally unresolvable from this run
- pLDDT/pTM scores reflect isolated peptide backbone confidence, not receptor-docked complex quality
- heuristic property estimates (aggregation, stability, BBB, half-life) are sequence-based approximations only
- the Glu-16/Arg310/Arg380 salt-bridge premise derives from the broader GLP-1R structural biology literature and is not directly confirmed in the retrieved abstract set
- homoglutamate (Hge) is a non-canonical amino acid; its helical compatibility is inferred from precedent, not directly modelled here
- the dominant determinant of semaglutide's in vivo exposure is albumin binding half-life, not raw receptor affinity — even a confirmed affinity improvement may not translate to pharmacodynamic benefit
- Verdict reclassified: DISCARDED → PROMISING. Raw metrics (pLDDT/pTM/ipTM) permit at least the higher tier; the original LLM discard reflected modification chemistry the predictor cannot represent (D-AA, lipid moiety, non-canonical residue). Per the metric-floor rule this is a caveat, not a verdict downgrade. Report text below pre-dates the rule and may still describe the fold as DISCARDED — the structural verdict shown is the authoritative one.

CITATIONS

1. **PMID —** (2015) — — Discovery of the Once-Weekly Glucagon-Like Peptide-1 (GLP-1) Analogue Semaglutide
2. **PMID —** (2024) — — Clinical Pharmacokinetics of Semaglutide: A Systematic Review
3. **PMID —** (2021) — — Safety of Semaglutide
4. **PMID —** (2020) — — Semaglutide lowers body weight in rodents via distributed neural pathways
5. **PMID —** (2023) — — Semaglutide for the treatment of obesity
6. **PMID —** (2022) — — Wegovy (semaglutide): a new weight loss drug for chronic weight management

SOLANA SIGNATURE 3n8A1GBzyn5Ltfm7DmrknNUahGKeEg3kdk99HXs6k8aNwmvwLRYhcEeTcXbvFekCBXtcryk7gfYJpyEc5YevU8jd

DATA SHA-256 283363ae7dbadc0c0b72c5c6e5be0ea4bb9e5e9ef780b0dd492b1862ba270631

VERIFY https://solscan.io/tx/3n8A1GBzyn5Ltfm7DmrknNUahGKeEg3kdk99HXs6k8aNwmvwLRYhcEeTcXbvFekCBXtcryk7gfYJpyEc5YevU8jd