SpoIIGA (P13801, BACSU) — Aspartic-type Endopeptidase Prediction Review

Hypothesis under review: GO-GPT (via BioReason-Pro) predicts **aspartic-type endopeptidase activity (GO:0004190)** for *Bacillus subtilis* SpoIIGA, the membrane protease that processes pro-σE to σE in the mother cell.

Focus type: computational_prediction · **Term:** GO:0004190 aspartic-type endopeptidase activity

---

Executive Judgment

Verdict: SUPPORTED (with a fold caveat).

Independent primary literature and curated database evidence converge on SpoIIGA being an **aspartic-type endopeptidase**:

- Direct mutational + homology-modeling study concludes SpoIIGA is *"a novel type of aspartic protease"* whose catalytic (C-terminal) domain **dimerizes like HIV-1 protease** (Imamura et al. 2008, P 18378688), reaffirmed in Imamura et al. 2011 (P 21362630).
- UniProt (SP2G_BACSU / P13801) annotates it as *"Sporulation sigma-E factor-processing peptidase"*, **EC 3.4.23.-** (the aspartic-endopeptidase EC class), keyword **"Aspartyl protease"**, and **MEROPS family A36.001** — SpoIIGA is the *holotype* of aspartic-peptidase family A36.
- The single UniProt-annotated **active site is Asp183**, embedded in a canonical aspartic-protease catalytic motif **D183-S184-G185 (D-S-G)** — directly homologous to the HIV-1 protease **D25-T26-G27 (DTG)** motif. Mutagenesis of residues 181/182/**183**/184 abolishes pro-σE cleavage (D183 abolishes cleavage of wild-type SigE).

The one caveat: the aspartic mechanism is established by **catalytic-aspartate mutagenesis + homology modeling**, not by an experimental crystal structure. SpoIIGA is a **divergent, dimeric (retropepsin/HIV-like), membrane-associated** aspartic protease, not a classical monomeric pepsin-fold enzyme. GO:0004190 does not require the pepsin fold — it requires the aspartic catalytic mechanism — so the term is appropriate. The prediction is therefore **correct**, and if anything is *more* specific and better supported than the historical "putative protease of unknown mechanism" status (Pfam still uses the legacy name **Peptidase_U4**, "U" = unknown catalytic type).

Evidence Matrix

#	Citation	Evidence type	Supports/ Refutes/ Qualifies	Claim tested	Key finding	Context	Co lim
1	Imamura, Zhou, Feig, Kroos 2008 — P 18378688	Structural modeling + mutagenesis + reconstitution	**Supports**	Is SpoIIGA an aspartic protease?	*"SpoIIGA is a novel type of aspartic protease whose C-terminal half forms a dimer similar to the HIV-1 protease"*; SpoIIR+SpoIIGA sufficient to process pro-σE in *E. coli*; catalytic-Asp mutations abolish cleavage	*B. subtilis* / *E. coli* reconstitution; sporulation	Hig asp me fol mo cry
2	Imamura, Kuwana, Kroos, Feig, Takamatsu, Watabe 2011 — P 21362630	Mutagenesis / substrate specificity	**Supports / qualifies**	Aspartic protease classification + substrate recognition	*"SpoIIGA is a novel type of membrane-associated aspartic protease… cleaving Pro-σ(E)"*; basic residues 245/284 (substrate binding, not catalytic) contribute to specificity	Cross-species *Bacillus* orthologs, *E. coli* co-expression	Hig dis cat fro rec res K2
3		Database record	**Supports**	Curated MF classification	EC **3.4.23.-**; keyword "Aspartyl		Hig cu syr

#	Citation	Evidence type	Supports/ Refutes/ Qualifies	Claim tested	Key finding	Context	Co lim
	UniProt **P13801** (SP2G_BACSU)				protease"; **active site Asp183**; MEROPS **A36.001** (family holotype); 5 N-terminal TM helices + C-terminal protease domain	Curated, reference proteome	inf ref
4	This analysis (sequence motif scan)	Computational (sequence)	**Supports**	Presence of aspartic catalytic motif	Asp183 sits in a **D-S-G** motif (positions 183–185), homologous to HIV-1 protease DTG dyad; C-terminal cytoplasmic domain	Computed on P13801 sequence	Hig pre co dir asp me
5	Fawcett, Melnikov, Youngman 1998 — P 9663680; Hofmeister 1998 — P 9573195	Localization / interaction	**Qualifies (context)**	Where/how SpoIIGA acts	SpoIIGA targets to the sporulation septum membrane; forms a complex with pro-σE	*B. subtilis* sporulation	Su (m sep sul int MF
6	Pfam **PF03419** "Peptidase_U4"	Database (legacy)	**Qualifies**	Historical mechanism uncertainty	Domain retains "U4" (unknown catalytic type) name	Domain family	Re asp a p ref

#	Citation	Evidence type	Supports/ Refutes/ Qualifies	Claim tested	Key finding	Context	Co lim
					predating 2008 reclassification to aspartic (MEROPS A36)		no su
7	This analysis — ortholog alignment (UniProt P13801, D5DQW6, Q45832)	Computational (evolutionary/ orthology)	**Supports**	Is the catalytic Asp motif conserved across the family?	Catalytic **D-[S/T]-G** motif conserved: *B. subtilis* D183-S-G, *P. megaterium* D183-S-G, *C. acetobutylicum* D174-T-G (DTG); flanking L-D-[S/T]-G-N pattern conserved	Bacillus → Priestia → Clostridium orthologs	Hig co ac ar lin art
8	This analysis — AlphaFold model AF-P13801-F1 (v6)	Computational (structural, predicted)	**Qualifies**	Is D183 in a confidently modeled protease fold?	Mean pLDDT 85; catalytic Asp183/DSG region pLDDT 80–93; res183=ASP. Monomer model cannot show the inter-subunit dyad	Predicted structure only; no experimental PDB	Mo su res ide co no dy

Evolutionary Conservation of the Catalytic Motif (computed this run)

Alignment of reviewed SpoIIGA orthologs (UniProt) confirms the catalytic aspartate sits in a conserved aspartic-protease **D-[S/T]-G** motif, independent of the *B. subtilis*-specific mutagenesis:

Ortholog	Organism	Length	EC	MEROPS	Catalytic motif (position)	Local context
P13801	*Bacillus subtilis* 168	309	3.4.23.-	A36.001	**D183-S-G**	`GLIDSGNQLYD`
D5DQW6	*Priestia (Bacillus) megaterium*	307	3.4.23.-	A36.001	**D183-S-G**	`GLIDSGNQLVD`
Q45832	*Clostridium acetobutylicum*	266	3.4.23.-	—	**D174-T-G**	`FLDTGNELRE`

The catalytic Asp and its flanking **L-D-[S/T]-G-N** pattern are conserved from *Bacillus* to the distant *Clostridium* ortholog (which carries the classical **DTG** variant, identical in type to HIV-1 protease). This is orthogonal evolutionary support that the aspartic catalytic residue is a genuine, maintained feature of the SpoIIGA family, not a lineage-specific artifact. (Provenance: UniProt REST retrieval + regex motif scan executed in Iteration 2.)

GO Curation Implications (leads — require curator verification)

- **Retain / assign GO:0004190 (aspartic-type endopeptidase activity)** as the SpoIIGA molecular-function term. It is biologically supported by direct mutational evidence (catalytic D183 in a DSG motif) plus curated MEROPS A36 / EC 3.4.23.- classification. The GO-GPT prediction is **correct and appropriately specific** (more informative than "peptidase activity" or "protein binding").
- **Recommended evidence basis:** because the aspartic mechanism rests on homology modeling + mutagenesis (not a solved structure), the most defensible support is a **sequence/**

mutation-based or **author-statement** basis citing

P 18378688

and 21362630, rather than a direct enzymological demonstration of aspartic-class inhibitor sensitivity. Curator may prefer ISS/IMP-flavored evidence with these PMIDs over IDA.

- **Companion terms that are also well supported** (for completeness, not part of this focus): BP **GO:0042174 negative regulation of sporulation / sigma-factor processing / GO:0030436 asexual sporulation** context via pro-σE processing; CC **GO:0005886 plasma membrane** / sporulation septum (5 TM helices, septal localization,

P 9663680

). MF could optionally note substrate = sigma-factor precursor.
- **Do not** downgrade to generic "protein binding" — a specific catalytic MF term is supported.

Mechanistic Scope

- **Direct molecular function (what GO:0004190 tests):** SpoIIGA is the endopeptidase that hydrolyzes a peptide bond in **pro-σE**, removing the 27-residue pro-sequence to release active σE. Catalysis uses an **aspartic mechanism**: the C-terminal cytoplasmic domain dimerizes so that two copies of the **D183(S184)(G185)** motif form the paired catalytic aspartate dyad (retropepsin/HIV-1-protease-like).
- **Distinct from downstream phenotypes:** the sporulation defect of *spoIIGA* mutants, mother-cell/forespore compartmentalization of σE activity, and the intercellular SpoIIR→SpoIIGA signaling are **downstream/contextual** consequences, not the molecular activity itself. The N-terminal 5-TM domain is a **membrane signal-receiving/regulatory module** (senses SpoIIR), not the catalytic center.

Conflicts and Alternatives

- **No primary evidence assigns SpoIIGA to a different protease class** (serine/cysteine/metallo). The only "conflict" is historical/nomenclatural: SpoIIGA was long a *"putative protease"* of unknown mechanism (Pfam **Peptidase_U4**, "U" = unknown), and some early models entertained that SpoIIGA might be a regulator of an unidentified protease. The 2008 *E. coli* reconstitution (SpoIIR + SpoIIGA **sufficient**) plus catalytic-Asp mutagenesis resolved this in favor of SpoIIGA being the aspartic protease itself.

- **Paralog/ortholog confusion:** low risk. SpoIIGA orthologs across *Bacillus* share the mechanism; the closest mechanistic analog is HIV-1 protease (fold analogy, not homology). Substrate-recognition residues (R245, K284) are sometimes mis-read as "catalytic" because their mutation abolishes cleavage, but 2011 data show these are **substrate-binding**, not the catalytic dyad.
- **In-vitro-only concern:** not applicable — activity is demonstrated in the physiological *B. subtilis* pathway and reconstituted heterologously.

Knowledge Gaps

1. **No experimental 3D structure.** Checked: UniProt lists only an AlphaFoldDB model (no PDB). I retrieved the AlphaFold model (AF-P13801-F1, v6): mean pLDDT = **85.0** (high confidence), C-terminal protease domain (151–309) mean pLDDT = 85.0, and the catalytic **Asp183/DSG** region is confidently modeled (pLDDT 80–93, res183 = ASP as annotated). However, the *functional* active site is an **inter-subunit dimer dyad**, which a monomer model cannot display — so the pepsin/HIV-like fold and the two-Asp active site remain **inferred from modeling + mutagenesis + conservation**, not observed. A crystal/cryo-EM structure of the dimer (ideally with a transition-state analog or pepstatin-class inhibitor bound) would confirm the aspartic dyad geometry. Matters because GO:0004190's mechanistic basis currently rests on indirect evidence.
2. **Second catalytic aspartate identity.** Checked: UniProt annotates only one active site (D183) because catalysis is provided by a **homodimer** contributing two D183 copies; whether an intramolecular second Asp also contributes is not experimentally pinned. Resolving this needs the structure or a mixed-dimer complementation assay.
3. **Inhibitor sensitivity.** No report (found) of classic aspartic-protease inhibitor (pepstatin A) inhibiting SpoIIGA. A positive result would provide orthogonal, direct enzymological confirmation of the aspartic class.

Discriminating Tests

- **Structure determination** of the C-terminal domain dimer (± substrate/inhibitor) to visualize the paired Asp dyad — the single most decisive test.
- **Pepstatin A / aspartic-protease inhibitor** sensitivity assay on reconstituted pro-σE cleavage (positive = confirms class; important because SpoIIGA is divergent and might be inhibitor-insensitive).

- **Catalytic-dyad complementation:** co-express two SpoIIGA D183 point mutants that could complement in trans within a dimer; restoration of activity confirms the shared dimeric aspartic active site.
- **AlphaFold2/3 model inspection** of the C-terminal domain for spatial apposition of two D183-S-G motifs at a dimer interface (fast, in-silico corroboration).

Curation Leads (verify before applying)

- **Action:** support/keep MF **GO:0004190** for SpoIIGA (P13801). Prediction is **correct**.
- **Candidate references (with snippets to verify):**
- P 18378688
  — *"SpoIIGA is a novel type of aspartic protease whose C-terminal half forms a dimer similar to the human immunodeficiency virus type 1 protease."*
- P 21362630
  — *"SpoIIGA is a novel type of membrane-associated aspartic protease that responds to a signal from the forespore by cleaving Pro-σ(E)..."*
- **Database:** UniProt P13801 (EC 3.4.23.-, "Aspartyl protease", active site D183); MEROPS **A36.001**.
- **Suggested qualifier/note for curator:** flag that the mechanism is **modeling** + **mutagenesis-based** (no crystal structure) and that the enzyme is a **divergent membrane-associated dimeric aspartic protease** (family A36), so evidence code should reflect sequence/experiment-inference rather than direct structural proof.
- **Suggested questions:** Is there any pepstatin-sensitivity or structural datum post-2011 that would upgrade the evidence to IDA-level? Should the substrate (sigma-factor precursor) be captured via a "with/from" or extension?
- **Suggested experiment:** pepstatin inhibition + cryo-EM of the SpoIIGA/pro-σE complex.

Bottom line

The aspartic-type endopeptidase prediction for SpoIIGA is **supported** by convergent mutational, database, and sequence-motif evidence. SpoIIGA is the founding member of MEROPS aspartic-peptidase family A36, carries a catalytic Asp183 in a D-S-G motif, and

functions as a dimeric HIV-protease-like aspartic endopeptidase — with the only reservation that the fold/dyad is established by modeling and mutagenesis rather than an experimental structure.
