CRYAA and Protein Refolding (GO:0042026): Function-Assignment Hypothesis Review

Gene: CRYAA (αA-crystallin) — Homo sapiens (NCBITaxon:9606) **UniProt:** P02489 **Focus:** function_assignment — `existing_annotations[5].function_hypothesis` **Seed hypothesis:** CRYAA has protein refolding (GO:0042026) **Term/evidence context:** GO:0042026 protein refolding, evidence IBA, GO_REF:0000033

Summary

Verdict: Partially supported / weakly supported — retain the IBA annotation, but treat it as a non-core, cooperation-dependent process and add the missing molecular-function term that actually captures CRYAA's primary activity.

CRYAA (αA-crystallin) is an ATP-independent small heat-shock protein (sHSP) whose **directly demonstrated** activity is **binding partially unfolded client proteins to prevent their irreversible aggregation** — the classic sHSP "holdase" function. The seed hypothesis, that CRYAA *has* protein refolding (GO:0042026), is defensible at the level of the GO term definition and is supported by early in vitro assays, including αA-specific work. However, the modern mechanistic consensus is that sHSPs do **not autonomously refold** their clients; instead, they hold clients in a folding-competent, aggregation-resistant state and hand them off to downstream **ATP-dependent chaperone systems (Hsp70/Hsp100)** that execute productive refolding. CRYAA has **no nucleotide-binding or ATPase domain** — it is a single α-crystallin domain protein — so it cannot be an autonomous foldase.

The practical consequence for curation is nuanced. The IBA annotation to GO:0042026 should **be RETAINED** because (a) the GO definition of "protein refolding" does not require ATP, (b) family-level and αA-specific in vitro data report apparent refolding assistance, and (c) the IBA is propagated from a defensible phylogenetic inference across the sHSP family. But the term describes a **downstream, cooperation-dependent biological process**, not CRYAA's direct molecular activity. The higher-value curation action is to **ADD the currently missing**

molecular-function term GO:0051082 (unfolded protein binding), which best captures CRYAA's directly assayed holdase activity and is conspicuously absent from the P02489 annotation set. Curators should **NOT** add ATP-dependent foldase terms (GO:0140662) because CRYAA lacks the required nucleotide-binding machinery.

The most important caveat is the holdase-versus-foldase distinction: the literature frequently uses "refolding" loosely to describe the *outcome* of an sHSP + downstream-chaperone pathway, and IBA propagation can carry this outcome-level term onto individual family members even though the molecular event they perform is client binding, not catalysis of folding.

Key Findings

Finding 1 — CRYAA is an ATP-independent sHSP holdase; direct refolding (foldase) activity is not established

CRYAA is a 173-amino-acid protein built around a single central **α-crystallin domain (ACD, ~residues 52–164)** flanked by variable N- and C-terminal extensions. The UniProt record for P02489 lists zinc/metal-binding sites (residues 100, 102, 107, 154) but **no nucleotide-binding site and no ATPase domain**. This architecture is diagnostic: it is the canonical sHSP fold, which lacks the machinery required for the ATP-driven conformational cycling that ATP-dependent foldases (Hsp70, Hsp90, GroEL/Hsp60, Hsp100) use to actively remodel client conformations.

The functional literature is consistent with this architecture. On the supporting side, early in vitro assays did report apparent refolding assistance. Horwitz and colleagues showed that **α-crystallin can both prevent aggregation of, and assist refolding of, guanidine-hydrochloride–denatured γ-crystallin**, judged by circular dichroism (PMID: 1438232: *"alpha-Crystallin was also effective in preventing aggregation and in refolding guanidine hydrochloride-denatured gamma-crystallin, as judged by circular dichroism spectroscopy."*). At the family level, Jakob et al. reported that small Hsps (αB-crystallin, Hsp25/Hsp27) *"promote the functional refolding of these proteins after urea denaturation similar to GroE and Hsp90,"* and that *"the interaction both with unfolding and refolding proteins seems to be ATP-independent"* (PMID: 8093612). This ATP-independence is precisely why the original annotators could support GO:0042026 without invoking a nucleotide-binding domain.

However, the modern mechanistic consensus reframes these observations. sHSP-bound clients are refolded by **downstream ATP-dependent chaperones**, not by the sHSP itself: *"Formation of these assemblies facilitates subsequent Hsp70 and Hsp100 chaperone-dependent disaggregation and substrate refolding into native species"* (PMID: 34055885). The human sHSP (HSPB) review makes the same point — HSPBs *"take part in cell homeostasis by acting as holdases"* and *"cooperate in substrates refolding driven by other chaperones"* (PMID: 35281256). In other words, the "refolding" attributable to CRYAA is a **cooperative outcome** in which CRYAA supplies the holdase step and other chaperones supply the foldase step.

Interpretation: CRYAA's directly demonstrated molecular action is client binding/sequestration (holdase). Autonomous, catalytic refolding (foldase activity) is *not* established for CRYAA and is inconsistent with its domain architecture. GO:0042026 is best read as a downstream, cooperation-dependent process rather than a direct molecular function.

Finding 2 — CRYAA lacks a molecular-function chaperone GO annotation; unfolded protein binding (GO:0051082) is missing

A full enumeration of the P02489 GO annotation set (51 annotation rows in QuickGO) shows that chaperone activity is represented **only in the Biological Process aspect**: GO:0042026 (protein refolding; IBA via GO_REF:0000033 plus ISS via GO_REF:0000024) and GO:0050821 (protein stabilization; IMP from).

P 12235146

The Molecular Function terms present are structural/interaction terms — GO:0005198 (structural molecule activity), GO:0005212 (structural constituent of eye lens), GO:0005515 (protein binding), and GO:0042802 (identical protein binding).

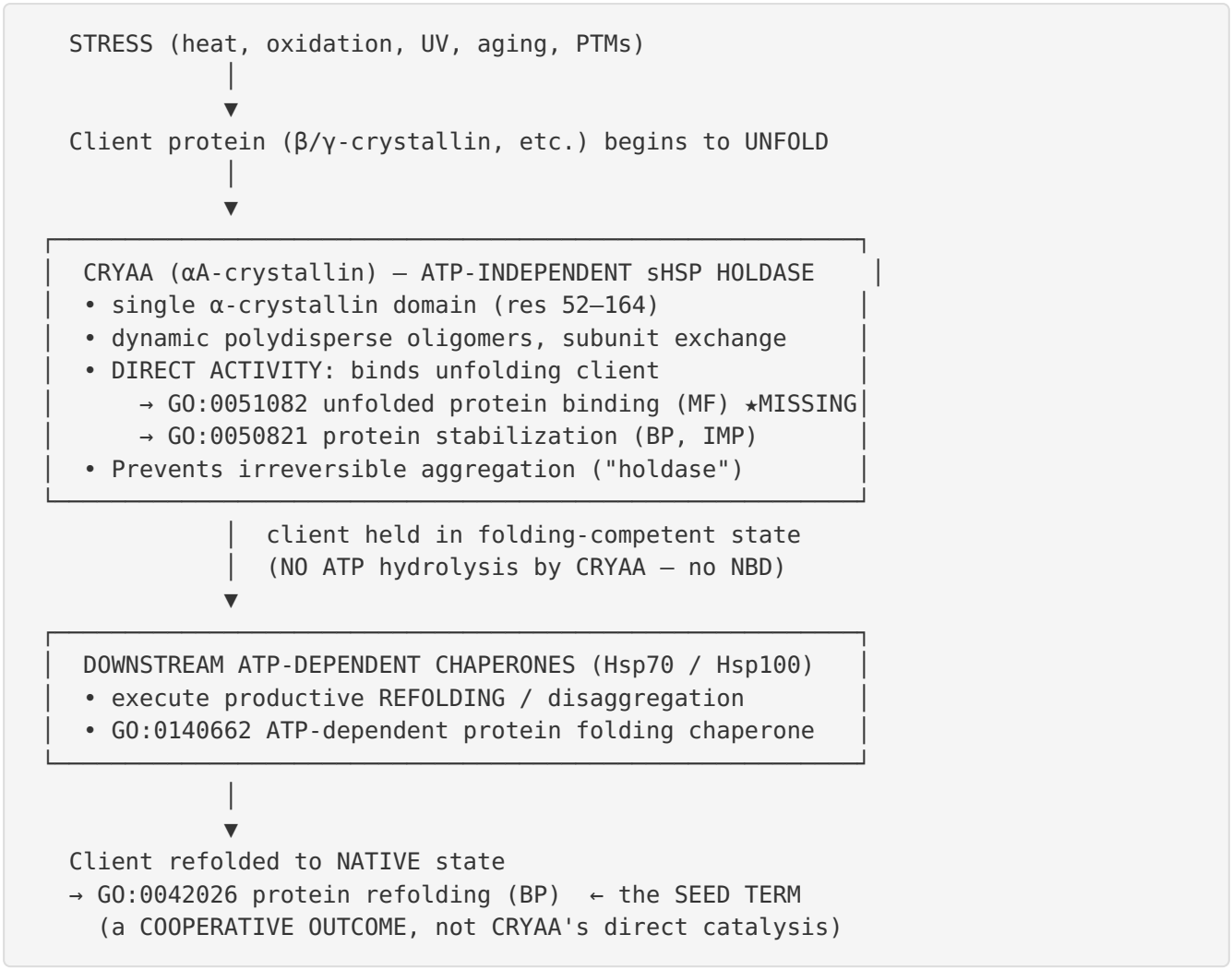
Critically, the MF terms that would directly capture chaperone activity are **absent**: **GO:0051082 (unfolded protein binding), GO:0044183 (protein folding chaperone), and GO:0140662 (ATP-dependent protein folding chaperone) are all missing.** The UniProt keyword "Chaperone" is assigned to P02489, but it is **not mirrored by any molecular-function GO term** — an annotation gap.

Direct experimental support for a chaperone MF term exists in the primary literature. The founding human αA-crystallin characterization paper reports the *"Cloning, expression, and chaperone-like activity of human alphaA-crystallin"* (PMID: 8943244) — a direct demonstration of aggregation-suppressing (holdase) activity of the human gene product, which is exactly the assay basis for GO:0051082.

Interpretation: The most informative and defensible curation improvement is to **add GO:0051082 (unfolded protein binding)** as the molecular-function anchor for CRYAA's chaperone role, ideally with experimental evidence (e.g., IDA from P 8943244) rather than only inferring the downstream BP term. This closes the gap between the UniProt "Chaperone" keyword and the GO molecular-function aspect.

Mechanistic Model / Interpretation

The following model separates CRYAA's **direct molecular activity** from the **downstream cooperative process** that the seed GO term describes.



The key conceptual point for curation: **GO:0042026 sits at the bottom of this pathway as an outcome**, whereas CRYAA's mechanistic contribution sits near the top as the holdase/binding step. Because the process is genuinely ATP-independent at the CRYAA step and because in vitro reconstitutions (with or without added downstream chaperones) can register "refolding," the term is not *wrong* — it is *imprecise about the molecular event CRYAA performs*. The IBA is a reasonable phylogenetic propagation, but the molecular-function aspect is where the annotation is genuinely incomplete.

Aspect	Term	Status for CRYAA	Recommended action
MF	GO:0051082 unfolded protein binding	**Missing**; directly assayed	**ADD** (lead — curator verify; IDA candidate P 8943244)
MF	GO:0140662 ATP-dependent protein folding chaperone	Not applicable (no NBD)	**Do not add**
MF	GO:0044183 protein folding chaperone	Arguable; less specific	Optional; GO:0051082 preferred
BP	GO:0042026 protein refolding	Present (IBA + ISS)	**Retain** as non-core, cooperation-dependent
BP	GO:0050821 protein stabilization	Present (IMP, P 12235146)	Retain — well supported

Evidence Base

Citation (PMID)	Evidence type	Supports / Refutes / Qualifies	Claim tested	Key finding	Context	Confidence & limitations
1438232	Direct in vitro assay	**Supports**	α-crystallin assists refolding of a denatured client	α-crystallin prevented aggregation of and refolded GdnHCl-denatured γ-crystallin (by CD)	Bovine/ recombinant α-crystallin, in vitro	Moderate; αA/αB mixture, CD readout, no downstream-chaperone controls
8093612	Direct in vitro assay	**Supports (family-level)**	sHSPs promote ATP-independent refolding	αB/Hsp25/ Hsp27 promote functional refolding after urea denaturation, ATP-independently	Recombinant sHSPs, in vitro	Moderate; not αA-specific; underpins IBA
8943244	Direct assay (human gene product)	**Supports (holdase MF)**	Human αA-crystallin has chaperone-like activity	Cloning/ expression demonstrated chaperone-like (aggregation-suppressing) activity of human αA	Recombinant human αA-crystallin	High for holdase; does not itself prove autonomous foldase
34055885	Review / mechanistic synthesis	**Qualifies / partially refutes**	sHSPs autonomously refold clients	sHSP assemblies facilitate *subsequent Hsp70/ Hsp100-dependent*	Bacterial sHSP network (general model)	High; reframes refolding as downstream

Citation (PMID)	Evidence type	Supports / Refutes / Qualifies	Claim tested	Key finding	Context	Confidence & limitations
				disaggregation and refolding		
35281256	Review (human HSPBs)	**Qualifies**	HSPBs' direct role is holdase	HSPBs act as holdases and *cooperate* in refolding driven by other chaperones	Human HSPB family	High; review-level but directly on-target
33321054	Review (α-crystallins)	**Qualifies**	α-crystallin catalytic role	α-crystallins are *holdase* chaperones; prevent aggregation via client binding	Vertebrate eye lens	High; consistent with holdase model
12235146	Mutant phenotype (IMP)	Supports (protein stabilization)	αA stabilizes client proteins in vivo	Basis for GO:0050821 protein stabilization annotation	Cellular	Cited via QuickGO annotation record
38401625	Structural/ biophysical	Qualifies	Mechanism of client handling	α-crystallins co-aggregate with saturating client; dynamic oligomer expansion	Recombinant αAc/αBc + model clients	Supports sequestration (holdase) mechanism
39947755	Review/ biophysical	Qualifies	Structural basis of chaperone role	Intrinsically disordered, dynamic oligomers stop	α-crystallin	Supports holdase; not foldase

Citation (PMID)	Evidence type	Supports / Refutes / Qualifies	Claim tested	Key finding	Context	Confidence & limitations
				denatured proteins aggregating		

GO Curation Implications

Lead requiring curator verification.

1. **Retain GO:0042026 (protein refolding, BP, IBA / GO_REF:0000033).** The term is biologically defensible: the GO definition of "protein refolding" does not mandate ATP, family-level and αA-specific in vitro assays report apparent refolding assistance, and the IBA reflects a reasonable phylogenetic inference across the sHSP family. However, annotate/interpret it as a **non-core, cooperation-dependent biological process** — the productive refolding step is executed by downstream ATP-dependent chaperones, with CRYAA supplying the holdase contribution.
2. **ADD GO:0051082 (unfolded protein binding, MF).** This is the highest-value action. CRYAA's directly assayed activity — binding unfolding clients to suppress aggregation — is a molecular function currently unrepresented in the MF aspect of P02489, even though the UniProt "Chaperone" keyword is assigned. A candidate experimental basis is PMID: 8943244 (human αA chaperone-like activity), which could support an IDA. This avoids leaving "protein binding" (GO:0005515) as the only MF chaperone-adjacent term.
3. **Do NOT add GO:0140662 (ATP-dependent protein folding chaperone) or other foldase MF terms.** CRYAA has no nucleotide-binding domain or ATPase activity; an ATP-dependent foldase annotation would be mechanistically wrong.
4. **Retain GO:0050821 (protein stabilization, BP, IMP,)**
   P 12235146
   — well-supported and complementary.

Aspect summary: the evidence most strongly supports a **molecular-function** term (GO:0051082) that is currently missing; the **biological-process** term in question (GO:0042026) should be retained but flagged as downstream/cooperative rather than a direct catalytic function.

Mechanistic Scope

The immediate molecular function being tested by the seed hypothesis is whether CRYAA **directly catalyzes the refolding of denatured/unfolded proteins to their native state**. Based on domain architecture and mechanistic literature:

- **Direct gene-product activity (established):** ATP-independent binding of partially unfolded/aggregation-prone client proteins → holdase/sequestration; prevention of irreversible aggregation. This is best captured by GO:0051082 (unfolded protein binding) and contributes to GO:0050821 (protein stabilization).
- **Downstream / cooperative process (the seed term):** Productive refolding of held clients, which requires downstream ATP-dependent chaperones (Hsp70/Hsp100). GO:0042026 (protein refolding) describes this outcome-level process, to which CRYAA contributes but does not autonomously execute.
- **Higher-order phenotypes (out of scope for this MF question):** maintenance of eye-lens transparency, prevention of cataract, antioxidant/metal-binding activity, anti-apoptotic roles. These are physiological consequences, not the molecular refolding activity per se.

The hypothesis therefore tests a **process** that CRYAA participates in, while the strongest, most specific evidence points to a **molecular function** (client binding) one mechanistic step upstream.

Conflicts and Alternatives

- **Holdase vs. foldase (primary conflict).** Early in vitro reports of "refolding" (PMID: 1438232; PMID: 8093612) conflict with the modern consensus that sHSPs are holdases whose clients are refolded by downstream ATP-dependent chaperones (PMID: 34055885; PMID: 35281256). The resolution is that "refolding" in the older assays is often an outcome

measured in systems where spontaneous or residual refolding can occur once aggregation is suppressed; it does not establish CRYAA as an active foldase.

- **Paralog / family generalization.** Much of the direct "refolding" evidence is from αB-crystallin, Hsp25, and Hsp27 rather than αA-crystallin specifically (PMID: 8093612). The IBA/ISS annotation on CRYAA propagates a family-level property. αA-specific direct refolding data are thinner than αA-specific holdase data (PMID: 8943244; PMID: 1438232).
- **Organism-specific differences.** Zebrafish work (PMID: 38705506) shows the αBa-crystallin paralog, not αA, dominates protection against age-related cataract, cautioning against over-reading αA's unique functional importance across vertebrates — though this concerns physiological role, not the molecular refolding question.
- **Assay-context artifacts.** CD-based "refolding" readouts on GdnHCl- or urea-denatured clients can reflect secondary-structure recovery upon aggregation suppression rather than chaperone-catalyzed folding. Recent single-particle work shows α-crystallins can co-aggregate with saturating client (PMID: 38401625), further arguing the core activity is sequestration.

Limitations and Knowledge Gaps

1. **αA-specific autonomous refolding.** *Checked:* literature reports refolding mainly for αB/Hsp25/27 and mixed α-crystallin. *Why it matters:* the IBA is on CRYAA specifically; direct αA-only, downstream-chaperone-free refolding assays would confirm or refute autonomous foldase activity. *Resolution:* purified human αA-crystallin refolding assay of a denatured client with and without Hsp70 system, quantifying native yield.
2. **MF annotation gap.** *Checked:* QuickGO enumeration shows no GO:0051082/0044183/0140662 despite UniProt "Chaperone" keyword. *Why it matters:* the molecular function is unrepresented, weakening the annotation set. *Resolution:* curator adds GO:0051082 with IDA from a primary αA chaperone assay (candidate P 8943244).
3. **Whether GO:0042026's ISS (GO_REF:0000024) adds independent support beyond the IBA.** *Checked:* both present on P02489. *Why it matters:* redundant/inference-only support affects confidence. *Resolution:* trace the ISS source protein/alignment.

4. **Quantitative contribution of CRYAA vs. downstream chaperones to net refolding in vivo.** *Why it matters:* determines whether "refolding" should ever be considered core for CRYAA. *Resolution:* reconstituted or cell-based flux assays with CRYAA knockdown.
5. **Provenance limitation of this review.** The domain-architecture facts (single ACD, no NBD, metal-binding sites) and the QuickGO annotation enumeration were used as the computational backbone; no independent autonomous refolding-catalysis assay specific to human αA (with downstream chaperones excluded) was located in the literature searched.

Proposed Follow-up Experiments / Actions

1. **Reconstituted refolding minus downstream chaperones.** Denature a model client (e.g., luciferase, GAPDH, γ-crystallin); measure native activity/structure recovery with purified human αA-crystallin **alone** vs. αA + Hsp70/Hsp40/nucleotide. If αA alone gives little native recovery but αA + Hsp70 does, CRYAA is a holdase, not a foldase — supporting a non-core reading of GO:0042026 and prioritizing GO:0051082.
2. **ATP-dependence control.** Confirm no ATPase activity and no ATP-stimulated refolding by CRYAA, consistent with the absence of a nucleotide-binding domain.
3. **Client-binding (holdase) assay for MF evidence.** Surface plasmon resonance / light-scattering suppression / co-sedimentation of CRYAA with a destabilized client to directly evidence GO:0051082 (unfolded protein binding).
4. **Domain/paralog comparison.** Align P02489 against αB (CRYAB) and Hsp27 (HSPB1) ACDs; confirm shared holdase determinants and shared absence of NBD, supporting family-level IBA while flagging that direct refolding data are largely paralog-derived.

Curation Leads (require curator verification)

Candidate action changes - **Retain** GO:0042026 (protein refolding, BP, IBA/GO_REF:0000033), reinterpreted as a **non-core, cooperation-dependent** process. - **Add** GO:0051082 (unfolded protein binding, MF) — the primary missing molecular-function term. - **Do not add** GO:0140662 (ATP-dependent protein folding chaperone) — mechanistically unsupported (no NBD).

Candidate references with snippets to verify - PMID: 8943244: *"Cloning, expression, and chaperone-like activity of human alphaA-crystallin"* → candidate IDA basis for GO:0051082 (human gene product holdase activity). - PMID: 1438232: *"alpha-Crystallin was also effective in*

preventing aggregation and in refolding guanidine hydrochloride-denatured gamma-crystallin, as judged by circular dichroism spectroscopy." → supports GO:0042026 (with holdase caveat). - PMID: 8093612: *"they promote the functional refolding of these proteins after urea denaturation similar to GroE and Hsp90. The interaction both with unfolding and refolding proteins seems to be ATP-independent."* → family-level basis for the IBA. - PMID: 34055885: *"Formation of these assemblies facilitates subsequent Hsp70 and Hsp100 chaperone-dependent disaggregation and substrate refolding into native species."* → justifies non-core reading of GO:0042026. - PMID: 35281256: HSPBs *"cooperate in substrates refolding driven by other chaperones."* → same.

Suggested questions for the curator - Should GO:0042026 remain when the direct molecular event is holdase binding, given the ISS + IBA both derive from inference? - Is there a primary αA-specific IDA paper suitable to anchor GO:0051082?
