

Bionamic Antibody Profiling

A developability, immunogenicity and humanness profile for a therapeutic antibody Fv, with every metric benchmarked against 906 clinical-stage molecules.

What goes in

A VH sequence, and a VL if the molecule is paired. Nothing else.

What comes back

A per-residue annotation matrix, 26 benchmarked molecule metrics, a predicted structure and four figures.

How long it takes

About two minutes for a paired Fv, with every model run offline against pinned weights.

WHAT IS MEASURED

Numbering & germline

Structure prediction

Solvent exposure

Aggregation

Surface hydrophobicity, stickiness & charge

PTM liabilities

MHC-II epitopes

Humanness

Isoelectric point & charge

CDR geometry

Clinical benchmarking

This document describes what BAP computes, which published method each number comes from, how the clinical-stage reference cohort was assembled, and how to read the result. It is written for scientists working on antibody discovery, engineering and developability assessment.

Contents

1 • Why profile developability before you commit	3
1.1 Where BAP fits in a campaign	3
2 • What BAP does	3
2.1 Inputs	4
2.2 Design commitments	4
3 • The annotation layer	4
3.1 Five numbering schemes at once	4
3.2 Regions and CDR boundaries	5
3.3 Germline assignment and the reverted variant	5
4 • Physicochemical properties	5
5 • Sequence liabilities	6
5.1 Grading a hit by its context	7
6 • Structure, exposure and aggregation	7
6.1 Folding the variable domain	7
6.2 Solvent accessibility	8
6.3 Aggregation propensity	8
6.4 Surface property maps	9
7 • Immunogenicity	9
7.1 The MHC-II scan	9
7.2 From windows to hotspots	10
7.3 The optional T-cell model	10
8 • Humanness	10
8.1 OASis: humanness against the natural repertoire	10
8.2 Sapiens: a proposed humanised variant	11
8.3 Germline identity	11
9 • Benchmarking against clinical-stage antibodies	11
9.1 The reference cohort	11
9.2 Percentiles and flags	12
9.3 The reference distributions	13
9.4 The comparison figures	13
10 • What a profile looks like	15
10.1 The per-residue matrix	15
10.2 The summary sheet	16
10.3 Attached files	16
11 • Worked example: trastuzumab	16
11.1 The benchmarked profile	16
11.2 What the benchmark cannot see	17
11.3 Humanness	18
12 • Reading a profile in practice	18
13 • Scope, assumptions and limitations	18
13.1 The structure is predicted	18
13.2 It is an Fv profile	18
13.3 Motifs are patterns, not rates	19
13.4 Presentation is not immunogenicity	19
13.5 The reference cohort is a population, not a standard	19
13.6 Percentiles are not thresholds	19
14 • Glossary	19
15 • References	20

1 • Why profile developability before you commit

A therapeutic antibody has to do two things: bind its target, and survive everything else. The second requirement is where programmes fail late and expensively. Molecules that express poorly, aggregate at formulation concentrations, isomerise in a CDR, carry an unexpected glycan, present T-cell epitopes, or clear too fast are usually discovered to have those properties long after the binding data made them look like winners.

Two studies define the field's reference frame. Jain and colleagues measured twelve biophysical assays across 137 clinical-stage antibodies and showed that approved molecules occupy a narrower region of property space than the clinical population as a whole, and that flagged outliers cluster in molecules that did not progress ¹. Raybould and colleagues then reduced that observation to five sequence- and structure-derived guidelines — total CDR length, patch hydrophobicity, patch positive and negative charge, and charge symmetry — computed from a model of the Fv and interpreted as percentiles against the therapeutic set rather than as hard thresholds ².

BAP follows that logic and extends it. Instead of asking whether a candidate passes a fixed cut-off, it asks **where this molecule sits relative to antibodies that have already been taken into humans**, on every axis at once, and **which residue is responsible**.

The two questions BAP answers

Is this molecule unusual? Twenty-six molecule-level metrics are reported as a value, the clinical median and interquartile range, and a percentile against 906 clinical-stage antibodies. Percentiles are context, not a verdict.

Where in the sequence does the risk sit? Every metric that can be localised — aggregation, exposure, MHC-II epitope coverage, humanness, PTM motifs, germline deviation — is also written per residue, aligned to five numbering schemes, so a finding points at a position you can mutate.

1.1 Where BAP fits in a campaign

BAP is a triage and characterisation instrument, not a decision procedure. It is inexpensive enough to run on every sequence that comes out of a screen, and detailed enough to brief a lead-optimisation discussion. Typical uses:

- **Panel triage.** Profile every hit from a campaign and rank by the number of tail and outlier flags before committing expression slots.
- **Liability review.** Identify CDR motifs that need a back-up variant designed before the lead is locked, when changing them is still cheap.
- **Humanisation support.** Compare OASis humanness and germline identity of a parental and humanised pair, and read the Sapiens-proposed sequence next to the original.
- **Developability dossier.** Attach the profile of a clinical candidate to a CMC or regulatory narrative, with every method traceable to a published reference.

The in-silico profile does not replace the wet-lab developability panel — HIC, SEC, DSF, AC-SINS, PSR, accelerated stability ^{1,35,36}. It decides which molecules are worth putting through it, and tells you what to look for when they come back.

2 • What BAP does

BAP takes the variable-domain sequences of an antibody and runs ten analysis stages over them, several of which require a three-dimensional model, and writes the result as a single workbook with a per-residue matrix and a molecule summary.

Table 1. The profiling pipeline. Every stage runs offline against pinned model weights, so a given sequence always produces the same profile.

Stage	What it does	Method
Numbering	Numbers both chains in five schemes and assigns the closest V and J germline alleles	ANARCI ⁴
Folding	Predicts the Fv structure (or the VHH domain) from sequence	ABodyBuilder2 / NanoBodyBuilder2 ¹⁰

Stage	What it does	Method
Exposure	Per-residue solvent accessibility on the model, absolute and relative	FreeSASA ¹¹ , Tien maxima ¹²
Aggregation	Per-residue structural aggregation propensity	Aggrescan3D ^{13,14}
Surface maps	Two-dimensional projections of hydrophobicity, stickiness and electrostatics, and per-residue surface patches	SURFMAP ¹⁶
Immunogenicity	MHC-II binding across a seven-allele DRB1 panel, tiled over the whole domain	MixMHC2pred ^{20,21}
Humanness	Nine-mer humanness against the natural repertoire, plus a humanised variant	BioPhi OASis and Sapiens ²⁴
Liabilities	Chemical-modification motifs, graded by surface exposure and germline context	Motif set, see §5
Physicochemistry	Mass, isoelectric point, net charge, Fv charge asymmetry	Sequence-derived, see §4
Benchmarking	Percentile and flag for 26 metrics against clinical-stage antibodies	Thera-SAbDab cohort ³

2.1 Inputs

The only required input is a heavy-chain variable domain. A light chain is supplied when the molecule is paired; omitting it puts BAP in single-domain mode, where the VHH is folded with NanoBodyBuilder2 and every light-chain metric is simply absent rather than imputed. Sequences are amino acids; whitespace and case are normalised. Constant regions are not required and are not modelled — the profile is an Fv profile.

Two options change what is produced. Surface mapping and the germline-reverted comparison are enabled when a content store is available to hold the rendered images. The TLImmuno2 immunogenicity model is off by default because it scores a 90 000-peptide background per allele and is slow without a GPU; when it is enabled its tracks appear alongside the MixMHC2pred ones rather than replacing them.

2.2 Design commitments

- **Offline and deterministic.** Every model runs against pinned weights with no network access. The same sequence gives the same profile today and in two years.
- **Versioned metrics.** Each metric family carries a version number that is written into the output, and the reference cohort records the versions it was computed with, so a benchmark is never silently compared across a definition change.
- **Nothing is imputed.** A metric that could not be computed — because folding failed, or a germline could not be assigned — is absent, not zero.
- **Per-residue first.** Molecule-level numbers are aggregates of a residue-level record that is kept and reported.

3 • The annotation layer

Everything downstream depends on knowing which residue is which. BAP numbers both chains before any metric is computed.

3.1 Five numbering schemes at once

ANARCI aligns each chain against germline HMMs and returns a numbering plus the closest V and J alleles ⁴. BAP runs it five times and keeps every scheme as its own track in the output matrix, so a position can be named in whichever convention the reader's team uses.

Table 2. Numbering schemes reported for every residue. IMGT is the scheme BAP's own region and CDR logic uses; the others are carried for cross-reference.

Scheme	Basis and typical use	Reference
IMGT	Structure-informed unique numbering, consistent across IG and TR V domains. The scheme BAP uses internally for CDR bounds, liability positions and all per-residue metrics.	Lefranc ⁵
Kabat	The classical sequence-variability scheme; still the lingua franca of patent and legacy literature.	Kabat ⁶
Chothia	Adjusts Kabat to structural loop positions, the basis of canonical class assignment.	Chothia & Lesk ⁷
Martin	Corrected and automated Chothia/Kabat numbering.	Abhinandan & Martin ⁸
AHo	Structurally aligned scheme with a fixed 149-position frame, convenient for loop-length comparison.	Honegger & Plückthun ⁹

3.2 Regions and CDR boundaries

Regions are assigned from the IMGT position number, which makes them independent of loop length: a residue is in a CDR if its IMGT number falls inside the CDR window, and in the framework otherwise.

Table 3. IMGT position windows used for region assignment. Insertion codes inside a window belong to that CDR, so the definition holds for loops of any length.

Region	IMGT positions	Clinical median length (H / L)	Notes
FR1	< 27	—	N-terminal framework; carries the pyroglutamate-forming residue in many molecules
CDR1	27 – 38	8 / 6	Length is strongly germline-determined
FR2	39 – 55	—	VH–VL interface residues
CDR2	56 – 65	8 / 3	CDR-L2 is 3 residues in essentially every clinical molecule
FR3	66 – 104	—	The longest framework stretch
CDR3	105 – 117	13 / 9	The diversity centre and the usual site of liabilities
FR4	> 117	—	J-derived framework

3.3 Germline assignment and the reverted variant

From the ANARCI V and J calls, BAP reconstructs the germline-encoded residue at every assessed position and reports two numbers per chain: the **somatic mutation count**, and **germline identity**, the fraction of assessed positions that still match the allele. Both are benchmarked. The per-residue germline track carries the germline residue at every assessed position and highlights the ones that differ from the candidate, which makes reversion candidates readable at a glance.

Germline context also feeds the liability grading in §5 — a motif that the germline itself encodes has been carried through human repertoires and is treated as less alarming than one introduced by engineering or somatic hypermutation.

When surface mapping is enabled, BAP additionally folds a **fully back-mutated variant** of the molecule and runs the same surface analysis over it. The workbook then shows each surface property twice, actual and germlined, so a hydrophobic or charged patch can be attributed to the somatic mutations that created it.

4 • Physicochemical properties

Four sequence-derived quantities describe the molecule's bulk behaviour. They are cheap, they are exact — no model is involved — and they correlate with properties that matter downstream: solubility, viscosity at high concentration, non-specific binding, and pharmacokinetics.

Table 4. Physicochemical metrics. The clinical column gives the median and interquartile range over the 906-molecule reference cohort.

Metric	Definition and why it matters	Clinical median (IQR)
Molecular weight	Monoisotopic mass of the variable domains, summed over both chains, from residue masses plus one water. A sanity check on the construct and the input for extinction-coefficient work.	24 889 Da (24 566 – 25 242)
Isoelectric point	The pH at which the Fv carries no net charge, found by bisection on the titration curve. Extremes of pI track with solubility minima near formulation pH and with faster clearance at the high end. ^{27,33}	8.26 (7.12 – 8.71)
Net charge @ pH 7.4	Titrateable charge at physiological pH. Strongly positive Fv charge is associated with non-specific tissue binding and faster clearance. ²⁷	+1.73 (–0.27 – +3.62)
Net charge @ pH 5.5	Charge at typical formulation pH, where high-concentration behaviour is decided.	+3.50 (+1.67 – +5.43)
Fv charge asymmetry	The product of the VH and VL net charges at pH 5.5. Large positive values mean both domains carry like charge; large negative values mean the domains are oppositely charged, which promotes self-association and elevated viscosity. ²⁷	+2.09 (–0.06 – +6.65)

Charges are computed from the Henderson–Hasselbalch relation over the titratable groups, with the termini included:

$$Z(pH) = \sum_{i \in \text{basic}} \frac{1}{1 + 10^{pH - pK_{a,i}}} - \sum_{j \in \text{acidic}} \frac{1}{1 + 10^{pK_{a,j} - pH}}$$

The pK_a set is the conventional one for polypeptides: N-terminus 8.6, C-terminus 3.6; Lys 10.8, Arg 12.5, His 6.5; Asp 3.9, Glu 4.1, Cys 8.5, Tyr 10.1.³³ These are sequence-intrinsic values — they do not account for the pK_a shifts a folded environment imposes, which is the standard trade-off for a metric that must be comparable across thousands of molecules.

Reading charge metrics

Net charge and charge asymmetry are **context** metrics, not **lower-is-better** ones. A strongly negative asymmetry and a strongly positive one both sit in the tails of the clinical distribution, and both have been associated with self-association. Read the percentile, and treat any position outside the 5th–95th percentile band as a question to answer experimentally rather than as a defect.

5 • Sequence liabilities

Certain short motifs are chemically labile. They deamidate, isomerise, fragment, oxidise or acquire a glycan during expression, purification, storage or circulation, and the result is charge heterogeneity, potency loss, or an unexpected species in the release assay. BAP scans both chains for fourteen such motifs.

Table 5. The liability motif set. Pattern notation is a regular expression over the amino-acid sequence; [^P] matches any residue except proline. Base severity is adjusted for structural and germline context before it is reported.

Type	Motif	Chemistry and consequence	Base severity	Reference
N-glycosylation	N[^P]S, N[^P]T	Asn-X-Ser/Thr sequon; an N-linked glycan is attached co-translationally unless X is proline. In a CDR this can change affinity and creates heterogeneity that is hard to control.	High	³⁰
Deamidation	NG	Asn-Gly is the fastest deamidating motif by an order of magnitude: succinimide formation gives Asp and iso-Asp, a –1 charge change and potential potency loss.	High	^{28,31}
Fragmentation	DG	Asp-Gly isomerises readily and the Asp-Gly bond is a hydrolysis hot spot; both give charge variants, the latter gives clips.	High	^{29,32}

Type	Motif	Chemistry and consequence	Base severity	Reference
Deamidation	NS, NN, NT, NH	Slower Asn deamidation contexts. Relevant when exposed and in a CDR, usually not otherwise.	Low	²⁸
Fragmentation	DS, DD, DN, DP	Asp isomerisation and acid-catalysed cleavage contexts; Asp-Pro is the classic low-pH clip site.	Low	²⁹
Oxidation	M	Methionine oxidises to the sulfoxide under peroxide and light stress; solvent-exposed Met in a CDR or at the VH–VL interface is the concern.	Low	³⁴
Oxidation	W	Tryptophan photo-oxidation gives kynurenine and related species, with colour and potency consequences.	Low	³⁴

5.1 Grading a hit by its context

A motif in the sequence is not automatically a liability. Whether it matters depends on whether the chemistry can reach it and whether nature already lives with it. BAP therefore grades each hit twice.

Surface. Every residue of the hit is looked up in the model's relative solvent accessibility. If any residue is exposed the hit is exposed; if all are buried it is buried; a mixture is partial. Buried motifs drop two severity levels, partial ones drop one — a buried methionine is not an oxidation risk in a folded domain.

Germline. The residues of the hit are compared against the assigned germline allele. A hit whose residues are all germline-encoded drops one further level, on the reasoning that the motif has been carried through human repertoires and is not something engineering introduced.

Table 6. Severity after context adjustment. Read down from the base severity of the motif and across by its context. Only motifs that reach **High** — an exposed, non-germline instance of a high-base motif — are counted in the benchmarked high-severity metric.

Base	Exposed, non-germline	Exposed, germline	Partial, non-germline	Partial, germline	Buried, non-germline	Buried, germline
High (NG, DG, sequons)	High	Medium	Medium	Low	Low	Benign
Low (M, W, slow motifs)	Low	Benign	Benign	Benign	Benign	Benign

Each hit is written into the per-residue matrix as `type:surface:germline` — for example `Deamidation:exposed:germline` — coloured by its adjusted severity, so the reason for the grade is visible next to the residue. Three counts are aggregated and benchmarked: total liabilities in the Fv, liabilities that overlap a CDR, and high-severity liabilities.

Why the CDR count is the one to watch

The clinical cohort carries a median of 13 motif hits per Fv, of which 3 fall in CDRs. Framework liabilities are ubiquitous and largely germline-encoded; they are not what separates molecules. A CDR liability sits in the binding site, so a modification there is the one most likely to cost potency — and the one most likely to be worth engineering out while a back-up variant is still cheap to make.

6 · Structure, exposure and aggregation

Several of BAP's metrics cannot be computed from sequence alone. They need a model of the folded Fv, and the quality of that model sets the quality of everything built on it.

6.1 Folding the variable domain

Paired Fv sequences are folded with ABodyBuilder2, and single-domain sequences with NanoBodyBuilder2 — the antibody-specific members of the ImmuneBuilder family.¹⁰ These are deep-learning structure predictors trained specifically on antibody variable domains; on their benchmark they reach roughly 1.0 Å RMSD on the framework and are competitive with much slower general predictors on the CDR loops, with CDR-H3 remaining the hardest

region. A prediction takes seconds rather than hours, which is what makes structure-dependent metrics affordable on a whole panel.

The predicted structure is kept and attached to the profile, so it can be inspected, coloured and re-used for docking or design work rather than being an invisible intermediate.

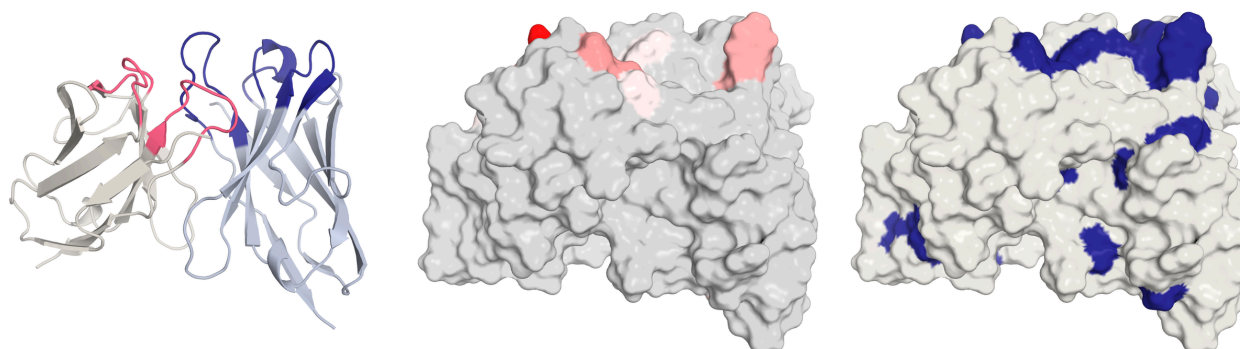


Figure 1. The same trastuzumab Fv model in three views. **Left:** chain trace with the IMGT CDRs highlighted — VH in blue, VL in pink. **Centre:** molecular surface coloured by Aggrescan3D score, grey for soluble and red for aggregation-prone; the strongest patch sits at the CDR apex. **Right:** solvent-exposed hydrophobic residues (rSASA > 0.075) in blue. All three are rendered from the model BAP produced, in identical orientation.

6.2 Solvent accessibility

Absolute and relative solvent-accessible surface area are computed per residue with FreeSASA¹¹ and normalised against the Tien et al. empirical maximum-accessibility set.¹² A residue counts as exposed above rSASA 0.075, the threshold that separates buried from accessible in that reference work.

Two aggregate metrics follow, and both are benchmarked:

- **Exposed hydrophobic residues** — the count of Ala, Ile, Leu, Met, Phe, Trp, Val and Tyr above the exposure threshold. Surface hydrophobicity drives self-association, HIC retention and non-specific binding.¹¹⁵ Clinical median 34, IQR 32–37.
- **Exposed CDR fraction** — the fraction of CDR residues that are exposed. A low value means an unusually occluded paratope; a high one, an unusually open one. Clinical median 0.725, IQR 0.698–0.755, and a remarkably tight distribution.

rSASA also does quiet work elsewhere in the profile: it is what decides whether a liability motif is exposed or buried (§5.1), and it is written as its own per-residue track so a surface argument can be checked residue by residue.

6.3 Aggregation propensity

Aggregation is scored with Aggrescan3D, which projects an experimentally derived aggregation-propensity scale onto the folded structure and averages it over each residue's spatial neighbourhood.^{13,14} The result is a per-residue score in which positive values mark aggregation-prone surface and negative values mark soluble surface. Because the averaging is spatial rather than sequential, A3D finds patches that sequence-only methods miss — residues that are far apart in sequence but adjacent on the surface — which is the same insight behind spatial aggregation propensity.¹⁵

Three metrics are reported: the **A3D total** over all positive residues, the **count of aggregation-prone residues** (score above zero), and the **A3D total restricted to CDRs**. The CDR-restricted number is the informative one for a binder, because an aggregation patch in the paratope cannot be engineered away without touching the binding site.

Table 7. Aggregation metrics against the clinical cohort. The distributions are wide — clinical antibodies tolerate a great deal of aggregation-prone surface — so a single high value is a prompt to look at the per-residue trace, not a rejection.

Metric	Clinical median	IQR	5th–95th percentile
Aggregation (A3D total)	14.15	10.62 – 18.16	6.85 – 24.67
Aggregation-prone residues	24	20 – 28	15 – 34
Aggregation in CDRs (A3D)	6.40	3.74 – 9.48	1.28 – 16.22

6.4 Surface property maps

Aggregation and exposure describe how much risky surface there is. SURFMAP describes **where it is**. It computes the molecular surface, evaluates a property at every surface particle, and projects the whole surface onto a single two-dimensional map, so patches that are hidden on the far side of a three-dimensional rendering are all visible at once.¹⁶

BAP renders three properties:

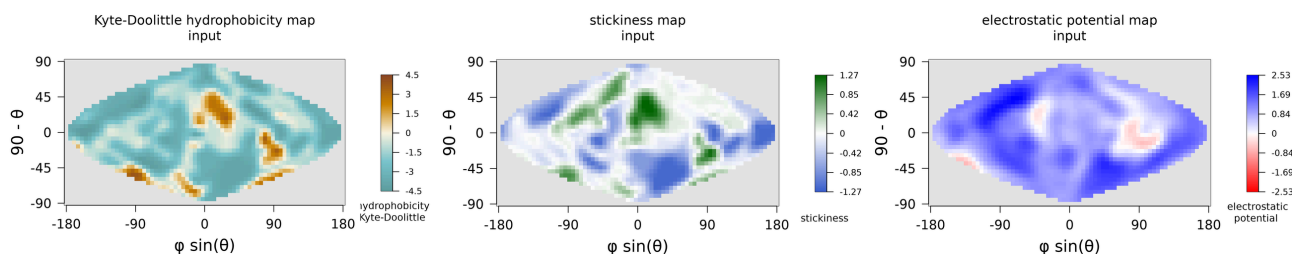


Figure 2. Three surface properties of the trastuzumab Fv projected onto a two-dimensional map. Each map covers the entire molecular surface; the axes are spherical coordinates, and contiguous coloured regions are patches. **Left:** Kyte–Doolittle hydrophobicity. **Centre:** stickiness, an empirical propensity for protein–protein contact. **Right:** Poisson–Boltzmann electrostatic potential.

- **Hydrophobicity** on the Kyte–Doolittle scale¹⁷ — the classical hydrophathy index, read here as surface patches rather than as a sequence plot.
- **Stickiness**¹⁸ — a residue-level propensity for non-specific protein–protein contact, derived from how often each amino acid is found in crystallographic interfaces relative to its surface abundance. It is the property most directly related to non-specific binding.
- **Electrostatics** — the Poisson–Boltzmann potential computed with APBS,¹⁹ which shows charge patches that the net-charge number in §4 averages away.

In the workbook these maps are also cut into **per-residue patches**: for every residue, the region of the map belonging to it is rendered as a small heat tile and placed in that residue’s column. The effect is that the surface property becomes another track alongside the sequence, and can be read against CDRs, liabilities and epitope coverage in the same view.

When germline reversion is available, every surface track is paired with the same track computed on the fully back-mutated molecule. A patch that is present in the candidate and absent in its germline ancestor was created by the engineering or by somatic hypermutation — and can usually be removed the same way.

7 • Immunogenicity

Anti-drug antibodies remain one of the least predictable failure modes in biologics development. The mechanistic prerequisite is well understood: the therapeutic must be taken up by an antigen-presenting cell, processed, and displayed on MHC class II to CD4⁺ T cells. BAP profiles that first step — which peptides of the variable domain a common set of HLA-DR alleles can present.

7.1 The MHC-II scan

The whole variable domain of each chain is tiled into 15-residue windows at every offset and each window is scored against a seven-allele DRB1 panel with MixMHC2pred, a predictor built by deconvolving motifs from mass-spectrometry immunopeptidomes rather than from binding assays.^{20,21} Scores are reported as a **%rank** — the percentile of the peptide against a background of random peptides for that allele — which is the conventional, allele-comparable scale.²²

DRB1*01:01 DRB1*03:01 DRB1*04:01 DRB1*07:01 DRB1*11:01 DRB1*13:01 DRB1*15:01

Figure 3. The default DRB1 panel. These seven alleles are the standard broad-coverage set and together account for roughly 90% of HLA-DR haplotypes in European-ancestry populations. A different panel can be supplied when a specific population matters.

Two thresholds are applied to the best rank across the panel: %rank ≤ 2 is a **strong binder**, %rank ≤ 10 a **weak binder**.²² Only strong binders contribute to the per-residue picture.

7.2 From windows to hotspots

Because windows are tiled at every offset, a residue is covered by up to fifteen of them. BAP counts, for each residue, how many strong-binding windows contain it — the **epitope coverage**. A residue covered by many strong binders is one that participates in several independently predicted epitopes, which is a stronger signal than any single window.

A run of consecutive residues that are each covered by five or more strong-binding windows is called an **epitope hotspot**, and is reported with its peak %rank and the allele responsible. Three molecule metrics follow, all benchmarked: strong binders, hotspots, and peak coverage.

Table 8. Immunogenicity metrics against the clinical cohort. Clinical antibodies are not epitope-free — the median molecule contains 28 strong-binding windows and two hotspots. The question is whether a candidate is unusual, and where its hotspots sit.

Metric	Clinical median	IQR	5th–95th percentile
MHC-II strong binders	28	24 – 31	18 – 37
Epitope hotspots	2	2 – 3	1 – 5
Peak epitope coverage	7	6 – 7	5 – 8

7.3 The optional T-cell model

MHC-II binding is necessary but not sufficient: a presented peptide only matters if a T-cell receptor responds to it. TLImmuno2 addresses that second step directly, using transfer learning from a binding model to predict immunogenicity of peptide–MHC-II pairs.²³ BAP can run it over 12-mer windows across the same seven alleles, producing a parallel set of coverage, hotspot and binder tracks that are shown **next to** the MixMHC2pred ones rather than replacing them — two models that disagree is itself informative.

It is off by default: it scores a large peptide background per allele and is slow without a GPU.

What an epitope prediction does and does not tell you

A hotspot says a peptide from this molecule is likely to be presented by a common DR allele. It does not say a patient will raise anti-drug antibodies. Real immunogenicity also depends on T-cell repertoire and tolerance, uptake and processing, dose, route, aggregation state, and the presence of impurities. Treat hotspots as ranked hypotheses for a T-cell assay — and give weight to hotspots that fall in CDRs or in engineered framework, where tolerance is least likely to have been established.

8 • Humanness

Human-derived sequence is the most reliable single lever on immunogenicity risk: clinical experience across formats shows lower anti-drug antibody incidence as sequence origin moves from murine towards fully human.²⁶ BAP measures humanness in three complementary ways.

8.1 OASis: humanness against the natural repertoire

OASis, from BioPhi, cuts the chain into overlapping 9-mer peptides and asks, for each one, what fraction of human subjects in the Observed Antibody Space repertoires²⁵ contain that peptide.²⁴ A peptide seen in at least 10% of subjects counts as human. Two numbers come out:

- **OASis identity** — the fraction of the chain's 9-mers that are human by that criterion. An absolute, interpretable measure.
- **OASis percentile** — where that identity sits in a reference distribution of therapeutic antibodies, on a 0–100 scale.

The per-9-mer breakdown is kept, so the workbook can show **which** stretches are unusual, as a tiled heat track under the sequence. Non-human stretches are typically short and clustered — and are exactly the stretches a humanisation campaign should target.

8.2 Sapiens: a proposed humanised variant

Sapiens is a deep language model trained on the natural repertoire; it proposes the substitutions that make a chain look more like its human neighbours.²⁴ BAP reports the count of proposed mutations per chain, marks each proposed position in the per-residue matrix, and writes the full humanised sequence into the summary. A low count means the chain is already close to the human distribution; a high count is a map of where a humanisation effort would start.

8.3 Germline identity

The third measure is the classical one — the fraction of positions still matching the assigned germline V allele (§3.3). Clinical medians are 0.907 for VH and 0.929 for VL, with the 5th percentile at 0.794 and 0.837 respectively. It is complementary to OASis: germline identity compares against a single ancestral allele, OASis against the whole observed repertoire, and a molecule can score well on one and poorly on the other.

Humanness is reported but not benchmarked

OASis and Sapiens numbers appear in the summary and in the per-residue matrix, but they are deliberately **not** part of the 26-metric clinical comparison. The reference cohort is itself dominated by humanised and human molecules with high OASis scores, so a percentile against it would compress the range and add little. Germline identity and somatic mutation count **are** benchmarked, and carry the humanness signal into the comparison.

9 • Benchmarking against clinical-stage antibodies

A developability number in isolation means very little. An A3D total of 15.6 is neither good nor bad until you know that half of all antibodies taken into the clinic score between 10.6 and 18.2. Every quantitative claim BAP makes is therefore made relative to a fixed cohort of molecules that have already been through the filter that matters.

9.1 The reference cohort

The cohort is derived from Thera-SAbDab, the therapeutic structural antibody database, which tracks antibody therapeutics through the clinic and links them to their sequences.³ BAP's snapshot retains molecules with both a heavy and a light chain, excludes chains belonging to bispecific arms, restricts to human sequence origin, and deduplicates on the VH/VL pair. That leaves **906 molecules**, each profiled with exactly the same code that profiles a candidate. The snapshot in use was taken on 14 June 2026 and its provenance and metric versions are recorded inside every profile.

Table 9. Cohort by molecular format.

Format	n
Whole mAb (IgG)	779
Whole mAb, ADC	55
Fab and Fab derivatives	18
Whole mAb, fusion	15
Fusion protein	14
scFv and Fv derivatives	10
Whole mAb, radiolabelled	8
Multispecific parent	5
Other	2

Table 10. Cohort by most advanced clinical stage.

Stage	n
Phase II	293
Phase I (incl. I/II)	178
Phase III (incl. II/III)	177
Approved	169
Not stated	68
Preregistration	19
Preclinical	2

Two properties of this cohort deserve to be stated plainly. It is **survivorship-weighted**: these are molecules that reached the clinic, so the distributions describe what is tolerable, not what is optimal. And it is **format-skewed** towards whole IgG, so a single-domain or fragment candidate is being compared against a population it does not belong to — informative, but read with that in mind.

9.2 Percentiles and flags

For each of the 26 benchmarked metrics, BAP reports the candidate's value, the cohort median and interquartile range, and the percentile — the share of cohort molecules at or below the candidate's value. A flag is then assigned from the tails of the cohort distribution:

Table 11. Flag rules. The bands are cohort percentiles of the metric, not of a theoretical distribution, so they make no normality assumption.

Flag	Condition	Reading
typical	within the 5th–95th percentile	Inside the range where most clinical molecules sit. Not evidence of quality — most metrics are typical for most molecules.
tail	below the 5th or above the 95th	Unusual. Worth an explanation, and worth checking against the per-residue trace to see whether one position is responsible.
outlier	below the 1st or above the 99th	Fewer than one clinical molecule in fifty looks like this. Treat as a finding that must be resolved experimentally or engineered out.

There is deliberately no overall score, no pass/fail and no weighting. Different programmes weight these axes differently — a subcutaneous high-concentration product cares about charge asymmetry in a way an intravenous oncology antibody does not — and a composite number would hide exactly the information that makes the profile useful.

9.3 The reference distributions

Table 12. The full clinical-stage reference distribution, n = 906. Metrics computed from the folded structure have n = 903, the three molecules for which folding did not complete. Germline identity is reported only where ANARCI assigned a V allele that is present in the bundled germline tables.

Metric	Median	Mean	IQR	5th–95th pct	n
Molecular weight (Fv)	24 889	24 922	24 566 – 25 242	24 154 – 25 783	906
Isoelectric point (Fv)	8.26	7.77	7.12 – 8.71	4.96 – 9.20	906
Net charge @ pH 7.4	1.73	1.43	–0.27 – 3.62	–3.27 – 5.82	906
Net charge @ pH 5.5	3.50	3.69	1.67 – 5.43	–1.47 – 8.29	906
Fv charge asymmetry	2.09	3.24	–0.06 – 6.65	–6.28 – 16.40	906
CDR-H1 length	8	8.15	8 – 8	8 – 10	906
CDR-H2 length	8	7.89	8 – 8	7 – 8	906
CDR-H3 length	13	12.78	11 – 15	8 – 19	906
CDR-L1 length	6	7.41	6 – 9	5 – 12	906
CDR-L2 length	3	3.00	3 – 3	3 – 3	906
CDR-L3 length	9	9.21	9 – 9	8 – 11	906
Somatic mutations (VH)	8	9.16	3 – 14	0 – 21.75	906
Germline identity (VH)	0.907	0.897	0.858 – 0.944	0.794 – 0.991	755
Somatic mutations (VL)	5	5.90	0 – 10	0 – 15	906
Germline identity (VL)	0.929	0.925	0.888 – 0.969	0.837 – 1.000	719
Liabilities (Fv)	13	13.25	11 – 15	9 – 18	906
Liabilities in CDRs	3	3.57	2 – 5	1 – 7	906
High-severity liabilities	0	0.16	0 – 0	0 – 1	906
MHC-II strong binders	28	27.53	24 – 31	18 – 37	906
Epitope hotspots	2	2.47	2 – 3	1 – 5	906
Peak epitope coverage	7	6.75	6 – 7	5 – 8	906
Aggregation (A3D total)	14.15	14.79	10.62 – 18.16	6.85 – 24.67	903
Aggregation-prone residues	24	24.14	20 – 28	15 – 34	903
Aggregation in CDRs (A3D)	6.40	7.20	3.74 – 9.48	1.28 – 16.22	903
Exposed hydrophobic residues	34	34.49	32 – 37	29 – 40	903
Exposed CDR fraction	0.725	0.727	0.698 – 0.755	0.667 – 0.795	903

9.4 The comparison figures

Two kinds of figure make the comparison visual. The **developability radar** places the candidate against the cohort on ten axes, scaled so that outward always means higher risk or more atypical. The grey web behind it is every one of the 906 reference molecules drawn individually, which shows the shape of the population rather than only its median.

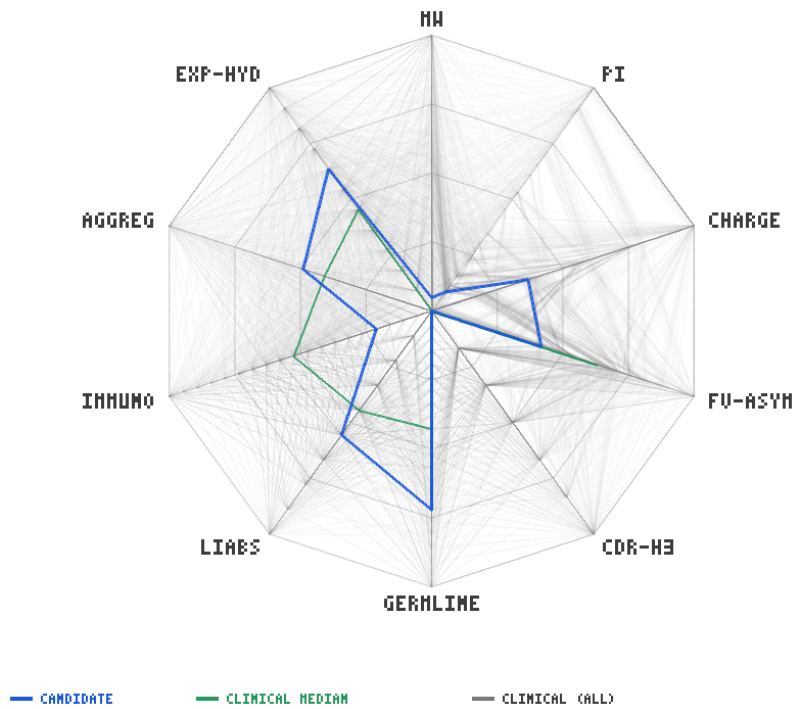


Figure 4. Developability radar for trastuzumab. Blue is the candidate, green the clinical median, grey the 906 individual reference molecules. Outward is higher risk or more atypical on every axis.

The **per-residue traces** do the same thing at residue resolution: the candidate's aggregation, exposure or epitope-coverage profile is drawn over the aligned profiles of the reference set, position by position, with CDRs shaded. A peak that rises above the reference envelope names the residue responsible.

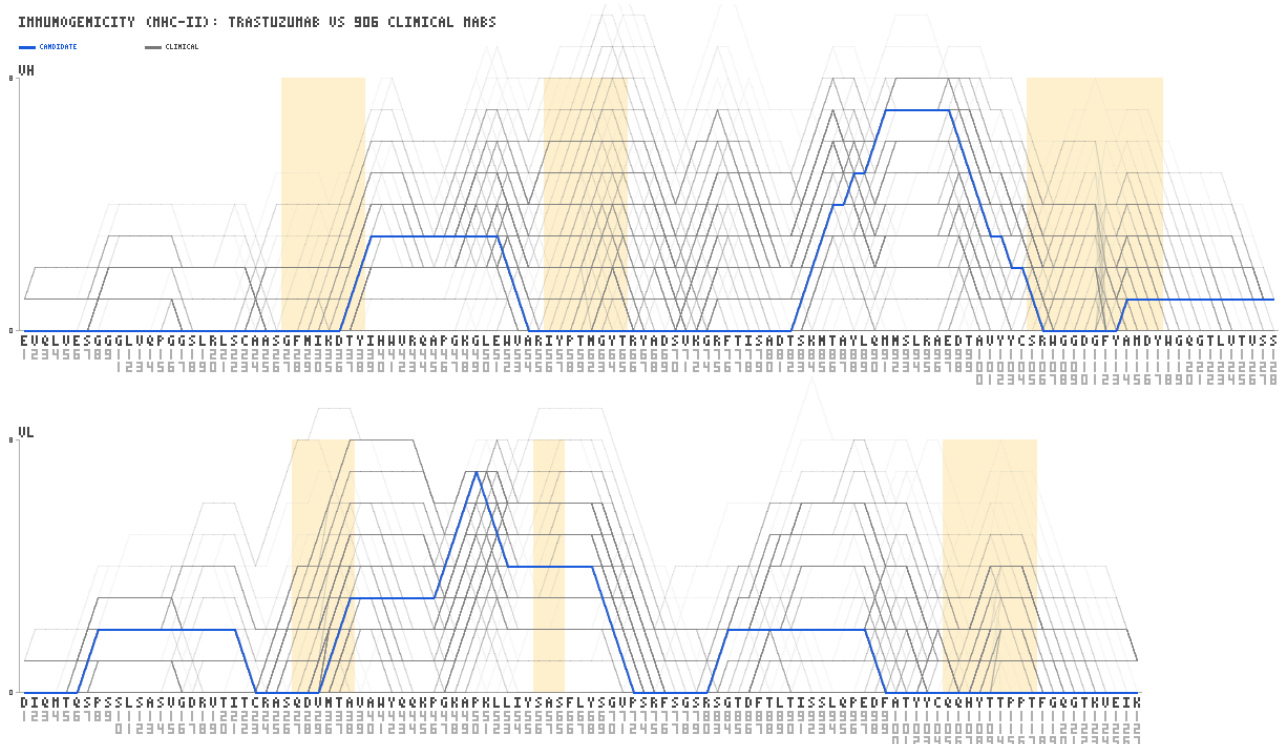


Figure 5. Per-residue MHC-II epitope coverage for trastuzumab (blue) over the aligned clinical reference set (grey), for VH above and VL below. Shaded bands are the IMGT CDRs. The tall central peak in VH is a framework-3 hotspot; the CDR-H3 region is comparatively quiet.

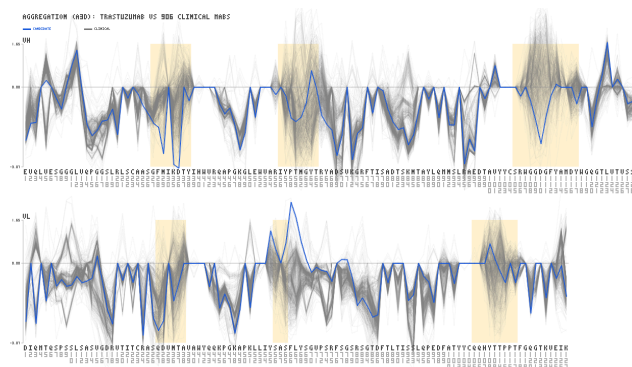


Figure 6. Aggregation trace.

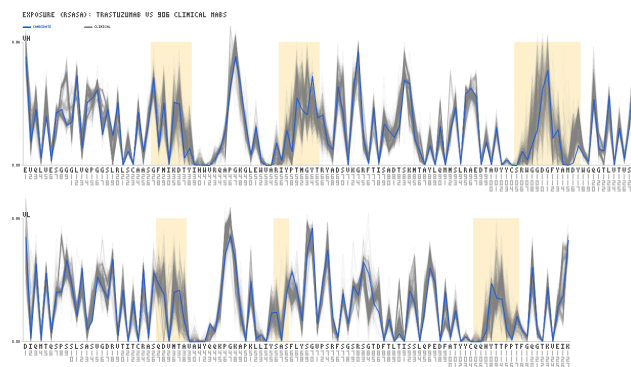


Figure 7. Exposure trace.

10 · What a profile looks like

A profile is delivered as a workbook with two sheets — a per-residue matrix and a molecule summary — plus a set of attached files.

10.1 The per-residue matrix

The matrix sheet has one column per residue and one row per track, grouped into sections. Every number that BAP localises appears here, aligned to the sequence, so the tracks can be read against each other.

Table 13. An authentic excerpt of the matrix: twelve residues of trastuzumab CDR-H3, reproduced from the workbook with its own colouring. Reading down — five numbering schemes (one shown), region, liabilities, MHC-II coverage, the residue, the germline residue, OAS 9-mer frequency, relative exposure, Aggrescan3D score, and the three SURFMAP surface-patch tiles. The window sits inside CDR-H3, whose merged region label begins to the left of it. The red Frag cells are the Asp-Gly at IMGT 110–111: an exposed, non-germline, high-severity fragmentation and isomerisation motif in the middle of the paratope.

IMGT	107	108	109	110	111	112	113	114	115	116	117	118
CDR (IMGT)												Fr4
Liability	Oxid			Frag					Oxid			Oxid
MHC-II coverage	0	0	0	0	0	0	0	1	1	1	1	1
Sequence	W	G	G	D	G	F	Y	A	M	D	Y	W
Germline												W
OAS residue freq %	1	16	17	7	15	2	25	15	16	84	42	99
rSASA %	4	15	25	53	66	19	25	1	0	2	13	7
Aggregation (A3D)	0.0	-0.6	-1.3	-2.2	-1.1	-0.3	0.1	0.0	0.0	0.0	-0.6	0.0
Hydrophobicity												
Stickiness												
Electrostatics												

Table 14. Sections and tracks of the matrix sheet. Tracks that depend on an optional stage are present only when that stage ran.

Section	Tracks
Numbering	IMGT, Kabat, Chothia, Martin, AHo — one row each
Features	CDR (IMGT) region assignment; Liability, labelled type:surface:germline and coloured by adjusted severity
Immunogenicity	MixMHC2pred hotspot (labelled with the peak allele) and MixMHC2pred coverage; the same pair for TLimmuno2 when enabled
Sequence	The residue
Germline	The germline residue at each assessed position, highlighted where it differs from the candidate, under the assigned V and J allele pair
Humanness	Sapiens proposed substitution; OAS residue frequency; nine offset rows tiling every OASis 9-mer with its subject frequency
Structure	rSASA %; Aggrescan3D score; SURFMAP hydrophobicity, stickiness and electrostatics patches, each paired with its germline-reverted counterpart

10.2 The summary sheet

The summary sheet is the one-page readout. It carries the input sequences, a codon-optimised DNA sequence for each chain, the humanness numbers and the Sapiens-humanised sequences, links to every attached file, and a machine-readable molecule record with CDR and liability ranges. Below that sits the `vs. clinical mAbs (n=906)` block: one row per benchmarked metric, formatted as **value · clinical median (IQR) · Nth percentile**, tinted and marked when the metric lands in a tail or beyond.

10.3 Attached files

Table 15. Files produced alongside the workbook.

Artifact	Contents
Predicted structure (PDB)	The folded Fv or VHH, ready for inspection, colouring, docking or design work
Aggrescan3D table (CSV)	Per-residue aggregation scores as produced by A3D
Developability radar (PNG)	The candidate against all 906 reference molecules on ten axes
Three trace plots (PNG)	Aggregation, exposure and immunogenicity per residue, over the aligned reference set
Metrics record (JSON)	Every molecule-level metric and every per-residue metric, with the version of each metric family and a list of any that were skipped
Molecule record (JSON)	The antibody as a structured value: chains, CDR ranges and liability ranges as features

The metrics record is the machine-readable form of the whole profile. It is the natural input to a ranking script, a dashboard, or a downstream model, and it is what makes a profile comparable across a panel without re-parsing workbooks.

11 · Worked example: trastuzumab

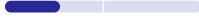
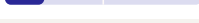
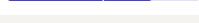
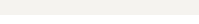
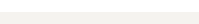
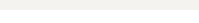
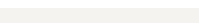
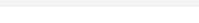
Trastuzumab is a useful test case because it is a humanised murine antibody with two decades of published characterisation behind it, so the profile can be checked against what is known. The input is its variable domains and nothing else.

```
VH  EVQLVESGGGLVQPGGSLRLSCAASGFMNIKDTYIHVWRQAPGKGLEWVARIYPTNGYTRYADSVKGRFTISADTSKNTAYLQMNSLRAEDTAVYYCSRWG-  
GDGFYAMDYWGQGLTVTVSS  
VL  DIQMTQSPSSLSASVGRVTITCRASQDVNTAVAWYQQKPGKAPKLLIYSASFVSGVPSRFSGRSGTDFTLTISLQPEDFATYYCQHHYTPPT-  
FGQGTKVEIK
```

11.1 The benchmarked profile

Table 16. Trastuzumab against 906 clinical-stage antibodies. The meter shows the percentile; the tick marks the cohort median. Every one of the 26 metrics falls inside the 5th–95th percentile band.

Metric	Value	Clinical IQR	Percentile	Flag
Molecular weight (Fv)	24 845	24 566 – 25 242	 46	typical
Isoelectric point (Fv)	8.54	7.12 – 8.71	 66	typical
Net charge @ pH 7.4	2.84	–0.27 – 3.62	 69	typical
Net charge @ pH 5.5	5.32	1.67 – 5.43	 70	typical
Fv charge asymmetry	6.92	–0.06 – 6.65	 77	typical
CDR-H1 length	8	8 – 8	 91	typical
CDR-H2 length	8	8 – 8	 95	typical
CDR-H3 length	13	11 – 15	 62	typical
CDR-L1 length	6	6 – 9	 60	typical
CDR-L2 length	3	3 – 3	 100	typical
CDR-L3 length	9	9 – 9	 82	typical

Metric	Value	Clinical IQR	Percentile	Flag
Somatic mutations (VH)	16	3 – 14	 83	typical
Germline identity (VH)	0.850	0.858 – 0.944	 21	typical
Somatic mutations (VL)	11	0 – 10	 83	typical
Germline identity (VL)	0.890	0.888 – 0.969	 28	typical
Liabilities (Fv)	14	11 – 15	 69	typical
Liabilities in CDRs	5	2 – 5	 86	typical
High-severity liabilities	1	0 – 0	 99	typical
MHC-II strong binders	22	24 – 31	 20	typical
Epitope hotspots	2	2 – 3	 51	typical
Peak epitope coverage	7	6 – 7	 76	typical
Aggregation (A3D total)	15.57	10.62 – 18.16	 60	typical
Aggregation-prone residues	22	20 – 28	 40	typical
Aggregation in CDRs (A3D)	2.92	3.74 – 9.48	 18	typical
Exposed hydrophobic residues	36	32 – 37	 71	typical
Exposed CDR fraction	0.700	0.698 – 0.755	 29	typical

Read as a whole, the profile says that trastuzumab is an unremarkable molecule on every bulk axis. Mass, isoelectric point, charge, CDR lengths, aggregation, exposure and epitope load all sit near the middle of the clinical population. The three metrics that sit furthest out are all telling the same story: 16 somatic mutations in VH and 11 in VL put it at the 83rd percentile for mutational load, and germline identity of 0.85 and 0.89 puts it at the 21st and 28th percentiles — the signature of a humanised murine antibody, whose CDRs are mouse-derived on a human framework.

One row deserves a second look. High-severity liabilities reads **1 • clinical 0 (0–0) • 99th percentile** — and is nevertheless unflagged, because the cohort's own 95th percentile is 1, so a single high-severity motif does not lie above the band. That is the clearest illustration of what a percentile does and does not say: it locates a molecule in a population, and when the population is overwhelmingly at zero, being at the top of it is common enough to be ordinary. The finding it points at is not.

Its aggregation profile is, if anything, better than typical: an A3D total in the CDRs of 2.92 sits at the 18th percentile, meaning the paratope carries less aggregation-prone surface than most clinical molecules — visible in Figure 1 as a small red patch confined to the CDR apex.

11.2 What the benchmark cannot see

No metric was flagged, and yet the liability scan finds three motifs in the CDRs that any developability review of this molecule would raise:

Table 17. CDR liabilities found in trastuzumab, with the position in IMGT numbering and the context that set the grade.

Chain	IMGT	Motif	Call	Severity
VH	110–111	DG in CDR-H3	Asp–Gly, exposed, not germline-encoded — isomerisation and fragmentation	High
VH	62–63	NG in CDR-H2	Asn–Gly, exposed, germline-encoded — the fastest deamidation context	Medium
VL	36–37	NT in CDR-L1	Asn–Thr, exposed, germline-encoded — a slower deamidation context	Low

These correspond to modifications that were characterised in the marketed product: the isomerisation of the CDR-H3 aspartate and the deamidation of the CDR-L1 asparagine are among the documented sources of its charge heterogeneity.³² A molecule can be entirely typical on every distributional metric and still carry a well-known chemical liability in the middle of its binding site — which is why BAP reports the residue-level scan and the population comparison side by side, and why neither alone is a developability assessment.

11.3 Humanness

OASis places VH at the 27.6th percentile with 0.69 of its 9-mers seen in human repertoires, and VL at the 52.5th percentile with 0.83. Sapiens proposes seven substitutions in VH and one in VL. For a molecule humanised by CDR grafting in 1992 this is exactly the expected result: the frameworks are human, the CDR-flanking 9-mers are not, and a modern repertoire-based humanisation would push a handful of framework positions further towards the human consensus.

Every figure and every number in this section was produced by running BAP on the two sequences above, with no manual editing of the output. The reference cohort it is compared against is the 14 June 2026 Thera-SAbDab snapshot described in §9.1.

12 • Reading a profile in practice

A profile is a set of hypotheses ranked by how unusual they are. The following order of reading gets to the useful ones quickly.

1. **Scan the flags first.** Any metric marked as an outlier is the headline; a tail is worth an explanation. A molecule with no flags is not necessarily safe — it is merely unremarkable in bulk, as §11 shows.
2. **Read the liability scan, restricted to CDRs.** High-severity CDR hits are the findings most likely to cost potency or create heterogeneity, and the cheapest to engineer out early. Framework hits, especially germline-encoded ones, rarely need action.
3. **Localise every flag.** For any flagged aggregation, exposure or immunogenicity metric, go to the corresponding trace plot and find the residues responsible. A metric driven by one exposed patch is a design problem; the same metric driven by a diffuse elevation across the domain usually is not.
4. **Check the germline-reverted comparison.** If a hydrophobic or charged patch is present in the candidate and absent in its back-mutated form, the somatic or engineered mutations that created it are visible in the germline track and are candidate reversions.
5. **Decide what the wet lab has to answer.** Convert what survives into assays: an aggregation flag into SEC and accelerated stability, an exposed-hydrophobic flag into HIC and AC-SINS, a charge-asymmetry flag into a concentration–viscosity series, an epitope hotspot into a T-cell proliferation assay.^{1,36}

Comparing molecules, not judging them

BAP is at its most useful across a panel. Profiling every hit from a campaign and sorting by the count of tail and outlier flags — then, within that, by CDR liabilities — turns a set of sequences into a ranked list with a stated reason for every position in it. Against a single molecule, the profile is a characterisation document; against twenty, it is a triage instrument.

13 • Scope, assumptions and limitations

Every number in a profile carries assumptions. The ones that matter most are these.

13.1 The structure is predicted

All structure-derived metrics — exposure, aggregation, surface maps, and the exposure term in liability grading — are computed on a single predicted model. ABodyBuilder2 is accurate on frameworks and good on most loops, but CDR-H3 remains the least reliable region, and it is also the region where these metrics matter most.¹⁰ A profile describes one static conformation: there is no ensemble, no dynamics, and no modelling of the conformational change that accompanies antigen binding. Treat a marginal exposure or aggregation call in a long CDR-H3 with corresponding caution.

13.2 It is an Fv profile

Constant domains, the hinge, the Fc, glycans, and any Fc engineering are outside the scope. That matters in two directions. Properties dominated by the constant region — whole-molecule isoelectric point, FcRn binding, effector function, hinge fragmentation — are not addressed at all. And properties of the assembled molecule that emerge

from the Fv, such as chain-pairing behaviour in a bispecific, are not modelled: each arm is profiled as if it stood alone.

13.3 Motifs are patterns, not rates

The liability scan matches sequence patterns. Real modification rates depend on the neighbouring residues, local backbone flexibility, pH, temperature, excipients and time, and they span orders of magnitude within a single motif class.²⁸ The grading in §5.1 uses exposure and germline context as a proxy for that, but a High call is a prompt for a forced degradation study, not a measurement. The converse also holds: chemistry that has no short motif — lysine glycation, some tryptophan oxidation, C-terminal processing — will not appear.

13.4 Presentation is not immunogenicity

The MHC-II scan predicts which peptides a panel of seven DRB1 alleles can present. It does not model antigen uptake and processing, HLA-DP and HLA-DQ, class I presentation, the T-cell repertoire, or central and peripheral tolerance — and tolerance is precisely what makes a self-like human framework peptide harmless despite binding well. Anti-drug antibody risk is also driven by dose, route, duration, aggregation state and product impurities, none of which a sequence can express.²⁶

13.5 The reference cohort is a population, not a standard

The 906 molecules are those that reached the clinic with a paired human-origin sequence in a single snapshot of Thera-SAbDab. The cohort is survivorship-weighted, so its distributions describe what has been tolerated rather than what is desirable. It is dominated by whole IgG, so fragment and single-domain candidates are compared against a population they do not belong to. And percentiles move when the snapshot is refreshed — which is why the metric versions and the cohort provenance are written into every profile.

13.6 Percentiles are not thresholds

There is no pass mark. A 96th-percentile value is unusual, not disqualifying; several approved antibodies sit in the tails of several metrics. Equally, an unflagged profile is not a clean bill of health, as §11 demonstrates. The profile ranks and localises risk; the decision, and the experiment that resolves it, stay with the programme.

14 • Glossary

A3D — Aggrescan3D. Structure-based per-residue aggregation propensity; positive is aggregation-prone.

ADA — Anti-drug antibody. The clinical readout of immunogenicity.

DRB1 — The most polymorphic and most commonly typed HLA class II beta chain locus.

Germline identity — Fraction of assessed positions still matching the assigned germline V allele.

IMGT — The international ImMunoGeneTics numbering scheme; structure-informed and length-independent.

OASis — BioPhi's humanness measure: the fraction of a chain's 9-mers seen in a given share of human subjects.

rSASA — Relative solvent-accessible surface area, normalised to the residue's maximum; > 0.075 is exposed.

Sequon — The Asn-X-Ser/Thr motif that directs N-linked glycosylation when X is not proline.

Thera-SAbDab — The therapeutic structural antibody database; the source of the clinical reference cohort.

AC-SINS — Affinity-capture self-interaction nanoparticle spectroscopy — a wet-lab self-association assay.

CDR — Complementarity-determining region. The six hypervariable loops that form the binding site.

Fv — The variable fragment: the VH and VL domains that carry the binding site.

HIC — Hydrophobic interaction chromatography. Retention time reports surface hydrophobicity.

OAS — Observed Antibody Space — a curated collection of natural antibody repertoire sequences.

%rank — A peptide's percentile against a random-peptide background for one allele; ≤ 2 is a strong binder.

Sapiens — BioPhi's deep-learning humanisation model; proposes substitutions towards the human repertoire.

Stickiness — An empirical per-residue propensity for non-specific protein–protein contact.

VHH — The single variable domain of a heavy-chain-only antibody; a nanobody.

15 • References

1. Jain T, Sun T, Durand S, et al. Biophysical properties of the clinical-stage antibody landscape. *PNAS* 2017;114(5):944–949.
2. Raybould MIJ, Marks C, Krawczyk K, et al. Five computational developability guidelines for therapeutic antibody profiling. *PNAS* 2019;116(10):4025–4030.
3. Raybould MIJ, Marks C, Lewis AP, et al. Thera-SAbDab: the Therapeutic Structural Antibody Database. *Nucleic Acids Research* 2020;48(D1):D383–D388.
4. Dunbar J, Deane CM. ANARCI: antigen receptor numbering and receptor classification. *Bioinformatics* 2016;32(2):298–300.
5. Lefranc M-P, Pommié C, Ruiz M, et al. IMGT unique numbering for immunoglobulin and T cell receptor variable domains and Ig superfamily V-like domains. *Developmental & Comparative Immunology* 2003;27(1):55–77.
6. Kabat EA, Wu TT, Perry HM, Gottesman KS, Foeller C. *Sequences of Proteins of Immunological Interest*, 5th ed. NIH Publication No. 91-3242, 1991.
7. Chothia C, Lesk AM. Canonical structures for the hypervariable regions of immunoglobulins. *Journal of Molecular Biology* 1987;196(4):901–917.
8. Abhinandan KR, Martin ACR. Analysis and improvements to Kabat and structurally correct numbering of antibody variable domains. *Molecular Immunology* 2008;45(14):3832–3839.
9. Honegger A, Plückthun A. Yet another numbering scheme for immunoglobulin variable domains: an automatic modeling and analysis tool. *Journal of Molecular Biology* 2001;309(3):657–670.
10. Abanades B, Wong WK, Boyles F, Georges G, Bujotzek A, Deane CM. ImmuneBuilder: deep-learning models for predicting the structures of immune proteins. *Communications Biology* 2023;6:575.
11. Mitternacht S. FreeSASA: an open source C library for solvent accessible surface area calculations. *F1000Research* 2016;5:189.
12. Tien MZ, Meyer AG, Sydykova DK, Spielman SJ, Wilke CO. Maximum allowed solvent accessibilities of residues in proteins. *PLoS ONE* 2013;8(11):e80635.
13. Zambrano R, Jamroz M, Szczasiuk A, Pujols J, Kmiecik S, Ventura S. AGGRESAN3D (A3D): server for prediction of aggregation properties of protein structures. *Nucleic Acids Research* 2015;43(W1):W306–W313.
14. Kuriata A, Iglesias V, Pujols J, Kurcinski M, Kmiecik S, Ventura S. Aggrescan3D (A3D) 2.0: prediction and engineering of protein solubility. *Nucleic Acids Research* 2019;47(W1):W300–W307.
15. Chennamsetty N, Voynov V, Kayser V, Helk B, Trout BL. Design of therapeutic proteins with enhanced stability. *PNAS* 2009;106(29):11937–11942.
16. Schweke H, Mucchielli M-H, Chevrollier N, Gosset S, Lopes A. SURFMAP: a software for mapping in two dimensions protein surface features. *Journal of Chemical Information and Modeling* 2022;62(7):1595–1601.
17. Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology* 1982;157(1):105–132.
18. Levy ED, De S, Teichmann SA. Cellular crowding imposes global constraints on the chemistry and evolution of proteomes. *PNAS* 2012;109(50):20461–20466.
19. Jurrus E, Engel D, Star K, et al. Improvements to the APBS biomolecular solvation software suite. *Protein Science* 2018;27(1):112–128.
20. Racle J, Michaux J, Rockinger GA, et al. Robust prediction of HLA class II epitopes by deep motif deconvolution of immunopeptidomes. *Nature Biotechnology* 2019;37(11):1283–1286.
21. Racle J, Guillaume P, Schmidt J, et al. Machine learning predictions of MHC-II specificities reveal alternative binding mode of class II epitopes. *Immunity* 2023;56(6):1359–1375.
22. Reynisson B, Alvarez B, Paul S, Peters B, Nielsen M. NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation. *Nucleic Acids Research* 2020;48(W1):W449–W454.
23. Wang G, Wu T, Ning W, et al. TLImmuno2: predicting MHC class II antigen immunogenicity through transfer learning. *Briefings in Bioinformatics* 2023;24(3):bbad116.
24. Prihoda D, Maamary J, Waight A, et al. BioPhi: a platform for antibody design, humanization, and humanness evaluation based on natural antibody repertoires and deep learning. *mAbs* 2022;14(1):2020203.
25. Olsen TH, Boyles F, Deane CM. Observed Antibody Space: a diverse database of cleaned, annotated, and translated unpaired and paired antibody sequences. *Protein Science* 2022;31(1):141–146.
26. Harding FA, Stickler MM, Razo J, DuBridge RB. The immunogenicity of humanized and fully human antibodies. *mAbs* 2010;2(3):256–265.
27. Sharma VK, Patapoff TW, Kabakoff B, et al. In silico selection of therapeutic antibodies for development: viscosity, clearance, and disposition. *PNAS* 2014;111(52):18601–18606.
28. Lu X, Nobrega RP, Lynaugh H, et al. Deamidation and isomerization liability analysis of 131 clinical-stage antibodies. *mAbs* 2019;11(1):45–57.
29. Vlasak J, Ionescu R. Fragmentation of monoclonal antibodies. *mAbs* 2011;3(3):253–263.
30. Gavel Y, von Heijne G. Sequence differences between glycosylated and non-glycosylated Asn-X-Thr/Ser acceptor sites: implications for protein engineering. *Protein Engineering* 1990;3(5):433–442.
31. Robinson NE, Robinson AB. Molecular clocks. *PNAS* 2001;98(3):944–949.
32. Harris RJ, Kabakoff B, Macchi FD, et al. Identification of multiple sources of charge heterogeneity in a recombinant antibody. *Journal of Chromatography B* 2001;752(2):233–245.
33. Bjellqvist B, Hughes GJ, Pasquali C, et al. The focusing positions of polypeptides in immobilized pH gradients can be predicted from their amino acid sequences. *Electrophoresis* 1993;14(10):1023–1031.
34. Ji JA, Zhang B, Cheng W, Wang YJ. Methionine, tryptophan, and histidine oxidation in a model protein, PTH: mechanisms and stabilization. *Journal of Pharmaceutical Sciences* 2009;98(12):4485–4500.
35. Xu Y, Wang D, Mason B, et al. Structure, heterogeneity and developability assessment of therapeutic antibodies. *mAbs* 2019;11(2):239–264.
36. Bailly M, Mieczkowski C, Juan V, et al. Predicting antibody developability profiles through early stage discovery screening. *mAbs* 2020;12(1):1743053.