PDBePISA for building biological assemblies

PDBePISA is a sophisticated service that analyses the interfaces between macromolecules in their crystal environment. Each X-ray structure is solved in a particular crystal symmetry and the PDB requires authors to deposit only the smallest or **asymmetric** unit from any crystal form. As a result the PDB entry in the archive often does not contain the assembly that is most likely to exist under physiological conditions. In many cases crystal symmetry operations need to be applied in order to obtain the biologically active form of a macromolecule or complex.

Conversely some macromolecules crystallise with several copies of the biologically relevant assembly in the asymmetric unit. This means that the deposited entry contains extra copies of the biological assembly. Contacts in the crystal hold these copies alongside each other but are too weak to survive in solution.

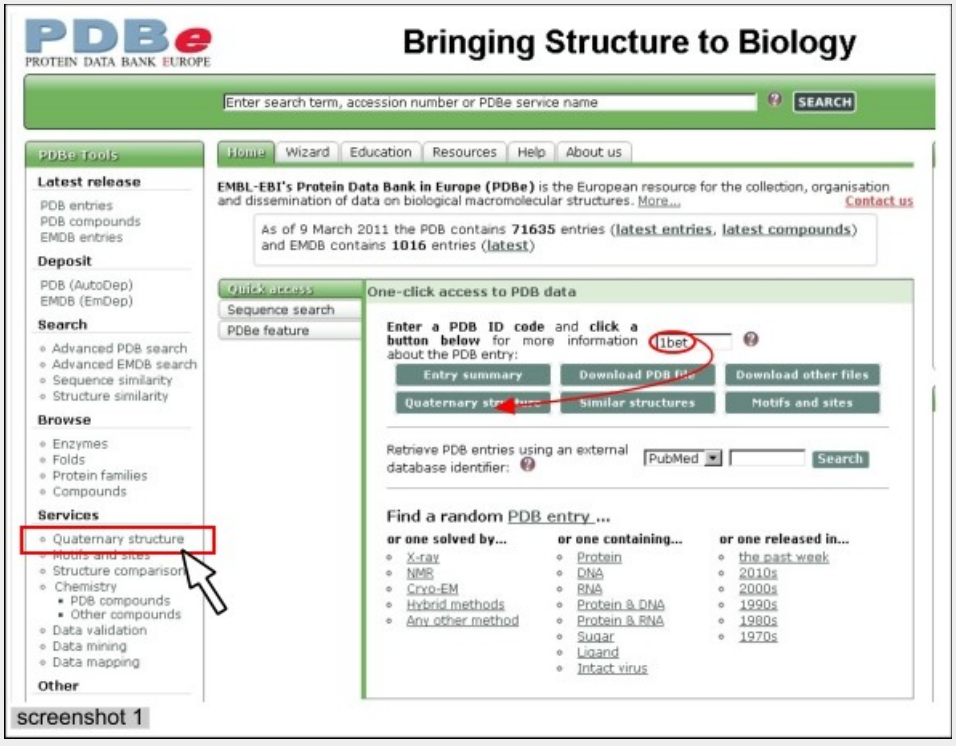
PDBePISA does a calculation on the likely strength of all contacts between macromolecules and will whittle down these assemblies to the most likely biologically relevant assembly. If crystal symmetry operations are needed to build up biological assemblies then **PDBePISA** will do that for you.

PDBePISA will let you analyse the interfaces between components of any assembly in a crystal structure. This is not a molecular docking prediction; **PDBePISA** uses only interactions and macromolecular contacts that are already present in the crystal structure. Unfortunately this means it is not applicable to structures determined by non-crystallographic methods.

Accessing **PDBePISA**.

Either enter the entry id in the one-click access field (circled in red) and click on Quaternary structure (red arrow).

Or click on **Services > Quaternary structure** (mouse arrow) for the **Interfaces and Assemblies** page.



Starting the PDBePISA service

The NGF structure **1bet** is a candidate for **PDBePISA** analysis since (as described in the accompanying **Quips** article) this crystal structure has only a single chain deposited in the PDB entry but is active as a dimer.

This minitutorial shows how to identify and access this assembly in the **PDBePISA** analysis of this entry.

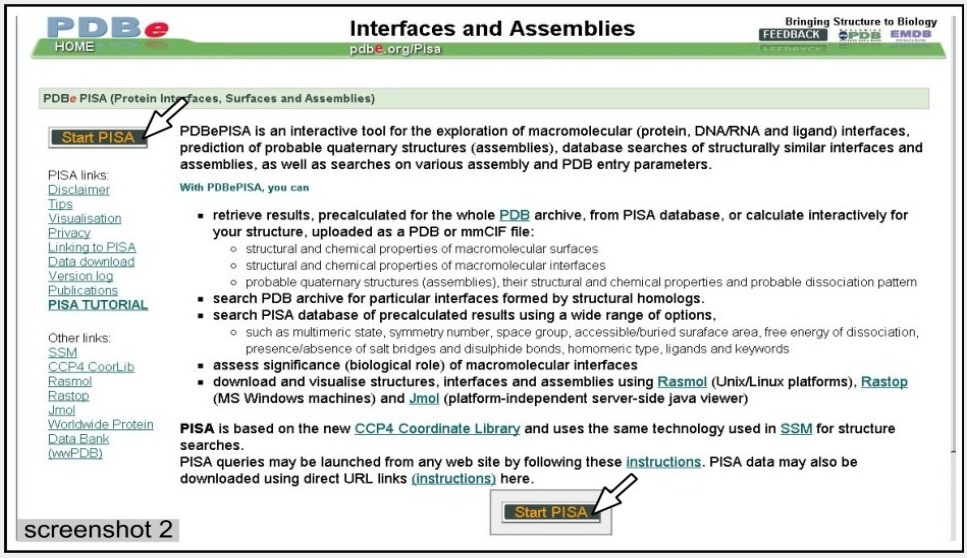
From the **PDBe** main page you can enter an entry identifier such as **1bet** directly in the the quick access box (circled in red in screenshot 1) and then hit the button for **Quaternary structure** button (arrowed in red). This takes you straight to a ready completed **PDBePISA** analysis form for the entry shown as screenshot 3.

Alternatively you can get a blank **PDBePISA** submission form by using the **Quaternary structure** link in the **Services** tab on the **PDBe** main page (red boxed link in screenshot 1). Click on either 'Start PISA' button (arrowed in screenshot 2) and fill in the PDB id code for the structure you are interested in - **1bet** in this case.

Interfaces and Assemblies page.

This is accessed from the **Services** tab on the main page.

Click on either of the 'Start PISA' buttons (mouse arrows) to get a submission form (see screenshot 3).



screenshot 2

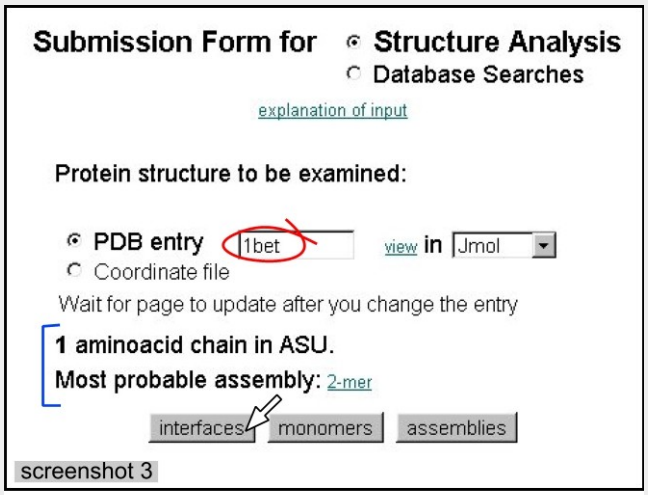
As this is an existing entry you will get immediate access to pre-calculated results - summarised as '1 aminoacid chain in ASU' and 'most probable assembly: 2mer' (highlighted with a blue bracket in screenshot 3).

(If you want a **PDBePISA** analysis for your own crystal structure to you can choose the option **Coordinate file** on this submission page and **PDBePISA** will process this file.)

PDBePISA submission form.

1bet entered in the entry field (circled in red) produces an immediate summary of the **PDBePISA** analysis (blue bracket).

Clicking on the **interfaces** button (mouse arrow) gives the Interfaces results (screenshot 4).



Interfaces formed by symmetry in PDB entry 1bet

To access the **1bet** results first click on the arrowed 'interfaces' button.

PDBePISA has listed five crystal interfaces for **1bet** in order of calculated stability (screenshot 4). The table of results gives details of each contact surface from a particular **Structure 1** (red highlighted) on one side of each interface to a **Structure 2** (green highlighted) on the other side.

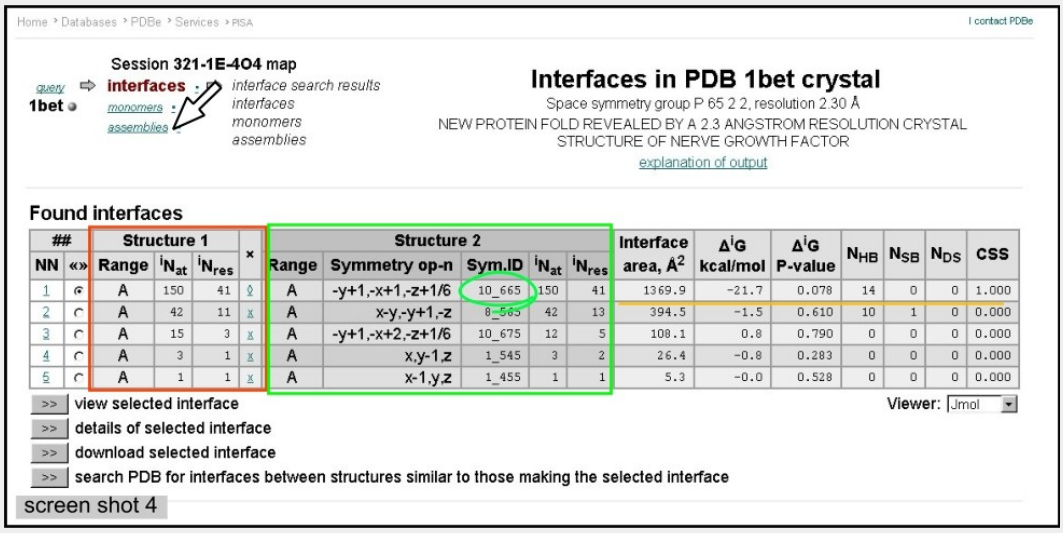
All **PDBePISA** pages give extensive help for the results so you can click on table headings to get a summary explanation.

Each **Structure 2** has been produced by applying a specific crystal symmetry operation to **Structure 1**. The symmetry is indicated in the **Structure 2** section - for example the top (most stable) assembly is generated by mapping coordinates (x,y,z) to (-y+1,-x+1,-z+1/6) and in addition a translation to position the **Structure 2** in an adjacent unit cell.

PDBePISA listing of **1bet** interfaces.

Structure 1 section is boxed in red. Structure 2 section is boxed in green. The symmetry code producing Structure 2 is circled in green. Details of the most stable interface are underlined in orange.

Click on 'assemblies' (mouse arrow) to see the dimer produced by this interface.



screen shot 4

You will now appreciate that it is good to have **PDBePISA** taking care of these details - who could have anticipated both rotation and a unit cell shift would produce the NGF biological assembly!

This assembly is present multiple times in the crystal unit cell - every single copy of NGF exists with a dimeric partner - **PDBePISA** just presents this for the one that the authors have chosen to deposit. The unit cell translation is hidden the **Sym.ID** four-number code **10_665** (circled in green). Here the first number relates to the symmetry operations possible for the crystal while the second set of numbers relate to translations.

1_555, a special code for doing nothing to a structure

One of the commonest symmetry operations is not to do anything and take the macromolecule 'as is': no rotation and no translation - this operation has the special code 1_555 so watch out for that when you check **PDBePISA** output. Code 1_555 next to a **Structure** means that this component is already present in deposited entry. This is not the case for any significant contacts in **1bet** - as the entry only has a monomer there can be no contacts with unrotated/untranslated molecules.

The interface analysis (underlined in orange, screenshot 4) for 10_665 **Structure 2** interface with **Structure 1** shows that it has a very sizable area (1369.9 Å ^{2} - anything approaching 1000 Å ^{2} represents a good interface between macromolecules) and that it

has a favourable calculated energy of -21.7 kcal/mol. Finally **PDBePISA** prints out the contribution of this interface to any predicted biological assemblies - this is CSS calculated as a fractional contribution to the assemblies in the crystal. A CSS value of 1.0 means that this interface represents 100% of the interactions for the predicted assembly. This is the NGF dimer we are interested in. Click on the 'assemblies' link (arrowed in screenshot 5) to examine this and optionally download it.

Clicking on the link 'view selected interface' will give a **Jmol** interactive graphics window showing the NGF dimer with the contact surface coloured up in red and green atoms.

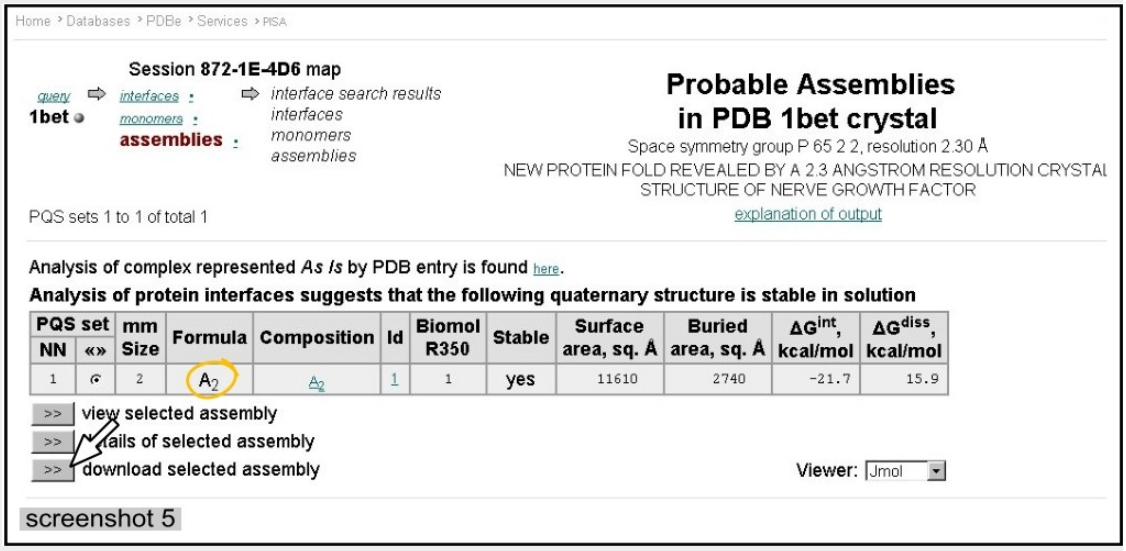
You might like to take the time looking through the other entries in the **PDBePISA** table of **Found interfaces** (screenshot 4). The calculated interface areas and energies should convince you that these are much weaker interactions. These contacts hold the crystal together but are not likely to persist in solution.

PDBePISA probable assemblies.

A dimer is the only likely assembly.

Click on download selected assembly (mouse arrow).

Click on details of selected assembly for next page.



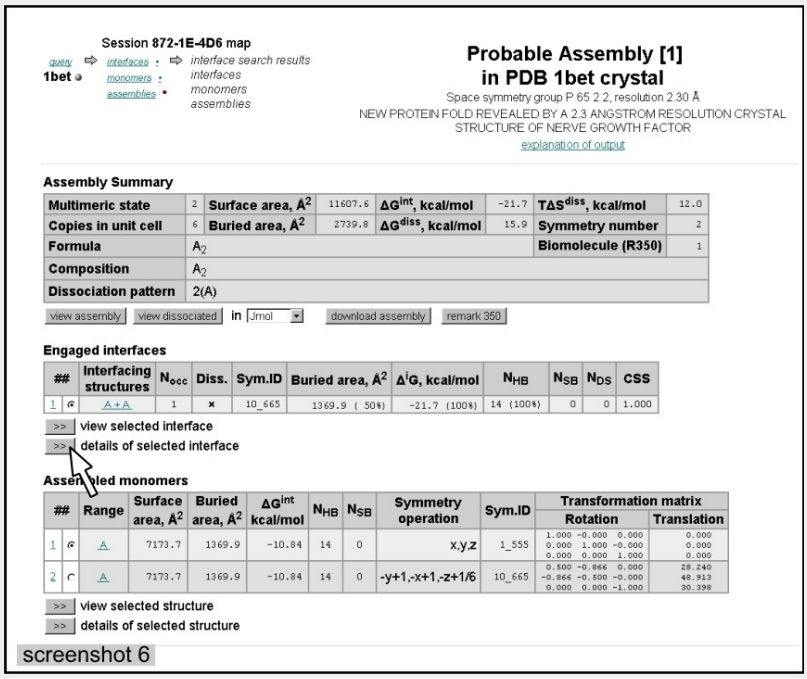
screenshot 5

Downloading and analysing the NGF dimer

Clicking on the arrowed link at the top of the screen (screenshot 4) to the **PDBePISA** 'assemblies' confirms that the **1bet** multimeric state is a dimer A_2 (circled in orange, screenshot 5). You can now download **PDBePISA**'s version of the dimeric assembly using the the arrowed button. Helpfully **PDBePISA** will have relabelled the second NGF molecule in the dimer as chain B so you can use the downloaded file for making pictures with the halves of the dimer coloured distinctly in Rasmol, Pymol, Jmol, or CCP4mg for example.
To get more details of the dimer assembly click just above the download button on the on the appropriately named 'details of selected assembly' button.

PDBePISA summary of the dimeric assembly.

Click on details of selected interface for a residue-by-residue and bond-by-bond listing.



Session 872-1E-4D6 map

query 1bet → interfaces • monomers • assemblies → interface search results interfaces monomers assemblies

Probable Assembly [1] in PDB 1bet crystal

Space symmetry group P 65 2 2, resolution 2.30 Å

NEW PROTEIN FOLD REVEALED BY A 2.3 ANGSTROM RESOLUTION CRYSTAL STRUCTURE OF NERVE GROWTH FACTOR

explanation of output

Assembly Summary

Multimeric state	2	Surface area, Å^2	11607.6	ΔG^{int}, kcal/mol	-21.7	$T\Delta S^{diss}$, kcal/mol	12.0
Copies in unit cell	6	Buried area, Å^2	2739.8	ΔG^{diss}, kcal/mol	15.9	Symmetry number	2
Formula	A_2					Biomolecule (R350)	1
Composition	A_2						
Dissociation pattern	2(A)						

view assembly | view dissociated | in Jmol | download assembly | remark 350

Engaged interfaces

##		Interfacing structures	N_{occ}	Diss.	Sym.ID	Buried area, Å^2	Δ^iG, kcal/mol	N_{HB}	N_{SB}	N_{DS}	CSS
1	σ	A+A	1	×	10_665	1369.9 (50%)	-21.7 (100%)	14 (100%)	0	0	1.000

>> view selected interface

>> details of selected interface

Assembled monomers

##		Range	Surface area, Å^2	Buried area, Å^2	ΔG^{int} kcal/mol	N_{HB}	N_{SB}	Symmetry operation	Sym.ID	Transformation matrix	
										Rotation	Translation
1	σ	A	7173.7	1369.9	-10.84	14	0	x,y,z	1_555	1.000 -0.000 0.000 0.000 1.000 -0.000 0.000 0.000 1.000	0.000 0.000 0.000
2	σ	A	7173.7	1369.9	-10.84	14	0	-y+1,-x+1,-z+1/6	10_665	0.500 -0.866 0.000 -0.866 -0.500 -0.000 0.000 0.000 -1.000	28.240 48.913 30.398

>> view selected structure

>> details of selected structure

screenshot 6

The first assembly summary page (screenshot 6) repeats details of the interfaces and components deriving from the **PDBePISA** analysis. A detailed analysis of the interface in the dimeric assembly is available by clicking on '>> details of selected interface' (arrowed) button. Unlike the Interfaces analysis shown earlier, which included all crystal interfaces (screenshot 4), the detailed analysis here (screenshot 7) relates only to contacts and bonds in the predicted biological assembly. Scrolling down from the the summary you can examine the residue-by-residue analysis shown in the screenshot 7 snippet. All residues involved in non-covalent and non-covalent bonds across the interface are listed. This analysis of the dimeric interface highlights that, unlike the NGF monomer, there are no disulfide bonds involved in stabilising the dimeric form (orange circled output).

What gives the NGF dimer its very high stability in solution? One important contribution is provided by the hydrogen bonds listed here. But, in addition, by checking through the residue-by-residue output you can start to see that hydrophobic sidechains such as those of tryptophan (Trp, one circled in blue) and phenylalanine (Phe) make a large contribution to the interface of the NGF dimer.

In the NGF dimer the same component is on each side of the interface, and so each **Structure 1** residue in the interface also contributes from the **Structure 2** side. The table below of residue-by-residue contributions reveals hydrophobic residues contributing (one example, Trp 21 is circled in dark blue). All residues making contacts in the interface are shown with a light blue background while those making hydrogen, or optionally disulfide or salt bridges, are coloured red. Each residue's contribution to the buried surface area, from which solvent is excluded, is calculated and shown graphically with bars.

Close-fitting hydrophobic sidechains such as Trp and Phe contribute to macromolecular interfaces by excluding solvent. It is the entropic gain of the excluded solvent moving into free solution that produces a 'hydrophobic interaction' stabilising the interface.

For a finer grain view of the interface the **level of detail** can be set to 'individual atoms' using the pulldown above this listing (which defaults to residue level).

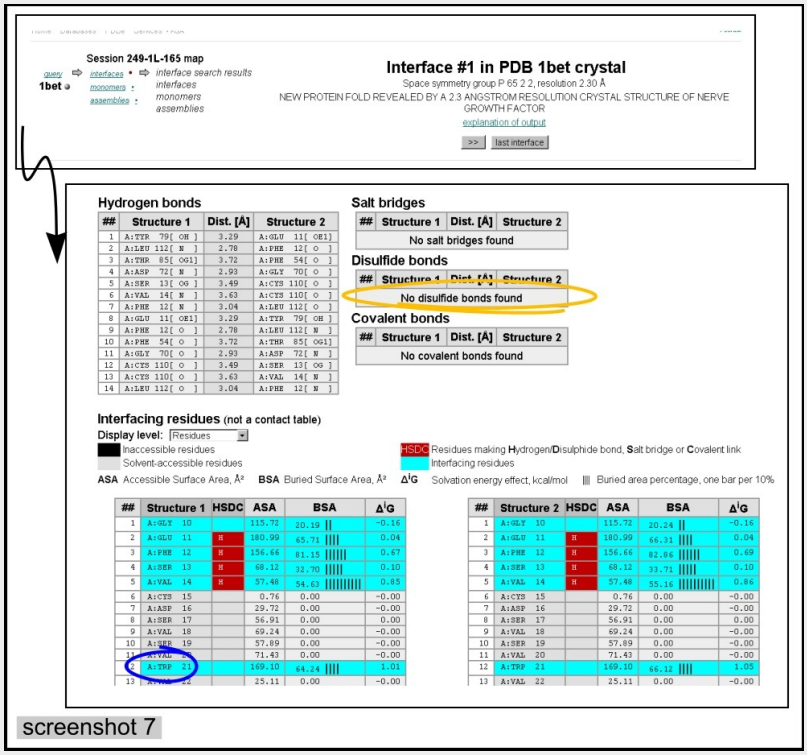
As you browse assemblies in the PDB archive you will frequently uncover heterodimeric

interfaces between different molecules. In more complicated output tables you need to check which assembly or interface you have selected for more detailed analysis - **PDBePISA** tables include a 'radio button' column (often labelled <<>>) - you need to click on the selection button in this second column of any table before proceeding to view details of, or download, your chosen assembly.

PDBePISA will help you drill down to view the fine detail in any crystal structure that you need to examine for interfaces and significant assemblies.

Extracts from the details of **1bet** probable assembly interface.

Scroll down (arrow) through summary text. List of interactions shows H-bonds across the interface and reveals no disulfide bonds (circled in orange).



screenshot 7