phenix.ensemble_refinement: a test study of apo and holo BACE1

B. Tom Burnley and Piet Gros
Crystal and Structural Chemistry, Bijvoet Center for Biomolecular Research, Utrecht University
Correspondence email: B.T.Burnley@uu.nl & P.Gros@uu.nl

Introduction

`phenix.ensemble_refinement` (Burnley et al. 2012) combines molecular dynamic (MD) simulations with X-ray structure refinement to generate ensemble models fitted to diffraction data. It is an evolution of the 'time-averaging' method first proposed by Gros et al. in 1990 (Gros, van Gunsteren, and Hol 1990) and now utilizes a maximum-likelihood target function, a dual explicit-bulk solvent model and introduces a TLS fitting procedure that absorbs rigid-body motions such that short MD simulations can be used to sample local disorder. The resulting ensemble model represents, as a Boltzmann-weighted population of structures, the simulation trajectory and provides implicit modeling of anisotropic and anharmonic motions. This family of structures provides two main advantages: 1) a reduction in R_{free} compared with traditional single structure models and 2) quantification and visualization of protein dynamics that have been demonstrated to correlate with biological function. Methodological implementation and testing is described in detail in Burnley et al. 2012; here we focus on a practical usage of the method.

For this test case we examine β-Secretase (β -site amyloid precursor protein-cleaving enzyme1; BACE1), in the apo and holo state. BACE1 is a transmembrane aspartic protease that cleaves β-amyloid precursor protein and is the putative rate-limiting enzyme in the amyloid β-peptide (Aβ) pathway (Sinha 1999). Aβ is likely responsible for the Alzheimer's disease cascade and therefore BACE1 is a primary target for the development of inhibitors to treat/prevent Alzheimer's disease (Xu et al. 2011). The importance of this protein has resulted in multiple entries in the PDB (Berman et al. 2000). We selected two 1.6 Å resolution datasets deposited by Xu and co-workers (Xu et al. 2011) of the protease ectodomain in the apo state and holo state complexed with a nonpeptide inhibitor (*N0*-[(2*S*,3*R*)-4-(cyclopropylamino)-3-hydroxy-1-phenylbutan-2-yl]-5-(methylsulfonylamino)-*N*-[(1*R*)-1-phenylethyl]benzene-1,3-dicarboxamide; BSIIV) for analysis by ensemble refinement (PDB codes 3TPJ and 3TPP respectively). Here we detail the phenix.ensemble_refinement process, from preparing input data through parameter optimization to structure analysis.

Preparing input data

`phenix.ensemble_refinement` uses a set of input files similar to `phenix.refine (Adams et al. 2010)`, i.e. structure (e.g. .pdb format) and reflection files (e.g. .mtz) are required and specified parameter definitions (e.g. .eff) and ligand restraints (e.g. .cif) should be supplied as needed. Ensemble refinement is dependent on the quality of the input structure therefore the input structure should be post-traditional refinement and thus ready for deposition, or as deposited, in the PDB. It should be noted that if TLS was used in the refinement protocol the TLS contributions must be present in the atom records (see `phenix.tls`).

For this test case the PDB structures and structure factors were downloaded from the PDB_REDO server (Joosten et al. 2010). One of the advantages of the PDB_REDO server is that it automatically builds and refines missing side-chains, and the usage of complete side-chains is recommended for ensemble refinement. However, we do not recommend building long sequences (>~4 residues) of highly disordered regions such as loops or termini if there is not sufficient electron density to support placement when using standard model building procedures. In the BACE1 structures loop 153-173 and loop 308-319 have missing residues, as do both the N and C termini and after examining in these regions. It was not possible to build any additional residues. `phenix.ready_set` was then used to generate ligand restraints and to add explicit hydrogen atoms to the PDB file. The structures were further refined using `phenix.refine` and TLS groups were determined using `phenix.find_tls`. Table 1 shows the refinement statistics. Prior to use with `phenix.ensemble_refinement` any alternative conformations should be removed and

Table 1. BACE1 apo and holo dataset and refinement statistics as deposited in the PDB and following re-refinement using `phenix.refine` and `phenix.ensemble_refinement`.

	Apo	Holo
PDB code	3tpj	3tpp
Space group	$C222_1$	$C222_1$
Unit-cell parameters		
a (Å)	104.4	104.5
b (Å)	128.7	128.2
c (Å)	76.7	76.5
Resolution range (Å)	35.8 - 1.6	49.1 - 1.6
PDB structure		
R_{work} (%)	17.8	15.5
R_{free} (%)	19.6	18.0
phenix.refine		
R_{work} (%)	14.4	14.0
R_{free} (%)	15.8	16.1
Geometric rmsd		
Bond lengths (Å)	0.0	0.0
Bond angles (°)	1.3	1.4
Dihedral angles (°)	12.5	13.1
Ramachandran		
Disallowed residues (%)	0.0	0.0
Allowed residues (%)	2.3	1.3
Favored residues (%)	97.7	98.7
phenix.ensemble_refinement		
R_{work} (%)	12.1	12.1
R_{free} (%)	14.7	14.6
Geometric rmsd (centroid)		
Bond lengths (Å)	0.008	0.007
Bond angles (°)	1.113	1.151
Dihedral angles (°)	8.38	8.51
Geometric rmsd (per structure)		
Bond lengths (Å)	0.014	0.013
Bond angles (°)	1.713	1.728
Dihedral angles (°)	17.12	17.22
Ramachandran (centroid)		
Disallowed residues (%)	1.6	0.5
Allowed residues (%)	1.3	1.9
Favored residues (%)	97.0	97.6
Ramachandran (per structure)		
Disallowed residues (%)	2.5	2.6
Allowed residues (%)	4.4	5.1
Favored residues (%)	93.1	92.3

occupancies re-adjusted. Alternative conformations will be sampled automatically during the simulation.

Both 3TPJ and 3TPP datasets were collected at 100 K and exhibit the $C222_1$ space group, also both have similar unit cell parameters and resolution ranges (Table 1). When possible it is advantageous to use analogous datasets such as these as it allows for more direct structural comparison. Care should be taken if, for example, datasets have different space groups or cell parameters as differences in crystal packing may influence atomic fluctuations and must be considered when comparing the resultant ensembles.

Running ensemble refinement

`phenix.ensemble_refinement` can be run from the command line or from the *PHENIX* GUI. There are several parameters that should be defined and/or optimized by the user. Parameter optimization is covered in the next section; here we examine the parameters defined for the BACE1 simulations that remain consist across all runs.

TLS groups used in the TLS fitting procedure should be defined by the user and do not correlate with those used in standard refinement (i.e. it is not recommended to use `phenix.find_tls` for ensemble refinement). Each defined TLS group will fit the rigid-body motion of that group such that they are not required to be sampled during the simulation. This reduces the conformational space that needs to be sampled in the simulation and simplifies the output ensemble. In practice it is best to use complete chains or domains for the defined groups (e.g. tls_group_selections = 'chain a') unless there is good reason to do otherwise. In the case of BACE1 it has a defined 2-lobe architecture so each lobe is assigned to a separate TLS group. The N terminal lobe occurs between residues -4 to 180 and the C-terminal lobe from residue 181 to 386. Both the apo and holo structures contain a number of non-solvent ligands and these must be assigned to the correct TLS group. Visual inspection in a molecular graphics program, e.g. PyMOL (Schrödinger 2010) is used to determine which lobe each ligand is closest to. For the apo dataset the following TLS selections were used to define two groups: tls_group_selections = 'resseq -4:180 or resseq 394 or resseq 395 or resseq 398:400 or resseq 402:406', tls_group_selections = 'resseq 181:386 or resseq 396:397 or resseq 401 or resseq 999'. The selection strings can be defined using the standard nomenclature. It is not necessary to include water atoms as these are automatically assigned to the nearest TLS groups during the simulation.

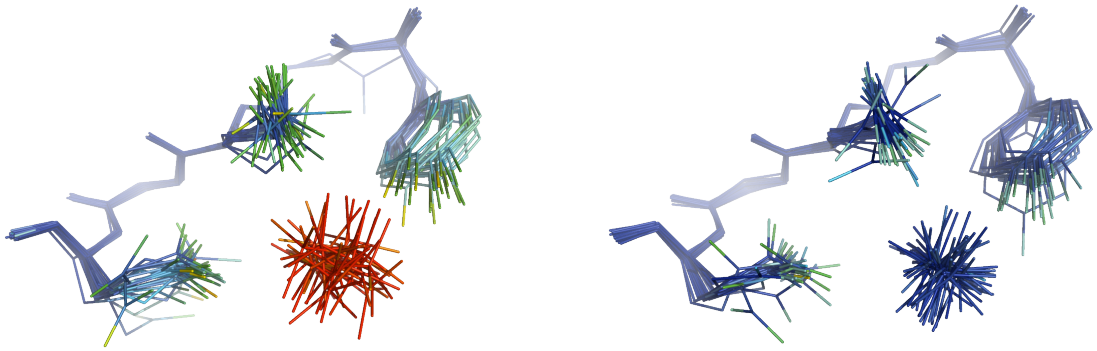

Figure 1. Effect of atom occupancy on kinetic energy. Two simulations of holo BACE1 were performed with difference occupancies for the SO_4^{2-} hetgroup, 1.0 on the left and 0.4 on the right. Atoms are colored by kinetic energy ranging from 0.0 to 2.0 kcal/mol (blue to red).

For weakly bound non-specific ligands it may be appropriate to add harmonic restraints to prevent stochastic displacement during the simulation. For apo BACE1 harmonic restraints we defined as: harmonic_restraints.selections = 'resname so4 or resname cl or resname ure'. This restrains all selected atoms to the x,y,z co-ordinate positions given in the input structure. To allow for any uncertainty in position the harmonic restraints have a 'flat bottom' so deviations less than a given slack value (distance in Å) will generate no force (e.g. harmonic_restraints.slack = 2.0). For the holo structure it is not necessary to add harmonic restraints for the high-affinity, well-ordered, inhibitor.

At present non-solvent ligands with non-unity occupancies cannot be replaced/reinserted in the simulation *a la* the explicit solvent atoms (see Burnley et al. 2012 for details). Atom occupancy is defined from the input structure and if applicable this should be refined, when possible, using single-structure methods prior to ensemble refinement. Errors in occupancy can be detected during ensemble refinement: an atom with too high occupancy will sample a greater amplitude and frequency of disorder than neighboring atoms and exhibit excessive kinetic energy. Atomic kinetic energies (kcal/mol x 10^{-1}) can be outputted in the final PDB ensemble in the occupancy column by setting the parameter: `output_running_kinetic_energy_in_occupancy_column=True`. Figure 1 shows a sulphate group in holo BACE1, where two simulations were performed with occupancies of 1.0 and 0.4 (0.4 was calculated using combined occupancy and ADP refinement in `phenix.refine`). An occupancy of 1.0 results in the sulphate group sampling excessively high kinetic energies whereas at the refined occupancy of 0.4 these atoms exhibit kinetic energies similar to that of the neighboring environment. For datasets where stable refinement of occupancies is not possible, e.g. at low resolution, we recommend running multiple simulations with different occupancy values (e.g. 0.3,0.4,0.5…) .

Finally the water picking parameters can be adjusted depending on the dataset resolution. As standard the water positions are assessed using cutoffs of 3.0 and 1.0 σ for mFo-DFm and 2mFo-DFm maps respectively. For structures with resolution better than ~1.7 Å, better results maybe obtained by lowering the mFo-DFm threshold (e.g. `ensemble_ordered_solvent.primary_map_cutoff = 2.5`). For both the apo and holo BACE1 structures better quality mFo-DFm difference maps and lower R_{free} values where found when using a cutoff of 2.5 σ was used (values of 2.5 and 3.0 σ were tested).

The simulation can be started from the GUI or from the command line using `phenix.ensemble_refinment`. The simulation length depends on several factors including: the number of atoms in the ASU, dataset resolution and hardware configuration. In our experience runtime is typically between several hours and a few days. Using a 1.9 GHz processor for BACE1 the CPU time was 3.6 days. While the simulations are running no user input is required and, due to the length of CPU time, it is highly recommended to test alternative parameters in parallel.

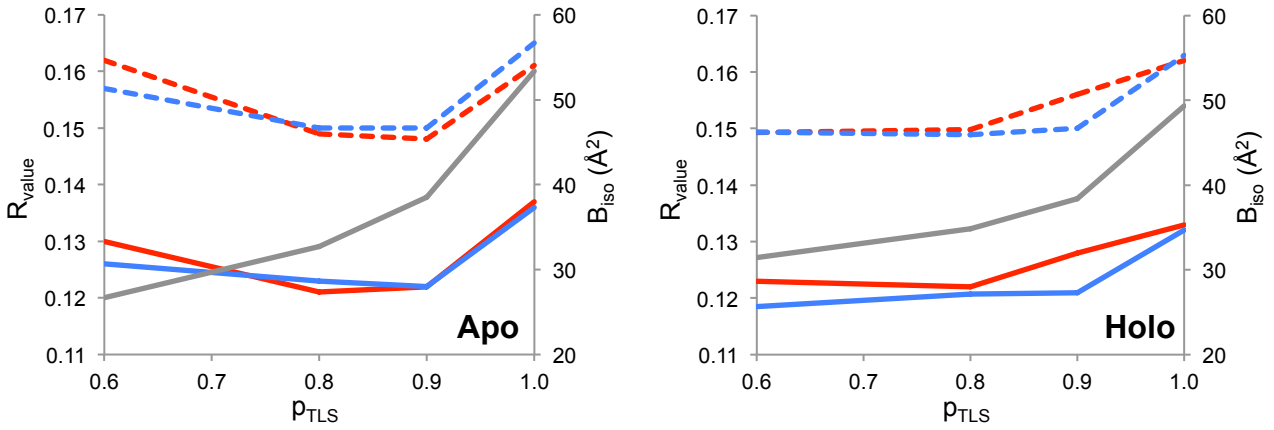

Figure 2. Ensemble refinement parameter optimization for apo and holo BACE1 datasets. An array of p_{TLS} values (0.6 – 1.0) and T_{bath} (2 and 10 K shown in red and blue respectively) was performed and the resulting R_{work} and R_{free} values are shown in solid and dashed lines respectively. The isotropic equivalent of the fitted TLS B-factors are shown in grey, averaged over all non-solvent, non-hydrogen atoms.

Optimizing simulation parameters

There are two main empirical parameters (p_{TLS}, T_{bath}) that affect the running and the outcome of ensemble refinement, and it is critical to optimize both on a dataset by dataset basis.

p_{TLS} (e.g. `ptls=0.9`) defines the fraction of atoms included in the TLS fitting procedure (for details see Burnley et al. 2012). The optimum value for this parameter cannot be determined *a priori*. Setting up a parallel array of simulations with different values (e.g. 1.0, 0.9, 0.8, 0.6) is recommended. The TLS fitting process requires that the number of non-hydrogen, non-solvent atoms per TLS group multiplied by the p_{TLS} fraction to be greater than 63 (enforcing a data:parameter ratio greater than 3 for TLS fitting). If this is not the case, p_{TLS} will be automatically increased. In the case that a TLS group has less than 63 non-solvent, non-hydrogen atoms then an isotropic model, based on the Wilson B-factor, is automatically applied.

T_{bath} (e.g. `wxray_coupled_tbath_offset=5`) set the temperature bath offset (K) and, counter intuitively, T_{bath} controls the X-ray weight. The X-ray weight is modulated in situ such that the simulation runs at the target temperature (e.g. `cartesian_dynamics.temperature=300`). The simulation uses velocity scaling to maintain the target temperature. This is coupled to a temperature thermostat which is set as: thermostat = cartesian_dynamics.temperature - wxray_coupled_tbath_offset. The non-conservative time-averaged X-ray force generates heat (Gros, van Gunsteren, and Hol 1990), thus larger offsets increase the X-ray weight. The default value is 5 K and, for example, values of 2 and 10 K may also be tested.

For both the apo and holo BACE1 datasets a parallel array of eight separate simulations was performed with a grid of p_{TLS} (1.0, 0.9, 0.8 and 0.6) and T_{bath} (2 and 10 K), as shown in Figure 2. A shallow optimum in R_{values} is observed and with optimum values of 0.9 and 2 K for p_{TLS} and T_{bath} respectively for the apo dataset and 0.6 and 2 K for the holo dataset.

Ensemble refinement is based on MD simulations that are stochastic by nature. Therefore to aid analysis it is useful to run a number of random seed repeats (e.g. `random_seed=26556257`). This seed is used in the initial velocity assignment and the alternative trajectories that arise can be used as a guide to the reproducibility of the simulation as a whole and also the individual atom fluctuations (see Burnley et al., 2012, Figures 4, 5 & 10). In total five random number seed repeats where run for both BACE1 datasets producing consistent global R_{values}, see Table 2, the best of which (as judged by R_{work} *and* R_{free} *values*) were then selected for the subsequent structural analysis and the refinement statistics for these are reported in Table 1. In this case we found that the R_{values} improve for both the apo and holo datasets after ensemble refinement.

Table 2. R_{values} and number of structures in output PDB ensemble model for five random number seed repeats of `phenix.ensemble_refinement` for apo and holo BACE1.

	Apo			Holo		
Repeat	R_{work} (%)	R_{free} (%)	*Structures*	R_{work} (%)	R_{free} (%)	*Structures*
1	12.1	14.7	134	12.1	14.6	107
2	12.1	14.7	178	11.9	14.7	134
3	12.2	14.8	107	12.1	14.8	134
4	12.1	14.8	134	12.3	14.9	89
5	12.2	14.8	134	12.4	15.0	134
Mean	*12.1*	*14.8*	*137*	*12.2*	*14.8*	*120*

Two sets of geometry statistics are reported in Table 1 for the ensemble structures: 'centroid' and 'per structure' (see Burnley et al. 2012 for complete definitions). This is because the ensemble as a whole is fitted to the data. Therefore the structures represent a distribution that reflects fluctuations around ideal geometric values, i.e. each individual restraint in each individual structure does not necessarily have to fit the ideal geometric value and a degree of variation is expected. However, considering the model as a whole ensemble, the restraints deviations can be expected to oscillate around an ideal value therefore 'centroid' statistics are provided. The geometric deviations are also calculated for each structure separately and then averaged across the ensemble giving the 'per structure' values. While large deviations can be observed when evaluated per structure, the centroids of the distribution show fluctuations around the correct value for the BACE1 datasets, see Table 1.

Structural analysis

By default phenix.ensemble_refinement outputs the following files: log (.log), map coefficients (.mtz), geometry of input structure (.geo) and the ensemble structure (.pdb.gz). As the ensemble PDB structure can contain hundreds of models it is zipped to reduce size and should be unzipped (using gunzip for instance) before use. During the simulation thousands of sets of coordinates are sampled, so to simplify the final model the minimum number of models is found that reproduces the R_{free} value of the whole trajectory within a 0.25 % tolerance. To generate these reduced ensembles every 1st,2nd,3rd...nth model of the total number of coordinates sets stored during the acquisition phase of the simulation are selected and the R_{values} of these concerted ensembles are recalculated. The relationship between the number of structures and the R_{values} is shown for apo and holo BACE1 in Figure 3. There is some fluctuation in the recalculations and as such the number of models in the final PDB can vary as is shown in Table 2 for the five number seed repeats.

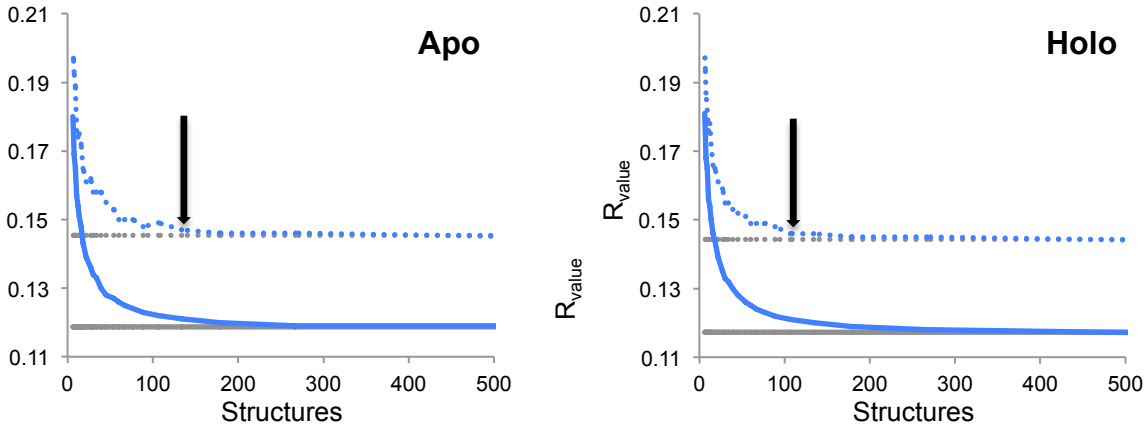

Figure 3. Determination of the number of models in the output ensemble PDB. For the best apo and holo simulation the charts shows the recalculation of R_{work} and R_{free} (solid and dashed line respectively) with different numbers of structures in the ensemble models (blue), the minimum number of models to reproduce complete trajectory R_{values} (shown in grey) within a 0.25% tolerance is selected for output (highlighted with arrow). In this case, apo and holo BACE1 contained 134 and 107 models.

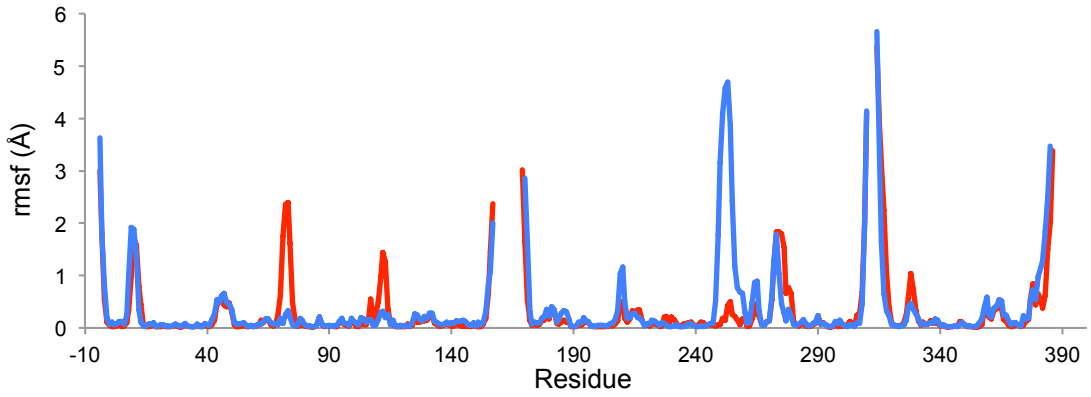

Figure 4. Backbone rmsf of apo (red) and holo (blue) BACE1 ensembles. Human aspartic protease residue numbering is used. Values were calculated used ens_tools.py script for PyMOL.

Visual analysis of the structures and maps can be done using standard programs like PyMOL and COOT (Emsley et al. 2010). When using PyMOL the models can be viewed sequentially as a movie using the play button or superimposed together via the show all states option (via Movie header or the command 'set all_states, 1'). If the user wants to view atomistic atom B-factors or kinetic energies (see above) for each individual model in the ensemble the additional option 'discrete=1' should be specified with the load command. Viewing in COOT may be improved by reducing the number of models in the ensemble to between 5 to 50 and this can be done using command line program `phenix.ensemble_reducer`, which allows the further truncation of the ensemble. Once the ensemble models have been produced no manual model building is possible. If, on inspection, an error has occurred, for instance a ligand has become displaced from its binding site than the simulation parameters need to altered, e.g. harmonic restraints maybe added or adjusted, and ensemble refinement should be re-run.

The structural fluctuation in the ensembles can be quantified by calculating the rmsf and this is shown per residue for apo and holo BACE1 in Figure 4. Many programs can process ensemble structures to calculate rmsf and other metrics, and we have provided a script for PyMOL that can also do this. 'ens_tool.py' is available from $PHENIX/cctbx_project/mmtbx/refinement/ensemble_refinement/ and the command 'ens_rmsf, selection' will return the rmsf values for selected atoms. Figure 4 shows the highest deviations for both apo and holo BACE1 datasets occur in the N and C terminal regions as well loop 153-173 and loop 308-319. These regions contain missing residues so it is unsurprising to observed high proximal mobility located here. A high rmsf value is returned for tyr71 in the apo structure. This residue is located in the 'flap' region, a hairpin loop found in the N-terminal lobe and spanning residues 67-75. Other structures deposited in the PDB show that this loop can adopt multiple conformations. Figure 5 shows the flap region in the apo and holo forms of BACE1. When the active site is empty the flap region is mobile, however, upon addition of the inhibitor BSIIV the flap folds over the binding pocket such that tyr71 can make weak hydrophobic contacts with the inhibitor and thus trapping the flap in a single conformation. Figure 6 shows the remainder of the binding pocket and here a general tightening of the residues surrounding the ligand is observed upon binding. In particular, the conserved catalytic dyad formed by asp32 and asp228 becomes more ordered in presence of the inhibitor, see Figure 7. Interesting, the holo structure is highly mobile between residues 248-256, a surface turn is located in the C-terminal lobe, whereas the apo structure is more rigid. This pattern is repeated for all five random number seed repeats for both apo and holo datasets, and could have functional role for instance in offsetting the loss of entropy upon binding and/or as a method of signaling.

In conclusion, `phenix.ensemble_refinement` is suitable for standard protein datasets and does not require specialist experimental techniques or hardware. For both apo and holo BACE1 it provides multi-structure models that fit well to the diffraction data. The TLS fitting procedure absorbs global motions and this gives the

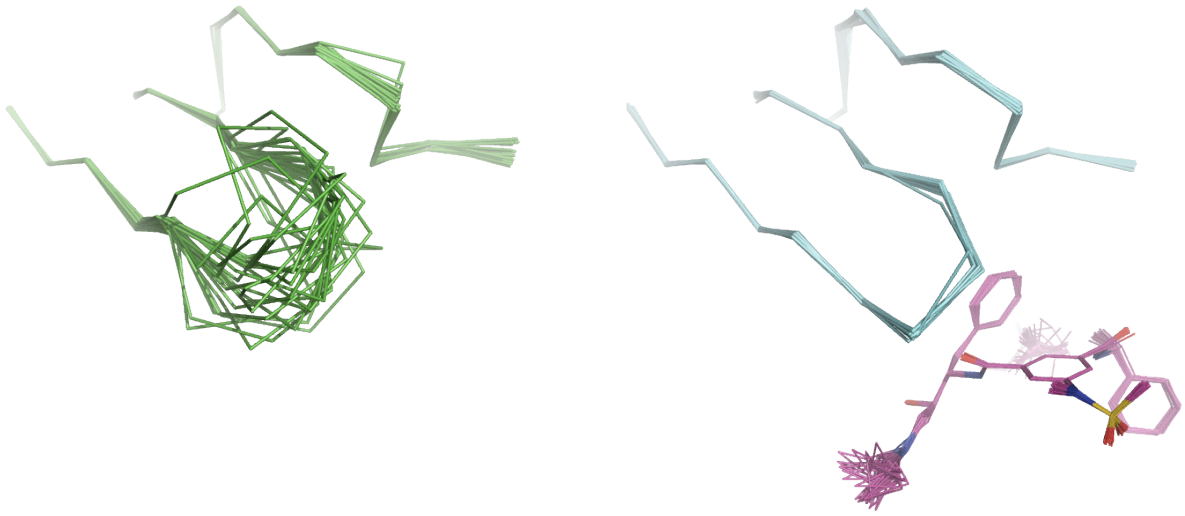

Figure 5. Comparison of the 'flap' region in the ensemble structures of apo (left, green ribbon) and holo (right, blue ribbon) BACE1. The hairpin loop (residues 67-75) becomes significantly more ordered in the presence of the inhibitor. 25 structures are shown for each ensemble.

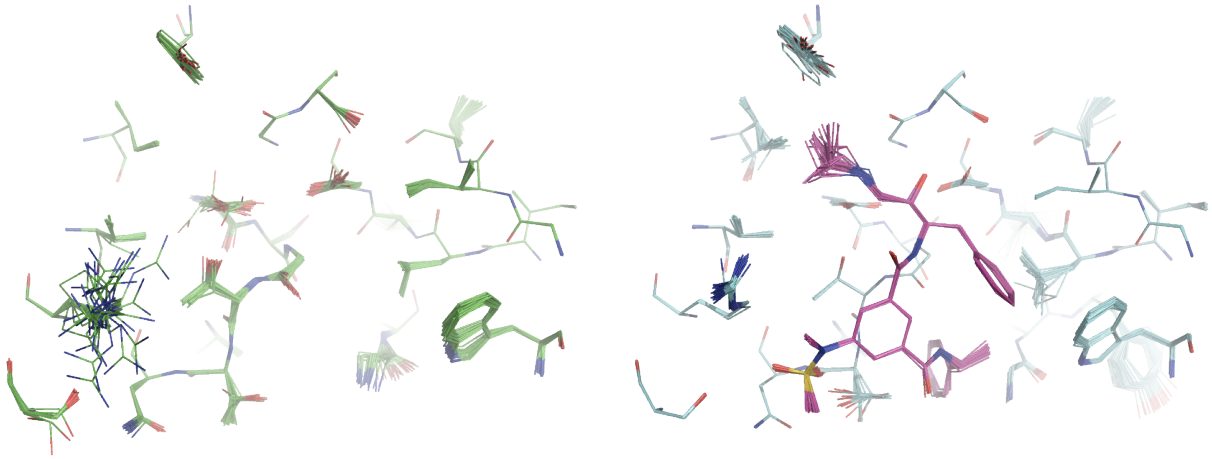

Figure 6. The binding pocket of apo (left, green carbon atoms) and holo (right, blue carbon atoms) BACE1 rigidifies in the presence of the inhibitor (right, pink carbon atoms). 25 structures are shown for each ensemble.

opportunity to visualize and quantify the intramolecular dynamics of the molecule. As such, it provides a tool to probe structure, function and dynamics from diffraction data.

Caveat Emptor

There are some important caveats that should be noted to prevent accidental misinterpretation of the models produced using ensemble refinement. 1) Ensemble refinement is recommended for use with datasets with a resolution of 3 Å or better. 2) The number of structures in the output ensemble is variable and should not be over interpreted. 3) The ensembles are a Boltzmann-weighted population and therefore contain high-energy conformations. 4) Dynamic rates cannot be calculated from the simulations. The X-ray force is non-conservative and accelerates atomic sampling. This allows movements that occur at timescales several orders of magnitude greater than a 100ps simulation, e.g. side chain sampling or loop motions, to be modeled in the (relatively) short simulation time but therefore the rate of conformation exchange cannot be derived. 5) No information regarding correlated motion is provided by the time- and spatially-averaged X-ray restraints although local correlations may be observed via short-range steric interactions.

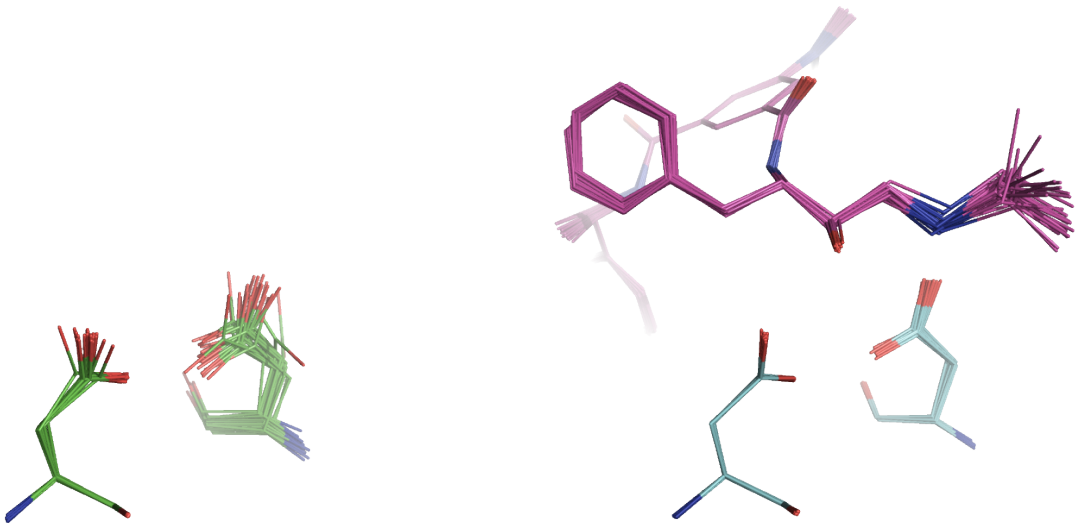

Figure 7. Both asp32 and asp228 residues that form the conserved catalytic dyad exhibit higher mobility in the absence of the inhibitor. Apo BACE1 ensemble is shown left (green carbon atoms) and holo ensemble is shown right (blue and pink for protein and inhibitor carbon atoms respectively). 25 structures are shown for each ensemble.

References

Adams, P D, Pavel V Afonine, Gábor Bunkóczi, Vincent B Chen, Ian W Davis, Nathaniel Echols, Jeffrey J Headd, et al. 2010. "PHENIX: a Comprehensive Python-based System for Macromolecular Structure Solution." *Acta Crystallographica Section D* 66 (2) (February): 213–221. doi:10.1107/S0907444909052925.

Berman, Helen M., John Westbrook, Zukang Feng, Gary Gilliland, T. N. Bhat, Helge Weissig, Ilya N. Shindyalov, and Philip E. Bourne. 2000. "The Protein Data Bank." *Nucleic Acids Research* 28 (1): 235–242. doi:10.1093/nar/28.1.235.

Burnley, B. T., P. V. Afonine, P. D. Adams, and P. Gros. 2012. "Modelling Dynamics in Protein Crystal Structures by Ensemble Refinement." *eLife* 1 (0) (January 1): e00311–e00311. doi:10.7554/eLife.00311.

Emsley, P., B. Lohkamp, W. G. Scott, and K. Cowtan. 2010. "Features and Development of *Coot*." *Acta Crystallographica Section D Biological Crystallography* 66 (4) (March 24): 486–501. doi:10.1107/S0907444910007493.

Gros, P, W F van Gunsteren, and W G Hol. 1990. "Inclusion of Thermal Motion in Crystallographic Structures by Restrained Molecular Dynamics." *Science (New York, N.Y.)* 249 (4973) (September 7): 1149–1152.

Joosten, R. P., T. A. H. te Beek, E. Krieger, M. L. Hekkelman, R. W. W. Hooft, R. Schneider, C. Sander, and G. Vriend. 2010. "A Series of PDB Related Databases for Everyday Needs." *Nucleic Acids Research* 39 (Database) (November): D411–D419. doi:10.1093/nar/gkq1105.

Schrödinger, LLC. 2010. "The PyMOL Molecular Graphics System, Version 1.3r1."

Sinha, S. 1999. "Cellular Mechanisms of Beta -amyloid Production and Secretion." *Proceedings of the National Academy of Sciences* 96 (20) (September 28): 11049–11053. doi:10.1073/pnas.96.20.11049.

Xu, Yechun, Min-jun Li, Harry Greenblatt, Wuyan Chen, Aviv Paz, Orly Dym, Yoav Peleg, et al. 2011. "Flexibility of the Flap in the Active Site of BACE1 as Revealed by Crystal Structures and Molecular Dynamics Simulations." *Acta Crystallographica Section D Biological Crystallography* 68 (1) (December 9): 13–25.