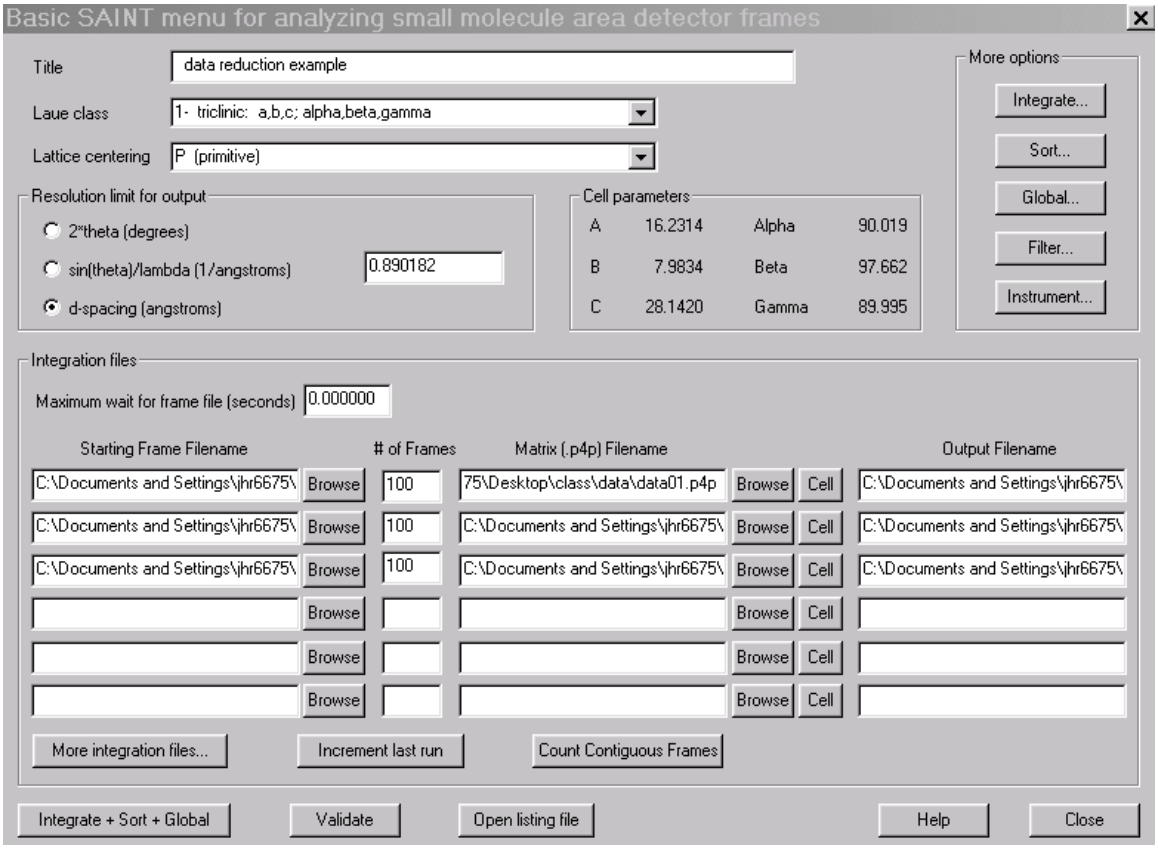
Data Reduction - Beginner

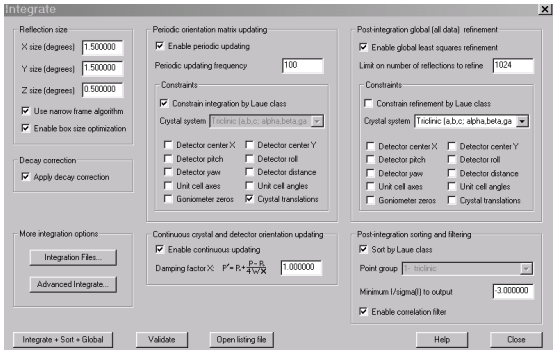
Saint ...

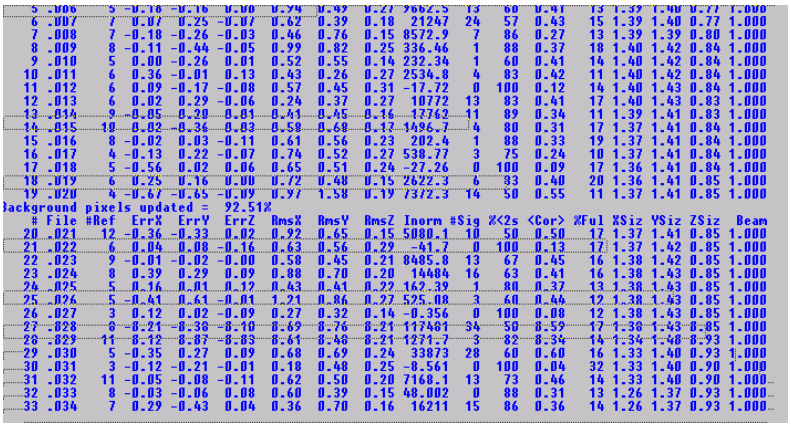
Start SAINT. Point to PROJECT/NEW find your data files, input a project name and open them.

Point to SAINT/INITIALIZE (answer any commands prompts with YES etc).
Point to SAINT/EXECUTE



Set Laue Class to 1- triclinic. Point to INTEGRATE. Check the enable box size optimization and the enable periodic updating. Check that the Laue Class is triclinic. Point to the integrate+sort+global button.





5	.006	5	-0.16	-0.16	0.00	0.94	0.49	0.27	9662.5	13	60	0.41	13	1.39	1.40	0.77	1.000
6	.007	7	0.07	0.25	-0.07	0.62	0.39	0.18	21247	24	57	0.43	15	1.39	1.40	0.77	1.000
7	.008	7	-0.18	-0.26	-0.03	0.46	0.76	0.15	8572.9	7	86	0.27	13	1.39	1.39	0.80	1.000
8	.009	8	-0.11	-0.44	-0.05	0.99	0.82	0.25	336.46	1	88	0.37	18	1.40	1.42	0.84	1.000
9	.010	5	0.00	-0.26	0.01	0.52	0.55	0.14	232.34	1	60	0.41	14	1.40	1.42	0.84	1.000
10	.011	6	0.36	-0.01	0.13	0.43	0.26	0.27	2534.8	4	83	0.42	11	1.40	1.42	0.84	1.000
11	.012	6	0.09	-0.17	-0.08	0.57	0.45	0.31	-17.72	0	100	0.12	14	1.40	1.43	0.84	1.000
12	.013	6	0.02	0.29	-0.06	0.24	0.37	0.27	10772	13	83	0.41	17	1.40	1.43	0.83	1.000
13	.014	9	-0.05	0.20	0.01	0.41	0.45	0.16	17762	11	89	0.34	11	1.39	1.41	0.83	1.000
14	.015	10	0.02	-0.36	0.03	0.58	0.60	0.17	1496.7	4	80	0.31	17	1.37	1.41	0.84	1.000
15	.016	8	-0.02	0.03	-0.11	0.61	0.56	0.23	202.4	1	88	0.33	19	1.37	1.41	0.84	1.000
16	.017	4	-0.13	0.22	-0.07	0.74	0.52	0.27	538.77	3	75	0.24	10	1.37	1.41	0.84	1.000
17	.018	5	-0.56	0.02	0.06	0.65	0.51	0.24	-27.26	0	100	0.09	17	1.36	1.41	0.84	1.000
18	.019	6	0.25	0.16	0.00	0.72	0.48	0.15	2622.3	6	83	0.40	20	1.36	1.41	0.85	1.000
19	.020	4	-0.67	-0.65	-0.09	0.97	1.58	0.19	7372.3	14	50	0.55	11	1.37	1.41	0.85	1.000

Background pixels updated = 92.51%

#	File	#Ref	ErrX	ErrY	ErrZ	RmsX	RmsY	RmsZ	Inorm	#Sig	%<2s	<Cor>	%Ful	XSiz	YSiz	ZSiz	Beam
20	.021	12	-0.36	-0.33	0.02	0.92	0.65	0.15	5080.1	10	50	0.50	17	1.37	1.41	0.85	1.000
21	.022	6	0.04	0.08	-0.16	0.63	0.56	0.29	-41.7	0	100	0.13	17	1.37	1.42	0.85	1.000
22	.023	9	-0.01	-0.02	-0.00	0.58	0.45	0.21	8485.8	13	67	0.45	16	1.38	1.42	0.85	1.000
23	.024	8	0.39	0.29	0.09	0.88	0.70	0.20	14484	16	63	0.41	16	1.38	1.43	0.85	1.000
24	.025	5	0.16	0.01	0.12	0.43	0.41	0.22	162.39	1	80	0.37	13	1.38	1.43	0.85	1.000
25	.026	5	-0.41	0.41	-0.01	1.21	0.86	0.27	525.08	3	40	0.46	12	1.38	1.43	0.85	1.000
26	.027	3	0.12	0.02	-0.09	0.27	0.32	0.14	-0.356	0	100	0.08	12	1.38	1.43	0.85	1.000
27	.028	8	-8.21	-8.30	-8.10	8.69	8.76	8.21	117401	34	50	8.59	17	1.38	1.43	8.85	1.000
28	.029	11	8.12	8.07	8.03	8.61	8.48	8.21	1271.7	3	82	8.34	14	1.34	1.48	8.93	1.000
29	.030	5	-0.35	0.27	0.09	0.68	0.69	0.24	33873	28	60	0.60	16	1.33	1.40	0.93	1.000
30	.031	3	-0.12	-0.21	-0.01	0.18	0.48	0.25	-8.561	0	100	0.04	32	1.33	1.40	0.90	1.000
31	.032	11	-0.05	-0.08	-0.11	0.62	0.50	0.20	7168.1	13	73	0.46	14	1.33	1.40	0.90	1.000
32	.033	8	-0.03	-0.06	0.08	0.60	0.39	0.15	48.002	0	88	0.31	13	1.26	1.37	0.93	1.000
33	.034	7	0.29	-0.43	0.04	0.36	0.70	0.16	16211	15	86	0.36	14	1.26	1.37	0.93	1.000

This is the typical output. The ErrX, ErrY and ErrZ should be less than |0.5|. Inorm will be positive %<2s should be less than 60 (most of the time). Xsiz, Ysiz and Zsiz should be greater than 0.5 and less than 3.0.

After data reduction you will see something like this

Coverage Statistics for data0m.raw

									Shell		
Angstrms		#Obs	Theory	%Compl	Redund	Rsym	Pairs	%Pairs	Rshell	#Sigma	%<2s
to	1.917	1081	1185	91.22	4.00	0.078	1003	84.64	0.078	47.99	72.9
to	1.522	2163	2330	92.83	4.14	0.086	2007	86.14	0.108	13.95	75.4
to	1.330	3206	3434	93.36	4.32	0.093	2967	86.40	0.137	7.55	77.7
to	1.208	4271	4589	93.07	4.06	0.098	3883	84.62	0.156	8.06	77.0
to	1.121	5287	5698	92.79	3.80	0.099	4692	82.34	0.114	7.95	79.4
to	1.055	6314	6848	92.20	3.60	0.101	5501	80.33	0.132	8.71	79.3
to	1.002	7244	7933	91.31	3.39	0.103	5976	75.33	0.160	5.39	80.5
to	0.959	8226	9074	90.65	3.17	0.103	6270	69.10	0.103	3.28	82.3
to	0.922	9189	10215	89.96	3.00	0.103	6589	64.50	0.104	2.85	81.7
to	0.890	10058	11306	88.96	2.87	0.103	6852	60.60	0.145	2.74	83.6

Look at the Rsym. This should be less than 0.2. The coverage in the 0.9 A range should be greater than 90. (The Rsym reported is for triclinic, it will be greater for other systems). Redund should be greater than 2.

At this point several *.raw files that contain your *hkl* values have been written, also several *.p4p files are included. All of these files are in the directory ...WORK. At this point move all of the *.raw and the *m.p4p to a new directory.

Rename the *m.p4p to your project name *.p4p. In this example we have the datam.p4p, we rename it to sucrose.p4p.

```
C:\Documents and Settings\jhr6675\Desktop\class>sadabs

SADABS-2006/1 - Bruker AXS area detector scaling and absorption correction
------------------------------------------------------------------------------

Expert mode (Y or N) [N]:
Enter listing filename [sad.abs]:

Laue group numbers:

[1] -1                          [8] -3m (rhombohedral axes)
[2] 2/m (Y unique)              [9] -31m (Z unique)
[3] mmm                         [10] -3m1 (Z unique)
[4] 4/m (Z unique)              [11] 6/m (Z unique)
[5] 4/mmm (Z unique)            [12] 6/mmm (Z unique)
[6] -3 (rhombohedral axes)      [13] m3
[7] -3 (Z unique)               [14] m3m

[0] to write list of equivalent indices for Laue/point groups to listing file

Enter Laue group number [2]: 2

Read reflection files written by EVALCCD (with extension .sad specified) or
by SAINT (extension .raw, default if no extension, or .ram for incommensurate
structures).  Either individual files for each scan or a single merged file
may be read.  It is important that all files are from the same crystal and
that reflections have been indexed consistently, i.e. that the orientation
matrices are similar (no rows with signs reversed)!  Note that XPREP can
re-index a .raw or .sad file transforming the direction cosines

Enter filename (/ if no more) [ ]: sucrose.raw

Mean and maximum errors in direction cosine check function =   0.000   0.000
The mean error should not exceed 0.005, and is usually caused by matrix
changes during data processing.

Approximate wavelength, cell and maximum 2-theta (from cosines etc.):
1.54179    7.755    8.705   10.864   90.006  102.942   90.002  117.97

==============================================================================

PART 1 - Refinement of parameters to model systematic errors

Thresholds should now be specified for excluding reflections from the
parameter refinement; these reflections may still be corrected and included
in the final output .hkl file

   4221 Reflections of which   1080 unique;    1.43 data per frame

Redundancy:          1     2     3     4     5     6     7     8     9+
Number of groups:    52   231   237   236   129    76    52    38    29

Mean(I/sigma):  -inf     0     1     2     3     5    10    15    20   +inf
Number of groups:    72    56    67    38    33   232   149   131   302

The following restraint esd could be increased for strong absorbers.
Restraint esd for equal consecutive scale factors [0.005]:
```

should all be about equal. Now Accept the results and continue. Enter an output file name

```
==============================================================================

PART 3 - Output Postscript diagnostics and corrected data

Write Postscript diagnostic file (Y or N) [Y]:
Enter name of Postscript file [sad.eps]:
Short (<21 chars) title for Postscript plots [Test]:

Repeat (R), write unmerged .hkl (U), merged .hkl (M), .sca (S), XD format (D),
testxtl.dat (BioXhit) (T) or quit (Q) [U]:
Reflection output file [sad.hkl]: sucrose.hkl
Lambda/2 correction factor (0 if no monochromator [0.0]:
   4077 Corrected reflections written to file sucrose.hkl
Estimated minimum and maximum transmission:  0.3820  0.7508
The ratio of these values is more reliable than their absolute values!
Repeat (R), write unmerged .hkl (U), merged .hkl (M), .sca (S), XD format (D),
testxtl.dat (BioXhit) (T) or quit (Q) [Q]:
```

when prompted (here I've inputted sucrose.hkl).

You should now have sucrose.p4p and sucrose.hkl. This will take us to the next step

For the next operation you will need to know the Laue Class. Open a command line. Navigate to your directory and type : sadabs at the prompt. Select the correct Laue Group and enter the *m.raw file when prompted.

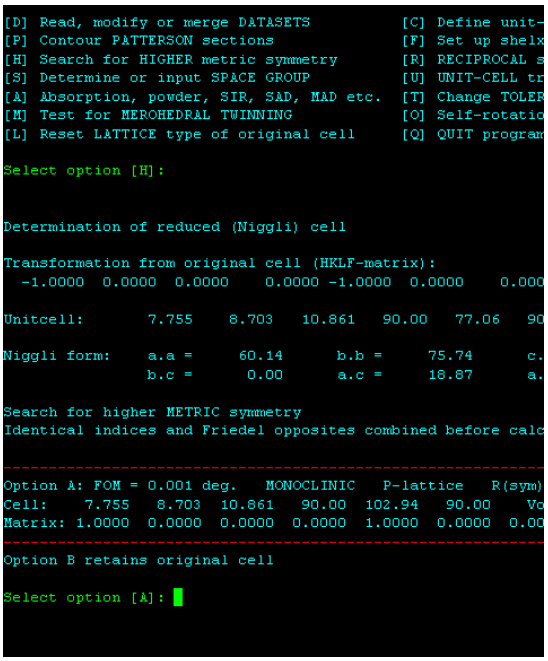
For the remaining questions use the value given. The R(int) should decrease upon refinement and the R(int) of the data runs should be less than 0.2. Look at the K values. They

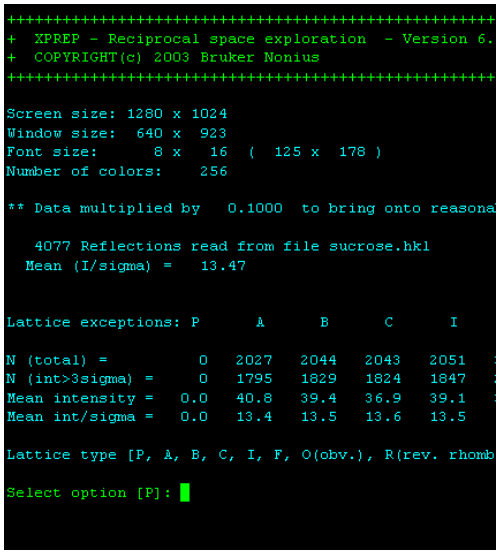
```
  3505 Reflections employed for parameter determination
Effective data to parameter ratio =    0.89

R(int) =  0.1866 (selected reflections only, before parameter refinement)

Cycle  R(incid)  R(diffr)  Mean wt.
   1    0.0968    0.0752    0.8823
   2    0.0554    0.0490    0.9238
   3    0.0436    0.0426    0.9337
   4    0.0403    0.0404    0.9378
   5    0.0395    0.0397    0.9403
   6    0.0390    0.0394    0.9418
   7    0.0388    0.0390    0.9432
   8    0.0382    0.0384    0.9445
   9    0.0377    0.0379    0.9457
  10    0.0371    0.0374    0.9468
  11    0.0367    0.0369    0.9478
  12    0.0363    0.0366    0.9485
  13    0.0360    0.0362    0.9493
  14    0.0358    0.0361    0.9500
  15    0.0356    0.0359    0.9509

R(int) =  0.0359 (selected reflections only, after parameter refinement)

Repeat parameter refinement (R) or accept (A) [A]:

==============================================================================

PART 2 - Reject outliers and establish error model

Rejected reflections are ignored in the statistics and Postscript plots
(except the detector diagnostics) and in the output .hkl file
Before applying rejections there are:

   4221 total and   1080 unique reflections assuming Friedel's law.
   4077 total and   1079 unique reflections left after |I-<I>|/su test

g = 0.0905  gives best error model.

Enter new value for g or <CR> to accept:

 Run 2-theta  R(int)  Incid. factors  Diffr. factors    K     Total I>2sig(I)
   1  -30.0  0.0306   0.686 - 1.627   0.953 - 1.142   0.761      330      320
   2  -30.0  0.0316   0.678 - 1.328   0.923 - 1.215   0.718      318      310
   3  -30.0  0.0408   0.749 - 1.655   0.906 - 1.265   0.921      325      311
   4  -30.0  0.0232   0.731 - 1.436   0.910 - 1.158   0.642      325      318
   5  -90.0  0.0398   0.649 - 1.200   0.907 - 1.267   0.663      613      550
   6  -90.0  0.0423   0.604 - 1.184   0.935 - 1.211   0.742      623      570
   7  -90.0  0.0426   0.669 - 1.280   0.906 - 1.236   0.627      602      550
   8  -90.0  0.0373   0.650 - 1.351   0.907 - 1.163   0.627      612      554
   9  -70.0  0.0478   1.125 - 1.633   0.908 - 1.093   0.758      329      302

su = K * Sqrt[ sigma^2(I) + (g<I>)^2 ]  where sigma(I) is estimated by SAINT

The above statistics are based on all non-rejected data, ignoring
reflections without equivalents when estimating R(int) and K.

Repeat parameter refinement (P) or accept (A) [A]:
```

XPREP. Prepare the SHELX instruction file. Assign space group and cell contents.

Type XPREP *myproject* here I will type XPREP sucrose

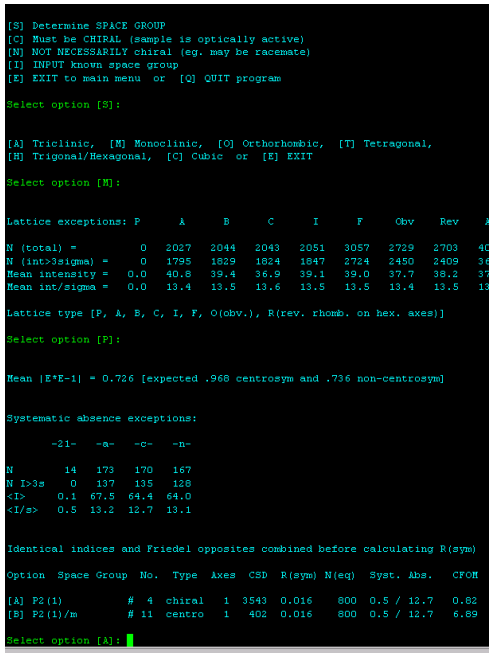
Choose the prompted values. Here I choose P for primitive and I went to higher symmetry





search. I select A for monoclinic.

Now its time to pick the space group.



First the lattice is Monoclinic [M] and Primitive [P]. The E*E-1 statistics in non-centroysmmetric. The systematic absence exceptions favour a 21 screw axis. Note the <I> for 21 is much less than the other reported <I> stats for -a- -c- or -n- as ins the <I/s> value. Also N I>3s is 0 for -21- and is large for -a-, -c- and -n-.
The spacegroup of P 21 (choice [A]) is acceptable. Continue hitting returns until you see

```
Select option [Q]: f

Output file name (without extension) [sucrose]:


File sucrose.ins set up as follows:

TITL sucrose in P2(1)
CELL 1.54178   7.7553   8.7028  10.8615  90.000 102.942  90.000
ZERR    2.00   0.0006   0.0006   0.0008   0.000   0.005   0.000
LATT -1
SYMM -X, 0.5+Y, -Z
SFAC C H O
UNIT 24 44 22
TEMP 0
TREF
HKLF 4
END

Do you wish to (over)write the intensity data file sucrose.hkl ? [N]: y
```

```
 Resolution  #Data #Theory
 Inf - 2.55      56     56
2.55 - 2.00      56     56
2.00 - 1.70      63     63
1.70 - 1.50      78     79
1.50 - 1.40      55     55
1.40 - 1.30      69     69
1.30 - 1.20     102    106
1.20 - 1.15      63     63
1.15 - 1.10      68     70
1.10 - 1.05      85     91
1.05 - 1.00     102    102
1.00 - 0.95     132    139
0.95 - 0.90     150    160
--------------------------
1.00 - 0.90     282    299
 Inf - 0.90    1079   1109

Merged [A],  lowest resolut

Enter <CR> to continue
```

Note the %complete is greater than 97% for inf-.9 This is good. You should see >90%. The redundancy is greater than 2. and the mean I/s is greater than 10 and the Rint is less than 4%.

Continue hitting returns until you see. Enter the empirical formula. In this case C12H22O11. The entry is case sensitive.

```
Select option [E]: F

Enter formula; numbers follow elements or brackets, 2nd character of element
name must be lower case, may include: Me, Et, Pp, Bu, Ph, or Cp:
C12H22O11


Tentative Z (number of formula units/cell) =   2.0  giving rho = 1.591,
non-H atomic volume =  15.5  and following cell contents and analysis:

C      24.00    42.10 %              H      44.00     6.48 %
O      22.00    51.42 %

[Z] change Z,  [F] new FORMULA,   [R] change RADIATION,
[E] EXIT to main menu  or  [Q] QUIT program

Select option [E]:
```

The Z will automatically select 2 (based on the rho value expected). You may need to adjust Z to match you compound.

Write the file when asked

You will now have Sucrose.ins and sucrose.hkl and you are ready for structure solution.

```
Select option [Q]: f

Output file name (without extension) [sucrose]:


File sucrose.ins set up as follows:

TITL sucrose in P2(1)
CELL 1.54178   7.7553   8.7028  10.8615  90.000 102.942  90.000
ZERR    2.00   0.0006   0.0006   0.0008   0.000   0.005   0.000
LATT -1
SYMM -X, 0.5+Y, -Z
SFAC C H O
UNIT 24 44 22
TEMP 0
TREF
HKLF 4
END

Do you wish to (over)write the intensity data file sucrose.hkl ? [N]: y
```